Cord Friebe · Meinard Kuhlmann
Holger Lyre · Paul M. Näger
Oliver Passon · Manfred Stöckler

The Philosophy of Quantum Physics

Springer

The Philosophy of Quantum Physics

Cord Friebe · Meinard Kuhlmann
Holger Lyre · Paul M. Näger
Oliver Passon · Manfred Stöckler

The Philosophy of Quantum Physics

Springer

Cord Friebe
Department of Philosophy
Universität Siegen
Siegen
Germany

Meinard Kuhlmann
Department of Philosophy
University of Mainz
Mainz
Germany

Holger Lyre
Chair of Theoretical Philosophy
University of Magdeburg
Magdeburg
Germany

Paul M. Näger
Department of Philosophy
WWU Münster
Münster
Germany

Oliver Passon
School for Mathematics and Natural Sciences
University of Wuppertal
Wuppertal
Germany

Manfred Stöckler
Institute of Philosophy
University of Bremen
Bremen
Germany

Translated by William D. Brewer, Institute of Experimental Physics, Freie Universität Berlin, Berlin, Germany

ISBN 978-3-319-78354-3 ISBN 978-3-319-78356-7 (eBook)
https://doi.org/10.1007/978-3-319-78356-7

Library of Congress Control Number: 2018936182

Printed on acid-free paper

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

Preface

The *Philosophy of Quantum Physics* grew out of an initiative by the Springer Spektrum publishers, for which we owe particular thanks to Dr. Vera Spillner. The task of coordinating the cooperation among the various authors and between authors and publisher was taken on by Cord Friebe—the other authors wish to express their hearty thanks to him for these sometimes tedious efforts!

Our guiding principle has been to fill a gap in the textbook market which exists between general introductions to this field and specialized monographs. The large number of popular articles on the subject documents the great interest shared by a broad audience in the epistemological and ontological implications of quantum theory. Our goal was thus to provide a current and well-founded introduction to the philosophy of quantum theory for advanced students of philosophy with an interest in physics. At the same time, this book confronts students and practitioners of physics with the philosophical implications of their field. This book can also serve to provide new impulses for teacher education in the fields of philosophy and physics.

That there is a close interrelation between the two fields is well known and needs no special justification. This relationship always becomes particularly intensive during periods of rapid scientific development. New physical theories can challenge the prevailing philosophical view of reality and even cause revisions to it. At the same time, philosophy can make its contributions to a more precise understanding and interpretation of the scientific results. The upheavals in physics in the early twentieth century due to the developments of quantum mechanics and the relativity theories provide substantial evidence for this.

In the second half of the twentieth century, a new development began, and it led the philosophy of physics to become a highly professional and very lively branch of the philosophy of science; it had been dominated earlier by authors in the Anglo-Saxon world. Motivations coming from the physicists continue to play a certain role, but the mainstream of research is now carried on by those philosophers who have solid backgrounds in physics and who concentrate their work on the fundamental questions and philosophical problems of the respective physical theories. The results of this research are published and discussed in specialized

journals. This professionalization has led to a situation in which the newer debates and the results of these discussions are hardly known within physicists' circles. This book thus aims to provide an approach to the current state of the important ongoing discussions.

The professional debates in journals and monographs usually presume that the reader has detailed mathematical, physical and philosophical foreknowledge. Here, again, our presentation in this book intends to bridge over this gap, and it presupposes essentially a knowledge of these fields only at the general school level. All the additional resources and concepts, including those from mathematics, are introduced at a basic level in the text. Depending on the extent of the reader's knowledge, this of course requires active participation and study, especially of the first two chapters.

It is a typical characteristic of quantum physics that even a century after its initial development, its consequences for our view of reality are still the subject of controversial discussions. Although the quantum theory permits the description and predictive calculation of many phenomena to an impressive extent, its relationship to the "objects" and the "properties" of the world remains unclear, in the sense that numerous different approaches compete for the exclusive right to explain this central relationship. This book offers an introduction to the numerous philosophical challenges provoked by the quantum theory. We retrace the course of various scientific–philosophical debates and classify them within the context of current research results. Fundamentally, however, we aim to provide a systematic account of the field.

The following résumé of the chapters in this book is intended to provide an orientation for the reader and to clarify the relations between its various parts. The first chapter provides a systematic access to the central, fundamental concepts of quantum theory in the sense of the overall conception of the book, including in particular the notion of "superposition", and furnishes successively the required mathematical apparatus. We forgo to a great extent the use of the differential calculus and differential equations. Instead, we presuppose only the basic fundamentals of coordinate geometry, vectors and linear algebra.

Following this introduction to the fundamentals, the second chapter describes the *minimal interpretation* and the "Copenhagen interpretation", which is still held as a standard by many physicists. The Copenhagen interpretation was, however, never strictly codified, and it is not without its own problems. In particular, its treatment of the measurement process and the role of the observer remain controversial even within this interpretation. Ghirardi, Rimini and Weber thus suggested a modification of the theory in 1986. Their interpretation, denoted by an acronym of the names of its originators as the "GRW approach", describes a "spontaneous collapse" of the quantum-mechanical wavefunction, and it is introduced at the end of the second chapter.

While the first two chapters consider only "single-particle states", the third chapter takes up "many-body systems" and their peculiarities. The surprising and characteristic property of the quantum theory represented by the fact of the empirical indistinguishability of similar quantum objects plays a decisive role here.

It leads in particular to quantum statistics, representing a significant revision of classical statistical mechanics. The empirical indistinguishability of quantum objects raises deep ontological questions of "identity" and "individuality", which have in turn led in recent times to a revival of the debate on the applicability of Leibniz's principle in quantum theory.

The fourth chapter treats the broad issues related to "entangled states" and "non-locality". Since entanglement involves a relation between two (or more) quantum systems, this chapter ties in with the concepts which were introduced in the third chapter. The distinctive feature of these systems lies in the fact that they apparently influence each other even when they are arbitrarily widely separated in space. This non-locality of the quantum theory is especially problematic because such an influence would have to propagate at more than the speed of light, in contradiction to our understanding of the special theory of relativity.

The discussion surrounding *entanglement* had its origin in a famous article published by Einstein, Podolsky and Rosen (EPR) in the year 1935, which made use of a thought experiment on entangled systems that still shapes today's discussions on the subject. EPR used their thought experiment to argue against the completeness of quantum mechanics, but they made the incorrect assumption that the quantum world is *local*. In the early 1960s, John Bell was able to show with his theorem that the thought experiment of EPR in fact demonstrates the non-locality of the quantum theory; even the introduction of "hidden variables" to complement the theory cannot rescue the locality demanded by Einstein. In the meantime, the EPR "thought" experiment has been carried out in practice in a number of places, with the result that non-locality is a verified basic feature of the quantum world. From Bell's argument, still more far-reaching consequences have been derived, and the fourth chapter contains a detailed discussion of these relationships and their justifications. Making use of causal graphs, the abstract discussion around Bell's theorem can be made more intuitively accessible.

The catchword "hidden variables" has already been mentioned, and in the first part of the fifth chapter, the best-known representative of this species of interpretations is introduced in the form of the "de Broglie–Bohm theory". Here, some of the more radical of the epistemological and ontological implications of, for example, the Copenhagen interpretation are avoided: Quantum objects in fact move along trajectories, according to this interpretation, and thus, in a formal sense, this theory is indeed deterministic. The price which must be paid for this lies in some of its properties, whose acceptability is still a matter of controversy. A similar situation is found in the "many-worlds interpretation" of quantum theory, which is treated in the second part of the fifth chapter. Its solution of the problem of measurement is as elegant as are its metaphysical implications extravagant. Both interpretations have the common feature that they dispense with the "collapse" of the wavefunction, and this is the reason why they are both treated here within one chapter.

In the sixth chapter, a bridge is built to the (relativistic) quantum field theories. Particle numbers now become variable (we speak, e.g., of the "creation" and "annihilation" of particles and quanta). Quantum field theories make it possible to treat the interactions between radiation and matter within the framework of the

quantum theory. Thus, old problems such as the wave–particle duality and the non-locality of the microscopic world can be attacked using new mathematical tools. To be sure, it becomes especially clear here what has been a source of difficulties in all the chapters: The question of how the mathematical formalism of the theory can be brought into coherence with the "real world" no longer admits of a simple answer, if one goes beyond the mere allocation of possible measured values to concrete measurement results, i.e. if one is not content with a minimal interpretation of quantum theory.

Finally, the seventh chapter rounds off the book by presenting a short chronology of the important steps in the development of quantum theory, in terms of both its physical–mathematical properties and also its interpretational questions. The essentially systematic structure of this book is complemented here by this historical treatment, and one can read the brief explanations of the milestones in the development of the theory like a glossary. Furthermore, some additional approaches to the interpretation of quantum theory are mentioned here, which could not be treated in detail in the rest of this book.

The many intensive discussions within the group of authors have shown that each of the six authors would have written a different book on the same topic. We hope that our mutual cooperation has led to the "best of all possible versions".

Bonn, Germany — Cord Friebe
Mainz, Germany — Meinard Kuhlmann
Magdeburg, Germany — Holger Lyre
Münster, Germany — Paul M. Näger
Wuppertal, Germany — Oliver Passon
Bremen, Germany — Manfred Stöckler
July 2014

Preface to the English and Second German Edition

The suggestion of an English and second German edition by Springer has provided us with a welcome opportunity not only to eliminate various typographical errors and minor flaws, but also to improve the presentation of the text and to make some additions to its content. Many helpful comments were made by students in our seminars in Wuppertal, Mainz, Saarbrücken and Bonn. The inclusion of exercises at the end of each chapter is intended to provide an aid to self-study as well as reference points for group discussions. Example solutions to the exercises are given at the end of this book.

Siegen, Germany — Cord Friebe
Mainz, Germany — Meinard Kuhlmann
Magdeburg, Germany — Holger Lyre
Münster, Germany — Paul M. Näger
Wuppertal, Germany — Oliver Passon
Bremen, Germany — Manfred Stöckler
December 2017

Contents

About the Authors

Cord Friebe studied philosophy, physics and mathematics in Freiburg (Germany) and Padua (Italy). He obtained his doctorate in 1998 with a thesis on "the Ontology of 'Indistinguishable' Objects in Quantum Mechanics and Quantum Field Theory". His research interests include Philosophy of Physics, Analytic Ontology, Kant's Theoretical Philosophy. He has been visiting scholar in Trient and Sydney; he has held teaching positions in Freiburg, Bonn, Saarbrücken, Cologne, Constance. Since 2018, he has been working as Professor for Theoretical Philosophy with Emphasis on Analytic Philosophy at the University of Siegen. His most recent books are *Zeit–Wirklichkeit–Persistenz. Eine präsentistische Deutung der Raumzeit* (2012); *Geld: eine philosophische Orientierung* (2015). His current research project (DFG) is on recent developments concerning the individual character of quantum objects.

Meinard Kuhlmann studied physics and philosophy at the Universities of Bochum and Munich (Germany), St. Andrews (Scotland) and Cologne (Germany), He obtained his Diploma in Physics at Cologne in 1995; doctorate in 2000; and *Habilitation* in Philosophy in Bremen 2008 (Germany). He was Visiting Scholar at the University of Chicago and UC Irvine (1998), Oxford University (2002/03), University of Pittsburgh (2010) and the London School of Economics (2011). Since 2012, he has been serving as Chairman of the Working Group *Philosophy of Physics* of the German Physical Society (DPG). He held Interim Professorships at the Universities of Hannover, Jena and Bielefeld in 2010–2014. Since 2014, he has been working as Interim Professor for the Philosophy of Science at the University of Mainz. His chief areas of interest include Philosophy of Science (in particular, ontology of quantum field theory and epistemology of complex systems theories in physics and economics) and Metaphysics (in particular, bundle theories, properties and causality). His monographs are *Ontological Aspects of Quantum Field Theory* (co-editor with H. Lyre and A. Wayne, 2002); *The Ultimate Constituents of the Material World—In Search of an Ontology for Fundamental Physics* (2010).

Holger Lyre studied physics, philosophy and computational neuroscience at the Universities of Marburg, Dortmund and Bochum (Germany). He obtained his Diploma in Physics in 1993 at Dortmund, doctorate in Philosophy in 1996 at Bochum and *Habilitation* in Bonn, 2003. Interim professorships at Bielefeld and Augsburg; since 2009, he has been working as Professor for Theoretical Philosophy/Philosophy of the Mind at the University of Magdeburg (Germany). He was Visiting Scholar at the University of Pittsburgh (1998/99) and UC San Diego (2014) and Founding President of the Society for Philosophy of Science in 2011–2016 (GWP, Düsseldorf). His principal research areas include Philosophy of Science (in particular, the philosophy of physics and theory of the cognitive sciences) as well as Philosophy of the Mind. His monographs are *Quantentheorie der Information* (1998); *Ontological Aspects of Quantum Field Theory* (edited with M. Kuhlmann and A. Wayne, 2002); *Informationstheorie. Eine philosophisch-naturwissenschaftliche Einführung* (2002); *Lokale Symmetrien und Wirklichkeit* (2004); *C.F. v. Weizsäcker: The Structure of Physics* (edited with T. Görnitz, 2006); *Kants 'Prolegomena': Ein kooperativer Kommentar* (edited with O. Schliemann, 2012).

Paul M. Näger studied physics and philosophy in Munich (2006 M.Sc. Physics, *LMU Munich*). Following a research visit in Oxford (2008/09), he was a Research Fellow at the University of Bremen (Germany, 2010–2013), where he obtained the doctorate in Philosophy with a thesis on the peculiar causal features of entangled quantum systems. Since then, he has carried out research and teaching at the *Westfälische Wilhelms-Universität Münster* (Germany; 2013-7: Post-Doc; since 2017: Assistant Professor). His principal research areas are in the Philosophy of Physics (in particular, quantum physics and the theory of relativity), Philosophy of Science (in particular, statistical methods and causal inference, explanations, reduction/emergence) and Metaphysics (in particular, space and time, causality, mereology). He is author of the book *Quantum Entanglement and Causation* (forthcoming, Springer) and a co-editor of the book *Peter van Inwagen: Materialism, Free Will and God* (2018, Springer, with L. Jansen).

Oliver Passon studied physics, mathematics, philosophy and education at the University of Wuppertal (Germany). He obtained his Master's (1998) and doctorate (2002) in Experimental Elementary-particle Physics with a thesis on the analysis of data from the DELPHI experiment at CERN, Geneva. He was Research Staff Member at the Jülich Research Centre in the Central Institute for Applied Mathematics (2004–2007). He completed his internship as a teacher of physics and mathematics (2nd State Examination 2008). He worked as teacher of physics and mathematics at the Carl-Duisberg-Gymnasium in Wuppertal (Germany; until 2013). Since 2013, he was been working as Lecturer in the research group for physics and its didactics at the *Bergische Universität Wuppertal*, and is a member of the Interdisciplinary Center for Science and Technology Research (IZWT) in Wuppertal. His principal research areas include Phenomenological Optics, Philosophy of Science and the Interpretation of Quantum Mechanics. His monograph is *Bohmsche Mechanik* (2004, 2nd revised edition 2010).

Manfred Stöckler studied physics and philosophy in Heidelberg and Gießen (Germany), and following his Master's in Theoretical Atomic Physics, he obtained the doctorate (Dr. Phil.) with a thesis on philosophical problems of relativistic quantum mechanics (1984). The subject of his *Habilitation* thesis was a philosophical investigation of elementary-particle physics and the fundamentals of quantum field theory. His academic career included being Research Staff Member at the Centre for Philosophy and the Foundations of Science in Gießen and at the Philosophical Seminar of the University of Heidelberg. Since 1991, he has been working as Professor for Theoretical Philosophy with Emphasis on Natural Philosophy and the Philosophy of the Natural Sciences at the University of Bremen (Germany). His principal research areas include Philosophical Problems of Modern Physics (in particular, quantum theory and cosmology), Philosophy of Time, and Methodological Problems in the Study of Complex Systems.

Chapter 1
Physical and Mathematical Foundations

Cord Friebe

This first chapter on the philosophy of quantum physics[1] will treat physical systems by ignoring their (possible) internal structure. Thus, we firstly consider single systems, e. g. single silver atoms or single electrons; the latter have in today's understanding in fact no internal structure at all. Thereby, we initially leave aside all those philosophical problems posed by composite or many-body systems: the loss of "individuality" of similar particles as well as the new kind of relationship between a whole and its parts, as exemplified in the famous EPR paradox. Both challenges are discussed later in their own chapters.

However, even a single system, as treated by quantum mechanics, gives rise to considerable problems of philosophical interpretation That said, we note that even macroscopic objects in the everyday world or in classical physics are not immune to philosophical controversies. On the contrary, the theoretical philosophy (epistemology, ontology) had its beginnings with Plato and Aristotle precisely in considerations of everyday objects which can be perceived by our ordinary senses. Numerically different persons, for example, can all be described as "courageous", as is done of Socrates in one of Plato's early dialogues. Then, one may ask what these persons have in common—*Courage* perhaps, i. e. a universal property which can be possessed, somehow, by the many? Namely either as a form *ante rem* from Plato's heaven, in which the concrete, individual persons can "participate", or else as something *in re*, which at one and the same time is "instantiated" in both this person and also that person. The problem of the relationship of an object or a person to its/his/her properties, which was raised here for the first time, is still a subject of current philosophical debate—quite independently of the developments of modern physics.

Or, let us take as a second example the early modern British philosopher David Hume and his sceptical argument against certain notions of causality: If we maintain

[1] Quantum physics, like classical physics, encompasses more than just mechanics, in particular also quantum field theory. Insofar as these other subfields are included, we will use the term "quantum physics". As a rule, however, this introductory chapter will deal with quantum *mechanics*.

C. Friebe et al., *The Philosophy of Quantum Physics*,
https://doi.org/10.1007/978-3-319-78356-7_1

that an event such as the motion of a billiard ball was caused by a collision with another moving billiard ball, then we apparently believe that the colliding projectile somehow *forces* the target ball to move, that some sort of power makes it inevitable that the effect occurs, necessitates its occurrence. What we in fact observe, however, according to Hume, is merely a temporal sequence and a spatial conjunction of two motions; simply a factual regularity and not a constraint, not a necessity. Are causal relations then nothing other than spatiotemporal regularities, or are there in the world necessary connections between events, over and above the merely observable? Again, this controversy is still continuing today—quite independently of quantum-physical phenomena.

A theme of particular importance in this chapter is therefore the following: To emphasize that a single quantum-mechanical system produces *additional* difficulties for every theoretical philosopher—regardless of whether he or she is a (modern) Aristotelian, a follower of Hume or of Kant: There are empirical phenomena in the realm of the microscopic, and theoretical consequences of quantum mechanics, which represent a particular challenge to philosophy, independently of the fundamental philosophical views of the particular philosopher—and that already considering a single system. The discussion of these phenomena and consequences can then prove philosophically fruitful in two ways: either as a fertilization of continuing philosophical discussions, whereby quantum mechanics is mobilized in support of an existing position; or else in the sense that it forces us to develop entirely new philosophical theories.

But taking up the topic in these ways could be considered to be biased: For, then, quantum mechanics would appear exclusively as a problem for philosophy, in that apparently only the question is posed as to which philosophical theory might be compatible with quantum mechanics. Physicists, but also many philosophers of physics, perhaps see this quite differently: When, for example, the physicist Niels Bohr introduced the concept of "complementarity" to aid in the interpretation of quantum mechanics, he was pursuing in particular the following goal: "Complementary", according to Bohr, are two quantities or two descriptions which on the one hand are mutually exclusive, but on the other hand *complete* each other, which may well seem at first view to be contradictory. However, what Bohr intended was by no means contradictory, so one may argue, since the two entities that he was considering, mainly the "particle view" and the "wave view", are *classically* mutually exclusive and *quantum-mechanically* completive. What he meant, accordingly, was the question of how quantum mechanics fits into the worldview of classical physics. We have apparently moved on past this question—who is still interested in the "worldview of classical physics"?—but in Bohr's sense, one still continues to ask: How is a realistic interpretation of quantum physics compatible with other parts of current physical theory, for example with the theory of special relativity? How are the theories of the microscopic realm related to the intuitive, everyday mesocosmic world (catchword: the measurement problem)? And how does the microcosm (in particular quantum field theory, or QFT) relate to the macrocosm (i. e. to the theory of general relativity, or GRT)? The question is therefore not which *philosophical* theory is most plausible in view of the results of modern physics, but rather whether and how quan-

tum physics can permit a unified theoretical structure of *physics* as a whole. This is not necessarily a contradiction, but it is quite a different matter whether one sees a philosophical problem only where physics itself is (still) inconsistent, as in the contradiction between QFT and GRT, or, by contrast, also where physics is already unified and consistent.

The focus of this textbook on the philosophy of quantum physics lies on this latter problem: In the first instance, we consider phenomena and theoretical consequences which can be considered in physics to be *well established*. This is even true of the next-to-last chapter on QFT, as long as it is not confronted with GRT; in particular, it holds for the important phenomenon of "indistinguishability" of similar particles (cf. Chap. 3), which represents a considerable philosophical challenge quite independently of the notorious measurement problem. Naturally, the theoretical structure of physics as a whole also plays an important role here, principally in the form of just that measurement problem; that is, the relationship between the microscopic and the macroscopic worlds which is considered in physics itself to be still an open problem. For its solution (or resolution!), newer theories must be considered—such as the (realistic) collapse interpretation of Ghirardi, Rimini and Weber (GRW; cf. Sect. 2.4), or the deterministic de Broglie–Bohm theory (cf. Sect. 5.1)—which are considered within physics to be *non*-standard. In this first chapter, however, we will begin by laying down the physical and mathematical foundations of the usual "standard quantum mechanics" of single systems.

1.1 Spin and Superposition

Seeing quantum physics at first view as a challenge to philosophical theories[2] would have an immediate consequence (didactically, as well): If we are not so much concerned with the theoretical structure of physics, then an introduction which treats quantum mechanics as emerging from classical physics and tries to give intuitive explanations of its unfamiliar phenomena in terms of notions of classical particles or waves would seem to be inappropriate. The historical context of discovery is perhaps not systematically relevant, so that we—in contrast to many popularizations on quantum mechanics—will begin neither with Planck's radiation quantization (discovered already in the year 1900), nor with the double-slit experiment or the photoelectric effect (1905); and thus *not* with those physical phenomena that suggest, according to Bohr, the dual nature of quantum systems as particles and waves: views which are classically mutually exclusive, but are ostensibly complementary within quantum mechanics.

Instead, we consider the quantum system from the beginning as an object in its own right, for which we—almost independently of the question as to whether it is

[2]That is, treating such problems as the relationship of a single object to its properties, or of a whole to its parts, or the relation between cause and effect and the questionable nature of persistence and temporal change in quantum systems.

mainly a classical particle or rather a wave or perhaps both—pose the questions as to whether it stands in relations of cause and effect, whether it persists over time, whether it can be individuated by intrinsic properties or merely relationally, etc. We therefore start with Bohr's second, quite different understanding of "complementarity", according to which two quantities such as position and momentum[3] rather are quantum-mechanically exclusive, though classically completive. As we shall see, this understanding can be precisely expressed within the mathematical formalism of quantum mechanics. And we will therefore begin with an experiment which was first performed in 1922, when quantum mechanics had thus already been "underway" for more than 20 years. This experiment demonstrated a property of quantum-physical systems for which there is *absolutely no classical analogy*; it can thus stand as a paradigm for the autonomy of quantum systems. We are referring to the "intrinsic angular momentum" or *spin*.[4] This spin demonstrates physically what is peculiar about quantum mechanics, and it directly motivates the vector space formalism of quantum mechanics which follows.

A warning is indeed appropriate here: In this treatment, the problem of embedding quantum mechanics within physical space is downplayed. That is quite appropriate for standard quantum mechanics. In later sections, however, we will see that this problem remains on the agenda of (the philosophy of) quantum physics. The spatiotemporal interpretation of quantum objects occupies the focus of the GRW, Bohm, and QFT approaches.

1.1.1 The Stern–Gerlach Experiment

In February 1922, the experimental physicists Otto Stern and Walther Gerlach carried out an experiment in Frankfurt am Main that, later, yielded an important theoretical contribution to the mature, internally consistent quantum mechanics, which was by that time still a mixture of classical elements and new principles (keyword: the Bohr model of the atom). Practically, the discovery of the "Raumquantisierung der magnetischen Momente in Atomen", as it is called on a plaque at the building of the Physical Society, has found application, e.g., in magnetic resonance tomography. In 1943, Stern was granted the Nobel Prize in physics for the discovery of the (electronic) spin

Using a small furnace to evaporate the metal, Stern and Gerlach produced a beam of *silver atoms*, which was then passed through an inhomogeneous magnetic field. After passing through this field, the beam was detected by collecting the atoms on a "screen" (Fig. 1.1).

In fact, they had not expected anything of special importance: All of the atoms in the beam of silver vapour are of course electrically neutral, so that they do not

[3]The momentum of an object is its velocity multiplied by its mass.

[4]Spin is an angular momentum which can be intuitively imagined as a rotation of the object around itself (which indeed for point-like objects becomes highly *non*-intuitive).

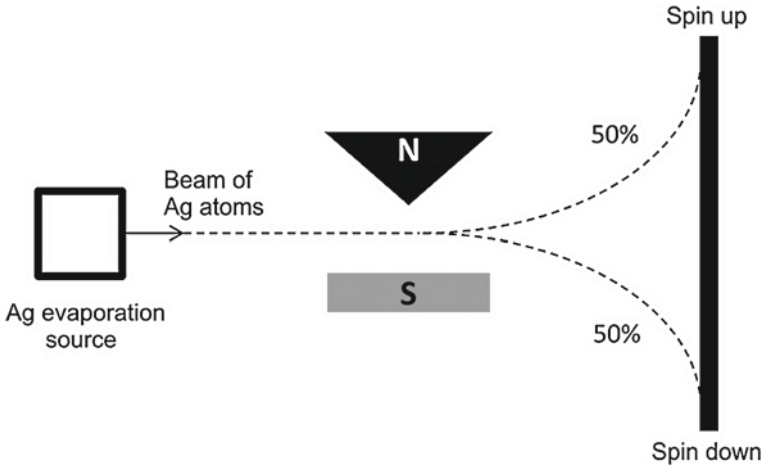


Fig. 1.1 The Stern–Gerlach experiment: a beam of silver atoms is split by an inhomogeneous magnetic field. This can be explained in terms of the electron's *spin*

experience a Lorentz force within the magnetic field. They could to be sure have magnetic moments which might be produced by the orbital motions of the electrons in the atomic shells. As was already known at that time, the magnetic moments in filled "energy shells" cancel each other, so that at most the single valence electron of the silver atoms might contribute. However, it is in the rotationally symmetric (5-)s orbital, so that its orbital angular momentum is zero—there was thus no magnetic moment to be expected. Nevertheless, Stern and Gerlach observed a splitting of the silver beam, leading to *two separate* silver traces on the screen, indicating that here a magnetic moment can have two discrete orientations. This could be explained by assuming that the valence electrons of the silver atoms—and thus electrons in general—have an additional property, up to that time undiscovered, which behaves like an intrinsic angular momentum that produces the magnetic moment and has only two possible values: spin up and spin down.[5]

In the intervening years, one has also discovered particles whose spin exhibits more than just two possible orientations. A beam of such particles, electrically neutralized to avoid additional magnetic or electrostatic forces, would thus show a splitting into several sub-beams in an inhomogeneous magnetic field. Furthermore, quantum objects are classified today into those with half-integer spins ("fermions", e. g. electrons) and those with integer spins ("bosons", e. g. photons). The latter classification will be important in the chapter on "indistinguishable" quantum particles, and the former should be kept in mind and should give rise to the question at an

[5] A note for readers with a physics background: Mathematically, the spin should also behave as an angular momentum; that is, its operators should obey the same commutation relations as those for the orbital angular momentum. That it behaves not only *analogously* to an angular momentum, but indeed *is* one, is shown by the fact that, in general, the quantum-mechanical orbital angular momentum is a conserved quantity only when it is coupled together with the spin.

appropriate point as to how the physics and mathematics of the two-valued spin can be generalized to higher-valued spin systems. At this point, only the following need be taken into account: The Stern–Gerlach apparatus always has a certain spatial orientation determined by the direction of its inhomogeneous magnetic field. Therefore, more precisely speaking, only the *projection* of the particle spin along a given axis is measured, and *it* has only two possible values in the case of electrons. In fact, the electron (like many other particles) has many other possible spin values, namely (for the electron) two each in *every possible direction* in space. Using a Stern–Gerlach apparatus, however, one can measure the spin of a quantum system at a given moment along only *one* spatial direction, which raises the philosophically important questions as to whether or not (for example) an electron "nevertheless" has at each moment all of these spin values which cannot be simultaneously measured; and if not, then what actually occurs in a Stern–Gerlach measurement?

The peculiar characteristics of this new property then become especially apparent if one carries out several Stern–Gerlach experiments in sequence.[6] Initially, following a measurement[7] in a certain spatial direction, we could carry out a second measurement relative to the same direction, which intuitively corresponds to a simple repetition. Or else we could carry out several spin measurements relative to *different* spatial directions, one after the other, which would perhaps yield unexpected results. Finally, the most important step would be to investigate what happens when we "reverse" a measurement which, apparently, has already been carried out: This leads directly to the principle of *superposition* of spin states and thus to the vector space formalism (for the following, cf. Albert (1992), in particular Chap. 1).

1.1.2 Sequential Spin Measurements

One of the most important problems in the interpretation of quantum mechanics is illustrated by properties which are *not* measured at a particular moment, and of which we cannot say *with certainty* whether they are present or not. We *could* draw the anti-realistic conclusion that we should not speak at all of real properties of a single quantum system, but instead perhaps only of macroscopic measurement apparatus and measurement outcomes, or only of quantum systems as *ensembles*, i. e. only of a large number of quantum particles. However, let us leave these interpretation threads aside for the moment and try to maintain for as long as possible the concept of a quantum system as an object which has properties, which persists over time, and thus can change its properties while maintaining its identity. We shall see how far we can go with this concept.

---

[6] This is of course only possible if one does *not* collect the particles on a screen.

[7] Here and in the following, we use the word "measurement" with as few preconceptions as possible. We neither presume that a macroscopic detector (screen) has to register events irreversibly, nor even that the quantum-mechanical system in some way "collapses".

Let us therefore first assume that an electron possesses the property spin up at a particular time (along the x direction), and at a later time, it has the opposite property spin down (along the x direction). This changeover of the spin property is assumed to be caused by an external magnetic field and to be perfectly predictable. The transition is thus causal and deterministic. Nevertheless, we can already discern a philosophical problem in it: Changes can in principle only occur *continuously*, so that such a discontinuous change could be only a "*sideroxylon*" or "wooden iron" (i. e. a contradiction in terms). Nature makes no jumps, as Aristotle already showed (cf. Aristotle 1988, *Physics*, Book VI): In the case of a discontinuous change, the object could be only either still in its initial state, or already in its final state, which would mean that it would have the change still to come or already past, without ever actually *passing through* it. One can react to this objection in two ways: Either one could demonstrate that Aristotle was in error, that there can indeed be discontinuous changes; or else one could assert that quantum-mechanical objects cannot actually *change*, but rather only being *exchanged*. For change requires something persistent, according to Kant (cf. Kant 1781/87, A 187/B pp. 230ff.), something of a self which survives over time, only whose properties (can) vary; and this is still indisputable today: "change needs identity as well as difference" (Mellor 1998, 89). Identity over a period of time, however, is (perhaps) not to be presumed of quantum-theoretical systems. The old particle may vanish, and the new particle with a different spin takes its place, perhaps because they are, as claimed by the so-called trope ontology (compare Sect. 6.5.2), only bundles of (particular) properties. Accordingly, at the one moment in time, the bundle [electronic charge, electron mass, spin up along the x direction, etc.] exists, and at a later moment in time, the numerically different bundle [electronic charge, electron mass, spin down along the x direction, etc.] appears, so that a discontinuous jump is possible. However that may be: We assume for the time being that quantum jumps no longer represent a philosophical problem.

The situation becomes truly problematic only when additional spin properties relative to *different* spatial directions are considered. If the inhomogeneous magnetic field of the Stern–Gerlach apparatus is turned, for example, into the y direction, then we once again have two possible measurement outcomes, which are supposed to correspond to two additional properties of the silver atoms or their valence electrons: spin up and spin down along the y direction. Following a measurement of the spin in the original x direction at a certain time, we now have two possibilities for continuing: We could on the one hand again measure the spin relative to the x direction, as shown in Fig. 1.2. It will be seen, as one might probably expect, that indeed every particle which has been shown to have its spin up (or its spin down) in a previous measurement must show *with certainty* the same result in a later measurement, independently of how much time has elapsed between the two measurements—as long as no external influences (such as magnetic fields) perturb the particles. One says that spin measurements are *repeatable*.[8]

[8]But how do the particles "exhibit" a particular spin value if they are not registered by a detector (e. g. a screen), showing that they were deflected upwards or downwards?—Those who are already

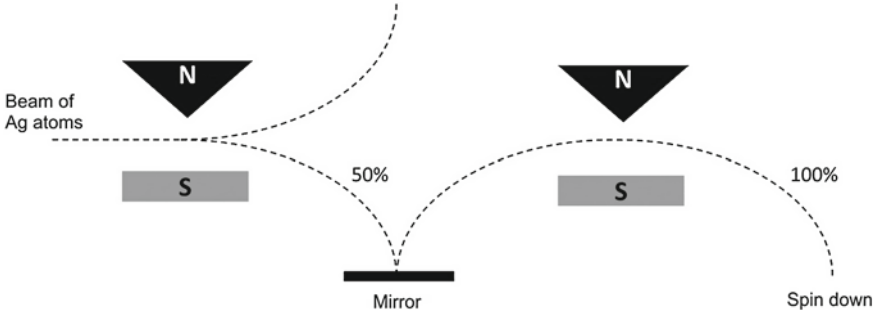


Fig. 1.2 A repeated measurement: 100% of all particles which previously exhibited spin down along a certain direction again show spin down in later measurements along the same spatial direction

If, on the other hand, we measure the spin relative to the y direction in the second measurement, then we find (which is perhaps more unexpected) that the results are now *not* predictable with any certainty. Both the spin up particles from the first measurement, as well as the spin down particles, proportionally show spin up and spin down along the y direction. In our example of a second measurement along a direction perpendicular to the first, we find more precisely that, respectively, exactly 50% of the particles will exhibit spin up and 50% spin down. In the case that the previous measurement had shown, e. g., spin up along the x direction, the change to, for example, spin up along the y direction takes place apparently *non-deterministically*. This is astonishing and should not be confused with a possible indeterminacy in the original measurement; in that case we would have had no kind of foreknowledge of the precise state of the particle (aside from the fact that it would be a silver atom or an electron). It was so to speak "a shot in the dark". In the second measurement, in contrast, we already know that the particles have the property spin up (or spin down) along the x direction. As we shall see, this information is even already *maximal* according to the formalism of quantum mechanics and its standard interpretation(s)—and thus the indeterminacy of the spin along the y direction with a given spin value along the x direction is *inevitable*. Einstein saw a serious problem here, as expressed in his famous slogan that "God doesn't play dice". What would be the explanation if the second measurement were to yield e. g. spin down along the y direction? Would the particle have possessed this property already previously to the second measurement, so to speak the *whole time*? Then more would be present in the world than we can maximally know, and the standard quantum mechanics would be *incomplete*. In any case, an assumed causal influence by the Stern–Gerlach apparatus on the particle does not yield a sufficient explanation, for that influence is symmetric, the opposite result spin up (along the y direction) would have been just as likely. The measurement along the y direction at most apparently "causes" the spin along

familiar with this problem may well ask this question. We, however, follow the particle in thought along its "trajectories" and show that this leads inevitably to inconsistencies.

the y direction to take on a certain value: *either* spin up *or* spin down. Or is there a probabilistic, non-sufficient causality, as is presumed today by many philosophers?

But let us leave aside not only the problem of the quantum jumps, but also that of the indeterminacy (for the moment). Our story is still far from its end. We have now carried out two measurements one after the other; first, we determined the spin value relative to the x direction (measured quantity: $\hat{S}_x$), and then we measured its value relative to the y direction (measured quantity: $\hat{S}_y$). At a still later moment in time, we can now either repeat the measurement of $\hat{S}_x$ or once again measure $\hat{S}_y$ Let us first consider the sequence $\hat{S}_x\hat{S}_y\hat{S}_y$: The second $\hat{S}_y$ measurement is, independently of how much time has elapsed since its first one, and as long as no external influences (magnetic fields) are present, clearly just a repetition of the first. Both $\hat{S}_y$ measurements follow one after the other *directly*, i. e. there are no further measurements or similar events between them. We can therefore say with certainty that particles with spin up (or spin down) along the y direction will again exhibit spin up (or spin down, respectively) along that same direction. What, however, will occur with the sequence $\hat{S}_x\hat{S}_y\hat{S}_x$ (compare Fig. 1.3)? Is the second $\hat{S}_x$ measurement now again just a repetition of its first, so that all those particles which leave the $\hat{S}_y$ apparatus, those with spin up along the y direction and those with spin down, will again show the same result as the first $\hat{S}_x$ measurement; those with spin up along the x direction again with certainty spin up, and those with spin down along the x direction again with certainty spin down? Do the two $\hat{S}_x$ measurements follow just as directly as the two $\hat{S}_y$ measurements before, since between them there is "only" a measurement of $\hat{S}_y$, that is a measurement of a quite different kind? Quantum mechanics (and of course the empirical results) say: No! *All* the particles behave after the $\hat{S}_y$ measurement in such a manner as if the previous $\hat{S}_x$ measurement had never occurred. Both those particles which originally exhibited spin up along the x direction—and now (in addition?) either spin up or spin down along the y direction; as well as those that originally showed spin down along the x direction—and now (in addition?) either spin up' or spin down along the y direction, yield in the second $\hat{S}_x$ measurement *in part* spin up and *in part* spin down along the x direction; precisely 50% each of spin up and spin down. The measurement of $\hat{S}_y$ clearly *destroys* the results of the previous $\hat{S}_x$ measurement, so that a further measurement of $\hat{S}_x$ is *not* a repetition of its first measurement.

The same conclusion holds in the reversed case, that is for the sequence $\hat{S}_y\hat{S}_x\hat{S}_y$; measurements of $\hat{S}_x$ and $\hat{S}_y$ mutually destroy each other. And all this holds independently of the elapsed time between the measurements (other influences are presumed to be negligible). This means, in particular, that immediately after the second measurement (in the figure: the measurement of $\hat{S}_y$), the third measurement will yield 50% spin up and 50% spin down, independently of what the first measurement showed. Neither the time sequence (i. e. whether we consider $\hat{S}_x\hat{S}_y\hat{S}_x$ or $\hat{S}_y\hat{S}_x\hat{S}_y$), nor the elapsed time between the measurements has a relevant influence on this decisive conclusion: the *destruction* of the result of a spin measurement by the following measurement along a different spatial direction. One concludes from this: The

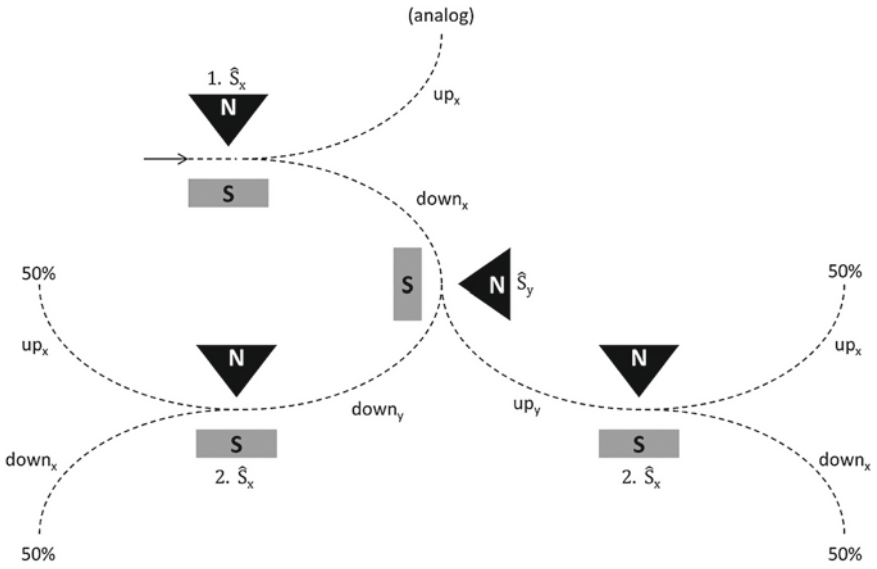


Fig. 1.3 Destruction of the result of a spin measurement of $\hat{S}_x$ by the following measurement along a different spatial direction, here $\hat{S}_y$

simultaneous measurement of the spin along different spatial directions is impossible *in principle*, just as it is impossible *in practice* to build a Stern–Gerlach apparatus whose inhomogeneous magnetic field points at the same time in two different spatial directions. At this point, as we have already remarked, one of the major questions of the interpretation of quantum mechanics raises its ugly head: Is this only an *epistemological* problem, or is it also an *ontological* one? Is it simply the case that our knowledge is (necessarily) limited, so that we will never be in a position to precisely *determine* simultaneous spin values in different directions? Or is it the case that the quantum-mechanical systems themselves cannot *possess* such properties at the same time, but rather always only one at a time, that is either only spin up (or spin down) along the x direction and *no spin value* along the y direction (and also none along any other different spatial direction)? At the moment of a spin measurement, e. g. during the measurement of $\hat{S}_y$, as a rule a real *changeover* of properties (e. g. from spin up along the x direction to spin up along the y direction) would then be *produced*, and non-deterministically at that!

1.1.3 The Superposition Principle

We will for the moment also set this problem aside. We simply mention the various philosophical problems which arise along the way, but in this introduction we are

initially more concerned with providing a motivation for the mathematical formalism of quantum mechanics. We can fulfil this purpose, most notably, by considering a further extension of our spin measurements on single systems. To this end, we extract from an original spin measurement, e. g. of $\hat{S}_y$, all of those particles which have (for example) their spin up, and send them through another Stern–Gerlach apparatus with perpendicular spatial direction, an $\hat{S}_x$ apparatus. One can expect that one-half of the particles will leave this $\hat{S}_x$ apparatus through the spin up exit, and the other half will leave it through the opposite spin down exit. The particles which have thus been separated are now passed to a system of "mirrors" and thereby mixed and again combined into one beam, so that we can no longer distinguish which particle had taken which path (cf. Albert 1992, pp. 7ff.).

What would now happen if we next perform another measurement of $\hat{S}_x$? "Another", since $\hat{S}_x$ had indeed already been measured once before.[9] And since a system of mirrors does not represent an additional spin measurement, the second $\hat{S}_x$ measurement also takes place *directly* after the first. Then, however, the second $\hat{S}_x$ measurement must be a *repeated measurement*, which with certainty (100%) must show a *particular* result. On the other hand, one is inclined to assert that 50% of the particles must show spin up and 50% spin down, since they were mixed by the system of mirrors and, therefore, the first $\hat{S}_x$ measurement has been *revoked*. And in fact, the second $\hat{S}_x$ measurement does not yield a unique spin projection, but rather one-half each of spin up and spin down. This seems to be trivial, but it is not completely without problems: It shows namely that "repeatability" cannot be simply assumed as a characteristic of what we call a measurement. The question of what is allowed to happen between two successive measurements of the same type so that the second counts as simply a repetition of the first is apparently not so easy to answer. Some physicists and philosophers conclude from this that the characteristic *quantum-mechanical* measurement procedure is a fundamentally problematical concept, which should best be avoided, as e. g. in the GRW approach (compare Sect. 2.4).

One could, however, react in a quite different manner: One could assert that the essence of a measurement lies not only in repeatability, but rather also encompasses *irreversibility*. A measurement would then in fact actually be a measurement only when it cannot be reversed without leaving traces. Since in the case under discussion, the ostensibly first $\hat{S}_x$ measurement can evidently be reversed by the system of mirrors, it would then not count as a true measurement, so that an ostensibly second $\hat{S}_x$ measurement could thus (in a nearly trivial manner) *not* be a repetition of the first. However, many people currently would doubt this interpretation, since with irreversibility one raises a *macroscopic* quantity to an essential property of a measurement, and so *presupposes* an essential difference between the macroscopic world and the quantum world which, if at all, should become apparent only within the framework of an adequate interpretation of quantum mechanics.

Quite independently, we see an indisputable result at this point, which must be taken into account by every interpretation of quantum mechanics, and which can

[9] At least in the sense that we, in thought, can follow the "paths" of the particles along which it is supposed to have certain spin values in the x direction.

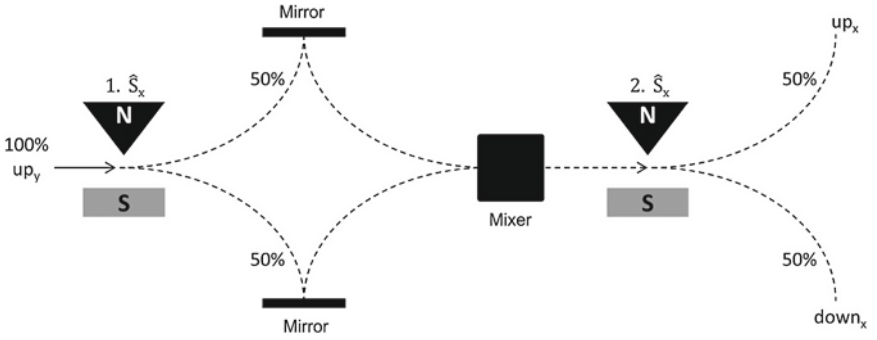


Fig. 1.4 The reversal of an (ostensible) spin measurement of $\hat{S}_x$ by remixing the particles

be considered to be the actual peculiarity of a single system, as well as of all other quantum systems: the possibility of the *superposition* of different states. Namely, instead of measuring $\hat{S}_x$ for the (ostensibly) second time as in Fig. 1.4, one could, after combining and remixing the particle beams, rather measure $\hat{S}_y$ for *in fact* the second time. We recall: A first, original $\hat{S}_y$ measurement (not shown in the figure!) had previously shown spin up in the y direction for all the particles which pass through the (ostensibly first) $\hat{S}_x$ apparatus and the mirror system. What, then, would a second $\hat{S}_y$ measurement *after* the remixing show? Would it show for all the particles once again with certainty spin up in the y direction?

The following considerations speak *against* this conclusion: *During* the passage through the mirror system, the particles take two separate "paths", along which they are supposed to have in each case spin up or spin down along the x direction, since the ostensibly first $\hat{S}_x$ measurement should at least have this effect and in fact does have it. Thus, if one were to place an additional $\hat{S}_x$ apparatus *into* one of these paths, then *that* apparatus with its factual second $\hat{S}_x$ measurement would show with certainty that all the particles have a certain spin—namely either only spin up or only spin down, depending on *which* path the second $\hat{S}_x$ apparatus is inserted into. However, particles which exhibit only spin up (or else spin down) in the x direction show upon measurement of the spin component along a perpendicular axis just 50% of spin up and 50% of spin down relative to that new axis. Therefore, if we were to place an $\hat{S}_y$ apparatus in one of the paths, we would measure 50% each of spin up and spin down (in the y direction). *Within* each of the paths, everything thus behaves just as before with the sequences $\hat{S}_x\hat{S}_y\hat{S}_y$ and $\hat{S}_x\hat{S}_y\hat{S}_x$, as was to be expected. The result obtained from the original $\hat{S}_y$ measurement, i. e. spin up in the y direction, is destroyed by the $\hat{S}_x$ measurement before the mirror system, in favour of a spin value along the x direction. If this is in fact the case, then after passing through the mirror system, 50% each of spin up and spin down in the y direction are combined, so that a following $\hat{S}_y$ measurement should show 50% each of spin up and spin down (Fig. 1.5).

Furthermore, if the $\hat{S}_y$ measurement were to give the same spin value as in the beginning (spin up in the y direction) *after the mixing* and *with certainty*, then

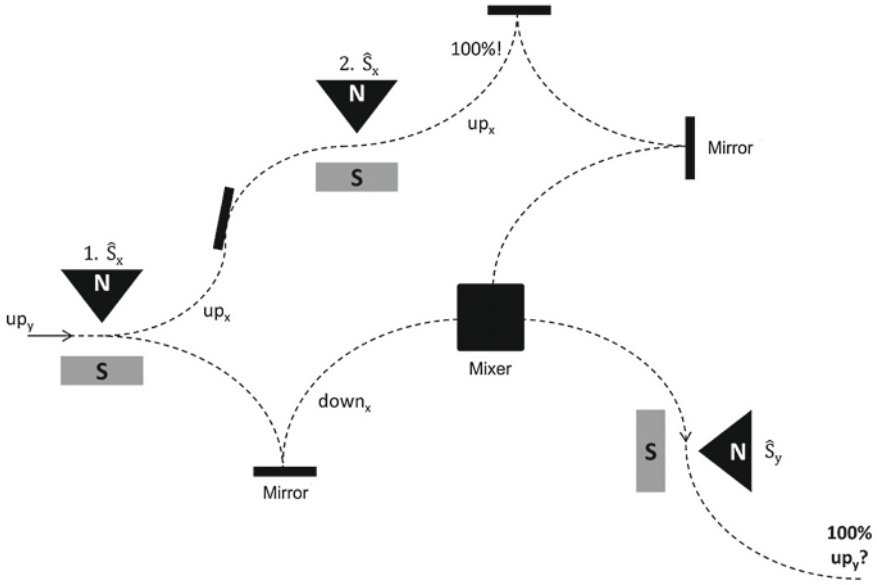


Fig. 1.5 Will the original $\hat{S}_y$ spin value "also" be restored by the mixing of the particles?

this second $\hat{S}_y$ measurement would be a simple repetition of the first, so that one would have to say that between the original $\hat{S}_y$ measurement and this second $\hat{S}_y$ measurement at the end of the whole passage through our system, nothing relevant had taken place. In particular, the $\hat{S}_x$ measurement before the mirror system would appear to have no relevant effect, since it is completely plausible to assume that the "mirrors" have no relevant effect on the spin value of the particles. However, it *cannot* be that the $\hat{S}_x$ apparatus also has no effect on the spin value of the particles; this *could* be verified by both $\hat{S}_y$ and also $\hat{S}_x$ measurements within the paths! And *yet* it is in fact the case that the second $\hat{S}_y$ measurement reproduces *with certainty* the result of the original measurement: All the particles after the mixing again show their original value—spin up along the y direction. One can only conclude from this that also this effect of an $\hat{S}_x$ measurement—namely the destruction of the spin value in the y direction—is *revoked* by the remixing of the particles. But while one can find an intuitively plausible explanation for the reversal of the $\hat{S}_x$ measurement in the sense of the destruction of the spin value in the x direction (or one at least believes that it can be found)—namely the mixing of particles with spin up in the x direction with particles which show spin down in the x direction—we cannot really understand how the mixing of the particles can reverse the effect of the $\hat{S}_x$ measurement also in the sense of the restoration of the original spin value in the y direction. After all, particles are "mixed" which (would) show along their respective paths one-half each of spin up and spin down in the y direction!

Both of these phenomena—on the one hand, the fact that in the end a second $\hat{S}_x$ measurement yields 50% each of spin up and spin down in the x direction, as if the

first $\hat{S}_x$ measurement had never occurred; and on the other hand, the fact that a second $\hat{S}_y$ measurement reproduces the result of the first, by yielding 100% spin up in the y direction, as if there had never been an $\hat{S}_x$ measurement in between—these facts can apparently be reconciled only by assuming that the $\hat{S}_x$ measurement can have only *one single effect* in general. It is simply not so that such a measurement could have *two different* effects, which could then be reversed completely independently of each other: one being that particles assume a particular spin value in the x direction and the other being that the spin value along any *different* axis is destroyed. In fact, we are dealing here with *one and the same* effect: The assumption of a spin value along one direction in space *is* at the same time and not additionally none other than the destruction of a spin value in all other directions. *Since*, in our case, the spin value in the x direction is again destroyed later, the spin value in the y direction is restored. These effects cannot be separated, because they are the same. One can say that the state of a particle with a particular spin value in a particular direction in space is at the same time none other than the *superposition* of two opposite spin values in some other direction. This does not mean that e. g. a particle with spin up in the y direction would have *both* spin up *and also* spin down in the x direction; that would be rather contradictory, especially considering that it would also have to have all the other spin values in all the other different directions. What is meant instead is unclear: "Superposition" is at this point simply an expression for something which has not yet been understood and which represents the main problem for the interpretation of all of quantum physics.[10] What is important at this point is that "superposition" motivates the mathematical formalism of (standard) quantum mechanics. It evidently behaves like the *linear combination* of vectors—each such linear combination yields a new vector in the same vector space, and each vector can be represented in infinitely many ways as a linear combination of other vectors—and this therefore motivates taking the mathematics of quantum mechanics to be the *theory of vector spaces*.

1.2 The Mathematical Formalism of Quantum Mechanics

We are seeking a mathematical formalism which is capable of expressing the phenomena described above, and which in particular can reproduce the decisive new property of the *superposition* of "states".[11] Such a formalism is presented by the structure of *vector spaces*, which we introduce here initially in a purely mathematical form. In order to keep the matter from becoming too abstract, we use an

[10] In physics, superpositions are also well known from the classical field theory in the addition of wave phenomena. Historically, the superposition principle was therefore introduced into quantum mechanics in order to describe the wave character of the particles. Here, however, it serves as a very abstract principle: Spin states are *not* waves.

[11] The concept of "states" will be discussed in more detail in the various sections on the interpretation of quantum mechanics. Intuitively, it somehow maps the set of properties which a quantum system possesses at a particular moment.

intuitively understandable geometric representation; the intended physical interpretation will also be kept in mind. However, we should note that it is important to distinguish carefully between the formalism and its application, in order to keep it clear just what is controversial and what is not, and where precisely the problems of the interpretation of quantum mechanics enter. Furthermore, we should note that in these introductory sections, we are considering only the standard formalism of ordinary quantum mechanics, as developed in John von Neumann's founding work of 1932 (cf. von Neumann 1932). Other descriptions—e. g. those used in the GRW approach (Sect. 2.4), or in the works of Bohm and of Everett (Chap. 5), or in QFT (Chap. 6)—raise additional questions and problems of interpretation. One could even ask whether such mathematical modifications are really interpretations of *the same* quantum physics, or rather already represent *alternatives* to standard quantum mechanics. Finally, as we have mentioned, many interpretation problems deal with "composite systems", with the "indistinguishability" of similar particles, as well as with the infamous EPR paradox: Their treatment requires distinctive mathematical features, which we will discuss later at suitable points in the text. Here, we limit our considerations again to single systems.

The following discussion is divided into four parts. The first two—"Vectors and their representation" and "Operators and their eigenvalues"—provide the necessary prerequisites for an understanding of the philosophical debates around quantum physics. They relate to school mathematics, expanded in terms of the central concepts of "(*Hermitian*) operators" (preliminary meaning: measurement apparatus; observable quantity), "eigenvalues" (measurement outcome; property) and "eigenvectors" (eigenstates). The section which follows them—"The Problem of multiple eigenvalues"—is more complex. It is necessary in particular for a deeper understanding of "indistinguishability" and "EPR/Bell" and can thus be skipped over by readers who are mainly interested in the chapters on the interpretations of quantum mechanics ("Copenhagen", GRW, Bohm, Everett). The fourth section, finally, deals with "special operators", which, depending on the problem at hand, may be required. We recommend that readers consult the corresponding subsections when they are referred to in the text.[12]

1.2.1 Vectors and Their Representation

A vector space $\mathcal{V}$ is a non-empty set whose elements, the *vectors*, can be added to one another and be multiplied by (real or complex[13]) numbers. This definition,

[12] One more comment: We have chosen the more abstract algebraic approach to quantum mechanics which, in particular, is suited for generalization to QFT. It is computationally less demanding; for example, no differential equations must be solved. For the (perhaps more intuitive, but computationally more difficult) calculus approach, we suggest that students of philosophy consult the book by Nortmann (2008).

[13] Complex numbers are an extension of the real numbers. The idea here is that quadratic equations can always be made solvable, i. e. even for negative numbers, by setting the imaginary quantity i

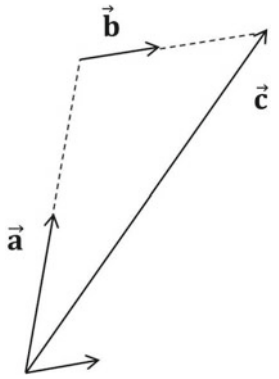


Fig. 1.6 Vector $\vec{c}$ is generated by taking a *linear combination* of the vectors $\vec{a}$ and $\vec{b}$ with the extension factors (coefficients or amplitudes) c_i: $\vec{c} = c_1\vec{a} + c_2\vec{b}$

which sounds rather abstract, is intended to describe the concept of a vector space as broadly as possible, so that also non-intuitive spaces are included within it. In the two-dimensional intuitive space, it simply corresponds to the well-known addition and elongation/contraction of arrows, as shown in Fig. 1.6.

At the same time, this visualization gives a first impression of how superposition is expressed mathematically: If quantum-physical states in some manner correspond to vectors, then the vector[14] $|C\rangle$ is a superposition of those states which are represented by $|A\rangle$ and $|B\rangle$.[15]

A vector space is called *Euclidian* (or when it is complex, *unitary*) if an *inner product* is defined on that space. This means that one can multiply two vectors by each other. Notation: $\langle A|B\rangle$ (referred to as "A times B"). To be sure, we do not obtain a new vector in this way, but instead simply a real (or complex) number. The intuitive meaning of the inner product is that it is the geometrical projection of one vector onto the other (see Fig. 1.7).

An inner product can be employed to determine the lengths of vectors—the norm: $\|\vec{a}\| = \sqrt{|\langle A|A\rangle|}$—or to measure the distances and angles between them. In particular, the inner product of two vectors which are mutually perpendicular is zero: $\langle A|B\rangle = 0$ This makes it possible to define the central concept of an *orthonormal*

as the solution to $x^2 = -1$ (this simplifies many computations). With the real numbers a and b, complex numbers in general take on the form $a + ib$, thus with a real part and an imaginary part, and they can be graphically represented in a plane, where one coordinate axis is taken to be imaginary.

[14] Instead of $\vec{c}$, it is convenient in quantum mechanics to write $|C\rangle$ for a state vector. The *dual* vector is written as $\langle C|$, so that, as we shall soon see, the inner product of two vectors is expressed simply as $\langle A|B\rangle$ (bra-ket notation due to P.A.M. Dirac).

[15] By addition and extension/contraction, one always obtains new vectors which are also elements of the *same* space as the original vectors. The superposition of two possible states of a particular single physical system is likewise always an additional possible state of the same system. For many particles of similar type, however, there is an important limitation (keyword: superselection rules; cf. Chap. 3).

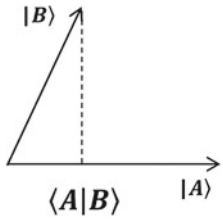


Fig. 1.7 The inner product $\langle A|B\rangle$ of two vectors is the geometrical projection of the one vector onto the other; it thus determines the angle between them

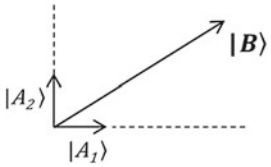


Fig. 1.8 In terms of the basis $\{|A_1\rangle, |A_2\rangle\}$, the vector $|B\rangle$ has the component representation $|B\rangle = 3|A_1\rangle + 2|A_2\rangle$, or $\binom{3}{2}$

basis (for short: basis).[16] In order to obtain such a basis, we first have to normalize the corresponding vectors to the same unit length and then choose the maximal number N of vectors which are pairwise mutually perpendicular. For these vectors, we have $\langle A_i|A_j\rangle = 0$; i, j are indices which run from 1 to N, with $i \neq j$. In the two-dimensional intuitive space, there are at most two mutually perpendicular vectors—a basis there defines a rectilinear coordinate system with two axes; and in general, it is found that N is precisely the *dimensionality* of the vector space in question.[17] Thus, *every* vector in our vector space can be represented in terms of a basis of that space, as shown in Fig. 1.8.

Thus far, we have considered vectors as abstract objects, for which we simply wrote $|A\rangle$; now we want to be more concrete. A vector is associated with certain numerical values as *components*, whose number is equal to the dimensionality of the vector space; in this way, the intuitive vector can be represented by: $|B\rangle = \binom{3}{2}$. The basis vectors themselves can also be written in terms of a component representation—in a real, two-dimensional space, they are $\binom{1}{0}$ and $\binom{0}{1}$. And then we can compute the inner products in a concrete manner, according to the following rule[18]:

$$\text{For } |A\rangle = \begin{pmatrix} a_i \\ \vdots \\ a_N \end{pmatrix} \text{ and } |B\rangle = \begin{pmatrix} b_i \\ \vdots \\ b_N \end{pmatrix}, \text{ we have}$$

[16] A set of vectors forms a *basis* of a vector space when all the other vectors in the space can be generated as linear combinations of them. We consider only orthonormal bases. It is important to note that a given vector space has an infinite number of such bases.

[17] In the extreme case, this dimensionality can be (countably) infinite.

[18] One can readily verify that the basis vectors given above are indeed normalized to length = 1 and are orthogonal (mutually perpendicular) to each other, i. e. *orthonormal*.

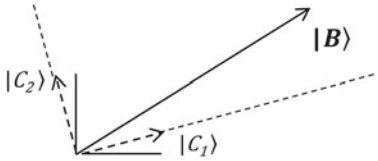


Fig. 1.9 Referred to a different basis $\{|C_1\rangle, |C_2\rangle\}$, *the same* vector has a different representation: $|B\rangle = 3.3\,|C_1\rangle + 1.2\,|C_2\rangle$ (numerical values are chosen so that the norm is conserved)

$$\langle A|B\rangle = a_1 b_1 + \cdots + a_N b_N = \sum_{i=1}^{N} a_i b_i \,. \tag{1.1}$$

Now there is, to be sure, not just one basis, but rather infinitely many of them. Referred to each such basis, the vectors have a unique set of components, so that the concrete numerical representation depends upon *which* basis is chosen (Fig. 1.9).

One and the same vector thus has infinitely many component representations which can be computed and converted into one another using so-called basis transformations. This variability of the representation will correspond to the fact that one and the same quantum-physical state—for example, the state with spin up in the x direction—is not only a superposition of the states with spin up and spin down in the y direction, but rather at the same time of infinitely many other superpositions of spin up and spin down along every other spatial direction. However, we have to remember that the choice of a particular basis is rather arbitrary: Strictly mathematically, it is purely conventional, and therefore, the quantities with a physical significance are rather those which are independent of the choice of basis, that is those which are *invariant* under basis transformations—which the components of the vectors are clearly *not*. Their lengths and the angles between the vectors should, in contrast, not depend upon which basis is chosen for their representation; and indeed, the inner product is an invariant quantity under basis transformations. Not least for this reason, the inner product has a special significance in physics, as well (Figs. 1.9 and 1.10).

The difference between the vector itself and its representation in terms of a particular basis can give us a first impression of the fact that the vector space formalism is basically a very abstract affair. The goal of an empirical theory such as quantum mechanics, however, is to confront the theory with experiments, i. e. with observational data. It is finally the success of predictions of the results of measurements that is the proving ground for empirical science. For *measurement outcomes*, however, there is only one possible mathematical representation, namely the *real numbers*, which we have not obtained here even after choosing a certain basis for the concrete representation of our vectors. Vectors correspond only to columns of numbers with possibly even a (countable) infinity of components, and furthermore in quantum mechanics, these in general are complex numbers. For this reason also, the inner product is of special significance: The square of its absolute value—$|\langle A|B\rangle|^2$—always yields usable real numbers, which, as we shall see, are not yet the measured values, but rather the *probabilities* of obtaining particular outcomes from a given measurement.

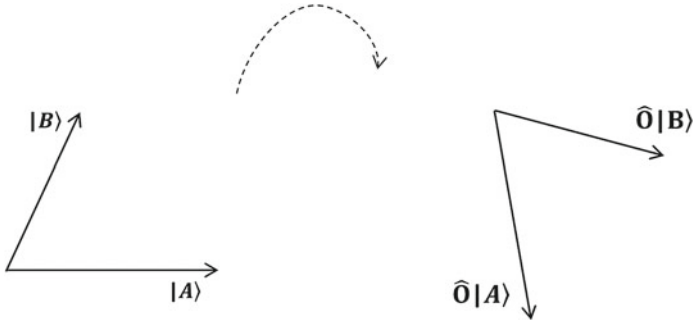


Fig. 1.10 The operator $\hat{O}$ "causes" a rotation of all vectors through a certain angle

In order to arrive at the measured values, we need to introduce one more mathematical object: the *operator*.

1.2.2 Operators and Their Eigenvalues

To start with, we make a didactic remark: This section presents the crucial extension of simple vector algebra to permit its application to quantum mechanics. The intended typical interpretation is suitable for the mathematical description of the experiments in the preceding subchapter, in particular the repeated measurement (cf. Fig. 1.2) and the destruction of the results of a spin measurement (cf. Fig. 1.3). We thus intend to represent a Stern–Gerlach apparatus or an observable quantity such as $\hat{S}_y$ by an *operator*, and the eigenvalues of that operator will then represent the measurement outcomes (or properties), such as spin up along the y-axis. However, not every mathematical operator is suitable for such a representation, since not every operator even possesses eigenvalues; thus we have to begin somewhat more generally, that is more abstractly.

More generally, an operator $\hat{O}$ is a mechanism which maps each vector from a given vector space to one and only one (as a rule different) other vector in that same space. We write: $\hat{O}|A\rangle = |A'\rangle$, and then say that the "application" of $\hat{O}$ to a given vector $|A\rangle$ leads to the vector $|A'\rangle$ in $\mathcal{V}$. In the two-dimensional intuitive space, for example, a rotation through the angle θ around a particular axis of rotation is a (geometrical) operation, which might be the *effect* produced by such an operator. Another example could be the lengthening (or shortening) of each vector by a factor λ.

In principle, there are a very large number of different operators; some of them have rather unfavourable computational properties. A class of operators which is particularly easy to apply are the *linear* operators. Linearity means mathematically that superpositions are conserved in the following sense:

$$\hat{O}\left(\lambda|A\rangle + \mu|B\rangle\right) = \lambda\left(\hat{O}|A\rangle\right) + \mu\left(\hat{O}|B\rangle\right). \tag{1.2}$$

Geometrically, we can say that they conserve *parallelity*, i. e. originally parallel lines (spanned by vectors) remain parallel.[19] Rotations (Fig. 1.10) or parallel shifts are thus examples of linear operations.

In view of our search for a mathematical representation of quantum-mechanical measurement outcomes, a further mathematical property of linear operators is helpful: Linear operators, which are at first abstract mathematical objects, just like the vectors themselves, likewise have—in a given basis—a component representation For this, we require N^2 components, where N is the dimensionality of the vector space. The component representation is namely the following: $\langle A_i|\hat{O}|A_j\rangle$—so that a linear operator can be represented as a square *matrix*. The application of an operator to a vector, i. e. $\hat{O}|A\rangle = |A'\rangle$, can then be computed as the product of its matrix with the vector according to the rule (given here for the two-dimensional case):

$$\begin{pmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} O_{11}a_1 + O_{12}a_2 \\ O_{21}a_1 + O_{22}a_2 \end{pmatrix}. \tag{1.3}$$

The rotation operator, which we again take as a concrete example, has the following matrix representation in two-dimensional space:

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{1.4}$$

so that for an angle of 90°, applied to a basis vector $\binom{1}{0}$, it yields:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.5}$$

As expected, the rotation operator projects the one basis vector precisely onto the other (orthogonal) basis vector in this special case.

Note that the axis of rotation here is perpendicular to the plane of the basis vectors, and thus lies outside the two-dimensional vector space itself. In view of our next step, however, it would be advisable to consider a rotation around an axis within the vector space in question, i. e. perhaps a rotation around the z-axis in the three-dimensional intuitive space. The corresponding matrix then has the following form:

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.6}$$

[19] In every interpretation of the formalism, parallel vectors correspond to *physically indistinguishable* states, so that linear operations bring no physical differences into play where none were present before.

This operator, which thus "causes" a rotation in a three-dimensional space through the angle θ around the z-axis, has a characteristic property: It leaves just one (normalized) vector unchanged, namely that vector which spans the axis of rotation. In general, vectors whose *directions* are conserved are *eigenvectors* of the corresponding operator, and the associated amplitude factors are called their *eigenvalues*. The operator possesses its eigenvectors and eigenvalues independently of the basis used to represent the matrix, so that they characterize the operator uniquely. Our rotation in the three-dimensional space is of course characterized above all by its axis of rotation; the vector which spans this axis is thus an eigenvector of the rotation operator with the eigenvalue 1. But take note that by no means all operators even have eigenvectors (and eigenvalues), as is the rule, for example, for the rotation operator in two-dimensional space, which has no eigenvalues (except for certain angles): It rotates all the vectors in the same way; i. e. it changes the directions of all vectors. Furthermore, in complex vector spaces—as in the case of quantum mechanics—eigenvalues are generally complex numbers, so that they can serve neither in a geometrically intuitive way as extension factors (amplitudes), nor physically as measurement outcomes.

Now, however, there is finally a special subclass of linear operators, so-called *self-adjoint* or *Hermitian* operators , whose matrices (making use of complex conjugation[20]) are *symmetric*. They have, as one can prove mathematically, always the maximum number—namely N—of eigenvalues, which all are also *real*. These real eigenvalues are what will become the mathematical representatives of the quantum-physical measurement outcomes, and indeed nearly independently of the particular interpretation of the mathematical formalism of quantum mechanics which may be proposed. Linear and self-adjoint operators thus represent measuring devices or observable quantities (property types),[21] and their eigenvalues represent the corresponding measurement outcomes, that is experimental data or concrete properties, whereby here in each case the "or" already suggests a certain leeway for interpretations. (In Bohm's interpretation in particular, it will be seen that while spin eigenvalues are observable data objects, they do not represent a property of the microsystem.) Thus, the first task is to solve the *eigenvalue problem*, namely the equation (where $\hat{O}$ is self-adjoint):

$$\hat{O}|A\rangle = \lambda|A\rangle, \tag{1.7}$$

which with certainty has real solutions for λ.

Let us consider a computationally simple example.[22] We wish to determine the eigenvalues of the following linear, self-adjoint operator in an abstract (because

[20] The complex conjugate of $z = a + ib$ is $z^* = a - ib$, and $z^2 = z^* z$.

[21] For single systems, we can say that *every* self-adjoint operator represents some sort of quantum-mechanical observable quantity such as spin, energy, even if it is not always easy to identify the concrete physical realization corresponding to a given mathematical operator. For many-particle systems, however, not every self-adjoint operator corresponds to an observable quantity (keyword: superselection rules; cf. Chap. 3).

[22] The following calculation is intended only to be illustrative. There is a standard mathematical procedure for calculating the eigenvalues and eigenvectors of self-adjoint matrices. More details on the mathematics of physics can be found in Räsch (2011), for those who lack experience in this area.

complex), two-dimensional vector space:

$$\hat{S} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} . \tag{1.8}$$

After inserting into Eq. 1.7, we find a first rearrangement:

$$\begin{pmatrix} -\lambda & -i \\ i & -\lambda \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} . \tag{1.9}$$

Such a system of equations (for a_1 and a_2!) is solvable, as is found in more-or-less elementary mathematics, when the so-called determinant of the matrix is 0. This leads to the following equation for λ:

$$\lambda^2 + i^2 = 0 . \tag{1.10}$$

The given operator thus has the eigenvalues 1 and -1, which, as one could perhaps already expect at this point, are supposed to correspond to the measurement outcomes of spin up and spin down along a certain spatial direction.[23]

The associated eigenvectors[24] are obtained by inserting the eigenvalues into Eq. 1.9; then, normalized to unit length, they are:

$$\begin{pmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{pmatrix} \text{ and } \begin{pmatrix} 1/\sqrt{2} \\ -i/\sqrt{2} \end{pmatrix} . \tag{1.11}$$

The inner product between these eigenvectors is 0 (remember complex conjugation!); i. e., they form an orthonormal basis of the underlying vector space; a result which can be generalized: *Every* linear and self-adjoint operator on a vector space of dimensionality N has on the one hand N (not necessarily different) real eigenvalues and on the other hand (at least) one basis composed of eigenvectors.

If we now represent the operator in a (here: *the*) basis composed of its eigenvectors—which was up to now *not* the case in our example—then its matrix becomes *diagonal*, and its eigenvalues can be found along the diagonal:

$$\hat{S} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{1.12}$$

Thus far, i. e. before we put the operator (or its matrix) into diagonal form, it, and its eigenvectors were represented in a *different* basis (compare Fig. 1.11).

---

[23] For electrons, the spin projection values are of course $\frac{1}{2}$ and $-\frac{1}{2}$, but numerical details are irrelevant here.

[24] Caution: Normally, each eigenvector has exactly one eigenvalue; while conversely, each eigenvalue belongs to *exactly* one eigenvector only when it (the eigenvalue) is *simple*, i. e. it occurs only once—which is in fact the case here. The problem of multiple eigenvalues will be treated in the following Sect. 1.2.3.

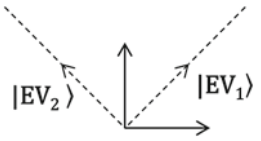


Fig. 1.11 Representation of the eigenvectors of $\hat{S}$ referred to the basis shown in boldface. They themselves form the eigenvector basis of a *different* operator, $\hat{S}'$

From the perspective of vector space theory, it may seem trivial that there are thus *many more* bases than those whose elements are eigenvectors of a given operator. However, from the viewpoint of their quantum-physical meaning, this is quite remarkable, since it fundamentally implies the incommensurability of two observable quantities and the superposition principle. In order to get a feeling for this, we once again choose the basis in which our operator was first represented: Its elements are then *not* eigenvectors of the given operator. However, it is the case that there *must* be *another* linear and self-adjoint operator whose eigenvectors form precisely this basis. Not only does every such operator have (at least) one basis formed from (some of) its eigenvectors, but also for every basis, there is an associated operator for which it is an eigenvector basis.[25] That operator which is associated with our basis must have the diagonal matrix form in that basis, and along the diagonal are its eigenvalues. As we shall see, it is just this matrix:

$$\hat{S}' = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{1.13}$$

It should not be confused with the operator given above (cf. the matrix 1.12), whose (identical) diagonal form occurs in a quite different basis. However, it does naturally have some common features with our original operator: Both of the real eigenvalues are likewise 1 and -1, corresponding again to the two (possible) measurement outcomes spin up and spin down in some particular spatial direction. It is in fact found that these two operators just correspond to our measurements of the spin along the x and the y directions.

Let us now again consider the two two-dimensional coordinate systems, or correspondingly, the two bases consisting of these two sets of eigenvectors: One operator is in each case represented as a diagonal matrix, while the other has a different form. This result can also be generalized: If two (linear and self-adjoint) operators have no *common* basis consisting of (some of) their eigenvectors,[26] then their matrices cannot be simultaneously brought into diagonal form, i. e. not in terms of the same basis (and *vice versa*). This formulation suggests that there are indeed *different* operators which can be simultaneously brought into diagonal form, because they in fact share

[25] Also, this holds for many-particle systems only with some limitations.

[26] Note that two operators can only be *different* if they have some different eigenvectors, because an operator is completely characterized by its eigenvectors. In order to have a common basis of eigenvectors they thus must have (at least) N eigenvectors in common and must have more than N eigenvectors in total, from which (at least) some are different. For more details see the following section.

a common basis of eigenvectors.[27] For any two operators which are currently not represented as diagonal matrices, where therefore the basis has not been chosen to consist of (some of) their eigenvectors, we can thus ask the instructive question as to whether they have *any common basis* of eigenvectors, that is whether their matrices can be simultaneously brought into diagonal form ("diagonalized").

In order to answer this question, we require the following purely computational rule for the multiplication of two matrices (here again in the two-dimensional case):

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} . \tag{1.14}$$

The multiplication of two matrices thus yields a matrix of the same dimensionality. It represents the successive application of two operators,[28] which together act as (another) linear operator: $\hat{A}\hat{B}|\Psi\rangle$. It is then of central importance that the effect of the sequential application of operators may depend on the *order* in which they are applied.[29] The rule for matrix multiplication as defined above is in general namely *not* commutative, as our example shows:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} ;$$

but in contrast[30]:

$$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} . \tag{1.15}$$

It can be shown that the multiplication of two matrices is *commutative* precisely when they can be simultaneously diagonalized, that is when the two operators represented thus have a common basis of eigenvectors. If they in contrast *do not* commute, they cannot be simultaneously diagonalized, because they have no common basis of eigenvectors. Since, as can be shown, the *commutator* of two matrices $\hat{A}$ and $\hat{B}$,

$$[\hat{A}, \hat{B}] = \hat{A} \cdot \hat{B} - \hat{B} \cdot \hat{A} \tag{1.16}$$

is invariant under basis transformations, it is mathematically significant, i. e. the question of whether two operators have a common basis of eigenvectors and thus are simultaneously diagonalizable has an unambiguous answer.[31] Interpreted physically,

---

[27] In the two-dimensional case, this statement is not very rich in content. We will see, however, that there are very interesting examples of this principle in higher-dimensional vector spaces.

[28] Casting a glance backwards: *Sequential* spin measurements are mathematically represented by successive application of (self-adjoint or Hermitian) operators.

[29] Consider, for example, two successive rotations in space around two different rotational axes.

[30] Note that the resulting matrices in this example are *not* Hermitian; i. e., they do not represent observable quantities.

[31] Therefore, in the following, we will speak simply of (non-)commuting *operators*, when the matrices which represent them (do not) commute.

the commutator becomes the most important equation in quantum mechanics: If it is not 0 (or not the null matrix)—as in our example, where it in contrast has the "value"

$$\begin{pmatrix} 0 & -2i \\ -2i & 0 \end{pmatrix}, \tag{1.17}$$

then a measurement of the one quantity destroys the results of a measurement of the other quantity, as shown in the first section for spin operators referred to different spatial axes; and this is independent of the temporal order in which they are applied and of the elapsed time between the two measurements. Realistically, we can choose the interpretation that the quantum-physical system cannot possess the properties corresponding to the two eigenvalues *simultaneously*. Non-commuting matrices (operators) thus form the precise mathematical analogue of incommensurable (or less clearly: complementary) observable quantities. If the commutator of two matrices is 0, in contrast, then the sequential application of the two operators they represent is independent of their order, and the physical measurements they represent do not mutually destroy each other's results. The properties associated with the corresponding eigenvalues can be attributed *simultaneously*, in a realistic interpretation, to the quantum-physical system.[32]

With this formulation, we should be aware of a certain ambiguity: It is uncontested that *if* the "state" of the quantum-physical system is represented by an eigenvector of a particular operator, a measurement must yield the associated eigenvalue *with certainty*. What is in fact disputable is the converse that when a measurement yields a certain eigenvalue of a particular operator, the system is then at least *immediately afterwards* in the state which is represented by a corresponding eigenvector. To assert *this* is equivalent to the assumption of the *eigenvalue–eigenvector link* (which is frequently treated in physics as a given)—namely von Neumann's "projection postulate". In philosophy, this link still is controversial (see for example van Fraassen 1991). Therefore, here we state only what is truly uncontested:

1. Every vector in the Hilbert space[33] is an eigenvector of some linear and self-adjoint ("Hermitian") operator. In relation to such a vector $|A_i\rangle$, the *expectation value* of this operator is identical with the associated eigenvalue[34]:

$$\langle A_i|\hat{O}|A_i\rangle = \lambda_i \,. \tag{1.18}$$

[32] Note that the commutation relation is not transitive: It can happen that A in fact commutes with B, and B with C, but not necessarily A with C.

[33] In quantum mechanics, the unitary vector space of at most countably infinite dimensionality which underlies the computations is denoted as the "Hilbert space".

[34] The expectation value of an operator is mathematically derived from the usual mean value. Geometrically, the vector that results from the application of $\hat{O}$ to $|A\rangle$ is projected back onto $|A\rangle$ Physically, this corresponds (uncontroversially) to just the mean value of numerous measurements of the observable quantity associated with $\hat{O}$.

Since the *variance* around this mean value is zero,[35] we can say that in this state, the measurement outcome corresponding to the eigenvalue λ_i will be obtained *with certainty*, or that the property corresponding to λ_i is present with certainty.

2. Every vector can at the same time be represented as a *superposition* of eigenvectors of a different operator $\hat{O}'$ that does not commute with $\hat{O}$[36]:

$$|A\rangle = \sum_i a_i |B_i\rangle \,. \tag{1.19}$$

Referred to $|A\rangle$, the expectation value of this other operator is not identical with one of its *own* eigenvalues μ_i (and, also, its variance is not zero!):

$$\langle A|\hat{O}'|A\rangle = (\sum_i a_i^* \langle B_i|)\hat{O}'(\sum_i a_i |B_i\rangle) = \sum_i \mu_i |a_i|^2 \,. \tag{1.20}$$

For each vector, there are thus operators corresponding to observable quantities whose measurement outcomes spread. They then do not yield a particular measured value which corresponds to one of the eigenvalues with *certainty*, but only with a certain probability. Every state thus has properties of the quantum-physical system for which it is not guaranteed that they will result from a measurement; there are only certain probabilities for their occurrences.

Thus far, the eigenvalues of a given operator were supposed not only to be all real, but also all different. Each such eigenvalue was associated with precisely one eigenvector. However, this is not necessary, and when an eigenvalue occurs more than once, interesting difficulties arise. The problem of multiple eigenvalues deserves its own section

1.2.3 The Problem of Multiple Eigenvalues

Multiple eigenvalues provide instructive peculiarities which are frequently neglected in the physical-philosophical literature. This correspondingly difficult section is directed in the main at students of philosophy or physics with ambitions to delve more deeply into the field of the philosophy of physics. In particular, it can be skipped over if the interest of the reader or the subject of the seminar is mainly directed towards the chapters on the interpretation of quantum mechanics.

Let us consider the following matrix in terms of the two-dimensional intuitive space:

$$\begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} . \tag{1.21}$$

[35] The variance (here: the standard deviation) is calculated from $\triangle\hat{O} = \langle A|\hat{O}^2|A\rangle - |\langle A|\hat{O}|A\rangle|^2$, which here yields zero.

[36] This representation is not unique; there are infinitely many such different representations.

It is diagonal; that is, it is in a basis consisting of eigenvectors of the operator that it represents, and as a result, it contains N—namely 2—real eigenvalues along its diagonal. In contrast to the matrices we have considered thus far, however, these two eigenvalues are *not* different, which in this two-dimensional example reflects the rather uninteresting fact that this operator, interpreted physically, measures nothing. It is mathematically in fact the operation which simply extends every vector by a factor of 5, so that in this special case, *every* vector is an eigenvector of this operator—it thus does not differentiate in any manner.[37]

If, on the other hand, in spaces of higher dimensionality there are *multiple*, but not all equal eigenvalues—for example two out of three in the three-dimensional case—then there are rather important and interesting consequences. We can say initially that a multiple eigenvalue is always an indication that the corresponding measurement outcome is not yet sufficiently differentiated. Mathematically, a multiple eigenvalue in fact does not belong simply to one (normalized) eigenvector, but rather to a whole *subspace* of the corresponding dimensionality; that is, for a double eigenvalue in a three-dimensional space, the subspace is a two-dimensional plane, within which every vector is an eigenvector with this double value as eigenvalue. Looking ahead, we therefore have already mentioned that every linear and self-adjoint operator has *at least* one basis composed of eigenvectors; we can now make this more precise: It has *exactly* one basis composed of its eigenvectors if all of its eigenvalues are different; and when some eigenvalues occur multiply, it has—not arbitrarily many—but yet infinitely many such different bases. This leeway indicates that in the case of multiple eigenvalues, there is (at least) one more, genuinely different operator which *commutes* with the given operator, and that there is, as a result, (at least) one more meaningful measurement device or observable quantity which can simultaneously be measured, so that—in an appropriate interpretation—the quantum-physical system has (at least) one additional property of a distinct kind at the same time.

In a two-dimensional space, however, apart from the trivial case of a uniform length change, all the operators have different eigenvalues and thus each one has just one basis consisting of its eigenvectors. There are thus no more genuinely different operators which commute with a given operator, which would then both have a common basis of eigenvectors (that could be only *the* one common basis). And therefore, a single electron cannot have two spin values of different types (i. e. along different spatial directions) simultaneously.[38]

Now, in contrast, let us consider the first excited state of the hydrogen atom, or the second energy level of the orbital model, which we all know from chemistry classes. As you may remember, one distinguishes there essentially between four different states, namely a state of spherical symmetry, the s orbital, and three "dumbbells", the p orbitals. They correspond to the eigenvalues of the orbital angular momentum operator $\hat{L}^2$: the s orbital to the eigenvalue 0 and the p orbitals to the eigenvalue 1.

[37] This has as its consequence that in every physical interpretation of the formalism, two vectors which differ only in their lengths correspond to physically indistinguishable states.

[38] As already stated, this last formulation is indeed standard, but nevertheless subject to different interpretations.

The eigenvalue 1 is thus *threefold*; it belongs to a three-dimensional *subspace* which contains all three p orbitals (and still more). There must therefore be still (at least) one more substantial operator which commutes with $\hat{L}^2$ We still lack one simultaneously observable quantity, one more measurement outcome, which (among others) can differentiate among the three p orbitals. Such an operator is (for example) $\hat{L}_z$, the component of orbital angular momentum along the z direction. It has three different eigenvalues, namely 1, -1 and 0 (twice!).

Since $\hat{L}^2$ and $\hat{L}_z$ must commute, they can be simultaneously diagonalized, so that along their diagonals, their eigenvalues are found—according to their multiplicities:

$$\hat{L}^2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \hat{L}_z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{1.22}$$

They have one common basis of eigenvectors, which, represented in intuitive s-pace, correspond to the four orbitals. The quantum-physical system—in a realistic interpretation—thus obtains *two* qualitatively different properties, for example the pair of values $(0, 0)$ for the s orbital or $(1, -1)$ for one of the p orbitals. Note that in fact, it has two more values (or properties), namely an energy eigenvalue (which was already mentioned as the "first excited state"), and a spin value of the bound electron along a particular spatial direction.[39]

The important question then arises as to how many pairs of mutually commuting operators are *at most* present in a given situation; how many observable quantities (or properties) are at most simultaneously guaranteed in a particular quantum-physical system. And it is still more important to constantly remind oneself that this maximal number can never mean *all* in quantum mechanics; that there are always still other operators which *do not* commute with (at least) one of the set of maximal pairs of mutually commuting operators, so that there is always (at least) one more basis of the vector space which is not an eigenvector basis of the given operators—and that therefore (according to the standard interpretation), the quantum-physical system does not always have an observable value corresponding to every type of property for which there is in principle an eigenvalue. In this example, such a quantity is, for instance, the *position*: A bound electron, like that in the hydrogen atom, has no well-defined position!

Therefore, multiple eigenvalues open the path to common observable quantities or the presence of qualitatively different properties. The example shows in addition that one must be very careful: We said that commuting matrices can be simultaneously diagonalized and that the operators they represent thus have (at least) one common basis of eigenvectors. Furthermore, it holds conversely that non-commuting matrices cannot be simultaneously diagonalized and the operators they represent have no

[39] If we wished to take these additional variables into account at the same time, we would have to base our calculations on a space with still more dimensions than the four-dimensional vector space considered here.

common basis consisting of eigenvectors. One could thus think that on the one hand, commuting operators would have *all* of their eigenvectors in common, and that on the other hand, non-commuting operators would have *no* common eigenvectors at all. In fact, this holds only in the case that all of their eigenvalues occur *singly*. *Multiple* eigenvalues imply, in contrast:

1. Commuting operators indeed have a common *basis* of eigenvectors, but by no means *all* of their eigenvectors are shared. Rather, an operator may well have many eigenvectors which are *not* eigenvectors of a certain other operator with which it commutes, "although" their observable quantities are commensurable.
   Thus, for example, $\hat{L}^2$ in the three-dimensional eigenspace of its threefold eigenvalue 1 has elements (i.e. eigenvectors) which instead of being eigenvectors of $\hat{L}_z$ are instead eigenvectors of $\hat{L}_x$, the component of the orbital angular momentum in the x direction, which do *not* commute with $\hat{L}_z$ One could say that the measurement outcome 1 of the operator $\hat{L}^2$ *means something different*, depending on which other commuting operator it is connected with, so that sometimes the properties of quantum-physical systems are *context-dependent*. But that is the subject of a more extended philosophical interpretation.
2. Non-commuting operators indeed have no common *basis* of eigenvectors, but they may well "nevertheless" have some common eigenvectors.
   Thus, for example, the s orbital—due to its spherical symmetry—is *not only* a common eigenvector (or rather its intuitive representation) of the commuting operators $\hat{L}^2$ and $\hat{L}_z$ with the eigenvalue pair (0, 0) The eigenvector which geometrically represents the s orbital is also *in addition* an eigenvector of the quantities $\hat{L}_x$ and $\hat{L}_y$, which are neither commensurable with $\hat{L}_z$ nor with each other—likewise with their (also double) eigenvalue 0 According to the standard reading, there is thus a special state in which the quantum-physical system has not only two values—as would correspond to the maximal[40] number of pairwise commuting operators—but instead *four* qualitatively different observable quantities (properties) at the same time. These are indeed numerically all equal—namely all 0—but qualitatively, they must be distinguished.

These facts—that therefore, commuting operators "nevertheless" may have non-common eigenvectors, and that incommensurable observable quantities "indeed" may have some common eigenvectors—should be kept in mind; in particular when considering the EPR paradox (cf. Chap. 4). The two-particle system that is considered there as a rule has a total spin of 0. However, the associated eigenvector, which is employed over and over—namely the so-called singlet vector—represents a very special state: On the one hand, it is a common eigenvector of the spin projection operators $\hat{S}_x$, $\hat{S}_y$ and $\hat{S}_z$, which are *not mutually* commuting, with the eigenvalues (0, 0, 0). And on the other hand, it is *not* an eigenvector of the single spin operators, which (depending on the spatial direction) *do* commute with these. What follows philosophically from this mathematical peculiarity of the singlet state?

[40] Energy and spin values are left out of consideration here.

1.2.4 Special Operators and the Position Space Representation

With the above, we have sketched out the mathematical formalism of ordinary quantum mechanics of single systems in its basic aspects. As we have already indicated with a view to further developments, the next major difficulty to be overcome arises in the transition to many-particle systems; their mathematical fundamentals will be presented at an appropriate point. In this section, we want to introduce some special operators which play a particular role for certain purposes. They are (with one important exception) all linear and self-adjoint and thus have a maximum number of eigenvalues which are all real and therefore can be interpreted physically as observable values. These operators are "special" in particular through their physical meanings, which we can therefore not leave unmentioned here, but also for their mathematical properties. They are characterized in addition to their linearity and symmetry[41] by (at least) one further, characteristic mathematical property.

These special operators become important in the context of specific problems, such as the statistical operator for the mathematical description of the measurement problem and of many-particle systems. In a first reading of this book, one could therefore skip over this section and then refer back to the corresponding subsections as they are referenced in other chapters. Those who are particularly interested in the problem of the spatiotemporal embedding of quantum objects can skip immediately to "time-evolution operators" and the "position representation".

Projection operators. *Projection operators* have the additional mathematical property of "idempotence", i. e. their repeated application to a previously projected vector has no further effect[42]: $\hat{P}\hat{P} = \hat{P}$ It follows from this that their only eigenvalues are 1 and 0. If $\hat{P}|\Psi\rangle = \lambda|\Psi\rangle$, then $\hat{P}\hat{P}|\Psi\rangle = \hat{P}(\lambda|\Psi\rangle) = \lambda^2|\Psi\rangle$ and thus $\lambda^2|\Psi\rangle = \lambda|\Psi\rangle$. These eigenvalues can be understood as answers to a yes/no question: Is a certain measured value (or a certain property) present or not?

One can therefore associate a projection operator $\hat{P}_{\lambda_i}$ with every eigenvalue λ_i of an arbitrary Hermitian operator $\hat{O}$. If λ_i is a *single* eigenvalue, then this holds also for the eigenvalue 1 of the corresponding projection operator; it thus projects onto the eigenvector $|\Psi_i\rangle$ with eigenvalue λ_i and can then be written as $\hat{P}_{\lambda_i} = |\Psi_i\rangle\langle\Psi_i|$. If the quantum-physical system is associated with this eigenvector, then the *expectation value* of this projection operator is of course 1; for every other vector $|\Psi\rangle = \sum_i c_i|\Psi_i\rangle$, it holds that:

$$\langle\Psi|\hat{P}_{\lambda_i}|\Psi\rangle = |\langle\Psi_i|\Psi\rangle|^2 = |c_i|^2\,. \tag{1.23}$$

[41] Keep complex conjugation in mind!

[42] This property should not be confused with the requirement of repeatability of a measurement.

One says that the expectation value of the projection operator of the eigenvalue λ_i simply gives the *probability* that a measurement of $\hat{O}$ will yield this eigenvalue as its measured result.[43]

When all of the eigenvalues a_i of a given operator $\hat{A}$ are single, then it can be expressed as a sum of projection operators:

$$\hat{A} = \sum_i a_i |A_i\rangle\langle A_i| \,. \tag{1.24}$$

Two operators $\hat{A}$ and $\hat{B}$ then commute if and only if the underlying projection operators $\hat{P}^{\hat{A}}$ and $\hat{P}^{\hat{B}}$ commute pairwise for all of their eigenvalues.

This can be generalized to *multiple* eigenvalues. In such a case, the eigenvalue 1 of the projection operator is also multiple, so that it no longer projects onto just one eigenvector, but instead onto a subspace of corresponding dimensionality. The question of when projection operators commute, if they project onto such multi-dimensional spaces, is non-trivial. The necessary condition is that the intersection set not be empty: Thus, the projection operator of $\hat{S}^2$ with eigenvalue 0 projects onto the singlet vector which lies in the plane onto which the projection operator of $\hat{S}_z$ with eigenvalue 0 also projects. This shows that the commutation relation is not transitive: In contrast to $\hat{S}^2$, $\hat{S}_z$ commutes with the single spin operator for the same spatial direction; the separable product vectors also lie in *its* subspace with eigenvalue 0. However, this condition is not sufficient: The singlet vector also lies in the intersection of the planes of $\hat{S}_x$ and $\hat{S}_y$ with eigenvalue 0, "although" they commute neither with each other nor with $\hat{S}_z$ Subspaces onto which the *commuting* projection operators project must obey certain orthogonality conditions.

The Statistical operator (the Density matrix). Projection operators possess the mathematical property that their eigenvalues are non-negative—namely 0 and 1. In the special case of one-dimensional projection operators,[44] in addition the sum of their eigenvalues is 1. These properties can be generalized as follows: A linear and self-adjoint operator is called "positive" when it has no negative eigenvalues, and it is called a *statistical operator* (or a *density matrix*) when in addition the trace of its matrix[45] is equal to 1. For its eigenvalues p_i, we thus have:

$$0 \leq p_i \leq 1 \qquad \text{and} \qquad \sum_i p_i = 1, \tag{1.25}$$

[43] Note that, as always, the vector $|\Psi\rangle$ is normalized to unit length, so that for its components, we find $\sum_i |c_i|^2 = 1$.

[44] "One-dimensional" refers to projection operators which project onto just one (normalized) vector, whose eigenvalue 1 is thus single.

[45] The "trace" (symbol: Tr) of a matrix is the sum of its diagonal elements. It is invariant under basis transformations, so that with a basis composed of eigenvectors, the trace is just the sum of the eigenvalues.

which justifies its interpretation as a "statistical" operator: Its eigenvalues behave like *probabilities*.

A statistical operator $\hat{\rho}$ can now be written as a weighted sum over projection operators as follows:

$$\hat{\rho} = \sum_i p_i |\Psi_i\rangle\langle\Psi_i| \,. \tag{1.26}$$

In the case that all the p_i are zero except for one, the statistical operator is simply the same as the corresponding projection operator, so that one-dimensional projection operators are special statistical operators. This suggests that statistical operators are correlated with the "state" of the quantum-physical system; in the special case with a "pure state", and in general with a "mixed state". The expectation value of an operator $\hat{O}$ relative to a pure state (cf. Eq. 1.20)[46]:

$$\langle\Psi|\hat{O}|\Psi\rangle = \sum_i \lambda_i |\langle\Psi_i|\Psi\rangle|^2 = \sum_i \lambda_i |c_i|^2 = \mathrm{Tr}\,(\hat{P}_{|\Psi\rangle}\,\hat{O}) \tag{1.27}$$

can be generalized to a mixed state as follows:

$$\langle\hat{O}\rangle_{\hat{\rho}} = \mathrm{Tr}\,(\hat{\rho}\;\hat{O})\,. \tag{1.28}$$

The convergence of pure and mixed states suggested by this, however, is problematic; in any case, it is at odds with the standard interpretation of an operator as an observable quantity. Then namely the statistical operator would represent an observable quantity or a measurement apparatus, with its eigenvalues as possible measurement results. It should be especially noted that in contrast to a pure state, which is correlated with the eigenvector of the corresponding projection operator and its (single) eigenvalue 1, a "mixed state" does not correspond to *any* vector in the Hilbert space. For example, in a two-dimensional space—with the eigenvectors $|up\rangle$ and $|down\rangle$ of a given spin operator—the so-called mixed state

$$\hat{\rho} = \frac{1}{2}|up\rangle\langle up| + \frac{1}{2}|down\rangle\langle down| \tag{1.29}$$

is not to be confused with the *superposition* of "spin up" and "spin down"

$$|\Psi\rangle = \frac{1}{\sqrt{2}}|up\rangle + \frac{1}{\sqrt{2}}|down\rangle\,, \tag{1.30}$$

which corresponds instead to a pure state.

One may have the impression that a "mixed state" is in fact a mixture of (pure) states, and that therefore the probabilities p_i simply correspond to our *subjective lack of knowledge* of the actual pure state. But *this* is certainly not the case: This

[46] With Eq. 1.24, we have $\hat{O} = \sum_i \lambda_i |\Psi_i\rangle\langle\Psi_i|$.

ignorance interpretation of the probabilities which occur as the eigenvalues of the statistical operator is namely inapplicable precisely where it would find its principal usage: for *composite* systems. As we shall see, parts of a whole—as, for example, in the EPR paradox, or in the measurement problem—are described as mixed states, in particular even when the whole can be considered to be in a pure state. One then often says that through neglecting the other parts—for example the measuring device—*information has been lost*. As we will discuss in more detail later, this is perhaps to some extent correct, in that the interference terms in expectation values and measurement probabilities which are characteristic of superpositions no longer occur for mixed states. It would be incorrect, however, to say that a subjective lack of knowledge now plays a role when it was previously not present. For if we consider, e. g., the statistical operator in Eq. 1.29, which comes into play in a similar way when the second particle in the EPR case is neglected, we see that it has a *double* eigenvalue, i. e. the given representation is not unique. This statistical operator has not only two, but in fact infinitely many eigenvectors which all belong to the same eigenvalue $\frac{1}{2}$; this reduces the ignorance interpretation *ad absurdum*.[47]

Unitary operators. The Time-evolution operator. Everything we have said thus far has been completely *timeless*. In all of the questions we have raised—whether an operator has real eigenvalues or not, which concrete values its eigenvalues and eigenvectors have, and whether operators commute pairwise—*time* has played no role. It was quite irrelevant at which time a quantum-physical system was investigated and how it evolves in time. This is indeed rather astounding, since it apparently implies that the mathematical formalism of quantum mechanics leaves a quantity more or less "by the wayside" which was and is of central importance for classical theories, including the theories of relativity: the time evolution and dynamics of physical systems.

If we now seek a mathematical operation which describes the dynamics of quantum-physical systems in Hilbert space, we must evidently require that all the relations described so far—e. g. those between a Hermitian operator and its eigenvectors—remain unchanged by the time evolution. We thus in particular must require that orthonormalized bases remain orthonormalized bases, and that in general the *inner product* of two vectors must be invariant; that is, the lengths of vectors and the angles between them remain unchanged. Operators which conserve the inner product "cause" (imaginary) rotations and are called *unitary*. Mathematically, a *unitary operator* $\hat{U}$ is characterized by

$$\hat{U}^* \hat{U} = \hat{1}\,, \tag{1.31}$$

where $\hat{U}^*$ is derived from $\hat{U}$ by exchanging rows and columns (transposition) accompanied by complex conjugation.[48]

[47] For a discussion of these problems, see in particular van Fraassen (1991) (pp. 157ff. and pp. 206/7).

[48] In the bra-ket notation, $\hat{O}^*$ is always the operator which "acts" to the *left*, so that for unitary operators, we have: $\langle \Psi | \hat{U}^* \hat{U} | \Psi \rangle = \langle \Psi | \Psi \rangle$ – as desired.

We can now show that with a continuous parameter t, a Hermitian operator $\hat{H}$ exists, and it characterizes the quantum system physically, so that the unitary *time-evolution operator* can be expressed as follows.[49]:

$$\hat{U}_t = \exp\left(-\frac{i}{\hbar} t \hat{H}\right), \tag{1.32}$$

and thus:

$$|\Psi\rangle_t = e^{-\frac{i}{\hbar} t \hat{H}} |\Psi\rangle_0 . \tag{1.33}$$

With this, we find a differential equation of first order in the variable *time* for the time evolution of the quantum-physical system. It is the famous *Schrödinger equation* if we choose the operator $\hat{H}$ suitably to represent the observable quantity *energy* ($\hat{H} = \hat{E}_{kinetic} + \hat{E}_{potential}$):

$$i\hbar \frac{d}{dt} |\Psi(t)\rangle = \hat{H} |\Psi(t)\rangle . \tag{1.34}$$

This Schrödinger equation has a *unique mathematical solution* for a given initial value $|\Psi(0)\rangle$, and—keeping complex conjugation in mind—it is invariant under reversal of the direction of the time axis ($t \to -t$). It is thus *deterministic* and *time-reversal invariant*, just like the fundamental equations of classical physics, that is Newtonian mechanics and Maxwellian electrodynamics. It may therefore seem surprising that in our representation—and in contrast to what is stated in many introductory physics textbooks—it is *not* at the centre of the theory or the mathematical formalism of quantum mechanics. Indeed, the commutation relation of Hermitian operators—$[\hat{A}, \hat{B}] = \hat{A} \cdot \hat{B} - \hat{B} \cdot \hat{A}$—has become the central point, and quite rightly so: For we must remember that the unitary time-evolution operator is *not* self-adjoint (only $\hat{H}$ in the exponent of $\hat{U}$ is self-adjoint). Similarly to the rotational matrices in the three-dimensional intuitive space which we have described above (cf. Equation/matrix 1.6), this rotation in a complex space in general does not have a maximal number of eigenvalues, and sometimes none of them are real. Therefore, the time-evolution operator does not correspond to a measurement device or to an observable quantity; the time evolution of the quantum-physical system that it describes cannot be directly observed. It indeed perhaps occurs only in the abstract Hilbert space—whatever that might mean, and although it has an effect on the probabilities of measurement outcomes.

A directly observable time evolution of the quantum-mechanical system rather takes place at most *during* the so-called measurement process. However, a measurement, which is frequently referred to as a "second dynamics", is in the first instance *not* described by the formalism of quantum mechanics—that is the reason for the notorious *measurement problem* of quantum physics, which will often demand our attention in the following chapters. In the meantime, many scientists have adopted the (controversial) position that it is best to do without such a second dynamics, which is

[49]The complex exponential function is periodic, so that $\hat{U}$ is analogous to a rotation matrix. $\hbar$ is the universal quantum-mechanical constant, the (reduced) *Planck's constant*.

in the end not describable mathematically; that is, there is nothing at all special about measurements (e. g. the physicists Ghirardi, Rimini and Weber; see Sect. 2.4). At this point, in any case, we can state: The mathematically describable time evolution of a quantum-mechanical system is not directly observable, and the directly observable dynamics during a measurement is not mathematically describable—a challenge for the interpretation of quantum mechanics!

Position and momentum operators. Both in philosophy as well as in classical physics, not only do the *time*, and changes of perceptible objects in time, play a central role, but so does *space*; their geometric form and their motions through space. Macroscopic objects, such as are dealt with in classical mechanics, occupy at every moment in time a certain spatial region, or as idealized "point-like particles", an exact position in space. Furthermore, at every moment in time they have an observable velocity or a certain momentum, so that they move along trajectories—for example the orbits of the planets—and thus mathematically along continuously differentiable curves in the three-dimensional intuitive space.

In quantum mechanics, this is quite different: Initially, the position and momentum operators are simply two particular operators among many; indeed, they are mathematically rather inelegant—and they are thus mentioned here only in this last subsection on "special operators". The three-dimensional intuitive space also no longer plays its previously distinguished role. Above all, position and momentum are incommensurable quantities in the sense that the more precisely the position of a particle has been determined, the less precisely is its momentum knowable (or *vice versa*): The position operator and the momentum operator do not commute! It follows in the usual way that quantum-mechanical objects do not move along trajectories; their "motion" in space is discontinuous, it occurs as a series of jumps, if they can be considered to be moving objects at all.[50]

Mathematically, the position and the momentum operators are problematic for another reason: If we proceed as we have so far, then for the position operator $\hat{Q}$—and analogously for the momentum operator[51] $\hat{P}$—an eigenvalue equation must hold:

$$\hat{Q}|q\rangle = x|q\rangle\,, \tag{1.35}$$

where "x" denotes the possible position coordinates of the particles. In contrast to what we have thus far described, however, the eigenvalues are *all real numbers*. While we have assumed up to now that the eigenvalues of self-adjoint operators were finite in number, or at least countably many—which corresponds to the idea of "quantum jumps"—the eigenvalues of the position operator are *continuous*, corresponding to

[50] In non-standard interpretations of quantum mechanics, however, one disagrees: Thus, for example, in the GRW variation, the spatial behaviour is eminently important, and in Bohm's mechanics, particles (again) move along trajectories.

[51] Strictly speaking, the position and the momentum operators of single particles naturally each have three components (which mutually commute). We neglect this fact here and in the following.

the intuitively understandable continuum of space.[52] However, this means that the basis consisting of eigenvectors of the operator must likewise be continuous, and this makes it hard to comprehend intuitively as an orthonormalized basis. And such eigenvectors pose problems in the realm of precise mathematics, as well: As we may remember, an eigenvector represented in "its own" eigenvector basis takes on the form of a (possibly infinitely long) column vector in which all the elements are 0 except for one, and this single component is 1. If we now make the transition to the continuum, then a "function" results, whose value is zero everywhere except for one point, where it is infinite—which is really no longer a well-behaved function.[53] And indeed, the Hilbert space of quantum mechanics is *bounded* and *separable*, which, put briefly, means that neither such operators nor such eigenvectors can occur within it; it has at most a countably infinite dimensionality.

Leaving aside such concerns for the moment, one usually continues by generalizing the representation of some arbitrary vector $|\Psi\rangle$ in terms of a discrete basis of eigenvectors of some operator—namely $|\Psi\rangle = \sum_i c_i |\Psi_i\rangle$—to a continuum, by letting the sum become an integral and the expansion coefficients become continuous variables:

$$|\Psi(x)\rangle = \int \psi(x')\delta(x - x')dx' \,. \tag{1.36}$$

This is the "position representation" of a vector in Hilbert space, and thus its most concrete representation, in terms of the eigenvector basis of the position operator.[54] It corresponds to the *wavefunction* in the Schrödinger formulation of wave mechanics.

In the position representation, the position operator "acts" simply like multiplication by real numbers, the position coordinates: $\hat{Q}_x = x$. And the momentum operator acts like the derivative with respect to position[55]: $\hat{P}_x = -i\hbar\frac{\partial}{\partial x}$. Then, a basis-independent commutation relation holds[56]:

---

[52] Since the time of Cantor, one distinguishes (at least) two types of infinite sets: those whose elements can be counted, which are thus no larger than the infinite set of the natural numbers or the set of the rational numbers; and those which are no longer countable, which thus appear to be greater, of higher cardinality, such as the real numbers. A continuum then forms a set of points which are more than countably infinite, and are in addition *dense*. More details can be found in the appropriate textbooks on calculus.

[53] In physics, such quantities are called "δ functions".

[54] Note that this "position space" corresponds to the ordinary three-dimensional intuitive space only for a single-particle system. In many-particle systems, we must in contrast operate in a $3N$-dimensional *configuration space*, corresponding to the particle number N; and it is no longer intuitively comprehensible. But our conclusion remains (for the time being): Only the eigenvalues (here: of the position operator) correspond to real observable values or to real properties of real quantum-physical systems.

[55] In the position representation, the eigenvectors of the momentum operator are found (after a short computation) to be *plane waves*, which could correspond in an intuitive manner to a complete spatial delocalization of the particle accompanying a precise value for its momentum.

[56] In the position representation, the commutation relation is found from $[x, -i\frac{d}{dx}]f(x) = -i(xf'(x)) - \frac{d}{dx}(xf(x)) = if(x)$.

$$[\hat{Q}, \hat{P}] = i\hbar\hat{1}\,, \tag{1.37}$$

from which *Heisenberg's uncertainty relation* between the position and the momentum results:

$$\triangle x \cdot \triangle p_x \geq \hbar/2\,. \tag{1.38}$$

It expresses a relation between the variance of measurement outcomes of the position and the momentum, from which (according to the standard interpretation) it then follows that quantum objects cannot move along trajectories in the usual sense.

With $\hat{E}_{kin} = -(\frac{\hbar^2}{2m})\frac{\partial^2}{\partial x^2}$ and $\hat{E}_{pot} = V(x)$, it follows in general for the time evolution[57] that

$$i\hbar\frac{\partial}{\partial t}\psi(x,t) = -\left(\frac{\hbar^2}{2m}\right)\frac{\partial^2}{\partial x^2}\psi(x,t) + V(x)\psi(x,t)\,, \tag{1.39}$$

which is the (perhaps more familiar) *time-dependent Schrödinger equation* in the position representation.

The problems mentioned for operators with a continuous spectrum of eigenvalues and a continuous basis of eigenvectors are now indeed no longer simply mathematical: If we postulate the eigenvalue–eigenvector link, then the quantum-mechanical system must be in a state following a precise measurement of its position which can be represented by an eigenvector of the position operator, i. e. by a δ function. Due to Heisenberg's uncertainty relation, in this state the momentum is maximally undetermined, which has the result that the δ function immediately spreads out and "melts away". This is verified by the time evolution according to the Schrödinger equation. It is even found that already for a second position measurement immediately after the first, the probability of finding the particle very distant from its originally measured position is, although small, in fact *not zero*. This subverts once again, and indeed definitively, the criterion of the *repeatability* of a measurement. In addition, it also calls into question the philosophically widespread concept of *persistence*, that is of an identity over time: It would appear as though the particle cannot be recognized and is thus not *perseverant*, as Kant would say. There is apparently no temporal identity that survives the change of position coordinates (and other properties). The time evolution of a quantum-physical system again proves to require interpretation to a severe extent.

But with that remark, we have already entered deeply into the philosophical interpretation debate; and many other problems of interpretation were already foreshadowed in the course of this mathematical introduction, so that it is now time for a first chapter on the *interpretations of quantum mechanics*.

[57] It holds furthermore that $\hat{H} = \hat{E}_{kin} + \hat{E}_{pot}$, where $V(x)$ is a potential that depends only on the position, such as the Coulomb potential in a hydrogen atom (and m is the mass of the particle).

Exercises

1. Niels Bohr introduced the concept of "complementarity" into the interpretation of quantum mechanics. Distinguish two readings of how it is to be understood.
2. In sequential spin measurements, we apparently distinguished finally between two effects which could be reversed by mixing the particles. Describe these two ostensible effects and explain why they are in fact only a single effect. What can we deduce from this?
3. Consider the expectation values of operators in regard to whether the physical system is represented by an eigenvector of the given operator or not. Compute the expectation values of the spin operators discussed earlier, relative to the various vectors described there. Explain the results by referring to the figures showing a repeated and a destructive measurement.
4. What does von Neumann's projection postulate state? Explain in particular to what extent this postulate goes beyond what we have considered to be well established in connection with expectation values.
5. In contrast to the general opinion of many philosophical schools of thought, and also of some alternative physical interpretations (e. g. GRW, Bohm), (intuitive or physical) *space* is not at the centre of standard quantum mechanics. Discuss this hypothesis, first informally and then by referring to the particular mathematical characteristics of the position operator.

References

Albert, David Z. 1992. *Quantum Mechanics and Experience*. Cambridge: Harvard University Press.
Aristotle. 1988. *Physics*. Books V–VIII, English translation.
van Fraassen, Bas C. 1991. *Quantum Mechanics. An Empiricist View*. Oxford: Clarendon Press.
Kant, Immanuel (1781/87). *Kritik der reinen Vernunft*. Cited according to the A- and B- editions. English translation: *Critique of Pure Reason* (available online at various sites).
Mellor, Hugh D. 1998. *Real Time II*. London: Routledge.
Nortmann, Ulrich. 2008. *Unscharfe Welt? Was Philosophen über Quantenmechanik wissen möchten*. Darmstadt: Wissenschaftliche Buchgesellschaft.
von Neumann, John (1932). *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer (Reprint 1968). English translation (1996): *Mathematical Foundations of Quantum Mechanics*. Princeton Landmarks in Mathematics.
Räsch, Thoralf (2011). *Mathematik der Physik für Dummies*. Weinheim: Wiley-VCH. English: See e. g. the "For Dummies" series, http://physicsdatabase.com/2012/03/08/for-dummies-series-for-maths/.

Chapter 2
The Measurement Problem. Minimal and Collapse Interpretations

Cord Friebe

Not only in philosophy, but even in physics itself, one depends on *interpretations*. Mathematical formalisms such as the one presented in basic form in the previous chapter are in themselves rather abstract; they say nothing about concrete reality. They require an interpretation, initially in the sense that the mathematical symbols and operations must be associated with elements of physical reality. While, however, in classical physics —that is in Newtonian mechanics, just as in Maxwellian electrodynamics—such an interpretation was fundamentally apparent, considerable difficulties appeared right from the beginning in the case of quantum mechanics. Hilbert space is, in contrast to, for example, the phase space of classical mechanics, a completely abstract vector space, whose vectors and operators cannot be automatically assigned to something in the real world. In quantum mechanics, there is much more freedom for interpretation than in classical physics: a wide spectrum, ranging from constructions which are very close to the usual von Neumann formalism, to far-reaching interpretations whose interventions into the mathematical apparatus are indeed dramatic.

If one tries to proceed systematically, then it is expedient to begin with an interpretation upon which everyone can agree, that is with an instrumentalist *minimal interpretation*. In such an interpretation, Hermitian operators represent macroscopic measurement apparatus, and their eigenvalues indicate the measurement outcomes (pointer positions) which can be observed, while inner products give the probabilities of obtaining particular measured values. With such a formulation, quantum mechanics remains stuck in the macroscopic world and avoids any sort of ontological statement about the (microscopic) quantum-physical system itself. Going one step further, we come to the *ensemble interpretation*: Here, the mathematical symbols indeed refer to microscopic objects, but only to a very large number of such systems. According to this view, quantum mechanics is a kind of statistical theory whose laws are those of large numbers. In regard to a particular system, this interpretation remains agnostic. This is not true of the "*Copenhagen interpretation*": The physicists Niels Bohr and Werner Heisenberg were the first to presume that

C. Friebe et al., *The Philosophy of Quantum Physics*,
https://doi.org/10.1007/978-3-319-78356-7_2

the formalism refers to particular quantum systems. This, however, caused a serious problem, since the question arose as to what would happen to such a system during a measurement. While Bohr remained reticent on this point and avoided discussing the details of the measurement process, Heisenberg emphasized the embedding of the measurement apparatus within an environment containing the observer as an essential element. At this point, the infamous *collapse* of the wavefunction comes into play; however, according to the Copenhagen interpretation, it is either merely methodological, or explicitly epistemological, but in any case not to be understood as ontological. Finally, then, the Copenhagen interpretation remains agnostic, or even anti-realistic, on a crucial point. In the meantime, in particular in quantum philosophy, several *realistic* collapse interpretations are therefore being taken very seriously, for example the one developed in 1986 by the physicists GianCarlo Ghirardi, Alberto Rimini and Tullio Weber. This *GRW theory* is also the first interpretation that we will introduce here that intervenes in the mathematical apparatus, in that it replaces the linear Schrödinger equation by nonlinear temporal dynamics.[1]

Step by step, we will thus become more and more "realistic" in the sense that more and more mathematical symbols and operations will be associated with real processes in the world. The *philosophical* interpretation goes as a rule even further, by asking whether, as a result of the so-called loss of individuality of similar quantum systems, the concept of "substance" in philosophy is not rendered obsolete—whether quantum-physical systems persist at all, i.e. whether they have an identity over time, and how the relationship of a whole to its parts can be determined in the light of the "entanglement" of states. But these problems will be taken up only in later chapters.

2.1 The Minimal Interpretation

We start by taking a look back at classical mechanics: We suppose that there is a system of N particles, each of which has 3 sharply measurable and existing components of position and of momentum. Classical mechanics at this point defines a "state space", namely a phase space of dimensionality $6N$: a set of points whose elements (q, p) can be interpreted directly as the positions and momenta of the N particles. Some sort of forces may act between the particles and can produce accelerations, and the particles have properties such as, for example, a certain kinetic energy or a certain angular momentum. All this has an apparent mathematical correlate: Functions which associate real numbers to the points in phase space—that is to the positions and momenta—correspond to measurable quantities, and the values of the functions—the associated real numbers—correspond to the respective measurement outcomes, which at the same time are considered to be *properties* of the physical system. Thus, for a single free particle, the function $f(\vec{\mathbf{q}}, \vec{\mathbf{p}}) = \frac{1}{2m}|\vec{\mathbf{p}}|^2$ gives its kinetic energy. In general, certain relations hold between the functions, in particular

[1] For overviews of the situation regarding interpretations of QM, see, e.g., Stöckler (2007); Esfeld (2012).

Hamilton's equations of motion, which are equivalent to the well-known Newtonian equation of motion $\vec{\mathbf{F}} = m\vec{\mathbf{a}}$. As mentioned at the outset, philosophy naturally also treats classical objects historically as well as systematically; problems of persistence, of causality, etc., arise not only in relation to quantum-physical systems. However, the particular problem of interpretation—the problem of whether we can associate mathematical symbols with reality at all, is quite simple to resolve classically, so that no philosophical discussions have arisen around it.

In quantum mechanics, the situation is quite different: Its state space, the *Hilbert space*, is considerably more abstract than phase space; its vectors and operators do not correspond in any simple way to elements of physical reality. One could indeed presume that, for example, the vector $|up\rangle_x$ indicates a physical system (a particle) which has the property spin up along the x direction. But as we know, each such vector can be represented in an infinite number of ways as linear combinations of other vectors —what, however, does such a representation signify? Does the particle perhaps possess not only the property spin up along the x direction, but also in addition some sort of superpositions of numerous other spin values? Should we in fact try to interpret such superpositions in a realistic manner? Not every mathematical operation in this vector space is automatically associated with a real correlate! Furthermore, to be consistent, a particle with the property spin up along the y direction should be represented by the vector $|up\rangle_y$, and a measurement of its spin along the y direction—which indeed destroys its spin value along the x direction—should produce a change in the spin value. But what corresponds mathematically to this transition? Do we perhaps need an operator which maps $|up\rangle_x$ onto $|up\rangle_y$ in 50% of the cases (and in the other 50% onto $|down\rangle_y$, i.e. spin down in the y direction)? Such an operator clearly does not exist. Not every process in the real world appears to be representable mathematically!

In the introductory chapter, however, it became clear what we can all agree upon: The *real* eigenvalues of Hermitian operators indeed represent the measurement outcomes in an uncontroversial way, where the term "measured outcome" here initially implies something like the pointer positions of macroscopic measurement apparatus, and not at the same time the *properties* of microscopic quantum-mechanical systems; this already would again be controversial. The *operators* whose eigenvalues represent measurement outcomes correspond to the measurable quantities, where again "measurable quantities" are understood initially simply as macroscopic measurement setups, such as, e.g., a Stern–Gerlach apparatus, and are not at the same time types of properties of microscopic systems such as, e.g., the spin of a particle along a certain spatial direction. This interpretation is accompanied by a serious limitation: In the Hilbert space, there are in fact many more operators than those that can be associated with measurable quantities. There are non-self-adjoint operators whose eigenvalues, if they have any at all, are not all real and can thus not represent measurement outcomes. And there are even nonlinear operators which very certainly do not correspond to anything physical. Such mathematical objects indeed occur in Hilbert spaces, but physically, they represent nothing—in the opinion of essentially all physicists. With this background in mind, it is already a strong hypothesis that now, however, *all* Hermitian operators are presumed to represent some sort of

measurement apparatus, even if one has no concrete information about how a theoretically prescribed operator might be realized in practice.[2] However, if we now assume that a certain Hermitian operator represents a realizable measurement setup, then *all* of its eigenvalues indicate physically possible, measurable values (pointer positions).

A further limitation can be seen immediately: The eigenvalue *equation* $\hat{O}|\Psi\rangle = \lambda|\Psi\rangle$ likewise does not correspond to any physical process; it serves merely to *compute* the eigenvalues and the eigenvectors. This is remarkable, since one could imagine that it would represent mathematically the *measurement process* in which these eigenvalues are observed as measurement outcomes. However, this cannot be the case: If operators are "applied" to vectors in Hilbert space—just as, for example, in the above eigenvalue equation—then this does *not* mean as a rule (perhaps with the exception of the application of the unitary time-evolution operator; cf. Sect. 1.2.4) that thereby some sort of real physical process was described. Let us consider, for example, the following operation:

$$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = i \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \tag{2.1}$$

It is the representation of $\hat{S}_x|up\rangle_y = i|down\rangle_y$ in the eigenvector basis of $\hat{S}_y$. The operator of the observable spin along the x direction is thus applied to the eigenvector with eigenvalue spin up along the y direction and maps it onto the eigenvector with the opposite eigenvalue, spin down along the y direction. Is this supposed to mean that the Stern–Gerlach apparatus in x direction produces a spin flip from spin up (along the y direction) to spin down (along the y direction)? But certainly not: Instead, it has the effect that the system exhibits either spin up or spin down *along the x direction*!

Hermitian operators therefore simply correspond to measurable quantities or measurement setups; their operations in Hilbert space, however, do not represent a *measurement process* in the real world. In addition to their eigenvalues, there is according to the instrumentalist minimal interpretation in fact only one other mathematical operation which refers to reality: the *inner product*, with which the *measurement probability* can be calculated. We assume that the quantum-physical system is correlated with the vector $|\Psi\rangle$, and the observable quantity $\hat{O}$ with eigenvectors $|\Psi_i\rangle$ is to be measured. We then define, using[3] $|\Psi\rangle = \sum_i c_i|\Psi_i\rangle$:

Born's Rule: In the state $|\Psi\rangle$, the *probability* of obtaining the outcome λ_i in a measurement of $\hat{O}$ is given by

$$Prob^{\hat{O}}_{|\Psi\rangle} = |\langle\Psi_i|\Psi\rangle|^2 = |c_i|^2 \quad . \tag{2.2}$$

[2] In fact, this holds for many-particle systems only with certain reservations.

[3] As already mentioned in Sect. 1.2.4, the measurement probability is identical to the expectation value of the corresponding projection operator. We likewise saw there how this can be generalized to mixed states.

This is the generalization of an interpretation of the wavefunction which was given by the theoretical physicist Max Born, i.e. of the vector $|\Psi\rangle$, in the position representation. It can also be seen that this is in fact a convention and not a property which could straightforwardly be read off directly from the formalism. Thus, the function $\Psi(\vec{\mathbf{x}}, t)$ was originally understood—for example by Erwin Schrödinger himself—as a delocalized mass or charge density; this proved to be problematic for many-particle systems, since the function does not actually describe a field in the three-dimensional intuitive space, but rather one in an abstract, generally higher-dimensional configuration space.[4] Born's interpretation of the function being the probability density was then (at least for the time being) generally accepted, with $|\Psi(\vec{\mathbf{x}}, t)|^2$ as the probability with which a position measurement at the time t would indicate the particle to be at the position $\vec{\mathbf{x}}$. We can now generalize and state that *if* the quantum-physical system is represented by the vector $|\Psi\rangle$ at the time t, then a measurement of $\hat{O}$ carried out immediately thereafter would yield the measurement outcome λ_i with the probability $|c_i|^2$. In the special case that the system is already represented *before* the measurement by the corresponding eigenvector $|\Psi_i\rangle$, then λ_i will be measured *with certainty*, and for every other eigenvector of $\hat{O}$, the probability of measuring this eigenvalue λ_i is zero—as could be expected. If the system is in contrast described by a *superposition* of the eigenvectors of $\hat{O}$, then the probability of observing each eigenvalue of $\hat{O}$ is neither 0 nor 1, but rather a precise value between those limits, so that the measurement outcomes will spread. Thus, Born's rule reflects essential features of the general understanding of the mathematical formalism of quantum mechanics.

The equation given above, however, notably permits—like the eigenvalue equation mentioned previously—only the *computation* of the measurement probability. It therefore by no means reflects the real process in which the physical system was projected from its one (initial) state $|\Psi\rangle$ onto another (final) state $|\Psi_i\rangle$. In contrast to the eigenvalue equation, which reflects a real process in *none* of the interpretations, at this point there has repeatedly been the temptation to interpret the geometric projection of the one vector onto the other as the *collapse* of the quantum-physical system. One is tempted also to say—above and beyond Born's rule —that *if* conversely a certain eigenvalue of a particular operator was in fact[5] measured, then the system must be correlated with the corresponding eigenvector at *later* times. However, the minimal interpretation which we have described here is agnostic in this respect; for it, the *vectors* of the Hilbert space have only an *operational* significance as quantities for carrying out computations. In order to be able to apply Born's rule, one must know the appropriate state vector, which is supposed to be "present" immediately *before* the measurement, and thus, one continues to calculate, for simplicity, with the "associated" eigenvector after a previous measurement; between the measurements, it follows the unitary time evolution $|\Psi\rangle_t = e^{-\frac{i}{\hbar}t\hat{H}}|\Psi\rangle_0$.

[4] For a discussion, see the historical articles, collected in Baumann and Sexl (1984).

[5] That is, in a repeatable and irreversible manner?

But these are only computational procedures, which need not be understood as describing real processes. Something real are (perhaps: merely) macroscopic measurement setups and measurement outcomes as well as measurement probabilities: a concept which appears to be excessively minimalistic to the majority of physicists. In particular, such a minimal interpretation says nothing at all about the significance of the measurement probabilities. Do they merely express a subjective lack of knowledge? Or do they refer to objective facts, relative frequencies for example or objective tendencies as dispositional properties of the quantum-physical system itself? Finding definite answers to these questions means that we have to go beyond this minimal interpretation.

2.2 The Ensemble Interpretation and the Copenhagen Interpretation

The first stage of interpretation of the mathematical formalism establishes the connection to the empirical world as far as needed for everyday physics in the laboratory or at the particle collider. Born's rule allows a precise prediction of the probabilities of observing particular outcomes in real, macroscopic measurements. The fact that this minimal interpretation makes statements *only* about macroscopic, empirically directly accessible entities such as measurement setups, particle tracks in detectors or pulses from a microchannel plate may be quite adequate for those who see the goal of the theory within an experimental science such as physics as being simply the ability to provide empirically testable predictions. For the metaphysics of science, this is not sufficient, and most physicists would also prefer to have some idea of what is behind those measurements and observational data, i.e. just how the microscopic world which produces such effects is really structured. In contrast to the instrumentalist minimal interpretation, however, every additional assumption which might lead to a further-reaching interpretation remains controversial.

The principal problem can be immediately recognized if one considers Born's rule somewhat more carefully (cf. also Held 2012):

$$|\langle\Psi_i|\Psi\rangle|^2 \text{ is the probability of obtaining the measurement outcome } \lambda_i \textit{ in a measurement} \text{ of } \hat{O}, \text{ presuming that the system is correlated with the state vector } |\Psi\rangle.$$

The question then arises as to whether or not the reference to a *measurement* at this point is really essential. If not, then Born's rule becomes "simply":

$$|\langle\Psi_i|\Psi\rangle|^2 \text{ is the probability that } \hat{O} \text{ has the value } \lambda_i, \text{ presuming that the system is correlated with } |\Psi\rangle.$$

This latter formulation suggests that the microscopic quantum system *possesses a property* corresponding to the eigenvalue λ_i—indeed, independently of whether that value is measured or not, and in particular, independently of any human observer. Correspondingly, the probability notion expresses a *subjective ignorance*, and as a result, quantum mechanics in the form it has assumed up to now must be *incomplete*. For only in the case that the system was already correlated with an eigenvector of $\hat{O}$ can we say with certainty which property the quantum system has *in fact*. Only then is the expectation value of the given operator sharply determined, and $\langle \Psi_i | \hat{O} | \Psi_i \rangle = \lambda_i$. In all other cases, the epistemic knowledge necessarily lags behind the ontological reality.

Let us once again consider our example in Fig. 1.11, the bisecting line(s) in a two-dimensional coordinate system which is generated by the eigenvectors of the operator $\hat{S}_x$. The angle bisectors themselves are then spanned by eigenvectors of the operator $\hat{S}_y$, which does not commute with $\hat{S}_x$, the measuring setup or property type of the spin along the y direction. Let us now assume that the system is correlated with one of these eigenvectors of $\hat{S}_y$—however we might know that—then it follows that its state relative to the eigenvectors and eigenvalues of $\hat{S}_x$ *superposes*. The inner product (or rather its absolute square) of an eigenvector $|up/down\rangle_y$ with an eigenvector $|up/down\rangle_x$ is in every case $\frac{1}{2}$, so that according to Born's rule, the probability that the system has the property that corresponds to an eigenvalue of $\hat{S}_x$—i.e. spin up or spin down along the x direction—is 50% for each case. Now, in the Hilbert space, no eigenvector of $\hat{S}_y$ coincides with an eigenvector of $\hat{S}_x$, so that within the framework of this formalism, no such vector is conceivable which would be correlated with the system in such a manner that we could predict the measurement outcomes of *both* observables with certainty. In other words, under the assumption that the expression "measurement" adds nothing essential to Born's rule, the microscopic quantum system would have *more properties* than could be predicted *with certainty* using the resources of the mathematical formalism. Ontologically—in the sense that real properties do in fact exist in the world—more would be present than could be determined epistemologically by making use of the vectors in the Hilbert space. Quantum mechanics would then be incomplete. This is the motivation of a number of physicists and philosophers who would prefer to modify standard quantum mechanics in the spirit of David Bohm, so that so-called hidden variables could be presumed to exist. Quantum physics would then be a kind of statistical mechanics with merely epistemological probabilities within an ontologically thoroughly determined and deterministic world (see Sect. 5.1 for Bohm's mechanics).[6]

If, in contrast, it is assumed to be *essential* to add the codicil *in a measurement* to Born's rule, then this could evidently mean that as a rule, the microscopic quantum system would by no means have had the corresponding property λ_i already at a time *before* the measurement of $\hat{O}$; instead, it would have *acquired* it only during

[6]However, one should note that in Bohmian mechanics, only the *positions* of the particles are introduced as additional properties, and not, for example, local spin values for the particles, as one might think at this point.

the measurement. In general, therefore, the quantum system will be *changed* by its interactions with the measurement apparatus, and this furthermore in a probabilistic fashion: A particle which, for example, exhibits spin up along the y direction, and thus(?) is describable by $|up\rangle_y$, acquires either the property spin up along the x direction or else spin down on interacting with $\hat{S}_x$—in either case with 50% probability. The problem is then that we do not actually know of any such interaction within physics: Electromagnetic, gravitational and nuclear forces act *deterministically*. Furthermore, such interactions, produced, for example, by magnetic fields, indeed enter into the continuous and deterministic Schrödinger equation[7]; they are thus already included within the time evolution which a vector in the Hilbert space undergoes when the unitary time-evolution operator is applied to it. But the discontinuous and indeterministic transition from an eigenvalue (and eigenvector?) of $\hat{S}_y$ to an eigenvalue (and eigenvector?) of $\hat{S}_x$ cannot be represented in this manner, since *linear* operators preserve superpositions in the sense of Eq. 1.2. The measurement process must therefore be *different* from all such continuous and deterministic time evolutions and thus from all those interactions which enter into the Schrödinger equation in such a way. The question is merely, *How* is it different?

Are there physical criteria for distinguishing a measurement from some other arbitrary interaction? We have already seen that; e.g., "repeatability" and "irreversibility" are highly problematic as candidates for such criteria: the latter presumes from the outset that there is a fundamental difference—still to be explained—between the microscopic and the macroscopic worlds, since irreversibility cannot occur in the quantum realm (the Schrödinger equation is invariant under time reversal). And the former has been found to be somewhat vague: What can be allowed to happen between two measurements of the same kind and what is not allowed, if the two measurements are still to be considered to follow each other *immediately*, and thus, the second counts as a repetition of the first? If, beyond what we have said so far, we treat the special case of a position measurement, then we must even insist that position measurements would not be measurements at all if they were required to be repeatable; for, as we have shown, the position operator in particular does not commute with the momentum operator, so that from Heisenberg's uncertainty relation, it follows not only that particles like electrons do not move along classical trajectories, but also that following a precise measurement of their positions—due to the resulting large uncertainty in their momenta —the correlated wavefunction (i.e. the corresponding eigenvector in the position representation) disperses very rapidly: every position measurement which follows immediately, even after a very short time, no longer leads with certainty to the same position as before (nor even a neighbouring position). There remains a certain (although small) probability for detecting the particle at a considerable distance from its previous position. Position measurements are thus *in principle* not repeatable.

Then, however, there is apparently no criterion for just what characterizes a measurement *physically*. The quantum-mechanical concept of measurement evidently

[7] namely within the potential $V(x)$ in $\hat{H} = -(\frac{\hbar^2}{2m})\frac{\partial^2}{\partial x^2} + V(x)$

requires a peculiar reference to a (non-physical, subjective) *observer* (see (Held 2012), p. 77). Only such a subject, who is capable of registering the pointer positions (or the like), seems to make it possible to distinguish the measurement process reliably from other interactions. Physics, which attempts to objectivize every knowledge, would thus have arrived at an insurmountable barrier, that is seen by many as an extremely unsatisfactory situation. However that may be: In the Hilbert space, in any case, such a measurement process does *not* occur, so that every interpretation which deems the reference to a measurement in Born's rule to be essential must then assume that there is an additional and primarily indeterministic dynamics, which is in the final analysis not described by standard quantum mechanics.

2.2.1 The Ensemble Interpretation

All of this would seem to become relatively harmless if one does not try to apply the realistic interpretation to particular systems, but instead only to a sufficiently large number of such systems, as is done in the *Ensemble Interpretation*.[8] A majority of physicists would no doubt accept this interpretation. It goes beyond the minimal interpretation in particular by giving a certain special meaning to the concept of probability as it occurs in Born's rule—just like Bohm's theory, as mentioned above. While according to Bohm, however, standard quantum mechanics is incomplete; that is, there are more properties of quantum-physical systems in the world than can be determined by the methods of the usual Hilbert-space formalism, so that the inevitable probability statements are *epistemological* in nature—in the ensemble interpretation, in contrast, the probabilities take on an *ontological* significance. They are now *relative frequencies* of occurrence.[9]

Casting a look back at the spin experiments discussed in the introductory chapter, and choosing, for example, the sequence $\hat{S}_x \hat{S}_y \hat{S}_x$—that is initially a spin measurement along the x direction, then following it a second measurement along the y direction, and finally another spin measurement along the x direction: All of the particles (electrons or silver atoms) which showed spin up (or spin down) along the x-axis after $\hat{S}_x$ should in the following $\hat{S}_y$ measurement show 50% of spin up and 50% of spin down along the y-axis. Finally, upon a renewed $\hat{S}_x$ measurement, half will show spin up and half spin down along the x-axis, so that the result of the first $\hat{S}_x$ will have been destroyed. As long as we consider a large number of particles, this result

[8] According to the ensemble interpretation, the state vector $|\Psi\rangle$ describes a large number of similarly prepared systems. Independently of this, and distinct from it, is the question of whether $|\Psi\rangle$ describes single- or many-particle systems. The version of the ensemble interpretation which has been most thoroughly worked out can be found in ((Ballentine 1998), Chap. 9).

[9] Caution: This contrast to Bohm's theory exists only when one holds quantum mechanics to be *complete* and at the same time does *not* wish to apply it to particular systems. In a certain sense, every type of Bohm's theory is an ensemble interpretation in its statistical part; we are, however, not referring to this when we speak of the "ensemble interpretation" here.

would appear to be relatively unproblematic: Out of perhaps 1 million electrons which exhibited spin up (or spin down) after the first $\hat{S}_x$, one-half, i.e. ca. 500000, give spin up along the y-axis in the following $\hat{S}_y$ measurement, etc. The number of positive results divided by the total number of particles is about $\frac{1}{2}$, that is a relative frequency of 50%, which is supposed to be the objective quantity in the world that is represented by the probability concept in Born's rule.

In fact, this interpretation is only *relatively* unproblematic; as we can see, its conclusions are only roughly valid: Only *roughly* half of all the particles exhibit the result sought, and not *exactly* half. One could of course presume that such a possible imprecision could be controlled by applying the law of large numbers. With an increasingly larger number of particles, the fraction of positive results would approach the value of 50% more and more closely, so that to a good approximation, the exact value predicted by quantum mechanics would finally be obtained. But, strictly speaking, even this conclusion holds only with a *very high probability*, but never with *certainty*. It is not impossible that in a real experiment, markedly different values would be obtained "by chance"; in an extreme case, it could happen that the first 1 million particles *all* indicated spin up along y-axis. One would then have to wait for a considerable time before the calculated result would be obtained *on average*. What is important here: In the interpretation of probability as objective in the sense of relative frequencies of occurrence, we have obviously not gotten rid of the reference to probability— which we should however do, since the intention here was to assert that probabilities are *nothing other* than just relative frequencies. However, they recur on the next-higher level in the form of that "highly probable" over and over again and thus cannot be reduced to simply factual circumstances, as desired. In philosophy, still more objections of this kind are taken very seriously as arguments against the interpretation of objective probabilities as relative frequencies (see, e.g., (Rosenthal 2003)), so that we must say that the ensemble interpretation, regarded philosophically, is rather unsatisfactory.

But also among physicists, many find this interpretation to be dissatisfying: First of all, one finds that the interpretation neither describes the measurement process, nor does it explain why precisely the observed relative frequencies occur. That might be acceptable as long as the ensemble interpretation is considered to be *preliminary*. In this view, quantum mechanics is not a fundamental theory, but is analogous to phenomenological thermodynamics, with the Schrödinger equation thus analogous to the ideal gas law; it does not hold for particular atoms or molecules. We still have to wait for the actual fundamental theory with which the relative frequencies can be explained. However, it threatens to become completely unacceptable if we expect the ensemble interpretation to be an interpretation of quantum mechanics *as a fundamental theory*. Then, one has to consider it to be a serious defect when such a physical theory describes only a large number of similar systems and takes a completely agnostic position with regard to particular systems. Correspondingly, it would be quite absurd to even ask the question as to which property—whether spin up or spin down along the x or the y direction—is possessed by a single arbitrarily chosen electron (or silver atom). Quantum mechanics, so understood, would appear to be in a much poorer position than the rest of physics, which indeed makes statements about

particular systems. Furthermore, this would also hold true when, for example, in a "repeated measurement"—i.e. after a measurement of $\hat{S}_x \hat{S}_x$—all the particles which previously were found to have spin up along the x-axis would again exhibit spin up in the x direction. One might think that it should be possible at least in this case to assert that each particular particle has the property spin up along x-axis—and not just the whole ensemble to the extent of 100%. If the ensemble interpretation, however, were to accept such a statement, it could no longer so readily reject as meaningless the question of what happens to this single particle —which presently exhibits spin up along the x-axis— when it later passes through a Stern–Gerlach apparatus oriented in the y direction. The consequence is that the ensemble interpretation of a complete and fundamental quantum mechanics in fact refers strictly only to a large number of microsystems. Ontologically, it takes the position that there are only ensembles in the world, and no particular systems, of which in fact such ensembles must consist. And this is in the end not a very convincing position.

2.2.2 The Copenhagen Interpretation(s)

Let us therefore consider a first interpretation, which focuses its attention onto particular microscopic quantum systems and insists upon the completeness of the Hilbert-space formalism (and thus upon the *objective* character of the probabilities in Born's rule). This is the *Copenhagen interpretation*. It dates back to the pioneering work of the quantum theoreticians Niels Bohr and Werner Heisenberg, and for a long time, it was considered to be *the* standard view of the physics community. It is, however, not quite clear just precisely what this interpretation states, so that our treatment here is itself an interpretation.

In the literature, there is controversy over whether the "Copenhagen interpretation" is even a unified point of view at all. In particular, a conflict between Bohr's and Heisenberg's views themselves has been emphasized[10]; it is supposed to consist in the fact that Heisenberg, but not Bohr, assumed a "second dynamics" of the measurement process, consisting of the notorious observer-induced "collapse of the wavefunction". In contrast, Bohr avoided considering the details of the measurement process at all; he set the limits of what could be known and explained more narrowly than Heisenberg did. Bohr thus maintained a sort of non-instrumentalist minimal interpretation and therefore an interpretation in which vectors and operators could indeed be associated with real properties of particular quantum-mechanical systems, but it remains open how the definite values come about. With an intentional reference to von Neumann's theory of measurement,[11] Heisenberg, in a (later) article in 1959 with the eponymous title "The Copenhagen Interpretation of Quantum Theory", established the concept that no isolated process can be a measurement, but instead only processes which are embedded in an environment to which the observer belongs

[10] See, e.g., Faye (2008).

[11] For this theory, see the section below on the "measurement problem".

in an essential way.[12] It is this variant which in the following will be discussed as the actual Copenhagen interpretation—with the didactic goal of dissociating it from realistic collapse interpretations.

The following three hypotheses appear to have played a central role for Bohr:

1. There is an ineluctable relationship between the microsystem and the measurement apparatus.
2. All experiments must be described in the language of classical physics.
3. "Complementarity" holds between spatiotemporal and causal descriptions, or among the descriptions of various experimental arrangements of non-commuting operators.

In contrast to the instrumentalist minimal interpretation on the one hand, and to the non-Bohmian ensemble interpretation on the other hand, this Copenhagen interpretation imputes a certain reality to the particular quantum-mechanical systems. However, their *independent* reality is contested, since the macroscopic measurement apparatus participates in the production of the quantum phenomena. It is then problematic as to just *how* definite measurement outcomes come about: Why, then, does not the entangled composed system remain in a superposition, as we will show in more detail below? For the dependence of the quantum-mechanical microsystem on the classical macrosystem (measurement setup), as maintained by Bohr, is indeed not to be confused with the recognized necessity of explaining classical phenomena on the basis of quantum phenomena. Whoever insists (cf. (Esfeld 2012), p. 89) that the macroscopic systems which surround us have well-defined positions *and* are composed of microscopic quantum systems, which as a rule do not have such well-defined positions, has the problem of explaining how this classical world emerges from such a quantum world—an apparently quite general question. According to Bohr, however, there are no quantum systems which exist independently of macroscopic systems and which make up those macroscopic objects around us. They—the quantum systems—exist and possess their properties on the contrary only in relation to a certain experimental setup and therefore only in relation to something macroscopic, which indeed for its own part, independently of the quantum systems, must have been real all along. The reference to the *measurement*, which according to Bohr is an essential part of Born's rule, evidently implies ontologically that the measured quantum systems and their properties are *not fundamental*, but instead depend ontologically on the macroscopic measurement apparatus. The measurement apparatus appears to create *them*, and not the converse: that the quantum systems produce the measurement apparatus which is composed of them.

Do we not overshoot our goal with this assertion? It does not seem to be in Bohr's sense, in any case. However, consider hypothetically a particular electron, whose spin component in the y direction is to be measured, and furthermore, we assume that Born's rule again yields a probability of 50% that spin up along the

[12] "For the measurement setup is worthy of that name only when it is in intimate contact with the rest of the world, when there is a physical interaction between the measurement setup and the observer" ((Heisenberg 1959), p. 41).

y-axis will be observed. According to Bohr, it cannot be the case that this electron possesses the corresponding property *already before* the $\hat{S}_y$ measurement, since then the information contained in the (from wherever) given state vector, and which enters into Born's rule, would be incomplete. It allows in fact only the prediction that the spin up property will be found with a probability of 50%. Attempting to apply Born's rule to particular quantum systems and at the same time maintaining that the reference to the measurement in the rule is essential—both of which are apparently done also by Bohr— leads inevitably to the conclusion that the particular particles do not by any means already have the relevant property, but rather *acquire* it only during or as a result of the measurement (cf. (Held 2012), p. 83), and indeed in an indeterministic manner. So far, so good; that is, so far, so more-or-less uncontroversial as an analysis of the Copenhagen interpretation. But does it follow from this that independently of the measurement, there exists *no* quantum system at all, as we have formulated it above?

One can namely object to this statement by pointing out that it does indeed depend on the measurement setup which *properties* are attributable to a microscopic quantum system and that, therefore, a particular quantum system *changes* during or even in virtue of a measurement in a characteristic manner, but that not necessarily the very *existence* of the quantum system is therefore dependent on the macroscopic measurement apparatus. And it would appear to be evidently correct that the existence of the particular electron does not depend on *this* particular $\hat{S}_y$ measurement which was carried out here and now; after all, it was previously—for example by means of an earlier $\hat{S}_x$ measurement—found to be in a state with spin up (or spin down) along the x-axis and thus to have that property. As we see from this example, however, we cannot so easily get rid of the reference to *some sort of* previous measurement—so that it becomes more plausible that every variant of the Copenhagen interpretation in fact implies the assertion that the microscopic quantum system depends *ontologically*—i.e. in terms of its reality status—on the macroscopic world.

What would be the alternative? We could at this point recall Aristotle, who distinguished between Socrates' being human as his *essential* property on the one hand, and his having snub nose as a merely *accidental* property on the other. Then, we could point out that quantum mechanics also distinguishes between properties which, for example, an electron must possess *in principle* and temporally varying state-dependent properties. Its charge, its mass and the fact that its spin component can have only one of two values are in this view essential properties of an electron—without them, no electron exists. Certain values of momentum, spin projection or kinetic energy are in contrast properties that are merely accidental and can thus vary with time and sometimes are not even present in quantum mechanics. The electron, defined in terms of its essential properties, charge, mass, and spin, is then a quantum system which has existed from the beginning, while in a measurement, "merely" its state-dependent properties are created. However, as will be shown in the chapter on the indistinguishability of similar quantum particles, electrons can neither be enumerated nor recognized through their "essential properties" alone; they are neither particulars nor are they individuals, characterized by state-independent properties alone. Then, however, the presumed bearer of the accidental properties by no means

fulfils its traditional functions. Furthermore, such a bearer is, considered by itself, completely counterintuitive, indeed anti-empirical: for an electron which is determined merely by its charge, mass and spin is neither spatially delocalized—as it is, rather, when its momentum is sharply determined—nor spatially localized—as it is, rather, when its position is sharply determined. Consequently, independently of its accidental properties, it is nothing at all spatially, which renders this interpretation rather unconvincing.

What is important here: If, with Bohr, we leave the details of the apparently so essential measurement process out of consideration, then the reality status of the quantum system in the end remains unclear. Furthermore, many interesting questions remain unanswered which arise in relation to a particular quantum-mechanical system that should be representable by mathematical symbols. Therefore, in the sense of the Copenhagen interpretation, Heisenberg goes beyond Bohr's explanation boundary and draws the consequence explicitly that the eigenvalues of Hermitian operators represent real properties of particular quantum systems. This evidently leads, together with the simultaneous assumption of the completeness of quantum mechanics and thus of an essential role of the measurement in Born's rule, to the conclusion that such properties are changed or newly created in a discontinuous and non-deterministic manner during or by the *measurement* itself; for composite and "entangled" systems, as in the EPR case, even over large distances. Going beyond Born's rule, Heisenberg in addition advances von Neumann's *projection postulate*, namely the eigenvalue–eigenvector link, so that temporally immediately *after* a measurement, the quantum-mechanical system is correlated with the eigenvector which "belongs" to the observed eigenvalue.[13] According to Heisenberg's Copenhagen interpretation, there are namely two temporal dynamics:

1. A continuous, deterministic and temporally reversible unitary time evolution of the vector $|\Psi\rangle$ governed by the Schrödinger equation, $i\hbar\frac{d}{dt}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle$.

2. A discontinuous, non-deterministic and temporally irreversible time evolution of $|\Psi\rangle$ into an eigenvector $|\Psi_i\rangle$ on measurement of $\hat{O}$ with the measurement outcome λ_i, i.e. a projection or a *collapse* with the *objective* probability $|\langle\Psi_i|\Psi\rangle|^2$ according to Born's rule.

But just as the Schrödinger time evolution occurs only in the abstract Hilbert space—since the time-evolution operator is not Hermitian—the mathematically unspecified collapse is also not a real process in the world, according to Heisenberg.[14] That is namely contradicted by the fact that the position representation of $|\Psi\rangle$ is not given in the three-dimensional intuitive or physical space, but rather in the abstract $3N$-dimensional configuration space. In addition, it is apparently contradicted by the

[13] Caution in the case of multiple eigenvalues, where (possibly several) other measurable quantities must first be determined.

[14] While that "which happens in an atomic process", namely the change of eigenvalues as properties, is supposed to be physical, for the state vector it is asserted that "the discontinuous change in the probability function occurs to be sure through the act of observation; for here, we are dealing with the discontinuous change of our knowledge at the moment of observation" ((Heisenberg 1959), p. 38). $|\Psi\rangle$ is thus construed epistemologically.

fact that the presumably spatial collapse of the wavefunction would have to occur instantaneously over large distances, which of course contradicts the principles of special relativity.[15]

Therefore, Heisenberg's Copenhagen interpretation, for its part, is now faced with a whole series of difficulties, which cause many physicists and philosophers to lean towards a realistic conception of the collapse, as maintained, for example, by the GRW theory which will be introduced later. The first problem concerns the status of the probability concept in Born's rule: Due to the essential reference to measurements—and thus[16] to the completeness of the formalism—an interpretation of the probabilities as subjective degrees of belief can be eliminated for all intents and purposes. The interpretation as objective relative frequencies of occurrence—as applied in the ensemble interpretation—can also not be considered, since the Copenhagen interpretation expressly intends to apply the probability statements of quantum mechanics to particular systems. As a result, they can be objectively only something like *propensities*, i.e. objective tendencies as real properties of the quantum system. However, this could, considered by itself, be regarded as all too metaphysical, since modal properties which are interpreted in such a realistic manner can no longer be distinguished from non-modal properties. In addition, one can ask towards what in the world these propensities tend, if the resulting collapse is in any case not real or will not be real, as it is to be understood as merely epistemological.[17]

The principal difficulty with Heisenberg's interpretation lies however elsewhere, given the general unease with it: In the spirit of the formalism of quantum mechanics, in the (indeed essential) measurement process, the observed quantum system and the macroscopic system that carries out the measurement must form a composite whole (cf. below, Sect. 2.3.1). If then the quantum system is represented by a superposition $|\Psi\rangle = \sum_i c_i |\Psi_i\rangle$ of eigenvectors of the observable $\hat{O}$ which is to be measured, then after its interaction with the measurement apparatus, the overall system is linked to a superposition. What we measure, however, is a sharply defined value which corresponds to a particular eigenvalue λ_i of the operator $\hat{O}$. This can apparently only be included within the Copenhagen interpretation if it carries out a so-called *Heisenberg cut*, i.e. determines an (objective?) boundary between the object measured and the measurement system. As is supposed to be shown by the popular thought experiments on Schrödinger's cat and Wigner's friend (cf. (Audretsch 2002; Baumann and Sexl 1984), Chap. 7), however, it is obviously arbitrary where precisely the cut lies: between Wigner's friend and the cat, or already between the cat and the

[15] Take note, however, of Maudlin's Lorentz-invariant realistic collapse interpretation, according to which the wavefunctions are hyperplane-dependent in four-dimensional spacetime; cf. ((Maudlin 1994), Chap. 7).

[16] Permitting ourselves a preview of the GRW theory, we note here that in that theory, Born's rule does *not* hold: In that theory, a (modified) quantum mechanics is *complete*, and there is "nevertheless" no measurement process—which is excluded *with* Born's rule.

[17] The currently favoured Bayesian interpretation of quantum mechanics is consistent, because it is consequently epistemological; see, for example, Fuchs and Peres (2000) and the critical balance given in Friederich (2011).

radioactive substance? Between the microscopic and the macroscopic system, one would perhaps say—but how many particles must it include for a system to become a macrosystem? The answer that a system is a macro (micro)-system when it follows the laws of classical (quantum) physics would in any case appear to be circular. No, unquestionably: There is *no physical* criterion which allows us to determine where the cut between the measured and the measuring systems must be made. The laws of physics and in particular the standard formalism of quantum mechanics give us no indication of where the distinction must lie.[18]

Now, however, we have already discussed at length the fact that there is no *physical* criterion such as, e.g., "repeatability" which can distinguish a measurement process physically from other (continuous and deterministic) interactions. And we have already noted there that a reliable (apparently non-physical!), strictly distinguishing criterion brings a conscious, perceiving observer into play, who reads off the results of the measurements.[19] Could we thus not argue along with Heisenberg that precisely at this point, physics or the mathematical expressibility of nature reaches its boundary, which perhaps forms a Kantian Subject? Could we not defend the Copenhagen interpretation (namely Heisenberg *and* Bohr) with Kant and say that nature has two sides: an appearance side that is experimentally, empirically accessible and physically mathematically understandable; and a way in which nature as a whole is *in itself* which cannot be captured mathematically and which ultimately refers to the Subject?

Thus far, however, it is not necessary to accept that there is a limit to what can be described mathematically, as will be seen from further developments, to be described below. The basis for any solution or dissolution of the measurement problem is today the so-called *decoherence*, which has finally taken an important step beyond the Copenhagen interpretation. But the question of how we in the end arrive at definite measurement outcomes cannot be answered by this programme either, so that the essential problem of interpretation still persists. Following a brief discussion of the decoherence programme, we next move on to *realistic* collapse interpretations, taking as an example the GRW theory, first published in 1986. Non-collapse interpretations are then treated in a later chapter.

2.3 The Measurement Problem and Decoherence

Many standard textbooks on physics, and probably all popular-scientific treatments of the philosophy of quantum physics, come to a stop at the Copenhagen

[18] Within the formalism of quantum mechanics, one can even demonstrate mathematically that under the assumption *that* a (real) collapse takes place, one can in principle not determine *when* it occurs, and thus where the (objective) cut must lie; see (Albert (1992), p. 91).

[19] Such formulations often produce abstruse misunderstandings: It is naturally not meant here that the "subject-independent", objective world is in a superposition so long and to such an extent (that is as a whole) until a transcendental ego "views" it from "outside".

interpretation.[20] That interpretation, however, is by no means still standard within the professional philosophy of science community, so that even an introductory book on the philosophy of quantum physics should go beyond the Copenhagen interpretation.

Let us recapitulate here the alleged flaws of that interpretation once more in different terms: If we place (a generalized version of) Born's rule—that is the probability interpretation of the vectors in Hilbert space, or the identification of their inner products as (conditional) probabilities—at the centre of the interpretation problem, then we can discern a conflict between *completeness* and *realism*. If it is *not* essential to include "in a measurement" within Born's rule, then we can apply a realistic interpretation which attributes all of its possible properties at all times to the quantum-physical system and presumes that the system evolves deterministically over time. Then, however, standard quantum mechanics must be incomplete, since there is evidently ontologically more at hand than can be determined epistemologically, and the formalism must be extended to include *hidden variables*. This, however, is not so simple, since following von Neumann (1932), via Bell (1964) and Kochen-Specker (1967) (see Chaps. 4 and 7), again and again so-called *No-go* theorems have been proposed. We shall see to what extent the de Broglie–Bohm theory can resist these theorems (cf. Sect. 5.1).

If, on the other hand, in Born's rule the codicil "in a measurement" is taken to be *essential*, we can maintain the *completeness* of standard quantum mechanics. However, if no more properties are ontologically present than can be epistemologically predicted *with certainty*, then clearly a *second dynamics* is needed, which in contrast to the first, i.e. the Schrödinger dynamics, is discontinuous, indeterministic, temporally irreversible, and not mathematically describable: the measurement process. Thus, with von Neumann's projection postulate, one has to accept the collapse of the state vector, which, however, according to Heisenberg and his followers occurs not in the world, but rather is supposed to take place only abstractly or epistemologically. *No ontological significance* is attributed to the vectors in Hilbert space.

One should emphasize, to be sure, that Heisenberg's Copenhagen interpretation is after all also realistic to the extent that it clearly attributes objective properties to particular microscopic systems, which they acquire, however, as a rule only during a measurement. If a quantum-mechanical system is correlated with an eigenvalue of a certain Hermitian operator, then that eigenvalue is given a realistic interpretation as an objective property of the particular object. More is not necessary in order to judge quantum mechanics to be complete—and thus epistemically on a par with other physical theories.[21] An undesirable anti-realism in Heisenberg's Copenhagen

[20] Everett's many-worlds interpretation is to be sure also popular; it lies mathematically very close to the standard formalism and is, in contrast to (Heisenberg's) Copenhagen interpretation, *realistic*. For Everett's theory, see Sect. 5.2.

[21] That the eigenvectors—in contrast to the eigenvalues—are not given a reality status is also not really a defect of this version of the Copenhagen interpretation, since in mathematical theories, such as in particular the Hilbert space formalism, there are many more symbols and operations which have been interpreted as at most instrumental, without anyone complaining.

interpretation is instead seen in the fact that the dependence on the measurement process is in the end a dependence on the *observer* or subject. Since a (discontinuous, indeterministic and irreversible) measurement cannot be identified by mathematical physical means as the special interaction that it must be, one must finally fall back on the observer who reads off the results of the measurements. In this subject dependence, an undesirable anti-realism is seen, because apparently only that can be *real* which is *independent of a subject*. However, this too is not automatically the case, or not automatically a defect of the theory, since being *dependent* on a subject must not immediately mean being *subjective*. Thus, Heisenberg and others were by no means proponents of subjectivism.[22] The principal problem is therefore evident that this subject dependence cannot be described by mathematical physics.[23] Getting rid of it, while maintaining the completeness of quantum mechanics, requires designing a *unified* temporal dynamics and thus setting up an alternative to the Schrödinger equation, so that the (macroscopic) measurement process could be made obsolete as the second dynamics. This was the goal of GianCarlo Ghirardi, Alberto Rimini and Tullio Weber, who claimed to have dissolved the notorious measurement problem with their theory, known in the meantime as the *GRW theory*. It presents an alternative to standard quantum mechanics.

2.3.1 The Quantum-Mechanical Measurement Problem

Let us first take a more precise look at the problem of the measurement process, which we wish to (dis)solve: According to Maudlin (1995), it essentially consists in the following trilemma:

1. Quantum mechanics is *complete*; i.e., the vector $|\Psi\rangle$ with which the quantum-mechanical system is correlated determines all the objective properties of the system being considered.
2. Vectors in Hilbert space are *always* subject to a *linear* temporal dynamics, namely (e.g.) that provided by the Schrödinger equation.
3. Measurements yield definite, well-defined results. Following a measurement, the apparatus thus indicates *exactly one* of the possible outcomes which are given by the eigenvalues of the corresponding operator.

We can see that the conjunction of these three assertions is *inconsistent*, so that a solution to the measurement problem must require refuting (at least) one of the three statements. Rejecting the first assertion implies, as indicated, that hidden variables

[22] Especially since the question of *which* properties the quantum-physical system acquires in a concrete case does not depend on the observer.

[23] Which in turn is a common feature with Bohr's Copenhagen interpretation, in the sense that Bohr avoids a description of the (physical) measurement process in terms of mathematical physics.

must be introduced, as in Bohm's mechanics.[24] The third assertion can be refuted only if one can tell a plausible story about why it *appears* to us as though measurements have definite results. We can understand the many-worlds interpretation (1957) of the American physicist Hugh Everett III in this sense. According to it, a measurement yields not only the particular result that we apparently register, but also all the other possible measurement outcomes—the others in different "worlds", which also contain us, or "copies" of us, each of whom then registers one of the other results; an interpretation which is in the meantime quite seriously supported. The second assertion, finally, can be refuted in one of two possible ways: on the one hand, by Heisenberg's Copenhagen interpretation, in which the linear Schrödinger evolution is *complemented* by an additional, second dynamics of the measurement process; and on the other hand, by GRW, in which the Schrödinger equation is *replaced* by a *nonlinear* time evolution. These two interpretations imply a collapse of the state vector— "Copenhagen" a methodological collapse brought about by the measurement process and GRW a realistic, spontaneous collapse. Maudlin's trilemma thus points up the spectrum of all the interpretation variants of quantum mechanics which are still being discussed today (and each of them has its place in this book: Bohm, the Copenhagen interpretation, GRW and Everett).[25]

GRW is thus "also" a collapse interpretation, but unlike Heisenberg's Copenhagen interpretation, it is a realistic one. In order to understand what is necessary to construct such a theory, we will have to investigate the precise reason why one cannot assume a unified time evolution according to (Bohr and) Heisenberg. To that end, we will try to describe the measurement process somewhat more formally: We thus assume that the measurement setup, for example a Stern–Gerlach apparatus, is describable with the methods of quantum mechanics.[26] In its initial state, i.e. at a time before the interaction with the quantum object that we wish to measure, the apparatus indicates "neutral", which is represented by a vector in some Hilbert space simply as $|M_0\rangle$. Furthermore, we presume that the quantity to be measured (of the quantum-mechanical system) can take on precisely two possible values, like our favourite example of the spin of an electron. Now, it *must* hold that *if* the quantum object was previously correlated with an eigenvector of the corresponding operator—thus either with $|up\rangle$ or with $|down\rangle$—then the measurement apparatus, at a *time after* the interaction with the quantum object, must show the value +1 (or −1) *with certainty*.[27]

---

[24] There is still another interpretation which also refutes assertion 1, but without assuming the existence of hidden variables. It is the modal interpretation of quantum mechanics (see (van Fraassen 1991), Chap. 9). However, its main problem—namely how to explain why repeated measurements (e.g. of the spin) lead with certainty again to the same result, although there has been no collapse nor have hidden variables guaranteed the result—has not been solved in a convincing way.

[25] Note that in this description of the measurement problem, we have not referred again to Born's rule. In fact, this probability interpretation of the state vectors holds neither in GRW theory nor in Everett's theory.

[26] Making this assumption means that we cannot balk here and like Bohr asserts dogmatically that macroscopic measurement apparatus is in any case only classically describable.

[27] Apart from the fact that real measurement apparatus also has a nonzero error rate.

We can therefore say that it is then in a state indicated by $|M_1\rangle$ (or by $|M_{-1}\rangle$). All together, the process can then be schematically represented as follows:

$$|M_0\rangle|up\rangle \longrightarrow |M_1\rangle|up\rangle$$

or

$$|M_0\rangle|down\rangle \longrightarrow |M_{-1}\rangle|down\rangle\,. \tag{2.3}$$

While the measurement apparatus still indicates "neutral" in the initial state, it then apparently adapts itself to the quantum-mechanical object via the interaction: Its pointer position in the final state corresponds to the then actually existing property of the quantum object, as should be expected of a valid measurement.[28] "Measuring" means in this case establishing macroscopic determinacy, which is thus perceivable for us, *without* changing the quantum object in the process,[29] i.e. only that kind of macroscopic determinacy which is in accord with a temporally immediately preceding microscopic determinacy.

As a rule, however, the object to be measured is not already correlated with one of the eigenvectors of the chosen operator, but rather with a *superposition* of these, so that the initial state of the composite system is rather the following:

$$|M_0\rangle(c_1|up\rangle + c_2|down\rangle)\,. \tag{2.4}$$

Now, the question arises as to which final state is to be expected *with the methods of the standard formalism* in this usual case. To be consistent, one would have to evolve this initial state with the temporal dynamics of the unitary time-evolution operator $\hat{U}_t = e^{-\frac{i}{\hbar}t\hat{H}}$. This appears at first to be technically unpromising, since the Hermitian Hamiltonian operator $\hat{H}$ which enters the unitary operator here represents the total energy of the composite system and is therefore hopelessly complex. That is because the measurement apparatus, supposing that it can be described by quantum mechanics at all, consists of an enormous number of particles, which must all be taken into account. But we can convince ourselves that precisely this time evolution converts the (previous) initial state $|M_0\rangle|up\rangle$ into $|M_1\rangle|up\rangle$[30]—and likewise $|M_0\rangle|down\rangle$ into $|M_{-1}\rangle|down\rangle$—so that the situation is considerably simplified: For in standard quantum mechanics, everything is indeed *linear*, and therefore, we can rearrange the initial state in the form:

---

[28] Caution: Tacitly, we have introduced here mathematically a product between vectors in different Hilbert spaces, which we had previously not discussed. In fact, the measurement apparatus and the quantum object form a *composite system*, which will be treated in detail only in the following chapter. Such a (pure) product, such as $|M_{-1}\rangle|down\rangle$, in any case reflects rather classically a whole whose properties are completely determined by its parts. One says that the state of the whole is *separable*, in that here the particle (quantum object) and the measurement apparatus each individually indicate $+1$ (the fact that they in the end always point in the same direction is a result of their interaction).

[29] In the ideal case! There are of course measurements which inevitably destroy the quantum object, even when it was already in an eigenstate of the corresponding observable.

[30] Otherwise, the measurement apparatus from the outset cannot accomplish what it is supposed to.

$$c_1|M_0\rangle|up\rangle + c_2|M_0\rangle|down\rangle\,, \tag{2.5}$$

and finally—due to the *linearity* of the unitary time-evolution operator—we find for the final state:

$$|\Psi_{final}\rangle = c_1|M_1\rangle|up\rangle + c_2|M_{-1}\rangle|down\rangle\,. \tag{2.6}$$

A fatal result: because now the final state of the composite system consisting of the measurement apparatus and the quantum object is in a superposition of different *macroscopically perceivable* states. Whatever the summation symbol in Eq. 2.6 may mean,[31] macroscopic definiteness with respect to the chosen observable has in any case not been achieved. $|\Psi_{final}\rangle$ is, as the sum of two pure product states, namely *measurably* distinct from these.[32] In fact, we nevertheless obtain *either* the measured value $+1$, with a probability of $|c_1|^2$, *or* the value -1, with a probability of $|c_2|^2$; measurements yield *definite* outcomes![33] The standard formalism of quantum mechanics thus leads to a contradiction with experience, since it predicts superpositions for *macroscopic* entities that we clearly do not observe.

At this point, one has often been tempted to adopt the following strategy: After the interaction, the overall system is in a *pure* state, which is given by the projection operator $\hat{P} = |\Psi_{final}\rangle\langle\Psi_{final}|$ and thus represents the undesired superposition; this is undoubted. For composite systems, however,[34] it holds that even when they are in a pure state, their subsystems are as a rule mostly in mixed states. If one then considers the subsystems by themselves—and, thus, "tracing out" mathematically the neglected part—one arrives at the statistical operator (density matrix) for the quantum-mechanical object:

$$\hat{\rho}_{quant} = |c_1|^2|up\rangle\langle up| + |c_2|^2|down\rangle\langle down| \tag{2.7}$$

as well as for the measurement apparatus:

$$\hat{\rho}_{meas} = |c_1|^2|M_1\rangle\langle M_1| + |c_2|^2|M_{-1}\rangle\langle M_{-1}|\,. \tag{2.8}$$

Through the interaction with the measurement apparatus, the quantum object has thus been *changed*, as was to be expected for this normal case; because before the interaction, the quantum object was correlated with a projection operator and, immediately thereafter, with the (diagonal) statistical operator, which physically

[31] The overall system is in no case in a state in which the measurement apparatus would indicate *both* 1 *and* -1, which would be contradictory.

[32] The superposition is itself instead the eigenvector of a different, incommensurable operator, analogously to the situation in two-dimensional spin space, even though here— in the macroscopic world—it is not so simple to specify the operator explicitly.

[33] A circumstance which Bohr avoided explaining right from the start.

[34] More details will follow in later chapters.

represents a different state.[35] The measurement apparatus for its part is now also in a mixed state, so that its components, which are those of a statistical operator, could be directly construed as probabilities. One is then tempted to say, "The quantum system itself is indeed *either* in the state $|up\rangle$, *or* in the state $|down\rangle$, and correspondingly, the measurement apparatus itself is in fact *either* in the state $|M_1\rangle$, *or* in the state $|M_{-1}\rangle$—since the probabilities that occur in the statistical operator are only *subjective degrees of belief*, i.e. they reflect only our lack of knowledge about the true state of the quantum object".

However, as already indicated in Sect. 1.2.4 on the statistical operator, the ignorance interpretation of its eigenvalues is, precisely at this point, *not* tenable; one can also say that the statistical operator here describes only an "improper" mixture. For if we assume (see (van Fraassen 1991), p. 207) that the real state of the quantum object were given by $|up\rangle$, what then excludes the possibility, on the basis of the ignorance interpretation—and it *is* to be excluded—that in the mixture *of the apparatus*, $|M_{-1}\rangle$ indicates the true state? And: Whence come the interference terms for the overall system, on the basis of the ignorance interpretation? Why is thus the whole (composite system) not also describable by a statistical operator? But *without* the ignorance interpretation, the entire measurement process is more like a step backwards; the determinacy which we wished to obtain has not yet been realized, due to the change of the quantum object— i.e. due to the transition from a pure (superposition) state to a mixed state. Instead, only the previously existing pure state has been destroyed.[36]!

2.3.2 The Decoherence Programme

By comparison, the programme of *decoherence* represents real progress.[37] Originally, it was even associated with the hope that the interpretation problem of quantum physics could be solved, and in fact, it delivers (at least locally) a *physical* explanation for the absence of superpositions of macroscopically distinguishable

[35]The corresponding projection operator would *not* be in diagonal form: $\hat{P}_{final} = (c_1|up\rangle + c_2|down\rangle)(c_2\langle down| + c_1\langle up|) = |c_1|^2|up\rangle\langle up| + |c_2|^2|down\rangle\langle down| + c_1c_2^*|up\rangle\langle down| + c_1^*c_2|down\rangle\langle up|$. This has the effect that the expectation values of operators, $\langle\Psi_{final}|\hat{O}|\Psi_{final}\rangle = \mathrm{Tr}(\hat{P}_{final}\hat{O})$, in general contain interference terms (cross terms, which exhibit intuitively the wave character of particles, as for example in the two-slit experiment). These are lacking in $\langle\hat{O}\rangle_{\hat{\rho}} = \mathrm{Tr}(\hat{\rho}_{quant}\hat{O})$.

[36]The original (superposition) vector was, like every vector of a pure state, indeed an eigenvector of a different (incommensurable) operator. Referring to the statistical operator that we have now obtained, the expectation value of every operator, however, is no longer the (former) eigenvalue and no longer free of variance.

[37]"Coherent", i.e. "interrelated", is the classical term for the condition which must be fulfilled by waves in order that they exhibit interference. Thus, one can also refer to quantum-mechanical superpositions as "coherent states". The decoherence approach then attempts to clarify the conditions under which a classical world, in which precisely such superpositions no longer occur, can emerge on the basis of quantum mechanics.

states.[38] The measurement problem is nevertheless not really solved. The decoherence programme, however, represents an important supplement of all the interpretation options which are still under discussion. Its details are mathematically quite involved, so that here, we can give only an indication of what progress it has brought and where its limits lie.

The fundamental idea is that the composite system consisting of the quantum object and the measurement apparatus is no longer to be treated as a *closed* system, as had been assumed previously. Instead, it is now taken into account that this system interacts with its environment in many ways—for example in that the measurement apparatus is constantly reflecting light. The actual final state is thus a much larger one which includes the entire surroundings E (the *environment*), namely (for example) a state like this one:

$$|\Psi_{final}\rangle = c_1|E_1\rangle|M_1\rangle|up\rangle + c_2|E_{-1}\rangle|M_{-1}\rangle|down\rangle\,. \tag{2.9}$$

If one could then succeed in showing that not merely neglect, but rather precisely the influence of the environment leads to the reduction, then the *open*, composite system consisting of the quantum object and the measurement apparatus would be released for physical reasons from its entanglement with its environment and would be converted to a mixed state:

$$\hat{\rho} = |c_1|^2|up, M_1\rangle\langle M_1, up| + |c_2|^2|down, M_{-1}\rangle\langle M_{-1}, down|\,. \tag{2.10}$$

The various macroscopically perceivable states then no longer interfere with each other. The difference from the previous situation could be that at *this* point, the ignorance interpretation of the probabilities of the statistical operator is indeed appropriate; that thus, due to the decoherence, a "proper" mixture is in fact obtained. Apparently, van Fraassen's arguments no longer hold: The environment includes here everything possible, so that it is not really important that orthogonal environment vectors $|E_1\rangle$ and $|E_{-1}\rangle$ be exactly correlated with the quantum-measurement apparatus system. There is in any case no measurable difference whether $|\Psi_{final}\rangle$ is described as in Eq. 2.9 as a pure state, or itself is already a mixture, because it refers to a much-too-large system. If this is convincing, then one can now say that at times *after* the interaction between the quantum object and the measurement setup, in fact either the macroscopic state which is indicated by $|up, M_1\rangle$ is present, *or else* the state which is represented mathematically by $|down, M_{-1}\rangle$, which is perceptibly different—as desired. This would provide a physical explanation for the fact that we do not perceive macroscopic systems whose states are superpositions ("environment-induced decoherence").

However, as self-critical physicists admit (cf. (Schlosshauer 2007), pp. 49 and 69), the ignorance interpretation of the statistical operator is at this point *still* not appropriate. Globally, coherence with the environment still holds, and is even amplified,

[38] Pioneering works are those by Zeh (1970); Zurek (1981); for the more recent state of the art, see Schlosshauer (2007).

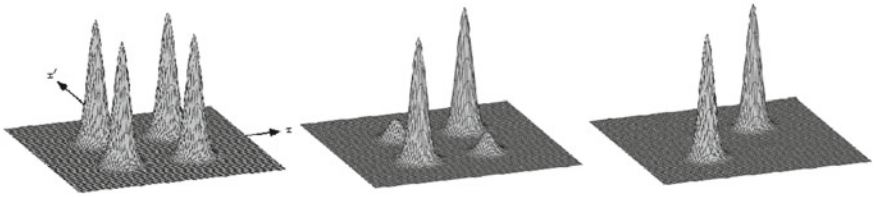


Fig. 2.1 The time evolution of the state (in the position representation) due to the influence of the environment: locally, the interference terms vanish (figure taken from (Schlosshauer 2007), p. 149)

so that in this case, also, the interferences disappear only because we intentionally avoid considering the environmental contributions. At least, however, the interference terms become significantly smaller *locally*, and this indeed for physical reasons, as shown in Fig. 2.1. Thus, we can at least explain why no superpositions *appear* to us as local observers.

Furthermore, a problem has been solved which we have thus far treated rather casually. For the measurement problem in fact has *two* parts, corresponding roughly to the observer in Heisenberg's Copenhagen interpretation, who plays *two* roles: Before he or she can register a measurement outcome, a measurement setup must first be chosen. Mathematically, the previous state vector could indeed be represented in infinitely many ways as a superposition—every different basis representation is mathematically equivalent. In the (mathematical) treatment of the measurement problem, however, we have represented the initial state from the beginning in the eigenvector basis of the quantity that was *to be measured*. What justifies physically this special status for the so-called pointer basis? Copenhagen "solves" this problem by appealing to an influence of the observer, which is unacceptable for many: A basis is selected by the *choice* of the measurement setup. Without a physical solution of the *preferred basis problem*, for example, Everett's many-worlds interpretation would be destined to fail. In this sense, it is an important result that the decoherence approach has actually solved *this* part of the measurement problem: According to the "tri-orthogonal uniqueness theorem" (cf. (Elby and Bub 1994)), the decomposition into orthogonal states of a threefold product space such as the one that encompasses the quantum system, the measurement apparatus and their environment is *unique*.

The dynamical distinction of a basis and the physical explanation of why we do not observe superpositions in the macroscopic world are thus the two areas of progress which have been achieved by the decoherence programme. But they on the one hand only determine the *type* of properties objectively and on the other hand only clarify that we cannot *perceive* superpositions. However, it is evidently important that no superpositions *exist*, but instead a (new) pure state. So, it must still be explained that in a measurement, we find a definite *value* of a property. Since we have not overcome superpositions ontologically, we are quite possibly still dependent on von Neumann's projection postulate. Furthermore: Even if the interference terms were all to disappear physically, and (therefore) the ignorance interpretation of the statistical operator were to be appropriate, we would still have obtained only an

(to be sure classically understandable) *either-or* situation, and thus, we would still have a case of indeterminacy. Even a "proper" mixture, to which the decoherence programme could have hypothetically led, fails to solve the measurement problem. There would then namely be subjective probabilities on a *fundamental* theoretical level. The ignorance interpretation indeed means that ontologically, there is a definite (new) pure state present, but that epistemologically, we do not know which state it is. The mathematical description ended with a mixture; objectively, however, there should be a pure state (or rather its corresponding eigenvalue): Then either standard quantum mechanics would be incomplete, or else one would have to turn to Everett's many-worlds interpretation; we would thus have to refute assertion 3 of Maudlin's trilemma.[39]

What is important here: If one does not wish to interpret superpositions in the sense of Everett, and does not accept a subject dependence of reality as with Heisenberg *et al.*, then there is obviously no path leading to definite measured outcomes that could be described by the methods of the standard formalism of quantum mechanics. What the attempts to get a grip mathematically on the measurement process then should show us is that one can make a plausible case that the cause of this failure could be the *linearity* of quantum mechanics—and that the latter can be overcome only by taking the mathematical description of the time evolution to be explicitly *nonlinear*. This, in any case, was the basic idea of the GRW theory: to replace the linear Schrödinger dynamics with nonlinear dynamics.

2.4 The Realistic Collapse Interpretation: GRW

The goal of the theory developed by Ghirardi, Rimini and Weber in 1986 was to find a unified dynamics for the microscopic and the macroscopic worlds and thereby to overcome the Copenhagen schism, thus resolving the measurement problem. This GRW theory is mathematically rather complex and has up to now not been treated in physics textbooks. Immediately after its publication, however, it was enthusiastically greeted by Bell (1987). In the philosophy of physics, it has been rather popular since its reception by Albert (1992). The debate surrounding the adequate GRW ontology has continued up to the present.

2.4.1 Nonlinear Dynamics

Already in the introductory sections, we pointed out that *linear* mappings preserve superpositions in the following sense:

[39] The significance of decoherence theory for Everett's interpretation is discussed in Sect. 5.2.4.

$$\hat{O}\left(a|A\rangle + b|B\rangle\right) = a\left(\hat{O}|A\rangle\right) + b\left(\hat{O}|B\rangle\right). \tag{2.11}$$

We can now state that a nearly instantaneous, discontinuous transition from a superposition of the eigenvectors of a particular operator to one certain eigenvector—as required by von Neumann's projection postulate—is possible only when the linearity of quantum mechanics is breached at a relevant point. Achieving this in a unified form, that is in a mathematically coherently describable manner, requires that the unitary Schrödinger evolution be replaced by a *new, nonlinear* equation.

This will not be possible without far-reaching effects. In Sect. 1.2.2, we furthermore stated that linear mappings imply mathematically that parallel lines (vectors) remain parallel. Since parallel vectors in Hilbert space are correlated with identical physical states, linear mappings imply physically that no physical difference can come into play where there was none "before". Applying this to the time evolution, we can deduce that it is *deterministic* when it is linear. A *nonlinear* temporal dynamics as expressed by the GRW equation (which replaces the Schrödinger equation) at least offers the possibility of an objective, lawlike indeterminism: Now, in the course of time, a physical difference can come into play where at previous times none was present. *One* physical state "splits" off by evolving nonlinearly in such a manner that at a later time, it has *either* already collapsed, *or* still not yet.[40] The nonlinear dynamics is then not only discontinuous, but also *indeterministic*. Finally, everything points to the fact that this process would also be *irreversible*; and thus that it would be considerably more improbable that two initial states which thus far correspond to the two (possible) final states, would evolve into the one final state (the initial state thus far). Indeed: The GRW dynamics is (sometimes) discontinuous, indeterministic and temporally irreversible.

This is naturally just what was wanted! The nonlinearity, however, has still other effects: Many characteristics of standard quantum mechanics are based on the linearity of the theory, for example the fact that the choice of a basis is simply a matter of convention, and thus that a vector in Hilbert space can be represented in (infinitely) many ways as a superposition (a linear combination) of other vectors. GRW, in contrast, distinguishes one basis as *preferred*: the position representation.[41] Furthermore, we had already found that, for example, the question of which eigenvalues (physically possible measurement outcomes) a given Hermitian operator has, and what its eigenvector basis (or bases) might be, is determined independently of time. From this, we *deduced* that the (measurement-independent) time evolution must be unitary, and thus in any case linear. If, however, the time evolution is now supposed to be nonlinear, this has repercussions for the concept of a quantum-mechanical

[40] An initial state can thus have *two* possible final states: Going in one direction of time, towards the future(?), the state can split apart.

[41] How does distinguishing this *particular* basis relate to the decoherence programme, one of whose steps forward consists in the fact that indeed one basis is dynamically distinguished? There is at least some tension between "decoherence" and GRW when it turns out that the special basis in certain situations is not the position representation. Concerning this problem, see (Schlosshauer 2007, pp. 349f.).

observable. No doubt: The GRW theory is a quite *different* theory from standard quantum mechanics.[42] In any case, with GRW, the time evolution again stands at the centre, namely at the beginning of the theory, on which, for example, the concept of observables or measurable quantities depends—while in standard quantum mechanics, in contrast, the time evolution is a consequence of fundamentally time-independent characteristics of operators, eigenvalues, and eigenvectors. The time evolution of the wavefunction, that is of a vector in the position representation, is thus the fundamental object of the GRW theory.[43]

How does this equation look explicitly? The idea is (cf. (Bell 1987)) that the wavefunction of N particles, that is

$$|\Psi\rangle = \psi(t, \vec{x}_1, \vec{x}_2, ..., \vec{x}_N)\,, \tag{2.12}$$

which would otherwise continue to evolve according to the usual Schrödinger equation dynamics, will collapse from time to time, as follows:

1. *Spontaneously*, i.e. not caused by some sort of external interaction such as in particular by a measurement which would require an observer;
2. *Stochastically*, both in terms of *when* the collapse occurs and also *where* it occurs;
3. And this finally in such a manner that with only a few particles, a collapse is very, very rare, while with a very large number of particles, a collapse occurs practically *immediately*. This guarantees that on the one hand, an isolated, single quantum system can evolve for quite a long time according to the Schrödinger equation, while macroscopically, practically no superpositions are present.

In order to achieve both features, two new *constants of nature* must be introduced; namely on the one hand, a constant with the unit of (inverse) time:

$$\frac{N}{\tau};\ N \text{ is the particle number and } \tau = 10^{15}\text{s}\,. \tag{2.13}$$

This quotient expresses the *probability* per time of a quantum jump occurring; clearly, it is very small for small numbers N of particles, but becomes very large for macroscopic particle numbers of the order of 10^{23}—as intended.

We can already see here a crucial difference from the usual quantum mechanics: While there, according to *Born's rule*, the wavefunction, or rather its absolute square, is understood to represent the (measurement) probability, here a newly introduced parameter carries the probability information. In this way, the probability is no longer *conditional*, especially not one which dependent on a measurement. Furthermore, it can be applied to single events without difficulties and thus to a single collapse. Therefore, in the literature, there is general agreement that the GRW probabilities

[42] It is perhaps even the case that GRW makes some predictions which deviate from those of standard quantum mechanics, so that future experiments might "prove" that GRW is also empirically more realistic—but they could also falsify the GRW theory!

[43] This is a common feature with Bohm's mechanics (which, however, is *deterministic*).

do not represent subjective degrees of belief, but rather objective chances (cf. (Frigg and Hoefer 2007), p. 376).

To answer the question of how and where a collapse occurs, a *localization operator* is introduced:

$$\hat{L} = \left(\frac{\alpha}{\pi}\right)^{\frac{3}{4}} \exp[-\frac{\alpha}{2}(\hat{\mathbf{q}}_k - \mathbf{x})^2] \,. \tag{2.14}$$

It has the shape of a Gaussian curve with a *randomly chosen* centre around the position (operator) of the kth particle. The degree of localization, that is the width of the Gauss curve, determines the second new constant of nature, which has the unit of a length:

$$\alpha = 10^{-5}\text{cm} \,. \tag{2.15}$$

Essentially, structurally,[44] the GRW equation then consists in adding a term to the Schrödinger equation which applies this Gaussian localization operator to the wavefunction. It then produces a so-called hit, i.e. a spontaneous localization around the randomly "chosen" centre, having the probability $\frac{N}{\tau}$ with respect to time.

Let us now consider again our measurement problem, or else Schrödinger's living and/or dead cat: The problem consisted of the fact that with the methods of the standard formalism, two macroscopically–perceivably different states would have to be in a superposition, contradicting the empirical result, whereby we do not observe such superpositions:

$$|\Psi_{final}\rangle = c_1|M_1\rangle|up\rangle + c_2|M_{-1}\rangle|down\rangle \,. \tag{2.16}$$

If we now, however, consider the components of the state of the measurement apparatus (or of the macroscopic components in the case of the cat) more carefully, i.e. remembering that the measurement apparatus (the cat) is composed of trillions of microscopic particles which are each—pointer position $+1$ compared to pointer position -1 (alive *vs.* dead)—located at significantly different positions, then the state can be written mathematically as follows:

$$|\Psi_{final}\rangle = c_1(|x_1\rangle_1|x_1\rangle_2...)|up\rangle + c_2(|x_{-1}\rangle_1|x_{-1}\rangle_2...)|down\rangle \,. \tag{2.17}$$

Each of the summands now contains essentially coupled wavefunctions which represent positions of the particles that suit the macroscopic pointer position $+1$ (or -1). If we then apply the GRW mechanism, the second particle, for example, could collapse to the state $|x_{-1}\rangle_2$—as "chosen" randomly. As a result of the coupling, then the overall state collapses—with a probability of $|c_2|^2$—to the macroscopically perceivable state:

$$|\Psi_{final}\rangle = |M_{-1}\rangle|down\rangle \,. \tag{2.18}$$

[44] The fundamental GRW equation is better represented by referring to the density matrix; compare ((Frigg and Hoefer 2007), p. 374).

The extremely large number N of particles guarantees—together with the new constants of nature—that such a collapse occurs almost *immediately*, so that macroscopically, nearly *always* one of the two product states is found and nearly *never* the superposition state, and this corresponds to our experience. We emphasize that this description, if it is convincing, indeed represents the *resolution* of the measurement problem: For with a very high probability, a single-particle state collapses—and *therefore* the overall state collapses—so that macroscopic determinacy is in fact obtained (and not, e.g., only an *either-or* situation, as with the decoherence approach). And this is independent of whether or not an *observer* registers the result.

From the viewpoint of the philosophy of science, however, an objection can be raised against this GRW approach, in that the introduction of new constants of nature was an *ad hoc* manoeuver: The *values* of those constants were chosen just so that—for a single-particle system as well as for the macroscopic whole—the right probabilities emerge. There is no independent procedure by which we could determine the values of α and τ empirically. They have furthermore not yet led to any new predictions; they serve only to explain the results of measurements which we already know in the context of conventional quantum physics. For comparison, we could consider, for example, the constant of the vacuum velocity of light c, which plays a decisive role in the theories of relativity: Quite independently of Einstein's theories, there are—indeed a number of—empirical methods for determining the precise quantitative value of c. In addition, it also plays a role in other physical theories. This is not the case for α and τ: Their values were fixed for no purpose other than the resolution of the measurement problem of quantum mechanics. Just such a procedure is termed an "*ad hoc* manoeuver" in the philosophy of science.

In addition, it should be borne in mind that the GRW mechanism is dependent upon the fact that macroscopically perceivable different states are always accompanied by different *spatial* positions. Only when the spatial locations of the particles which reflect macroscopically different states are significantly different does the argument function properly, stating that, owing to the coupling of the particle positions, the very probable collapse of just a few single-particle states will lead immediately to a perceivable, determined macrostate. However, one could perhaps invent experiments in which perceivable measurement outcomes are *not* associated with macroscopically differing spatial locations (see (Albert 1992), p. 103). Leaving such objections aside, we are philosophically interested above all in the *ontological* consequences of the GRW theory.

2.4.2 GRW Ontologies and Their Criticisms

The special status for the position basis and the fundamental character of the time evolution make GRW *prima facie* philosophically–ontologically more acceptable than standard quantum mechanics: If we agree that one goal of modern metaphysics of science is to treat—and ideally to solve—the conflict between our (thoroughly enlightened) everyday worldview and the pictures drawn by the most current scien-

tific theories such as quantum physics, then the prominence of spatial location and time should evidently be welcomed. By no means simply naive intuitions clearly support the view that relationships in space and evolutions in time belong to the fundamental phenomena of the world in which we live, so that a scientific theory which places them at its centre will tend to reduce the conflict of worldviews. Standard quantum mechanics, in contrast, with its abstract Hilbert space in which on the one hand the position operator is only one among many, and in addition a mathematically unlovely one; and on the other hand, time evolution is not an observable— is obviously further removed from our intuitions, which also should always be kept in view by the metaphysics of science.

Nevertheless, it should be emphasized that GRW claims to give a *unified* dynamics for microscopic as well as for macrophysical phenomena, so that the fundamental wavefunction is as a rule not that of a single quantum system, but rather that of a very large number of particles. The "position" space which seems to be distinguished here ontologically is thus not simply that of the three-dimensional intuitive space, but instead the $3N$-dimensional "configuration space", where N is the particle number. This space thus has an extremely high dimensionality and is correspondingly counterintuitive. In spite of this, serious support (cf. (Albert 1992), pp. 92f., and the criticisms in (Maudlin 2010; Monton 2006)) have been given to the notion that physical reality in fact takes place within this configuration space, so that the wavefunction and especially its collapse would represent real objects and events in configuration space—and that as a result, our intuitive understanding of a merely three-dimensional space is deceptive. Conversely, it has also been maintained again and again (see, e.g., (Bell 1987), pp. 204f.) that configuration space is simply an abstract, mathematical construction which must be distinguished from a physically real space. But we can very well ask, Why, actually? Must physical reality be at all different from its mathematical representation? And if so, *whereof* does their well-justified difference consist? One should not forget that the realism defended here is marshalled against "Copenhagen" and thus in particular against the idea that physical reality could depend upon the *subject* as a measuring observer. Every reference to our experience, which is presumed to play a constitutive role in distinguishing physics from mathematics, is thus inappropriate. Therefore, this notion of "the realism of configuration space", referred to as "wavefunction ontology", may well be the first fully valid variant of the GRW ontology.

Every other variant answers the question of how to overcome the direct, literal interpretation of the GRW equation as the equation of motion of a field in configuration space, in order to make our way out of that mathematical space and into the unique, physically real intuitive space. This includes in particular Bell's classical interpretation of the *local beables* (see (Bell 1987), p. 205): Bell supports the position that the wavefunction itself indeed "lives" only in that mathematical space and is thus not a real physical field, but that a GRW jump is centred on a normal spacetime point $(\vec{\mathbf{x}}, t)$. Whenever a collapse takes place, something therefore occurs in our intuitive space, or in spacetime. A macroscopic object is then "a galaxy of such events" ((Bell 1987), p. 205). This ontology of galaxies of flashes remains today the most promising variant (cf. (Maudlin 2010), p. 139), especially because it, in

contrast to the competing third variant—the density-of-stuff ontology[45]—appears to be generalizable to the relativistic case.[46] The main problem with Bell's ontology is apparently that[47] the flashes cannot be fundamental; they depend after all on the wavefunction in configuration space, so that in our real, physical world such events would occur discontinuously—and basically inexplicably, in any case not explained by anything which exists in our world. As long as, and always, when the wavefunction has not collapsed and is in a superposition, precisely nothing exists physically, but rather only something mathematical in the $3N$-dimensional configuration space. As a result, Bell has inherited an anti-realistic touch from "Copenhagen", to the discomfort of those who share the hope that GRW can deliver a universal realism.

Thus, it has recently been suggested (cf. (Dorato and Esfeld 2010)) that a real, spatiotemporal foundation should be provided to Bell's stochastic events. The "spatial" superpositions in configuration space thereby represent *dispositional* properties in the real physical spacetime, produced causally by the spontaneous Bell flashes. According to Dorato/Esfeld, this could take us out of the mathematical abstract and into the physical concrete realm by giving superpositions a *modal* character and construing them as *metaphysical* powers. The fundamental idea originates from an—in the meantime well-established— position from the philosophical debate over the status of natural laws: Hume's conception according to which natural laws supervene over certain regularities of categorial (purely qualitative) properties without modal capacities is opposed by the position that fundamental properties in physics (charge, mass) are essentially dispositional (modal-causal) properties.[48] This position, which would appear to be excessively metaphysical, can be defended by saying that it is not possible to distinguish physical structures from mathematical structures without the interpretation of properties as modal or dispositional. In any case, Dorato/Esfeld argue in this way in relation to the GRW theory. If their argument is found to be convincing, it will provide a real spatiotemporal, causal foundation for Bell's flashes, which certainly appears to be desirable. The problem, however, is that dispositions—such as, e.g., the solubility of sugar in water—usually require an external trigger before they manifest themselves; they must be set off from outside the system. However, GRW localizations are *spontaneous*; the GRW collapse characteristically requires no external interactions (in particular, no measurement). It follows for its realistic interpretation that the superpositions represent spatiotemporal, dispositional properties, which "manifest themselves spontaneously" ((Dorato and Esfeld 2010), p. 44); this apparently recalls a spontaneous *causa sui*. In intermediate phases, real properties of spatiotemporal quantum systems would have no definite values, but they would be "mind-independently and probabilistically dis-

[45] In this approach, a continuous "material" field in concrete spacetime corresponds to the wavefunction in configuration space. It is supported by Ghirardi himself, among others.

[46] Physical difficulties of the GRW theory relate to the "indistinguishability" of similar particles, the treatment of counterfactual dependencies in the EPR problem, and in particular its compatibility with the theory of (special) relativity. This last topic is also problematic for Bohm's mechanics.

[47] In the literature, a so-called *counting anomaly* has also been discussed, which, however, we need not consider here.

[48] A Hume interpretation of GRW is defended by Frigg and Hoefer (2007).

posed to become definite" ((Dorato and Esfeld 2010), p. 45): During the collapse, they give themselves their own values by self-acting, although probabilistically. One gains the impression that the subject, as a measuring observer, can be eliminated using GRW only *animistically*, in that now the world as a whole realizes ("manifests") itself spontaneously, like a subject. The alternative is wavefunction ontology and thus the identification of the physical with some segment of the mathematical.

With this, we have arrived at the end of this first chapter on the interpretations of quantum physics. It dealt with uncontroversial but inadequate interpretations (the minimal interpretation and the ensemble interpretation), as well as with the more controversial collapse interpretations (Copenhagen, GRW). Before we discuss the *no-collapse* variants (Bohm, Everett) in a further chapter on interpretations, we first treat the two most important special problems of quantum mechanics: similar particles and their "indistinguishability", and the EPR paradox. Individuation of objects and the relationship of a whole to its parts belong among the prominent topics of ontology, which in this regard is faced with quite unique and new challenges—already from standard quantum mechanics.

Exercises

1. Distinguish between two readings of Born's rule, depending upon whether the reference to a measurement in it is essential or not.

2. According to the Copenhagen interpretation (in Heisenberg's version), there are two temporal dynamics of the state vector. Describe them in your own words. How is the second dynamics related to Born's rule and to von Neumann's projection postulate? What is problematic about it?

3. The interpretation problem in quantum mechanics may be regarded as a trilemma. Explain the three statements and show that they are inconsistent when taken together. What is the advantage of this description as compared to the conventional one, which is guided by Born's rule?

4. The decoherence programme is an essential step forward. Highlight the ways in which all the interpretation variants could profit from it. Why, however, can the programme not solve the measurement problem in the end?

5. Formulate in your own words what is accomplished by the GRW theory (in the opinion of its supporters). Defend the standard view of physics against it.

References

Albert, David Z. 1992. *Quantum Mechanics and Experience*. Cambridge: Harvard University Press.
Audretsch, Jürgen (ed.). 2002. *Verschränkte Welt. Faszination der Quanten*. Weinheim: Wiley-VCH.

Baumann, Kurt, and Roman U. Sexl. 1984. *Die Interpretationen der Quantentheorie*. Braunschweig: Vieweg.

Ballentine, Leslie E. 1998. *Quantum Mechanics: A Modern Development*. Singapore: World Scientific.

Bell, John S. 1987. Are there quantum jumps?. In: Bell. In *Speakable and Unspeakable in Quantum Mechanics (Revised edition 2004)*, ed. Speakable Bell, 201–212. Cambridge: Cambridge University Press.

Dorato, Mauro, and Michael Esfeld. 2010. GRW as an ontology of dispositions. *Studies in History and Philosophy of Modern Physics* 41: 41–49.

Elby, Andrew, and Jeffrey Bub. 1994. Triorthogonal uniqueness theorem and its relevance to the interpretation of quantum mechanics. *Physical Review A* 49 (5): 4213.

Esfeld, Michael (2012). Das Messproblem der Quantenmechanik heute: Übersicht und Bewertung. In Esfeld (ed.). *Philosophie der Physik*, 88–109. Frankfurt/M.: Suhrkamp.

Faye, Jan (2008). Copenhagen Interpretation of Quantum Mechanics. In E.N. Zalta (ed.). *The Stanford Encyclopedia of Philosophy (Fall 2008 edition)*. http://plato.stanford.edu/archives/fall2008/entries/qm-copenhagen/

van Fraassen, Bas C. 1991. *Quantum Mechanics: An Empiricist View*.Oxford: Clarendon Press.

Friederich, Simon. 2011. How to spell out the epistemic conception of quantum states. *Studies in History and Philosophy of Modern Physics* 42: 149–157.

Frigg, Roman, and Carl Hoefer. 2007. Probability in GRW theory. *Studies in History and Philosophy of Modern Physics* 38B: 371–389.

Fuchs, Christopher A., and Asher Peres. 2000. Quantum theory needs no interpretation. *Physics Today* 53 (3): 70–71.

Ghirardi, GianCarlo, Alberto Rimini, and Tullio Weber. 1986. Unified dynamics for microscopic and macroscopic systems. *Physical Review D* 34: 470–491.

Heisenberg, Werner. 1959. Die Kopenhagener Interpretation der Quantentheorie. In *Physik und Philosophie*, ed. Heisenberg, 27–42. Stuttgart: Hirzel.

Held, Carsten. 2012. Die Struktur der Quantenmechanik. In *Philosophie der Physik*, ed. M. Esfeld, 71–87. Frankfurt/M.: Suhrkamp.

Maudlin, Tim. 1994. *Quantum Non-Locality and Relativity*. Oxford: Blackwell.

Maudlin, Tim. 1995. Three measurement problems. *Topoi* 14: 7–15.

Maudlin, Tim. 2010. Can the world be only wavefunction? In *Many Worlds? Everett, Quantum Theory, and Reality*, ed. S. Saunders, *et al.,* 121–143. Oxford: Oxford University Press.

Monton, Bradley. 2006. Quantum mechanics and $3N$-dimensional space. *Philosophy of Science* 73: 778–789.

von Neumann, John. 1932. *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer. (Reprint 1968).

Rosenthal, Jacob (2003). *Wahrscheinlichkeiten als Tendenzen. Eine Untersuchung objectiver Probabilitätssbegriffe*. Paderborn: mentis.

Schlosshauer, Maximilian. 2007. *Decoherence and the Quantum-to-Classical Transition*. Berlin/Heidelberg: Springer.

Stöckler, Manfred (2007). Philosophische Probleme der Quantentheorie. In: A. Bartels and M. Stöckler (eds.). *Wissenschaftstheorie. Ein Studienbuch*, 245–264. Paderborn: mentis.

Zeh, Heinz-Dieter. 1970. On the interpretation of measurement in quantum theory. *Foundations of Physics* 1: 69–76.

Zurek, Wojciech H. 1981. Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse? *Physical Review D* 24: 1516–1525.

Chapter 3
Quantum Identity and Indistinguishability

Holger Lyre

This chapter stands conceptually between Chaps. 1 and 6. In Chap. 1, we introduced single-particle quantum mechanics in a Hilbert space $\mathcal{H}$, while the present chapter treats n particles in a many-particle Hilbert space $\mathcal{H}_n$; and Chap. 6 deals with variable particle numbers using creation and annihilation operators in a Fock space $\mathcal{H}_F = \oplus\mathcal{H}_n$. The present chapter consists of two parts, of which Sect. 3.1 is more physical, while Sect. 3.2 has a stronger philosophical orientation.

3.1 The Quantum Theory of Similar Objects

3.1.1 Statistical Mechanics

Quantum theory has its historical origins in large part in thermodynamics. The problem of black-body radiation motivated Max Planck in 1900 to formulate a new *ad hoc* rule for the energy distribution of the radiation field of a black body as a function of the radiation frequency, and thereby to introduce his "quantum of action" (Planck's constant) as a new constant of nature. This achievement is generally considered to represent the birth of quantum theory, and one can say that the problem of black-body radiation could be seen as an anomaly for classical physics (in the sense that the classical Rayleigh–Jeans radiation law leads to an infinite radiation energy density with increasing frequency, known as the "ultraviolet catastrophe"). Another anomaly was presented by the entropy of mixing of identical gases—and it leads us to the topic of this chapter: Are physical objects *individuals*, i.e. do they possess both a diachronic identity which persists over time, and also a synchronic identity in the sense of being distinguishable from other objects at a particular moment in time?

C. Friebe et al., *The Philosophy of Quantum Physics*,
https://doi.org/10.1007/978-3-319-78356-7_3

Consider two vessels filled with gas and separated by an (impermeable) membrane. Pressure and temperature are assumed to be the same in both vessels. When the membrane is removed, the gases mix, but the internal energy of the overall system remains constant. If the two gases are chemically different, the process is irreversible, and the entropy, a measure of the number of microstates in a given macrostate, increases.[1] A mixture of identical gases should, however, not lead to any increase in the entropy, since the membrane can again be inserted reversibly. This result is in accord with the macroscopic view of phenomenological thermodynamics. But from the microscopic viewpoint of the kinetic theory of gases, or of statistical thermodynamics, the gas consists of molecules, its entropy is derived from Maxwell–Boltzmann statistics, and it depends upon the number of microstates within the macrostate. When two gases are mixed, it is evidently microscopically irrelevant whether or not the molecules belong to different chemical species; in either case, the molecules are individually mixed. According to this viewpoint, one should obtain an increase in the overall entropy due to the entropy of mixing, even when both vessels initially contained chemically identical gases—this however contradicts the result expected from phenomenological thermodynamics. This was first pointed out by Josiah Willard Gibbs near the end of the nineteenth century; the problem is thus known as the *Gibbs paradox*.

Formally, the contradiction can be resolved by using the so-called *Gibbs correction factor* $\frac{1}{N!}$ when counting the microstates (whereby N denotes the number of particles). Evidently, this factor signifies that the number of microstates must be divided by the number of their possible permutations, which means that the impossibility of capturing individual microstates must be accounted for.

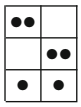
This can be illustrated in an intuitive manner: Consider the combinatorially possible distributions of two particles a and b over two states, e. g. states of different energies, represented here as boxes:

$$\begin{array}{ll} (1) & \boxed{ab}\,\boxed{} \\ (2) & \boxed{}\,\boxed{ab} \\ (3) & \boxed{a}\,\boxed{b} \\ (4) & \boxed{b}\,\boxed{a} \end{array} \tag{3.1}$$

In the first and second cases, both particles are in the same energy state, while in the third and fourth cases, they have differing energies. Counting the occupation possibilities now depends on how we weight cases (3) and (4). If, as is suggested by the Gibbs correction factor, the objects a and b can be determined only up to a permutation, then cases (3) and (4) are not distinguishable. Weighting with $\frac{1}{2!}$, they

[1] The macroscopic characterization of the state of a gas is defined by the state parameters pressure, volume and temperature; a microscopic description requires, in principle, knowledge of the positions and momenta of all of the individual molecules.

contribute on average only *one* state to the count. All together, instead of the four possibilities in (3.1), there are then only three possible occupations:

(1) (2) (3)  . (3.2)

The notation • indicates that the previously suggested unique identification of the two objects *a* and *b* has been abandoned. This means *de facto* that the two objects are empirically indistinguishable (at least for purposes of counting physical occupation possibilities).

Applying the Gibbs correction factor was found to be necessary in all areas of statistical mechanics in the following years, in order to obtain results which were in agreement with experiments. In particular, it was found to be necessary for constructing the quantum mechanics of many particles in the 1920s to propose an *indistinguishability postulate* of the following form:

> *The application of particle permutations to a many-particle state leads formally to a state which is physically indistinguishable from the original state.*

The mathematical implementation of this postulate is realized by requiring symmetry under permutations for all quantum-mechanical states. We will consider this further in the following two sections.

3.1.2 Many-Particle Tensor Products

In Chap. 1, we treated the quantum mechanics of a single particle, or more generally, of one object. We now consider the generalization to systems of arbitrarily many objects. We will concentrate our attention on particles of the same kind, such as a quantity of electrons or photons, or of comparable objects. In general, the states of a quantum-mechanical object with n eigenstates can be represented by vectors in an n-dimensional Hilbert space. If an n- and an m-dimensional object are composed into a larger object, a compound system, then the larger object generally has states defined on an $n{\cdot}m$-dimensional Hilbert space. We want to understand the peculiarities of this composition.[2]

We first consider the composition or combination of objects in classical physics. Here, the state space of one object is a six-dimensional phase space. It is isomorphic to $\mathbb{R}^6$ and is spanned by three position and three momentum coordinates. The phase space of a system consisting of two objects is correspondingly 12-dimensional and is given by the direct product $\mathbb{R}^6 \times \mathbb{R}^6 = \mathbb{R}^{12}$. In set theory, the direct product

[2]The presentation in the following two paragraphs is taken up again in Sect. 6.3.3 and extended there. For the purposes of the present chapter, the introduction of the permutation operator (3.4) is especially relevant.

corresponds to a Cartesian product, i. e. to the set of all ordered pairs: $A \times B = \{(a, b) \mid a \in A \wedge b \in B\}$. For a finite number of vector spaces, the direct product is the same as the direct sum of the vector spaces: $\mathbb{R}^n \oplus \mathbb{R}^m = \mathbb{R}^{n+m}$. This illustrates that the vectors from the individual subspaces of the product space are independent of each other and not correlated. Precisely this is different in a quantum-mechanical composite system.

In quantum mechanics, the combination of systems is described by the tensor product of their Hilbert spaces. Let $\{\vec{e}_i\}$ and $\{\vec{h}_j\}$ be basis systems of two n- and m-dimensional Hilbert spaces, $\mathcal{H}^n$ and $\mathcal{H}^m$; then the expression $\mathcal{H}^n \otimes \mathcal{H}^m$ denotes the *tensor product* of these two vector spaces. The product space $\mathcal{H}^{n \cdot m}$, in contrast to classical physics, has the dimensionality $n \cdot m$, and is spanned by the basis vectors $\vec{e}_i \otimes \vec{h}_j$. The crucial point is that a general vector $\vec{\psi} = \sum_{i,j} \alpha_{ij} \vec{e}_i \otimes \vec{h}_j$ of a tensor-product space is not written as the product of the basis vectors $\vec{e}_i$ and $\vec{h}_j$, but rather only as a linear superposition of them. In combining the subsystems to give a composite system via the tensor product, there are therefore relationships between the states of the subsystems, so-called correlations, which occur due to the superposition of the product states of the parts of the compound. This particular feature of quantum mechanics is called *entanglement*, and it leads to numerous unusual phenomena which will be treated in particular in Chap. 4. The present chapter focuses on the construction of many-particle states.

From what was said thus far, we can see immediately that an n-dimensional Hilbert space can be written as a tensor product of one-dimensional Hilbert spaces. The state space of n particles in quantum mechanics is thus the tensor product of n single-particle Hilbert spaces, $\mathcal{H}$:

$$\mathcal{H}^n = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes ... \otimes \mathcal{H}_n = \bigotimes_{i=1}^{n} \mathcal{H}_i . \tag{3.3}$$

The simplest basis vectors $\psi \in \mathcal{H}^n$ of an n-particle Hilbert space can be written as

$$\psi = \psi(1, 2, 3...n) = \psi_1 \otimes \psi_2 \otimes ... \otimes \psi_n, \quad \psi_i \in \mathcal{H}_i .$$

Because of the indistinguishability postulate introduced in Sect. 3.1.1 and justified heuristically there, it now holds that every permutation of identical particles in a many-particle state leads to a state which is physically indistinguishable from the original one. If $\hat{P}_{ij}$ denotes exchange of the particles i and j, then the state

$$\hat{P}_{ij}\psi(1, 2, ...i, ...j, ...n) = \psi(1, 2, ...j, ...i, ...n) \tag{3.4}$$

cannot be physically distinguished from the state $\psi(1, 2, ...i, ...j, ...n)$.[3] Due to the non-commutativity of the tensor product, the state of a many-particle system must have the general form

$$\Psi(1, 2, ..n) = \sum_P \frac{C(n)}{\sqrt{n!}} \psi(1, 2, ...n) \,,$$

where on the right-hand side, the sum is taken over $n!$ permutations (with a factor $C(n)$, which we shall not specify in detail here; it depends on the type of particles and may be a complex number).

The set of all permutations of an ordered set with n elements of dimensionality $n!$ is called the *symmetric group* S_n, or also the *permutation group*. The invariance of a state under the permutation group indicates that every physical observable $\hat{O}$ commutes with every permutation operator $\hat{P}$; no physically measurable quantity or observable can thus distinguish between a permuted and a non-permuted state. Formally, this means that

$$\langle \psi | \hat{O} | \psi \rangle = \langle \hat{P}\psi | \hat{O} | \hat{P}\psi \rangle = \langle \psi | \hat{P}^{-1} \hat{O} \hat{P} | \psi \rangle \,, \quad \text{or} \quad [\hat{O}, \hat{P}] = 0. \tag{3.5}$$

3.1.3 Quantum Statistics

As suggested in Sect. 3.1.1, the Maxwell–Boltzmann statistics of classical statistical mechanics must be replaced by a new kind of quantum statistics that takes the indistinguishability postulate into account. At the level of elementary particles, this requirement is realized in two ways—depending upon whether the particles have half-integer or integer spins.

In 1924, the Indian physicist Satyendranath Bose succeeded in deriving Planck's formula for the energy density of black-body radiation without referring to classical electrodynamics (in contrast to Planck himself), but instead based only on the assumption of light quanta with phase-space volumes of h^3 (cf. Darrigol 1991). He made use of the Gibbs correction factor in computing the occupation numbers of the states. Bose asked Einstein, who had introduced the light-quantum hypothesis in 1905, for help in publishing his results. Einstein recognized the significance of Bose's work and added some extensions to it. Following the introduction of Schrödinger's wave mechanics in 1926, the generality of the resulting *Bose–Einstein statistics* for many-particle wavefunctions was recognized; it describes particles with integer spins. Formally, Bose–Einstein statistics is based upon the indistinguishability postulate, which leads to the permutation invariance of bosonic states under the transformation (3.4).

[3] The permutation operator $\hat{P}_{ij}$ is self-adjoint and has the special property that its eigenvalues are 1 and -1 (compare, analogously, Sect. 1.2.4).

Since not ψ, but rather $|\psi|^2$ is related to observable quantities, the wavefunction is determined only up to a phase factor. Furthermore, for the permutation operator, it holds that $\hat{P}_{ij}^2 = 1$; it thus has eigenvalues ± 1 (see Footnote 3). Along with (3.4), this allows for the possibility that a permutation, in contrast to (3.4), leads to a change of *sign* and is thus *anti*symmetric:

$$\hat{P}_{ij}\psi(1, 2, ...i, ...j, ...n) = -\psi(1, 2, ...j, ...i, ...n). \tag{3.6}$$

It is again found empirically that antisymmetric wavefunctions belong to a particular sort of particles, namely those with half-integer spins. In order to give a quantum-mechanical explanation of atomic structure, Wolfgang Pauli in 1925 stated that no two electrons within an atom can have identical values of all four quantum numbers (which serve to define the state of an orbital electron). This so-called *Pauli exclusion principle* leads, together with the indistinguishability postulate, to the requirement that the many-electron wavefunction be antisymmetric under exchange of two electrons, as in (3.6).

The generalization of this requirement to arbitrary particles of half-integer spin, so-called fermions, leads to *Fermi–Dirac statistics*. It implies that for instance in the case of a system of two similar particles only the third possible state in (3.2) is allowed, whereby the wavefunction is to be written in antisymmetric form:

$$\Psi_A = \frac{1}{\sqrt{2}}\Big(\psi_a\psi_b - \psi_b\psi_a\Big)\,. \tag{3.7}$$

In contrast, Bose–Einstein statistics allows all three possibilities in the form of symmetrized wavefunctions:

$$\Psi_S^{(1)} = \psi_a\psi_a, \tag{3.8}$$

$$\Psi_S^{(2)} = \frac{1}{\sqrt{2}}\Big(\psi_a\psi_b + \psi_b\psi_a\Big), \tag{3.9}$$

$$\Psi_S^{(3)} = \psi_b\psi_b\,. \tag{3.10}$$

Bose–Einstein and Fermi–Dirac statistics are the two new types of quantum statistics which are needed for the computation of occupation numbers for quantum-mechanical systems; they replace the classical Maxwell–Boltzmann statistics. Here, we make two additional remarks: First, the requirement that exchange of two particles in a many-particle wavefunction be either symmetric (3.4) or antisymmetric (3.6) does not imply that all possible two-particle transpositions within a many-particle state occur only either symmetrically or antisymmetrically. In a purely formal

manner, we could consider also mixed-symmetric permutations, i.e. those in which some particles transform symmetrically, and others antisymmetrically. This leads to so-called parastatistics (cf. Messiah and Greenberg 1964). Such mixed-symmetric representations of the permutation group are, however, not realized in nature. The possibility of parastatistical states will therefore not be pursued further here; a many-particle wavefunction will be written in either totally symmetrized or else in totally antisymmetrized form.

Secondly, it can be seen—likewise purely empirically—that as in (3.4), symmetrized state functions describe particles with integer spins, that is bosons, and as in (3.6), antisymmetrized functions describe particles with half-integer spins, that is fermions. A more fundamental, theoretical justification of the connection between spins and statistics was given only at the end of the 1930s by Fierz and Pauli, in the form of the *spin statistics theorem* (and was later derived in different versions by various authors). The proofs are typically based on very general assumptions of relativistic quantum field theory such as locality and positive definiteness of the energy, which go beyond the present context (see Chap. 6).

The empirical relevance of Fermi–Dirac and Bose–Einstein statistics is demonstrated tangibly in scattering experiments with particles, as was shown in an intuitively clear manner by Feynman in his famous *Lectures* (Feynman *et al.* 1964, Chap. 4). Consider two empirically distinguishable particles a and b. Particle a scatters in direction 1 with an amplitude of $a_1 = \langle a|1\rangle$, while particle b scatters in direction 2 with an amplitude $b_2 = \langle b|2\rangle$. The probability for the mutual double scattering is the product of the individual probabilities, $|a_1|^2|b_2|^2$. If, conversely, particle a scatters in direction 2 and b in direction 1, we would obtain $|a_2|^2|b_1|^2$, and for the overall probability

$$P = |a_1|^2|b_2|^2 + |a_2|^2|b_1|^2 \,.$$

Now we assume that the directions 1 and 2 approach each other; then the above expression is reduced using $a = a_1 = a_2$ and $b = b_1 = b_2$ to

$$P = 2|a|^2|b|^2 \,.$$

When a and b are indistinguishable bosons, we cannot distinguish experimentally between the double scattering process a in direction 1, b in direction 2; and the exchanged process a in direction 2, b in direction 1. We then obtain (following the rule "sum before you square")

$$P_b = |a_1 b_2 + a_2 b_1|^2 = 4|a|^2|b|^2 \,,$$

i. e. a probability which is twice as large as in the case of distinguishable particles. For *fermions*, on the other hand, the total transition amplitude must be taken to be antisymmetrized,

$$a_1 b_2 - a_2 b_1 \,,$$

so that in the case when directions 1 and 2 have approached closely, it follows that

$$P_f = 0 \,.$$

It is impossible to observe two indistinguishable fermions in the same final scattering state!

A further notable consequence of Fermi–Dirac statistics is that due to the exclusion principle, one can derive the so-called Fermi pressure or degeneracy pressure of a fermion gas at a high density. In a white dwarf star, this Fermi pressure opposes further compression by gravitational forces and thereby stabilizes the star. At least two interesting physical-philosophical questions can be posed in this connection, which we can only mention here without giving their answers. First: In what manner is the Fermi pressure really a pressure (in the mechanical sense)? Its origins lie finally in a symmetry requirement, not in the existence of an explicit interaction (in contrast to e. g. gravitational pressure). And secondly: To what extent could this already point to an ontological difference between fermions and bosons?

3.1.4 Symmetric Group

This section is directed at mathematically interested readers and can be skipped over in a first reading. Its principal purpose is to characterize our considerations thus far in terms of group theory.

The indistinguishability postulate acts as a "superselection rule" on the state space (3.3) of similar objects. It states that $\mathcal{H}^n$ decomposes into subspaces or sectors corresponding to the representations of the permutation group. In the interplay with the two previous empirical remarks, the indistinguishability postulate leads to the following *symmetrization postulate* (cf. Messiah 1979, Chap. 14):

> *The state space* (3.3) *in the quantum theory of n identical objects decomposes into the subspaces of either all completely symmetric, bosonic state functions or of all completely antisymmetric, fermionic state functions.*

Group Representations and Selection Rules

The *vector space-representation* of a *group* G refers to a homomorphic mapping of G onto an automorphism group of non-singular operators on a vector space $\mathbb{V}$. Under the action of the group, $\mathbb{V}$ decomposes into invariant subspaces (sectors), which means that the application of the group operators on states which span the representation space does not lead out of that space; thus the states transform only among themselves. There are therefore no possible state transitions from one sector to another. The representation (or the representation space) is called *irreducible* if no other subspaces exist which are invariant under the action of the group (except for the null vector and $\mathbb{V}$ itself, as trivial subspaces).

If the representation space is *degenerate*, it is called a *multiplet*. A state is degenerate when one eigenvalue of the Hamiltonian belongs to more than one eigenfunction. A multiplet is thus a set of degenerate states with the same energy eigenvalue. The degree of degeneracy corresponds to the dimensionality of the representation. The Hamiltonian is then invariant under the group, i. e. it commutes with all the operators in the group. For the permutation group, we have already seen this at the end of Sect. 3.1.2.

Selection rules tell us whether or not particular transitions between states are allowed. Forbidden transitions correspond to vanishingly small transition probabilities for the corresponding perturbation operators $\hat{A}$:

$$\langle\psi|\hat{A}|\phi\rangle = \langle\phi|\hat{A}|\psi\rangle = 0.$$

Hindrance of state transitions can be due to the fact that the state space obeys certain symmetries and is thus characterized by conservation laws. For example, conservation of the angular momenta and spins of atoms determines the various selection rules for computing the strengths of spectral lines in atomic physics. Selection rules thus also characterize the decomposition of the Hilbert space into incoherent sectors, whose states cannot form superpositions. When such rules apply not just to certain operators, but instead to all the measurable observables $\hat{A}$, they are called *superselection rules*. Superselection rules apply to strictly conserved quantum numbers – in the case of permutation symmetry, the boson and fermion numbers. In the framework of supersymmetry, these quantum numbers are abandoned and, instead, transitions between bosons and fermions are postulated. This is, however, thus far only a hypothetical symmetry.

Compact explanations of group representations and selection rules are given in the grey box on p. 81. Here, we consider a simple example: Since a quantum object with only two states corresponds to an irreducible representation of the group $SU(2)$, the states (3.7) and (3.8)–(3.10) span the subspaces of the tensor product of two irreducible $SU(2)$ representations. One can then see that the tensor-product space decomposes into a one-dimensional and a three-dimensional, irreducible subspace. Or, expressed differently, the tensor product of two fundamental $SU(2)$ representations is decomposed into the direct sum of a singlet and a triplet, schematically denoted as

$$(2) \otimes (2) = (1) \oplus (3)\,.$$

Now, we are interested here in the representations of S_n; they are related to $SU(n)$. While (3.7) forms a one-dimensional (antisymmetric) representation of S_2, (3.8)–(3.10) span a three-dimensional (symmetric) representation space. The latter is reducible, since each of the states (3.8)–(3.10) corresponds to a one-dimensional irreducible representation. In general, it holds that all of the totally symmetric and totally antisymmetric irreducible representations of S_n are one-dimensional (in our example, the application of the permutation operator to each of the states (3.7)–(3.10) does not lead out of the one-dimensional ray $c\psi$, where ψ is one of the states (3.7)–(3.10)). In contrast, the irreducible mixed-symmetric representations which occur from S_3 on are higher-dimensional. Thus, S_3 decomposes the space of states $\mathcal{H}^3$ of three particles irreducibly into one each of a totally symmetric and a totally antisymmetric representation, as well as two two-dimensional mixed-symmetric representations, that is doublets (which we will not consider further here).[4]

What is suggested here using the example of the connection between S_2 and $SU(2)$ holds quite generally: The multiplicity of an irreducible representation of S_n is equal to the dimensionality of the irreducible representation of $SU(n)$, and vice versa.[5] The representations of S_n and $SU(n)$ can be elegantly illustrated graphically by so-called Young schemata (or Young tableaux; the interested reader is referred to Messiah 1979, Appendix D.4, and also Weyl 1950, Chap. V, Sect. 13).

3.2 Ontology of Quantum Theory

In Sect. 3.1, the consequences of the empirical fact of the physical indistinguishability of particles or objects for the formal apparatus of quantum theory were treated. Now, we want to discuss the implications of that fact in regard to ontology (or also

[4] See Exercise 1.

[5] Here are some additional examples without further comments: The tensor product of three fundamental $SU(2)$ doublets decomposes into a doublet and a quartet: $(2) \otimes (2) \otimes (2) = (2) \oplus (4)$. The tensor product of two fundamental $SU(3)$ triplets decomposes into a triplet and a sextet: $(3) \otimes (3) = (3) \oplus (6)$. And for the tensor product of three fundamental $SU(3)$ triplets, we obtain: $(3)\otimes(3)\otimes(3) = (1)\oplus(8)\oplus(8)\oplus(10)$. Correspondingly, S_3 has 1 one-dimensional antisymmetric, 8 two-dimensional mixed-symmetric and 10 one-dimensional symmetric irreducible representations.

contemporary metaphysics). Ontology is that branch of philosophy which investigates the nature of Being and its basic categories; that is, what exists, and in what modes and manners does it exist.[6]

3.2.1 Identity and Leibniz's Principle

3.2.1.1 Identity and Individuality

At the centre of ontology are questions of identity and individuality[7]; not a few philosophers even link the very possibility of an ontology quite closely to the question of identity, or of how entities can be individuated. In this connection, Willard van Orman Quine's well-known dictum, "No entity without identity" (Quine 1969, p. 23) is notorious.

As a preliminary, we make a few remarks on the terminology: An *entity* denotes in philosophy any sort of form of being, whether concrete or abstract.[8] For physicists, it is often of primary importance to describe *physical systems*. The term *system* is mostly used in this connection in the sense that a system can consist of subsystems. Elementary systems which cannot be further decomposed are, for example, *elementary particles*. If one wishes to avoid adopting a particle ontology from the beginning, then it is better to speak more generally of physical *objects*. The question treated in the present chapter, quite generally expressed, is what comprises the numerical distinctness of quantum objects, and how they can be individuated.

We must also be mindful of the usage of the concept *identity*. In physics texts, one often sees the term *identical particles* instead of *indistinguishable particles*. Such a manner of speaking is confusing, for if several particles are identical, then we are dealing not with several, but rather with *one* particle. David Lewis expresses this very concisely in his inimitable manner:

> Identity is utterly simple and unproblematic. Everything is identical to itself; nothing is ever identical to anything else except itself. There is never any problem about what makes something identical to itself; nothing can ever fail to be. And there is never any problem about what makes two things identical; two things never can be identical. (Lewis 1986, pp. 192–193)

---

[6] See Loux (1998), a highly recommendable introduction to modern ontology, Castellani (1998) as a useful collection of texts, and especially French and Krause (2006), a comprehensive account of the questions of quantum ontology which are treated in the following sections.

[7] These two concepts are used here as synonyms to a great extent (however, caution is recommended for interpreting the quotes in Sect. 3.2.2).

[8] Entities are termed *concreta* or particularia when they are localizable in spacetime; this includes in particular all of the everyday objects around us, as well as other physical things. *Abstracta*, in contrast, do not exist in space and time, and typically are causally inert. Among the standard examples, we could mention sets, numbers, propositions, possible worlds or abstract concepts such as love, God or the Good. But also properties, insofar as they are conceived as universalia, are abstract. They however possess the possibility of being realized at locations in spacetime (i. e. "instantiated").

In this sense, we are concerned with empirically indistinguishable objects and the related concept of object-identity—speaking of *similar objects* is therefore more accurate, although unfortunately less commonly used.

There are three modes of individuation of physical objects, namely by means of

(1) sets of properties;
(2) spatiotemporal localization, or spacetime trajectories;
(3) primitive (irreducible) metaphysical identity.

Now let us consider two particles a and b, with spacetime trajectories γ_a and γ_b, respectively, within a given volume of space. If these particles are of different species, i. e. they are of different types, then there is at least one property with respect to which they differ.[9] In order to avoid difficulties with composite properties or those derived from more elementary properties, we can concentrate here on the fundamental properties of elementary particles, i. e. rest mass, charge and spin. Distinguishing, let us say, electrons from photons presents no difficulties owing to their different fundamental properties; this would be an example of individuation in the sense of (1).

What, however, is the situation when a and b are similar, i. e. for example two electrons? We can then try to make use of their spatiotemporal localization to achieve a synchronic differentiation. Individuation in the sense of (2) signifies that the spacetime trajectory can serve to establish diachronic identity. Schopenhauer spoke, following Kant, of space as *principium individuationis*. But Kant himself already noted that spatiotemporal individuation presupposes a further assumption, namely that of the *impenetrability* of the objects considered.[10]

---

[9] *Properties* are to be understood here initially as empirically accessible in principle (although not necessarily directly observable). In an empirical science such as physics, it appears ostensibly that only and always such properties are meant; in metaphysics, however, properties can be more fundamentally distinguished, beyond their empirical nature, as for example the distinction between properties as *universalia* and as *tropes* indicates – this will be treated more thoroughly later on. Initially, however, properties in the sense of physics are understood to be universalia which are instantiated at spatiotemporal points, that is *in re*. By way of explanation: Since universalia, as emphasized in footnote 8, are abstracta, then in principle also non-instantiated properties, so-called universalia *ante rem*, can also be considered. The property of being a unicorn would be an example. Strict universalia-realists incorporate such properties into their ontology (for reasons which we cannot discuss here). But in this chapter, we will refrain from considering such possibilities.

[10] In his *Lectures on Metaphysics* (1790), Kant writes about space as an individuation principle: "Objects in space are therefore already *plura*, because they are in space" (28:569–570; Cambridge edition, edited and translated by K. Ameriks and S. Naragon). This is followed by the consideration of two droplets of water, which can also be found in the Supplement on the "*Amphiboly of Concepts of Reflection*" in Kant's *Critique of pure Reason* (A 1781/B 1787):

> ...the difference of the places of these appearances at the same time is still an adequate ground for the numerical difference of the object (of the senses) itself. Thus, in the case of two drops of water one can completely abstract from all inner difference (of quality and quantity), and it is enough that they be intuited in different places at the same time in order for them to be held to be numerically different. (A263–264/B319–320; Cambridge edition, edited and translated by P. Guyer and A. W. Wood)

Let us assume that the objects *a* and *b* could completely interpenetrate each other, so that they would henceforth occupy the same position in spacetime, or, in the case of extended objects, the same spacetime region (cf. also Della Rocca 2005). Call the entity in this position or region *X*. What justification would we then have to say of *X* that it consists of two objects, and not of one—or else a thousand—objects? If we individuate exclusively in the sense of (2), we would have to say, strictly speaking, that *a* and *b* have lost their identities at the moment of complete interpenetration, and instead the new object *X* was created at that moment. We could avoid this only by requiring impenetrability of the objects in addition to (2). That furthermore the topology of the spacetime also plays a role, or conventionalist elements related to the basis of physical topology and geometry, will be treated at the end of the following section.

3.2.1.2 Leibniz's Principle

One could raise the objection that locations and distances, that is the localization of an object in space, belong as well to the *properties* of the object. Individuation in the sense of (2) would then be reduced to individuation in the sense of (1). We therefore want to look into the idea that identity can be determined by the equivalence of properties in more detail. This idea forms the basis of Gottfried Wilhelm Leibniz's well-known principle of the identity of indiscernibles (usually abbreviated as PII: *principium identitatis indiscernibilium*)[11]:

> If, for all properties F, object x has property F if and only if object y has F, then x and y are identical.

In formal notation, the PII reads:

To be sure, Kant's position can be understood in detail only in terms of his transcendental philosophy, which we cannot discuss further here, and according to which physics does not refer to things-in-themselves, but rather to their appearances on the basis of space and time as pure forms of intuitions and the categories of understanding. On the topic of impenetrability, one finds in Kant's *Metaphysical Foundations of Natural Science* from 1786 the following Proposition 3 in the second part on dynamics:"Matter can be compressed to infinity, but can never be penetrated by a matter, no matter how great the compressing force of the latter may be" (AA IV:501; Cambridge edition, edited and translated by M. Friedman).

[11] In his *Primae veritates*, Leibniz writes: "Sequitur etiam hinc non dari posse duas res singulares solo numero differentes" (it follows even that there can be no two individual things which are merely numerically distinct). The continuation of this quote shows immediately that Leibniz considered his PII to be a consequence of what he held to be a still more fundamental principle, the *Principle of sufficient reason*: "utique enim oportet rationem reddi posse cur sint diversae, quae ex aliqua in ipsis differentia petenda est" (for it must be possible to state a reason why they are distinct, which must be sought in some sort of difference between them). The significance of the PII was illustrated by Leibniz as he related it in a quite intuitive way to the ladies of the Herrenhausen Palace by challenging them to find two identical leaves, which they could not do (cf. C.I. Gerhardt (ed.), *Gottfried Wilhelm Leibniz: Philosophische Schriften*. 7 volumes, Berlin 1875–1890. Reprint by Olms, Hildesheim, 1960, p. 214).

$$\forall x, y: \quad (\forall F: \; Fx \leftrightarrow Fy) \;\Rightarrow\; (x = y). \tag{3.11}$$

Leibniz understood his principle in such a way that the properties which serve to quantify the objects are monadic and intrinsic. A property is intrinsic when it is attributable to an object independently of the existence of other objects and properties; in the converse case, it is extrinsic. Generic (although by no means uncontroversial) candidates for intrinsic properties are the masses, charges and spins of elementary particles. Extrinsic or relational properties, briefly: relations, depend on more than one entity and are therefore many-place (n-adic or n-ary). The relations "larger than" or "brother of" are paradigmatic examples of binary relations; "lies between" is ternary. Monadic properties, in contrast, are one-place or unary.

In the light of the above quote from Lewis, there is an air of paradox in our formulation of the PII, since initially we speak of *two* objects x and y, which are then asserted to be identical. The PII should therefore rather be expressed in its logically equivalent, contrapositive formulation of the *dissimilarity of the diverse*:

> Objects x and y are distinct in the case that x possesses at least one property which y does not possess, or *vice versa*.

Or still more simply: *No two objects share all their properties*; formally:

$$\nexists x, y: \quad (x \neq y) \wedge (\forall F: \; Fx \leftrightarrow Fy), \tag{3.12}$$

which is equivalent to (3.11).[12]

The question now arises as to what extent the PII represents a self-evident and possibly *a priori* comprehensible metaphysical posit. It is instructive in this connection to take a look at the logical converse of the PII, i.e. the *principle of the indiscernibility of identicals*: $\forall x, y: \; (x = y) \Rightarrow (\forall F: \; Fx \leftrightarrow Fy)$, or

$$\forall x, y: \quad (\forall F: \; Fx \not\leftrightarrow Fy) \;\Rightarrow\; (x \neq y)\,. \tag{3.13}$$

In words: *Discernable things can never be identical*. This requirement seems indeed to be evident. It is hard to imagine or construct possible worlds which violate it.[13] It

[12] This can be shown as follows:

$$\begin{array}{ll} & \neg\exists x, y: \neg(x = y) \wedge (\forall F: \; Fx \leftrightarrow Fy) \\ \Longleftrightarrow & \forall x, y: \quad \neg\big(\neg(x = y) \wedge (\forall F: \; Fx \leftrightarrow Fy)\big) \\ \Longleftrightarrow & \forall x, y: \quad (x = y) \vee \neg(\forall F: \; Fx \leftrightarrow Fy) \\ \Longleftrightarrow & \forall x, y: \quad (\forall F: \; Fx \leftrightarrow Fy) \;\Rightarrow\; (x = y) \end{array}.$$

[13] David Lewis, in his modal realism regarding possible worlds, maintains the view that in other such worlds there are "counterparts" to the entities in our world (e. g. to every reader of this book). Unlike, for instance, Plantinga, he therefore denies trans-world identity, the view that, e. g., a person in our world is identical with persons in other possible worlds (of which we are speaking when we say things like, "I was almost hit, but I managed to jump to one side at the very last moment"). Counterparts can be arbitrarily similar to each other, but not identical. Supporters of trans-world identity, according to Lewis, violate the principle of the indiscernibility of identicals (see Lewis 1986, pp. 198ff; Loux 1998, pp. 166ff).

thus represents a candidate of a principle which is valid in all metaphysically possible worlds.

This seems however not to hold for the PII, since it is quite possible to describe worlds in which the PII is violated. Such a world, a counterexample to PII, was discussed by Max Black using a well-known scenario (which in effect is a variation of Kant's water droplets; see Footnote 10). In Black's own words:

> Isn't it logically possible that the universe should have contained nothing but two exactly similar spheres? We might suppose that each was made of chemically pure iron, had a diameter of one mile, that they had the same temperature, colour, and so on, and that nothing else existed. Then every quality and relational characteristic of the one would also be a property of the other (Black 1952, p. 156).

Black's scenario begins with the assumption that spacetime locations cannot be used for the individuation of the two spheres. Such an assumption is fulfilled for example in a world in which space is relational. *Relationalism* about space holds that space is nothing more than the set of all possible relations between bodies, and that therefore, empty space is impossible. The opposite view, *substantivalism*, considers space or its constituents to be entities *sui generis*. In a relational space, locations and distances represent relational properties of the objects within it; in contrast, a point in substantivalist (absolute) space possesses its location *intrinsically*. In principle, there thus exists an absolute frame of reference in such a space.[14]

As already indicated at the end of the previous section, still another subfield of the philosophy of spacetime is associated with our topic: Spacetime *conventionalism*. The idea is that the geometry (and possibly also the topology) of the physical world are not empirical facts in themselves, but only the conjunction of spacetime geometry and the set of all physical laws. In order to determine the empirical geometry, we have to stipulate first certain conventions about the behaviour of measuring rods and clocks upon transport through spacetime. A given convention (e. g. the assumption that measuring rods act as rigid bodies with constant lengths) can, in principle, be replaced by any other convention, as long as corresponding adjustments are made in other parts of physics (e.g. in optics and electrodynamics concerning the paths of light). Thus, spacetime geometry is empirically underdetermined, and only the combination of geometry plus laws of nature is subject to empirical test.

Hacking (1975) points out that by making a suitable choice of the spatiotemporal structure, the PII can be defended against Kant's water droplets and Black's spheres. Black's world indeed allows only the following description: One sphere with intrinsic properties Q is given; from it, a sphere with properties Q at a distance of z sphere-diameters can be reached along a straight line. This scenario can now be described either by a world which contains two spheres within a Euclidean space, or else by a world which contains *one* sphere within a cylindrical space with a circumference of z (cf. Adams 1979, p. 15). Black's scenario of two spheres can thus be reinterpreted qua suitable convention into a scenario with only a single Leibniz-individuated sphere. Hacking claims that every objection to the PII can be reinterpreted such that the PII is in fact fulfilled.

[14] For an introduction to the philosophy of spacetime theories, see Dainton (2001).

In the following, we dispense with conventionalist provisos (see French 1975 for an explicit criticism of Hacking). Black's thought experiment then undermines the PII as a fundamental metaphysical principle for defining identity. One might concede that he considers only an imaginable world, but not the actual world. But even if it were true that within our actual world, no two things can be found which share all their intrinsic properties (which is precisely the essential point of the indistinguishability postulate of quantum theory, as we shall see in the next section); still, even the mere conceivability of such a scenario threatens the claim of the PII to being a fundamental metaphysical principle. A world in which a fundamental metaphysical principle is violated should not be conceivable at all, i. e. it should not be a logically possible world. Black's spheres, however, seem to be perfectly possible.

3.2.1.3 Bundle Ontology, Trope Ontology and Haecceitism

However, not only Leibniz's PII, but along with it a prominent account of an object ontology, the so-called *bundle ontology*, is threatened here, since it appears to be dependent on the PII. Bundle ontology maintains that objects are nothing other than bundles of properties. For the individuation of an object as a bundle of properties, the advocate of this position may refer to nothing more than just precisely those properties. In addition to it, properties such as "*a* is identical with itself" should be excluded as question-begging.

So far we have taken properties to be *universalia*. As noted in Footnote 9, this is however not imperative. A prominent opposing position considers properties as *tropes*, i. e. as particularized and individuated. Proponents of a trope ontology, especially nominalists, try to avoid a characteristic of universalia which they see as obscure: namely their abstractness or, in the case of instantiated universalia, their multiple localizability. For tropes are not only particular, but also numerically distinct, that is they are already individuated in an ontologically primitive sense. Tropes are property individuals. Thus, no one electron would have the same charge trope as any other electron, but only a precisely similar one. Since advocates of trope ontology typically propound a bundle theory as well, they conceive of things as bundles of tropes. Leibniz's principle is thereby trivially fulfilled: Since tropes are property individuals, no two things possess the same tropes (cf. also Sect. 6.5.2 on the trope-ontological interpretation of QFT).

Some manifest advantages and disadvantages of tropist and universalist ontologies stand in direct opposition to each other. In the framework of trope ontology, it is a disadvantage that, for example, the fact that every electron in the universe carries precisely the same charge as every other is inexplicable and therefore must be regarded as an ontological *factum brutum*; while universalists may see a natural explanation of this fact in the participation of every electron in the same charge universal. The price for this is paid by the universalist theory in assuming the existence of abstract entities with ominous instantiation conditions.

As we have seen, both conceptions of properties can be combined with a notion of bundles. A concrete physical thing, a *particular*, is then simply the bundle of all

its properties. For proponents of the trope theory, this option is virtually mandatory, since each individual trope can already be regarded as a *particular*. Basically, the theory of bundles must then consider the fact that at a point in spacetime (or within a very small and compact region of spacetime), numerous properties must evidently always be co-present and co-localized, in order to constitute an object bundle, as a *factum brutum* (e. g. in the case of the trope-bundle theory, the presence of the elementary-charge trope, the spin $\frac{1}{2}$ trope, and the electron-mass trope at the location of each electron). This is one of the reasons why in ontology, within the counter-position to bundle theory, *particularity*—and with it, and at the same time, *identity*, at least in the sense of numerical distinctness—is ascribed to the presence of a propertyless *carrier*, i. e. in substance or substrate theory (this carrier "bears" the properties of each concrete object and holds them together). In the philosophical tradition, numerous terms have been coined in order to address this position. It is sometimes referred to as *haecceitism*. The term *haecceitas* (from Latin *haec*: this) can be translated as "thisness", and is due to Duns Scotus. Other terms are "Lockean substance" (cf. French 1989), "primitive thisness" or "primitive identity" (Adams 1979), "transcendental individuality" (Post 1963), or simply "bare particular".

Using Black's spheres as an example, we can illustrate this idea. Since the spheres exhibit no intrinsic differences, but are nevertheless *two* spheres rather than one, they would seem to possess their identities *solo numero*: We can attribute to them a multiplicity or a cardinality, even though the spheres are not individually numerable, so that they have no ordinality. Another common example is the set of points of a spatiotemporal manifold. *Manifold substantivalism* is a variant of spatiotemporal substantivalism according to which the ontological constituents of spacetime are the points within the spatiotemporal manifold. As points, they themselves have no sort of properties; they are completely homogeneous. For their individuation, manifold-substantivalists therefore consider spacetime points as haecceistic entities.[15]

Haecceitism represents the third option in the framework of the above list (p. 84). Individuation in the sense of (1) and (2) turned out to be problematic, as we have seen. In both cases, the attempt is made to reduce identity to other quantities such as sets of properties, or spatiotemporal behaviour. Individuation in the sense of (3) sees identity as ontologically irreducible and primitive. From an haecceitist point of view, it is reasonable to ask whether or not a certain individual exists in another possible world, without referring to properties or spatiotemporal behaviour; and thus without reducing individuality to (1) or (2). Haecceistic differences between possible worlds are therefore differences which are not based upon distinctions in properties or behaviour. But precisely this aspect makes haecceitism appear rather obscure to every empirically oriented metaphysicist. In any case, it seems, the concept of identity has its price.[16]

---

[15] Leibniz argued precisely against this view in his well-known debate with Clarke (and Newton), appealing to the PII, and thereby to the overriding authority of the *Principle of sufficient reason*.

[16] In the light of trope ontology, it is seen that along with property individuation in the sense of (1), whereby properties are to be understood as universalia, and haecceistic individuation in the sense of (3), a third option arises: that is individuation via tropes. Evidently, it is a kind of combination

3.2.2 Leibniz's Principle and Quantum Theory

Let us consider once more the example given above of the individuation of two particles a and b with spacetime trajectories γ_a and γ_b, respectively. We wish in addition to assume that the spatiotemporal behaviour of the particles is *chaotic*; i. e. that the trajectories γ_a and γ_b are thus practically unpredictable. Let us also assume that the locations of the particles are accessible to the observer only at discrete time intervals, then there is evidently no chance of recognizing a and b again. Numerous interpretations of quantum theory, most prominently the Copenhagen interpretation, deny the existence of well-defined spacetime trajectories.[17] For quantum objects, neither an individuation in the sense of (1) nor in the sense of (2) seems to be possible.

Such an argumentation scheme can still be found today in numerous quantum mechanics textbooks; for example, such influential authors as Landau and Lifschitz write:

> In classical mechanics, identical particles (electrons, say) do not lose their 'individuality', despite the identity of their physical properties. [...]
>
> In quantum mechanics the situation is entirely different... We have already mentioned several times that, by virtue of the uncertainty principle, the concept of the path of an electron ceases to have any meaning. If the position of an electron is exactly known at a given instant, its co-ordinates have no definite values even at an infinitely close subsequent instant. Hence, by localising and numbering the electrons at some instant, we make no progress towards identifying them at subsequent instants; if we localise one of the electrons, at some other instant, at some point in space, we cannot say which of the electrons has arrived at this point.
>
> Thus, in quantum mechanics, there is in principle no possibility of separately following each of a number of similar particles and thereby distinguishing them. We may say that, in quantum mechanics, identical particles entirely lose their 'individuality'. (Landau and Lifschitz 1965, p. 209).

It is decisive that the argumentative focus lies here not on the question of the simultaneous distinguishability of similar particles, but instead on their recognizability over time. In a systematic sense, we can distinguish between *synchronic identity* (at a certain moment in time) and *diachronic identity*, that is the persistence of an object over time. This distinction is tentatively based upon the alternative of an individuation in the sense of (1) as opposed to (2). The problem of re-recognizability was discovered early on by the founding fathers of quantum theory; an explicit connection with Leibniz's principle was made by Hermann Weyl, in particular. His discussions of

---

of (1) and (3), since tropes are on the one hand properties, but on the other they are irreducible and primitive. Astoundingly, this option is not considered in the debate on quantum identity and Leibniz's principle explicitly anywhere in the literature. The reason is no doubt that differences in similar types of tropes—for instance the elementary-charge tropes of two electrons—represent not empirical, but merely metaphysical differences (similarly to haecceities).

[17] In Bohm's theory, the situation is different, as shown by Brown *et al.* (1999). To be sure, here the possibility of maintaining individuation in the sense of (2) comes at the price of building the topologies of non-crossing spacetime trajectories directly into the structure of the configuration space and of Bohm's guidance equation. The requirement of mutual impenetrability is in this sense ontologically primitive.

this topic are at first glance somewhat enigmatic; Muller and Saunders (2008) also stumble over them. If one traces these discussions historically, it becomes clear how the view of Leibniz's principle in relation to quantum mechanics changed over time. In his *Theory of Groups and Quantum Mechanics* (German original from 1928), Weyl argues:

> ...the possibility that one of the identical twins Mike and Ike is in the quantum state E_1 and the other in the quantum state E_2 does not include two differentiable cases which are permuted on permuting Mike and Ike; it is impossible for either of these individuals to retain his identity so that one of them will always be able to say "I'm Mike" and the other "I'm Ike." Even in principle one cannot demand an alibi of an electron. In this way the Leibnizian principle of *coincidentia indiscernibilium* holds in quantum mechanics (Weyl 1950, p. 241).

The conclusion reached in the last sentence initially appears unintelligible: If Mike and Ike are indeed indistinguishable and "it is impossible for either of these individuals to retain his identity", and we are nevertheless dealing with *two* individuals, is then Leibniz's principle not violated? If we look at later publications, this confusion seems even to increase; in 1949, Weyl writes:

> The upshot of it all is that the electrons satisfy Leibniz's *principium identitatis indiscernibilium*, or that the electronic gas is a 'monomial aggregate' (Fermi–Dirac statistics). In a profound and precise sense physics corroborates the Mutakallimūn; neither to the photon nor to the (positive and negative) electron can one ascribe individuality. As to the Leibniz–Pauli exclusion principle, it is found to hold for electrons but not for photons (Weyl 1949, pp. 247).

A confirmation of Leibniz's principle is again mentioned, this time however with the remark that the electron gas (and, based on the preceding passages, in a similar sense the photon gas as well) forms a "monomial aggregate", that is, a whole. In addition, Weyl's use of the term "Leibniz-Pauli exclusion principle" is remarkable (we take up this topic in more detail below).

We can however indeed find a consistent reading of these passages. Weyl evidently does not have Black's scenario (as a counterexample to PII) in mind; rather, he is considering the question of re-recognizability over time, in a similar manner to that of Landau and Lifschitz later; he thus discusses the diachronic identity in contrast to the synchronic identity. Now, the two scenarios of Mike and Ike *before* and *after* their exchange are empirically indistinguishable, and therefore they must be counted physically as a single scenario—in the sense of the PII.

The numerical distinctness of the electrons at a certain moment in time, i. e. their synchronic identity, is not at all questioned by Weyl. Perhaps he has tacitly assumed an haecceistic position; this remains open in the passages quoted. In his later book "Symmetry", in 1952, he writes:

> I told you that Leibniz had given the geometric notion of similarity this philosophical twist: Similar, he said, are two things which are indiscernible when each is considered by itself. Thus two squares in the same plane may show many differences when one regards their relation to each other; for instance, the sides of the one may be inclined by 34° against the sides of the other. But if each is taken by itself, any objective statement made about one will hold for the other; in this sense they are indiscernible and hence similar (Weyl 1952, pp. 127–128).

Once again: The two squares can indeed be distinguished relationally insofar as they are two; in view of their "squareness", however, they are essentially similar, that is indistinguishable. If Weyl had followed up the question of the relational distinction of the squares, if he had thus shifted the focus onto their synchronic identity, then he would quite possibly have arrived at considerations which closely approach those of the later debates, such as those described in Sect. 3.2.3. Instead, in the early debates on quantum identity, the emphasis lies on a holistic aspect which consists in the fact that a multiplicity of similar quantum objects forms a "monomial aggregate", in the sense that no physical distinction can be made between the state $|\text{Mike}\rangle + |\text{Ike}\rangle$ and its permutation $|\text{Ike}\rangle + |\text{Mike}\rangle$; and they therefore, according to Leibniz, form a *whole*. Correspondingly, any physically reasonable talk about the (individual) parts of such wholes is prohibited. Max Born in 1927 writes that particles "[are] in many cases not identifiable at all as individuals, e. g. when they combine into an atomic union" (Born 1927, p. 240). Ernst Cassirer takes this theme up in 1937:

> The impossibility of delimiting different electrons from one another, and of ascribing to each of them an independent 'individuality', has been brought into clear light through the evolution of the modern quantum theory, and particularly through the considerations connected with the 'Pauli exclusion principle'. Considered solely from the standpoint of its methodological significance in the construction of the quantum theory, Pauli's exclusion principle is strangely analogous to the general principle introduced into philosophy by Leibniz under the name of *principium identitatis indiscernibilium*. This principle states that there cannot be two objects which completely correspond to each other in every determining characteristic, and thus are indistinguishable except by mere number. There are no things that differ from each other 'solo numero'; rather every true difference must be definable as a qualitative difference, a distinction of the attributes and conditions that constitute the object. The Pauli principle is, as it were, the *principium identitatis indiscernibilium* of quantum theory. It characterizes every electron within the atom by a definite complex of conditions, by ascribing to it four quantum numbers that completely determine its orbit. Furthermore it states the conclusion that electrons that show no differences in this respect are to be regarded as a single physical entity. (Cassirer 1956, p. 184–185, Footnote 17; German original 1937).

Just like Weyl, Cassirer finds an analogy between the principles of Leibniz and Pauli. According to the Pauli principle, no two fermions can have precisely the same quantum numbers, i. e. properties. This looks like an analogy to the PII—and indeed, also in the sense of synchronic identity. Since however a similar exclusion principle does not hold for bosons, we must make a distinction here. The quoted passages, both of Weyl and also of Cassierer, are unsatisfying in this regard. This is related not only to the distinction between fermions and bosons, but also to the question of whether we can apply Leibniz's principle to diachronic or to synchronic individuation. In the further course of the debate, the latter shifts into the centre of the discussions, with the result that now, to some extent contradictory consequences are drawn (at least in a literal interpretation of the statements made).

Henry Margenau, in 1944, speaks explicitly of a contradiction between the Pauli principle and Leibniz's principle, and of a violation of the latter. He first says:

> This conclusion recalls Leibnitz' principle of the identity of indiscernibles; indeed physicists have occasionally thought that the E.P. [exclusion principle] implies this principle with regard to elementary particles of the same species.

Unfortunately, Margenau leaves open the question of which physicists he is referring to (but it is most likely that he was thinking among others of Weyl). He then continues:

> ... the E.P., so far as it goes, contradicts Leibnitz [...] two particles, as we have seen, differ in no observable respect. Nevertheless quantum mechanics would lead to entirely erroneous results if they were treated as a single entity. The particles, though they can not be labelled individually, can be counted. If and only if identity were understood as not implying numerical identity, then two electrons in an atom could be said to be identical (Margenau 1944, p. 202).

Clearly, the synchronic question is finally addressed, and with it the question of how it is possible that two particles, although they differ "in no observable respect" (i.e. they are empirically indistinguishable), nevertheless are still *two* particles. This apparently admits of only one conclusion: In synchronic terms, the PII is violated.

In the following years, precisely this viewpoint—especially within the philosophy of science—finds an increasing number of adherents (see e. g. Post 1963, Cortes 1976, Teller 1983, French and Redhead 1988 and Butterfield 1993). Castellani and Mittelstaedt write: "[I]t is also commonly held that a form of the principle of the identity of indiscernibles is valid in the domain of classical physics, while the principle is inapplicable in the quantum case"; and they add in a footnote: "This is undoubtedly the prevailing position in the literature" (Castellani and Mittelstaedt 2000, p. 1589). In this respect, we must agree with Steven French when he refers to the violation of Leibniz's principle in quantum theory in his encyclopaedia article on "Identity and Individuality in Quantum Theory" (French 2011) as a "received view" (for the continuation of the discussion of this problem complex in QFT, see 6.4.2).

If, however, the PII is violated in quantum theory, what follows from this? A series of questions has to be dealt with separately, among others:

1. In which sense can quantum objects (not) be individuals?
2. Do quantum objects violate Leibniz's principle?
3. Is there an ontological difference between fermions and bosons?
4. In which sense does a many-particle state consist *de facto* of many particles (or must it not be regarded instead as a whole)?

The motivation for these questions appears in various forms in the confusing interplay of the quotes given above. They will be treated in more detail in the following. In connection with the third question, we also return in Sect. 3.2.3 to the second question. They will be affirmed only provisionally here. The fourth question is taken up again in Sect. 3.2.4. Let us initially turn to the first question.

Are quantum objects non-individuals? What is that supposed to mean? Indeed, quantum theory would seem to let Black's scenario become reality: Similar quantum objects possess a cardinality, but no ordinality. This is a direct consequence of the postulate of indistinguishability—quantum objects of the same type are in states for which it is possible to know the overall number of objects, although the objects themselves are empirically indistinguishable. To use a didactic picture: The amount of change in my pocket is Leibniz-individuated in terms of the individual coins of various denominations. This, however, does not hold for the same sum of money in

my bank account.[18] It remains unclear whether this is a permissible analogy or not; whether mathematical conditions can be mapped onto the world (see also Sect. 3.2.4 on this question), and whether we have captured the essence of cardinality in this way. Quantum objects would appear to possess a kind of distinctness *solo numero*—in contradiction to Leibniz's principle. Does it follow from this that quantum theory verifies haecceitism? Do quantum objects then possess a primitive identity, which can be expressed neither in terms of their spatiotemporal history nor in terms of their properties? Interestingly, only a few authors have explicitly drawn this conclusion, although the conviction that quantum mechanics violates the PII has become predominant (cf. French and Redhead 1988). In recent times, several authors have even turned to a revision of logic, or to set theory (see Sect. 3.2.4).

3.2.3 Weak Discernability

Both Weyl and Cassirer have indicated an analogy between the principles of Leibniz and of Pauli. Pauli himself was by no means in agreement with this; in 1949, he writes to Fierz correspondingly, that Leibniz's principle, as a metaphysical principle, indeed can have no empirical consequences; and then, literally:

> That would indeed be a curious principle in the philosophy of Leibniz, *which doesn't hold for all objects* (e. g. not for photons, as expressly emphasized by Weyl), but instead only for *some* objects.[19]

Weyl also receives a letter from Pauli. The letters in fact show that Pauli was not very well informed about the PII, but he touches upon the important question dealt with above, whether there is a difference between bosons and fermions in regard to the PII and to ontology. This point deserves a more detailed treatment.

The Pauli principle states that no two fermions can be in precisely the same state, that they thus must differ in at least one of their quantum numbers (i. e. in at least one property). Thus, it seems that Leibniz's principle is obeyed by fermions. Why, then, do Margenau and the "received view" come to the opposite conclusion? The application of a permutation operator to a fermion state (3.6) leads merely to a change of sign; otherwise it is similar to the application to a boson state (3.4). Since

---

[18] Schrödinger made use of this illustration in 1949: "... the shillings and pennies in your bank account are not individuals". (quoted after French and Krause 2006, p. 122), and Mary Hesse writes:

> With pounds, shillings, and pence in a bank balance, however, it is not merely the case that we cannot in practice re-identify a given pound appearing in the credit column, but that there is no sense in speaking of the self-identity of this pound, and of asking where it reappears in another column or whether it is the pound paid over the counter yesterday (Hesse 1966, p. 49–50).

[19] Translated from von Meyenn (1987); see in particular the second section on the exclusion principle and on the discernability of particles.

furthermore every physical observable $\hat{O}$ commutes with every permutation operator $\hat{P}$ (3.5), the expectation values of all operators for a single fermion in a many-fermion state are the same. In this regard, there is no difference between fermions and bosons: empirically, we can neither distinguish between single bosons in a many boson-state nor between single fermions in a many-fermion-state.

In spite of this at first apparently convincing argumentation in support of the "received view", the debate over the status of Leibniz's principle in quantum theory has received quite unexpectedly a new thrust in the early 21st century. A decisive impulse was given by Simon Saunders (2003, 2006). Saunders takes up earlier work of Quine on weak discernability. So far, the assumption has been made that, in the PII, one should quantify over monadic, intrinsic properties. This is, however, an unnecessarily stringent requirement. Relaxing this requirement, the Leibniz principle can be formulated with varying degrees of strength, depending upon what sort of properties serve for the quantification. Besides intrinsic properties, we could also consider relational properties, whereby in particular ordering relations and irreflexive relations lead to interesting extensions of the concept of indiscernability.

According to Quine (1976), we can characterize three kinds of discernability: *absolute, relative and weak discernability*. They are defined as follows:

- *Absolutely discernable objects* differ in at least one of their monadic properties.
- *Relatively discernable objects* are different with respect to at least one ordering relation.
- *Weakly discernable objects* are different with respect to at least one irreflexive relation.

Some paradigmatic examples: The natural numbers are absolutely discernable. In contrast, the moments in time along the time axis (the "arrow of time") are in fact intrinsically equivalent, and thus not absolutely discernable, but rather relatively discernable with respect to an "earlier–later" relation. Weakly discernable objects require irreflexive relations, whose definition we first give here:

> A relation R is reflexive, if for all x in the domain, $R(x, x)$ holds. In the case that $\neg R(x, x)$ holds, R is irreflexive.

Black's spheres at a distance d define an irreflexive distance relation: Every sphere is at a distance d from the other, but not from itself.

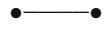
The various forms of discernability can also be illustrated by using graphs: In the simplest case of a "labelled graph" with two nodes and an edge

$$a \bullet \cdots\cdots \bullet b\,, \tag{3.14}$$

the two nodes are absolutely discernable. Here, "……" can be either a directed or an undirected edge. The unlabelled and directed graph

$$\bullet \longrightarrow \bullet \tag{3.15}$$

represents an example of relative discernability of the nodes; the unlabelled and undirected graph



(3.16)

is an example of weakly discernable nodes. In contrast, the two nodes of the edgeless graph

(3.17)

are not even weakly individuated.

The three variants of discernability illustrated above correspond to three forms of Leibniz's principle (with differing degrees of strength; each one is formulated in the contrapositive form employed above):

Strong PII: There exist no two individuals which are not absolutely discernable.
Moderate PII: There exist no two individuals which are not relatively discernable.
Weak PII: There exist no two individuals which are not weakly discernable.

In the previous sections, Leibniz's principle was considered in the strong sense only. Saunders (2006) and Muller and Saunders (2008) now show that fermions indeed violate the strong PII, but not the weak PII. Consider for example the antisymmetrized state function (3.6) of two fermions; in this concrete example, it could represent a spin singlet of two electrons, which otherwise share all their remaining quantum numbers,[20] with the spin orientations $|\uparrow\rangle$ and $|\downarrow\rangle$

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big(|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle\Big)\,. \tag{3.18}$$

The two electrons obey the irreflexive relation $R =$ "possess spins oriented oppositely to each other, but not to themselves". According to Saunders, the electrons, or more generally, fermions, are therefore weakly discernable; their identity is grounded in R via the weak PII.

If one follows this argumentation, then fermions are weakly discernable in the same way as Black's spheres; but not bosons. Saunders holds this to be unproblematic, since the most elementary fermions—leptons and quarks—are the fundamental matter particles, while elementary bosons appear only as gauge particles and as the Higgs boson. As such, they should be regarded not as objects, but rather only as excitation modes of quantum fields: "We went wrong in thinking the excitation numbers of the mode, because differing by integers, represented a count of things; the

[20] Such a state represents a didactic simplification, which is indeed commonly used, but which can lead to false conclusions. In particular, antisymmetry and EPR entanglement should not be confused. For the complete description of the state of an electron must also include the spatial degrees of freedom in addition to the spin degrees of freedom; otherwise, the state (3.18) could be misunderstood to imply that the two spins are located at the same point in space, which is obviously impossible for electrons. For particles, the complete state is a wavefunction in spin+position space. There, however, states which are formed by direct antisymmetrization of product states are not yet EPR entangled (in the sense of Chap. 4). Antisymmetry and EPR entanglement must be separated from each other conceptionally, as shown in detail by Ghirardi *et al.* (2002) and Friebe (2014).

real things are the modes" (Saunders 2006, p. 60). It is however questionable as to what extent such a separation of fermions and bosons is ontologically plausible; it is also questionable whether gauge bosons (in both massless and massive forms) and Higgs bosons are ontologically on a par, since the Higgs does not play the role of an intermediary field for the fundamental interactions.

But we need not delve into such questions, if we follow the argumentation of Muller and Seevinck (2009); they show that Saunders' result can be extended to bosons and, in fact, to all quantum objects. The basic idea is that quantum objects in a many-object state must necessarily obey certain Heisenberg commutation relations, e. g. the irreflexive relation "to have mutually complementary positions and momenta", or more generally "to contain canonically conjugated variables", as a result of the non-commutative algebra structure of quantum theory. Hence, it plays no role whether we consider fermions or bosons; all quantum objects obey the weak PII (see also Huggett and Norton 2014 for a refinement of these arguments).

The publications of Saunders and of Muller on weak discernability have stimulated anew the discussion of the status of the PII, and more generally of the ontology of quantum theory; but they have not been without controversy. This leads us over to the next section.

3.2.4 Outlook

Let us briefly recapitulate: The empirical result of the physical indistinguishability of the application of a particle permutation to a many-particle state, in the sense of the indistinguishability postulate, is beyond dispute. Likewise its formal consequences within the mathematical apparatus in the form of the symmetrization postulate and permutation invariance. In the discussion around Leibniz's principle, Weyl and Cassirer found it initially to be verified in diachronic terms, while following Margenau, it was generally agreed that in synchronic terms, the (strong) PII is violated in quantum theory. The work of Saunders and Muller rehabilitated Leibniz's principle, even in synchronic terms, but not in its strong form; instead, only as the weak PII. But these conclusions are not uncontroversial, as we shall finally discuss.

The idea of grounding identity or individuality with the help of Leibniz's principle via properties can be seen as reductionistic. Identity supervenes on properties and is thus a derived concept. Haecceitism, in contrast, is an anti-reductionistic or primitive form of identity grounding. In the case of weakly discernable objects, identity reduces to purely relational properties. Since, however, the concept of a "relation" also presupposes that of "*relata*", we are threatened here with a certain circularity. Katherine Hawley (2009) argues that weak discernability is grounded in the fact that the objects considered *are* already distinct, i. e. different, and that one therefore cannot conversely regard discernability as the basis of their distinctness. In brief: no relations without *relata*. It would seem that the debate between reductionism and anti-reductionism of identity has reached a stalemate. Dorato and Morganti (2013) even suggest a pluralistic strategy.

A different type of critique of the rehabilitation of the weak PII in quantum theory by Saunders and Muller was presented by Dieks and Versteegh (2008). These authors emphasize the manifest differences between the classical and the quantum world. They agree that Black's spheres are weakly discernable, but they reject a straightforward application of this scenario to the quantum case. In their view, a "many-particle state" in quantum theory must not be seen as consisting of actually individual objects—but rather as indicating potentially possible measurements. In the words of the authors:

> There is no sign within the standard interpretation of quantum mechanics that 'identical fermions' are things at all; there is no ground for the supposition that the quantum relations 'between fermions' connect any actual physical objects. The irreflexivity of these relations does not help us here. Quantum relations have a standard interpretation not in terms of what is actual, but rather via what could happen in case of a measurement (Dieks and Versteegh 2008, p. 934).

The supporters of the weak quantum PII could, as a countermove, point out that the question remains as to why it is quantum-theoretically possible to attribute a particle number to a many-particle state, if indeed the parts do not actually exist (see also Ladyman and Bigaj 2010 for a reaction to Dieks). On the other hand, it seems remarkable that the difference between classical and quantum physics should play no role in questions of individuation.

On at least four points, the debate over quantum ontology touches in a remarkable way with the debate on *structural realism*, a moderate variant of scientific realism with an overriding bearing on modern physics (cf. also Sect. 6.5.1 for the structural realist interpretation of QFT). Structural realists, especially ontic structural realists, consider the fundamental entities of the world as structurally individuated. Precisely how this should be understood is the subject of continuing discussions and can be only implicitly clarified in the following (cf. Lyre 2010 and French 2014 for comprehensive accounts). Steven French (1989) argues that the quantum theory is compatible with *both* the assumption that quantum objects are not individuals (in the sense of Leibniz, owing to the violation of the strong PII), *as well as* with the contrary assumption that they are individuals in the haecceistic sense. We are thus dealing with an underdetermination of metaphysics itself. According to French and Ladyman (2003), this metaphysical underdetermination indicates that an object-oriented ontology is doomed to failure and should be replaced by a structural metaphysics.

French and Krause (2006) go still further and try to develop a revised set theory for quasi-objects. This is the second point of contact between quantum ontology and structural realism. A third point of contact is connected with the observation that the permutation invariance of quantum theory has a corresponding invariance in general relativity in the form of diffeomorphism invariance; more precisely: According to John Stachel (2002), both invariances aim at the abstract property of certain theories of being *generally permutable*. A theory T is generally permutable if models of T can be regarded as equivalent which differ only in the question of which objects take on which positions or roles in a network of relations. In general relativity, diffeomorphism invariance refers to the general permutability of spacetime points; for quantum theory, Caulton and Butterfield (2012) argue that it is generally permutable if one

considers the full symmetric group, i. e. including the mixed-symmetric representations or parastatistics. Whether such objects in fact occur in nature is secondary; it is decisive only that the indistinguishability and symmetrization postulates of quantum theory allow for this possibility.

Fourthly, the weak PII seems to be perfectly compatible with structural realism: Objects or *relata* are individuated merely up to irreflexive relations. The above problem of relations not grounded in *relata* can be construed by structuralists such that either relations and their *relata* stand ontologically on a par, or else that the *relata* are merely nodes within a network of relations. Structuralism (both in physics and in mathematics) is in fact best *defined* as a position that grounds the individuality or cardinality (i.e. numerical distinctness) of objects in their positions or roles within a structure or network of relational properties. Thus, the projects of structural realism and the defence of the weak PII meet up precisely at the point where a grounding of relations is sought, which leads neither to a strong object conception as stipulated by the strong PII, nor to an haecceitism or even eliminativism (in the sense of relations without *relata*), but instead gets by with a thin conception of objects or *relata* in the sense of primitive numerical distinctness ("thin objects"; cf. French and Ladyman 2011).[21]

The debate over the ontology of quantum theory shows in an exemplary manner how progress in philosophy is frequently achieved: New insights are gained and the discussion is raised to a higher and more abstract level. But no less dodgy questions result, and the debates remain open. Just how open the debate remains in the case of quantum ontology has been demonstrated by this final outlook.

Exercises

1. Construct the fully symmetrized basis functions in the state space of three similar objects a, b and c (Hint: As presented in Sect. 3.1.4, these are the one-dimensional irreducible representations of S_3).
2. What is the content of the Leibniz PII in the contrapositive formulation?
3. To what extent is a bundle ontology naturally associated with PII?
4. What do the concepts of "synchronic" and "diachronic identity" refer to?
5. Define the three types of discernability according to Quine.
6. Discuss to what extent both the weak PII and structural realism lead to related object conceptions.

[21] According to Leitgeb and Ladyman (2008), the world could correspond to an edgeless graph, without this leading to haecceitism, but rather still in agreement with structuralism. In terms of graph theory, an edgeless graph (3.17), just like its counterpart with edges (3.16), obeys the same non-trivial automorphisms; both are structural invariants under node permutation. Nevertheless, the distinctness of the nodes in (3.17) is not grounded in any sort of relations, not even by weakly discernible, irreflexive relations.

References

Adams, Robert M. 1979. Primitive thisness and primitive identity. *Journal of Philosophy* 76: 5–26.

Black, Max. 1952. The identity of indiscernibles. *Mind* 61: 153–64.

Born, Max. 1927. Quantenmechanik und Statistik. *Naturwissenschaften* 15: 238–242.

Brown, Harvey, Erik Sjöqvist, and Guido Bacciagaluppi. 1999. Remarks on identical particles in de Broglie-Bohm theory. *Physics Letters A* 251: 229–235.

Butterfield, Jeremy. 1993. Interpretation and identity in quantum theory. *Studies in History and Philosophy of Science* 24: 443–476.

Cassirer, Ernst. 1956. *Determinism and Indeterminism in Modern Physics*. New Haven: Yale University Press.

Castellani, Elena (ed.). 1998. *Interpreting Bodies: Classical and Quantum Objects in Modern Physics*. Princeton: Princeton University Press.

Castellani, Elena, and Peter Mittelstaedt. 2000. Leibniz's principle, physics, and the language of physics. *Foundations of Physics* 30 (10): 1587–1604.

Caulton, Adam, and Jeremy Butterfield. 2012. Symmetries and paraparticles as a motivation for structuralism. *British Journal for the Philosophy of Science* 63 (2): 233–285.

Cortes, Alberto. 1976. Leibniz's Principle of the identity of indiscernibles: A false principle. *Philosophy of Science* 43 (4): 491–505.

Dainton, Barry. 2001. *Time and Space*. Chesham: Acumen.

Darrigol, Oliver. 1991. Statistics and combinatorics in early quantum theory II: Early symptoms of indistinguishability and holism. *Historical Studies in the Physical Sciences* 21: 237–98.

Della Rocca, Michael. 2005. Two spheres, twenty spheres, and the identity of indiscernibles. *Pacific Philosophical Quarterly* 86: 480–492.

Dieks, Dennis, and Marijn A.M. Versteegh. 2008. Identical quantum particles and weak discernibility. *Foundations of Physics* 38: 923–934.

Dorato, Mauro, and Matteo Morganti. 2013. Grades of individuality. A pluralistic view of identity in quantum mechanics and in the sciences. *Philosophical Studies* 163 (3): 591–610.

Feynman, Richard P., Robert B. Leighton, and Matthew Sands. 1964. *The Feynman Lectures on Physics*, vol. 3. Reading: Addison-Wesley. http://www.feynmanlectures.caltech.edu/.

French, Steven. 2011. Identity and individuality in quantum theory. *The Stanford Encyclopedia of Philosophy* (Summer 2011 Edition). http://plato.stanford.edu.

French, Steven. 1975. Hacking away at the identity of indiscernibles: Possible worlds and Einstein's principle of equivalence. *Journal of Philosophy* 92 (9): 455–466.

French, Steven. 1989. Identity and individuality in classical and quantum physics. *Australasian Journal of Philosophy* 67: 432–446.

French, Steven. 2014. *The Structure of the World*. Oxford: Oxford University Press.

French, Steven, and Decio Krause (eds.). 2006. *Identity and Individuality in Modern Physics*. Oxford: Oxford University Press.

French, Steven, and James Ladyman. 2003. Remodelling structural realism: Quantum physics and the metaphysics of structure. *Synthese* 136 (1): 31–56.

French, Steven, and James Ladyman. 2011. In defence of ontic structural realism. In *Scientific Structuralism*, ed. P. Bokulich, and A. Bokulich. New York: Springer.

French, Steven, and Michael Redhead. 1988. Quantum physics and the identity of indiscernibles. *British Journal for the Philosophy of Science* 39: 233–246.

Friebe, Cord. 2014. Individuality, distinguishability, and (non-)entanglement: A defense of Leibniz's principle. *Studies in History and Philosophy of Modern Physics* 48: 89–98.

Ghirardi, GianCarlo, Luca Marinatto, and Tullio Weber. 2002. Entanglement and properties of composite quantum systems: A conceptual and mathematical analysis. *Journal of Statistical Physics* 108: 49–122.

Hacking, Ian. 1975. The identity of indiscernibles. *Journal of Philosophy* 72 (9): 249–256.

Hawley, Katherine. 2009. Identity and indiscernibility. *Mind* 118 (1): 101–119.

Hesse, Mary B. 1966. *Models and Analogies in Science*. Notre Dame: University of Notre Dame Press.
Huggett, Nick, and Josh Norton. 2014. Weak discernibility for quanta, the right way. *British Journal for the Philosophy of Science* 65 (1): 39–58.
Ladyman, James, and Tomasz Bigaj. 2010. The principle of the identity of indiscernibles and quantum mechanics. *Philosophy of Science* 77 (1): 117–136.
Landau, L.D., and E.M. Lifschitz. 1965. *Quantum Mechanics. Non-relativistic Theory*, vol. 3, 2nd ed., Course of Theoretical Physics Oxford: Pergamon Press.
Leitgeb, Hannes, and James Ladyman. 2008. Criteria of identity and structuralist ontology. *Philosophia Mathematica (III)* 16: 388–396.
Lewis, David K. 1986. *On the Plurality of Worlds*. Oxford: Blackwell.
Loux, Michael J. 1998, ^{3}2006. *Metaphysics: A Contemporary Introduction*. London: Routledge.
Lyre, Holger. 2010. Humean perspectives on structural realism. In *The Present Situation in the Philosophy of Science*, ed. F. Stadler. Dordrecht: Springer.
Margenau, Henry. 1944. The exclusion principle and its philosophical importance. *Philosophy of Science* 11 (4): 187–208.
Messiah, Albert. 1979. *Quantenmechanik, Band 2*. Berlin: W. de Gruyter (French original: *Mécanique quantique*, Paris, 1959).
Messiah, Albert, and Oscar Greenberg. 1964. Symmetrization postulate and its experimental foundation. *Physical Review* 136 (1B): 248–267.
Meyenn, Karl von. 1987. Pauli's belief in exact symmetries. In *Symmetries in Physics (1600–1980)*, ed. M. Doncel, 329–358. Barcelona: Universitat Autònoma de Barcelona.
Muller, Fred A, and Simon Saunders. 2008. Discerning fermions. *British Journal for the Philosophy of Science* 59: 499–548.
Muller, Fred A, and Michiel P. Seevinck. 2009. Discerning elementary particles. *Philosophy of Science* 76: 179–200.
Post, Heinz. 1963. Individuality and physics. The. *Listener* 70: 534–537.
Quine, Willard van Orman. 1969. *Ontological Relativity and Other Essays*. New York: Columbia University Press.
Quine, Willard van Orman. 1976. Grades of discriminability. *Journal of Philosophy* 73 (5): 113–116.
Saunders, Simon. 2003. Physics and Leibniz's principles. In *Symmetries in Physics: Philosophical Reflections*, ed. K. Brading, and E. Castellani. Cambridge: Cambridge University Press.
Saunders, Simon. 2006. Are quantum particles objects? *Analysis* 66: 52–63.
Stachel, John. 2002. 'The relations between things' versus 'the things between relations': The deeper meaning of the hole argument. In *Reading Natural Philosophy: Essays in the History and Philosophy of Science and Mathematics*, ed. D.B. Malament, 231–266. La Salle, IL: Open Court.
Teller, Paul. 1983. Quantum physics, the identity of indiscernibles and some unanswered questions. *Philosophy of Science* 50 (2): 309–319.
Weyl, Hermann. 1950. *The Theory of Groups and Quantum Mechanics*. Dover edition, Courier Corporation.
Weyl, Hermann. 1949. *Philosophy of Mathematics and Natural Science*. Princeton: Princeton University Press.
Weyl, Hermann. 1952. *Symmetry*. Princeton: Princeton University Press.

Chapter 4
Entanglement and Non-locality: EPR, Bell and Their Consequences

Paul M. Näger and Manfred Stöckler

4.1 Introduction and Overview

The problems that we shall discuss in this chapter formally originate in the way in which composite systems are described by quantum theory (cf. Sect. 3.1.2). In such descriptions, there are, on the one hand, product states, e. g.

$$|\phi\rangle = |\uparrow_z\rangle_1 |\downarrow_z\rangle_2 \,. \tag{4.1}$$

This state describes a system composed of two objects, which are of different kinds. The indices 1 and 2 outside the brackets indicate the subsystem to which the state in the brackets corresponds. For example, $|\uparrow_z\rangle_1$ means that subsystem 1 is in the state spin up relative to the spatial direction z; $|\downarrow_z\rangle_1$ means correspondingly that it is in the state spin down along that direction. Analogously to classical physics and to our intuition, one can uniquely assign a state to each of the subsystems.

In quantum physics, however, one cannot generally characterize composite systems through product states alone, but only by *superpositions* of product states instead, i. e. by so-called "entangled states", such as

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}\Big(|\uparrow_z\rangle_1 |\downarrow_z\rangle_2 - |\downarrow_z\rangle_1 |\uparrow_z\rangle_2\Big) \,. \tag{4.2}$$

In Sect. 4.2.4, we analyse such state vectors in more detail, including their formal aspects, whereas in this introduction, we first give an intuitive overview. A state such as (4.2) describes an overall system which one can consider to be composed of two subsystems (as indicated by the indices). In contrast to a product state, it cannot, however, be brought into product form. This means that when a state $|\psi^-\rangle$ is present, one can neither uniquely assign a spin state to system 1 nor to system 2. That is, the individual subsystems 1 and 2 can neither be described correctly by the state spin up, nor by the state spin down, nor by a superposition state of these two. The composite state (4.2) does not specify the states of the subsystems.

C. Friebe et al., *The Philosophy of Quantum Physics*,
https://doi.org/10.1007/978-3-319-78356-7_4

If, however, one measures the quantity spin on the systems 1 and 2, one finds statistically distributed but definite outcomes in terms of spin up and spin down, and in particular, also *correlations* between these outcomes: If the measurement yields that system 1 is in the state $|\uparrow_z\rangle_1$, then the measurement on system 2 will result in state $|\downarrow_z\rangle_2$ *with certainty* (and *vice versa*); likewise, $|\downarrow_z\rangle_1$ and $|\uparrow_z\rangle_2$ are perfectly correlated. According to quantum mechanics, these correlations are still found when the measurement apparatuses are at great distance from each other; indeed, even then when the spatiotemporal order of events is such that not even a signal travelling at the speed of light could have produced the correlations. Entanglement establishes a special connection between the subsystems, which ignores all the usual spatiotemporal limitations.

Such entangled systems are the reason behind nearly all of the central problems of the interpretation of quantum theory. We have already met up with them in the discussion of the measurement process (Sect. 2.3.1) and of the quantum-mechanical description of indistinguishable particles (Chap. 3). Erwin Schrödinger called the possibility of entangled systems *the* characteristic feature of the quantum theory, which forces a decisive deviation from classical thinking in physics (Schrödinger 1935b, p. 555).

In this chapter, we concentrate on the correlations, and thus on a particular consequence of the existence of entangled systems, which was emphasized for the first time in a famous thought experiment by Albert Einstein, Boris Podolsky, and Nathan Rosen (1935). (The acronym of their last names, "EPR", became the eponym for the whole thematic area connected with this approach to entanglement.) The situation that EPR describe is often called the "EPR paradox", since the correlations due to an entangled state are classically quite unexpected. The article in which they describe the paradox gave rise to a conspicuously large number of publications.[1] Einstein and his co-authors in particular intended to show, in opposition to Niels Bohr, that quantum mechanics is still *incomplete*; i. e. it does not describe all properties of a system that are present in the physical world. In the ensuing discussions, the philosophical analyses—especially after a pioneering article by John Stewart Bell in 1964—concentrated on the specific *non-locality* which is characteristic of quantum mechanics. Entangled systems, which are discussed as a result of the EPR article, appear to imply that physical systems can still influence each other when they are separated in such a way that they could only be connected by signals faster than light. This, however, seems to contradict a fundamental principle of the theory of relativity.

We first cast a view onto the beginnings of the debate over entangled states (Sect. 4.2). In later sections, we will then describe the current debate on entanglement and non-locality, which is based for the most part on Bell's proof and the relevant experiments. Bell's argument shows that at least one of the ontological and methodological principles on which classical physics is based would have to be abandoned (Sect. 4.3). Indeed, quantum mechanics predicts the correlations arising due

[1] For the controversies of the first 50 years after the EPR paper, see Stöckler (1984); for contemporary discussions, see Fine (2013).

to entanglement, but their generation cannot be causally explained in the usual sense. We will investigate which of these classical principles on which such explanations are based can plausibly be considered to be violated. In particular, we will concern ourselves with the question of what it could mean that we abandon the central locality assumption, and whether or not non-locality is compatible with relativity (Sect. 4.4). Finally, we will discuss what consequences a violation of the other principles might entail (Sect. 4.5).

We remark for the reader who is pressed for time that there are several possibilities for taking shortcuts in this not-very-brief chapter. The EPR argument, seen historically, was at the origin of the debates over entangled states, but today, it is hardly viewed as a convincing argument. Whoever is less interested in the historical sources of the current systematic discussion can therefore skip over Sect. 4.2 (on the EPR argument). (For those who do address this section, we mention that Sect. 4.2.4 is mainly intended for friends of formal considerations, and can be initially skipped over by all others.) The systematic centre of this chapter consists of the Sects. 4.3 (on Bell's theorem) and 4.4 (on non-locality). (Readers without a formal mathematical background can leave out Bell's original proof in Sect. 4.3.2 and concentrate on the somewhat more intuitive description as a strategy game in Sect. 4.3.3; readers with mathematical skills may want to follow the inverse route.) The discussion of alternative solutions in Sect. 4.5 completes the systematic treatment, but is of mainly supplementary character.

4.2 The EPR Argument and Its Consequences

4.2.1 The EPR Argument: An Overview

The original formulation of the EPR argument was part of the debates between Bohr and Einstein on the status of quantum mechanics ("Bohr–Einstein debates"). Einstein, Podolsky and Rosen intended to show in their 1935 article that quantum mechanics is incomplete; i. e. there are properties within physical reality which have no corresponding quantities in the theory. EPR were less interested here in introducing completely new properties than in asserting that properties such as position and momentum, which due to Heisenberg's uncertainty relations cannot simultaneously have precise values according to quantum physics, are *in fact* simultaneously present with sharply defined values. Einstein, in particular, was not content with the new quantum mechanics, because it had broken with some of the central assumptions of classical physics. Only by assuming that the indeterminate properties (according to quantum mechanics) must in reality have clearly defined values, can one think of the apparently indeterministic quantum-mechanical processes as in fact deterministic (compare Bohm's theory; see Sect. 5.1); and the apparently non-local occurrences during collapse of the wavefunctions can be seen to be merely changes of knowledge (instead of real physical processes).

For the position that there are such additional properties or a more precise determination of the values of properties ("hidden variables"), which are not described by quantum mechanics, the concept of "realism" is often used in the context of the EPR debate (and in combination with the assumption of locality, it is called "local realism")—but this is misleading. Since "realism" can have a number of different meanings, we need to make clear here that EPR are referring to what philosophers call "metaphysical realism": Physical objects exist and their properties are clearly determined, independently of what human observers may know or think about them. While different variations of metaphysical realism are conceivable, the term "realism" in the EPR debate is to be understood as the position that—in contrast to what is suggested by quantum mechanics—properties which are found in a superposition are not indeterminate, but instead, through the hidden variables, they in fact always have certain, uniquely determined values. "Metaphysical", by the way, does not mean here that the properties are not accessible to physical measurements; rather, a typical assumption of realists in this context is that measurements reveal the previously determined values of the properties. Instead, "metaphysical" is used here as opposed to "epistemic": "Epistemic realism" denotes the claim that we conceive of the world (approximately) as it in fact is.

The assertion that our best scientific theories describe the world in an approximately correct manner is called "scientific realism" and requires both metaphysical realism and epistemic realism with respect to these best theories. In particular, scientific realism demands that the central concepts of the theory correspond to something within (mind-independent) reality. A realistic interpretation of quantum theory would then be an interpretation in which the wavefunction represents an existing physical object (cf. the GRW theory, Sect. 2.4), in contrast to anti-realistic or instrumentalistic interpretations, in which the wavefunction is merely a useful instrument for carrying out computations, but corresponds to nothing in physical reality (cf. epistemic interpretations and quantum-Bayesian theories, Sect. 7.3). In the former case, changes in the wavefunction describe physical processes; in the latter case, they represent merely changes in the state of our *knowledge*. EPR are thus aiming at a *metaphysical realism* (since they hold the existence and especially the properties of fundamental objects to be clearly defined facts at each moment, independently of our knowledge of them), but they also aim at an *epistemic/scientific anti-realism in reference to the quantum theory* (because they do not wish to understand the wavefunction as describing a real physical object).

EPR's argument is based on a thought experiment in which they consider a system composed of two entangled objects. In contrast to later presentations, they did not contemplate spin states, in particular not the state (4.2), but instead an entangled state in position space:

$$\psi(x_1, x_2) = \int \mathrm{e}^{i(x_1 - x_2 + x_0)p/\hbar}\, dp\,. \tag{4.3}$$

Such a situation occurs when two systems interact with each other over a certain time and are then separated and can (presumably) no longer affect each other. EPR

now ask the question of what will happen when one measures the position or the momentum of the particles when they are separated at a great distance. Making use of the theory describing composite systems which von Neumann had published for his analysis of the measurement process (cf. Sect. 2.3), Einstein, Podolsky and Rosen presume that as a result of the entanglement of the wavefunctions, a position or a momentum measurement on particle 1 allows the corresponding state of particle 2 to be determined. If one measures a certain momentum for particle 1, then one can immediately predict the momentum of particle 2 with certainty. If one finds particle 1 at a certain location, one can immediately predict the position of particle 2 with certainty.

Since, as EPR assume, the measurement on particle 1 cannot change the state of particle 2 (the measurements might take place at a large distance and their temporal succession could be arranged such that not even an influence travelling at the speed of light could connect the two), particle 2 must be ascribed with both a well-defined momentum and a well-defined position. This, however, is not represented in the quantum-mechanical description of the entangled state, from which it follows that quantum mechanics must be incomplete.

Einstein and his co-authors understand the measurement epistemically, i. e. as a change in our knowledge of the system:

> We see therefore that, as a consequence of two different measurements performed upon the first system, the second system may be left in states with two different wavefunctions. On the other hand, since at the time of measurement the two systems no longer interact, no real change can take place in the second system in consequence of anything that may be done to the first system.
>
> (Einstein, Podolsky and Rosen 1935, p. 779)

EPR thus presume that the two separated subsystems can have no further physical effects on each other (which can be seen as a sort of locality assumption). This is clearly stated in a letter from Einstein to Karl Popper, dated 11.9.1935:

> Now it is unreasonable to assume that the physical state of B may depend upon some measurement carried out upon a system A which by now is separated from B [so that it no longer interacts with B] and this means that two different ψ-functions belong to one and the same physical state of B. Since a complete description of a physical state must necessarily be an unambiguous description […], it is therefore not possible to regard the ψ-function as the complete description of the state of the system […].
> One can therefore hardly avoid the conclusion that the system B has indeed a definite momentum and a definite position co-ordinate. For if, upon freely choosing to do so […], I am able to predict something, then this something must exist in reality.
>
> (Einstein, letter to Popper 1935, reprinted in Popper, 2002, pp. 483–484)

This result makes the state description in quantum mechanics appear to be incomplete, since quantum mechanics can ascribe to a system only *either* a sharply defined momentum, *or* a sharply defined position. This incompleteness suggests that the correlations could be explained by introducing hidden variables; the authors, however, do not take an explicit position on this point:

> While we have thus shown that the wave function does not provide a complete description of the physical reality, we left open the question of whether or not such a description exists. We believe, however, that such a theory is possible.
>
> (Einstein, Podolsky and Rosen 1935, p. 780)

4.2.2 Analysis of the EPR Argument

The short EPR article (four pages in the journal *Physical Review* in 1935)—in particular its focus on an entangled wavefunction, the special proposed experimental setup, certain definitions as well as its line of reasoning—produced a multiplicity of reactions and discussions which are still continuing today. It belongs among the most influential works on the philosophy of the quantum theory. In the following, we consider the structure of this article in more detail, among other reasons because it makes a good case study of the mutual relations between the mathematical formalism, physical interpretations and philosophical principles.[2]

Viewed from today's vantage point, the EPR argument is flawed, owing to at least one false premise: the assumption of locality. The strategy of the argumentation is in addition not very transparent. The premises, which are of different types (definitions, elements of the mathematical formalism of quantum mechanics, and philosophical assumptions), are not always clearly stated. The formal structure of the argument is not very clear and appears to take unnecessary detours.

The conclusion of the argument is the assertion that quantum mechanics is incomplete. This thesis is defended in two steps which constitute the two parts of the article. The authors make it clear right from the beginning of the first part that only a complete physical theory can, in their opinion, be considered satisfactory. They stipulate that a complete theory must fulfil at least the following conditions:

> **(C) Condition of completeness**: "[E]very element of the physical reality must have a counterpart in the physical theory."
>
> (Einstein, Podolsky and Rosen 1935, p. 777)

In order to make this condition operational, there must be a criterion for when it is allowable to speak of an element of physical reality, which then must have a corresponding element in the physical theory. The authors formulate a sufficient condition for the case that such an element of reality exists:

> **(R) Criterion of physical reality**[3]: "If, without in any way disturbing a system, we can predict with certainty (i.e. with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity."
>
> (Einstein, Podolsky and Rosen 1935, p. 777)

---

[2] A valuable aid to this study is provided by the material and the careful analyses in Kiefer (2015).

[3] This somewhat misleading term is due to Bohr (1935).

Evidently, it is presumed here as a matter of course that there is an objective (mind-independent) reality ("metaphysical realism"), and that physical theories should represent this reality in the most complete manner possible. (C) and (R) are statements which refer to the relationship between theory and reality ("semantic assumptions"). They make sense only in reference to a background of certain philosophical assumptions, and they reflect Einstein's understanding of "realism". Additional explicit assumptions about "realism" are not given in the article.

By way of explanation, Einstein, Podolsky and Rosen consider quantum theory and remind us that in the case of physical quantities whose operators do not commute, like those of position and momentum, knowledge of the precise values of the one quantity (e. g. the exact position of a particle) excludes precise knowledge of the exact value of the other quantity. The usual opinion is that quantum theory is nevertheless complete, since a precise knowledge of the position of a particle makes it impossible to ascribe physical reality to any precise value of the momentum. This interpretation, the majority opinion which was established by Niels Bohr in his Solvay talk in 1927, is contradicted by the authors of the EPR article; they wish to show that a particle, at least under certain conditions, indeed can simultaneously have a well-defined position *and* a well-defined momentum.

The formal structure of the argument takes up this goal in a manner which is initially not very obvious. Einstein, Podolsky and Rosen conclude in an intermediate step that quantum mechanics and their semantic assumptions (C) and (R) entail for quantities with non-commuting operators the following statement composed of two partial statements (1) and (2):

> **(A1)**: "[E]ither (1) the description of physical reality given by the wavefunction in quantum mechanics is not complete; or (2) these two quantities cannot have simultaneous reality."
>
> (Einstein, Podolsky and Rosen 1935, p. 777)

If, indeed, both of the quantities were simultaneously elements of reality, then they would have to have corresponding elements within a complete theory. The derivation of (A1) presumes elements of quantum theory as well as semantic principles such as (C) and (R), but no assumptions on the locality or non-locality of quantum theory.

In the *second part* of the article, the entangled wavefunction (4.3) for two particles is used in order to demonstrate the incompleteness of quantum mechanics. The authors describe their argument by the following structure in terms of propositional logic (in the third-to-last section of the article, p. 780): The above statement (A1) (either (1) or (2)) is the first premise. The second premise is statement (A2), which is derived from a longer series of considerations and with additional assumptions from the entangled wavefunction ((1) and (2) stand for the same statements as in (A1)):

> **(A2)**: From the negation of (1), the negation of (2) follows.

(A2) thus states that it follows from the completeness of the quantum-mechanical description that two non-commuting physical quantities must be considered as simul-

taneous elements of reality. From (A1) and (A2), the conclusion (Cn) that statement (1) is true then follows in a logically correct way[4], i. e.:

(Cn): The quantum theory is not complete.

To be sure, it does not become really evident why the authors believe that they have derived the implication (A2). It is namely not clear at which point in the second part of the article the assumption of the completeness of theories plays a role (cf. Fine 2013, Sect. 1.2). In contrast, the assertion that by making use of (R), it was shown that (2) is false (compare the central EPR argument sketched in the overview Sect. 4.2.1), and that due to (A1), the statement (1) must be true (cf. Kiefer 2015, p. 40) would be more readily comprehensible—but EPR do not write this assertion. An explanation for the lack of transparency of the argumentation could be differences between Einstein's actual intentions and Podolsky's formulations, which the latter wrote down after many discussions among the authors of the article (for reasons of language, as often claimed; cf. Kiefer 2015, pp. 44–45, and Fine 2013, Sect. 1).

So much for the rough formal structure. Let us now look in more detail at the central EPR argument, already sketched in the previous Sect. 4.2.1, which is allegedly formalized by (A2). We have already shown that using the wave function (4.3) for a system of two particles, it can be demonstrated how it is possible that through measurements on particle 1, a certain value for a position measurement or a momentum measurement on particle 2 can be predicted.[5] Making use of quantum theory, including the assumption that a measurement causes a discontinuous change of the state vector ("collapse", also referred to as the "reduction of the wavepacket" by EPR, p. 779),[6] the conclusion is derived that after the measurement, a product state is present for the overall system and thus a well-defined state can be found for each subsystem. Due to the entanglement of the original state vector (4.3), it follows that depending on which measurement is carried out on the first particle, different state vectors can be associated with the second particle; possibly even state vectors which are eigenstates of non-commuting operators (p. 779).

Up to this point, EPR have used only assumptions which are in agreement with quantum theory.[7] On p. 779, however, a further premise is introduced, almost incidentally: "[…] since at the time of measurement the two systems no longer interact, no real change can take place in the second system in consequence of anything that

[4]Premise A1: Either (1) is true or (2) is true. Premise A2, reformulated by contraposition, yields: (2) implies (1). The demonstration of the validity of this conclusion is a variant of the classical dilemma: Either (1) (and thus Cn) is true. Or (2) is true, but in that case, the premise A2 requires that (1) also be true.

[5]For details, cf. Kiefer (2015), pp. 37–39.

[6]This makes von Neumann's theory of measurements important for EPR (cf. Kiefer 2015, pp. 23 and 38).

[7]Perhaps EPR understand this application of the quantum-theoretical state description to a single system as a completeness assumption in the sense of the negation of (1). Einstein, for example, presented his own statistical ensemble interpretation at the Solvay Congress in 1927 in opposition to an interpretation (which is evidently attributable to Bohr) according to which the wavefunction is a "complete theory of individual processes" (see Howard 1990, p. 92).

may be done to the first system." This is a locality assumption; it has underlying physical and philosophical reasons, which we will discuss in more detail later in this chapter. We shall see that from a contemporary point of view the locality assumption is false and, hence, the EPR argument is deficient. This assumption is necessary for arguing that the measurement on the one object does not disturb the state of the other object. Only under this condition does EPR's conclusion follow logically.

In a realistic reading, however, the description of states in the quantum theory evidently contradicts this locality assumption, especially when one considers the collapse of an entangled state vector at measurement. On the one hand, the authors follow quantum theory when they compute the state vectors for the subsystems after the measurement; on the other hand, they argue that the real state of one of the subsystems (e. g. the second particle) is not affected by the measurement at the other. The most plausible reading seems to be that the authors presuppose that quantum theory makes empirically correct predictions, but they do not assume that the quantum-mechanical description of states should be understood realistically. Since one can prepare several systems in the same entangled state, and can carry out a variety of measurements on the first particle without changing the real state of the second particle under the given premises, one arrives—under the assumptions of the EPR article—at the conclusion that quantum theory describes the same real state (viz. that of the second particle) by different wavefunctions (p. 779).[8]

Several authors have come to the conclusion as a result of their detailed studies of the pre-history and the reception of the EPR argument that in the article of 1935, Einstein's intentions were not clearly stated, and that he later expressed those intentions in a more clear-cut manner.[9] Thus, Don Howard (1990) proposes the thesis that for Einstein, even before 1935, the non-locality of quantum mechanics was the main reason for his dissatisfaction with the theory. In later statements by Einstein, the condition of completeness and the reality criterion of the EPR article were not emphasized. Instead, Einstein stresses the central importance of the assumption that the two subsystems must have their own states (because they are at some distance from each other), and that the state of the one system cannot be changed by a measurement on the second system, if the two are at such a distance that physical interactions are no longer possible (at least not in a certain time interval). Einstein's principal point was that the completeness of the quantum-mechanical state description and the requirement of locality are not mutually compatible. Locality means for Einstein here in particular the assumption of the independence of spatially separated objects (especially also in view of the theory of relativity as a field theory). What we have briefly and in a preliminary way termed "locality" will be more precisely formulated in the following sections (especially in Sects. 4.3.1 and 4.4.5).

For Einstein, it was a question of a conflict between the fundamental assumption of locality, which for him was in no case to be doubted, and a particular interpretation of quantum theory (which in the course of the argumentation and in the literal text of

[8] EPR thus do not need to argue counterfactually. They explicitly deny that the location and momentum of the objects can be simultaneously predicted or measured (p. 780).

[9] See Fine 2013, Sect. 1.3; cf. also Held (2006), and Kiefer 2015, p. 50.

the EPR article does not really become clear). The following developments were to show that a more precise locality assumption is not compatible with the experimental predictions of quantum theory.

To conclude, let us return once more to the EPR article. In its fourth-to-last paragraph, a more stringent argumentation is foreshadowed. There, an example is used to show that there are elements of reality (namely the simultaneous presence of a well-defined position and a well-defined momentum for one of the particles of the system), which according to quantum theory should not appear simultaneously (as exemplified by Heisenberg's uncertainty relations). Their simultaneous presence is proven by the fact that the position and the momentum can be predicted with certainty without perturbing the system. From (C) and (R), the incompleteness of the quantum theory then follows. This manner of argumentation, however, is misleading because the non-commutativity of the quantities (and hence the violation of the uncertainty relations) is not so much the significant point. In the same manner, one could namely argue that one could associate different momentum values simultaneously to the second particle if one makes suitable measurements on the first particle. Making use of the structure of the EPR argument, one could then derive the absurd conclusion that the second system in fact has arbitrary values of momentum. This consequence shows that at least one premise is false, or that the set of premises is inconsistent. Then however, one can no longer derive evidence from the argument that quantum objects can have a sharply defined position and simultaneously a sharply defined momentum.

Rational reconstruction of the EPR argument

As discussed in the main body of the text, EPR's argument is complex, at some points unclear, and—even apart from the very likely false locality assumption—presumably not completely free of flaws. It does not become clear in detail just how their argument should be understood. To be sure, one can indeed make use of their assumptions and considerations to construct a clear-cut argument (cf. Redhead 1987, pp. 78–81):

(P1) A perturbation of an object A cannot influence another, distant object B faster than by speed of light. (locality)

(P2) If one can predict the value of a physical quantity of a system with certainty without influencing that system in any way, then there is an element within physical reality (a property) which corresponds to that quantity. (criterion of physical reality)

(P3) In a complete physical theory, every element within physical reality must have its corresponding element in the theory. (condition of completeness)

(P4) Quantum-mechanical formalism

(C1) A measurement on object A of a (perfectly) entangled state permits the outcome of measuring the same quantity on the distant object B to be predicted with certainty. (from P4)

(C2) A measurement on object A of a (perfectly) entangled state does not influence the distant object B (or at least, not before the measurement at B has been completed). (from P1)

(C3) The value of the quantity measured on B corresponds to an element of reality. (from P2, C1 and C2)

(C4) The element of reality at B must have already been present before the measurement on A. (from P1 and C3)

(C5) According to quantum theory, before the measurement on A, the property of B is not yet determined. (from P4)

(C6) Therefore, quantum mechanics is not complete. (from P3, C4 and C5)

This argument is valid (i. e. logically correct). It is in fact simpler than the argument which EPR had in mind, because it requires measurements of only a single quantity (on both A and B), instead of two quantities.

Since, due to Bell's argument (see Sect. 4.3), we can today assume with a great deal of certainty that the quantum world is non-local, at least one of the premises of the EPR argument, namely (P1), is not true. The argument does thus not succeed in providing a good reason for assuming its conclusion. Of course, the failure of the argument does not imply that quantum theory is complete. Whether this is the case is an open question (see the debate about interpretations of quantum theory in Chaps. 2 and 5).

4.2.3 The Debate over the EPR Article and Its Aftermath

EPR's article is a clear attack on the still-young quantum theory. Correspondingly, the founders of the theory felt obligated to react quickly: Bohr writes a response to the EPR argument in the same year (1935). His article takes up the question raised by EPR—under the same title ("Can the quantum-mechanical description of physical reality be considered complete?")—and attempts to defend the theory on the basis of his own special interpretation. He does not engage the mathematical part or the entanglement, but instead concentrates on a correct understanding of the quantum-mechanical state description. He disputes the assertion that two quantities such as position and momentum can both be considered real at the same time. Nevertheless, he holds the quantum-mechanical description not to be incomplete, since it does not arbitrarily dispense with further information, but rather additional knowledge

is excluded *in principle*. According to Bohr's interpretation of quantum theory, a unique description of quantum phenomena must categorically include a description of the measurement apparatus used ("contextuality"). Under these assumptions, the EPR argument in its original form is no longer conclusive. In this respect, Bohr's reply was successful.

Bohr's answer, however, also contains some difficulties in itself. First of all, a simplified EPR argument can get along with measurements of only *one* quantity (see the grey box "Rational reconstruction of the EPR argument" in Sect. 4.2.2), and Bohr's answer would not shed light on that argument. Secondly, Bohr asserts that the collapse is not a physical process ("no question of a mechanical disturbance", Bohr 1935, p. 700), which in fact shows that he had not understood the central point of the EPR argument. For if the collapse is not physical, then it follows *immediately* from the thought experiment of EPR that quantum mechanics is incomplete (which, however, Bohr intended to dispute). Thirdly, later discussions of the EPR correlations begin with measurements and probabilities of measurement outcomes. Bohr, however, has no plausible explanation for the strong correlations. His remark—that the possibilities for description within quantum mechanics are limited in comparison to classical mechanics because there are uncontrollable quantum-like interactions between an atomic system and a measurement apparatus—does not help him further in the case of the EPR situation, because for the realization of the measurement outcomes, physical interactions between the distant measurements are not possible (cf. Hooker 1972, pp. 194ff., pp. 222ff.).

Erwin Schrödinger (1935b) also very quickly took up a position on the EPR article and emphasized that according to the entangled state vector used there, the separated systems cannot be associated individually with a state vector (for Schrödinger: with a wavefunction). For that, the state of the overall system would have to be representable as a product of the states of the individual systems. Schrödinger speaks here for the first time of the "entanglement" (*Verschränkung*) of the systems, which can again be broken up by a measurement, accompanied then by the typical correlations. He does not follow Einstein's assumption of the incompleteness of quantum mechanics, but he finds that, evidently in contrast to Bohr, entanglement as a consequence of the quantum-mechanical description of many-body systems is unsatisfactory. Altogether, he treats the problems connected with entangled systems more from the point of view of a mathematician and an instrumentalist.

In another article in the same year (Schrödinger 1935a), he makes use of entangled states—this time in a very intuitively clear manner—to demonstrate that superpositions of microscopic objects can indeed have macroscopic consequences. To show this, he constructs the "burlesque case" of a radioactive substance which, after some time, is in a superposition of "no atom as yet decayed", $|\uparrow\rangle_1$, and "one atom decayed", $|\downarrow\rangle_1$. In his thought experiment, Schrödinger assumes that the decay of one atom—via a "diabolic device", i.e. via a Geiger counter which releases poison when it detects a radioactive decay—can lead to the death of a cat. The overall system is described by the state vector

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big(|\uparrow\rangle_1|\uparrow\rangle_2 + |\downarrow\rangle_1|\downarrow\rangle_2\Big)\,, \tag{4.4}$$

in which the cat is neither in the state "dead", $|\downarrow\rangle_2$, nor in the state "alive", $|\uparrow\rangle_2$. The cat, like all systems which are part of an entangled composite system, has no clearly defined state by itself as an individual. Only the state of the overall system consisting of the cat and the radioactive substance is well-defined. Schrödinger asserts that the maximal knowledge of the overall system, given by (4.4), does not include maximal knowledge of the subsystems. With the example of the cat, Schrödinger especially wants to show that this remains true even when one subsystem is macroscopic. In Schrödinger's thought experiment, the conflict with classical intuitions becomes especially clear, since one can apparently have a precise conception of the cat's state of health during the whole course of the experiment.

The EPR argument did not lead to abandoning the usual quantum-physical description. It was unclear just how a "complete" local theory with hidden variables might look, and in spite of the puzzling correlations, the success and explanatory power of quantum theory were too great to allow a thought experiment to shake confidence in it. Physicists simply continued to develop the existing quantum theory further. Only in the 1950s did new interest in theories with hidden variables arise (due to Bohm's theory, see Chap. 5), and in the ensuing discussions, the EPR argument took on a new significance.

Since the 1950s, the EPR argument has no longer been discussed using the wave-function of Einstein, Podolsky and Rosen, but instead with an entangled state which was suggested by David Bohm in his textbook on quantum theory (1951, p. 616) and with which we have already met up in Eq. (4.2). Measurements with a spatiotemporal setup like that of EPR, but using this simpler entangled quantum state, are called "EPR/B experiments". In the debates over the EPR argument, since Bohm's introduction of this mathematically much simpler entangled state (in comparison to EPR's original entangled state (4.3)), it is used exclusively for the discussions. According to Bohm, this state vector describes the spin state of a system consisting of two atoms (each with spin $\frac{1}{2}$), which are initially bound in a molecule, and after the breakup of the molecule, have taken on this special spin state. While such molecular breakup processes are relatively difficult to observe experimentally, it has been found that nature has provided numerous realizations of this mathematical structure. For example, entangled polarization states of photons are today relatively easily accessible to experiments and are therefore often used in contemporary experiments with entangled systems. We will analyse such states in more detail in Sect. 4.2.4.

The entangled state (4.2) was also the starting point of the Irish physicist John Stewart Bell, who gave the argument of Einstein, Podolsky and Rosen quite a new twist. In a pioneering article from 1964, he argues, against EPR, that the correlations from experiments with entangled states cannot be explained in terms of local processes even if one assumes the existence of hidden variables: The assumption of hidden variables is not sufficient to explain the strong correlations in a local manner. In Sect. 4.3, we will encounter more details of this work.

It follows from Bell's proof that every correct theory of the microscopic world must contain a certain kind of non-locality. Bell could show this without limiting himself to particular theories of the quantum world. His proof requires merely the correlations between measurement outcomes, as they are found in EPR/B experiments, as well as some general, plausible background assumptions and some probability theory.

If Bell's theorem indeed shows that the quantum world is non-local (and not that one of his background assumptions is false), then the locality premise in the EPR argument is flawed and the truth of EPR's conclusion cannot be maintained: Entangled states do not speak compellingly for the incompleteness of quantum mechanics (a realism with respect to fundamental properties), since the quantum world is non-local. The most that one can then make out of the EPR argument is to no longer presume locality and to reformulate the argument towards a disjunctive conclusion: Quantum mechanics is incomplete, or it is non-local. From Bell's argument, we know today with near certainty that non-locality is the correct alternative. Whether in addition quantum mechanics is complete (the topic of realism) is still an open question (compare the various interpretations of quantum mechanics).

Because it contains a false premise, one has to consider the EPR argument to be a failure in terms of its content. Conceptually, in contrast, considering the introduction of entangled states, the measurement setup, the clarity of its principles and its chain of thought, it has had enormous effects, as we have described here, and must be considered to be one of the milestones in the history of science. This shows how fertile false arguments can be when they are clearly thought out. Finally, it must be considered as an irony of history that Einstein initiated a debate through his considerations of entangled states—that had the goal of creating a local theory of the microscopic world, following his local theory of gravitation—which, via Bell's demonstration of non-locality, finally refuted just this intuition.

Correspondingly, in modern discussions, EPR experiments are no longer considered to provide evidence for the incompleteness of quantum mechanics, but rather as evidence that a large class of theories with hidden variables—namely the class of local theories—can be excluded. The focus of the discussions has shifted from the question as to whether or not the wavefunction contains all the features of reality, towards the significance of the specific non-locality which shows itself in the EPR correlations predicted by quantum mechanics. This non-locality has become an accepted property of quantum mechanics and requires further examination.

4.2.4 Analysis of the Singlet State

In this section, we want to formally investigate the paradigmatic entangled spin state (4.2), usually called "singlet state", in more detail and to work out its characteristics. Those readers who are not friends of formal considerations can skip over this section on a first reading.

The spin state of the composite system is represented by a state vector in the Hilbert space of the tensor product $H_1 \otimes H_2$, which is formed from the spin Hilbert spaces of the subsystems. $|\uparrow_z\rangle_1$ is the eigenvector of the operator $\hat{S}_{z1}$ with the eigenvalue $+\hbar/2$ in the spin space of particle 1, and $|\downarrow_z\rangle_1$ is the eigenvector with the eigenvalue $-\hbar/2$. $\hat{S}_{z1}$ is the operator for the observables "spin projection along the z-direction". Likewise, for the second particle, $|\uparrow_z\rangle_2$ is the eigenvector of the operator $\hat{S}_{z2}$ with the eigenvalue $+\hbar/2$, and $|\downarrow_z\rangle_2$ is the eigenvector with the eigenvalue $-\hbar/2$. The pairs of eigenvectors each form a complete set of basis vectors in the spin spaces of the individual particles.

The vectors $|\uparrow_z\rangle_1|\uparrow_z\rangle_2, |\uparrow_z\rangle_1|\downarrow_z\rangle_2, |\downarrow_z\rangle_1|\uparrow_z\rangle_2$ and $|\downarrow_z\rangle_1|\downarrow_z\rangle_2$ form a complete basis in the Hilbert space of the tensor product. Let us return to the state (4.2), the much-quoted singlet state. This state vector is a common eigenvector of the spin projection operators $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$ of the composite system.[10] It can also be written as

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big(|\uparrow_x\rangle_1|\downarrow_x\rangle_2 - |\downarrow_x\rangle_1|\uparrow_x\rangle_2\Big)\,. \tag{4.5}$$

One can say intuitively that the singlet state is rotationally symmetric around the axis along which the particles move apart.

The state (4.2) is at the same time an eigenvector of the total spin

$$\hat{S}^2 = (\hat{S}_{x1} + \hat{S}_{x2})^2 + (\hat{S}_{y1} + \hat{S}_{y2})^2 + (\hat{S}_{z1} + \hat{S}_{z2})^2\,. \tag{4.6}$$

$\hat{S}_z$ and $\hat{S}^2$ form a maximal system of commuting operators (cf. Sect. 1.2.3) for this state; i. e. one can attribute to the system in this state at most the value (the property) of the total spin (here 0) and at the same time the value (the property) of the spin projection along the z-direction (here also 0). The state (4.2) is in fact not an eigenstate of the operator $\hat{S}_{z1} \otimes I$, which is defined over the whole spin space by the property that $\hat{S}_z$, like $\hat{S}_{z1}$, acts on the first factor of the basis vectors in the product space and the second factor remains unchanged. One thus cannot attribute an eigenvalue of the operator $\hat{S}_{z1}$ to the individual particles, and therefore no well-defined value of the spin projection. Intuitively, this corresponds to the fact that the state vectors of the individual particles cannot be characterized as pure states. From this mathematical result, one can go in various directions. If one wants to maintain the concept that the two particles in the singlet state should each be characterized by its own state, then one can only resort to the statistical operator (density operator):

$$\hat{\rho} = \frac{1}{2}|\uparrow\rangle\langle\uparrow| + \frac{1}{2}|\downarrow\rangle\langle\downarrow|\,. \tag{4.7}$$

It shows that the subsystems instantiate a so-called "mixed state" (cf. Sects. 1.2.4 and 2.3.1). From the so-called "reduced states" for each single particle that follow from the density operator of the composite system (4.7), one cannot determine the correlations between the measurement outcomes; therefore, it is said that the composite

[10] Although these operators are non-commuting; cf. Sect. 1.2.3.

state is not determined by the states of the subsystems. The density operators for the two particles are in this case indeed the same, so that some authors have arrived at the conclusion that Leibniz's Principle is violated here (cf. Sect. 3.2.3). Staying closer to the implications of the mathematical formalism, it would appear that—in contrast to states which can be written as products—in an entangled state like (4.2) one should not speak at all of "states of the components" or of "properties of the components".

In Bohm's simplified version of the EPR thought experiment, it is assumed that a measurement of the spin projection of each particle along the direction a is performed at distant points in geometrical space. Quantum mechanics predicts that the outcomes of spin measurements by the measurement apparatuses 1 and 2, each along the a-direction, will exhibit perfect anti-correlations. If one measures spin up with device 1, then device 2 must measure spin down (and *vice versa*). After the measurement, the system either is in state $|\uparrow_a\rangle_1|\downarrow_a\rangle_2$, or in state $|\downarrow_a\rangle_1|\uparrow_a\rangle_2$.

These correlations can be found for the singlet state in measurements along arbitrary directions, owing to its rotational symmetry. The probability that device 1 will measure, e. g. spin up is $\frac{1}{2}$; the probability that device 2 will measure spin down is also $\frac{1}{2}$. The probability that device 1 will measure spin up *and* device 2 will measure spin down is also $\frac{1}{2}$, however. The conditional probability of measuring spin down with device 2 when spin up has been measured by device 1 is 1. These correlations were later confirmed by experiments.

In the states which are present after the measurement, one can again assign a particular spin state to the individual particles. The states after a measurement by apparatus 1 along the z-direction are eigenvectors of the operator $\hat{S}_{z1} \otimes I$ (with the eigenvalues $+\hbar/2$ or $-\hbar/2$), and at the same time eigenvectors of the operator $\hat{S}_z$ (with the eigenvalue 0). The states after the measurement are no longer entangled. There are several equivalent criteria for this (compare the careful investigation in Ghirardi *et al*. 2002, especially Sect. 4.1). For us, it is important that in non-entangled systems (typically represented as product states in the tensor-product space), the subsystems can be associated with pure states and they exhibit no non-classical correlations like those that we have met up with in the example of the singlet state. In non-entangled states, one can again speak of two particles with individual properties (the complications in the case of indistinguishable particles will be considered next). Sometimes one refers to non-entangled states as "separable".

The criterion that states are not entangled if and only if they can be written as product states can lead to difficulties in the case of indistinguishable particles (more accurately: particles of the same kind), since in such systems even the appropriate symmetry requirements necessarily lead to superpositions of product states. When we noted the possible states after the measurement, we tacitly assumed that measurement apparatus 1 measures the state which belongs to H_1, that is, it has the index 1. If, however, the two particles are indistinguishable, the state following the measurement must be symmetric with respect to exchange of the indices 1 and 2, and the state after symmetrization can no longer be represented as a product in the given basis. Indeed, one cannot even determine for the spin measurements with apparatus 1 whether the outcome spin up can be associated with particle 1 or with particle 2. To deal with this

problem, there are various technical expedients (see Ghirardi *et al*. 2002; Ladyman *et al*. 2013), which we cannot treat in detail here; above all since they yield no essential new insights for our fundamental considerations. One of the suggestions asserts that, considered intuitively, states which are superpositions of product states are nevertheless *not* entangled if they were produced via a symmetrization of the indices from a product state. Another method is described by Audretsch (2007, p. 158) for the entanglement of the polarization states of photons. Here, as basis states for the measurement outcomes, state vectors are defined from the beginning in which no particle indices occur; instead, there are only specifications as to whether the left-hand or the right-hand measurement apparatus was used for the measurement, and which polarization state was found there.

Our analysis of Bohm's variant of the EPR situation up to now was simplified in still another respect. We limited ourselves to spin space and neglected the propagation of the particles in geometric space. This simplification has a certain justification, since the correlations are determined by the state vector in spin space, and according to the quantum theory, they must also be measurable at arbitrarily large distances between the two measurement setups. On the other hand, a measurement outcome spin up is found at a certain point in geometric space, and spin down at another point, and this must be taken into account in the formulation of the state after the measurements. Furthermore, with indistinguishable particles, the position space parts of the state vectors have to be taken into account in the symmetrization following the measurements. Details on this topic can be found in the comprehensive and precise study by Ghirardi *et al*. 2002. In this case, also, one can gain a first impression of the fundamental philosophical consequences without delving all too deeply into the technical details.

The most important physical consequence of our considerations up to now is the existence of non-classical correlations between the measurement outcomes. In analogy to the original EPR argument, one can presume that through the measurement of the spin projection along the a-direction or along the b-direction on system 1, the spin state of system 2 can be 'steered into' one of the states $|\uparrow_a\rangle_2$ or $|\downarrow_a\rangle_2$, or else into one of the states $|\uparrow_b\rangle_2$ or $|\downarrow_b\rangle_2$, without there being a classical interaction between the systems. In the following sections, we will try to investigate whether and how this can be made understandable within a more general framework.

To conclude our considerations of entangled systems in the EPR context, we want to cast a brief glance at possible implications for the ontology of quantum theory. Given the entangled state before the measurement, it is not possible, as we have seen, to ascribe particular eigenvalues and eigenvectors, and thus particular properties, to the subsystems. That is only again possible with the states *after* the measurement. Thus, it is not actually clear in what sense one can even speak of subsystems. Possibly, there are reasons for the assumption of sub-objects which are due to the background assumptions, e. g. the idea that some kind of objects emerge from a source and propagate away from each other in opposite directions. There is

an idea that in (4.2), each subsystem has its state only in relation to the other system. If we look only at the type of state description, a more likely assumption is that in a state such as (4.2), no components or parts exist at all. These parts are formed only during a measurement, which separates the composite system described by a state in the tensor-product space, so that one obtains components which can be associated with state vectors such as $|\uparrow_a\rangle_1$ or $|\downarrow_b\rangle_2$ in the sub-spaces (cf. Friebe 2004, for a defence of this viewpoint). However, quantum mechanics gives no sort of indications of what happens in detail during a measurement and why, after the measurement, we can again speak of two individual systems. Quantum mechanics gives us only the states before and after the measurement, and the probabilities for the occurrence of particular measured values.

All of this makes it difficult to trace the ontological consequences of the EPR situation. In Sect. 4.3, we will first choose a route which concentrates exclusively on the measurement outcomes and their statistical relationships, and in Sect. 4.4 we will once again discuss the implications for a holistic world view.

4.3 Bell's Proof

In this section we shall take a closer look at Bell's famous theorem from a systematic perspective. We will attempt to comprehend his argumentation, to work out his implicit assumptions and premises, and to discuss the consequences of his theorem. Even 50 years after its formulation, Bell's theorem is still one of the most-discussed topics in the philosophy of quantum physics. In spite of its elegance and simplicity, its correct interpretation is still controversially debated today. In a very immediate manner, i.e. to a great extent independently of any particular interpretation, it addresses a central issue of the quantum world: the problem of embedding quantum objects in space and time.

The theorem shows that there is a contradiction between the phenomena of the quantum world and the assumptions of a local, classical worldview. "Local" means here that physical processes never propagate faster than light, as apparently implied by the theory of relativity. The contradiction reveals in the fact that from a local classical worldview, one can derive an upper limit for the strength of correlations between distant events ("Bell's inequality"), which, however, is violated by outcomes of experiments with entangled quantum objects ("violation of Bell's inequality"). At the time of the formulation of Bell's theorem (Bell 1964), it was at first not clear which side of this contradiction should be considered to be false, because the outcomes of the experiments in which the quantum phenomena that we have mentioned occur were accessible only through quantum-mechanical predictions (thought experiments based on quantum theory). In the ensuing decades, however, different variants of real experiments were carried out, and they showed that the predictions of quantum physics were indeed correct. This made clear that it was the local, classical worldview that contains at least one untenable assumption. Most authors thus argue that it is the locality assumption which is violated in the quantum world. Since locality is

apparently a fundament of our accepted relativistic theories of spacetime, Bell's theorem thus shakes the foundations of our understanding of space and time and the causal processes which occur within them.

We first turn to the experiments, which are the empirical basis of the theorem.

4.3.1 Experimental Foundations

The experiments which are at the basis of Bell's theorem are realizations of the thought experiment suggested by Einstein *et al.* (1935) and simplified by Bohm (1951); they are therefore termed "EPR/B experiments". In their most modern variants, these experiments are performed using photons (light quanta). Such an experiment is typically carried out as follows: A suitable source C is caused to emit a pair of photons whose polarization states are entangled. The entangled polarization state, which is structurally similar to the spin singlet state discussed in the previous section (for the relation between the two, see the grey box "Bell states"), might for example read:

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}\Big(|+_z\rangle_1|+_z\rangle_2 + |-_z\rangle_1|-_z\rangle_2\Big)\,. \tag{4.8}$$

Here, "$|+_z\rangle$" stands for a polarization along the z-direction, while "$|-_z\rangle$" represents a polarization perpendicular to z. Analogously to the entangled states that we have already encountered, the individual photons in this state have no unique polarization. This description of the quantum state, however, serves here only as an explanation of which kinds of states are required for the experiments; nothing in Bell's argument depends upon this theoretical description; it depends only on the setup and the outcomes of the experiments, which we now want to characterize more closely.

Bell states

The polarization state of two quantum objects can be written in general as

$$|\psi\rangle = c_1|+_z\rangle_1|+_z\rangle_2 + c_2|+_z\rangle_1|-_z\rangle_2 + c_3|-_z\rangle_1|+_z\rangle_2 + c_4|-_z\rangle_1|-_z\rangle_2\,, \tag{4.9}$$

where the c_i are complex coefficients with $\sum_i |c_i|^2 = 1$. In other words, the product states $|+_z\rangle_1|+_z\rangle_2$, $|+_z\rangle_1|-_z\rangle_2$, $|-_z\rangle_1|+_z\rangle_2$ and $|-_z\rangle_1|-_z\rangle_2$ form an orthonormal basis of the corresponding Hilbert space.

The four special cases

$$|\phi^{\pm}\rangle = \frac{1}{\sqrt{2}}\Big(|+_z\rangle_1|+_z\rangle_2 \pm |-_z\rangle_1|-_z\rangle_2\Big)\,, \tag{4.10}$$

$$|\psi^{\pm}\rangle = \frac{1}{\sqrt{2}}\Big(|+_z\rangle_1|-_z\rangle_2 \pm |-_z\rangle_1|+_z\rangle_2\Big) \tag{4.11}$$

are called "Bell states" and likewise form an orthonormal basis of the Hilbert space. They violate a Bell inequality (presuming suitable measurements) in the maximal possible way (and are called "maximally entangled"); therefore, in typical EPR/B experiments for testing Bell's theorem, such Bell states are prepared. (All other entangled states are termed "partially entangled" and can generate only weaker violations of a Bell inequality.)

Among the Bell states, only $|\phi^+\rangle$ and $|\psi^-\rangle$ are rotationally symmetric. This can also be seen from the probabilities for certain measurement outcomes α and β (given certain measurement settings a and b and the preparation of one of the Bell states), which for the two rotationally symmetric states mentioned above depend only on the *relative* angle:

$$P(\alpha\beta|ab\phi^{\pm}) = |\langle\alpha_a, \beta_b|\phi^{\pm}\rangle|^2 = \tfrac{1}{2}\cos^2(a \mp b)\,, \tag{4.12}$$

$$P(\alpha\beta|ab\psi^{\pm}) = |\langle\alpha_a, \beta_b|\psi^{\pm}\rangle|^2 = \tfrac{1}{2}\sin^2(a \pm b)\,. \tag{4.13}$$

("$|\alpha_a, \beta_b\rangle$" denote here the state in which photon 1 has the polarization α along the a-direction, and photon 2 has the polarization β along the b-direction.)

While the states $|\phi^+\rangle$ and $|\psi^+\rangle$ are symmetric under exchange of indices, the states $|\phi^-\rangle$ and $|\psi^-\rangle$ are antisymmetric. Photons are bosons (spin = 1) and must have a symmetric total wavefunction. Nevertheless, they can also be found in antisymmetric Bell polarization states, since combined with an antisymmetric position or momentum part of the wavefunction (not included in the above equations) these still yield a symmetric total wavefunction.

Analogous facts hold for the spin states $|\phi_s^{\pm}\rangle$ and $|\psi_s^{\pm}\rangle$ of quantum objects with the notation $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$ instead of $|+_z\rangle$ and $|-_z\rangle$, however, with one important difference. Since the angles in spin space and in geometric space are scaled differently, different angular dependencies are found for the probabilities:

$$P(\alpha\beta|ab\phi_s^{\pm}) = \tfrac{1}{2}\cos^2\left[\tfrac{1}{2}(a \mp b)\right], \tag{4.14}$$

$$P(\alpha\beta|ab\psi_s^{\pm}) = \tfrac{1}{2}\sin^2\left[\tfrac{1}{2}(a \pm b)\right]. \tag{4.15}$$

After emission from the source, the photons propagate in opposite directions towards two polarization-measurement devices (see Fig. 4.1). Each of the two devices A and B has a pointer that allows adjusting the direction a or b in which the polar-

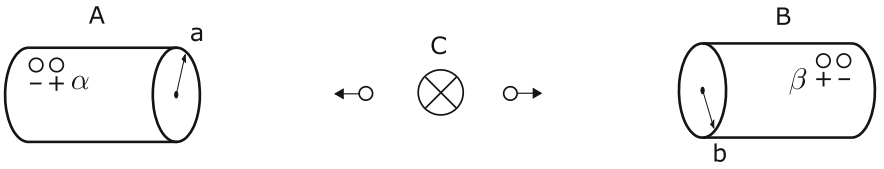


Fig. 4.1 Schematic setup of an EPR/B experiment

ization is to be measured.[11] This measurement setting is chosen arbitrarily for each measurement apparatus from one of three possible angles; for example, the pointer is adjusted to an angle of 0°, 30°, or 60° with equal probabilities. In the most stringent versions of the experiments, the choice and adjustment of the settings are carried out *while the photons are in flight* (i.e. after their emission at the source and before their arrival at the measurement devices). This is intended to guarantee that the choice of the angle can have no influence on the state of the photons emitted by the source (we will see below why this is important).

When a photon arrives at the measurement apparatus, the latter detects whether the photon is polarized along the chosen direction (+) or perpendicular to it (−). At each detector, there are thus two possible measurement outcomes, $\alpha = \pm$ or $\beta = \pm$, respectively. According to this, a complete cycle of the experiment is characterized by five variables: the preparation of the source (which determines the state of the photons emitted by the source), the two measurement settings, and the two measurement outcomes. A typical laboratory record for an experiment (with a fixed preparation procedure for the photons, which is therefore not noted here) looks like the one shown in Table 4.1.

These apparently innocuous columns of numbers in fact are the empirical basis for shattering the local classical worldview that we are accustomed to: All of the far-reaching consequences which can be drawn from the application of Bell's theorem—non-locality, non-separability, holism, etc.—are based on such simple data which can be collected from EPR/B experiments. An initial indication of how special these data are is obtained when one analyses them statistically. Three types of correlations are found:

1. (Nearly) Perfect correlations: When the angles of the two apparatuses are the same (cf. measurement cycles 1, 3, 8, 10, … in Table 4.1), the measurement outcomes agree in approximately, but not exactly 100% of the cases. It is, however, highly plausible to assume that these correlations are indeed perfect (as quantum theory predicts) because one can explain the measured deviation from perfectness by the fact that the detectors are not ideal measurement devices, but exhibit inefficiencies.[12]

[11] According to the usual convention, the angles are indicated relative to the z-axis. In principle, any other axis could be chosen as reference direction without the state $|\phi^+\rangle$ in (4.8) changing its basic form, since $|\phi^+\rangle$ is rotationally symmetric.

[12] Another consequence of the detector inefficiencies is the detection loophole (see the end of this section).

Table 4.1 A typical laboratory record with measured values from an EPR/B experiment

	a	α	b	β
1	30°	+	30°	+
2	0°	+	30°	+
3	60°	−	60°	−
4	60°	−	0°	+
5	60°	+	0°	+
6	30°	+	0°	−
7	60°	−	0°	+
8	0°	−	0°	−
9	30°	+	60°	+
10	30°	−	30°	−
…	…	…	…	…
1000	0°	−	30°	+

2. Imperfect correlations 1: When the angles of the two apparatuses differ by 30° (cf. measurement cycles 2, 6, 9, …, 1000 in Table 4.1), the measurement outcomes agree in 75% of the cases.
3. Imperfect correlations 2: When the angles of the two apparatuses differ by 60° (cf. measurement cycles 4, 5, 7, … in Table 4.1), the measurement outcomes agree in 25% of the cases.

These correlations of the measurement outcomes for each of the given measurement settings demand explanation. A first observation is that they are unexpectedly strong and depend only on the relative angle between the measurement settings. This is surprising, since the experiment is indeed set up in such a way that possible correlations are *minimized*: First of all, the measurement settings are chosen *randomly* and *independently* of each other. Secondly, the spatial arrangement and the temporal order of the experiments are chosen in such a way that many of the events in the experiments *cannot* influence each other by normal standards. The underlying idea is that, according to the theory of relativity, no influences between events are possible when their spatiotemporal arrangement is such that they could be connected only by processes propagating faster than the speed of light. In the given setup, there should thus be no possible influences between the state of the photons at the source and the measurement settings, since the latter are chosen only *after* the emission of the photons (but before the photons' arrival at the detectors). For the same reason, there should be no interactions between the measurement events (settings and outcomes) in one wing of the experiment (a, α) and those in the other wing (b, β).

Since these spatiotemporal considerations are of central importance for the interpretation of Bell's theorem, let us have a closer look at them. In the framework of the theory of relativity, such spatiotemporal relationships are illustrated by spacetime diagrams, in which one typically plots one of the three spatial dimensions on the horizontal axis and the time dimension on the vertical axis. For each spacetime point

P, the so-called light cone, the path taken by light beams towards P or away from P, divides spacetime into three regions (see the picture in Fig. 4.2a):

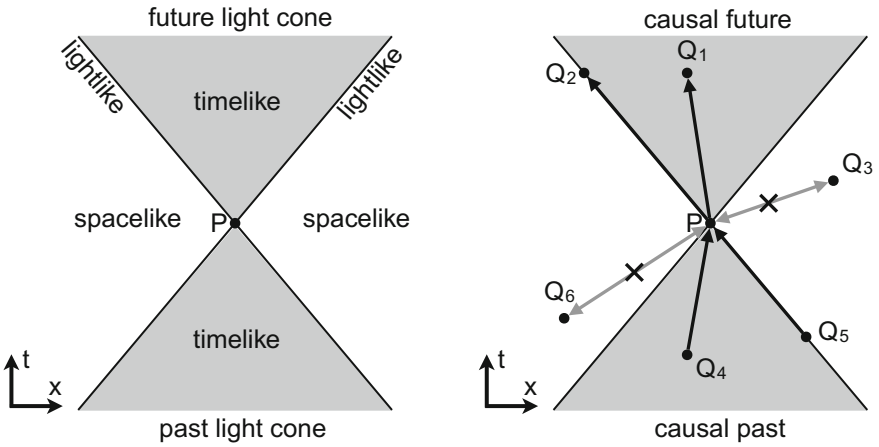
- Interior of the light cone ("timelike"): Points in this region can be connected with P by processes which propagate more slowly than light.
- Surface of the light cone ("lightlike"): Points in this region can be connected with P by processes which propagate at the velocity of light.
- Exterior of the light cone ("spacelike"): Points in this region can be connected with P only by processes which propagate faster than light.

One says that events which are located on or within the light cone of P are "local to P". Spacelike-separated events, in contrast, are said to be "non-local to P".

The crucial point for our considerations here is that, according to the standard interpretation of the theory of relativity the following principle holds:

> **Global Einstein locality**: There are no causal processes which propagate faster than light.

A causal process is a chain of causes and effects, i. e. a succession of events in which each event in the chain is an effect of the preceding event and a cause of the subsequent event (e. g. a sequence of falling domino tiles). Equivalently to the principle, one also says that according to the theory of relativity, only local processes can occur, or that the theory of relativity is a local theory. If this principle holds, it is impossible that spacelike-separated events can influence each other (see Fig. 4.2b). The region of spacetime with events that can influence events at P is then given by the one half of the light cone of P, the "past light cone" or the "causal past"; and the region with events that can be influenced by events at P is given by the other half, the "future light cone" or the "causal future". (Note that the light cone includes its edge.)



(a) Spatio-temporal relations (b) Possible causal relations

Fig. 4.2 Spacetime diagrams

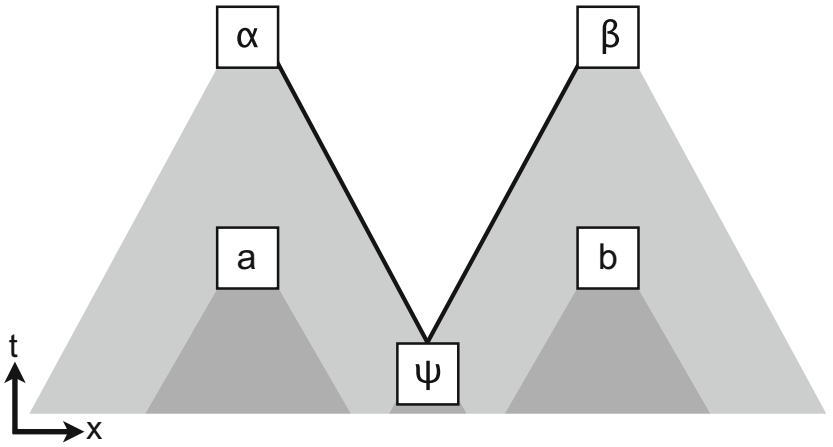


Fig. 4.3 Spacetime diagram of an EPR/B experiment

We now apply these considerations to the EPR/B experiment as described above. In Fig. 4.3, we have sketched the spacetime diagram of the experiment (cf. Bell 1975). The surface of the future light cone of the entangled state at the source ψ (heavy black lines) describes the path of the photons from the source to the detectors. This makes it clear that the state at the source ψ can influence only the measurement outcomes α and β, but not the measurement settings a and b, which lie outside of this light cone (and, *vice versa*, ψ cannot be influenced by a or b). The past light cones of the settings are depicted as dark grey areas, whereas the past light cones of the measurement outcomes are shown as light grey areas (including the dark grey regions). Thus, for example, the measurement outcome α can be influenced only by events within the light grey area (including the enclosed dark grey area) which has its apex at α, that is by the measurement settings in the same wing a and by the state of the photons at the source ψ. In particular, it should not be influenced by the setting b or the outcome β in the other wing! Correspondingly, according to the theory of relativity, no sort of influence should be possible from the one measurement outcome to the other, since such an influence would have to propagate faster than the speed of light.

The remaining possible influences are shown in Fig. 4.4. In this diagram, causal influences are indicated by arrows between variables. Variables represent properties (of objects or events). Such diagrams, which represent a particular causal structure, are called "causal graphs".[13] The causal structure shown here emerges from the fact that only influences between events that are locally related to each other are permitted. One can see especially that no direct connections between the two spacelike-separated measurements are allowed. The only connection between the measurement on the

[13]While for our considerations here, we initially require only an intuitive concept of causal influences; we will render this concept more precise in a formal way in Sect. 4.3.4 (see in particular the causal Markov condition there).

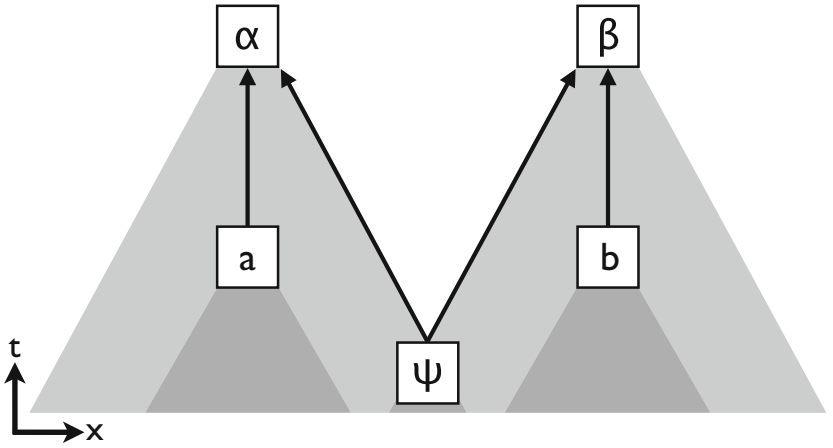


Fig. 4.4 Local causal relations in an EPR/B experiment

left and that on the right is the state of the photons emitted by the source, which is a common cause of the measurement outcomes. In this sense, we wrote above that the spatiotemporal arrangement of the experiment is intended to minimize correlations between the two measurements.

In spite of all of these limiting conditions, the strong EPR/B correlations are in fact observed, and we shall see in the next section that they are too strong to be explained in any usual manner. In contrast to the usually accepted constraints of the theory of relativity, it appears that there are influences between the two wings of the measurement setup which do not obey the relativistic limitation to less-than-or-equal-to the velocity of light.

Before we turn to possible explanations, we want to mention briefly that it was the research group of Alain Aspect who were the first to carry out these experiments in a convincing manner (Aspect, Dalibard and Roger 1982). Since the photons are very fast, it is technically quite challenging to adjust the measurement settings within the short time required by the photons to propagate from the source to the detectors, and Aspect and his co-workers were the first to find a solution to this problem. Weihs *et al*. 1998 then succeeded in addition in performing a choice of the settings by a random-number generator within this very short time interval. These sophisticated experimental setups, in which the measurement settings are chosen so late that they are spacelike with respect to both the more distant measurement as well as to the emission at the source, are intended to eliminate the possibility that the choice of a setting can have a local influence on the distant measurement (which closes the so-called "locality loophole"), as well as the possibility that a setting can influence the state of the photons at the source (or *vice versa*; closing the so-called "independence loophole" or "freedom-of-choice loophole"[14]).

[14]We do not accept the misleading interpretation suggested by the term "freedom-of-choice loophole", which is unfortunately widespread, that experimenters must *freely* choose the settings. Rather,

In the past 15 years, many new variations of this experiment have been performed. In particular, the group of Anton Zeilinger has made outstanding achievements in this field (Walther *et al*. 2006; Gröblacher *et al*. 2007; Paterek *et al*. 2007). For example, the measurements have been carried out at greater and greater distances between the detectors, and in the meantime, distances of over 100 km have become possible (Ursin *et al*. 2007). Thus far, there is no indication that the correlations weaken with increasing distance. Quantum theory also does not predict this, and correspondingly, one assumes that entangled systems can produce the correlations over arbitrarily great distances.

There is still another loophole in these experiments, which we have thus far not mentioned. It is due to the fact that the detectors, even when they are very well constructed, cannot register all of the quantum objects; i. e. every detector has a less-than-perfect detection efficiency. In real EPR/B experiments, this yields runs according to which only one of the detectors registers a photon, and such runs are as a rule eliminated from the statistical analysis of the measurement record. If the non-registration of a photon is a purely random event, this procedure will not falsify the overall statistics of the experiment. If, however, the non-registration of a photon depends in some way upon which measurement settings have been chosen, then it can be shown that even in a completely local world, a Bell inequality could be ostensibly violated, if in a suitable manner, i. e. only for certain combinations of settings the detector would fail to register (Clauser and Horne 1974; Fine 1982). This fact is referred to as the "detection loophole" or the "fair-sampling loophole".

In order to maintain a genuine violation of the Bell inequality, one thus requires an additional assumption, the "fair-sampling assumption", i. e. the measurement runs in which two photons are registered give a *representative* sample of *all* runs (including those cycles in which one or even both of the photons were not registered). The higher the efficiency of the detectors, the more difficult it becomes to simulate a violation, and there is a threshold value for the detector efficiency (81.8 %, see Maudlin 2011, Chap. 6) above which such a simulation becomes impossible. In order to avoid having to make the fair-sampling assumption, efforts have been made to increase the detector efficiencies to above this threshold. In experiments in which this has been achieved (the first were performed by Rowe *et al*. 2001), the Bell inequalities were still consistently violated. Thus, the detection loophole has been closed just as the locality loophole—however, in separate experiments. Recently, a research group has had the great success of being able to close *both* loopholes in a *single* experiment (Hensen 2015). The violation of the Bell inequalities by experimental data is a very well confirmed fact.

---

it is simply a question of making the settings *statistically independent* of the state of the photons at the source—freedom is not required.

4.3.2 Bell's Original Theorem

In his original article from Bell 1964, Bell took the experimental situation we have described above as the basis for his considerations. The important question for him was whether or not quantum mechanics can be converted into a local and deterministic theory by the introduction of hidden variables. The result of his considerations was that this is not possible. In the following, it will soon become clear that the reason for this is not his assumption of *determinism* (this assumption can be dropped in more stringent proofs), but rather the assumption of *locality*. Thus, no theory with hidden variables which reproduces the predictions of quantum mechanics can be local, in the sense that measurements on one system cannot be influenced by the measurement settings for another, distant system.

We will illustrate the fundamental ideas of Bell's proof here since they can demonstrate the physical core of Bell's strategy without going deeply into the mathematics. In the sections that follow, we will then treat Bell's chain of reasoning in a more systematic way, and, making use of statistical considerations, we will analyse it more precisely. Those who are not friends of mathematical formulae will find an approach in the following Sect. 4.3.3 which almost entirely dispenses with formalism, so that they can skip over the present Sect. 4.3.2 on a first reading.

Bell considers an experimental setup similar to that shown in Fig. 4.1, where, however, the source does not emit photons, but instead spin-$\frac{1}{2}$ particles, e. g. electrons, and the detectors are the well-known Stern–Gerlach apparatus from Chap. 1. Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be the measurement directions in the left-hand and the right-hand detector, respectively.[15] The outcome of a measurement $A(\boldsymbol{a})$ with the left-hand detector can be spin up along the $\boldsymbol{a}$-direction (with the measured value $+1$), or spin down along the $\boldsymbol{a}$-direction (with the measured value -1). The outcome of a measurement $B(\boldsymbol{b})$ with the right-hand detector can be spin up along the $\boldsymbol{b}$-direction (with the measured value $+1$), or spin down along the $\boldsymbol{b}$-direction (with the measured value -1).

Bell now assumes that the measurement outcomes at the left or the right detector are determined only by the respective measurement direction $\boldsymbol{a}$ or $\boldsymbol{b}$, as well as by a hidden variable λ; i. e. $A(\boldsymbol{a}, \lambda) = \pm 1$ or $B(\boldsymbol{b}, \lambda) = \pm 1$, respectively. The hidden variable λ is not specified explicitly by Bell; for his proof, he requires no concrete theory of a mechanism which would describe just how λ influences the outcomes of the measurements. The only assumption about λ is a distribution function $\rho(\lambda)$, which determines the probabilities with which the various values of λ occur, and which is normalized—as is usual for probabilities—to $\int d\lambda \rho(\lambda) = 1$.

The decisive point to note here is that Bell's precondition that every measurement outcome is a function of only the local measurement direction and the hidden variable implicitly contains a locality assumption, since in this way the measurement outcomes are functions of purely local quantities. (In particular, a measurement outcome depends neither on the outcome at the distant detector, nor on its chosen

[15] In this section, we print the variables $\boldsymbol{a}$ and $\boldsymbol{b}$ in bold font in order to denote them, as Bell did, to be vectors. In the other parts of this chapter, it suffices to regard a and b as scalar quantities which describe the angles of the measurement settings (wherefore they are printed non-bold there).

measurement direction.) This limitation to local dependencies is motivated by the relativistic considerations suggested in Sect. 4.3.1, that only events which are locally related can influence each other. This assumption implies that each individual particle carries locally all of the information necessary to determine the outcome of its measurement.

Based on these simple assumptions, Bell defines the central quantity for his proof, the expectation value of the product of the two measurement outcomes:

$$E(\boldsymbol{a}, \boldsymbol{b}) = \int d\lambda \rho(\lambda) A(\boldsymbol{a}, \lambda) B(\boldsymbol{b}, \lambda) \,. \tag{4.16}$$

It is essentially a sum over all the products, each weighted by $\rho(\lambda)$. From the expectation value $E(\boldsymbol{a}, \boldsymbol{b})$, one can read off whether and how the two measurements are correlated.

With these premises, Bell can now carry out his proof. If we, like Bell, distinguish three directions $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ along which the spin components can be measured in the two detectors, then the following inequality can be derived; it is the original Bell inequality:

$$|E(\boldsymbol{u}, \boldsymbol{v}) - E(\boldsymbol{u}, \boldsymbol{w})| \leq 1 + E(\boldsymbol{v}, \boldsymbol{w}) \,. \tag{4.17}$$

To derive this relation, Bell used only simple geometrical considerations and estimates of the measurement outcomes and their expectation values.

In the next step, Bell shows that the expectation values which follow from quantum mechanics for the typical EPR entangled state violate this inequality. For the expectation value of the product of two spin measurements along the $\boldsymbol{a}$-direction for the left detector and along the $\boldsymbol{b}$-direction for the right detector, one finds from quantum mechanics:

$$E_{\mathrm{QM}}(\boldsymbol{a}, \boldsymbol{b}) = -\boldsymbol{a} \cdot \boldsymbol{b} \,. \tag{4.18}$$

Here, $\boldsymbol{a} \cdot \boldsymbol{b}$ is the scalar product of the measurement directions, which is dependent on $\cos(\phi)$, i. e. on the angle ϕ between the two directions. If, for example, both detectors measure along the same direction, the expectation value $E_{\mathrm{QM}}(\boldsymbol{a}, \boldsymbol{a})$ is equal to -1; i. e. on one side, one finds spin down, and on the other side, spin up.

Making use of a special choice of the directions[16] $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$,

$$\boldsymbol{u} = (\boldsymbol{v} - \boldsymbol{w})/|\boldsymbol{v} - \boldsymbol{w}| \quad \text{and } \boldsymbol{v} \text{ perpendicular to } \boldsymbol{w} \text{ (i. e. } \boldsymbol{v} \cdot \boldsymbol{w} = 0 \text{)}, \tag{4.19}$$

one can arrive at the condition that the corresponding expectation values yield a contradiction when inserted into Bell's inequality, namely

$$\sqrt{2} \leq 1 \,. \tag{4.20}$$

[16] See Jammer's description of Bell's proof (1974, p. 307).

It follows from this that no theory which fulfils the locality condition for the expectation values in (4.16), and therefore fulfils Bell's inequality (4.17), can predict the expectation values (and thereby the correlations between the measurement outcomes) which follow from quantum mechanics. If now, as shown by the measurements, the expectation values of quantum mechanics prove to be empirically correct, then every theory which contains the assumption (4.16) must contradict empirical experience.

The presentation of the theorem in this section accommodates the expectations that are natural from a physical point of view. In the next sections, we will analyse the theorem once again on the basis of conditional probabilities for measurement outcomes and statistical considerations. In that treatment, Bell's theorem will be generalized and derived from weaker assumptions; in particular, the assumption that the hidden variables fix the outcome of the measurements in a deterministic manner can be dropped. One can then derive a Bell inequality for the case that the hidden variable λ, together with the measurement directions chosen for the detectors, determine only *probabilities* for the respective measurement outcomes.

4.3.3 Bell's Theorem as a Strategy Game

With the simple statistical relations for distant measurements which we illustrated in Sect. 4.3.1, one can show that the world in which these relations are measured must be non-local. One cannot see at first glance from the statistics how significant they are, and Bell's proof, which derives these far-reaching consequences from the data, is a paradigm of simplicity and elegance in the development of scientific argumentation.

The principal result of Bell's proof is that the experimentally measured statistics cannot be explained if there are only local influences (and if the usual background assumptions hold). In other words, if one assumes on the basis of the theory of relativity that global causal Einstein locality holds, i. e. influences which propagate faster than the velocity of light are excluded, then one obtains a contradiction with the empirical data. Purely local theories cannot reproduce the correlations. If we maintain the usual background assumptions (see Sect. 4.5), there must be non-local influences in some sense. In Fig. 4.4, we showed the maximal set of causal relations which can occur in the EPR/B situation under the assumption of global Einstein locality. It is the result of Bell's theorem that such structures cannot explain the correlations and therefore, they do not adequately represent the real situation.

A decisive characteristic of Bell's theorem is that one need not assume any special theory with particular state descriptions or a certain dynamics. In particular, the theorem does not refer to quantum mechanics, the quantum-mechanical wavefunction, or to any of its interpretations. It is based on the (for the most part) theory-independent measurement outcomes from EPR/B experiments, and its arguments are carried out on an abstract and general level. Therefore, its result is very strong: *All* those theories which are limited to local action cannot be correct. This result holds even when one allows that a photon at the source can carry an arbitrarily great amount

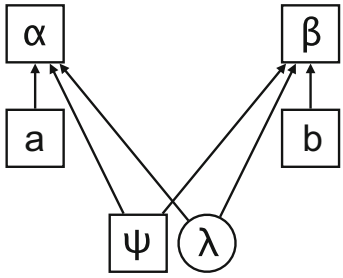


Fig. 4.5 Local causal structure with hidden variables λ

of information about the respective other photon. That assumption permits us to go beyond the quantum-mechanical state description and thus to specify the state of the photons more precisely by using so-called hidden variables. Critics of quantum mechanics have expressed their hope again and again that there might be such hidden variables which make the description of quantum objects in terms of their quantum-mechanical state more precise, and finally show that the quantum world is deterministic and local. Bell wanted to keep the door open for this possibility, and he thus assumed that the state of the photons at the source could be described not only by their quantum state, but also by an additional variable λ which is not empirically accessible. Like the quantum-mechanical state ψ at the source, it can play only the role of a common cause of the measurement outcomes in the causal structure since global Einstein locality holds. The extended causal diagram with the latent common cause λ is shown in Fig. 4.5 (we dispense here and in the following with drawing in the light cone). Even this stronger local structure (with *two* common causes) cannot explain the correlations, according to Bell's theorem.

Before we present an analysis of Bell's proof which is sound and clear by the standards of modern philosophy of science, we want in this section to first describe his argument in an intuitive form. Tim Maudlin (2011, Chap.1) has found a penetrating and illuminative analogy for Bell's argument. He compares the situation of the photons in EPR/B experiments, which, after leaving their common source, can no longer interact with each other (since they are spacelike-separated from that point onwards), with a situation in which two persons are initially in the same room and are then put separately into two different rooms. As long as they are together in one room, they can come to agreements and adopt conventions as much as they like. (This corresponds to the fact that the photons can possess arbitrary amounts of information about each other due to their common origin in the photon source.) After their separation in different rooms, the persons can no longer communicate with each other. (This reflects the experimental setup in which the photons move apart at the velocity of light and therefore can no longer mutually interact). In the separate rooms, each of the persons is then asked randomly one of three questions, which he or she must answer with "yes" or "no". (These questions correspond to the measurement settings of the setups where the photons are detected, and the answers correspond to the behaviour of the photons, which produces one of two possible measurement outcomes). For example, one of the persons is asked the question "30°?" and answers with "no", while the

Table 4.2 Possible strategies for perfect correlations

Strategy	Answer to "0°?"	Answer to "30°?"	Answer to "60°?"
1	Yes	Yes	Yes
2	Yes	Yes	No
3	Yes	No	Yes
4	No	Yes	yes
5	Yes	No	No
6	No	Yes	No
7	No	No	Yes
8	No	No	No

other person is asked the question "0°?" and answers with "yes". This procedure and questioning are repeated many times over with different pairs of persons. The results are shown in a table of a form which is analogous to the laboratory record of EPR/B experiments (cf. Table 4.1; the only formal difference is "yes" instead of "+" and "no" instead of "−"). The goal of the persons is to answer the questions in such a way that the results of the questioning follow the *same statistics* as the measurements of photons in an EPR/B experiment. That is, whenever the two persons are asked the same question, their answers must agree, and when the questions differ by 30°, the answers must agree in 75% of the cases; and with a difference of 60° in the questions, the answers must agree in 25% of the cases. Can they succeed in this task?

The difficulty consists in the fact that the persons who are answering know neither the question being posed to the other person, nor that person's answer. This means that if they are to have any chance at all of reproducing the statistics, they must agree upon a strategy before they leave the common room, regulating how they will answer the later questions when they are separated (for the photons this corresponds to hidden variables). Which strategy is the most promising? In order to fulfil the first statistical requirement, answering perfectly similar questions with perfectly similar answers *with certainty*, they must consent in advance what their answers will be to each of the three possible questions (deterministic strategy). For each pair of test persons, there are thus eight possible strategies, of which each determines a unique answer to each of the questions (Table 4.2).

If each pair of test persons chooses one of these strategies, then it is certain that they will produce perfect correlations when asked perfectly similar questions. Since the answers to perfectly similar questions vary statistically between different cycles of the experiment, it is furthermore clear that pairs of persons in different cycles must choose different strategies. This would have to be done in such a clever way that the two other correlations for differing questions would result. In what ratio would the persons have to mix their strategies in order to yield the known statistics?

We will now show that there is no such mixture that can guarantee the correct statistics. To that end, we consider the possible mixtures quite generally, i. e. without making any particular assumptions, and denote the fraction of cases in which the person chooses strategy 1 as f_1, the fraction of cases in which strategy 2 is chosen

as f_2, etc. From these weights, one can then derive the resulting statistics with the help of Table 4.2. For example, the probability that person A answers with "yes" and person B with "no" when A is asked the question "0°?" and B the question "60°?" is given by $f_2 + f_5$. To find this, we have simply added the weights of all the strategies which yield these answers to the corresponding questions. We record this fact in the usual notation as $P(\alpha = +, \beta = -|a = 0°, b = 60°) = f_2 + f_5$. The probability for the same answers to the questions "0°?" and "30°?" or to the questions "30°?" and "60°?" is found from Table 4.2 to be $P(\alpha = +, \beta = -|a = 0°, b = 30°) = f_3 + f_5$, or $P(\alpha = +, \beta = -|a = 30°, b = 60°) = f_2 + f_6$, respectively. Since the weights are all positive or 0, it is simple to see that these three probabilities must obey an inequality:

$$f_2 + f_5 \leq f_3 + f_5 + f_2 + f_6\,, \tag{4.21}$$

$$P(\alpha = +, \beta = -|a = 0°, b = 60°) \leq P(\alpha = +, \beta = -|a = 0°, b = 30°) + \\ + P(\alpha = +, \beta = -|a = 30°, b = 60°)\,. \tag{4.22}$$

This last inequality is one of the so-called Wigner–Bell inequalities, a type in the class of Bell inequalities. Now, the decisive point of the present argument is that this inequality, which followed from the assumption of the strategies and their weights, contradicts the measured statistics. The probability on the left-hand side has the value 37.5%, according to the measured statistics (it corresponds to half the value of 75% of measurement results which do not agree in the case of a difference angle of 60°; the other half are cases in which A answers with "no" and B with "yes"). The probabilities on the right-hand side are each 12.5% (each half of the non-agreement fraction of 25%). This yields $37.5\% \leq 12.5\% + 12.5\%$, and that is an obvious contradiction: *The empirical statistics violates Bell's inequality*.

Since we have made no special assumptions about the weights, this contradiction means that no distribution of weights, however constructed, would yield a strategy which could reproduce the measured statistics. The idea that persons could determine a strategy for answering which would produce the observed statistics has proved to be impossible. If the two persons do not yet know the questions when they agree upon a strategy, and can no longer communicate when they know the questions, they cannot give their answers in the peculiarly correlated manner that is observed in the EPR/B experiments.

This result can be applied nearly unchanged to the situation of the photons: If the photons do not yet 'know' the measurement settings when they are emitted from the source, and after leaving the source cannot 'communicate' any longer, then it is impossible that they can produce the given statistics. But we can observe empirically that they *do*! Therefore, at least one of the assumptions made must be false. It is most probable that they can indeed mutually 'communicate' after leaving the source although they are arranged in such a way that such an influence could only occur at a velocity faster than the speed of light. Influences which occur faster than light are

called "non-local" in the context of the theory of relativity, and they are forbidden according to the usual interpretation of that theory. To what extent this non-local connection between entangled objects is compatible with the theory of relativity, i. e. whether one can imagine that entangled objects are embedded in a relativistic spacetime, is the central problem of entangled systems. We will discuss it further in Sect. 4.4.

First, in the following Sect. 4.3, we want to bring the intuitive treatment of Bell's theorem as presented here into a clear form that meets usual standards in philosophy of science. This means, on the one hand, that we eliminate the anthropomorphic terminology of "strategies", "communication" and "knowledge", since all of these are descriptive terms which do not apply to photons. Instead we will introduce concepts such as "probabilistic dependencies", "causal influences" and "hidden common causes". On the other hand, we want to make the implicit assumptions in the situation as described more transparent in order to gain an overview of what precisely is at stake and which possible reactions to Bell's theorem are indeed appropriate.

4.3.4 Bell's Theorem, More Precisely

A precise treatment of Bell's theorem, which we want to develop in this section, does not correspond to Bell's original article (1964) in two respects: Firstly, it has become clear through the discussion over the course of years that the theorem can also be derived from a much weaker set of assumptions. Recognizing that some of the original assumptions can be dropped has limited possible reactions to the theorem and has rendered its significance more and more clearly. Clauser *et al*. (1969) derive a Bell inequality without assuming perfect correlations, Bell (1971) demonstrates a derivation without determinism. Here we base our discussion on the standard version of Bell (1975), which dispenses with both of these original assumptions. Further strengthenings of the theorem in recent times will be discussed below.

Secondly, we are giving here not a mere repetition of Bell's arguments, but rather a philosophical reconstruction. The latter differs from Bell's treatment in particular in our attempt to make all the implicit assumptions of content and method explicitly apparent. Only in this way can we guarantee a reliable interpretation of this theorem with its far-reaching consequences. Bell's proof is at its core a mathematical argument, which is formulated in its most stringent versions in the language of conditional probabilities. In its mathematical precision lies its strength, on the one hand (if its premises are correct, then its conclusion follows strictly); on the other hand, the mathematical expressions naturally require an interpretation if we wish to draw conclusions about our world—and this transition from formal expressions to physical or metaphysical facts has proved to be the central problematic point of the argument. While the formal facts are generally accepted, there is a not-insignificant amount of dissent around the question of the appropriate interpretation and evaluation of the premises and also of the conclusion. In the following treatment, we want to pay special attention to these interpretational transitions.

The task of the strategy game discussed above, to reproduce the measured photon statistics, has a deeper reason. This concerns one of the central principles of scientific inquiry, namely that correlations require explanation. If we find that two variables are mutually correlated, for example that a head cold is associated in, say, 70% of cases with a sore throat, then we attempt to explain this statistical connection. As a rule, we do that by trying to establish a causal relationship; that is, we assume that either one of the variables is the cause of the other, or that both have a common cause. The explanation for the correlation of head colds and sore throats is, of course, that cold viruses are their common cause.

In the case of the EPR/B experiments, we likewise have found correlations which must be explained (perfect correlations of measurement outcomes when measurement settings are parallel, and correspondingly weaker correlations when the settings are rotated relative to each other). We have seen above that the theory of relativity suggests, through its principle of global Einstein locality, that the correlations should be explained in a *local* manner. When one makes the typical background assumptions, then the strongest structure which can be assumed for the explanation of the correlations and which is consistent with global Einstein locality is the local common cause structure as shown in Fig. 4.5.

Among the background assumptions are first of all the preconditions that causal relations are always directed in the forward sense of time ("no backwards causation"), and secondly that the variables of the experiment which are controlled by the experimenter, namely the choice of the measurement settings and of the quantum state, are not effects of other variables within the setup ("intervention assumption"). We also mention that the hidden variable λ is to be understood as describing all of the possibly present hidden common causes of the measurement outcomes; that is, above and beyond λ, there are no additional hidden common causes (this is, to be sure, a part of the definition of λ, and not an assumption). Each of these two assumptions is very plausible and natural, but as we shall see, the discussion surrounding Bell's theorem is so tricky that even the most natural-appearing assumptions must be treated as potentially doubtful (cf. Sect. 4.5).

One can understand Bell's theorem as an argument which excludes the whole class of local causal explanations of the EPR/B correlations. It does so in an indirect manner: One assumes the strongest local structure, derives from it the statistical consequences (namely a Bell inequality which sets an upper limit for the strength of correlations resulting from such structures), and then shows that this inequality is contradicted by the measured statistics (the measured correlations are stronger than allowed by the Bell inequality). In order to derive statistical consequences from the local causal structure, one requires a translation principle between causal structures and statistical facts. The central translation principle between these two areas, which we will count as the third background assumption of the theorem (Spirtes *et al*. 1993; cf. also Pearl 2000), is called the

Causal Markov condition: Let X be a variable in a causal graph; Z denotes the direct causes of X, and Y includes all the variables which are not effects of X. Then X is statistically independent of Y given Z: $P(X|YZ) = P(X|Z)$.

This condition needs some explanation: Direct causes of a variable X in a given causal graph are all those variables from which an arrow leads to X (for example, in Fig. 4.5, a, ψ and λ are direct causes of α). (The indirect causes of X, in contrast, are those from which a path of *at least two* arrows leads to X.) The set of variables which are not effects of a variable X consists of all the causes of X (direct and indirect), and all those variables which are neither causes nor effects; i. e. variables which are causally connected to X only via common effects or are not connected to X at all. (For example, in Fig. 4.5, a, b and λ are not effects of ψ because they are only connected to ψ via common effects: a via α, b via β and λ via both outcomes.)

$P(X|YZ)$ is the conditional probability of X, given Y and Z; and $P(X|YZ) = P(X|Z)$ is the definition of the statistical independence of X and Y, given Z. The equation is meant to hold for *all* values of the variables X, Y and Z. Statistical independence of X and Y, given Z, means that if information about the value of Z is available, then knowing the value of Y yields no additional information about which value X may take (since information about Y does not change the probability of X). With these elaborations of the concepts, we can now formulate the statement of the Markov condition as follows: If one knows whether the direct causes Z of a variable X have occurred, then knowledge of its indirect causes (some part of Y) and of the variables which are connected with X only via common effects or not at all (the remaining part of Y) can give no additional information on whether or not X has occurred.

Evidently, the causal Markov condition is a translation principle, which sets causal factors in relation to statistical independencies. As such, it is the central bridge principle between causal and statistical facts. The causal structures are as a rule given in terms of causal graphs, in which the nodes are variables. Pairs of nodes can be connected by arrows which symbolize causal relations. Furthermore, such graphs must be acyclic, in the sense that arrows in them must not form circular structures. We have already used such graphs in an intuitive manner (see e. g. Fig. 4.5).

The causal Markov condition can be applied to arbitrarily complex causal graphs, but it is instructive to note that it implies statistical independencies in the following three paradigmatic fundamental structures: (i) Variables which are causally connected only through common effects, $A \rightarrow C \leftarrow B$, are independent: $P(A|B) = P(A)$; (ii) Variables within a causal chain, $A \rightarrow C \rightarrow B$, are independent given the middle variable, $P(A|BC) = P(A|C)$; (iii) Variables which are connected only through a common cause, $A \leftarrow C \rightarrow B$, are independent given the common cause C, $P(A|BC) = P(A|C)$.

The last case, (iii), makes it clear that the causal Markov condition contains Reichenbach's principle of the common cause (Reichenbach 1956) as a special case: If two statistically correlated variables X and Y are not directly causally connected, then there is a common cause Z which statistically screens off X and Y, which means that X and Y become statistically independent given Z: $P(X|YZ) = P(X|Z)$. While Reichenbach's principle allows the validation of their statistical independence only in simple cases with common causes, the causal Markov condition goes beyond the Reichenbach principle, in that it implies the independence for arbitrary structures (as long as they are acyclic).

Before we apply the causal Markov condition as a translation principle to the local causal structure, we should mention that the Markov condition is also a far-reaching methodological principle. It states that all correlations are caused; i. e. there can be no uncaused correlations. It is the probabilistic generalization of the classical principle of causality, which maintains that every event has a sufficient cause. While the classical formulation clearly fails when indeterminism holds or is useless when only statistical data are given, the Markov condition is well suited for such cases. Thus, the Markov condition is the methodological principle which forms the foundation of scientific research into causes, as well as of our considerations in this chapter. This methodological statement is not obvious in the above formulation of the Markov condition, but one can show that the condition is equivalent to it. In its two functions as a methodological requirement and as a bridging principle between causal and statistical facts, the Markov condition is the central principle of causal explanations.

The prominent role played by the Markov condition in Bell's proof can be seen from the fact that it is required in order to translate the requirement for a local common cause structure (Fig. 4.5) to conditional probabilities. This structure seems to be the natural explanation of the EPR/B correlations, if one maintains the worldview of classical physics, in which all processes are local and the normal methodological principles of scientific practice remain valid. The genial insight of Bell was then to see that such a structure cannot possibly explain the observed correlations. This can be clearly seen by deriving the following statistical facts from the graph, making use of the Markov condition[17]:

Local factorization (LF): $P(\alpha\beta|ab\psi\lambda) = P(\alpha|a\psi\lambda)P(\beta|b\psi\lambda)$

Measurement independence (MI): $P(\lambda|ab\psi) = P(\lambda)$

The factorization condition formalizes the fact that in the given structure, each of the measurement outcomes is influenced directly only by its local measurement setting, by the quantum state at the source, and by the hidden variables, but not by the other measurement outcome or the distant measurement setting. Measurement independence,[18] in contrast, results from the fact that there are no influences between the hidden variable λ on the one hand and the measurement settings as well as the quantum state ψ on the other.

These two probabilistic conditions form the mathematical basis of Bell's proof, from which he derived his inequalities. This derivation can dispense with the assump-

[17] Measurement independence follows directly from the Markov condition. For the local factorization condition, an additional intermediate step is necessary:

$$P(\alpha\beta|ab\psi\lambda) = P(\alpha|\beta ab\psi\lambda)P(\beta|ab\psi\lambda) = P(\alpha|a\psi\lambda)P(\beta|b\psi\lambda)\,. \tag{4.23}$$

The first step follows from the product rule of probability theory, and the second step follows from the Markov condition for the causal graph.

[18] Sometimes the condition is also called "autonomy", "no conspiracy" or "freedom-of-choice". Especially the latter two names already suggest a certain interpretation, which would require further assumptions; for the sake of generality, we would like to avoid this here.

tion of perfect correlations, which we had to use for the derivation of the Bell–Wigner inequality (4.22) above. This is an important point, since it makes the set of assumptions weaker. For the explicit derivation, the reader interested in formal details is referred to the grey box. The resulting inequality is noted there in (4.28). It belongs among the technical details that this Bell inequality takes on a somewhat different form from the Bell–Wigner inequality: The principal difference is that the former is formulated in terms of expectation values (instead of conditional probabilities).

Derivation of a Bell inequality

Perhaps the most elegant derivation of a Bell inequality from the assumptions of local factorization and measurement independence was given by Abner Shimony (1990). He begins with the following lemma: If x, y, x', y' are numbers within the interval $[-1, 1]$, then the inequality

$$-2 \leq xy + xy' + x'y - x'y' \leq 2 \tag{4.24}$$

holds. He then chooses $x = \sum_\alpha \alpha P(\alpha|a\psi\lambda)$, $y = \sum_\beta \beta P(\beta|b\psi\lambda)$, $x' = \sum_\alpha \alpha P(\alpha|a'\psi\lambda)$, and $y' = \sum_\beta \beta P(\beta|b'\psi\lambda)$, and applies the distributive law:

$$\begin{aligned} -2 \leq & \sum_{\alpha,\beta} \alpha\beta P(\alpha|a\psi\lambda) P(\beta|b\psi\lambda) + \sum_{\alpha,\beta} \alpha\beta P(\alpha|a\psi\lambda) P(\beta|b'\psi\lambda) + \\ & + \sum_{\alpha,\beta} \alpha\beta P(\alpha|a'\psi\lambda) P(\beta|b\psi\lambda) - \sum_{\alpha,\beta} \alpha\beta P(\alpha|a'\psi\lambda) P(\beta|b'\psi\lambda) \leq 2\,. \end{aligned} \tag{4.25}$$

Now, in each summand, we can use the local factorization condition:

$$\begin{aligned} -2 \leq & \sum_{\alpha,\beta} \alpha\beta P(\alpha\beta|ab\psi\lambda) + \sum_{\alpha,\beta} \alpha\beta P(\alpha\beta|ab'\psi\lambda) + \\ & + \sum_{\alpha,\beta} \alpha\beta P(\alpha\beta|a'b\psi\lambda) - \sum_{\alpha,\beta} \alpha\beta P(\alpha\beta|a'b'\psi\lambda) \leq 2\,. \end{aligned} \tag{4.26}$$

This is an inequality for the common expectation value on the *hidden* level. In order to obtain an empirically testable inequality, we must eliminate λ. We multiply the inequality by $P(\lambda)$, integrate over λ, and apply measurement independence in a suitable manner, so that λ is marginalized out. We demonstrate this for the first summand:

$$\int \sum_{\alpha,\beta} \alpha\beta P(\alpha\beta|ab\psi\lambda) P(\lambda) d\lambda = \sum_{\alpha,\beta} \alpha\beta \int P(\alpha\beta|ab\psi\lambda) P(\lambda|ab\psi) d\lambda =$$
$$= \sum_{\alpha,\beta} \alpha\beta \int P(\alpha\beta\lambda|ab\psi) d\lambda = \sum_{\alpha,\beta} \alpha\beta P(\alpha\beta|ab\psi) =: E(\alpha\beta|ab\psi) \,. \tag{4.27}$$

In the last step, we have introduced a notation for the common expectation value of the measurement outcomes, given the measurement settings and the quantum state. With this notation, the resulting inequality is given as:

$$-2 \leq E(\alpha\beta|ab\psi) + E(\alpha\beta|ab'\psi) + E(\alpha\beta|a'b\psi) - E(\alpha\beta|a'b'\psi) \leq 2 \,. \tag{4.28}$$

This is a Bell inequality. The present prominent form was first derived (but in a different way than presented here) by Clauser *et al.* (1969) and is called the "CHSH inequality".

The decisive point for the argument, however, is that this new inequality is also violated for some chosen measurement settings by the empirical statistics found in experiments. There is thus a contradiction between the empirical statistics and the theoretical assumptions which we made in a plausible way to explain them. Among those assumptions, we can distinguish two levels. First, at least one of the probabilistic assumptions, local factorization or measurement independence, must be false. Second, at least one of the assumptions which we used to derive the probabilistic assumptions, namely global Einstein locality or one of the background assumptions, must be false. Here is the structure of the argument once more as an overview:

(P1) Global Einstein locality (GEL) and a set of classical background assumptions (BA) imply measurement independence (MI) and local factorization (LF): (GEL) $\wedge$ (BA) $\rightarrow$ (MI) $\wedge$ (LF)

(P2) Measurement independence and local factorization imply Bell inequalities (BI): (MI) $\wedge$ (LF) $\rightarrow$ (BI)

(P3) Bell inequalities are violated: $\neg$(BI)

(C1) Measurement independence or local factorization fails: $\neg$(MI) $\vee$ $\neg$(LF) (from P2 and P3, *modus tollens*),

(C2) Global Einstein locality or at least one of the background assumptions are invalid: $\neg$(GEL) $\vee$ $\neg$(BA) (from P1 and C1, *modus tollens*).

Bell's argument thus has a negative result: It is a typical no-go theorem. It proves that the EPR/B correlations cannot be explained under the assumptions which we would normally make. In this sense, it shows that the quantum world is not compatible

with a classical worldview. At least one of the *prima facie* plausible assumptions which we have made must be false.

The explicit structure of the argument makes it once more clear that there is a mathematical core to the argument, namely the conclusion of the argument from (P2) and (P3) to (C1). (P1), which makes the conclusion (C2) possible, then takes on the status of an interpretive premise that confers a causal interpretation on the formal argument. In contrast to the impression which one could gain from articles on the subject in the physics literature, it is this interpretive framework which gives the formal argument its far-reaching significance. The mathematical core is for the most part undisputed, but the interpretive framework is the subject of controversial discussions: What exactly are the consequences of Bell's theorem?

Before we seriously discuss those consequences, we should first lay aside one 'consequence' which is often alleged but which is in fact due to a misunderstanding. Over and over, it has been and is still being claimed by some authors that Bell's theorem implies an *anti-realism*. For example, in 1979 in an article in *Scientific American*, one could read that the doctrine that the world consists of objects which are independent of human consciousness is not compatible with quantum mechanics nor with experimentally confirmed facts (d'Espagnat 1979, p. 128). Similarly as within the context of the EPR argument (see Sect. 4.2.1 for various meanings of "realism"), "anti-realism" is understood here to be the hypothesis that physical properties in the world are not independent of the knowledge or thinking of observers (i.e. experimenters), which we have called "metaphysical anti-realism with respect to fundamental properties". Only through measurements do the properties of quantum systems obtain clearly defined values.

From our explicit reconstruction of Bell's theorem which was presented above, it should be clear, however, that conversely to this widespread opinion, anti-realism is *not* a possible consequence of Bell's theorem. For the conclusion of Bell's theorem (in a strong version, which we have presented here) is either non-locality or violation of one of the background assumptions— i.e. anti-realism is not among the possible consequences. If one holds the background assumptions to be plausible, then Bell's theorem permits only the conclusion of non-locality—and of nothing else. Anti-realism does not occur at all as a possible conclusion—the theorem is just tacit about the matter. The reason for this is that realism (with respect to the fundamental properties of quantum systems) is not required as an assumption for the derivation of the Bell inequalities. The only realistic assumption which one requires is that we are convinced of the independent existence of the involved measurement apparatuses and their properties, since only then can the EPR correlations—and with them the violation of the Bell inequality (P3)—obtain the status of objective facts, but this assumption has thus far not been seriously called into question by anyone.

It is not quite clear how this far-reaching misunderstanding came about. A possible explanation is the following: An original motivation for opponents of the quantum theory was to manifest the possibility of local, realistic theories, in contrast to the quantum theory with its strange characteristics. The discussion around EPR and Bell was originally based upon the question as to whether local realistic theories are possible, but Bell's theorem negates this possibility. This might suggest that one of

its assumptions, either locality or realism, must be false.[19] This inference, however, forgets that the assumption of realism is not required for the derivation of the Bell inequalities. One can derive a Bell inequality using only the assumption of locality (plus background assumptions); i. e. the violation of the BI excludes *all* local theories, the realistic as well as the anti-realistic theories. Thus, only non-local theories remain possible—but Bell's theorem makes no statement as to whether the former or the latter are to be preferred. Bell's theorem is *neutral* in regard to the realism question.

4.4 Non-locality

4.4.1 Locality Versus Background Assumptions

Which of the assumptions for the derivation of the Bell inequalities is invalid? We distinguished two levels within the derivation, a probabilistic and a causal level, and on each of them, at least one assumption must be false. It seems to be nearly impossible to find an argument on the purely probabilistic level as to which assumption is plausibly violated. Practically, all the authors who discuss this problem refer at least implicitly to the causal level, even when they sometimes claim otherwise explicitly, since only there can one find criteria for rejecting an assumption.

This will also be the path that we follow here: In this section, we first evaluate the locality assumption, and in the following section the background assumptions, both on the more substantial causal level. For one thing, we have to investigate for every principle whether its infringement would in fact be able to explain the violation of Bell's inequality. The failure of at least one of the principles is indeed only a *necessary* condition for the violation of the inequality. For another, we must discuss for each of the principles for which this is the case whether it is plausible to dispense with that principle, and what consequences doing that would have. Here, we should not expect that we will arrive at a unique solution which is impervious to all doubts. Gerd Graßhoff, in a lecture, compared the procedure in the debates surrounding Bell's inequalities to a detective story: There are various persons who are under suspicion, and the task is to find out who is the murderer. Only in very rare cases can the detective *prove* in a strict sense who committed the crime, but often, we can be convinced by good evidence.

The great majority of authors interpret the violation of the Bell inequalities as a sign of non-locality: There must be a connection of some sort between the two wings of the experiment, which acts at a superluminal velocity. This hypothesis follows from Bell's argument, if one presupposes that those assumptions hold which we have termed "background assumptions". What are the reasons for abandoning the locality assumption and not one of the background assumptions? In the evolution of this position, it certainly plays a great role that two main positions advanced for

[19] This is also suggested by the article on Bell's theorem on the English Wikipedia pages, https://en.wikipedia.org/wiki/Bell's_theorem (accessed on 21st Oct 2017).

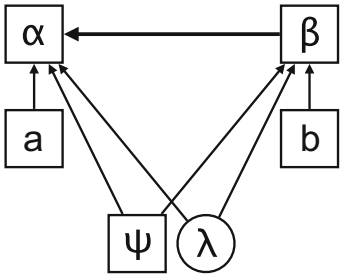
solving the quantum-mechanical measurement problem (see Sect. 2.3.1)— the GRW theory (see Sect. 2.4) and the de Broglie–Bohm theory (see Sect. 5.1)—are explicitly non-local. No matter how theory-independent Bell's project may be in terms of its assumptions, in the evaluation of the question as to which of the assumptions is false, most authors seem to take existing theories into account. This, however, on the other hand, also prevents the solution from being labelled as *ad hoc*.

In the following section, we will therefore discuss what it would mean if one of the background assumptions were to be abandoned, instead of the locality assumption. Such proposals rather play the role of alternative solutions within the discussion. The fact that they pertinaciously persist, however, shows that the suggestion of non-locality is itself not without problematic consequences. In particular, the question of compatibility with the theory of relativity, as we shall see, is still a serious topic. Nevertheless, the assumption of non-locality is the widely accepted consequence of Bell's theorem. The fact that it will turn out that all the alternative suggestions for a solution likewise exhibit great (if not even greater) problems is an additional reason to seek the solution in a non-locality.

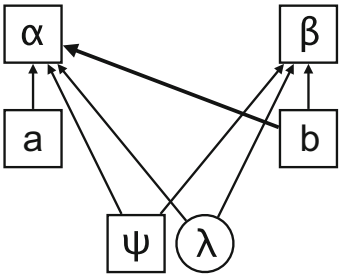
4.4.2 Areas of Conflict with Relativity

What precisely does it mean that the locality assumption is violated by the quantum world? As we have already mentioned, non-locality means that there is an influence between variables which are spacelike relative to one another so that this influence must have propagated faster than light (see Sect. 4.3.1). In EPR/B experiments, the detectors are arranged intentionally in such a way that each measurement is spacelike relative to the other measurement in order to exclude influences between the two measurements. Our argumentation, however, seems to have brought us to a point at which we have to abandon this assumption. It would appear that—in a way still to be determined—an influence between the two wings of the experiment *must* be present. For such an influence, there are three prototypes (see Fig. 4.6): Either there is (a) a direct influence from one measurement outcome on the other (which is labelled as "outcome dependence" in the debate); or there is (b) a direct influence between one of the measurement settings and the distant outcome ("parameter dependence"); or else there is (c) an indirect influence from one measurement setting on the distant outcome, mediated by the hidden common cause λ. (In the game of the persons this would mean that, when asked the question, one of the persons has information about (a) which answer the other person has given or (b) which question she has been asked, or that (c) one of the questions was known when the strategy was chosen.)

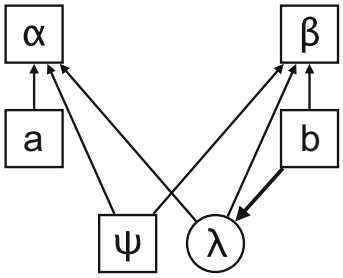
Concerning this last case, we should mention that λ must no longer describe a state at the source, owing to the abandonment of the locality assumption, and no longer necessarily refers to hidden variables of the photons. From an abstract causal point of view, λ was from the very beginning simply a hidden common cause of the measurement outcomes and is otherwise a sort of 'joker variable' in the diagram. In a local world, a common cause of the measurement outcomes must naturally lie within a section of the past light cone of the measurement outcomes, and the most plausible scenario is that λ is located at the source, and that it describes hidden variables of the



(a) Direct structure



(b) Non-local common cause



(c) Indirect structure

Fig. 4.6 Prototypes of non-local causal structures (important causal relations are shown here and in many of the following figures in boldface to make the image clearer)

photons. In a non-local world, in contrast, the states which are described by λ need not lie within the past light cone of the outcomes; instead, they can be found anywhere in the spacetime region between the measurement wings[20] (and thus could also be spacelike relative to the outcomes); and nevertheless, they could be the common cause of those outcomes.

The principal problem of the non-local model is that it seems to be in serious contradiction with the concepts of space and time which the theory of relativity suggests. We made use above of the standard interpretation of the theory of relativity, in which the principle of global Einstein locality holds (see Sect. 4.3.1). If, as claimed by this interpretation, it is correct that the principle of global Einstein locality is implied by the theory of relativity, then relativity is truly inconsistent with non-local models. Supporting a non-local model for EPR/B experiments would then mean that

[20]Considered precisely, λ must not necessarily lie between the measurement wings, even though this may seem most plausible; it must be located merely at some position outside the future light cones of α and β.

one would have to abandon the theory of relativity in a certain sense. Such a radical consequence, however, is avoided by most authors. Their idea is that a solution can be sought which on the one hand saves non-locality, in order to explain the EPR/B correlations, but on the other does not violate the theory of relativity.

Such suggestions have to assume that the theory of relativity does not strictly imply the principle of global Einstein locality. In a certain, limited sense, the theory of relativity must be compatible with the concept that there are influences that propagate faster than light. In the following, we now want to test whether non-local models can be brought into agreement with a relativistic concept of space and time. To this end, we have to look more closely at the question as to what extent the principle of global Einstein locality, which is violated by such models, is indeed founded on the theory of relativity.

First of all, we should say that in classical, pre-relativistic concepts of space and time (cf. Newton's view), it represents no difficulty if things influence one another at a velocity greater than the speed of light. As long as the velocity with which the influences propagate is finite, there can always be a continuous process from the cause to the effect. Influences between distant events which occur *simultaneously*, however, exclude such continuous processes; these are examples of classical action at-a-distance (such classical actions at-a-distance are also often called "non-local"; but here, we do not wish to 'water down' our concept of "non-locality"). Whether actions at-a-distance are admissible in classical physics has been under debate for a long time.

This debate came to an end as a result of the formulation of the theory of relativity. Firstly, in the theory of relativity, every action which propagates with a superluminal velocity is non-local; and secondly, all such actions would appear to be forbidden by the theory. The result is that the theory of relativity is a local theory through and through.

We should explain these two points in somewhat more detail. Why, according to relativity, are all superluminal influences considered to be non-local? According to our definition, such influences are connected in a spacelike manner. The decisive point is now that the temporal ordering of spacelike-separated events is no longer an objective fact within the theory of relativity. Simultaneity becomes a concept which is dependent on the frame of reference (or on the observer), and for every pair of spacelike-separated events A and B, there are reference frames in which A occurs before B, but also reference frames in which B occurs before A; and precisely one reference frame in which the two events are simultaneous. Therefore, for every action which proceeds at a velocity faster than the speed of light, there is precisely one system in which its effect is simultaneous. This explains also why such influences within the theory of relativity are termed "non-local".

Considering the second point, that the theory of relativity forbids all superluminal causal influences (global Einstein locality), we should look somewhat more carefully at the reasons why this is assumed to be the case:

1. According to the theory of relativity, the speed of light is an upper limiting velocity for the motion of matter and energy. It is excluded that matter or energy transport

can propagate faster than light. If, then, causal processes always involve matter or energy transport, there can be no causal action that propagates faster than the speed of light.

2. It is a consequence of the relativistic structure of spacetime that signals which are sent faster than the speed of light could be used to produce a signal loop, and such loops lead to paradoxes. If, therefore, a causal connection can be used for sending signals, then it cannot exist between spacelike-separated events.
3. According to the theory of relativity, for spacelike-separated events A and B, there are reference frames in which A occurs earlier than B, but also reference frames in which B occurs earlier than A. Therefore, if a causal connection must act forward in time in all reference frames, there can be no causal connections between spacelike-separated events.
4. The principle of relativity requires that all frames of reference be equivalent. Non-local connections, however, distinguish a particular frame of reference, in which the two events occur simultaneously. Therefore, there can be no non-local connections.

These are the reasons which speak against non-local causal connections in a relativistic spacetime. If we wish to explain the violation of the Bell inequalities in terms of such a non-locality, we must take a stand with respect to these arguments. Either we must show that they are invalid, or we must show that the suggested connection does not have the problematic properties.

It is clear that argument 1 is not conclusive: There could be causal connections which are simply not based on the transport of matter or energy. Even though the typical causal connections with which we are familiar are not of this type, many authors seem to admit this possibility. This argument, however, hardly plays a role in the debate, probably also because according to the theoretical descriptions which we have in the quantum theory and in the de Broglie–Bohm theory, the relation between the two wings of the measurements in an EPR/B experiment are not connected by any transport of matter or energy.

4.4.3 Signals, Causality and Fine-Tuning

Argument 2 is considerably more influential in the debate. It forbids non-local connections with which one could send signals. The belief that the theory of relativity forbids superluminal signals is common sense, but not all authors make it clear just why this should be so. The most strict reason, as emphasized by Arntzenius (1994), is that superluminal signals in a relativistic spacetime would permit us to construct self-contradictory signal loops—thus, such signals cannot be allowed. This can be seen as follows (Fig. 4.7): Two observers A and B (heavy grey bars) move apart, and each observer has an apparatus with which superluminal signals can be emitted in her rest frame (a or b, resp.), and also a detector with which such signals can be received (α or β, resp.). With the rules "A sends a signal if he receives a signal from

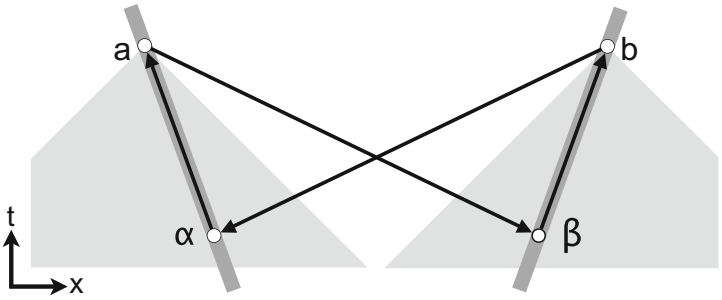


Fig. 4.7 Signal loops due to superluminal signals

B" and "B sends a signal if she receives no signal from A", we find a contradiction: Assuming that A received a signal from B, then A also sends a signal to B, and then B does not send a signal to A—contradicting our assumptions.

In order to avoid contradictions and paradoxes, superluminal signals must therefore be forbidden in relativistic spacetime. Since, according to certain theories of causality, causal connections imply the possibility of exchanging signals (e. g. Salmon's theory of mark transmission 1984), or can even be identified with that possibility, it can be argued that non-local connections in EPR/B experiments cannot be causal. By the way, superluminal signals (i. e. signals which propagate faster than light) are forbidden not only by the theory of relativity; it is also an empirical fact that in EPR/B experiments, no signals can be sent at speeds greater than that of light.[21]

But if it is not a causal connection, then what is it? The characterizations of many authors with respect to this question have to be understood as tentative attempts at expressing something unusual, which is unknown to our daily experience. Somehow, the connection is supposed to explain the observed correlations (that is, it should have the explanatory virtues of a causal relation), but without being usable for transmitting signals, and it should preferably also be symmetric. Shimony, for example, writes that the connection between the wings of the EPR/B experiment, with which one cannot send signals, is not to be considered to be an "action at-a-distance", but rather to be a "passion at-a-distance" (Shimony 1984, p. 224). Therefore, this type of non-locality in EPR/B experiments can "peacefully coexist" with the theory of relativity. Other authors support similar hypotheses, e. g. Jarrett (1984), or Redhead (1983, 1987).

[21] In order to be able to send superluminal signals, there would have to be a correlation between a controllable variable in the one wing and a detectable variable in the other. This, however, is not the case: On the one hand, the measurement outcomes are correlated, but neither of them is controllable (each outcome varies statistically from one measurement cycle to the next). On the other hand, the measurement settings can indeed be controlled, but there is no correlation between the choice of a measurement setting in one wing and any variable in the other. In particular, the choice is (marginally) independent of the distant measurement outcome. (Given a local measurement outcome, there is to be sure a correlation between the choice of setting and the distant outcome, but the local outcome cannot be controlled, so that one cannot make use of this conditional dependence to send signals.)

To be sure, it remains to a certain extent enigmatic just how such a special connection could explain the phenomena observed in EPR/B experiments. For one thing, one would require clear, generally valid principles for a convincing explanation; they should state under what circumstances such a connection implies which statistical facts, quite analogously to what the causal Markov condition accomplishes for causal connections. Since this connection would appear to be unique to entangled states, it is difficult to arrive at general principles which would not be immediately labelled as *ad hoc*.

Furthermore, also for a non-causal connection (just as for a causal one) the central problem remains that it would, on the one hand, have to be able to produce correlations in order to explain the strong observed EPR/B correlations (between the measurement outcomes), but on the other hand, it should *not* produce correlations (between a setting and its distant outcome), in order to prevent the superluminal transmission of signals. This dilemma thus occurs not only for causal connections but also for *every* type of connection which claims to be able to explain the statistical facts. Appealing to non-causal connections by no means eliminates this fundamental problem.

Since the statistical independencies that guarantee the impossibility of sending superluminal signals are an empirical fact, the latter problem persists even if one ignores the spatiotemporal limitations which follow from the theory of relativity. Independently of spatiotemporal relations, we are dealing here in the first instance with the question as to whether or not one can specify a structure from causal (or non-causal) connections which, together with non-arbitrary principles, explains the rather awkward observed statistics. This fundamental problem was only recently formulated (Näger 2016; see also Wood and Spekkens 2015) and is referred to as "the causal problem of entanglement" (in contrast to the classical spatiotemporal problem of entanglement, which concerns the question of how to understand the non-locality in accord with the theory of relativity).

The causal problem of entanglement can be very clearly stated by making use of the theory of causal graphs. This theory is based on three axioms: (i) Causal graphs are directed and acyclic (axiom of representation); (ii) causally unconnected variables[22] are statistically independent (causal Markov condition); and (iii) causally connected variables are statistically dependent (causal faithfulness condition). If these three axioms hold, one can show that a contradiction results for the statistics of EPR/B experiments: Either the strong EPR/B correlations cannot be explained, or else the statistical independencies which make superluminal signals impossible cannot be explained. At least one of the three assumptions thus cannot be correct. (We recall: A similar problem is found also for non-causal connections when one tries to formulate clear principles for the question under which circumstances such connections imply correlations or independencies).

[22]Two variables A and B in a causal diagram are connected if and only if A is the cause of B, or B is the cause of A, or if both have a common cause. In particular, A and B are not causally connected if they have only a common effect.

In this trilemma, it seems most plausible to abandon the causal faithfulness condition (Näger 2016), for possible violations of this condition are well known.[23] For example, there is the case that two causal paths exist from the cause to the effect, which, however, just mutually compensate each other (similarly to how the forces in a tug-of-war can exactly compensate each other). This requires a precise balance between the two paths, a fine-tuning of the so-called "causal parameters".[24]

In such cases, there are causal connections, but no statistical dependencies between cause and effect. Thus, manipulations of the cause do not modify the effect, so that one can send no signals, although a causal connection is present. Such violations of the faithfulness condition have precisely the properties which are required in order to explain the statistics of EPR/B experiments. Indeed, one can show that the quantum-mechanical formalism also contains a violation of the faithfulness condition.[25]

From this viewpoint, one need not introduce any new, mysterious non-causal relations in order to explain the EPR/B phenomena. The clear concepts and principles of an established causal theory are sufficient, if a certain fine-tuning of the causal paths occurs. The appeal to a fine-tuning makes this explanation of the EPR/B phenomena special, but not incomprehensible, even though an explanation of the fine-tuning itself is still not at hand. In summary, one can say that since causality does not necessarily imply the possibility of signal transmission, conflict point 2 does not prevent us from presuming that a causal connection exists between the two wings of the experiment.

4.4.4 Outcome Dependence Versus Parameter Dependence

The debate over a possible signal transmission and the question of the essential character of the non-local relations is in most cases combined with a related discussion in which one attempts to find out precisely which of the variables are non-locally connected with one another. This debate is based on the probabilistic core of Bell's argument, for which J. Jarrett (1984) introduced a both famous and misleading analysis. Jarrett reasons from Bell's theorem that on the probabilistic level, the local factorizability condition must be violated; and he shows that its violation is math-

[23] Violations of the causal Markov condition have also been claimed (see e. g. van Fraassen 1982; Cartwright 1988), but they are controversial (see e. g. Hausman and Woodward 1999). One can also show that they would not be sufficient to explain the strong EPR/B correlations (Näger 2013; see Sect. 4.5.1 below).

[24] In a typical causal system, every arrow in a causal diagram is associated with a causal parameter. It describes, roughly speaking, how strongly the causal variable influences the effect variable.

[25] Normally, such fine-tunings are unstable with respect to external perturbations, because a perturbation as a rule affects only one of the two paths and thus destroys the balance between them. Then, cause and effect become dependent and one could send signals. The quantum-mechanical formalism, however, demonstrates how it is possible that such a fine-tuning can be stable with respect to external perturbations: In quantum-mechanical non-faithfulness, the paths are so closely interwoven that external perturbations always act on *both* paths, and the laws for these perturbations guarantee that both paths are always perturbed in such a way that they remain in balance (Näger 2016).

ematically equivalent to the disjunction of the assumption of a correlation between the two measurement outcomes ("outcome dependence") and the assumption of a correlation between the measurement setting ("parameter") on the one side and the measurement outcome on the other ("parameter dependence"). These two correlations are then stylized as the two grand alternatives: "Outcome dependence or parameter dependence?" would then appear to be the decisive question, and a long debate has resulted, which has continued up to the present, discussing which of these dependencies is in fact relevant.

The definitions of outcome dependence and parameter dependence

Jarrett (1984) shows that the local factorization condition, which is a complex statistical-independence condition, is equivalent to the conjunction of the following pairwise statistical independencies:

Outcome independence:

$$P(\alpha|\beta ab\psi\lambda) = P(\alpha|ab\psi\lambda)$$

Parameter independence:

$$P(\alpha|ab\psi\lambda) = P(\alpha|a\psi\lambda) \quad and$$
$$P(\beta|ab\psi\lambda) = P(\beta|b\psi\lambda)$$

(The names of the mathematical conditions are due to Shimony (1984). Jarrett's original terms were "completeness" and "locality"; they were criticized early on and were never generally accepted). Since Bell's argument hinges on the *violation* of the local factorization condition, the negation of at least one of these independencies must hold (i.e. at least one set of values must exist for which at least one of the three equations is invalid). These are called correspondingly "outcome dependence" or "parameter dependence".

Almost unanimously, many authors plead in favour of outcome dependence and against parameter dependence; there are two principal arguments for this view. First, a violation of parameter independence would cause a conflict with the theory of relativity because parameter dependence would make possible in principle a signal transmission at superluminal velocity. Outcome dependence, in contrast, should be compatible with the theory of relativity, since it does not allow signal transmission at superluminal velocity. Second, the standard interpretation of quantum mechanics speaks for a violation of outcome independence, since it asserts that the photons can occupy a pure state only together, and only through the measurement does this state collapse (Butterfield 1989).

In this treatment, we can not penetrate more deeply into this very technical debate, but we should keep in mind the following points: First of all, it is not true that with

parameter dependence, one could in every case send signals. The correlation depends on the hidden variable λ, and if one cannot control it, as for example in the de Broglie–Bohm theory, then one cannot send signals with parameter dependence, not even in principle (Maudlin 2011, 88).

Secondly, Jones and Clifton (1993) argue that it is wrong to choose outcome dependence over parameter dependence on the basis of the criterion of possible signal transmission. If one leaves aside the difficulties with control of the hidden variable λ and takes the standpoint that parameter dependence would allow the transmission of signals, then outcome dependence would also allow signal transmission, if in addition one of the measurement outcomes depends on its local measurement setting (which is the case for partially entangled states[26]). Then, the measurement outcome, by itself uncontrollable, can be influenced by the choice of the local measurement setting (that is, its probability can be varied), and thus one could also send signals with outcome dependence. In other words: If one can send signals with the structure shown in Fig. 4.6b, then just as well with the one shown in Fig. 4.6a (cf. also Glymour 2006). In the latter case, the signal does not propagate directly from b to α, but instead indirectly via β.

Maudlin (2011, Chap. 6) and Näger (forthcoming a) criticize the still more fundamental point that the options upon which the debate is based, i. e. outcome dependence or parameter dependence, are misleading, since one cannot avoid a certain dependence on the distant measurement setting in any case. A dependence between the measurement outcomes alone is too weak to explain the violation of a Bell inequality. At least one of the measurement outcomes must depend also on the distant measurement setting. In particular, theories whose only non-local dependence is outcome dependence are excluded by this argument. With the causal Markov condition as bridging principle, these results can be transferred to the causal level (Näger 2013): Direct causal structures (as in Fig. 4.6a), which assume an influence from one measurement outcome to the other as the only connection between the wings of the experiment, are excluded. What was considered for many years to be the standard solution to the problem, a statistical dependence and the assumption of a quasi-causal connection between the measurement outcomes, has proven to be untenable. In order to violate a Bell inequality, at least one of the measurement outcomes must be an effect of *both* of the chosen measurement settings, i. e. at least one of the measurement settings must have an effect on the *distant* measurement outcome. This can happen either directly (as in Fig. 4.6b), or indirectly via the hidden variable λ (as in Fig. 4.6c), but not indirectly via the local measurement outcome (as in Fig. 4.6a).

Finally, from what we have said above, we can see that since quantum mechanics violates Bell inequalities, it cannot have the structure shown in Fig. 4.6a, as is often claimed. Indeed, an analysis of the quantum-mechanical formalism shows that it has a variant of the indirect structure as in Fig. 4.6c (for the exact structure, see Fig. 4.8c).

[26] While for experiments with the perfectly entangled Bell states (see the box in Sect. 4.3.1), the measurement outcomes are independent of each local measurement setting, this apparent independence, which results from symmetry, vanishes for partially entangled states.

4.4.5 Causal Non-locality Versus Non-separability

Up to now, we have argued that the violation of the Bell inequality requires non-locality in the general sense; i. e. there must be at least one *causal process* connecting the spacelike-separated wings of the experiment. We have represented causal processes in the diagrams by arrows, and in particular, non-local processes were indicated by arrows between variables which describe spacelike-separated states. We now wish to take a more detailed look at what these arrows represent and at what varieties there are in our understanding of non-locality.

Causal graphs are a simplified representation of actual processes. Every arrow stands for a causal process, but the graph tells us nothing more. In fact, a causal process from A to B is in general a series of states which are causally related to one another ($A \to C_1 \to C_2 \to \cdots \to C_n \to B$); and the graph represents the intermediate states C_i only implicitly through the arrow between A and B. For example, in a chain of 100 dominos, the toppling of the first domino (A) is the cause of the toppling of the last domino (B), and could be indicated in the diagram by $A \to B$. But of course the causal process between A and B could be more exactly described by the toppling of all the dominos in between. Leaving out the intermediate states in causal processes is permitted, because the conclusions which one can draw from the graphs regarding the statistics remain correct even when the intermediate steps are left out,[27] and this is important, because the complexity of real processes often makes it difficult to know the processes in detail or to represent them in complete form.

In the case of the non-local arrows in causal graphs, it is worth taking a closer look at how the underlying causal processes could be implemented. To this end, we have to take a step back and introduce some new concepts. We pointed out above that a local worldview requires that all causal processes take place more slowly—or just as fast—as the propagation of light (the principle of global Einstein locality). This principle is fundamental for a local worldview, as expressed in the theory of relativity and the relativistic field theories. Einstein (1948) shows that this principle is fulfilled if and only if the two conditions—which we will soon explain—of *(causal) locality* and *separability* are valid. The idea here is that *both* components of causal processes—states and the causal relations which connect them—must be local; only if this is the case is global Einstein locality fulfilled. Let us look more closely at this assertion. The following is meant by the condition of (causal) locality:

> **Causal Einstein locality**: There are no fundamental causal relations between spacelike-separated events.

The difference with respect to the *global* Einstein locality, which we defined earlier, is not easy to recognize. It is signalled by the word "fundamental": A and B are connected by a fundamental causal relation if there is no chain of other states which

[27] One must not, however, leave out any state which is a common cause ("causal sufficiency").

transmit the causal influence from A to B, i. e. if A acts *immediately* on B.[28] The principle of causal Einstein locality thus guarantees that the basic *dynamics* of physical systems is local—while global Einstein locality guarantees that physical processes are local *as a whole*. This terminology is somewhat unfortunate, in that the concept of "locality" stands both for the limitation of fundamental causal relations and also for the limitation of causal processes as a whole. Therefore, we qualify these different statements about locality throughout this text by adding the adjectives "causal" or "global". We note for the reader that this is not done in all texts relating to the debate: "locality" or "Einstein locality" can stand sometimes for the one condition, but sometimes for the other.

Einstein makes it clear that causal Einstein locality is not sufficient in order that a world could be considered to be local. One must also require that the *states* be *localized* in the following sense (cf. Howard 1989):

> **(Spatiotemporal) separability**: For every pair of non-overlapping spacetime regions A and B, it holds that:
>
> **1.** Each of the regions has its own separate state,
> and
> **2.** the joint state, i. e. the state of the overall region $A \cup B$, is determined by the state of A and the state of B, and the spatiotemporal relations between the regions.

Since this is supposed to hold for every pair of spacetime regions, this principle in the end means that each point in spacetime has its own state, with its intrinsic properties (in contrast to relational properties), and that the states of extended spacetime regions are determined by the states of their point-like constituents and the spatiotemporal relations between them. One can also say that the state of an extended spacetime region "supervenes" on those of the points and their relations (which means approximately that the latter states determine the former, but not *vice versa*).[29] This principle is realized in classical field theory, according to which, for example, the electric and the magnetic fields have well-defined values at every point in space, and the electromagnetic field is determined by the values at all the individual points. Einstein attributes a deep methodological significance to this principle:

> Without the assumption of such an independence of the existence (of the "being-thus") of entities spatially distant from each other, which initially originates in everyday life, physical thinking in the usual sense would not be possible. One cannot see how physical laws could be formulated and tested without such a clear-cut separation. (Einstein 1948, p. 321, transl. by the authors)

---

[28] "Fundamental" is not supposed to mean here that the causal relation could not be analysed in terms of non-causal concepts.

[29] The rather technical concept of supervenience originated in the philosophy of mind and in meta-ethics. Cleland (1984) introduced a variant of the original concept into the debate over space and time, and it is this variant which is used in the debate on entangled quantum systems (cf. for example, French 1989, or Esfeld 2004). The definition is: A dyadic relation R supervenes over a determinable, non-relational property P if and only if (i) each of the relata of R instantiates the property P in a determined manner, and (ii) the instantiations of the property P determine the relation R. A simple example: The more-massive-than relation supervenes over the masses of physical objects.

From the philosophical viewpoint, also, this principle is an attractive requirement, since its validity would imply that one can ontologically reduce all the physical states of extended regions to the intrinsic properties of their spacetime points. This would conform to the principle of ontological parsimony, which is generally considered to be an attractive characteristic. David Lewis introduced a similar principle under the label of "Humean supervenience" into the philosophical literature:[30]

> Humean supervenience is named in honour of the great denier of necessary connections. It is the doctrine that all there is to the world is a vast mosaic of local matter of particular facts, just one little thing and then another. […] We have geometry: a system of external relations of spatiotemporal distance between points. Maybe points of spacetime itself, maybe point-sized bits of matter or aether or fields, maybe both. And at those points we have local qualities: perfect natural intrinsic properties which need nothing bigger than a point at which to be instantiated. For short: we have an arrangement of qualities. And that is all. There is no difference without difference in the arrangement of qualities. All else supervenes on that.
>
> (Lewis 1986, pp. ix–x)

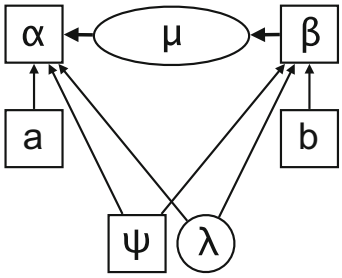
So much for the meaning of the separability principle.

If one has been convinced by Bell's argument that there must be non-locality at a fundamental level, it is then an interesting and important question as to whether or not this non-locality comes about through a fundamental non-local causal relation (violation of the principle of causal locality), or else through a non-local state (violation of the principle of separability). The former corresponds to what we have thus far tacitly presumed, that the states, i. e. the variables in the causal diagrams, are in fact localized; they thus occupy well-defined, limited spacetime regions. (Insofar as they are macroscopic variables which extend over a finite region of spacetime, one must presume that they supervene on the states of their sub-regions). In such a case, separability would be valid, and it would be some of the causal relations which would have a non-local extension. For example, the boldface arrows in Fig. 4.6 might be taken to represent causal relations as fundamental, non-local relations.

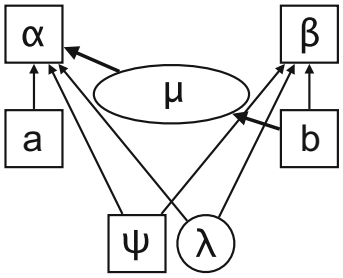
The second suggestion, in contrast, would lead to a case that we have so far not discussed, namely that all of the causal relations are purely local, but some of the variables in the causal structure are *not* localized. This would mean that there are states which extend over a large region of spacetime, but are not describable in terms of the states of their sub-regions. Such states are called "spatiotemporally non-separable".

The prototypes of non-local causal explanations which involve a non-separability are illustrated in Fig. 4.8. In Fig. 4.8a and b, we have introduced an intermediate

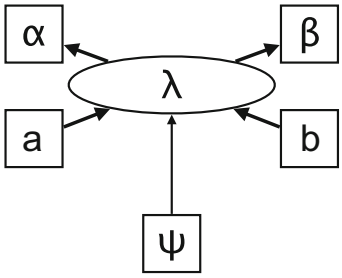
[30] Although they are similar, these two principles do not completely overlap. Humean supervenience is, on the one hand, stronger than the principle of non-separability, since it requires that *everything* supervenes on the states of spacetime points, and thus not only the states of extended spacetime regions, but also entities which must not necessarily be located in space and time, such as, for example, mental states or numbers. Humean supervenience is, on the other hand, also weaker than non-separability, since, in contrast to the latter, it does not require that the supervenience be *local*. The question as to whether or not for example an event in region A causes another event in region B is determined according to Lewis not by the state of the corresponding spacetime regions, but instead by the total state of all spacetimes in all possible worlds.



(a) Direct structure



(b) Non-local common cause



(c) Indirect structure

Fig. 4.8 Prototypes of non-separable causal structures

variable μ, which has a non-local extension. As we saw above, according to the formal rules of causal graphs, it is always possible to insert an additional variable into a causal path, since causal graphs do not need to list all the intermediate variables along a causal path (and as a rule do not do so). Nevertheless, from the physical point of view, the question remains as to which physical state is actually supposed to be represented by μ. Furthermore, the structure in Fig. 4.8a, as we explained in Sect. 4.4.4, is not possible because it is too weak to violate Bell inequalities. The most plausible among these structures would thus seem to be the indirect structure in Fig. 4.8c. The hidden common cause λ here no longer plays the role of a hidden variable at the source, but instead that of a non-local state which develops from the state ψ at the source (correspondingly, we have introduced some minor modifications to the prototypical indirect structure in Fig. 4.6c: ψ no longer influences the outcomes directly, but rather, mediated through λ, the two settings influence the outcomes only indirectly via λ).

Should we now understand the non-locality of the quantum world mainly as a result of non-local causal relations, or as a result of non-separability? In fact, most authors who participate in the debate presume that we are dealing with non-

separability, and many suggest that this can be derived from Bell's theorem. It is, however, difficult to justify this view in a convincing manner from the result of Bell's argument. Bell's argument is, as we have shown, an argument about which variables depend on which others; thus, finally, it is a *causal* argument. That there must be a dependence between spacelike-separated variables tells us nothing about whether this dependence comes about through a violation of locality or through a violation of separability. In Bell's argument itself, one can thus find no justification for either viewpoint.

Some authors have attempted to argue that it must be a case of non-separability, because it is an empirical fact that one cannot send signals via the connection. This, however, is a misunderstanding, for the question of whether or not one can send signals depends (as we have argued in Sect. 4.4.3) on whether the (causal) parameters are fine-tuned; that, however, has nothing to do with the present question of which type the non-local relations along these paths are. For example, with the non-separable causal structure in Fig. 4.8b, one could just as well send signals if there were no fine-tuning of the causal parameters as with the corresponding non-local structure in Fig. 4.6b. In particular, a consideration which often stands behind such arguments has proven to be dubious, namely that an explanation of the EPR/B correlations which involves a non-separability (and not on causal non-locality) would itself represent a non-causal explanation. The question of whether a causal non-locality or a non-separability occurs is relevant only to the spatiotemporal realization of the processes considered; whether such a process is causal in general is completely independent of this question. It is in fact the case that a non-separable state which is part of an otherwise causal process as a variable changes nothing in the overall causal nature of the process (it would, to be sure, be a causal process with unusual spatiotemporal properties).

On an abstract level, it appears that in the end, only one argument that favours non-separability remains, namely the conflict with the theory of relativity mentioned above under Point 3. It warns that a non-local causal connection would propagate forward in time in some reference frames, but backwards in others. This conflict with the asymmetry of causality seems to suggest that the non-local relation should be symmetric, and thus not causal. Non-separability, in contrast, can be a perfectly symmetric relation and thus avoids this point of conflict.

This plausible but not completely convincing argument obtains support from a different train of thought, which calls on the quantum theory. Here, we leave the level of Bell's theory-independent, abstract proof and refer explicitly to a specific theory. Indeed, it seems to be true that some of the authors who have maintained that they can derive non-separability from Bell's argument have had the quantum theory in mind all along and have let themselves be guided by it in their deliberations about the abstract topics of Bell's argument.

According to the quantum theory, the entangled quantum state in Eq. (4.8) extends between the two wings of the experiment (and on out to infinity). This can be seen from the fact that a measurement on either of the two wings can change the state (it then collapses to one of the terms in the superposition); i. e. the quantum state as a whole can be locally influenced from any point within its extension. If we denote

the spacetime region around the left-hand wing as "A", and the region around the right-hand wing as "B", then the quantum theory clearly violates the first condition of the principle of separability: It is not the case that A and B have their own separate states, because there is a state, namely that of the entangled pair of photons, which extends over both regions. (A violation of the first condition, incidentally, also implies a violation of the second condition for logical reasons: If the regions do not have separated, independent states, one cannot say that their overall state supervenes over the separate states.) Stated briefly: Non-local states violate separability; thus, the non-local quantum state represents a non-separability. One can show that the quantum-mechanical formalism has a causal structure like that shown in Fig. 4.8c. All the causal relations are local, and it is the *non-local quantum state* which mediates the action from one wing to the other.

It has sometimes been claimed that an explanation of the EPR/B correlations that makes use of such a non-separability would mean dispensing with a causal explanation. Our considerations, however, show that this is not the case: The non-separable variable is embedded within a causal structure in a clear manner, like every other variable. That is, on the level of the abstract causal structure, there is no difference to usual variables. Its special feature is due to its peculiar and unusual spatiotemporal embedding—this, however, is not a causal property. In this respect, non-separability does not represent a break with causal principles, but instead a break with spatiotemporal assumptions, namely that variables or states must always be localized.

If, then, quantum mechanics is the correct theory of our world, there is non-separability. For this insight, we would not, however, have required Bell's theorem in particular—a solid interpretation of the quantum-mechanical formalism would have been sufficient. If one wishes to base one's discussion solely on Bell's argument and not to give preference to any particular theory, then in contrast one cannot say reliably which type the resulting non-locality must be: A different theory of the microscopic world from quantum theory might possibly implement the unavoidable non-locality via non-local causal relations. This is, for example, the case in some interpretations of the de Broglie–Bohm theory (namely if one does not ascribe reality to the non-separable quantum potential).

4.4.6 Holism

The spatiotemporal non-separability of quantum theory is often mentioned together with another of its characteristics, which is termed "holism". Holism can likewise be derived from the consequences of Bell's argument only with considerable difficulties. Instead, it is—like non-separability—based on an analysis of the quantum-mechanical formalism. Roughly speaking, holism is the hypothesis that a whole is more than the sum of its parts. Sometimes the concept of "holism" is used simply as a synonym for "non-separability". This is insofar reasonable, in that a non-separable state does imply a type of spatiotemporal holism: A non-localized state extends over

a whole spacetime region, so that the states of all the sub-regions depend upon the state of the whole region.

More frequently, however, "holism" is used to denote a related but not identical hypothesis which refers not to spacetime regions, but instead to physical *systems*. Analogously to separability, one can define the opposite of holism as follows (cf. Teller 1989):

Particularism: For each pair of different physical objects S and T, it holds that:

1. Each one has its own, separate state,
   and
2. the state of the overall system is determined by the state of S and the state of T, and their spatiotemporal relations.

This definition of "particularism" is quite analogous to the definition of "separability", except that the things considered are not spacetime regions, but instead they are physical systems. Therefore, instead of "particularism", one could also speak of "system separability" (in contrast to "spatiotemporal separability"). The two concepts are differentiated in that the various systems need not be located in different spacetime regions. Particularism leads to the view that the world consists of individuals with intrinsic properties, and all the relations between individuals (except for the spatiotemporal relations) supervene on those intrinsic properties. Notably, particularism also states that individuals do not depend on each other, and (insofar as they are parts of a system) also not on the overall system of which they are a part.

Holism would then imply a violation of one of those conditions which define particularism, and according to quantum theory, indeed both conditions are violated. One can best see this by considering an entangled state, for example, the singlet state (4.2), which we note here again:

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}\Big(|\uparrow_z\rangle_1|\downarrow_z\rangle_2 - |\downarrow_z\rangle_1|\uparrow_z\rangle_2\Big)\,. \tag{4.29}$$

According to this state, neither of the objects has its own intrinsic spin state. Rather, it is only the entangled state of the overall system that describes the expected behaviour which would result from spin measurements.

One can interpret such entangled states ontologically in two ways. The first possibility is to assume that the subsystems do not actually exist when the entangled state is present. An entangled state would then be a whole without parts, which has the disposition to disintegrate into several subsystems when measured. These dispositions are represented by the specific combination of state vectors for the subsystems. Then, these combinations of state vectors may not be understood as describing properties that the system actually realizes ("categorical properties"), but instead, it should be seen as describing *potential* properties which are only realized when certain conditions are fulfilled (e. g. a suitable measurement).

The standard view, in contrast, is that, secondly, this construction out of subsystems is to be taken seriously in the ontological sense; that is, one presumes that the

entangled system does in fact consist of parts. Since these parts or subsystems do not have their own spin states, according to the description of the state in (4.29), this implies the remarkable situation that the overall system has a well-defined state, but the subsystems do not. The individual subsystems can be described only in terms of the state of the overall system, and in this sense, there is an ontological dependence of the parts on the whole. This is a clear-cut violation of the first condition for particularism and represents the special feature of quantum holism according to the standard interpretation (Esfeld 2004).

Furthermore, a violation of the first condition naturally also implies a violation of the second condition for particularism: Where there are no independent states of the individual systems, the state of the overall system cannot supervene on them. Many authors consider the entangled state as a relation between the subsystems. Teller (1986, 1989) writes in this sense of a "relational holism", Esfeld (2004) of a "metaphysics of relations". Others consider the state, on the contrary, to be an intrinsic property of the overall system (Healey 1991). Thus or so, it is clearly a non-supervenient property. In the quantum world, systems have properties on a higher level, which cannot be derived from their subsystems; i. e. a reduction of the overall system to its parts at a given instant ("synchronic microreduction") is in general not possible (Hüttemann 2005).

Since entanglement is an omnipresent phenomenon in the quantum world, one must presume that our world is holistically structured through and through on a fundamental level. The older picture of the world, successful for a long time, which held that it is constructed of very small particles that exist independently of each other and only are connected by interactions, is thus very probably incorrect. In the quantum world, the objects are woven into a whole which cannot be reduced to its components. It is important to emphasize that the holism of the quantum world is not some sort of "everything-depends-on-everything-else" assertion, but—as shown—it is conceptually readily comprehended and obeys clear mathematical rules. From this new viewpoint, it remains a serious question as to why we have the impression that the mesoscopic objects which we perceive in everyday life exist so relatively independently of one another.

4.4.7 Non-locality and the Relativity Principle

Thus far, everything has fit together quite smoothly. We were able to show that three of the four points of conflict with the theory of relativity could be avoided if we assume that: (1.) Non-locality does not require any transport of matter or energy; (2.) it also does not permit the transmission of signals at superluminal velocities, since the causal parameters can be correspondingly fine-tuned; and (3.) non-locality is realized in terms of a symmetric non-separability (instead of an asymmetric causal relation). Furthermore, such a non-separability is suggested by quantum theory (in a realistic interpretation such as GRW theory).

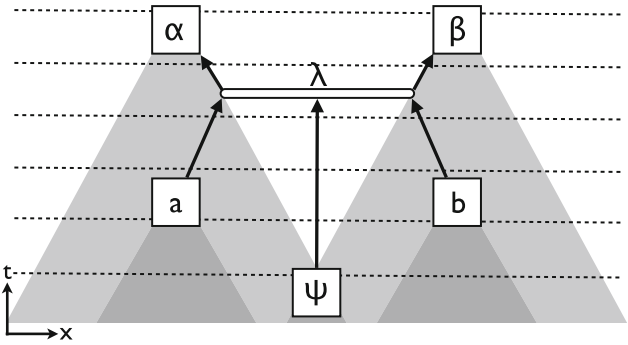
Now, however, we must turn to the fourth point of conflict with the theory of relativity, which deals with the fact that non-locality, on the one hand, has a tendency to prefer a particular frame of reference, the one in which it is simultaneous; and on the other, the principle of relativity requires that all reference frames must be equivalent. For this conflict to arise, it is unimportant whether we are dealing with a non-local causal relation or with non-separability; the conflict occurs for both types of non-locality. This conflict remains the greatest challenge for supporters of non-locality. Until today, there has been no generally accepted solution to it, and it is not clear whether the conflict can ever be resolved in a convincing manner. We will see that it goes so deep that it has been seriously suggested to abandon the principle of relativity on the fundamental level.

Before we go into the details of this topic, we should mention two things. First, there is already a conflict between non-locality and the special theory of relativity, so that in the following, we can concentrate on that theory (and leave aside the general theory of relativity).[31] Second, the tension between non-locality and the (special) relativity principle has its roots in an even deeper conflict: Non-locality violates the fundamental symmetry of the special theory of relativity, that is the Lorentz invariance of the laws of physics. Lorentz invariance is a stronger requirement than the principle of relativity, and Lorentz invariance is also the essential requirement of the (special) theory of relativity, its core. (For a theory is fully relativistic if and only if the laws that it involves are Lorentz invariant.) Lorentz invariance, however, is a very mathematical requirement, and we cannot delve into such technical questions here. Fortunately, the central idea of the conflict can also be illustrated using the weaker (and only necessary, not sufficient) condition of the principle of relativity. Now and then, where necessary, we shall refer to the stronger requirement of Lorentz invariance.

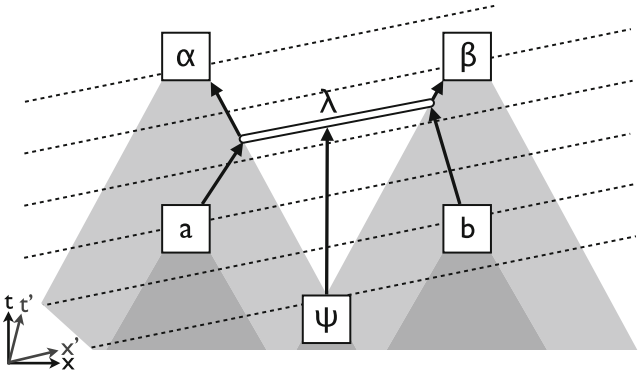
Let us now return to the conflict between non-locality and the principle of relativity. The fact that non-locality prefers a particular frame of reference can be exemplified with the help of spacetime diagrams. For purposes of illustration, let us assume that the quantum-mechanical description is correct and that its non-locality consists in a spatiotemporally non-separable quantum state λ. In Fig. 4.9, we show three (of infinitely many) possibilities of how such a non-separable state can be embedded in spacetime. The state is non-locally extended, and one can say that it lies along a spacelike hyperplane ("spacelike", since the points along this plane are spacelike-separated from each other, and "hyper-", since it is a plane within four-dimensional spacetime). For every non-local state λ, there is just one corresponding frame of reference whose simultaneity planes[32] all lie parallel to it; i. e. in this frame of reference,

[31] While in the general theory of relativity, *all* frames of reference are considered to be equivalent, this holds in the special theory of relativity only for *inertial* frames, i e. non-accelerated frames (the principle of special relativity).

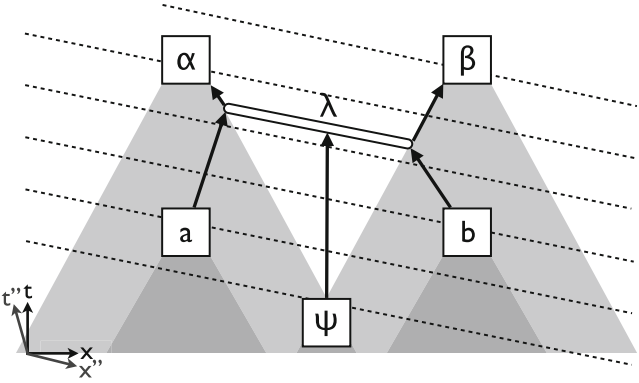
[32] A simultaneity plane of a frame of reference through a point P is the set of all spacetime points which are simultaneous with P *in this frame of reference*. Simultaneity in the theory of relativity is, in contrast to its role in classical theories, not an absolute, objective fact, but rather it depends upon the frame of reference. Every inertial frame of reference can be uniquely determined by specifying one of its simultaneity planes.



(a) λ simultaneous in the laboratory frame S



(b) λ simultaneous in the frame S′ (which is moving in the $+x$ direction relative to S)



(c) λ simultaneous in the frame S″ (which is moving in the $-x$ direction relative to S)

Fig. 4.9 Various orientations of a non-local state

all changes of the non-local state propagate instantaneously. It is this frame which is distinguished by the corresponding state. Distinguishing such a frame of reference now contradicts one of the two axioms of the theory of relativity, i. e. the principle of relativity, which states that all frames of reference are equivalent. This principle means that physical processes can be described from all frames of reference by the same physical laws. Therefore, there can be no facts which nullify the equivalence of all frames of reference— and non-local states would seem to be precisely such facts.

A resolution of this central point of conflict between quantum theory and the theory of relativity could now consist in specifying *along which hyperplane the non-local quantum states can lie without violating the relativity principle*. T. Maudlin (2011) discusses this topic in detail, and we give here a brief summary of his line of thinking. If one were to say that all non-local states lie along a family of parallel hyperplanes, then the corresponding frame of reference would be distinguished—this possibility can thus be eliminated. A natural suggestion would be that the hyperplane could be chosen *at random*. If every hyperplane would have exactly the same probability of being chosen for every entangled system which was being formed, a 'democracy of reference frames' would be guaranteed. Unfortunately, however, as Maudlin proves, there is no Lorentz invariant possibility for attributing a measure of probability to hyperplanes—so that this possibility must also be eliminated.

Finally, there remains only the possibility of determining the hyperplane on the basis of the distribution of matter. The particular matter distribution in a certain spacetime region is not governed by physical law, but is rather a contingent fact, and since the relativity principle refers specifically to laws, it would not be threatened if one were to make use of the matter distribution in order to choose a hyperplane. One suggestion could then be to determine the hyperplane via the matter distribution of the entire universe. For example, the centre of gravity of all the matter in the universe could determine a hyperplane. The problem with this suggestion is that the centre of gravity is the result of an averaging process over the mass distribution *at a given time*. In order to find the centre of gravity, one would thus require an objective concept of simultaneity—but according to the theory of relativity, precisely such a simultaneity does not exist. And for the same reason, choosing a frame of reference via any other quantity averaged over the whole universe would also fail.

Among the various possibilities which Maudlin discusses, and which are all rejected, we will pick out just one more, which at first glance appears quite natural: One could think that the hyperplane could be determined by the state of motion of the source of the entangled system. It corresponds to the plane of simultaneity in the rest system *of the source*. The problem with this suggestion is that there is empirical evidence that even photons which are emitted by *different* sources can still violate Bell inequalities. If these sources are moving relative to each other, it is then completely unclear which of the corresponding hyperplanes should be chosen for the entangled photon state. Maudlin concludes that there is no imaginable possibility of choosing exactly *one* hyperplane for each entangled system without violating the relativity principle.

If we wish to maintain consistency with the theory of relativity, according to Maudlin there are finally only two possibilities, which both demand a heavy price, however: hyperplane dependence or a GRW theory with a flash ontology. Hyperplane dependence is a suggestion which originated with G. Fleming (1986, 1989 jointly with Bennett), in his attempt to obtain a fully Lorentz invariant quantum theory. The background was the observation that quantum field theory (cf. Chap. 6), which is often presented as the relativistic generalization of quantum mechanics, is not a Lorentz invariant theory of *all* microscopic events. The Lorentz invariance of quantum field theory is limited to processes *between* measurements, but the collapse of the wavefunction due to measurements, and with it the non-locality of the microscopic world, remains excluded. Fleming wanted to find a Lorentz invariant theory which *included* the measurement process. From our considerations above, it is clear that it will be difficult to find a criterion for choosing the *one* hyperplane along which a given quantum state lies. Fleming therefore takes a different path. His answer to the question of along which hyperplane an entangled quantum state lies is: along *all* of them! A quantum state becomes an object which is defined only *relative to hyperplanes* (thus "hyperplane dependent"). A given entangled quantum system then has a certain quantum state relative to each hyperplane. Depending on the hyperplane along which one considers the system, it exhibits a different quantum state at every given spacetime point.

While Fleming's theory represents a technically clean solution to the problem (the theory is fully Lorentz invariant and yields correct predictions), it is clear that this theory is highly counter-intuitive. From the philosophical perspective, one has to say that the number of entities has grown boundlessly (infinitely many quantum states for each quantum system). We are not exactly dealing with an ontologically parsimonious position. To be fair, however, one must also say that the explosion of entities does not occur without reason: We are searching for the resolution of a very deep conflict. Is the price paid by this theory too high? Many would indeed say "yes".

In the past few years, the theory of hyperplane dependence is no longer without alternatives. There is now another theory which describes all quantum processes in a fully Lorentz invariant manner, and in particular therefore does not distinguish a special frame of reference: the GRW theory with a flash ontology (for short: GRWf; Tumulka 2006a, b). We introduced this theory in Sect. 2.4 and already noted there that it has two decisive disadvantages: Firstly, at present it has been formulated only for non-interacting particles, and it is not certain whether there will ever be an extended version with interactions. Therefore, in contrast to the theory of hyperplane dependence, we can so far speak only of the first step towards a Lorentz invariant theory. Secondly, this theory buys its Lorentz invariance at a high ontological price: The suggested ontology consists in the evolution of the quantum state within the many-dimensional configuration space, and only now and then, when a collapse occurs according to the stochastic GRW dynamics, do events manifest themselves as so-called flashes in our normal four-dimensional spacetime. Between the flashes, there exists in spacetime—nothing! Normal things like tables and rabbits are then not spatiotemporally continuous objects, but in reality only a galaxy of flashes, which, however, occur so rapidly in sequence that we believe we are seeing continuous

objects. The time resolution of our perceptions is simply too poor. If one is willing to pay this ontological price, then one is on a track that might not be completely hopeless. Many authors, however, consider this price also to be too high.

If one is not willing to condone the counter-intuitive ontologies which follow from hyperplane dependence or the GRW-flash theory, then from today's viewpoint, there remains only the possibility which is described by Maudlin as perhaps even the most plausible way out, namely to assume that there is in fact a preferred frame of reference in the structure of spacetime. This would mean assuming that in addition to the relativistic structure of spacetime, there is also the structure of a preferred foliation, which determines planes of simultaneity globally. It would, however, not require that we explain relativistic effects in terms of a Lorentzian ether wind or some such exotic mechanism. The relativistic effects could be explained in the usual way; it would merely be the case that spacetime would contain an *additional structure*, which would allow quantum systems to generate the empirical results of EPR/B experiments without resorting to an extravagant ontology. The disadvantage of this suggestion is that the elegant relativistic symmetry and structures would be broken by this additional structure. Furthermore, the assumption of an additional structure is rather strong, and it would also have the character of an *ad hoc* assumption, if it were found not to contribute to the solutions of some other problems.

In the end, we can only emphasize that all three suggestions— hyperplane dependence, GRW-flashes, and a preferred frame of reference—all are bought at a high cost. Given that it is not obvious which should be considered the most plausible solution, it is indeed acceptable to argue in favour of each of these three positions. But it is not particularly surprising that no one of the three solutions has as yet been generally accepted. Only one thing is certain: The discussion of the compatibility between the theory of relativity and quantum-non-locality will continue.

4.5 Alternative Solutions

Since none of the possible solutions to the compatibility of the theory of relativity with non-locality has been generally accepted, some authors have taken a quite different route: They argue that instead of the locality assumption, one of the other assumptions which are required to derive Bell's theorem, and which we called the "background assumptions", is violated. Here, almost every conceivable position has been adopted, which makes the debate complex and confusing. In the following, we will briefly sketch for each background assumption what abandoning it would imply, whether abandoning it can explain the violation of Bell's inequalities, and if it is plausible to abandon it.

4.5.1 The Causal Markov Condition

The causal Markov condition played a central role in the derivation of Bell's theorem. On the one hand, it was the bridge principle which allowed us to infer statistical facts from causal structures; on the other, it was the methodological principle which demanded that all correlations must have a causal explanation.

Could the Markov condition be violated? There has been a notable discussion about the validity of Reichenbach's principle of the common cause, in which possible counter-examples have been suggested and discussed. Since this principle, as we have already noted above, represents a special case of the causal Markov condition, it is clear that counter-examples opposing Reichenbach's principle would also be arguments against the Markov condition. We cannot discuss all the alleged counter-examples against Reichenbach's principle here (for a more detailed discussion, see e. g. Arntzenius 2010), but the decisive point is that one type of the counter-examples has good prospects of success (which, however, is also controversial; see Hausman and Woodward 1999). The examples in question are common causes in indeterministic worlds which do not screen off their effects (van Fraassen 1982; Cartwright 1988), i. e. conditionalizing on the common cause does not make its effects statistically independent.

The example given by Cartwright is a molecule which breaks apart into two equal-sized fragments (smaller molecules or atoms), which move off in opposite directions owing to the conservation of linear momentum. In such a case, the state of the molecule before its indeterministic decay does not screen off the motion of its fragments afterwards, although it is the common cause of their motion (and, according to the Markov condition, common causes must screen off their effects statistically). If there are no hidden variables, then the entangled quantum-mechanical state (4.8) of the photons at the source is also a common cause which does not screen off its effects ("interactive common cause" as opposed to a usual "conjunctive common cause", see the grey box). In other words, if quantum mechanics is complete, then Reichenbach's principle and the Markov condition are violated. In Fig. 4.10, we have illustrated the local structure which the quantum state obtains as an interactive common cause (symbolized by the arc between the two emerging arrows). But could such a structure with a quantum state that does not screen off explain the EPR/B correlations?

The answer is no. Interactive common causes alone cannot violate Bell inequalities. One can show that the correlations which are produced by an interactive common cause have only a similar strength to those of two usual common causes which together screen off their effects (Näger 2013). Since in the latter situation, however, Bell inequalities are implied, a violation of the causal Markov condition by an interactive common cause is not sufficient to explain the observed EPR/B correlations.

Quantum mechanics, for example, violates the causal Markov condition in the sense mentioned, *and* it furthermore violates the locality assumption. We have seen that the structure of quantum mechanics is that shown in Fig. 4.8c (and thus *not* that given in Fig. 4.10). According to quantum mechanics, the non-local quantum

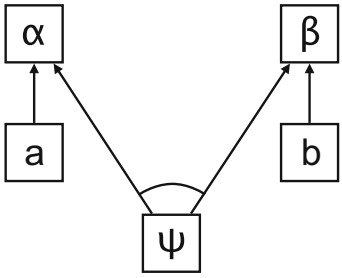


Fig. 4.10 Local structure with an interactive common cause

state λ before the collapse does not screen off the correlations, and, to be correct, this should also be noted in the diagram (e. g. by an arc analogous to the one in Fig. 4.10). However, the decisive reason why the quantum-mechanical structure can violate a Bell inequality is the fact that the common cause λ is influenced by the measurement settings, so that there is a causal path from at least one of the settings to the distant outcome. Since such a path is not present in the structure shown in Fig. 4.10, the fact that the common cause does not screen off its effects is not sufficient to guarantee a violation of Bell inequalities.

We should emphasize that such a violation of the causal Markov condition must not necessarily imply a break with the implicit methodological principle that all correlations must have a cause. Instead, one can formulate a generalized causal Markov condition which takes into account such interactive common causes and maintains the methodological principle (Näger 2014). Insofar, the violation of the causal Markov condition discussed here is only a kind of weak violation. It does not contradict the basic idea of the principle.

Common causes and statistics

Screening off: A variable Z screens off two variables X and Y statistically from each other if the latter are only marginally statistically dependent ($P(X_1|Y_1) \neq P(X_1)$ for at least one pair of values X_1, Y_1), but are independent given Z ($P(X|YZ) = P(X|Z)$, i.e. for all values of the variables).

Conjunctive common cause: A common cause Z of X and Y, $X \leftarrow Z \rightarrow Y$, is conjunctive, when the statistical pattern of screening off holds. This is the typical case of common causes, for which the Markov condition holds.

Interactive common cause: A common cause Z of X and Y is interactive, when it does not screen off its effects. Such cases violate the causal Markov condition.

Decaying molecule: Let S be the state of the molecule before its decay, and assume that the molecule decays with equal probability

$$P(p, -p|S) = \frac{1}{2} = P(p', -p'|S) \tag{4.30}$$

along one of two directions, where $p, -p$ or $p', -p'$ are the states (here: momenta) of the fragments after the decay. That is, the states of the fragments after the decay are perfectly anti-correlated, but the state of the molecule S does not screen off the correlation:

$$P(p, -p|S) = \frac{1}{2} \neq P(p|S) \cdot P(-p|S) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}. \tag{4.31}$$

Hence, S is an interactive common cause.

Entangled quantum state: Very similar to the decaying molecule, entangled quantum states are interactive common causes of their measurement outcomes. For instance, when measured along the same direction z, the entangled state $|\phi^+\rangle$ from Eq. (4.8) decays with the probability $\frac{1}{2}$ either into the state $|+_z\rangle_1|+_z\rangle_2$ or into the state $|-_z\rangle_1|-_z\rangle_2$. Mutatis mutandis, Eq. (4.31) holds (replace p by $|+_z\rangle_1$, $-p$ by $|+_z\rangle_2$, p' by $|-_z\rangle_1$, $-p'$ by $|-_z\rangle_2$, and S by $|\phi^+\rangle$).

We could, however, imagine stronger violations which would lead us to accept the EPR/B correlations as causally unexplained correlations. While Butterfield (1989) holds this to be a mysterious but unavoidable fact, Fine (1989) argues that our ideal of an explanation, that is to explain all the correlations, is obsolete. One can also accept causally unexplained basal correlations, and the EPR/B correlations are a good candidate for such a case. Given that it is the central methodological assumption of all empirical sciences that correlations require explanation, accepting unexplained correlations would be a rather radical consequence. (In the game of the persons, this would mean that there are correlations which are stronger than one could get by prior agreement on a strategy, but nevertheless one would assume that this fact does not require explanation.) Thus, such a conclusion should probably be seen as a last resort, to be adopted only when all the other suggested solutions have proved to be untenable.

4.5.2 The Intervention Assumption

In the causal structures which we have thus far examined, the choice of measurement settings and of the quantum state exhibited only outgoing arrows, never incoming arrows; i. e. they were only causes, but never effects of the other variables. Such variables are called "exogenous". The justification for the exogeneity of these variables is that their values were determined through external interventions by the experimenter independently of each other and of other variables. "Intervention" is a technical term here and means that we are dealing with a variable which is (i) the direct cause of precisely one variable (and of no others); which (ii) determines the value of that variable; and which (iii) itself is not an effect of any of the variables considered (Spirtes *et al*. 1993, Sect. 3.7.2; Pearl 2000, Sect. 3.2).

If the actions of the experimenter who prepares the quantum state and chooses the measurement settings are interventions, then this means that due to (ii), the settings and the quantum state cannot be *direct* effects of any other variable. For if the experimenter, for example, determined the setting, then that setting cannot be influenced directly by any other of the variables of the experiment; determination prevents a direct influence through other factors. The variables which are controlled by the experimenter could then at most be indirectly influenced by other variables, namely if a variable of the experiment influences the intervention of the experimenter—but that is excluded by (iii). Therefore, the controlled variables must be exogenous if the actions of the experimenter are indeed interventions.

But are the actions of the experimenter in fact interventions? The following principle formulates the usual view:

> **Intervention assumption**: The experimenter (or a machine which is set up and programmed by the experimenter) can control the settings of a macroscopic apparatus by means of an intervention.

This assumption is a part of the everyday scientific methodology of controlled experiments. If an experimenter carries out a measurement and in the process controls certain variables of the involved measurement apparatuses whose values he chooses independently, then he naturally assumes that these variables are exogenous. There is no particular reason to assume that this principle should not hold for quantum-mechanical experiments: The measurement devices in such experiments are quite normal macroscopic setups, whose settings are controlled by the experimenter.[33]

Nevertheless, the possibility that this principle is violated in EPR/B experiments has been seriously raised. The reason is that abandoning this principle seems to be the only possibility of providing a causal explanation that is straightforwardly compatible with a relativistic space-time: It would appear to be the only possibility

[33] Note that also the preparation of the quantum state is controlled by an experimental apparatus, which, if it is correctly constructed, can be started by a corresponding macroscopic mechanism (e. g. a pushbutton, a laser beam, an electrical pulse)—and thus falls within the area of validity of the intervention assumption.

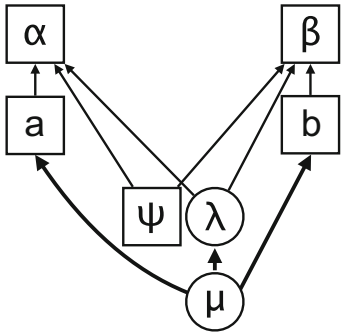


Fig. 4.11 Hidden common cause of the measurement settings (violation of the intervention assumption)

of giving a local causal explanation of the correlations which acts forward in time. On a probabilistic level, such models violate measurement independence.

The basic model is the one shown in Fig. 4.11 (cf. e. g. Suppes and Zanotti 1981). Here, there is an additional hidden variable μ which is a common cause of the hidden variable of the quantum state λ and of the measurement settings; indirectly, namely via λ, μ is also a common cause of the measurement outcomes. (In the game of the persons this would mean that, e. g. the scientist conducting the experiment both assigns a strategy to the test persons and predetermines the questions of the enquirers. It is clear that by suitably adjusting questions and strategy, one can generate stronger correlations than in the case that questions and strategy are chosen independently.) To keep the influence of μ on the settings local, μ must lie within the common past cone of the settings; that is, temporally before the emission of the photons by the source. μ could then not be a property of the photons, since they do not yet exist at this time; instead, it would have to be considered as a hidden property of the source (or even as property of another object, if the source does not yet exist at this point in time; the extreme case would be that μ is a property of the Big Bang). It would seem extremely strange to assume that before the emission (possibly even before the preparation of the photons) a causal process emerges from the source which can influence the distant measurement settings, but we do not wish to discuss this assumption further here. More importantly, we need to ask what it might mean that μ can influence the measurement settings, given that they are determined by the experimenter. It can mean only that the experimenter is herself influenced by μ to choose a certain setting. According to this model, the experimenters could not make a (libertarian) free decision about which setting to choose.

Most authors assess this solution of the problem to be rather implausible. For one thing, it is *ad hoc* and improbable that the actions of experimenters can be crucially influenced or even determined by hidden states of, e. g. a source of quantum objects. The suggestion sounds more like a bizarre fantasy world than a serious, scientifically well-founded explanation. Indeed, it has been noted that this approach would amount to a "cosmic conspiracy". For it would be not only the scientists who must be influenced by the quantum states; everything, whatever determines the measurement settings—in modern experiments for example an optical switch or a computer programme which operates on the basis of a random-number generator—

would have to be correspondingly influenced by the hidden state. Such an influence, which acts in an undifferentiated manner on whatever determines the measurement settings, would seem to be highly implausible.

Finally, we should also note that science, as we know it and have evidently practiced it with some success, would no longer be possible for such quantum systems: It is one of the basic principles of the scientific method that the question of what is to be measured can be decided or varied independently of which state the system to be measured is in. If, however, the system itself dictates the questions which can be asked of it, then it is impossible to investigate it in a deeper sense. Large parts of its properties might then remain undiscovered, and normal causal conclusions would no longer be possible for such systems.

4.5.3 Backwards Causation

Finally, we made a further implicit assumption in the derivation of Bell's theorem: Namely, that causal relations between timelike-separated events are always directed forward in time ("no backwards causation"). Only by excluding backwards causation and with the locality assumption can we be certain that none of the variables of the one wing of the experiment could influence a variable of the other wing. For if such an influence were possible, one could have, for example, an influence of b on α, which could be mediated in their common past by some variable μ (see Fig. 4.12). For the derivation of Bell's theorem, we therefore had to assume:

> **No backwards causation**: Effects do not occur earlier than their (timelike-separated) causes.

If one allows backwards causation, it is then possible to generate the same causal structures as by abandoning the locality assumption, except that the causal paths would have a different embedding in spacetime. In other words, each of the non-local causal graphs described above could be produced just as well by replacing the non-local causal relations that they contain by an action which first goes backward in time and then forward. For this reason, such scenarios are also called "zigzag causation".

In principle, every influence between the wings can thus be generated by paths involving backwards causation, but naturally not all of them are equally plausible. The structure shown in Fig. 4.13 is one of the typical proposals for explaining the correlations by backwards causation. Here, the chosen measurement setting influences the hidden variable λ of the photons at the source. To ensure that this is indeed backwards causation (and not a non-local effect), the influence must occur only at the moment when the photon arrives at the detector. Price (1994) suggested such a model and has developed his approach further in recent years (e. g. Price 2012). Cramer (1980, 1986) has even given a detailed mathematical description of such a theory, according to which the detector sends a confirmation wave back to the source

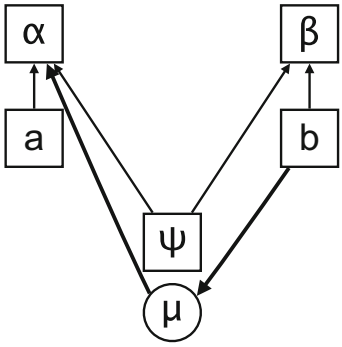


Fig. 4.12 The causal connection between the two wings via backwards causation

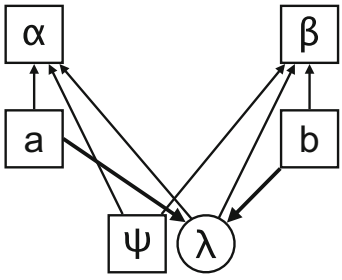


Fig. 4.13 A typical structure with backwards causation

at the moment when the quantum-mechanical wavefunction arrives. In this way, it is guaranteed that the source will emit objects in an appropriate state so that they produce the corresponding correlations in the measurements. In the strategy game with two players (see Sect. 4.3.3), this would mean that the players already know the question that will be asked while they are still in the same room, so that they can agree on an appropriate strategy in advance. Under these circumstances, it is obviously simple to produce the correlations. Models with backwards causation can produce the observed EPR/B correlations. But are they also convincing?

The great advantage of models with backwards causation is that they do not contradict the theory of relativity. The theory of relativity would appear to have a certain conflict with non-local relations, but not with causal actions which act backwards in time. Independently of its compatibility with the theory of relativity, backwards causation has been strongly criticized, however. Some authors claim that backwards causation is conceptually impossible (i. e. it is self-contradictory). Among them are the supporters of a dynamic theory of time, according to which the state of the world *just now* produces the immediately following future state. Such concepts of time are incompatible with backwards causation, since the past no longer exists, and can therefore not be influenced by the present. Only assuming a static theory of time is a concept such as backwards causation conceivable at all. Causal explanations are then not metaphysical explanations of how the correlations come about, but instead are epistemic histories which merely concern the explanation of the correlations.

But even given a static theory of time, there are serious objections to backwards causation. For example, also supporters of a causal theory of time—who assume that the direction of time is determined by the direction of causality—hold backwards causation to be a conceptual impossibility. Others admit its conceptual possibility, but assert that backwards causation leads to paradoxical causal loops (A at t_1 causes B at t_2 causes not-A at t_1), and thus cannot be realized in the real world (Mellor 1981). T. Maudlin (2011, Chap. 7), for example, argues that Cramer's model of the EPR/B correlations is inconsistent in this latter sense.

All these arguments, however, do not exclude explanations making use of backwards causation under certain circumstances. If there are limitations on what can influence what, then there could nevertheless be consistent theories with backwards causation. The decisive point, however, is that we have no sort of indications that backwards causation might exist in fact. On the one hand, the experimental data give no evidence for backwards causation. With normal causality, there are as a rule asymmetries in the statistics which favour a particular direction, and backwards causation should show similar asymmetries, but simply in the reversed direction. But neither in the statistics of EPR/B experiments, nor in any other kind of statistics, have such asymmetries ever been observed. A supporter of backwards causation in the EPR/B case could respond that the alleged backwards causation acts via the hidden variable λ and therefore leaves none of the usual tracks in the statistics. As we have already noted: possible, but not particularly convincing.

In the end, still more important for most authors is the fact that there are also no theoretical indications of backwards causation. All known theories which consider the observed probabilities to be objective, and thus in particular the GRW theory and the de Broglie–Bohm theory, contain a non-locality. In sum, the impression remains that backwards causation rather is an *ad hoc* explanation for the violation of Bell's inequalities.

4.5.4 Conclusion: Alternative Suggestions

The alternatives to the locality condition discussed in this section were born out of desperation in view of the problems that arise from the incompatibility between the theory of relativity and non-locality. The existing compatibility suggestions are not convincing, at least not on first view. In contrast to the state of the discussions around 30 years ago, however, it would appear that compatibility between non-locality and the theory of relativity is no longer considered to be impossible. The current state of desperation is thus not quite as acute as it was previously, and it is questionable whether in view of that background, any of the alternative solutions will have sufficient persuasive power to be considered seriously.

4.6 Résumé

This chapter tells the story of fascinating insights, great surprises and disappointed hopes. Albert Einstein and his co-authors wanted to show that quantum mechanics is incomplete. They did not attain that goal, but their pioneering article raised a new problem—unintentionally— which contradicted Einstein's intuitions still more intensely than accepting the completeness of quantum mechanics would have done: The quantum world is non-local.

It was J. S. Bell who drew the right conclusions from the EPR article and developed a new approach with his inequality, which made the non-locality in the special EPR/B correlations apparent. The empirical confirmation of those correlations is currently carried out with experimental setups which are essentially the same as the one suggested by the EPR thought experiment. These experiments have confirmed that the predictions of quantum mechanics are correct. Measurements on the components of a system that is described by an entangled state exhibit strong correlations which are still observable even when the detectors are set up at a considerable distance from each other.

The results of J. S. Bell have also destroyed hopes based on the idea that the quantum-mechanical state is incomplete and must be characterized more precisely. Even theories with hidden variables cannot explain the predicted and observed correlations unless they are non-local. The generality of his proof eliminates the possibility that we will find a new solution through technical refinements of the quantum-mechanical formalism.

Do we have an explanation for the empirical phenomena of entangled systems? On the one hand, quantum mechanics explains the correlations which occur in the sense that they can be correctly derived from its formalism. On the other hand, Günther Ludwig writes about the question of how the correlations come about: "The answer, being disappointing perhaps, is that quantum mechanics can say nothing about it" (Ludwig 1971, p. 312), and he evidently means that quantum mechanics provides no mechanism (for example, in the sense of Machamer *et al*. 2000) which could explain the occurrence of the correlations and thus could lead to an intuitively clear understanding. The best that one can do in the spirit of the mechanistic programme is to indicate a causal structure which would relate the entangled quantum state to the measurement settings and the measurement outcomes, but the entangled quantum state itself cannot be further disassembled into interacting parts. Furthermore, one can hardly imagine a simple and intuitive spatiotemporal description of such a mechanism which would be unproblematically compatible with the theory of relativity. Here, two hopes have been destroyed: The idea that we could carry out a thorough spatiotemporal analysis of the microworld, and the expectation that familiar mechanistic ideals of explanation should always be applicable in fundamental physics. The particular value of these insights is comparable to the benefits gained when one is forced to give up one's illusions: Thereafter, the remaining problems can be attacked in a less biased manner.

Bell's proof shows that at least one of the plausible methodological and metaphysical assumptions which have proved themselves in classical physics must be abandoned. While the discussion during the more than four decades since Bell's discovery has achieved ever deeper insights into the preconditions, scope and consequences of Bell's proof, it has not yet led to a generally accepted and readily comprehensible result; in particular not with respect to the question of how the non-locality of quantum mechanics could be reconciled with relativistic spacetime. This remains the central question in the debates surrounding entangled states.

Exercises

1. In the EPR article, there is an assumption of major significance which we have called the "locality assumption": "Since at the time of measurement, the two systems no longer interact, no real change can take place in the second system in consequence of anything that may be done to the first system". How does this assumption relate to the other concepts of locality which we introduced in Sect. 4.4.5: Does it imply global and causal Einstein locality as well as spatiotemporal separability?
2. Assume that an EPR/B experiment were correctly described by a local causal structure with hidden variables λ (see Fig. 4.5). One can then show that the existence of perfect correlations implies that measurements must proceed *deterministically*. Try to formulate a suitable argument.
3. List the minimal set of assumptions which are required to derive a Bell inequality and sketch out what they state.
4. Apply the causal Markov condition to the local causal structure in Fig. 4.5 and note the resulting statistical independencies.
5. Outline the four fields of conflict of a non-local theory with the theory of relativity.
6. Discuss the two known possibilities for non-locality which are not in conflict with the principle of relativity. Take into account both physical and ontological consequences.

References

Arntzenius, Frank. 1994. Space-like connections. *British Journal for the Philosophy of Science* 45 (1): 201–217.

Arntzenius, Frank. 2010. Reichenbach's common cause principle. In *The Stanford Encyclopedia of Philosophy*, ed. E.N. Zalta. http://plato.stanford.edu/entries/physics-Rpcc/

Aspect, Alain, Jean Dalibard, and Gérard Roger. 1982. Experimental test of Bell's inequalities using time-varying analyzers. *Physical Review Letters* 49: 1804–1807.

Audretsch, Jürgen (ed.) (2006). *Entangled World: The Fascination of Quantum Information and Computation*. Weinheim: Wiley-VCH. Trans. Ehrlich, Rudolf. (originally published in German as *Verschränkte Welt. Faszination der Quanten*, 2002)

Audretsch, Jürgen. 2007. *Entangled Systems: New Directions in Quantum Physics*. Weinheim: Wiley-VCH. (originally published in German as *Verschränkte Systeme. Die Quantenphysik auf neuen Wegen*, 2005)

Bell, John S. 1964. On the Einstein-Podolsky-Rosen paradox. *Physics* 1 (3): 195–200.

Bell, John S. 1971. Introduction to the hidden-variable question. In *Foundations of Quantum Mechanics: Proceedings of the International School of Physics 'Enrico Fermi', Course XLIX*, ed. B. d'Espagnat, 171–181. New York: Academic Press. (Reprinted in Bell 1987, 29–39).

Bell, John S. 1975. The theory of local beables. *TH-2053-CERN*. (Reprinted in Bell 1987, 52–62).

Bell, John S. 1987. *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Mechanics*. Cambridge: Cambridge University Press.

Bohm, David. 1951. *Quantum Theory*. Englewood Cliffs NJ: Prentice-Hall.

Bohr, Niels. 1935. Can the quantum-mechanical description of physical reality be considered complete? *Physical Review* 48: 696–702.

Butterfield, Jeremy. 1989. A spacetime approach to the Bell inequality. In Cushing and McMullin 1989, 114–144.

Cartwright, Nancy. 1988. How to tell a common cause. In *Probability and Causality: Essays in Honor of Wesley C. Salmon*, ed. J.H. Fetzer, 181–188. Dordrecht: Reidel.

Clauser, John F., Michael A. Horne, Abner Shimony, and Richard A. Holt. 1969. Proposed experiment to test local hidden-variable theories. *Physical Review Letters* 23: 880–884.

Clauser, John F., and Michael A. Horne. 1974. Experimental consequences of objective local theories. *Physical Review D* 10 (2): 526–535.

Cleland, Carol. 1984. Space: An abstract system of non-supervenient relations. *Philosophical Studies* 46: 19–40.

Cramer, John G. 1980. Generalized absorber theory and the Einstein-Podolsky-Rosen paradox. *Physical Review D* 22: 362–376.

Cramer, John G. 1986. The transactional interpretation of quantum mechanics. *Reviews of Modern Physics* 58: 647–687.

Cushing, James T., and Ernan McMullin (eds.). 1989. *Philosophical Consequences of Quantum Theory: Reflections on Bell's Theorem*. Notre Dame: University of Notre Dame Press.

Einstein, Albert. 1948. Quanten-Mechanik und Wirklichkeit (Quantum mechanics and reality). *Dialectica* 2: 320–324.

Einstein, Albert, Boris Podolsky, and Nathan Rosen. 1935. Can the quantum-mechanical description of physical reality be considered complete? *Physical Review* 47: 777–780.

Esfeld, Michael. 2004. Quantum entanglement and a metaphysics of relations. *Studies in History and Philosophy of Modern Physics* 35 (4): 601–617.

d'Espagnat, Bernard. 1979. The quantum theory and reality. *Scientific American* 241 (5): 128–141.

Fine, Arthur. 1982. Some local models for correlation experiments. *Synthese* 50 (5): 279–294.

Fine, Arthur (1989). Do correlations need to be explained? In Cushing and McMullin 1989, 175–194.

Fine, Arthur. 2013. The Einstein-Podolsky-Rosen argument in quantum theory. In *The Stanford Encyclopedia of Philosophy* ed. E.N. Zalta. http://plato.stanford.edu/entries/qt-epr/

Fleming, Gordon N. 1986. On a Lorentz-invariant quantum theory of measurement. In *New Techniques and Ideas in Quantum Measurement Theory*, ed. D.M. Greenberger, 574–575. New York: New York Academy of Sciences.

Fleming, Gordon N., and Harry Bennett. 1989. Hyperplane dependence in relativistic quantum mechanics. *Foundations of Physics* 19: 231–267.

French, Steven. 1989. Individuality, supervenience and Bell's theorem. *Philosophical Studies* 55: 1–22.

Friebe, Cord. 2004. Teilen, Trennen und Vereinen: EPR ohne Holismus. *Journal for General Philosophy of Science* 35 (2): 261–281.

Ghirardi, GianCarlo, Luca Marinatto, and Tullio Weber. 2002. Entanglement and properties of composite quantum systems: A conceptual and mathematical analysis. *Journal of Statistical Physics* 108: 49–122.

Glymour, Clark. 2006. Markov properties and quantum experiments. In *Physical Theory and Its Interpretation: Essays in Honor of Jeffrey Bub*, ed. W. Demopoulos, and I. Pitowsky, 117–125. Dordrecht: Springer.

Gröblacher, Simon, Tomasz Paterek, Rainer Kaltenbaek, Caslav Brukner, Marek Zukowski, Markus Aspelmeyer, and Anton Zeilinger. 2007. An experimental test of non-local realism. *Nature* 446: 871–875.

Hausman, Daniel M., and James Woodward. 1999. Independence, invariance and the causal Markov condition. *The British Journal for the Philosophy of Science* 50 (4): 521–583.

Healey, Richard A. 1991. Holism and nonseparability. *Journal of Philosophy* 88: 393–421.

Held, Carsten. 2006. The Bohr-Einstein debate and the fundamental problem of quantum mechanics. In Audretsch 2006, 65–90.

Hensen, B., H. Bernien, A.E. Dréau, A. Reiserer, N. Kalb, M.S. Blok, J. Ruitenberg, R.F. Vermeulen, R.N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M.W. Mitchell, M. Markham, D.J. Twitchen, D. Elkouss, S. Wehner, T.H. Taminiau, and R. Hanson. 2015. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* 526 (7575): 682–686.

Hooker, Cliff A. 1972. The nature of quantum mechanical reality: Einstein versus Bohr. In *Paradigms & Paradoxes: The Philosophical Challenge of the Quantum Domain*, ed. R.G. Colodny, and A. Fine, 67–302. Pittsburgh: University of Pittsburgh Press.

Howard, Don. 1985. Einstein on locality and separability. *Studies in History and Philosophy of Science* 16: 171–201.

Howard, Don. 1989. Holism, separability, and the metaphysical implications of the Bell experiments. In Cushing and McMullin 1989, 224–253.

Howard, Don. 1990. Nicht sein kann, was nicht sein darf, or the prehistory of EPR 1909–1935: Einstein's early worries about the quantum mechanics of composite systems. In *Sixty-two Years of Uncertainty: Historical, Philosophical, and Physical Inquiries Into the Foundations of Quantum Mechanics*, ed. A.I. Miller, 61–111. New York: Plenum.

Hüttemann, Andreas. 2005. Explanation, emergence, and quantum entanglement. *Philosophy of Science* 72: 114–127.

Jammer, Max. 1974. *The Philosophy of Quantum Mechanics: The Interpretations of Quantum Mechanics in Historical Perspective*. New York: Wiley.

Jarrett, Jon P. 1984. On the physical signifance of the locality conditions in the Bell arguments. *Noûs* 18: 569–590.

Jones, Martin R., and Robert K. Clifton. 1993. Against experimental metaphysics. In *Philosophy of Science*, ed. P.A. French, T.E. Uehling, and H.K. Wettstein, 295–316. Notre Dame: University of Notre Dame Press.

Kiefer, Claus (ed.). 2015. *Albert Einstein, Boris Podolsky, Nathan Rosen: Kann die quantenmechanische Beschreibung der physikalischen Realität als vollständig betrachtet werden?* Berlin: Springer Spektrum.

Ladyman, James, Øystein Linnebo, and Tomasz F. Bigaj. 2013. Entanglement and non-factorizability. *Studies in History and Philosophy of Modern Physics* 44 (3): 215–221.

Lewis, David. 1986. *Philosophical Papers*: *Volume II*. Oxford: Oxford University Press.

Ludwig, Günther. 1971. The measuring process and an axiomatic foundation of quantum mechanics. In *Conceptual Foundations of Quantum Mechanics*, ed. B. d'Espagnat, 287–315. Menlo Park CA: Benjamin.

Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. Thinking about mechanisms. *Philosophy of Science* 67: 1–25.

Maudlin, Tim. 2011. *Quantum Non-locality and Relativity: Metaphysical Intimations of Modern Physics*. 3rd ed., Oxford: Wiley-Blackwell. (1st ed., Oxford: Blackwell, 1994).

Mellor, D.Hugh. 1981. *Real Time*. Cambridge: Cambridge University Press.

Näger, Paul M. 2013. Causal graphs for EPR experiments. Preprint. http://philsci-archive.pitt.edu/9915/

Näger, Paul M. 2014. *Entanglement and Causation*. Dissertation, Universiy of Bremen. (to be published as Näger forthcoming b).

Näger, Paul M. 2016. The causal problem of entanglement. *Synthese* 193 (4): 1127–1155.

Näger, Paul M. (forthcoming a). A stronger Bell argument for (some kind of) parameter dependence. *Studies in the History and Philosophy of Modern Physics*. (Preprint available at http://philsci-archive.pitt.edu/11947/).

Näger, Paul M. (forthcoming b). *Quantum Entanglement and Causation*. Boston Studies in the Philosophy of Science. Dordrecht: Springer.

Paterek, Tomasz, Alessandro Fedrizzi, Simon Gröblacher, Thomas Jennewein, Marek Zukowski, Markus Aspelmeyer, and Anton Zeilinger. 2007. Experimental test of nonlocal realistic theories without the rotational symmetry assumption. *Physical Review Letters* 99: 210406.

Pearl, Judea. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Popper, Karl R. 2002. *The Logic of Scientific Discovery*. London/ New York: Routledge. (originally published in German as *Logik der Forschung*, Wien: J. Springer, 1935; 1st English ed. London: Hutchinson, 1959).

Price, Huw. 1994. A neglected route to realism about quantum mechanics. *Mind* 103: 303–336.

Price, Huw. 2012. Does time-symmetry imply retrocausality? How the quantum world says "maybe". *Studies in History and Philosophy of Modern Physics* 43 (2): 75–83.

Redhead, Michael L.G. 1983. Relativity, causality, and the Einstein-Podolsky-Rosen paradox: Nonlocality and peaceful coexistence. In *Space, Time and Causality*, ed. R. Swinburne, 151–189. Dordrecht: Reidel.

Redhead, Michael L.G. 1987. *Incompleteness, Nonlocality, and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics*. Oxford: Clarendon Press.

Reichenbach, Hans. 1956. *The Direction of Time*. Berkeley: University of California Press.

Rowe, M.A., D. Kielpinski, V. Meyer, C.A. Sackett, W.M. Itano, C. Monroe, and D.J. Wineland. 2001. Experimental violation of a Bell's inequality with efficient detection. *Nature* 409: 791–794.

Salmon, W.C. 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.

Schrödinger, Erwin. 1935a. Die gegenwärtige Situation in der Quantenmechanik. Die Naturwissenschaften 23, 807–812, 823–828, 844–849. (English translation in *Quantum Theory and Measurement*, ed. John A. Wheeler, and Wojciech H. Zurek, 1983, 152–167. Princeton: Princeton University Press).

Schrödinger, Erwin. 1935b. Discussion of probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society* 31 (4): 555–563.

Shimony, Abner. 1984. Controllable and uncontrollable non-locality. In *Foundations of Quantum Mechanics in the Light of New Technology*, ed. S. Kamefuchi, 225–230. Tokyo: The Physical Society of Japan.

Shimony, Abner. 1989. Search for a worldview which can accommodate our knowledge of microphysics. In Cushing and McMullin 1989, 25–37.

Shimony, Abner. 1990. An exposition of Bell's theorem. In *Sixty-two Years of Uncertainty: Historical, Philosophical, and Physical Inquiries into the Foundations of Quantum Mechanics*, ed. Arthur I. Miller, 33–43. New York: Plenum.

Spirtes, Peter, Clark Glymour and Richard Scheines. 1993. *Causation, Prediction, and Search*. New York: Springer. (2nd ed., Cambridge MA: MIT Press, 2000).

Stöckler, Manfred. 1984. *Philosophische Probleme der relativistischen Quantenmechanik*. Berlin: Duncker & Humblot.

Suppes, Patrick, and Mario Zanotti. 1981. When are probabilistic explanations possible? *Synthese* 48: 191–199.

Teller, Paul. 1986. Relational holism and quantum mechanics. *British Journal for the Philosophy of Science* 37: 71–81.

Teller, Paul. 1989. Relativity, relational holism and the Bell inequalities. In Cushing and McMullin 1989, 208–223.

Tumulka, Roderich. 2006a. Collapse and relativity. In *Quantum Mechanics: 'Are There Quantum Jumps?' and 'On the Present Status of Quantum Mechanics'*, ed. A. Bassi, D. Dürr, T. Weber, and N. Zanghì, 340–352. Berlin: American Institute of Physics.

Tumulka, Roderich. 2006b. A relativistic version of the Ghirardi-Rimini-Weber model. *Journal of Statistical Physics* 125 (4): 821–840.

Ursin, Rupert, *et al*. 2007. Entanglement-based quantum communication over 144 km. *Nature Physics* 3 (7): 481–486.

van Fraassen, Bas C. 1982. Rational belief and the common cause principle. In *What? Where? When? Why? Essays on Induction, Space and Time, Explanation*, ed. R. McLaughlin, 193–209. Dordrecht: Reidel.

Walther, Philip, Kevin J. Resch, Caslav Brukner, and Anton Zeilinger. 2006. Experimental entangled entanglement. *Physical Review Letters* 97: 020501.

Weihs, Gregor, Thomas Jennewein, Christoph Simon, Harald Weinfurter, and Anton Zeilinger. 1998. Violation of Bell's inequality under strict Einstein locality conditions. *Physical Review Letters* 81 (23): 5039–5043.

Wood, Christopher J., and Robert W. Spekkens. 2015. The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning. *New Journal of Physics* 17 (3): 033002.

Chapter 5
No-Collapse Interpretations of Quantum Theory

Oliver Passon

In Sect. 2.3.1, the measurement problem was formulated in the form of a trilemma. In this view, either (i) the wavefunction is not a complete description; or (ii) the time evolution is not a continuous unitary process; or (iii) measurements do not lead to well-defined results. The GRW theory described in Sect. 2.3.1 chooses alternative (ii); it adds a nonlinear term to the Schrödinger equation, which models a physical mechanism for the "actual" collapse of the wavefunction. The Copenhagen interpretation also denies a continuous time evolution which follows the Schrödinger equation; in contrast to the GRW theory, this process is however not given a *realistic* interpretation.

In this chapter, we treat the most prominent advocates of those strategies which either deny the completeness of the wavefunction (de Broglie–Bohm theory), or question the uniqueness of the measurement results (Everett's or the many-worlds interpretation). In these theories, the state vector is thus subject to a continuous unitary time evolution. Their common feature is dispensing with the "collapse" of the wavefunction; only the *appearance* of this non-unitary change of state needs to be justified in these interpretations. Thus, the name *no-collapse interpretations* has become common as a generic label for these theories.

5.1 The de Broglie–Bohm Theory

Within the debates over the interpretation of the quantum theory—especially in view of the measurement problem—the question of whether or not quantum mechanics in its present form is simply *incomplete* is immediately raised. The statistical interpretation of quantum mechanics suggests that it must be based on an additional structure, whose elucidation would give the interpretation of the theory a completely new direction. Since this additional structure is unknown in the present version of

C. Friebe et al., *The Philosophy of Quantum Physics*,
https://doi.org/10.1007/978-3-319-78356-7_5

quantum mechanics, this research programme was originally called "the search for 'hidden' variables".

In 1952, David Bohm published his article *"A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables"* (Bohm 1952). At the time, he was unaware that Louis de Broglie had introduced a mathematically equivalent formulation of this theory already in 1927 at the 5th Solvay Conference (de Broglie 1927). For this reason, we refer to this interpretation as the "de Broglie–Bohm theory" (DBB theory).[1] de Broglie himself referred to the interpretation as the "theory of pilot waves" (*l'onde pilote*). The conference proceedings of the 5th Solvay Conference have been accessible in English only since 2009 (Bacciagaluppi and Valentini 2009). Antony Valentini and Guido Bacciagaluppi not only undertook the translation, but also, in their knowledgeable commentary, they discuss the role of this conference for the interpretation of the quantum theory in general. According to their analysis, it is misleading to reduce the significance of the 5th Solvay Conference to the (unquestionably important) debates between Bohr and Einstein. Bacciagaluppi and Valentini argue in favour of a re-evaluation of the role of de Broglie within the early interpretation debates, and in that connection, they state:

> Today, pilot-wave theory is often characterized as simply adding particle trajectories to the Schrödinger equation. An understanding of de Broglie's thought from 1923 to 1927, and of the role it played in Schrödinger's work, shows the gross inaccuracy of this characterization: after all, it was actually Schrödinger who removed the trajectories from de Broglie's theory (Bacciagaluppi and Valentini 2009, p. 78).

A discussion of the priorities in the early development of wave mechanics can and should not be carried out here. We have cited this thought-provoking passage mainly because it expresses the basic idea of the de Broglie–Bohm theory in such a simple and clear-cut manner. This is a theory which alleges the incompleteness of the usual quantum mechanics and adds "particles" in the literal sense to the wavefunction. As we have already indicated above, the term "hidden variables" has been adopted for these additional determining quantities. This term is, to be sure, somewhat misleading, since even the harshest critics cannot deny that particles, and their locations in particular, are directly observable (and thus in this sense not at all hidden). Instead, it is simply the *wavefunction* which is not susceptible to direct observation.[2]

For reasons which of course must be explained in more detail in the following, the de Broglie–Bohm theory succeeds in this way in describing the measurement process as a normal interaction which leads to a uniquely defined final state. At the same time, it is (in the technical sense) a *deterministic* theory—while in addition, it can also reproduce all the predictions of the quantum theory. However, this theory

[1] Bohm's lack of knowledge of the earlier work is understandable if one is aware that de Broglie himself did not develop his theory further, but instead became a supporter of the "conventional" quantum theory. Only after reading Bohm's publication of 1952 was his interest in these questions again aroused.

[2] We shall see that the lack of knowledge (and control) of the *precise* initial conditions plays an important role in the DBB theory. This aspect of the additional variables can indeed be considered to be "hidden". Furthermore, the concept of "hidden variables" also refers to the fact that they do not occur in the standard interpretation.

makes *no* new predictions which deviate from those of the quantum theory, so that experimentally, there is no way to decide between the two.[3]

In Bohm's formulation of 1952, we are dealing with an extension of non-relativistic quantum theory. We will take up the question of a relativistic generalization in Sect. 5.1.7. The following description of the theory makes use at various points of a comparison with the "standard interpretation" or the "usual textbook version" of quantum mechanics. These concepts are naturally not strictly defined, and the reader can think here of the Copenhagen interpretation or a textbook description of quantum mechanics, which do not deal with the problems treated in this book.

5.1.1 Mathematical Description

The de Broglie–Bohm theory is an extension of the standard quantum theory. Among the relations which define the theory mathematically, we thus find the usual Schrödinger equation:

$$i\hbar\frac{\partial\psi}{\partial t} = -\left(\frac{\hbar^2}{2m}\right)\nabla^2\psi + V(r)\psi\,. \tag{5.1}$$

Here, V refers to the potential which characterizes the corresponding system (see also Eq. (1.39) in Sect. 1.2.4; there, the Schrödinger equation was introduced for only one spatial dimension). We have chosen the positional representation not by chance, since it is, as we shall see, in fact distinguished within the de Broglie–Bohm theory. In the standard interpretation, ψ is presumed to contain the complete description of the system, and from its absolute square $|\psi|^2$, the probability of observing a particle by a measurement within a particular spatial region can be obtained. In the standard interpretation, one however cannot speak of a particle's trajectory or orbit, i.e. that which brought it to the position where it was observed.

In the de Broglie–Bohm theory, the concept of "particle" is taken so seriously that at each moment in time (i.e. even *without* measurements), it is associated with a well-defined position. A quantum-mechanical N-particle system is thus no longer described by the wavefunction alone, but rather by the *pair* consisting of the wavefunction and the position coordinates of the particles: $(\psi, Q(t))$. Here, $Q(t) = (Q_1(t), \cdots, Q_N(t))$, where $Q_i : t \to \mathbb{R}^3$ denotes the trajectory of the ith particle. $Q(t) \in \mathbb{R}^{3N}$ is called the configuration of the system, and $\mathbb{R}^{3N}$ is its so-called configuration space.[4]

[3] This statement holds strictly under the so-called quantum equilibrium hypothesis (see Sect. 5.1.2). Without that assumption, predictions which differ from the ordinary quantum theory may result (cf. Cushing 1995 and Valentini 2004).

[4] Configuration space is of central importance even in conventional quantum theory, because the wavefunction is likewise defined on this space.

For the particle positions $Q(t)$, one must specify an equation of motion, i.e. a (differential) equation which describes the temporal and spatial evolution of the particle positions under the influence of the given external conditions. This prescription must reproduce—on average—the statistical predictions of quantum theory. There have been various suggestions for the motivation of this equation of motion (cf. Passon 2010, pp. 32–36). In the following, we will make use of the analogy between quantum theory and hydrodynamics, which was pointed out as early as 1926 by Erwin Madelung (cf. Madelung 1926). Let us therefore briefly consider a liquid (or a gas) with a mass density of ρ_m. Under the assumption that the mass is a conserved quantity, the mass density within a certain region in space can then change its magnitude only if fluid flows out of or into that region. In order to describe the flow of the fluid, we define the "current-density vector" or, for short, the "current density", as the product of the mass density and the flow velocity of the fluid: $j_m = \rho_m v$. The x component of j_m denotes the amount of fluid which flows per unit time through a unit surface element (perpendicular to the x-axis) and correspondingly for the y and z components. Then the conservation of mass is represented by the following mathematical expression:

$$\underbrace{\frac{\partial \rho_m}{\partial t}}_{\text{time rate of change}} = \underbrace{-\nabla \cdot j_m}_{\text{spatial rate of change}} . \tag{5.2}$$

Here, the symbol "∇" denotes the *divergence*, i.e. the sum of the spatial changes over all three directions. This *equation of continuity* from hydrodynamics expresses—as explained—the conservation of the fluid mass.

We now turn back to quantum theory, in which likewise an equation of continuity holds—but now for the "probability density" $\rho = |\psi|^2$. This equation is formally identical[5] to the hydrodynamic equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot j = 0 . \tag{5.3}$$

At this point, naturally, the mathematical details should not be so much the subject of our considerations as the structural relations. The decisive point is that this equation can be *derived* from the Schrödinger equation, and for the probability current density, we find the following (somewhat complicated) expression:

$$j = \frac{\hbar}{2mi} \left[\psi^* (\nabla \psi) - (\nabla \psi) \psi^* \right] . \tag{5.4}$$

[5]There is, however, a decisive difference with respect to the hydrodynamic equation of continuity: While the mass density ρ_m is defined on real position space, the probability density $\rho = |\psi|^2$ is a function on configuration space. A naive identification of $|\psi|^2$ with a matter density thus appears to be impossible.

In the usual textbook descriptions of quantum theory, the equation of continuity (5.3) is interpreted as an expression of the "conservation of probability". Probability (like mass within hydrodynamics) can be neither "created" nor "destroyed".

In the de Broglie–Bohm theory, one takes a step further, since the goal is finally to arrive at an equation of motion for the "Bohmian particles". The expression ρ in quantum theory is interpreted as the probability density of the real particle configuration, and we recall that in hydrodynamics, the relation $j = \rho v$ holds. If we put in the corresponding quantum-mechanical expressions for ρ and j (and use the "polar representation" $\psi = Re^{\frac{i}{\hbar}S}$ for the wavefunction), we find, after a simple computation, the equation of motion for the particle positions $Q(t)$ that we were seeking (for its velocity, we have of course $v = \frac{dQ}{dt}$):

$$\begin{aligned} v &= \frac{j}{\rho} \\ \frac{dQ}{dt} &= \frac{\nabla S}{m} . \end{aligned} \tag{5.5}$$

This Eq. (5.5) is called the *guidance equation* of the de Broglie–Bohm theory. Pictorially speaking, the particle trajectories are thus guided by the wavefunction (or rather by its phase S). Treating a physical problem with the help of the DBB theory thus means first of all solving the Schrödinger equation (as in the usual quantum mechanics). In Sect. 5.1.4, we will discuss concrete applications.[6]

The validity of the equation of continuity (5.3) has still another important consequence for the de Broglie–Bohm theory. It follows from this equation namely that a configuration once distributed according to $|\psi|^2$ retains this property under Bohmian dynamics. This observation is the key to the fact that the de Broglie–Bohm theory reproduces all the predictions of the usual quantum theory, since naturally a differential equation fixes the motion only through its boundary and initial conditions. If one now chooses the initial configuration $Q(t_0)$ at random according to the probability distribution $|\psi_{t_0}|^2$ for a system that is described by the wavefunction ψ, then the configuration $Q(t)$ will *remain* distributed according to $|\psi_t|^2$ at each later moment in time, t. In other words, according to Born's rule, all the predictions of the usual quantum theory will be reproduced.[7] This condition is called the "quantum equilibrium hypothesis", and we will take a closer look at it in Sect. 5.1.2.

The three relations which define the de Broglie–Bohm theory mathematically are thus

1. The **Schrödinger equation**: $i\hbar\frac{\partial\psi}{\partial t} = -\left(\frac{\hbar^2}{2m}\right)\nabla^2\psi + V(\mathbf{r})\psi$
2. The **guidance equation**: $\frac{dQ}{dt} = \frac{\nabla S}{m}$

[6] In fact, the condition of being able to reproduce the statistical predictions of quantum mechanics does not fix the dynamics *uniquely*. In this sense, there are indeed infinitely many "de Broglie–Bohm-like" theories. In these theories, the individual trajectories do *not* follow Eq. (5.5), but they however reproduce the same statistics (Deotto and Ghirardi 1998).

[7] The equivalence to quantum mechanics presumes that all predictions can be uniquely described in terms of position coordinates—e.g. by "pointer positions" of a measurement apparatus.

3. The **quantum equilibrium hypothesis**: The position distribution ρ of states with the wavefunction ψ is given by the probability density $\rho = |\psi|^2$.

The second and the third relations deserve a more careful consideration, since they signal the differences relative to conventional quantum theory.

5.1.2 The Quantum Equilibrium Hypothesis

According to the quantum equilibrium hypothesis, the positions of the particles of a state which is described by the wavefunction ψ are distributed in accord with the probability density $|\psi|^2$. The occupation probability within a spatial region V is calculated by integration, $\int_V |\psi|^2 dV'$.

If this initial condition is fulfilled at one time, it follows from the equation of continuity (5.3) that Born's rule will remain valid at all later times. Furthermore, the quantum equilibrium hypothesis guarantees that the particle positions cannot be more precisely controlled. Bell writes on this topic:

> Note that the only use of probability here is, as in classical statistical mechanics, to take account of uncertainty in initial conditions (Bell 1980, p. 156).

If thus follows that the Heisenberg uncertainty relations can also not be violated within the de Broglie–Bohm theory! At the same time, one might be tempted to call the "determinism" of the de Broglie–Bohm theory "fictitious". In its descriptive content, the de Broglie–Bohm theory does not differ from the standard interpretation of quantum mechanics, and it likewise can make only statistical predictions. The quote from Bell however indicates a conceptual difference. Within the de Broglie–Bohm theory, the statistical character of the predictions is attributable to our lack of knowledge and is thus epistemic in nature. Within the standard interpretation of quantum mechanics, the ignorance interpretation of the probability is not possible; it is thus an ontic probability.

Let us now turn to the question of how this equilibrium distribution can be justified. The first attempt dates back to Bohm (cf. Bohm 1953), who gave a *dynamical* explanation of the $|\psi|^2$ distribution. His approach was developed further by Valentini (1991). In Valentini and Westman (2005), one finds for example numerical simulations of systems which, under the dynamics of the guidance equation, lead from a non-equilibrium distribution to the quantum equilibrium distribution. In the framework of this approach, it would seem reasonable to consider systems in "quantum *non*-equilibrium"—together with all possible deviations of the predictions between conventional quantum mechanics and the de Broglie–Bohm theory (cf. Valentini 2004). Another strategy—for which Bell seems to express support at various times—consists in simply *postulating* the quantum equilibrium hypothesis. This would give it the status of a fundamental law.

In contrast, Dürr *et al.* (1992) argue that neither postulating the quantum equilibrium condition, nor its dynamic justification is reasonable or convincing. At the core,

the question is namely how—within a deterministic theory—probability statements can occur at all. This problem is naturally much older than the de Broglie–Bohm theory, and it has dominated the discussion on the relation between (Newtonian) statistical mechanics and classical thermodynamics since the nineteenth century. In their justification of the quantum equilibrium distribution, Dürr *et al.* therefore hark back to a concept introduced by Ludwig Boltzmann (1844–1906), namely that of "being typical" for a physical event. "Being typical" has a terminological meaning here, namely the appropriateness for the "overwhelming majority" (as defined by measure theory) of initial configurations (Dürr 2001, pp. 49ff).

The application of this concept to de Broglie–Bohm theory is now carried out in two steps. First, the authors clarify the question of under which conditions subsystems can be associated with a wavefunction at all. This can naturally not be expected of arbitrary subsystems, owing to interactions with their environment. In principle, the de Broglie–Bohm theory thus holds for the wavefunction of the universe, Ψ. The concept of the "wavefunction of the universe" sounds presumptuous. In fact, it does *not* mean that the de Broglie–Bohm theory claims universal validity. Rather, it is the wavefunction of a system in which probability statements can no longer be explained in terms of an "external influence", i.e. by the existence of a still larger system in which the system considered is embedded. For the *fundamental* justification of probability statements, this standpoint thus *must* be adopted.

For the wavefunction of the universe, however, the assertion that its position coordinates are distributed according to $\rho = |\Psi|^2$ appears problematic. After all, there is only *one* universe,[8] and a test of this probability statement by measurements of relative frequencies of occurrence is impossible. For the wavefunction of the universe, one cannot ascribe the meaning of a probability density to the expression $|\Psi|^2$, at least not in an operational sense. Instead, Dürr *et al.* suggest that we see in it a measure of what a "typical" initial condition (in Boltzmann's sense) for the universe would be like. They justify their choice with the "equivariance" of the distribution, i.e. with the fact already mentioned that a configuration which at one moment is distributed according to $|\psi|^2$ will retain this property. The choice of any other (non-equivariant) distribution as the measure of "typical" initial configurations would have to distinguish a particular moment in time, and the moment at which precisely that distribution was present in an unnatural way.

In addition, there is a class of subsystems which can be described by using "effective wavefunctions". This means that the particle dynamics of these subsystems are almost completely determined by that effective wavefunction.[9]

---

[8] Speculations about "multiverses" change nothing in this situation—for there, also, as a rule any contact to the other "universes" is forbidden.

[9] The effective wavefunction $\psi(x)$ of a subsystem with the variables x on configuration space, which belongs to the overall system $\Psi(x, y)$, is defined as a part of the following decomposition: $\Psi(x, y) = \psi(x)\Phi(y) + \Psi^{\perp}(x, y)$. Here, Φ and $\Psi^{\perp}$ have disjunct carriers, and the configuration of the environment (Y) lies in the carrier of Φ. For the overall system, one could think for example of subsystem + environment or, concretely, subsystem + measurement apparatus. The above decomposition occurs namely during a measurement interaction: If the configuration of the measurement setup corresponds to Y (this could be a particular "pointer position" of the measurement apparatus),

Finally, Dürr *et al.* can *prove* that subsystems with an effective wavefunction ψ within a "typical" universe fulfil the quantum equilibrium hypothesis. In this sense, the deterministic de Broglie–Bohm theory obtains the appearance of randomness, and the empirical distributions correspond to the quantum-mechanical predictions. If one accepts this "Boltzmann argument", then the quantum equilibrium condition becomes even a *theorem* of the de Broglie–Bohm theory.[10]

5.1.3 The Guidance Equation

Thus far, we have considered only the single-particle case. The general form of the guidance equation for an N-particle system is given by[11]:

$$\frac{dQ_i}{dt} = \frac{\hbar}{m_i} \Im \frac{\nabla_i \psi}{\psi} = \frac{\nabla_i S}{m_i} \,. \tag{5.6}$$

Here, m_i denotes the mass of the ith particle, $\Im$ the imaginary part of the following expression and ∇_i is the gradient with respect to the spatial coordinates of the ith particle. In case the wavefunction is a spinor, i.e. $\psi : \mathbb{R}^{3N} \to \mathbb{C}^{2N}$, the probability current is changed, so that one obtains the following guidance equation:

$$\frac{dQ_i}{dt} = \frac{\hbar}{m_i} \Im \frac{\psi^* \nabla_i \psi}{\psi^* \psi} \,, \tag{5.7}$$

where $\psi^*\psi$ is the scalar product on $\mathbb{C}^2$. The latter equation is mentioned here not only for completeness, but also because it will be used in the treatment of the measurement problem in Sect. 5.1.5.

The existence and uniqueness of the solutions of the guidance equation for all the relevant types of potentials have been demonstrated (see Teufel and Tumulka 2005). Two points should be emphasized: First, the order of the guidance equation (as well as the resulting general properties of its solutions); second, its so-called non-locality. The next two subsections are devoted to these two issues.

---

the x system is guided by the wavefunction $\psi(x)$. The remaining parts of Ψ are then irrelevant for the particle dynamics, and in this way, an "effective collapse" is described (cf. Sect. 5.1.5).

[10] Our treatment here could of course only roughly sketch the train of reasoning, and it suppresses many mathematical details. Thus, an impression of circularity may have (falsely) arisen: One postulates the $|\Psi|^2$ distribution of the universe and obtains the $|\psi|^2$ distribution of subsystems. See Dürr (2001, p. 201) for more on this topic.

[11] The following is naturally difficult to understand for those readers who are not well versed in mathematics. The decisive point is that the position (and velocity) of the Bohmian particles are mathematically determined by the *wavefunction*.

General Properties of the Trajectories

Since the guidance equation is a differential equation of first order, *one* initial condition $Q(t_0)$ already determines the trajectories uniquely. In configuration space, the paths are thus not overlapping. It follows for the single-particle case, in which position space and configuration space are identical, that the trajectories within the DBB theory do not intersect each other. If they are in fact identical at one point, then they must be identical at all points. In many cases, this information alone allows us to visualize a qualitative picture of the trajectories.

Non-locality

The guidance equation determines the trajectory of the ith particle essentially by taking the derivative of the wavefunction (more precisely: by taking its gradient). The wavefunction is however defined on configuration space and is evaluated at the position $Q(t)$. In other words, the change of position of each particle at the time t depends on the positions of all the other particles at the *same moment in time*. Since these influences do not propagate through space in the sense of a short-range interaction, one speaks of a non-local influence, or of the non-locality of the de Broglie–Bohm theory. However, it is precisely this non-locality which permits the de Broglie–Bohm theory to violate the Bell inequalities (in agreement with experiments; see Chap. 4). At the same time, the quantum equilibrium hypothesis guarantees that this non-locality cannot be used for the transmission of signals, since it is in the end a question of stochastic events. The evident problem of the relativistic generalization of this theory will be addressed in Sect. 5.1.7.

5.1.4 Applications of the de Broglie–Bohm Theory

We now turn our attention to the obvious question of which form the particle trajectories take, whose existence distinguishes the de Broglie–Bohm theory from the usual quantum theory. The guidance equation has in fact been solved numerically for various problems. For those who favour this theory, the *existence* of these trajectories is notably more important than their concrete characteristics or their numerical simulation. Dürr writes on the question of whether Bohmian trajectories should be calculated at all:

> Roughly speaking, no! Sometimes, however, the asymptotic behaviour of the trajectory – essentially that of free particles – can be quite useful. [...] All that we learn from the trajectories is indeed only that at every time t, particles are present whose positions are distributed according to the quantum equilibrium hypothesis as $|\psi(q,t)|^2$ (Dürr 2001, p. 142).

In the following, we nevertheless consider explicit Bohmian trajectories for the tunnel effect, for interference from gratings (and from the double slit), as well as for the hydrogen atom; thus for several examples of quantum phenomena which, in the usual understanding, can not possibly be explained in terms of continuous trajectories.

The Tunnel Effect

A spectacular prediction of quantum theory is the "tunnel effect". It consists in the fact that quantum-mechanically described particles can overcome a potential barrier, although the energy height of the barrier is *greater* than the energy of the particles. Radioactive α decay, as well as nuclear fusion in the interior of the Sun, is understandable only in terms of the tunnel effect.[12] Pictorially speaking, the particles pass *below* the barrier—they thus "tunnel" beneath it.[13] In an orthodox manner of speaking, there is a finite probability that the particles will be detected on the other side of the barrier. Within the de Broglie–Bohm theory, a continuous particle trajectory must naturally lead from inside the potential barrier to a position outside the barrier.

Figure 5.1 shows the paths taken by some of these trajectories. The y-axis corresponds to the position coordinate (in arbitrary position units) and the x-axis to the time coordinate. A Gaussian wave packet ψ was assumed as initial condition, and it approaches the barrier from below in the figure. This potential barrier is located at $0.72 \leq y \leq 0.78$, and it is twice as high as the average energy of the wave packet.[14] Then the Schrödinger equation is solved numerically and input into the guidance equation. In this way, the course of the trajectories can be computed. One can first recognize how all the particles are braked within the barrier (the slopes of the trajectories in Fig. 5.1 correspond to the particles' velocities). The tunnel effect occurs for those particles which reach the barrier first, while those arriving later are reflected earlier and earlier. If this were not the case, the particles' trajectories could intersect. Thus, the property of being *intersection-free* determines the shape of the trajectories already qualitatively.

This description of the tunnel effect notably opens up the possibility of calculating the "tunnelling time". The obvious question of the time required by a particle in order to overcome the tunnel barrier cannot even be reasonably asked of conventional quantum theory, since time is not an observable. Cushing (1995) discusses the possibility of an experimental test of the de Broglie–Bohm theory on this basis.

Two-Slit Interference

Diffraction and interference of an electron beam by a double slit and the pattern of the typical interference fringes (see Fig. 5.2, left) were successfully observed by Claus Jönsson in 1959 (see Möllenstedt and Jönsson 1959). Particularly impressive

[12] In the case of α decay, helium nuclei overcome the potential barrier at the surface of the decaying nucleus, although their energies, considered classically, are too small to permit this. In the case of nuclear fusion in the interior of the sun, hydrogen atoms combine to form helium. Here again—considered classically—the pressure and temperature are too low to overcome the repulsion of the positively charged hydrogen nuclei.

[13] This manner of speaking, "tunnelling" or "passing beneath", is naturally to be understood as metaphorical, since the "height" of the potential barrier is not a spatial quantity, but rather an energy.

[14] The wavefunction is Gaussian, with its centre initially at 0.5 and a width (variance) of 0.05. The density of the trajectories between 0.66 and 0.68 was increased in order to be able to study the oscillatory behaviour within the barrier more precisely (see Dewdney and Hiley 1982).

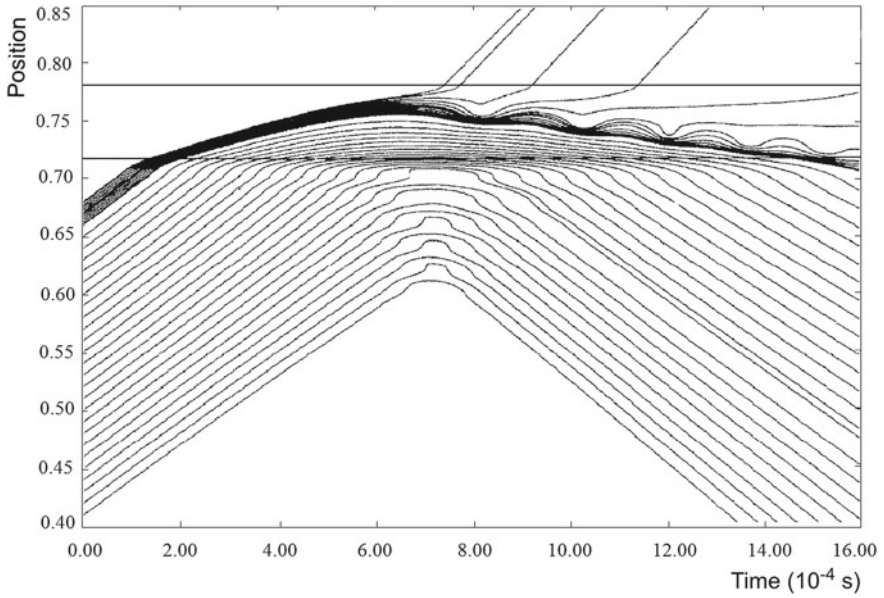


Fig. 5.1 A numerical simulation of some trajectories in the 1-dimensional tunnel effect (from Dewdney and Hiley 1982, reprinted with kind permission of the *Springer Verlag*). The x-axis corresponds to the time coordinate and the y-axis to the position coordinate

are the experiments in which the particle beam has such a low intensity that the formation of the interference pattern can be observed over a longer period of time. Then, *point-like* detected particles on the detection screen are seen to gradually build up the interference pattern.

This experiment would appear to illustrate with unusual clarity that the concept of particle trajectories is not applicable in quantum theory. If—thus runs the usual argument—the particle trajectories pass through the upper *or* the lower slit, it should be irrelevant whether at that moment the *other* slit were opened or closed. The result should be that the distribution, after passing through a double slit, should correspond to the sum of those from each of two single slits.

However, the observed pattern is evidently quite different. Popularizations occasionally claim that the particle has passed "through both slits". This paradoxical formulation is apparently intended to suggest that particle trajectories in the classical sense can no longer be considered to exist.

Within the de Broglie Bohm theory, this problem is resolved in a simple manner. The particle trajectories naturally each pass through only one of the openings of the double slit (or of the grating). The trajectories are however led by the wavefunction, according to the guidance equation. The wavefunction encodes the information on the slit geometry and steers the trajectories correspondingly towards the interference maxima. Here, it becomes clear how within the de Broglie–Bohm theory, the

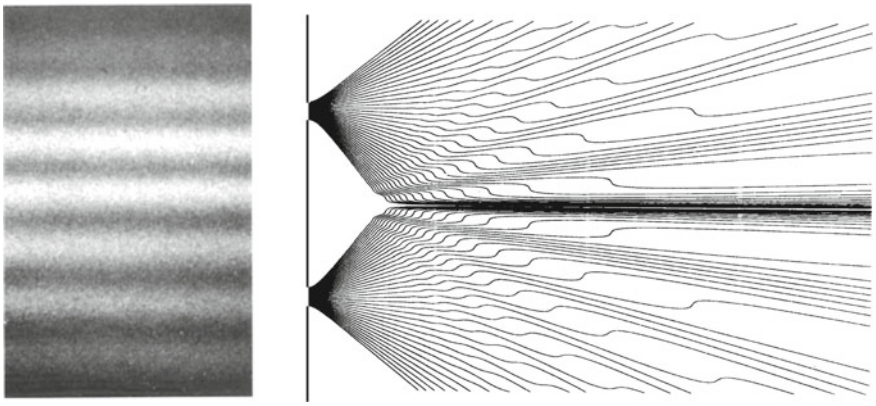


Fig. 5.2 Left: Measurement of the interference fringes of electrons from a multiple-slit arrangement (from Möllenstedt and Jönsson 1959). Right: A numerical simulation of some of the trajectories through a double slit (from Philippidis *et al.* 1979). Both of these images are reprinted with the kind permission of the *Springer Verlag*

wavefunction no longer represents a "probability wave", but instead a real physical effect.[15] Every reference to wave–particle dualism thus becomes superfluous.

If the initial values of the particle are distributed according to the quantum equilibrium hypothesis, the DBB theory exactly reproduces the occurrence frequency distribution of quantum theory. A numerical simulation of some of the corresponding trajectories can be seen on the right in Fig. 5.2. Again, it can be clearly discerned that the trajectories do not intersect. At the same time, they exhibit a completely "non-classical" behaviour, in that they show abrupt changes of direction (in—classically—"field-free" regions). Here, one can already see that momentum or energy conservation do not hold on the level of individual particles, since they obey Bohmian mechanics and not Newtonian mechanics. In Sect. 5.1.5 (see also Footnote 19), this aspect is explained in more detail.

The Hydrogen Atom

The discrete and characteristic spectra of excited atoms provided important impulses to the early development of the quantum theory. The successful description of the discrete energy levels of the hydrogen atom belongs among its early triumphs.

The solution of the Schrödinger equation for this problem (i.e. for the potential $V = -\frac{e^2}{r}$) is mathematically rather involved and will not be repeated here. The decisive point is that one is led to eigenstates of the energy for which the wavefunction is a product of a real function and the expression $e^{i(m\phi - Et/\hbar)}$. In the ground state, m (the "magnetic" or "directional" quantum number) is zero, so that the phase is given by $S = -Et$. Inserting this expression into the guidance equation (5.5), we obtain for the velocity field of course everywhere zero; we have computed the spatial

[15] On the status of the wavefunction; see however Sect. 5.1.7, and also Dürr *et al.* (1996).

derivative of an expression which has no spatial dependence. In other words, the particle in the ground state is at rest, at positions which are distributed according to the quantum equilibrium condition for the associated wavefunction. One might call this result counter-intuitive—but it must be admitted that no one has an "intuition" of the processes within an atom.[16]

5.1.5 The Solution of the Measurement Problem

At its core, the measurement problem consists in the fact that following a measurement, the measurement apparatus indeed *shows a result*. After the measurement, the apparatus should (considered quantum-theoretically) thus be in an eigenstate of the corresponding operator.

In general, the microscopic state (on which the measurement is carried out) is described as a superposition of various components, which each correspond to a different "pointer position" of the measurement apparatus. Under the dynamics of the linear Schrödinger equation, the measurement apparatus should also take on a superposition state and not an eigenstate. In reality, however, a superposition of macroscopic states is neither readily imaginable, nor has one ever been observed.[17]

The solution of the measurement problem by the de Broglie–Bohm theory can be illustrated in a completely non-technical and nevertheless appropriate way. It is based on the idea that it is only the *pair* consisting of the wavefunction *and* the configuration which makes up the complete description of a system and not just the wavefunction alone. Due to the definite particle positions, *every* system is in a well-defined state at *every time*. This therefore holds also for the measurement apparatus after a measurement has taken place. The different pointer positions of the measurement apparatus are in the end none other than different configurations, $Q(t)$. In other words, in the de Broglie–Bohm theory, the "wavefunction of the measurement apparatus" will in general be in a superposition state. The configuration however indicates the result of the measurement which is actually realized. That part of the wavefunction which "guides" the particle(s) can be reasonably termed the *effective* wavefunction. All the remaining parts can be ignored, since they are irrelevant for the particle dynamics. As a result of decoherence effects (see Sect. 5.2.4), the probability that they will produce interference effects with the effective wavefunction is vanishingly small. In

[16] In the excited states, where $m \neq 0$, the azimuthal-angle ϕ is time-dependent, and the Bohmian particle orbits around an axis (see Passon 2010, pp. 87f). Note that also this motion does not correspond to Bohr's atomic model, which is well known (and rebutted) in school physics.

[17] Formally speaking, we are considering here the superposition of several states of the overall system, consisting of a measurement object ($\psi = \sum c_i \psi_i$) and a measurement apparatus (Φ_i). In the case that the measurement apparatus is initially at the position Φ_0, during the measurement interaction it undergoes a time evolution $\psi \otimes \Phi_0 \rightarrow \sum c_i \psi_i \otimes \Phi_i$. Here, Φ_i denotes the state of the apparatus after the measurement, on measuring the property which is associated with the state ψ_i.

this sense, the de Broglie–Bohm theory describes an "effective collapse" (see also Footnote 9). In the words of Dürr:

> This 'collapse' is not a physical process, but rather an act of convenience. It takes place only through the choice of description [...] because the price of forgetting about the other, non-effective parts of the wavefunction is extremely low, since future interferences are practically excluded (Dürr 2001, p. 160).

This solution of the measurement problem makes an additional tacit assumption: All the results of measurements must be uniquely characterizable in terms of position coordinates. Think for example of the "pointer positions", or of the positions of inked pixels on a sheet of paper.[18] This however does not mean that only the measurement of particle *positions* can be described by the de Broglie–Bohm theory. Naturally, this solution of the measurement problem applies also to spin, momentum or the measurement of any other "observable". Their status is however drastically re-interpreted in this theory, as is described by the keyword "contextuality".

Contextuality

Already in Sect. 1.1.1, we dealt with the Stern–Gerlach experiment for the measurement of the spin component of an electron. A beam of silver atoms is passed through an inhomogeneous magnetic field, so that the spins of their outer electrons lead to a splitting of the beam.

Here, also, we are dealing with a measurement whose definite result is described by the de Broglie–Bohm theory. The discussion is made more complicated by the fact that the Schrödinger equation cannot describe particles with spin. Instead, one has to resort to the so-called Pauli equation. This modification of the Schrödinger equation describes spin-$\frac{1}{2}$ particles using a 2-component wavefunction. A guidance equation for the particles is found analogously to the case of the Schrödinger equation (this procedure was already described in Sect. 5.1.3, Eq. 5.7). This however yields no conceptional differences relative to the above discussion. Figure 5.3 gives a naive representation of how the results of the measurements are determined in the de Broglie–Bohm theory. If the particle coordinate is *above* the plane of symmetry (like the small black dot under the magnifying glass in the figure), a deflection into the upper branch of the wavefunction occurs ("spin up"), and *vice versa*. It is thus the particle's *location* which determines the result of a spin measurement! The property "spin" is not attributed to the particle itself, but instead, it is a property of the wavefunction.[19]

---

[18] At this point, it again becomes clear that the term "hidden variables" for the particle positions is misleading. It is just their *un-hiddenness* which qualifies them to describe the observable results of a measurement!

[19] The same is true of all other physical quantities. The particles in the de Broglie–Bohm theory have no properties besides their positions and their velocities. Even mass, momentum or charge cannot reasonably be attributed to the particles; think for example of quantum-mechanical interference experiments in which the influence of gravity or an electromagnetic interaction can (in principle) modify the wavefunction. Therefore, we have thus far avoided referring to the "Bohmian particles" as "electrons", "atoms", etc. However, in Holland (1993) as well as in Bohm and Hiley (1993), a possible spin variable is discussed. Our treatment here follows Bell (2001, pp. 5ff) and Dürr (2001).

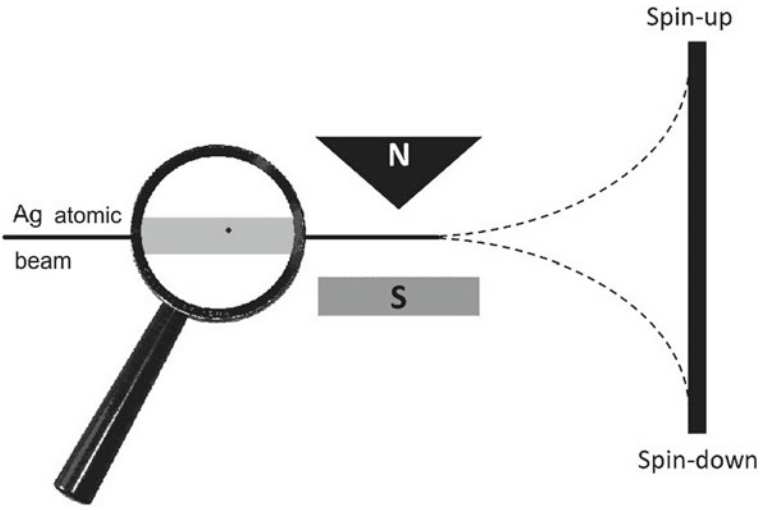


Fig. 5.3 Deflection of silver atoms in an atomic beam by an inhomogeneous magnetic field (Stern–Gerlach experiment). The initial position (indicated by the small dot under the magnifying glass) is decisive for the result of the spin (-component) measurement in the de Broglie–Bohm theory

We could argue in a similar way about the measurement of energy, momentum or other observables. For all of these quantities, the de Broglie–Bohm theory thus introduces *no* additional "hidden variables" which would describe their actual values. Instead, their values are determined by the wavefunction, the particle position *and* the particular implementation of the measurement. Taking the example of the Stern–Gerlach experiment, we can illustrate the influence of the particular measurement setup in an intuitively clear manner: If the orientation of the magnetic field in Fig. 5.3 were reversed, we would measure the *opposite* spin for the *same* particles! The de Broglie–Bohm theory thereby composes a picture in which only position measurements yield a value that was already present within the system *before* the measurement and is thus a *property* of the particle in the narrow sense. All other measurements owe their outcomes to the "context" of the implementation of the measurement. The terms "measurement" and "observable" are rather misleading here. This property of the de Broglie–Bohm theory is called "contextuality". Indeed, this concept has a somewhat more extended meaning in the discussions and includes those mutual influences which occur in *combined* measurements of different quantities.

The relations treated here can be formulated rather concisely by making use of the terminology which has been developed in philosophy for the description of various types of properties. The spin, or all other properties aside from the position, are not categorial properties within the de Broglie–Bohm theory, but rather they are

dispositions.[20] The contextuality of dispositions is however not remarkable; it is simply a part of their definition (cf. Pagonis and Clifton 1995).

Proofs of the Impossibility of Hidden Variables

This contextuality of the de Broglie–Bohm theory explains why the numerous *"no-go" theorems* or "proofs of impossibility" relating to theories with hidden variables do not apply to it. The best known of these theorems is due to von Neumann. A generalization was formulated by Kochen and Specker; see Mermin (1990) and the references there. These theorems are based on the intuition that hidden variables fully encode the (only apparently) statistical outcome of the measurements. The proofs demonstrate the impossibility of a mapping which ascribes to every state a unique value in regard to *every* possible measurement—indeed, without taking the context into account. The de Broglie–Bohm theory does not claim even the *existence* of actual values in regard to every physical quantity which can be measured; for only the position is a categorial property of this theory. Think again of the example given above of the measurement of the spin component: The particle is not associated with any fixed orientation of its spin, independently of a concrete "measurement"; when the direction of the magnetic field is changed, the spin can even take on the reversed value. According to Daumer *et al.* (1996), the understanding of measurements which is based for example on such *no-go* theorems reveals a "naive realism" in relation to the role of operators. These authors understand this as the usual identification between operators and observables, bound up with the widespread manner of speaking that "operators can be measured". This expression is however highly misleading, since the influence of the experimental context on a measurement as described above is not taken into account.

5.1.6 The Schools of the de Broglie–Bohm Theory

The *Compendium of Quantum Physics* (Greenberger *et al.* 2009) contains two entries on the subject of this chapter. One of them is entitled the "*Bohm Interpretation of Quantum Mechanics*", while the other is called simply "*Bohmian Mechanics*". One is left with the suspicion that "Bohmian mechanics" is not identical to "Bohm's interpretation of quantum theory". This impression is correct and deserves a closer look—even if only to facilitate the orientation of the reader in studying the relevant literature.

The article on "*Bohmian Mechanics*" was written by Detlef Dürr, Sheldon Goldstein, Roderich Tumulka and Nino Zanghì. This group supports a version of the theory which was formulated by John Bell, beginning in the 1960s. Our own treatment is closely related to this version. At its centre stand the guidance equation and the re-interpretation of the concept of *observables* (keyword: "contextuality").

[20] While categorial properties are associated with an object *without* any reference to its environment (e.g. "being round"), dispositions describe those properties which manifest themselves only in certain special contexts (e.g. "being fragile").

The author of the article "Bohm Interpretation of Quantum Mechanics" is Basil Hiley. He was a colleague and close coworker of David Bohm at Birkbeck College; and together with Chris Dewdney, Chris Philippidis and others, this "English group" strongly supports Bohm's formulation of the theory from 1952. Bohm and his coworkers referred (or refer) to this theory notably as an "ontological" or a "causal" interpretation of the quantum theory. In this variant of the theory, the concept of the "quantum potential" plays a central role. Let us consider the derivation of the guidance equation in this respect, as it was given by David Bohm in 1952. He chose a path for the derivation which invokes an analogy to the Hamilton–Jacobi equation of classical mechanics. In the classical case, the Hamilton–Jacobi theory contains the relation $p = \nabla S$ (with the momentum $p = mv$ and the action S). Bohm could show in his work that a similar relation holds in quantum theory, if the action is replaced by the phase S of the wavefunction. This then led him to the well-known guidance equation, $v = \frac{dQ}{dt} = \frac{\nabla S}{m}$. Indeed, this theory can then be represented in such a way that it appears to be a modification of Newtonian mechanics:

$$m\frac{\mathrm{d}^2 Q(t)}{\mathrm{d}t^2} = -\nabla(V + U_{\text{quant}}) \tag{5.8}$$

with the classical potential V and the additional quantum potential

$$U_{\text{quant}} = -\frac{\hbar^2\nabla^2|\psi|}{2m|\psi|}\,. \tag{5.9}$$

Note, however, that in contrast to Newtonian physics, the velocity is already fixed via the guidance equation. The representation in terms of a second-order differential equation is thus misleading, since it suggests that position and momentum may be chosen independently.

In fact, the quantum potential has completely non-classical properties, which allow the adherents of this "causal viewpoint" to justify the uniqueness of quantum phenomena. They find for example that wavefunctions which differ only through a complex factor lead to the same quantum potential, since in U_{quant}, the wavefunction enters both into the denominator and the numerator. Here, Bohm and Hiley (1993, p. 31) introduce the concept of "active information", and they find in it the justification for a new kind of "holism" (see also Hiley 1999).

Although these two readings of the de Broglie–Bohm theory are *mathematically* equivalent, and the real contradiction between the usual quantum theory and these variants of the DBB theory holds for both of them, the rivalry of these schools is considerable. Hiley writes:

> It should be noted that the views expressed in our book (Bohm and Hiley 1993) differ very substantially from those of Dürr *et al.* (1992), who have developed an alternative theory. It was very unfortunately that they chose the term 'Bohmian mechanics' to describe their work. When Bohm first saw the term he remarked, 'Why do they call it 'Bohmian mechanics'? Have they not understood a thing that I have written'? He was referring [...] to a footnote in his book *Quantum Theory*, in which he writes, 'This means that the term 'quantum mechanics' is a misnomer. It should, perhaps, be better called quantum nonmechanics.' It would have

> been far better if Dürr *et al.* had chosen the term 'Bell mechanics'. That would have reflected the actual situation far more accurately. (Hiley 1999, p. 119)

The acrimony in this dispute is essentially due to the fact that the "ontological interpretation" associates far-reaching natural-philosophical speculations to its concept of the quantum potential, while supporters of "Bohmian mechanics" see the strength of the theory in its being able to *eliminate* philosophical speculations from the formulation of the theory. Characteristically, the title of an article by Dürr and Lazarovici (2012) is "Quantum physics without quantum philosophy".

5.1.7 Criticism of the de Broglie–Bohm Theory

John Bell, who, beginning in the 1960s contributed to the popularization of the de Broglie–Bohm theory with a number of articles, writes concerning the topic of this section:

> It is easy to find good reasons for disliking the de Broglie–Bohm picture. Neither de Broglie nor Bohm liked it very much; for both of them, it was only a point of departure. Einstein also did not like it very much. He found it 'too cheap', although, as Born remarked, 'it was quite in line with his own ideas'. But like it or lump it, it is perfectly conclusive as a counter example to the idea that vagueness, subjectivity, or indeterminism are forced on us by the experimental facts covered by nonrelativistic quantum mechanics (Bell 2001, p. 152).

According to Bell, all the counter arguments cannot reduce the importance of the theory in principle. Nevertheless, in the following we will take a closer look at some of those arguments. Heisenberg expresses the opinion that the identical descriptive content of the theory (relative to standard quantum mechanics) disqualifies it as an independent theory. He writes (Heisenberg 1959, p. 106):

> From a strictly positivistic point of view, one could even say that we are dealing here not with an alternative suggestion to the Copenhagen interpretation, but instead with an exact repetition of it, only with different terminology.

In the face of the conceptional differences between the de Broglie–Bohm theory and the usual quantum theory, this statement would seem to be overly strong. Furthermore, Heisenberg naturally presumes here that the Copenhagen interpretation offers a convincing solution to the measurement problem. Closely related are references to "Ockham's razor". According to the prevailing opinion, when two theories are equivalent, the one which requires the lesser number of premises should be preferred. Does Ockham's razor thus "cut off" the guidance equation as superfluous ballast from a theory which offers no additional predictions? This demand fails to notice that the additional equation in the de Broglie–Bohm theory renders unnecessary all of the postulates concerning the outcome of a measurement and the interpretation of the wavefunction.

While the previous arguments take the considerable *similarity* of the DBB theory to the quantum theory as an object of criticism, others see the reason to repudiate the

DBB theory in their radical *dissimilarity*. They find fault with the explicit distinction of the position, and the lack of symmetry between, e.g., the position and the momentum spaces (see the objection of Pauli in Myrvold 2003). In the DBB theory, with its first-order equation of motion, momentum and energy at the level of individual particles are however no longer conserved quantities. The justification of the demand for symmetry between position and momentum can thus be reasonably questioned.

Still other critics are bothered by the (double) role of the wavefunction: It fixes the particle dynamics and is at the same time (i.e. its absolute square) a measure for the equilibrium distribution. In addition, it acts upon the particle's motion *without* any back-reaction effects. Another point of criticism refers to the fact that according to the de Broglie–Bohm theory, the world is populated by innumerable "empty" wavefunctions. This is indeed somewhat inelegant.

The role or the status of the wavefunction is also the object of a discussion among those scientists who work with the DBB theory. Originally, the wavefunction was taken to represent a real, physical field. Dürr *et al.* (1996) suggest, in contrast, that it should play a "nomological" role (i.e. with a law-like character). The wavefunction would then more closely correspond to, e.g., the Hamilton function in classical mechanics than to the usual type of physical field. This could reduce the weight of the criticisms of the lack of reaction effects and of the "empty" wavefunctions. While the interested reader can find a more detailed discussion of the criticisms of the DBB theory in Passon (Passon 2010, pp. 117ff), we will now turn to the principal objection against it: The question of the possibility of a *relativistic generalization* of the theory.

The particle dynamics of the de Broglie–Bohm theory connects positions at arbitrary distances. This non-locality would appear to violate Einstein's postulate of the speed of light as an upper limiting velocity. However, the DBB theory discussed so far is an extension of non-relativistic quantum mechanics. The allusion to the fact that it is not compatible with the requirements of the special theory of relativity is thus not really a criticism, but rather simply a statement of fact. This rejoinder is however too superficial, since it is indeed just the non-local dynamics which allows the de Broglie–Bohm theory to explain the violation of the Bell inequalities (cf. Chap. 4, and there in particular Sect. 4.4).

As a rule, the criticism of the non-locality of the DBB theory is primarily associated with doubts as to whether or not the theory can be relativistically *generalized*. At the same time, there is an orthodox relativistic quantum theory (the Dirac theory) and a relativistic quantum field theory (see Chap. 6), so that the final (and negative) judgment about the DBB theory seems to be passed. However, this argument would be significantly more convincing if those (orthodox) relativistic theories had *no* measurement problem. But there, also, e.g. the question of definite measurement outcomes is controversially debated.

Thus, the development of a "Bohm-like" relativistic quantum theory (and quantum field theory) without the foundational problems of the conventional formulation is part of the current research program of scientists who work in this field. Some of the approaches discussed there apply notably not *particle ontologies*, but instead *field*

ontologies. Furthermore, some of the relativistic generalizations dispense even with a deterministic description.[21]

It turns out that not only the dynamics of a generalized guidance equation pose a problem for a relativistic generalization, but also the (generalized) quantum equilibrium distribution. This requirement distinguishes a frame of reference, namely that in which the distribution is determined. The equivalence of all inertial systems is, however, at the very core of special relativity, according to the usual understanding. There are some approaches in which the "distinguished" frame of reference has no experimentally accessible influence and which can reproduce all the predictions of relativistic quantum theory. A new evaluation of the relationship between quantum theory and relativity is however certainly bound up in such approaches. The supporters of the DBB theory recall in this connection quite rightly that (as mentioned above) it is precisely this non-locality, as expressed in the violation of the Bell inequalities, which is also a part of conventional quantum mechanics and quantum field theory. Therefore, from the viewpoint of many adherents of the DBB theory, the conventional interpretations of quantum mechanics and quantum field theory likewise have this same problem, but they mask it by their vague formulations (cf. also Bricmont 2016, pp. 169–173) for more on this topic).

We shall now leave the de Broglie–Bohm theory for a while and turn to Everett's work, another controversially discussed interpretation of quantum theory. In Sect. 5.3, we will then come back to the DBB theory within the framework of a comparison between various interpretation approaches.

5.2 Everett's Interpretation

In 1957, the American physicist Hugh Everett III (1930–1982) published his "*relative state*" formulation of quantum mechanics (see Everett 1957). It contains the results of his doctoral thesis, mentored by John A. Wheeler at Princeton University. Its goal was a re-formulation of the theory, in which the discontinuous change of state ("collapse") would be superfluous, and instead, a unitary time evolution of the wavefunction would hold throughout. In contrast to the de Broglie–Bohm theory, however, the *completeness* of the quantum-mechanical description is asserted, and thus the third statement of Maudlin's trilemma (Sect. 2.3.1) is denied: Measurements *appear* to have only one definite outcome in Everett's approach, although in fact the wavefunction (with its superposition states) contains a complete description.

Everett's guiding concept was to derive the interpretation from the mathematical formalism.[22] He was motivated explicitly by the measurement problem, or by the distinction of an *external* observer in the usual formulation:

[21] The existing approaches and attempts to find a Lorentz-invariant generalization of the DBB theory are discussed in Passon (2006) and in Tumulka (2007).

[22] Everett himself writes of his methodology: "The wavefunction is taken as the basic physical entity with no *a priori* interpretation. Interpretation comes only *after* an investigation of the logical

> No way is evident to apply the conventional formulation of quantum mechanics to a system that is not subject to *external* observation. The whole interpretive scheme of that formalism rests upon the notion of external observation (Everett 1957, p. 455).

But at the latest when considering cosmological problems, the standpoint of an *external* observer can no longer be reasonably assumed, and the applicability of the quantum theory would appear to be frustrated by that fact.

Along with the justification for how—in view of the superposition states—the *appearance* of definite measurement outcomes is produced, Everett must furthermore explain how and why the statistics of those measurement results follows Born's rule (i.e. $|c_i|^2$ corresponds to the probability of occurrence of the given outcome). Now, Everett's work has itself become the object of an interpretation debate. Jeffrey Barrett writes on this subject:

> The fact that most no-collapse theories have at one time or another been attributed to Everett shows how much the no-collapse tradition owes to him, but it also shows how hard it is to say what he actually had in mind (Barrett 1999, pp. 90f).

In the following, we will trace roughly how the open technical and conceptional questions relating to the 1957 article have led to the development of variants and modifications. We begin however with a description of Everett's basic idea.

5.2.1 The Basic Idea

Everett's re-interpretation of the measurement problem is just as surprising as it is brilliant. This problem results, as is well known, from the application of the quantum theory to the measurement process. In general, a superposition state results from this process, consisting of, e.g., *different* "pointer states", while our experience tells us that measurements lead to *unique* results. Superposition states appear under these circumstances not to give an appropriate description of the physical situation. The drastic consequences which result ("either the Schrödinger equation is false or it is not complete" (Bell 2001, p. 173)) are avoided by Everett with the aid of the following consideration: Under the premise that the quantum theory is applicable also to the observation process, the observer must therefore also enter into a "superposition state"—and *this* superposition undermines the reliability of the judgement that caused us to doubt the appropriateness of superposition states as a description in the first place! Instead, Everett suggests that we identify *every* term of the superposition with an (equally weighted) state of the observer.[23] The evolution of measurements (or observations) can then be described as follows:

---

structure of the theory. Here as always the theory itself sets the framework for its interpretation" (Everett 1957, p. 455).

[23] He models the "observer" by a physical system, in the concrete case a machine which has access to sensors and storage media.

> We thus arrive at the following picture: Throughout all of a sequence of observation processes there is only one physical system representing the observer, yet there is no single unique *state* of the observer (...). Nevertheless, there is a representation in terms of a *superposition*, each element of which contains a definite observer state and a corresponding system state. Thus 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 given sequence of observations (Everett 1957, p. 459).

In which sense Everett can still consider just *one* observer ("one physical system representing the observer"), who is simultaneously in the *multiplicity* of states as described, is initially unclear. The various different answers to this question lead essentially to the different variants of the Everett interpretation which were mentioned in the above quote from Barrett.

5.2.2 The Many-Worlds Interpretation

Bryce DeWitt and Neil Graham (1973) popularized the Everett theory through their anthology "*The Many-Worlds Interpretation of Quantum Mechanics*" and coined the catchy name with their choice for its title. They interpret the branching of the wavefunction mentioned in the Everett quote above in a completely realistic manner, as a real splitting into different "worlds", and write[24]:

> The universe is constantly splitting into a stupendous number of branches, all resulting from the measurement-like interactions between its myriads of components. Moreover, any quantum transition taking place on every star, in every galaxy, in every remote corner of the universe is splitting our local world on earth into myriads of copies of itself (DeWitt and Graham 1973, p. 161).

"World" means here the totality of all the (macroscopic) objects, and the human observer likewise is subject to this splitting into a manifold of "copies".

David Wallace (Wallace 2010, p. 4) illustrates this astounding idea by means of an analogy with classical electrodynamics. Imagine an electromagnetic configuration $F_1(r, t)$ which describes a pulse of light that is propagating from the Earth to the Moon. A second configuration $F_2(r, t)$ could describe a light pulse underway from Venus to Mars. How, asks Wallace, should one now interpret the configuration

[24] It is very questionable as to what extent this suggestion corresponds to Everett's own understanding of the theory. Since Everett worked in the strategic planning department of the Pentagon after finishing his doctorate, and published no more work on quantum theory, this question can be answered only by consulting his sporadic correspondence and papers from his estate. These sources give the impression that Everett *did not* have a splitting up into different "worlds" in mind, whose definition would seem to make a connection with classical concepts. In some respects, the current version of the many-worlds interpretation, which we will discuss more detail in the following sections, appears to be more similar to Everett's original conception. However, he did not categorically reject the language of DeWitt—especially since he was very grateful to the latter for the popularization of his ideas. See Barrett (2011), and the essay by Peter Byrne in Saunders *et al.* (2010), for more on this subject.

$$F(r,t) = \frac{1}{2} \cdot F_1(r,t) + \frac{1}{2} \cdot F_2(r,t) \ ? \tag{5.10}$$

Does it describe a light pulse which is moving *simultaneously* between the Earth and the Moon as well as between Venus and Mars, since it occurs as a superposition? This is of course nonsense; instead, Eq. (5.10) does not describe a "strange" light pulse in a superposition state, but rather two "ordinary" light pulses at different locations. Wallace continues:

> And this, in a nutshell, is what the Everett interpretation claims about macroscopic quantum superpositions: they are just states of the world in which more than one macroscopically definite thing is happening at once. Macroscopic superpositions do not describe indefiniteness, they describe multiplicity (Wallace 2010, p. 5).

Here, however, we are not dealing with a spatial separation (as in the example from electrodynamics), but instead—as Wallace expresses it—with a *dynamic* separation. This means that the parallel worlds have no mutual interactions, i.e. described pictorially, they are "transparent" to one another. The innumerable "worlds" are located unperturbed in the same, single spacetime region.[25]

The interpretation of Everett's construction given by DeWitt and Graham has entered into the popular-scientific literature and has since ignited the fantasy of (not only) laypersons interested in physics and science fiction authors. In an obvious sense, the measurement problem is resolved by this construction, since in each "world", an eigenstate of the measurement apparatus in fact exists. Whether this condition suffices for a complete resolution of the measurement problem is however questioned by Maudlin (2010). In Sect. 5.2.6, we will discuss this criticism of Everett's interpretation. The situation regarding the question of non-locality is similar: While Bacciagaluppi (2002) supports the view that the violation of Bell's inequality (see Chap. 4) can be explained here *without* action at-a-distance, Allori *et al.* (2011) argue that the many-worlds interpretation produces this appearance of locality only because of its imprecise formulation. In Allori *et al.* (2011), a modification of the many-worlds interpretation is suggested, which likewise contains action at-a-distance (cf. Sect. 5.2.6).

In the version that we have thus far sketched, the theory however does not appear to be complete. Leslie Ballentine has pointed out that the meaning of probability statements within the Everett interpretation is unclear. Finally, all possible events do actually occur (see Ballentine 1971, pp. 233–235). Furthermore, the branching is subject to an ambiguity with respect to the choice of basis. This problem of the "preferred" basis will be discussed first, in the following section.

[25] This spacetime is subject to splitting only when the many-worlds idea is applied to theories of quantum gravitation.

5.2.3 The Problem of the Preferred Basis

Let us consider a typical example of the superposition of various spin states (e.g. those of a silver atom): $|\Psi\rangle = \frac{1}{\sqrt{2}}\,(|\uparrow_x\rangle + |\downarrow_x\rangle)$. If one wishes to determine the orientation of the spin along the x direction, one would investigate this state using a correspondingly oriented Stern–Gerlach magnet. At the end of the measurement, the state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_x\rangle|M_{\uparrow_x}\rangle + |\downarrow_x\rangle|M_{\downarrow_x}\rangle\right) \tag{5.11}$$

is present. This state thus describes—according to the many-worlds interpretation—two "worlds", in which the x component of the spin is either $\uparrow_x$ or $\downarrow_x$. The decomposition into basis vectors is however in general not unique and could be just as well carried out with eigenvectors with respect to some other measurable quantity. For example, the following linear combination could be considered[26]:

$$|\uparrow_z\rangle = \frac{1}{\sqrt{2}}(|\uparrow_x\rangle + |\downarrow_x\rangle) \qquad |\downarrow_z\rangle = \frac{1}{\sqrt{2}}(|\uparrow_x\rangle - |\downarrow_x\rangle)$$
$$|M_{\uparrow_z}\rangle = \frac{1}{\sqrt{2}}(|M_{\uparrow_x}\rangle + |M_{\downarrow_x}\rangle) \qquad |M_{\downarrow_z}\rangle = \frac{1}{\sqrt{2}}(|M_{\uparrow_x}\rangle - |M_{\downarrow_x}\rangle)\,.$$

With respect to this basis, the state (5.11) now has the following representation:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_z\rangle|M_{\uparrow_z}\rangle + |\downarrow_z\rangle|M_{\downarrow_z}\rangle\right)\,. \tag{5.12}$$

If the two "worlds" branch in terms of *these* basis vectors, the spin along the x direction would not have a well-defined value, and instead, its z component[27] would be well defined. The choice of a basis within the quantum theory is to be sure purely conventional and should have no physical relevance. A factual difference between the representations in (5.11) and (5.12) must therefore be separately justified. In other words, the choice of a "preferred basis" is necessary. One might object at this point that the choice of a specific measurement setup leads to precisely such a distinction of the pointer basis (5.11). In the other basis (5.12), in contrast, in every term there is a superposition of the various states of the x measurement apparatus. The non-occurrence (or rather the non-observability) of superpositions of macroscopically different states was however just what we were trying to explain with the Everett interpretation—it should thus not be a *precondition* of the investigation. Furthermore, such a distinction of a particular basis for the measurement process contradicts the

[26] The ambiguity of the representation is the subject of the "biorthogonal decomposition theorem" (cf. Bub 1997, p. 151). The decomposition is unique if and only if all the components have different and nonzero coefficients.

[27] Note that there are no *common* eigenvectors of σ_x and σ_z.

spirit of an interpretation which merely wishes to let the formalism remain valid and in which the observation plays no fundamental role.[28]

In today's view, the suggestions for solving this problem fall into two classes: The older ones, which do not refer to decoherence, and those which make use of the mechanism of decoherence. A brief treatment of those Everett variants which are currently considered to be obsolete since the advent of the approaches based on decoherence is desirable with a view to the discussion of the concept of probability (Sect. 5.2.5). We will therefore first cast a brief glance at these older approaches before considering the role of decoherence theory in Sect. 5.2.4.

David Deutsch's Variant of Everett's Interpretation

David Deutsch, in his early works, suggested a mechanism for distinguishing a basis (cf. Deutsch 1985).[29] He extends the quantum-theoretical formalism in terms of an algorithm which produces the corresponding basis. This depends (without going into the details here) only on the corresponding physical state and its dynamics. The choice is limited by the requirement that in the case of "measurements", the relevant basis in fact corresponds to the "pointer basis". This guarantees that after a measurement, a unique result is in fact obtained (Deutsch 1985, pp. 22f).

Wallace (2010, p. 7) calls this variant of the Everett interpretation the "many-exact-worlds" interpretation. In Sect. 5.2.4, we will see that in the meantime, there are more promising candidates for the solution of the problem of the preferred basis, and David Deutsch himself has also rejected this interpretation since the end of the 1990s. First, however, we will consider yet another variant.

The Many-Minds Interpretation

The many-worlds interpretation includes the act of observation within the physical description. This apparently presumes that mental states are also a part of the physical world and are subject to the laws of quantum theory.[30]

In this sense, a many-worlds interpretation *appears* to always imply a theory of branching consciousness states (the exception will be discussed below). This evident significance is however *not* meant when one speaks of the *many-minds* interpretation.

---

[28]The problem treated here thus occurs in other interpretations of quantum mechanics as well, and it shows that the measurement problem actually consists of two sub-problems: (i) The problem of the preferred basis and (ii) the problem of the definite outcome of a measurement. Within, e.g., the Copenhagen interpretation, however, (i) can be resolved by specifying the measurement setup (choice of direction).

[29]Deutsch mentions here (on p. 2) that he has taken up an idea of Everett's, based on private conversations with him.

[30]This position is called "physicalism". Physicalism (expressed in simplified form) asserts the metaphysical hypothesis that everything which exists is *physical*. It can be understood as a further development of materialism. In particular, it rejects any kind of dualism between physical and mental ("mind") states. The relation between physical and mental states is not necessarily an identity, however. In the philosophy of the mind, the viewpoint is widespread that these two property areas are connected through a "supervenience relation". The supervenience of A over B is understood to mean that (in "slogan" form) "no change in A is possible without a change in B". This also permits speculations on a possibly non-reductionistic physicalism.

A prominent suggestion of this variant is due to Albert and Loewer (1988). They were motivated by the problem of the preferred basis, as well as the difficulty of understanding the significance of probability statements within the many-worlds interpretation (this problem will be treated in more detail in Sect. 5.2.5).

The point of departure of the *many-minds* interpretation is the statement that mental states can never be in superpositions, according to our introspective experience. Loewer and Albert reason from this that mental states (i.e. beliefs, intentions, memories, etc.) are *not* physical.[31]

They then postulate that every observer is outfitted with an infinite number of "*minds*". While in the case of a measurement or an interaction, the physical brain states take on a superposition state, a probabilistic time evolution leads to a state in which a certain portion of these *minds* corresponds in each case to the perception of *one single* outcome for the experiment. This process takes place within *one* world.

Now, how does this interpretation deal with the problem of the "preferred basis"? In an evident sense, the choice of basis vectors for the evolution of a state has no physical significance, since in the *many-minds* interpretation, there is only *one* world. An ambiguity with respect to the splitting into "many worlds" thus cannot arise here. However, Barrett (1999, p. 195) has pointed out that the "basis" of the consciousness states plays a comparable role.[32]

Both in the *many-minds* interpretation and also in the interpretation of Deutsch (1985), a preferred basis must thus be postulated. This common strategy is accompanied by a common difficulty: All attempts to introduce a preferred basis *ad hoc* must *postulate* properties which should in fact be *explained* in a fundamental theory (cf. Wallace 2010, p. 8). In the next section, we will treat the theory of decoherence. With it, one associates the hope that a convincing solution of the problem of the preferred basis can be found, since it does without such *ad hoc* assumptions.

5.2.4 The Role of Decoherence Theory

As a rule, physics investigates "isolated systems", i.e. it considers the influence of the "environment" to be a negligible perturbation and, above all, an unnecessary complication. We now find that within the quantum theory, precisely the *inclusion* of the interactions with the environment can lead to conceptional progress in describing measurements, as well as the classical limits of the theory. The research which has been accomplished since the early 1970s in this field has not been associated with any particular interpretation of quantum theory and makes use simply of the mathematical properties of the standard formalism. Pioneers in this field of

[31] One may consider the astute self-observation on which this conclusion is based not to be a particularly powerful tool for philosophical reflection. But for questions involving our conscious minds, it is however our *only* tool!

[32] Just how Barrett means this for a non-physicalistic conception of the mind remains unclear.

"decoherence"[33] were Zeh (1970) and Zurek (1981). Already in Sect. 2.3, we have discussed the decoherence programme. We make use here of the concepts introduced there, amplify them, and codify the results within the context of the Everett interpretation.

In Sect. 5.2.3, we have already explained how the decomposition of a state into its basis vectors can be ambiguous. The decompositions (5.11) and (5.12) are mathematically equivalent—their physical differences must therefore be justified.

The first step towards the resolution of this problem can now be accomplished through a purely mathematical consideration: If we look at the entanglement with a *third* system E (the *environment*, in our example likewise represented by a two-dimensional Hilbert space with states $|e_i\rangle$), we will be led to a state of the form

$$|\Psi\rangle = \frac{1}{\sqrt{2}}|\uparrow_x\rangle|M_{\uparrow_x}\rangle|e_{\uparrow_x}\rangle + \frac{1}{\sqrt{2}}|\downarrow_x\rangle|M_{\downarrow_x}\rangle|e_{\downarrow_x}\rangle\,. \tag{5.13}$$

Andrew Elby and Jeffrey Bub (1994) were able to show that this decomposition into orthogonal states on a *triple* product space is *unique*.[34] It thus eliminates the ambiguity in the choice of a basis, in a formal sense (and also that of the associated physical measurand). Naturally, this purely mathematical argument as yet yields no indication of *which* basis is to be distinguished—especially since the extremely detailed states of the environment are unobservable. In this situation, physical criteria for the identification of this unique ("preferred") basis must still be developed, as Schlosshauer mentions:

> The decoherence programme has attempted to define such a criterion based on the interaction with the environment and the idea of robustness and preservation of correlations. The environment thus plays a double role in suggesting a solution to the preferred-basis problem: it selects a preferred pointer basis, and it guarantees its uniqueness via the tridecompositional uniqueness theorem (Schlosshauer 2005, p. 1279).

These criteria were thus not *postulated*, but rather they follow from the quantum-theoretical investigation of the dynamic influence of the environment. For this purpose, one treats complicated models of the environment. The interactions between it and the measurement apparatus take place as a rule via force laws which contain powers of the spatial distance (e.g. the Coulomb force $\propto r^{-2}$). It follows that the unique decomposition as a rule distinguishes the basis of *positional space*, and in the case of a measurement, the "pointer basis" is the relevant basis. Schlosshauer summarizes this approach, called *environment-induced superselection*, as follows:

---

[33] The adjective "coherent" in the general vocabulary means "connected". The physical-terminological expression "coherent" is usually applied to optics, and it describes, roughly speaking, the precondition that must be fulfilled by different wavetrains in order that they may be able to interfere with each other. Expressed non-technically, the decoherence program thus attempts to clarify the conditions and preconditions under which quantum states lose this "non-classical" property.

[34] This *tridecompositional uniqueness theorem* is valid under rather general conditions. The existence of the decomposition is by the way not guaranteed. The proof of this theorem can also be found in (Bub 1997, Sect. 5.5).

> The clear merit of the approach of environment-induced superselection lies in the fact that the preferred basis is not chosen in an *ad hoc* manner simply to make our measurement records determinate or to match our experience of which physical quantities are usually perceived as determinate (for example, position). Instead the selection is motivated on physical, observer-free grounds, that is, through the system-environment interaction Hamiltonian. The vast space of possible quantum-mechanical superpositions is reduced so much because the laws governing physical interactions depend only on a few physical quantities (position, momentum, charge, and the like), and the fact that precisely these are the properties that appear determinate to us is explained by the dependence of the preferred basis on the form of the interaction. The appearance of *classicality* is therefore grounded in the structure of the physical laws – certainly a highly satisfying and reasonable approach (Schlosshauer 2005, pp. 14f).

This quote once again emphasizes that the results of decoherence are not tied to any particular interpretation of the quantum theory, i.e. that they can be applied within every interpretation.[35]

Since the interaction with the environment is described quantum-mechanically (i.e. via a unitary time evolution), the combination of the [object + measurement apparatus + environment] remains in a so-called pure state. This overall state will in general contain both a superposition of various "pointer positions" and also interference terms. The exact state of the environment is not only not susceptible to influences, but as a rule also not to observation. If one computes the predictions for the real observables in the subsystem [object + measurement apparatus], one obtains a result in which there are *practically* no more interference terms.[36] This part of the programme is termed the *environment-induced decoherence* and consists—in summary—in the fact that from a *coherent* superposition (a "pure state"), an *incoherent* (or "decoherent"—thus the name) superposition with respect to a uniquely defined basis emerges. Due to an apparent reason, this process *alone* does not constitute a solution to the measurement problem, for it still cannot explain *which* branch of this now decoherent superposition corresponds to the outcome of the measurement. In Footnote 28, the measurement problem was divided up into two sub-problems: (i) "Preferred basis" and (ii) "definite outcome". The decoherence theory thus merely solves the first sub-problem.

For the Everett interpretation, this question is naturally not relevant: In its context, the basis which is preferred in this manner *defines* the splitting into independent "worlds". These are however not "exact" (as, e.g., in the suggestion of Deutsch 1985), but are rather merely approximations. In the end, the preferred basis is approximately distinguished by a dynamic process.

According to David Wallace (2010, p. 11), since the mid-1990s there has been a broad consensus among physicists that the problem of the preferred basis has been solved by environment-induced decoherence. Only in some areas of the philosophy of science is there still criticism of the fact that the approximate dynamic process of

[35] For the adherents of the de Broglie–Bohm theory, the results of decoherence, for example, permit a more exact justification of the so-called effective collapse of the wavefunction (cf. Sect. 5.1.5).

[36] Expressed technically, one takes the trace of the density matrix over the degrees of freedom of the environment. This makes it (in the preferred basis) *approximately* diagonal. The off-diagonal elements are however just what give rise to the interference effects.

decoherence is used to define objects which one then "takes seriously in the ontological sense". In Wallace's opinion, the quasi-classical branches of the wavefunction are "emergent structures", whose ontological status corresponds, for example, to that of the temperature in statistical mechanics (see the essay by Wallace in Saunders *et al.* 2010, p. 53).

The Everett interpretation has experienced a considerable revival through these results, since the decoherence-based modification is ontologically certainly less extravagant than the versions of DeWitt and Graham (1973), Deutsch (1985) or Albert and Loewer (1988).[37] The definition of the"worlds" is based here on a dynamic process which can be analysed using the methods of the standard formalism. Furthermore, this approach can be relativistically generalized in a manifest way. The significant open question which remains is that of the status of probability statements, to which we will now turn our attention.

5.2.5 Probability in Everett's Interpretation

Within the Copenhagen interpretation, if we consider a state $|\Psi\rangle = \sum_i c_i |\psi_i\rangle$, the square of the amplitude $|c_i|^2$ denotes the probability of obtaining the state $|\psi_i\rangle$ as the result of a measurement of the corresponding observable on the system $|\Psi\rangle$. In the de Broglie–Bohm theory, the same is true – but there, on the grounds that the configuration of the particle selects out this part of the wavefunction. In the GRW theory, finally, this is the probability that the dynamic collapse of the wavefunction of the measurement apparatus will lead to this state. In all of these cases, there are two preconditions for the practicable application of the probability concept: Various *possible* outcomes and the *lack of knowledge* of the actual result. Within the many-worlds interpretation, however, *all* of the results will occur *with certainty*. It thus initially appears unclear just what the probability statements could *refer to* in this connection (the "incoherence problem")—not to mention why *these* probabilities should correspond to $|c_i|^2$ (the "quantitative problem"). Precisely these two aspects (which are however closely related) are singled out in the discussion of the probability problem.

The status of probability statements within the Everett interpretation has led to a technically and conceptionally highly complex debate. Some of the important contributions to this discussion will be treated in the following. Here, again, it is seen that the advent of the decoherence theory marked a division point within the overall debate.

The Incoherence Problem

Naturally, the square of the amplitude $|c_i|^2$ still retains the mathematical properties within the Everett interpretation that qualified it to be a measure of probability (over

[37] Now and then at scientific conferences, surveys are conducted (not always with complete seriousness) about which interpretation of quantum theory is favoured by the conference attendees. Tegmark (1998) reports the result of such a survey at a workshop on quantum theory; it found that the Everett interpretation was the preferred alternative to the Copenhagen interpretation.

the set of all branchings). However, the c_i are just "branching amplitudes", and *every* branch claims the same reality in this interpretation. Both Everett, as well as later DeWitt and Graham, appear to have appreciated this difference insufficiently, since they claimed that they could even derive Born's rule:

> The conventional probability interpretation of quantum mechanics thus emerges from the formalism itself (DeWitt and Graham 1973, p. 163).

This claim is supported by DeWitt and Graham on the basis of the following mathematical result[38]: If one considers a series of N measurements of a superposition state with coefficients c_i, then in the limit $N \to \infty$, the state of the overall system ($= N$ measurement apparatus $+ N$ systems) converges towards an eigenstate of the so-called relative frequency operator for the measured value i. This operator measures—as its name implies—just the relative frequency with which the experiment yields the outcome i. The associated eigenvalue is then indeed given by $|c_i|^2$. However, to see a proof of Born's rule in this fact is a failure to recognize that in real experiments, the value of N must always remain finite, and therefore, branches occur with *statistical deviations*. Now, one can justifiably expect that their squared amplitudes remain "small". The assertion that these events thus also occur with small *probabilities* is however correct only if the squares of the branching amplitudes are indeed identifiable with probabilities. This however renders the argument circular, for just this identification is what was supposed to be *justified* (cf. Barrett 1999, p. 163; Deutsch 1985, p. 20; or Ballentine 1971, p. 234).

A genuine solution to the incoherence problem was suggested by David Deutsch in the same article in which he also treated the question of the preferred basis. It is based on the intuition that the *most probable* outcome should also be the *most frequent*. While with DeWitt, *individual* worlds branch off, Deutsch postulates an (uncountable) infinity of identical copies of the *same* world (see Axiom 8 in Deutsch 1985, p. 20). In the case of a measurement (with i possible measured values), a relative fraction p_i branches off into worlds with the corresponding experimental outcome. This fraction then corresponds to the probability of occurrence of the event i (in "my" world). Deutsch thus solves the incoherence problem by means of an extension of the "ontology" of the theory.

The *many-minds* interpretation of Albert and Loewer (1988) proceeds identically in a structural sense. As we have seen, there also, each observer state is associated with infinitely many *minds*. In the case of a measurement (with i possible measured values), these are supposed to assume the "consciousness content" that "the event has occurred", likewise with the fractional weight p_i.

If we now set this fraction p_i of the *minds* or the *worlds* (Deutsch), respectively, equal to the squared amplitude $|c_i|^2$, we also obtain an (*ad hoc*) solution to the quantitative problem.[39]

---

[38]This theorem was discovered by Neil Graham in 1970, during his doctoral work which was mentored by DeWitt. Already in 1968, James Hartle had proved an equivalent result (Hartle 1968).

[39]The *many-minds* interpretation buys the solution to the probability problem at the price of a substance dualism, which is accepted in modern philosophy of the mind by only a small minority

These two suggestions are of course based on special solutions of the problem of the preferred basis (cf. Sect. 5.2.3), which at latest since the advent of decoherence-based approaches is regarded as obsolete. We thus find here the curious situation that precisely the most convincing solution (in the eyes of many physicists) to the problem of the preferred basis leads to the result that the concept of probability once again appears to be a "foreign body" within the Everett interpretation.

Now, there exist various approaches which—as escape routes out of this dilemma—attempt to justify a concept of "uncertainty" or "indeterminacy" within the Everett interpretation. This concept appears to many authors to be a necessary condition for making it possible that probability statements can reasonably be made at all.

Vaidman (1998) undertook such an attempt. He considers a measurement whose possible outcomes are denoted as A and B. It is true, maintains Vaidman, that in the world A, the probability for the occurrence of the outcome could be trivially $A = 1$; however, it could also be that an experimenter *in* the world A might have no knowledge of this circumstance—for example as long as that observer in world A had not yet read off the result from the measurement apparatus. See also Vaidman *et al.* (2008)

Whether this type of "lack of knowledge" suffices to give the concepts of "probability" and "chance" a reasonable meaning however remains unclear. David Albert (see Albert 2010, pp. 367f) objects that this uncertainty is on the one-hand avoidable, and on the other, it occurs only *after* the experiment has been carried out.

Simon Saunders has developed a stronger version of this "subjective indeterminacy", which according to its claims can also be applied to situations *before* a measurement has been carried out. He argues that the branching into different worlds occurs in a way that is subjectively indeterministic. On the basis of a specific definition of "personal identity", Saunders sees in every "copy" of the observer a future self of the original observer. In this sense, the person should experience uncertainty *before* a measurement as to which person he or she will become *after* the measurement (cf. Saunders 1998). Another justification of subjective indeterminacy in the Everett interpretation is due to David Wallace, who makes the semantics of probability statements his starting point (cf. Wallace 2005). These results however remain the objects of a controversial debate (see, e.g., Greaves 2004 for a criticism of these positions). At the end of the next subsection, we will meet up with one more suggestion for treating the incoherence problem.

The Quantitative Problem

Let us postpone for the moment the incoherence problem and turn to the question of why probability statements within the Everett interpretation should obey Born's rule in particular. The thought offers itself that when a splitting into N worlds has occurred, each branch should be associated with the *same* probability, $\frac{1}{N}$. After all,

---

of philosophers. This problem motivated Lockwood (1996) to suggest a variant of the *many-minds* interpretation, which dispenses with dualism and a probabilistic dynamics. Ironically, it is however controversial as to whether or not Lockwood's theory permits a plausible probability interpretation at all (see Barrett 1999, pp. 206–211).

their "reality" is alleged to be equivalent. This strategy is however not allowed in the decoherence-based Everett interpretation, since no counting arguments can be utilized for the dynamically and only approximately defined worlds.

Some authors however challenge the justification of demanding a positive substantiation of Born's rule in the Everett interpretation (see Saunders 1998, p. 384, as well as the contribution by Papineau in Saunders *et al.* 2010). Born's rule should be able to be postulated here just as in conventional quantum theory (and like analogous propositions in other theories)—the status of probability statements would then be just as secure (or insecure) as in other areas of physics.

A completely new twist was given to this discussion by the publication of Deutsch (1999) (later, his approach was rendered more precise by Wallace (2003)). In this article, David Deutsch transferred the methods and results of decision theory to a quantum-theoretical context, and even claimed that he could derive Born's rule.

The (classical) decision theory models decision processes which are carried out by "rational agents" in uncertain situations. Probabilities are thus construed here as functional, namely as factors which guide behaviour. The fundamental concepts of this theory are "states of the world" ($s_i \in \mathcal{S}$), "actions" ($A, B, \ldots$), their "consequences" ($\mathcal{C}$), as well as "preferences", which an agent ascribes to possible actions. These preferences define an ordering within the set of actions: $A \geq B \geq C \cdots$ (in words: "action A is preferred relative to B; both are preferred over C, etc.").

Formally, actions are mappings between the states of the world and the consequences ($A(s) \in \mathcal{C}$). The agent considered has only incomplete knowledge of the actual state of the world—and therefore of the consequences of his or her actions. Decision theory can now prove the so-called representation theorem: If the preferences for actions are subject to so-called *rationality conditions*,[40] those preferences can be expressed in terms of a unique utility function U for the consequences, as well as a probability measure p for the states:

$$EU(A) = \sum_{s_i \in \mathcal{S}} p(s_i) \cdot U(A(s_i)) \,. \tag{5.14}$$

In this expression, $EU(A)$ stands for the *expected utility* of the action, and the preference of action A over action B as chosen by the agent is translated into the condition $EU(A) > EU(B)$. Greaves summarizes this relation as follows:

> This result guarantees an operational role for subjective probability: any rational agent will (at least) act as if she is maximizing expected utility with respect to some probability measure (Greaves 2007, p. 113).

These relations are often illustrated in an economic context, for example, as the rational behaviour for choosing how to bet a sum of money in a wager.

[40] The concept of "rationality" is used here in a very narrow or weak sense. Decision theory investigates logical limitations of the preferences and makes no claim to determine them with regard to content. A typical rationality requirement is the transitivity of preferences: If I prefer action A over action B, and B over C, then A must be preferred over C.

David Deutsch and David Wallace were able to prove an analogous result for the Everett interpretation by utilizing the following correspondences: "states of the world" correspond to the number of branchings after carrying out a particular measurement; "actions" correspond to the wagers on particular measurement outcomes (in a "quantum game"); and "consequences" correspond to the winnings (or losses) in the case that a certain *single* event occurs. Making use of analogous "rationality conditions", it was then possible to prove a representation theorem like that in Eq. (5.14). For the probability measure, one finds just the squares of the amplitudes, $p_i = |c_i|^2$. The rational agent will thus behave in such a way *as if* the multiple branchings represented alternatives whose occurrence frequency is given by Born's probability rule.[41]

In the eyes of the supporters of this position, the probability concept is even better accommodated within the Everett interpretation than in all other physical theories. Instead of posing a special problem, the role of probability would now even represent a strong argument *in favour* of the many-worlds interpretation.

This result however by no means ended the discussion, since there is no unanimity over the question of how conclusively the *premises* for the proof can be justified. Some authors doubt that in fact only *non-probabilistic* parts of decision and quantum theories enter into the proof. That would of course invalidate the alleged *proof* of the probability rule (cf. Hemmo and Pitowsky 2007).

Likewise problematic is the fact that decision theory investigates actions "in uncertain situations". Its applicability thus depends again on the question of whether "uncertainty" is present or not in the Everett interpretation (or whether its subjective appearance can be conclusively justified). This is at its core of course simply once again the incoherence problem of the previous section. Here, Hilary Greaves now takes a radical position: She admits freely that genuine probability and subjective uncertainty indeed have no place within the Everett interpretation. She adopts the position (Greaves 2004) that in the framework of the decision-theoretical programme, this is not at all necessary, and she argues that the rationality conditions can also be justified in the context of (deterministically) branching worlds. The associated measure $p(s_i)$ cannot however be reasonably called "probability" here. Greaves suggests instead the term *caring measure* and describes its meaning as follows:

> We might instead call it the agent's 'caring measure', since the measure quantifies the extent to which (for decision-making purposes) the agent cares about what happens on any given branch (Greaves 2007, p. 118).

The rational agent thus behaves in such a way that the expected utility is maximized over all the branches of the wavefunction with respect to the $|c_i|^2$ measure, *because* he or she knows that all the results will in fact occur.

A further objection to the decision-theoretical justification of probability relates to the fact that this programme *presupposes* that the rational agent accepts the validity

[41] Therefore, the Everett variant of the representation theorem makes an even stronger statement than its counterpart in classical decision theory. The latter determines the probability measure only relative to the corresponding preferences of the agent. There are however several such preferences which fulfil the rationality conditions!

of the Everett interpretation. Which arguments speak for it? This question refers to the so-called *evidence problem* of the Everett interpretation, that is, the question of how a verification of this theory could be obtained from measurements. Within the Everett interpretation, a series of measurements leads to a splitting into branches of the wavefunction which correspond to every *arbitrary* statistical distribution of the measured values. The occurrence of a distribution which deviates strongly from the Born prediction would thus not represent a reason for doubting the quantum theory; instead, it would be rather expected. A suggestion for solving this problem was made by Greaves and Myrvold (see Saunders *et al.* 2010, pp. 264ff). According to those authors, the decision-theoretically justified "branching weights"[42] would likewise play a confirmation-theoretical role.

5.2.6 Criticism of Everett's Interpretation

The problematic status of probability statements within the Everett interpretation was already discussed in the previous section. Let us now turn directly to another obvious objection to the interpretation, namely its extravagance. David Wallace notes at the end of his essay on the Everett interpretation:

> I have left undiscussed the often-unspoken, often-felt objection to the Everett interpretation: that it is simply unbelievable. This is because there is little to discuss: that a scientific theory is wildly unintuitive is no argument at all against it, as twentieth century physics proved time and again (Wallace 2010, p. 23).

Against this succinct remark, we could answer that the Everett interpretation carries its application of "scientific realism" further than other theories in modern physics. The scientific realist supports the view that the success of a scientific theory can best be explained by assuming that the objects and properties that it *postulates* do in fact *exist* (cf. Bartels 2007). This hypothesis thus refers expressly to not-directly-observable objects such as quarks or black holes.[43] The adherents of the Everett interpretation reason in precisely the sense that the branching of the wavefunction implies the existence of parallel worlds. This circumstance is described with notable accuracy by Ballentine:

> Rather than deny that a state vector can be a complete model of the real world, Everett and DeWitt choose to redefine 'the real world' so that a state vector [...] can be a model of it (Ballentine 1971, p. 232).

The modern (decoherence-based) approaches would seem however to have markedly improved the ontological status of the many-worlds interpretation. The almost arbitrary and unlimited multiplication of universes (or *minds*) within the earlier variant

[42] The concept of "probability" is thus again avoided here.

[43] In fact, variants of scientific realism are also possible which attribute a valid claim of truthfulness to certain *theories*, while the *entities* in question are not considered to be realistic (see Russell's position in Hacking 1983, p. 27).

of David Deutsch, or the *many-minds* interpretation of Albert and Loewer, become superfluous in these newer approaches.

The solution of the measurement problem in the many-worlds interpretation is based on an additional strong metaphysical assumption: In order to eliminate an "external observer", the measurements and observations are referred to the worlds which are continually branching off. This presumes that the mental states of the observer can likewise be described quantum-theoretically.[44] This physicalism is indeed a widespread position, but it is controversial within the philosophy of the mind as to whether it provides a solid basis for explaining the qualia problem or the typical intentionality of mental states. Tying the solution of the measurement problem to this precondition would seem to be maladroit, at the very least.

A still much more fundamental criticism was expressed by Tim Maudlin (cf. Maudlin 2010). He doubts that the Everett interpretation in fact offers a solution to the measurement problem. According to the usual view (e.g. according to Maudlin 1995!), the measurement problem consists essentially in interpreting the superposition of macroscopically different states (i.e. different pointer positions, living and dead cats, etc). From this reading, a measurement on an *eigenstate* would be unproblematic. Let $|M_0\rangle$ be the state of a measurement apparatus before the measurement, and $|\psi_1\rangle$ the eigenstate of a system relative to the quantity which is to be measured by M. Then after its measurement, the overall state $|M_1\rangle|\psi_1\rangle$ is found. Maudlin now expresses doubt as to the alleged simplicity of this special case and poses the question of in which sense a state (e.g. $|M_1\rangle$) in a high-dimensional vector space can *at all* represent the well-defined *spatial* state of a measurement apparatus, e.g. "pointer at the position 1". He criticizes the usual manner of speaking, according to which the wavefunction is defined on the *configuration space*, since the "spatial configuration" of all the parts that are represented by a point within this configuration space is not at all a component of all the interpretations of the quantum theory. While the spatial configuration of all the parts on $\mathbb{R}^3$ is an explicit component of the description in the de Broglie–Bohm theory, in a "wavefunction-monistic" theory, in contrast, this concept can not even be referred to (see Maudlin 2010, pp. 126f). The adherents of the Everett interpretation (and the same applies to some variants of the GRW theory) thus, according to Maudlin, lack the resources that would be required to establish a connection to the localized objects of our four-dimensional spacetime:

> For if the result of a measurement consists in, say, a pointer pointing a certain way, and if a pointer is made of particles, then if there are no particles there is no pointer and hence no outcome. All of this talk of a wavepacket 'representing' an outcome is unfortunate: what the wavefunction monist has to defend is that the outcome just *is* the wavefunction taking a certain form (in some high-dimensional space) (Maudlin 2010, p. 130).

According to Maudlin, the technical discussions about the concept of probability within the Everett interpretation thus obscure a decisive point: The probabilities sought after must not only be probabilities for the occurrence of physical events, but also of the *right* physical events. "Right" refers here of course to the ability

[44] Albert and Loewer (1988) formulate, in contrast, a dualistic position in their *many-minds* interpretation.

to establish a connection to elements of our everyday physical world in the sense described above.

In fact, the work of Allori *et al.* (2011) provides a variant of the many-worlds interpretation which takes Maudlin's objection into account (although, curiously, Maudlin's work has not been cited). As we shall however see, this modification dispenses with the fundamental assumption of previous many-worlds interpretations that a physical system is to be described by the wavefunction *alone*. But since it is this *formal* simplicity which is emphasized by the supporters of the many-worlds interpretation as its principal distinctive characteristic, the Allori suggestion is naturally not considered from *their* perspective to be an alternative which can be taken seriously.[45]

Let us take a brief look at this work, which also makes use of an interesting scientific-historical allusion. At the beginning of the work of Allori *et al.* (2011), the original suggestion of Schrödinger (1926) of interpreting the wavefunction "realistically" is namely analysed. In that view, in the single-electron case, the charge density is given by the expression $e \cdot |\psi|^2$. For the many-electron case, Schrödinger formulates a prescription which involves integration over the additional coordinates in configuration space. Then, as is well known, the dynamics of the Schrödinger equation leads in general to a spreading of this charge density over a large region of space within a short time. This is also the reason why Schrödinger rapidly discarded this interpretation of the wavefunction.[46] Allori *et al.* however apparently find this step to have been premature, for while Schrödinger's idea indeed stands in contradiction to point-like charges within a "one-world" theory, it can be re-interpreted in an evident way in terms of a *many-worlds theory*. Instead of the charge density, Allori *et al.* notably make use of the mass density $m(x, z)$ for technical reasons (see Footnote 1 on p. 4 in their article):

$$m(x,t) = \sum_{i=1}^{N} m_i \int d^3x_1 \cdots d^3x_N \delta(x - x_i) |\psi(x_1, \cdots, x_N)|^2 \,. \tag{5.15}$$

The mass density at a point x is thus obtained by integrating the probability density $|\psi|^2$ over all of the rest of the configuration space (this is quite analogous to the prescription of Schrödinger 1926). The many-worlds character of this theory is now obvious: If, for example, the wavefunction of Schrödinger's cat branches into the disjunctive parts ψ_{alive} and ψ_{dead}, then (5.15) will lead to *interaction-free* mass densities m_{alive} and m_{dead}. The objects described by these mass densities can, pictorially speaking, be considered to be "reciprocally transparent" (Allori *et al.* 2011, p. 7).

Allori *et al.* refer to the additionally introduced mass density $m(x, t)$ (in comparison to the usual many-worlds interpretation) as the *primitive ontology* (PO) of their

[45] Thus, we have here a conceptional similarity to the de Broglie–Bohm theory, which is not surprising if one casts a glance at the list of authors: with Valia Allori, Sheldon Goldstein, Roderich Tumulka and Nino Zanghì, we find here several prominent supporters of Bohmian mechanics.

[46] This difficulty was pointed out to Schrödinger by Hendrik Antoon Lorentz in a letter from March, 1926 (Jammer 1974, p. 31).

theory. They point out the necessity of such a structure, in order (as in Maudlin's argument) to describe material objects in real space via a physical theory.[47] As mentioned at the outset, the guiding idea of Everett's, of working with the wavefunction *alone*, is intentionally disregarded in this theory. While thus Maudlin's criticism of the many-worlds interpretation is formally invalidated by this variant, it aims in its content at the converse, for there is now no reason to prefer this interpretation over the de Broglie–Bohm theory.

5.3 The Relation Between the Various Interpretations

We conclude this chapter with a brief summary, which in particular establishes some relations among Bohm, Everett, and the interpretations introduced in Chap. 2 (the ensemble and the Copenhagen interpretations).

Both the de Broglie–Bohm theory and the Everett interpretation of quantum mechanics dispense with a discontinuous change of state ("collapse") of the wavefunction. Both interpretations thus in fact contain all the branches of the wavefunction which accumulate through splitting off as a result of every interaction. The non-observability of superpositions of macroscopically different states (e.g. during the act of measurement—but a measurement is of course a typical example of an interaction with a macroscopic object) must be explained in both interpretations. They however choose different strategies for solving this problem.

The "Bohmian solution" of the measurement problem consists in the fact that the additional spatial configuration of the "Bohmian particle" distinguishes precisely that part of the wavefunction which corresponds to the output of the measurement apparatus.[48] There can thus be no ambiguity in the "pointer position", since each state of a measurement apparatus is characterized by a unique configuration of these Bohmian particles. Applying suitable initial conditions, this allows the theory to reproduce all of the statistical predictions of quantum mechanics. In this sense, the de Broglie–Bohm theory complements the ensemble interpretation of quantum mechanics by a mechanism which describes the behaviour of the members of the ensemble.

With the exception of position measurements, however, one finds here no properties of the quantum objects which were already present *before* the measurement. In a way, this form of contextuality could be seen as the detailed development of a remark of Bohr's, which can be found for example within the following quote: "The procedure of measurement has an essential influence on the conditions on which the

[47] In Sect. 5.2.2, we have already mentioned that this variant of the many-worlds interpretation is non-local. The problem of the preferred basis and the role of probability statements can likewise be treated differently in this theory.

[48] The description of the "effective collapse" of the wavefunction in addition profits from the results of the work on decoherence.

very definition of the physical quantities in question rests" (Bohr 1935, p. 1025).[49] The "production" or "establishment" of the result through and in the act of measuring is likewise part of the Copenhagen interpretation. In contrast to the Copenhagen interpretation, the de Broglie–Bohm theory however offers a *physical mechanism* which explains this process realistically. That naturally says nothing yet about the plausibility of this mechanism.

On a quite different level, we can establish a parallel between the de Broglie–Bohm theory and the Copenhagen interpretation: A characteristic of the de Broglie–Bohm theory is its description of physical reality in terms of the *pair* consisting of the wavefunction and the configuration (formally: (ψ, Q)). As we have mentioned in Sect. 2.2.2, the Copenhagen interpretation claims that there is an "indissoluble connection" between the microscopic system and the measurement apparatus (i.e. the macroworld). In this sense, the Copenhagen interpretation thus also describes the physical world in terms of a pair—expressed formally, for example, as (ψ, 'macroworld').[50] In the de Broglie–Bohm theory, the second element of this pair is therefore replaced by the objects which according to this theory represent the constituents of the macroscopic world.

In the case of the Everett interpretation, all the possible outcomes of a measurement are realized in fact. This however is not subject to observations, since each observer is likewise subject to the splitting up of the worlds. The integration of a plausible concept of probability and the justification of Born's rule (i.e. the observable relative frequencies of occurrence) remain problematic, as we have discussed in Sect. 5.2.5. However, the results of decoherence theory have made it plausible how the pointer basis of a measurement apparatus is in fact distinguished. This "decoherence-based" version of the many-worlds interpretation thus dispenses with a good deal of ontological ballast which had been held against its earlier formulations.

The splitting into infinitely many worlds of course still appears radical and eccentric. With this background, one may tend to prefer the de Broglie–Bohm theory, at least with reference to the solution of the measurement problem. Numerous authors have however pointed out that the latter likewise contains all of the branches of the wavefunction which split off with every interaction. A more plausible resolution of the measurement problem is thus only then possible within the de Broglie–Bohm theory if it defines the ontological status of the wavefunction in a corresponding manner. In Brown and Wallace (2005), the question is discussed as to which difficulties are faced by this strategy. The suggestion of Dürr *et al.* (1996), already mentioned in Sect. 5.1.7, that the wavefunction should be seen as nomological, is criticized

[49] Bohr however saw no sort of causal relationship here; instead, he compared the influence of a measurement on its outcome with the connection between the frame of reference and the observations within the special theory of relativity.

[50] This "pairing" is intended to illustrate that even within the Copenhagen interpretation, a complete description of the physical world with reference to the wavefunction alone is not possible. The classic textbook of Landau and Lifschitz formulates this relation in a particularly pointed way: "Quantum mechanics thus occupies a rather remarkable position among physical theories: It contains classical mechanics as a limiting case, and at the same time, it requires this limiting case for its own justification" (Landau and Lifschitz 2011, p. 3).

for depending in a speculative manner on cosmological considerations. According to Brown and Wallace, the "empty wavefunctions" in the de Broglie–Bohm theory likewise correspond to real worlds—the solution of the measurement problem in the de Broglie–Bohm theory therefore must (and can) make no decisive reference at all to the particle and is congruent with that in the many-worlds interpretation. Brown and Wallace thereby emphasize the dictum of David Deutsch, who characterizes the guidance-field theories as "parallel-universe theories in a state of chronical denial" (Deutsch 1996, p. 225). A reply to this accusation is to be found, for example, in Maudlin (2010). In the section on criticisms of the Everett interpretation (Sect. 5.2.6), we have already cited this work, which casts doubt on the possibility of finding a solution to the measurement problem at all, as long as the spatial configuration is not taken into account. This points out an important and still open problem for the de Broglie–Bohm theory: The status of the wavefunction is not completely clarified in that theory either, and this signals a further line of separation between various schools within the de Broglie–Bohm theory (cf. Sect. 5.1.6).

Exercises

1. The de Broglie–Bohm theory is frequently called a theory of "hidden variables". This term implies the criticism that the theory introduces in principle unobservable quantities into its description. Write a brief dialogue between an advocate of the de Broglie–Bohm theory and a supporter of the Copenhagen interpretation, in which the former defends the theory against this criticism and accuses the "Copenhagen" advocate of actually making this error herself. In the course of this debate, additional arguments *pro* and *contra* could be introduced!
2. Explain why within the de Broglie–Bohm theory, the uncertainty relation $\Delta x \cdot \Delta p \geq \frac{\hbar}{2}$ is not violated!
3. Compare the solutions to the measurement problem in the de Broglie–Bohm and the Everett interpretations. Give examples of structural similarities and differences between them.

References

Albert, David. 2010. Probability in the Everett picture. In Saunders, *et al.,* 355–368.

Albert, David, and Barry Loewer. 1988. Interpreting the many-worlds interpretation. *Synthese* 77: 195–213.

Allori, Valia, Sheldon Goldstein, Roderich Tumulka, and Nino Zanghì. 2011. Many worlds and Schrödinger's first quantum theory. *British Journal for the Philosophy of Science* 62: 1–27.

Bacciagaluppi, Guido. 2002. Remarks on spacetime and locality in Everett's interpretation. In *Nonlocality and Modality*, ed. T. Placek, and J. Butterfield., NATO Science Series Dordrecht: Kluwer.

Bacciagaluppi, Guido, and Antony Valentini. 2009. *Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference*. Cambridge: Cambridge University Press.

Ballentine, Leslie. 1971. Can the statistical postulate of quantum theory be derived? - a critique of the many-universes interpretation. *Foundations of Physics* 3 (2): 229.
Barrett, Jeffrey A. 1999. *The Quantum Mechanics of Minds and Worlds*. Oxford: Oxford University Press.
Barrett, Jeffrey A. 2011. Everett's pure wave mechanics and the notion of worlds. *European Journal for the Philosophy of Science* 1: 277–302.
Bartels, Andreas. 2007. Wissenschaftlicher Realismus. In *Wissenschaftstheorie*, ed. A. Bartels, and M. Stöckler. Paderborn: Mentis.
Bell, John S. 1980. de Broglie–Bohm, delayed-choice, double-slit experiment, and density matrix. In Bell *et al.,* 2001: 94.
Bell, Mary, Kurt Gottfried, and Martinus Veltman (eds.). 2001. *John S. Bell on the Foundation of Quantum Mechanics*. Singapore: World Scientific.
Bohm, David. 1952. A suggested interpretation of the quantum theory in terms of 'hidden' variables. *Physical Review* 85: 166 (1st part); 180 (2nd part).
Bohm, David. 1953. Proof that probability density approaches $|\psi|^2$ in causal interpretation of the quantum theory. *Physical Review* 89: 458.
Bohm, David, and Basil J. Hiley. 1993. *The Undivided Universe*. London: Routledge.
Bohr, Niels. 1935. Quantum mechanics and physical reality. *Nature* 136: 1025–1026.
Bricmont, J. 2016. *Making Sense of Quantum Mechanics*. Heidelberg: Springer.
Brown, Harvey, and David Wallace. 2005. Solving the measurement problem: De Broglie-Bohm loses out to Everett. *Foundations of Physics* 35: 517.
Bub, Jeffrey. 1997. *Interpreting the Quantum World*. Cambridge: Cambridge University Press.
Cushing, James T. 1995. Quantum tunneling times: a crucial test for the causal program? *Foundations of Physics* 25: 296.
Daumer, Martin, Detlef Dürr, Sheldon Goldstein, and Nino Zanghì. 1996. Naive realism about operators. *Erkenntnis* 45: 379–397.
de Broglie, Louis. 1927. La structure atomique de la matière et du rayonnement et la méchanique ondulatoire. *Journal de Physique VI* 8: 25.
Deotto, Enrico, and GianCarlo Ghirardi. 1998. Bohmian mechanics revisited. *Foundations of Physics* 28: 1.
Dewdney, Chris, and Basil J. Hiley. 1982. A Quantum potential description of one-dimensional time-dependent scattering from square barriers and square wells. *Foundations of Physics* 12 (1): 27–48.
DeWitt, Bryce S., and Neil Graham (eds.). 1973. *The Many-Worlds Interpretation of Quantum Mechanics*. Princeton: Princeton University Press.
Deutsch, David. 1985. Quantum theory as a universal physical theory. *International Journal of Theoretical Physics* 24 (1): 1–41.
Deutsch, David. 1996. Comment on Lockwood. *British Journal for the Philosophy of Science* 47 (2): 222–228.
Deutsch, David. 1999. Quantum theory of probability and decisions. *Proceedings of the Royal Society of London A* 455: 3129–3137.
Dürr, Detlef. 2001. *Bohmsche Mechanik als Grundlage der Quantenmechanik*. Berlin: Springer.
Dürr, Detlef, Sheldon Goldstein and Nino Zanghì. 1992. Quantum equilibrium and the origin of absolute uncertainty. *Journal of Statistical Physics* 67: 843. arXiv:quant-ph/0308039 (the page numbers refer to those in the *arXiv* version).
Dürr, Detlef, Sheldon Goldstein, and Nino Zanghì. 1996. Bohmian mechanics and the meaning of the wave function. In *Experimental Metaphysics - Quantum Mechanical Studies in Honor of Abner Shimony*, ed. R.S. Cohen, M. Horne, J. Stachel, *et al.* Boston Studies in the Philosophy of Science Dordrecht: Kluwer.
Dürr, Detlef, and Dustin Lazarovici. 2012. Quantenphysik ohne Quantenphilosophie. In *Philosophie der Physik*, ed. M. Esfeld. Berlin: Suhrkamp.
Elby, Andrew, and Jeffrey Bub. 1994. Triorthogonal uniqueness theorem and its relevance to the interpretation of quantum mechanics. *Physical Review A* 49 (5): 4213.

Everett, I.I.I., Hugh. 1957. "Relative state" formulation of quantum mechanics. *Review of Modern Physics* 29 (3): 454.
Greaves, Hilary. 2004. Understanding Deutsch's probability in a deterministic multiverse. *Studies in History and Philosophy of Modern Physics* 35: 423.
Greaves, Hilary. 2007. Probability in the Everett interpretation. *Philosophy Compass* 2 (1): 109–128.
Greenberger, Daniel, Klaus Hentschel, and Friedel Weinert (eds.). 2009. *Compendium of Quantum Physics*. Dordrecht: Springer.
Hacking, Ian. 1983. *Representing and Intervening*. Cambridge: Cambridge University Press.
Hartle, James. 1968. Quantum mechanics of individual systems. *American Journal of Physics* 36: 704–712.
Heisenberg, Werner. 1959. *Physik und Philosophie*. Frankfurt/M.: Ullstein.
Hemmo, Meir, and Itamar Pitowsky. 2007. Quantum probability and many worlds. *Studies in History and Philosophy of Modern Physics* 38: 333–350.
Hiley, Basil J. 1999. Active information and teleportation. In *Epistemological and Experimental Perspectives on Quantum Physics*, ed. D. Greenberger, *et al.* Dordrecht: Kluwer Academic Publishers.
Holland, Peter R. 1993. *The Quantum Theory of Motion*. Cambridge: Cambridge University Press.
Jammer, Max. 1974. *The Philosophy of Quantum Mechanics*. New York: Wiley.
Landau, Lev D. and Jewgeni M. Lifschitz. 2011. *Lehrbuch der theoretischen Physik (Vol. 3) Quantenmechanik*. Unedited reprint of the 9th edition, 1986. Frankfurt/Main: Harri Deutsch.
Lockwood, Michael. 1996. 'Many minds' interpretations of quantum mechanics. *British Journal for the Philosophy of Science* 47: 159–188.
Madelung, Erwin. 1926. Eine anschauliche Deutung der Gleichung von Schrödinger. *Naturwissenschaften* 14 (45): 1004.
Maudlin, Tim. 1995. Three measurement problems. *Topoi* 14 (1): 7.
Maudlin, Tim. 2010. Can the world be only wavefunction? In Saunders, *et al.,* 2010: 121–143.
Mermin, David. 1990. Simple unified form for the major no-hidden variables theorems. *Physical Review Letters* 65: 3373.
Möllenstedt, Gottfried, and Claus Jönsson. 1959. Elektronen-Mehrfachinterferenzen an regelmäßig hergestellten Feinspalten. *Zeitschrift für Physik* 155: 472–474.
Myrvold, Wayne C. 2003. On some early objections to Bohm's theory. *International Studies in the Philosophy of Science* 17 (1): 7–24.
Pagonis, Constantine, and Rob Clifton. 1995. Unremarkable contextualism: dispositions in the Bohm theory. *Foundations of Physics* 25 (2): 281.
Passon, Oliver. 2004. What you always wanted to know about Bohmian mechanics but were afraid to ask. *Physics and Philosophy* 3.
Passon, Oliver. 2010. *Bohmsche Mechanik*, 2nd ed. Frankfurt/Main: Harri Deutsch.
Philippidis, Chris, Chris Dewdney, and Basil J. Hiley. 1979. Quantum interference and the quantum potential. *Il Nuovo Cimento* 52 B: 15.
Saunders, Simon. 1998. Quantum mechanics, and probability. *Synthese* 114: 373.
Saunders, Simon, Jonathan Barrett, Adrian Kent, and David Wallace (eds.). 2010. *Many Worlds? Everett, Quantum Theory, & Reality*. Oxford: Oxford University Press.
Schlosshauer, Maximilian. 2005. Decoherence, the measurement problem, and interpretations of quantum mechanics. *Review of Modern Physics* 76: 1267–1305.
Schrödinger, Erwin. 1926. Quantisierung als Eigenwertproblem (Vierte Mitteilung). *Annalen der Physik* 81: 109–139.
Tegmark, Max. 1998. The Interpretation of quantum mechanics: many worlds or many words. *Fortschritte der Physik* 46: 855–862.
Teufel, Stefan, and Roderich Tumulka. 2005. Simple proof for global existence of Bohmian trajectories. *Communications in Mathematical Physics* 258: 349–365.
Tumulka, Roderich. 2007. The 'unromantic pictures' of quantum theory. *Journal of Physics A: Mathematical and Theoretical* 40: 3245–3273.

Vaidman, Lev. 1998. On schizophrenic experiences of the neutron or why we should believe in the many-worlds interpretation of quantum theory. *International Studies in the Philosophy of Science* 12: 245–261.

Vaidman, Lev. 2008. Many-worlds Interpretation of quantum mechanics. In *The Stanford Encyclopedia of Philosophy*, ed. E.N. Zalta, (Fall 2008 Edition).

Valentini, Antony. 1991. Signal-locality, uncertainty, and the subquantum H-theorem. *Physics Letters A* 156 (1–2): 5 (I); *Physics Letters A* 158 (1–2): 1 (II).

Valentini, Antony. 2004. Universal signature of non-quantum systems. *Physics Letters A* 332: 187–193.

Valentini, Antony, and Hans Westman. 2005. Dynamical origin of quantum probabilities. *Proceedings of the Royal Society of London A* 461: 253–272.

Wallace, David. 2003. Everettian rationality: defending Deutsch's approach to probability in the Everett interpretation. *Studies in the History and Philosophy of Modern Physics* 34: 415–438.

Wallace, David. 2005. Language use in a branching universe. http://philsci-archive.pitt.edu/2554/.

Wallace, David. 2010. The Everett interpretation. In *The Oxford Handbook of Philosophy of Physics*, ed. Batterman (in preparation, cited from the preprint version at http://philsci-archive.pitt.edu).

Zeh, Hans Dieter. 1970. On the interpretation of measurement in quantum theory. *Foundations of Physics* 1: 69–76.

Zurek, Wojciech H. 1981. Pointer basis of quantum apparatus: into what mixture does the wave packet collapse? *Physical Review D* 24: 1516–1525.

Chapter 6
Quantum Field Theory

Meinard Kuhlmann and Manfred Stöckler

Quantum field theory (QFT) shares many of its philosophical problems with quantum mechanics. This applies in particular to the quantum measurement process and the connected interpretive problems, to which QFT contributes hardly any new aspects, let alone solutions. The question as to how the objects described by the theory are spatially embedded was already also discussed for quantum mechanics. However, the new mathematical structure of QFT promises new answers, which renders the spatiotemporal interpretation of QFT the pivotal question. In this chapter, we sketch the mathematical characteristics of QFT and show that a particle as well as a field interpretation breaks down. Eventually, we discuss alternative ways for interpreting QFT.

6.1 Characterization of Quantum Field Theory

Quantum field theory (QFT) shares many of its philosophical problems with quantum mechanics. Among these are the problem of the measurement process and the interpretation problems which are related to it, to which QFT can make no new contributions. The question of how the objects described by the theory are embedded in spacetime is also discussed in quantum mechanics. The new mathematical structures of QFT lead us to expect new answers, however, so that the spatiotemporal interpretation of the theory has become an important topic. QFT would appear also to open up a new view on the questions of indistinguishability and identity of the quantum objects and on the applicability of Leibniz's principle (cf. Chap. 3), which extends beyond that of quantum mechanics. The question of which kinds of objects and properties are addressed by QFT is especially significant for classical topics of natural philosophy, since as a fundamental theory, QFT is particularly relevant to ontology. For the philosophy of science, QFT is appealing because it is a theory whose development has not yet been completed, and whose provisional character is much more typical of how physicists actually work than are the textbook descriptions which often form the basis for philosophical analyses.

C. Friebe et al., *The Philosophy of Quantum Physics*,
https://doi.org/10.1007/978-3-319-78356-7_6

QFT is the mathematical and conceptual framework in which the physics of elementary particles is formulated. This is the starting point for those who are investigating which picture of matter best conforms to the view of contemporary physics. However, the answers to such questions cannot be simply read out of the mathematical formalism. Neither within physics nor in philosophy is there a consensus as to what types of entities are in fact dealt with by this theory. This question stands at the centre of current philosophical debates on QFT, and it is also at the focus of the present chapter.

From a mathematical point of view, one can arrive at QFT by applying the same heuristic procedures onto classical field theories, which led from classical particle mechanics to the Schrödinger equation. In this view, QFT is a quantum theory of systems with infinitely many degrees of freedom. Degrees of freedom refer here to modes of motion which are generally independent of one another. Thus, for example, the three degrees of freedom of a point-like particle are associated with its three independent position coordinates. The number of degrees of freedom determines the number of quantities which are required in order to characterize the state of a system. For a single classical particle, specifying the three components of its position (vector) and the three components of its momentum will suffice. For the characterization of a field, one requires at every location the field amplitude and an associated field momentum. With respect to mathematics, a field is thus similar to a system of infinitely many particles which are distinguished by their spatial coordinates. For many purposes, the relation between quantum mechanics and QFT can in this sense be seen as the transition from a finite number to an infinite number of degrees of freedom.

This transition is in fact already necessary if one makes use of the Schrödinger equation to compute atomic spectra. The reason is that quantum mechanics is incomplete in a characteristic manner: the interaction of electromagnetic radiations and matter—which indeed makes a test of the theory possible in the first place—is either excluded completely or is taken into account in a semi-classical manner. Already in the 1920s, Max Born, Werner Heisenberg and Pascual Jordan, as well as Paul Dirac, were working on approaches to a quantum theory of the electromagnetic field and its interactions with matter, i.e. a "field quantization". It soon became evident that in QFT, matter (such as electrons) and radiation (such as light) can be described in a mathematically quite similar way. They combined within the framework of QFT earlier, mostly heuristic models of the particle character (which, for instance, light can exhibit under certain circumstances) and the wave character (which matter exhibits under certain circumstances) into a unified mathematical formalism. At the same time, it became clear that the old division of matter as spatially localized particles and radiation as spatially continuously distributed fields would have to be abandoned. It is particularly notable that QFT can also describe systems and processes with variable particle numbers, i.e. processes in which particles are created or annihilated. There is no possibility of such a description within quantum mechanics.

In the following Sect. 6.2, we will first, in preparation, have a look at how classical physics describes the spatiotemporal embedding of the fundamental objects, i.e. of particles and fields, mathematically. We shall see why these classical concepts are just as unsuited to the mathematical structure of QFT as to that of quantum mechanics.

To this end, we must take a more precise look at QFT in its different alternative forms. Then, we will be able to analyse and evaluate various suggestions for the interpretation of QFT.

6.2 The Spatiotemporal Description of Processes

Before we analyse the spatiotemporal embedding of processes in QFT, we provide a summary of how classical physics deals with natural processes in the framework of space and time. We will begin by concentrating on their placement within space. One can distinguish two types of spatial placement, which are associated with two concepts of the objects considered, thus leading to different ontologies: The particle ontology and the field ontology—or, alternatively, the particle picture and the field picture.

The assumption of a particulate structure of matter lies in the tradition of mechanistic thinking, and this assumption has proven to be very successful both on a large scale (e.g. for the description of the orbits of the planets) and on a small scale (e.g. in the kinetic theory of heat). In the investigation of the newly discovered electron, it was quickly recognized that electrons follow paths and that their masses and charges are always multiples of an elementary mass and an elementary charge. Thus, electrons were classified as *particles*.

In classical mechanics, the fundamental objects are particles whose spatial extension is vanishingly small in the ideal case. The state of such a point particle is determined when one specifies its three position coordinates and the three components of its momentum. The particles (corpuscles) are localized at all times, and their locations in space follow a continuous path over time. Such a path (trajectory) can be described by a function $x(t)$ which specifies the location x of the particle at each time t. This path also permits the rediscovery and identification of the same particle at a later time. It thus becomes possible to distinguish particles and to follow their motions as "individuals". Insofar as these are not composite particles, they are indestructible. Within the mathematical formalism of classical mechanics, the creation or annihilation of a particle cannot be described.

In the nineteenth century, it became clear that not all phenomena in nature can be encompassed within a corpuscular world view. Thus, alongside the particle model, a second, fundamentally different way of describing processes developed: *field physics*. In field physics, the state of a physical system is determined by specifying the field strength (or more precisely, the suitably defined field quantity and its associated field momentum) at every point in spacetime. Here, there are no individuals that would be analogous to the point particles in mechanics, which are the carriers of properties and events. In a certain sense, space itself becomes the carrier of properties. Whereas, for example, in hydrodynamics one assumes, in spite of its field description, a discontinuous particle substructure, the electromagnetic field is a standard example of a fundamental field. In the theory of electric fields, every point in spacetime (x, t) is associated with a field strength $E(x, t)$ which specifies which force would act on

a test charge at that point (x, t). In the debates about the nature of light, the wave or field picture had become accepted, because it successfully explains interference and diffraction phenomena. While according to classical ideas, two particles can never occupy the same spacetime point, and the energy of a particle is always concentrated at its spatial location, *fields* can be superposed, as the case requires, and can even cancel each other out in certain spatial regions due to interference. Their continuously distributed energy can thus be diluted over a large region of space.

The particle and the field ontologies are in a certain sense the descendants of older ideas from natural philosophy, namely on the one hand *atomistic* concepts and on the other *continuum* notions. Towards the end of the nineteenth century, the contention between these two pictures of the world appeared to have been reconciled: The particle ontology describes matter, and the field ontology the forces, for example the electromagnetic field. Although at first glance particles and fields appear to be essentially different, there are nevertheless common features in their mathematical descriptions. In this framework, one can regard a classical field as the limiting case of a mechanical system with infinitely many degrees of freedom, whereby the position function x in the field quantity $\Psi(x, t)$ formally takes on the role of particle indices i for the coordinates $q_i(t)$.

With the transition from classical mechanics to quantum mechanics, the difficulties with the spatial embedding of the objects described by physics began. The spatial aspects of the interpretation of quantum theory were initially discussed under the heading "wave-particle duality". The empirical starting point was the problems raised by considering interactions of radiation with matter. Einstein (1905) showed that at low intensities, radiation (light) behaves in a thermodynamic sense as if it consisted of independent energy quanta $h\nu$, and that this picture also permits an elegant explanation of the photoelectric effect. Thus, the idea became established that both matter and the electromagnetic field behave under some conditions as though they are composed of particles, and under other conditions as though they are fields. Early on, physicists also asked the question as to whether or not a spatiotemporal description of processes in microphysics had generally reached its limits (cf. Jammer 1966, p. 326). For Niels Bohr, the applicability of causal and spatiotemporal descriptions had become mutually exclusive. In his view, processes cannot simultaneously fulfil dynamic conservation laws and be integrated into a spatiotemporal framework.[1]

In the framework of "classical" quantum mechanics, neither the wave–particle duality nor the spatial embedding of quantum objects could be truly clarified. Such questions played an important heuristic role in the development of QFT, however. In this situation, most physicists adopted an instrumentalist conception of physical theories, which was very elegantly formulated by Hans Reichenbach: "…, and if you ask of him (the physicist) whether it is really a question of material particles, he will reply that this is a delicate question, which he would rather not answer. That means that more philosophy belongs to this question than the physicist requires for his technical investigations" (Reichenbach 1955, p. 84).

[1] See, e.g. the 1925 Como lecture of Bohr (1961), in particular p. 54.

From the philosophical point of view, the questions thus remain as to whether QFT deals with objects in space and time at all, whether its mathematical apparatus (although it is nominally a field theory) is closer to a particle or to a field ontology, or whether yet other ontological models should be considered. How can one find answers to such questions? The evident way is to investigate whether the mathematical structure of QFT (or of one of its formulations) is still one of a particle theory, or rather that of a field theory. The implementation of this project shows that one can find some properties of QFT which can be associated with aspects of the classical characteristics of particles, as well as some which belong to the characteristics of fields. In order to obtain an insight into the current discussions on the spatiotemporal interpretation of QFT, we must therefore first take a more detailed look at its mathematical apparatus.

6.3 The Mathematical Structure of Quantum Field Theory

The two pillars of the formalism of quantum physics, that is of quantum mechanics as well as of QFT, are *states* and *observables*. In classical point particle mechanics, the state of a particle is determined at every moment in time by specifying its position and its momentum. While the concept of state in classical mechanics is not problematic and is seldom a topic of discussion, the role of states in quantum physics is more complicated: (pure) states contain the maximum possible information about the time-varying properties of the system considered—i.e. apart from its permanent properties, which define the type of system, e.g. an electron or a photon. In contrast to classical mechanics, this does not mean that for each dynamic observable a particular value is specified, however. The uncertainty relations say that there are limits for the simultaneous attribution of incompatible properties, such as in particular position and momentum (cf. Sect. 1.2.4, especially Eq. 1.37). This fact is expressed with special clarity by the *commutation relations* of quantum physics (cf. Sect. 1.2.2, Eq. 1.16). Indeed, the commutation relations are so fundamental that with their help, the corresponding observables can be "mathematically defined" in a certain sense. In the present chapter, we shall see that commutation relations also play a central role in QFT. Moreover, as in quantum mechanics, the representation of properties by operators leads to difficulties of interpretation. The field quantities are themselves operators in QFT, which determine the value of the associated properties of a system only after their application to a state vector (cf. Sect. 1.2.2).

In order to make an understanding of QFT as accessible as possible, we will begin from an historical or heuristic viewpoint, and thereby introduce some essential elements of *conventional QFT*, insofar as their interpretation can be discussed in a more or less well-founded manner. The conventional formulations of QFT, which suffice in practice for computations, are not entirely satisfactory from a strict mathematical viewpoint, however. For this reason, there have been various attempts to produce a more exact formulation. We will take up one of these attempts: In Sect. 6.3.5, we give a brief introduction to the conceptually most important axiomatically oriented

reformulation, namely algebraic QFT. At first, however, in Sect. 6.3.1, we treat fundamental questions of the so-called quantization, and in Sect. 6.3.2, a concrete example of the quantization of a theory will be introduced, along with the space of states of QFT, which will then be more precisely analysed in Sect. 6.3.3. In Sect. 6.3.4, we investigate how operators and states are related to measurable quantities. Then, we will have collected all the preliminaries we will need in order to discuss various interpretations of QFT in Sect. 6.4.

6.3.1 The Quantization of Fields

In this section, we want to give a more detailed treatment of the heuristic programme which carries out the transition from classical fields to quantum fields, in a manner analogous to the earlier transition from classical mechanics to quantum mechanics. Seen from this point of view, the most important difference is that quantum mechanics and classical particle mechanics apply to systems with a finite number of degrees of freedom (i.e. to the properties of a finite number of particles), whereas QFT and classical field theories treat systems with an infinite number of degrees of freedom. As already mentioned, in both cases, the transition from the classical theory to the corresponding quantum theory has the consequence that the fundamental physical quantities are no longer represented by "normal" numerical values, but rather by *operators*, that is by unsaturated mathematical expressions which must act on another expression in order to yield definite values (cf. Sect. 1.2.2).

The formal transition from a classical theory of point particles to quantum mechanics can be described in such a way that the classical quantities position and momentum are subject to certain commutation relations. More precisely, this means that, e.g., the operators which represent the position and the momentum of a particle must satisfy the commutation relations

$$\begin{aligned} [q_m, p_n] &\equiv q_m p_n - p_n q_m = i\hbar\delta_{m,n} \\ [q_m, q_n] &= [p_m, p_n] = 0 \end{aligned} \tag{6.1}$$

(compare Eq. 1.16, Sect. 1.2.2). The lower equations mean that the commutator of the position operators and the commutator of the momentum operators for *different* particles are always equal to zero (the corresponding measurements can thus simultaneously yield well-defined values). The upper equation means that the commutator between the position and the momentum operator for the *same* particle ($m = n$) is equal to 1 (here times $i\hbar$), but for *different* particles ($m \neq n$), it is equal to 0. In this process of *quantization* of a classical theory, elements of the classical theory (e.g. mathematical expressions representing position, momentum and energy) are taken over into quantum physics, but there, they play new roles. The term *quantization* is due to the fact that the quantum-physical operators frequently do not have arbitrary eigenvalues; thus, for example, hydrogen atoms can take on only discrete values of the eigenvalues of their energy or angular-momentum operators, according

to quantum mechanics. This is new as compared to classical mechanics, where all the dynamic quantities can in principle take on arbitrary, continuous values.

The following considerations are directed primarily at readers who already have a somewhat better knowledge of physics. They introduce a special formulation of mechanics, called the "Lagrange formalism", which permits the field quantization to be presented in an especially transparent manner. Equation 6.1, the so-called canonical commutation relations, already belong to a formulation of mechanics that makes use of the generalized coordinates q and the corresponding "conjugate" momenta p. The general concept of a "conjugate" or "canonical" momentum is defined as $p = \frac{\partial L}{\partial \dot{q}}$, where L is the Lagrange function $L = T - V$, with the kinetic energy T and the potential V; the dot above q indicates the time derivative. Just why the momentum is defined in this way can be understood if we consider the special case of a Lagrange function for which the potential V depends only on the position, so that (in Cartesian coordinates)

$$\frac{\partial L}{\partial \dot{x}} = \frac{\partial}{\partial \dot{x}} \left(\frac{1}{2} m \dot{x}^2 \right) = m\dot{x} = p_x \, .$$

In this case, the generalized momentum is identical to the usual momentum of Newtonian mechanics. The Lagrange function L characterizes the system being considered, so that, for example, a torsion pendulum and a simple (mass and string) pendulum would have different Lagrange functions. If one knows the Lagrange function, within classical physics and with given boundary conditions it is possible to calculate all relevant quantities and their time evolution. Lagrange functions can be applied to mechanical systems, but also to fields. In the latter case, one uses Lagrangian *densities*, i.e. functions defined on real space from which one can obtain the Lagrange function by means of integration over the whole volume of space.

Within the framework of the Lagrange theory of classical fields one associates, corresponding to the generalized coordinates q and momenta p in mechanics, with every field ϕ a conjugate field (the "field momentum")

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \, , \tag{6.2}$$

which is defined by a partial derivative of the Lagrangian density $\mathcal{L}$.

This Lagrange formulation now permits us, in complete analogy to the quantization of classical mechanics, to quantize classical fields also in accord with (6.1), so that the canonical commutation relations

$$\begin{aligned} \big[\phi(\mathbf{x}, t), \pi(\mathbf{y}, t)\big] &= \mathrm{i}\hbar\delta^3(\mathbf{x} - \mathbf{y}) \qquad \text{and} \\ \big[\phi(\mathbf{x}, t), \phi(\mathbf{y}, t)\big] &= \big[\pi(\mathbf{x}, t), \pi(\mathbf{y}, t)\big] = 0 \end{aligned} \tag{6.3}$$

are fulfilled, but now for the field ϕ and its corresponding conjugate field π. The delta function δ means that ϕ and π commute when they are taken at different positions $\mathbf{x}$ and $\mathbf{y}$ (their commutator is then 0). The commutation relations refer in each case to fixed times t. We recall here once more that this is an heuristic procedure, whose

results must be checked in each individual case. The essence of quantum fields can therefore not consist of the quantization of classical fields, since for many quantum fields, there is no classical analogue.[2]

In QFT, the "field values" $\phi(\mathbf{x}, t)$ associated with spacetime points $(\mathbf{x}, t)$ are operators. This means, thus, that the "field values" are no longer be determinate properties, such as electromagnetic field strengths. In order to arrive at definite measurable quantities, the operator-valued quantum fields must act on states. In quantum mechanics, too, one requires operators *and* states in order to arrive at measurable values or probabilities for their occurrence. In Sects. 6.3.2 and 6.3.3, we will learn more about the nature of field states and in Sect. 6.3.4 about their relation to experiments. The classical field concept loses its meaning in QFT in any case, due to the operator-valuedness of quantum fields: The field operators do not attribute determinate physical properties to the points in spacetime. The question of whether or not a quantum field itself has any ontological status at all will be taken up in Sect. 6.4.3.

The structure and the interpretation of QFT are essentially determined by their manifold relations to classical precursor theories. Figure 6.1 gives an overview of the common aspects and differences along the path from classical point mechanics to a relativistic QFT. Essential structural differences lie in the two steps from above to below, i.e. in the transition to infinitely many degrees of freedom ($N \rightarrow \infty$) and in taking account of the invariance requirements of special relativity theory (& SRT). The transitions from left to right stand for the "quantization" steps which we have just encountered.

Figure 6.1 describes possible formal and heuristic relations between theories. It does not reflect the historical path to QFT, along which the possibility of a non-relativistic QFT played no role, for example. In classical electrodynamics, a relativistic theory was already available as the starting point for the quantization.

6.3.2 The Simplest Example of a Quantum Field Theory

Early in the first phase of development of quantum mechanics, it had become obvious that the relativistically invariant theory of the electromagnetic field does not fit into the framework of (non-relativistic) quantum mechanics.[3] Although quantum objects

[2] A note for more advanced readers: A manifestly relativistically invariant notation is also possible. Furthermore, we should also mention that the commutation relations (6.3) are valid for bosonic fields only, such as the electromagnetic field in particular. In the framework of QFT, however, one can in addition to these interaction fields also describe material "particles", with half-integer spin, in terms of fields. For such fermionic fields, e.g. the Dirac field for electrons, one requires *anti-commutation relations* instead of Eq. 6.3. In the following, we also make use of the so-called Heisenberg picture, i.e. we will be working with time-dependent operators.

[3] The Schrödinger equation violates the requirement of special relativity which states that the laws of nature must maintain their form when one goes via Lorentz transformations from the inertial frame of reference of one observer to that of another. The Maxwell equations, for example, fulfil this requirement: light, in particular, has the same velocity c in vacuum for all these observers.

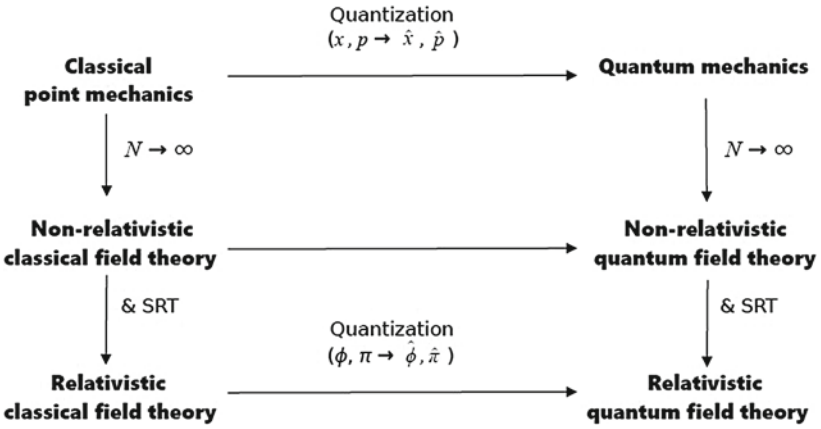


Fig. 6.1 In terms of the requirements of *quantization*, the *transition to infinitely many degrees of freedom*, and *relativistic invariance*, one can characterize structural relationships between various theories (SRT = Special Relativity Theory)

could take up and release energy only in the form of quanta even in Bohr's first atomic model, the electromagnetic radiation field itself was treated by simply using the classical theory. During the early development of quantum theory, there was a search, on the one hand, for a relativistic version of the Schrödinger equation, and on the other hand, for a quantum-physical analogue of the electromagnetic field. Both were found to be unexpectedly difficult to achieve, since the apparently direct methods ran afoul of deep-seated problems which could not be solved by simply modifying the known framework of the theory. Finally, this led to the development of QFT. In the following, we will review briefly the most important milestones along the path from quantum mechanics to the simplest example of a QFT.

One approach to combining special relativity and quantum mechanics together into a relativistically invariant wave equation consists in employing the relativistic energy-momentum relation

$$E^2 = p^2c^2 + m^2c^4 \tag{6.4}$$

for a particle of mass m, to construct an operator equation which can act upon the wavefunction $\phi(\mathbf{x}, t)$. Replacing the energy and the momentum by operators using the standard rules of quantum mechanics, $\hat{E} = \mathrm{i}\hbar\frac{\partial}{\partial t}$ and $\hat{\mathbf{p}} = -\mathrm{i}\hbar\nabla$ (see Sect. 1.2.4), where the vectorial "Nabla operator" $\nabla \equiv (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$ is used, one directly obtains the relativistic wave equation

$$\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 + \frac{m^2c^2}{\hbar^2}\right]\phi(\mathbf{x}, t) = 0\,. \tag{6.5}$$

This is the well-known *Klein–Gordon equation* (for the case of no interactions, since all possible interactions with other objects or fields have been neglected here[4]). One can see the relativistic invariance of this equation still more clearly if it is written more compactly by using the wave operator $\Box \equiv \partial^\mu \partial_\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$ (the "d'Alembert operator"), as well as the abbreviated four-dimensional notation $x = (\mathbf{x}, t)$ and the choice of units common in particle physics (which simplifies many formulas), i.e. $\hbar = 1$ and $c = 1$:

$$(\Box + m^2)\phi(x) = 0\,. \tag{6.6}$$

This free Klein–Gordon equation is a wave equation for the field ϕ. In contrast to the Schrödinger equation, it fulfils the requirements of special relativity. It is the simplest example of an equation for relativistic quantum mechanics, because it transforms like a scalar under the Lorentz transformations, since only absolute squares of the four vectors and intrinsically scalar quantities occur in it. The Klein–Gordon equation could thus describe "scalar particles", e.g. massive bosons with spin 0, such as pions. However, it is obvious then that it is not suitable for describing fermions, in particular electrons. They obey the Dirac equation, which is the second famous result of the search for a relativistic generalization of the Schrödinger equation.

This heuristic derivation of the Klein–Gordon equation was originally called "first quantization". The reason for this term is the circumstance that the Klein–Gordon equation could not be interpreted as a relativistic (quantum-mechanical) wave equation for one particle (the electron), as originally hoped.[5] As a way out of this dilemma, Eq. (6.5) is no longer considered to be the dynamical equation for a wavefunction (and m is correspondingly not taken to represent a particle mass), but rather of a classical scalar field, for which a "second quantization" according to the procedure described in Sect. 6.3.1 should be carried out. There, $\phi(x)$ is made into an operator by imposing the canonical commutation relations for fields. This procedure was finally found to be successful, leading to a physically reasonable and in fact realized result.

The concept of a "second quantization" is misleading, however, in that it suggests that for the transition to QFT two quantizations are necessary.[6] Historically, or heuristically, it is in a certain sense true in the special case of the Klein–Gordon field, but it is not in fact necessary. The "first quantization" was a failure in the sense intended, but considered *ex post facto*, it yielded the simplest example of a classical

[4] As usual in QFT, we will call it the "free theory" in the following.

[5] In particular, there are solutions with negative energies, which would lead to endless cascades towards energetically more favourable states at lower energies. In retrospect, one can argue that it cannot be expected that relativistic processes could be described by a single-particle theory, since the energy-mass equivalence $E = mc^2$ in special relativity permits the formation of particle-antiparticle pairs (Peskin and Schroeder 1995, Chap. 2). An additional problem consists in the fact that the normalization of the states determined by the Klein–Gordon equation is no longer time-independent, which undermines their interpretation as probability densities. The Klein–Gordon equation indeed fulfils the requirements of special relativity, but not those of quantum mechanics (Srednicki 2007, Chap. 1).

[6] Peskin and Schroeder (1995), Footnote on p. 19, and Redhead (1988).

field which could be quantized using the prescription described above. In modern treatments, one therefore often begins directly with the Lagrangian density

$$\mathcal{L}_{\mathrm{KG}} = \frac{1}{2}\partial^{\mu}\phi\,\partial_{\mu}\phi - \frac{1}{2}m^2\phi^2 \tag{6.7}$$

for the classical Klein–Gordon field, from which (by insertion into the Euler–Lagrange equation) one obtains the field equation directly, i.e. here the Klein–Gordon equation (6.5).[7] Without the heuristic derivation of the Klein–Gordon equation given above, the Lagrangian density (6.7) would of course seem to fall from the heavens. However, it can at least be motivated independently of the desired result, and it is a paradigm for the construction of other quantum field theories.

A particular solution of the Klein–Gordon equation (6.5) yields plane waves $\mathrm{e}^{\mathrm{i}\mathbf{p}\cdot\mathbf{x}}$, which propagate in the direction of the momentum $\mathbf{p}$, where $\mathbf{p}$ can be chosen arbitrarily. The general (real) solution of the Klein–Gordon equation is given by

$$\phi(\mathbf{x},t) = \int \frac{\mathrm{d}^3 p}{f(\omega_{\mathbf{p}})}\left[a(\mathbf{p})e^{\mathrm{i}\mathbf{p}\cdot\mathbf{x}-\mathrm{i}\omega t} + a^{\dagger}(\mathbf{p})e^{-\mathrm{i}\mathbf{p}\cdot\mathbf{x}+\mathrm{i}\omega t}\right] \tag{6.8}$$

with the corresponding wave amplitudes $a(\mathbf{p})$ and $a^{\dagger}(\mathbf{p})$, and a factor $f(\omega_{\mathbf{p}})$, which depends upon the normalization and which we will not discuss further here (Peskin and Schroeder 1995, pp. 20–22).[8] The solution (6.8) is a decomposition of $\phi(\mathbf{x},t)$ in terms of plane waves, or, physically speaking, a superposition of plane waves with a continuum of different frequencies (a so-called Fourier integral) .

Let us now carry out the quantization procedure for the field described by the Klein–Gordon equation (6.5), or (6.6). We thus assume that we are dealing with a classical field, where the quantity m at this point is simply an uninterpreted parameter. For the quantization, ϕ and its corresponding conjugate field π are considered to be operators which obey the commutation relations (6.3). An important consequence is that the factors $a(\mathbf{p})$ and $a^{\dagger}(\mathbf{p})$ introduced for the general solution likewise become operators, which fulfil the following commutation relations:

$$\begin{aligned}[a(\mathbf{p}), a^{\dagger}(\mathbf{p}')] &= f(\omega_{\mathbf{p}})\delta^3(\mathbf{p}-\mathbf{p}')\\ [a(\mathbf{p}), a(\mathbf{p}')] &= [a^{\dagger}(\mathbf{p}), a^{\dagger}(\mathbf{p}')] = 0\end{aligned} \tag{6.9}$$

containing the factor $f(\omega_{\mathbf{p}})$ as above, which we will not evaluate explicitly.

---

[7] In the Lagrange formalism, the relevant equation of motion (for fields the field equation) can be derived by subjecting the Lagrangian density to a variation procedure which is expressed by the Euler–Lagrange equation (or more precisely: which leads to that equation). Thus, for example, inserting the Lagrangian density of electrodynamics into the Euler–Lagrange equation permits one to derive the Maxwell equations.

[8] The frequency $\omega = 2\pi\nu$ is connected with the momentum via $\hbar\omega_{\mathbf{p}} = \sqrt{|\mathbf{p}|^2c^2 + m^2c^4}$. In many textbooks on QFT, the solutions of the Klein–Gordon equation are not formulated in terms of the momenta $\mathbf{p}$, but rather using the wavevectors $\mathbf{k}$, which are related via $\mathbf{p} = \hbar\mathbf{k}$ (and, if we use the above-mentioned unit convention $\hbar = 1$, are in fact identical).

The operators in (6.9) and their commutation relations are decisive for the characterization of the space of states for the field. The structure of these commutation relations is already known from the quantum-mechanical theory of the harmonic oscillator. One can show that the operator

$$N(\mathbf{p}) = a^\dagger(\mathbf{p})a(\mathbf{p})$$

has discrete whole integers $n(\mathbf{p}) = 0, 1, 2, \ldots$ as its eigenvalues. Furthermore, one can also show that there must be a state $|0\rangle$ (the "vacuum state") for which

$$a(\mathbf{p})|0\rangle = 0$$

holds (this guarantees that the norm of the eigenstates $|n(\mathbf{p})\rangle$ cannot become negative).[9] The discrete eigenvalues of the operator $N(\mathbf{p}) = a^\dagger(\mathbf{p})a(\mathbf{p})$ belong to eigenstates which are obtained when the operator $a^\dagger(\mathbf{p})$ is applied as many times to the vacuum state $|0\rangle$ as the integer eigenvalue belonging to that eigenstate, i.e.

$$[a^\dagger(\mathbf{p})]^{n(\mathbf{p})}|0\rangle = |n(\mathbf{p})\rangle\,, \tag{6.10}$$

where we have ignored the normalization.[10] Thus, for example,

$$a^\dagger(\mathbf{p})|0\rangle = |1\rangle$$

and

$$a^\dagger(\mathbf{p})|1\rangle = [a^\dagger(\mathbf{p})]^2|0\rangle = |2\rangle\,.$$

Repeated application of the operator $a^\dagger(\mathbf{p})$ to the vacuum state thus creates new states in integral steps; they are the eigenstates of the operator $N(\mathbf{p})$ with the eigenvalues 0, 1, 2,…Therefore, these operators are also called *creation operators*. As we shall see in the following section, the states thus "created" form a basis of the space of states, i.e. one can produce all the state vectors within the space by multiple applications of various creation operators and composing suitable linear combinations. Conversely, the application of the operator $a(\mathbf{p})$ leads to a reduction of the integer eigenvalue by 1, that is, for example,

$$a(\mathbf{p})|4\rangle = |3\rangle.$$

Therefore, these are called *annihilation operators*. Making use of the creation and annihilation operators, one can describe the creation and annihilation of particles in the processes of particle physics. The operator $N(\mathbf{p})$ is called the *particle-number*

[9] The "0" on the right-hand side denotes the null vector, which should not be confused with the vacuum state $|0\rangle$!

[10] In fact, at this point we are already working with the so-called Fock space representation of the commutation relations, which we will introduce systematically in the next Sect. 6.3.3.

operator (or simply "number operator"), owing to its integer eigenvalues and for reasons which will be discussed below.[11]

For a more precise characterization of the field states, we first consider the vacuum state $|0\rangle$, which is physically and philosophically interesting and important. In Brown and Harré (1988), some of these aspects are discussed. Classically, the vacuum is simply an empty space in which neither matter nor energy is present. In quantum mechanics, the harmonic oscillator has an energy ground state (i.e. the state with the lowest possible energy) whose energy eigenvalue is *not* equal to 0, and for related reasons, the vacuum of QFT is not simply a state in which nothing is present. However, due to the particle-number operator introduced above, it is a state with the particle number 0. Another remarkable property of the vacuum of QFT is that it depends on the frame of reference, i.e. different observers will find different vacua.

In order to further illustrate the significance of the operators $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$—and thus of the field operators—we will investigate the energy eigenvalues of the Klein–Gordon field in more detail. Because the Klein–Gordon field ϕ is operator-valued, the Hamiltonian H (or "Hamilton operator") which describes its energy is likewise operator-valued, since ϕ enters into it. For this Hamiltonian, we obtain:

$$H_{\mathrm{KG}} = \int \mathrm{d}^3 p \; \omega_{\mathbf{p}} \, a^\dagger(\mathbf{p}) a(\mathbf{p}). \tag{6.11}$$

In this relation, one can find a superposition of infinitely many harmonic oscillators, each with a different frequency—due to the already-mentioned analogy of the commutation relations to corresponding operators which play a central role in the quantum-mechanical theory of the harmonic oscillator. These harmonic oscillators are themselves well known in quantum mechanics (see also Footnote 11). The creation operator $a^\dagger(\mathbf{p})$ adds energy quanta to a given oscillator in integer steps $\hbar\omega_{\mathbf{p}}$. It is quite common to say that these results complete "the justification for interpreting $N(\mathbf{k})$ as the number operator, and hence for the particle interpretation of the quantized theory" (Ryder 1996, p. 131). In this view, $a^\dagger(\mathbf{p})$ is a creation operator for "particles" (i.e. bosons with spin 0; see below), which have momentum $\mathbf{p} = \hbar\mathbf{k}$ and energy $\hbar\omega_{\mathbf{k}}$. This can be seen from the fact that the single-particle states $a^\dagger(\mathbf{p})|0\rangle$ as well as the many-particle states $[a^\dagger(\mathbf{p})]^{n(\mathbf{p})}|0\rangle$ are eigenstates of the Hamiltonian (which represents the energy). The associated eigenvalues are just the relativistic energies for one or many non-interacting particles (see also Fraser 2008, pp. 845f).

In Sect. 6.4.2, we will discuss in detail the admissibility of this standard interpretation. Before we do that, we will ourselves use it provisionally in the following, in order to keep the treatment as simple as possible and to employ the usual terminology. Whenever the term "particle" appears, it should therefore be understood as a provisional manner of speaking.

Furthermore, the Klein–Gordon equation would not describe a single particle of course, as was originally intended. Equation (6.10) would mean that we would obtain

[11] For the relationship between commutation relations and the space of states, cf. Mandl and Shaw (2010), Sects. 1.2.2 and 3.1.

a state with $n(\mathbf{p})$ spin 0 bosons, each of which would have a momentum $\mathbf{p} = \hbar\mathbf{k}$ and an energy $\hbar\omega_{\mathbf{k}}$, if we allow the creation operator $a^\dagger(\mathbf{p})$ to act $n(\mathbf{p})$ times on the vacuum state $|0\rangle$. Correspondingly, one could interpret $N(\mathbf{p})$ in a literal sense as the operator for the particle number and $n(\mathbf{p})$ as the *occupation number* for bosons with momentum $\mathbf{p}$, and analogously $a(\mathbf{p})$ as their annihilation operator.

6.3.3 Occupation-Number Representation

In single-particle quantum mechanics (Chap. 1) as well as in many-particle quantum mechanics (Chap. 3), we have so far always dealt with a fixed number of quantum objects. In particle physics, which deals in particular with processes in which particles collide at such high energies that they can be annihilated and other, new particles can be formed, the framework of many-particle quantum mechanics is evidently no longer sufficient. We need a description that allows us to describe an infinite number of degrees of freedom and treat variable particle numbers.[12]

We have seen in the previous section how one can obtain the quantum-mechanical analogue of a classical relativistic field (without interactions) by field quantization. Solving the resulting field equation, we could write the operator-valued quantum field as an infinite sum of terms which itself contained expressions with remarkable properties. In particular, we found creation operators which produce new states in integer steps by repeated application to a unique ground state $|0\rangle$, the vacuum state. Their energies and momenta are related to each other as expected for relativistic particles. Furthermore, we saw that the states thus generated are eigenstates of an operator which can be interpreted as a particle-number operator, in the sense that its eigenvalues reproduce exactly the number of times that the creation operator must be applied in order to obtain the corresponding eigenstate. These results can be employed to construct an especially useful representation of the space of states of the quantized free Klein–Gordon field. It is due to a special choice of the basis in the space of the field states, which we have in fact already used in the previous section.

First, the basis of single-particle states is generated. It contains all states which are produced by a single application of the creation operators corresponding to the different values of momentum. Thus, for example, the creation operator $a^\dagger(\mathbf{p}_i)$ produces the basis state

$$|\phi_i\rangle \equiv a^\dagger(\mathbf{p}_i)|0\rangle = |0, 0, \ldots, 1_i, \ldots\rangle \,,$$

[12] The possibility of describing variable numbers of particles does not entail that fundamental interactions themselves are described. We will continue to work with the so-called free theory from the previous section. Creation and annihilation operators do not describe dynamic processes in that theory. In fact, *Haag's theorem* even says that the description of interactions is generally excluded within the framework of this theory. This limitation to a free theory has important consequences for its interpretation (see Sect. 6.4.2). For the actual treatment of scattering processes, this is not as important as it might seem, since there we are mainly dealing with asymptotically free states, i.e. far from the scattering process, and this "far from the scattering process" is attained almost immediately following the interaction.

where $|0, 0, \ldots, 1_i, \ldots\rangle$ means that there is a 1 at the ith position and otherwise only zeroes (with the abbreviation $|0\rangle \equiv |0, 0, \ldots\rangle$, which we in fact used above already). Physically, this is commonly interpreted as a state in which there is precisely one particle (a boson), which has the ith momentum $\mathbf{p}_i$—or alternatively, as a field which is excited in its ith mode.[13]

The basis of the two-particle state is formed by applying a creation operator twice to the 0-particle state $|0\rangle$. Here we have two variants, namely applying two creation operators for two different momentum values, or a twofold application of an operator for the same momentum value. In this manner, basis vectors of the type $|0, 0, \ldots, 1, \ldots, 1, \ldots\rangle$ or $|0, 2, \ldots, 0, \ldots\rangle$ are generated, for which the sum of the occupation numbers is 2. In general, the basis vectors in this "occupation-number representation" (also called the "Fock space representation") are given by

$$|n_1, n_2, ...n_i, \ldots\rangle \,,$$

which expresses the fact that n_1 particles are in the state $|\phi_1\rangle$, n_2 particles in the state $|\phi_2\rangle$, etc. These states are created by multiple applications of the creation operators to the vacuum state. Thus, the basis states of an n-particle system consist of all the states created in this way with $\sum n_i = n$. As we have already mentioned at the beginning, one essential goal of the Fock space representation is to describe states with varying particle numbers. This goal is achieved now by combining the n-particle state spaces for all $n \in (0, 1, 2, \ldots)$ into the overall state space of a bosonic Klein–Gordon field, represented mathematically by the direct sum

$$\mathcal{F}_{\text{bos.}} = \bigoplus_{n=0}^{\infty} \left\{ |n_1, n_2, \ldots\rangle, \Sigma n_i = n \right\}, \tag{6.12}$$

where the first index indicates that this Fock space is the state space of a *bosonic* field. In contrast to an n-particle Hilbert space, in a Fock space, it is thus possible to describe states with different particle numbers. In addition, the direct sum of n-particle state spaces includes linear superpositions of n-particle states with different n, so that one also has states with an undefined particle number.

Now, how does this description of many-particle systems in a Fock space fit together with the description that we encountered and used in Chaps. 3 and 4? There, particles were initially always labelled by an index, and the indistinguishability of particles of the same type was implemented by the symmetrization requirement for the allowed states (see Sect. 3.1.3). In the occupation-number representation which we are introducing here, indices for the particles are, in contrast, no longer used. The only index that we use here refers to the discrete series of momentum eigenstates which are each n_i-fold occupied. This index-free (often called "*unlabelled*")

[13]This alternative means that the ontological significance of the states created is here not (yet) determined.

representation automatically obeys the symmetrization requirement, since it is not even possible to specify that particle 'a' has the momentum value $\mathbf{p}_i$ and particle 'b' the momentum value $\mathbf{p}_j$; i.e.

$$|\phi\rangle = |\phi_i^a\rangle \otimes |\phi_j^b\rangle \,. \tag{6.13}$$

The symmetric two-particle state

$$\frac{1}{\sqrt{2}}\Big(|\phi_i^a\rangle \otimes |\phi_j^b\rangle + |\phi_i^b\rangle \otimes |\phi_j^a\rangle\Big) \tag{6.14}$$

of the labelled Hilbert-space formulation is simply expressed in terms of the state

$$|0, 0, \ldots, 1_i, \ldots, 1_j, \ldots\rangle \tag{6.15}$$

in the occupation-number representation (for bosons).[14] Here, *one* particle is in state ϕ_i and *one* particle in state ϕ_j. The indistinguishability of quantum-mechanical particles of the same type—including the relevant symmetrization requirement—is thus automatically anchored within the occupation-number representation, which naturally results from the field quantization. As we shall see in Sect. 6.4.2, this observation is an essential foundation of the so-called quanta interpretation. However, in the discussion of that interpretation, we will also see that although the Fock space representation is a useful choice of basis for many purposes in the space of the field states, we cannot derive any good arguments for a particular interpretation of QFT from it.

For an appropriate discussion of the quanta interpretation, we still require a deeper investigation into the relation between the Fock space representation (which we have just introduced) and the representation in terms of the labelled tensor product, many-particle formalism which we encountered in Sect. 3.1.2. In the latter representation, we can generate the space of states of an n-particle system in terms of the n-fold tensor product of single-particle Hilbert spaces $\mathcal{H}$; that is,

$$\mathcal{H}^n = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_n = \bigotimes_{i=1}^{n} \mathcal{H}_i \,, \tag{6.16}$$

where the lower index n enumerates the single-particle state spaces. Correspondingly, the simplest basis states $\psi \in \mathcal{H}^n$ of an n-particle Hilbert space can be written as

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \ldots \otimes |\psi_n\rangle, \quad |\psi_i\rangle \in \mathcal{H}_i \,.$$

[14] Since, according to Eq. 6.9, $a^\dagger(\mathbf{p})$ and $a^\dagger(\mathbf{p}')$ commute, the two-particle state $a^\dagger(\mathbf{p})a^\dagger(\mathbf{p}')|0\rangle$ is identical to the state $a^\dagger(\mathbf{p}')a^\dagger(\mathbf{p})|0\rangle$ with exchanged creation operators. Many-particle states are thus symmetric under permutations of the creation operators. Furthermore, arbitrarily many "Klein–Gordon particles" can be created with the same momentum or in the same field mode $\mathbf{p}$. As we have seen in Chap. 3, this means that "Klein–Gordon particles" must be *bosons*.

We have also seen that the tensor-product space (6.16) is characteristic of quantum mechanics, since its elements also include superpositions of several product states. In contrast, in the many-particle state spaces of classical mechanics, the states of the particles are completely independent of each other and are simply summed (see Sect. 3.1.2). We also saw, however, that the tensor-product space is so full of structure that in many important cases states can be described as distinct, even though nothing distinct in nature corresponds to them. Concretely, this occurred for many-particle states which differed only through permutations of particles of the same type. The result was that depending on the relevant type of particles—bosons or fermions—only symmetric or antisymmetric states are allowed, while the tensor-product space also contains non-symmetric states which do not occur in nature.

Since, in the case of our field system discussed above—with bosonic commutation relations (6.9)—we are also dealing with indistinguishable particles of the same type, $\mathcal{H}^n$ here also contains states that are not allowed, due to symmetry reasons.[15] In the framework of the labelled Hilbert-space description, it is possible now to reduce $\mathcal{H}^n$ to such a degree that the non-symmetric states are excluded and only the symmetric states (for bosons) or the antisymmetric states (for fermions) remain. For our purposes, we construct the symmetric subspace of the n-particle Hilbert space in Eq. (6.16) by means of

$$\mathcal{H}^n_{sym} = \mathcal{H}_1 \otimes_s \mathcal{H}_2 \otimes_s \ldots = \bigotimes_{i=0}^{n} {}_S\mathcal{H}_i \,, \tag{6.17}$$

where $\otimes_s$ denotes the tensor product that allows only symmetric states.[16]

In order to obtain all of the states for the case of the Klein–Gordon field, this is no longer sufficient, however. There, bosonic particles with arbitrary values of momentum can occur, or alternatively speaking, arbitrarily many field modes can be excited. The state space must therefore be extended to include variable particle numbers, by summing over all possible n-particle Hilbert spaces. Then, the Fock space introduced in Eq. (6.12) can also be constructed within the labelled tensor-product many-particle formalism as

$$\mathcal{H}_{F,\text{boson}} = \bigoplus_{n=0}^{\infty} \mathcal{H}^n_{sym} \,, \tag{6.18}$$

where the vacuum state is written simply as $\mathcal{H}^0$ for computational reasons, analogously to the n-particle Hilbert spaces. The Fock space can thus be written out as

[15] Of course there are also systems in quantum physics with *distinguishable* particles, namely many-particle systems with different types of particles. These are also described by the Hilbert-space formalism and therefore also permit the typical superpositions.

[16] We use here $h_1 \otimes_s h_2 \equiv h_1 \otimes h_2 + h_2 \otimes h_1$.

$$\mathcal{H}_{F,\text{boson}} = \bigoplus_{n=0}^{\infty} \left(\bigotimes_{i=0}^{n}{}_S \mathcal{H}_i \right). \tag{6.19}$$

Here, corresponding to (6.12), the direct sum of all the symmetrized n-particle Hilbert spaces is taken. In Sect. 6.4.2, the equivalence of the Fock space representation and the corresponding symmetrized representation in the labelled tensor product, many-particle formalism will play an important role in evaluating the quanta interpretation.

6.3.4 Quantum Field Theory and Experiments

The presentation of the mathematical structure of QFT given thus far has not yet shown explicitly how this theory is to be interpreted in a spatiotemporal sense. In a general treatment much remains initially unspecified, for example the question of how the states on which the field operators act are to be constructed. For the application of the theory, the connection between its mathematical expressions and the results of measurements must be clarified. Such results are obtained from measurement apparatus which register events at a particular location, e.g. in the form of "particle tracks". At the latest for applications to experiments, it must therefore be possible to interpret QFT spatiotemporally; at least to the extent that the results of experiments can be compared with predictions of the theory.

In the following, we will sketch a typical application, namely the analysis of a scattering process. When QFT is applied to particle physics, it is almost always employed to calculate scattering processes and their experimental verification. In a typical scattering experiment, quantum objects are prepared in a particle accelerator so that they all have the same momentum. In an intuitively clear and popular manner of speaking, which is also used in experimental physics, one could say that particles are accelerated which all move in the same direction with the same energy. This "beam of particles" is, e.g., directed at a target (scattering centre) which consists of other quantum objects. An alternative scenario is that two such particle beams are caused to collide. Usually, one cannot observe the processes in the immediate interaction zone. Using detectors, however, one can register how the momenta of the quantum objects have changed and which particles have been produced in the collisions. By evaluating the data obtained from the different detectors, one can determine properties such as the mass, charge and energy (or momentum) of the quantum objects which emerge from the interaction centre. These quantities can be compared with the calculations of the theory, which predict, for example, how many particles should be detectable in various directions and in particular energy intervals.

The usual calculations presume that at the beginning and the end of the scattering process, only widely separated, non-interacting ("free") quantum objects are found (compare Fig. 6.2); this corresponds to the experimental situation. A scattering operator S is introduced—it describes how the initial state is modified, i.e. which field

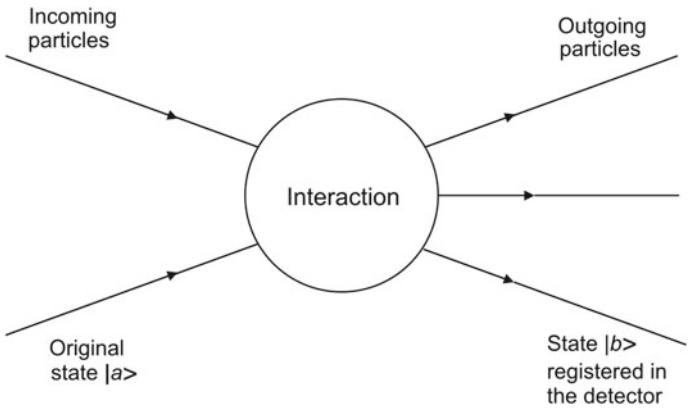


Fig. 6.2 Scheme of a scattering experiment

state will evolve from the prepared initial state (the latter describes, e.g., the momenta and energies of the particles that are going to collide).

The goal of such a calculation is to determine the probability

$$W_{ab} = |S_{ab}|^2 \tag{6.20}$$

that a certain initial state $|a\rangle$ will be transformed into a certain final state $|b\rangle$. This expression plays a similar role for scattering processes as does Born's rule in quantum mechanics. Here, the S_{ab} are the elements of the *scattering matrix*

$$S_{ab} = \langle a \,|\, S \,|\, b \rangle. \tag{6.21}$$

The basis for the calculation of the elements of the scattering matrix is an equation for the time evolution of the states, which takes into account only the evolution arising from interactions, and for which a kind of Schrödinger equation holds, whose Hamiltonian $H_{\mathrm{Int}}(t)$ contains only those elements of the Lagrange operator ('Lagrangian') that describe the interaction between the various quantum fields:

$$\mathrm{i}\hbar \frac{\partial}{\partial t} \,|\psi(t)\rangle = H_{\mathrm{Int}}(t) \,|\psi(t)\rangle. \tag{6.22}$$

The operator $H_{\mathrm{Int}}(t)$ is thus system specific and contains descriptions of the corresponding types of interactions (e.g. the interaction of a photon with an electron) in terms of field operators. Newly discovered types of particles (or quantum fields) would thus lead to additional terms in $H_{\mathrm{Int}}(t)$. For the calculation of the scattering operator S, there is a path to a solution which is determined by the operator $H_{\mathrm{Int}}(t)$. S can be calculated stepwise with successively improving approximations. One obtains the value of an element of the scattering matrix as the sum of contribu-

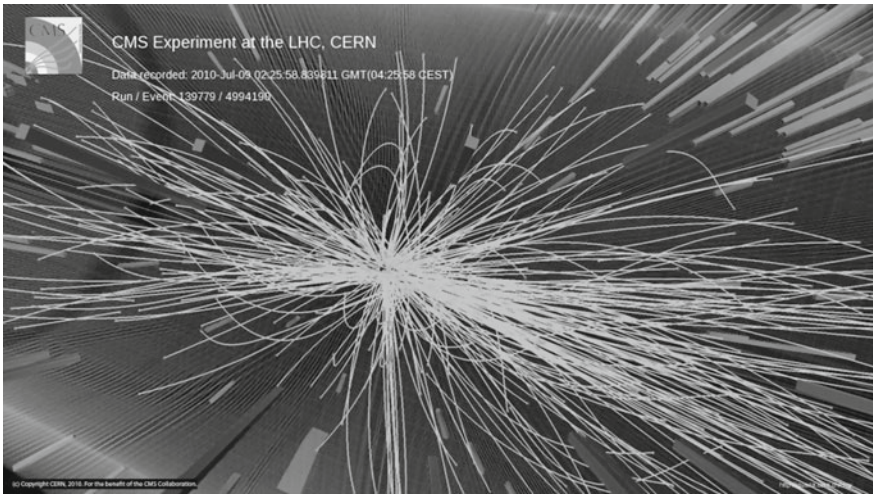


Fig. 6.3 Results of a scattering experiment

tions which become smaller and smaller (i.e. as a series expansion in the sense of perturbation theory).

The transition probabilities W_{ab} permit a test of the theory, since from them, it follows, for example, how many events are to be expected in a certain detector, or how many particles will emerge in a particular direction and with which momenta (Fig. 6.3). In order to relate the theory to a concrete experiment, one must not only know something about the field operators which enter into S, but also about the field states $|a\rangle$ and $|b\rangle$. With regard to experiments, it suggests itself to describe the states of ingoing particles and the products of interaction processes in terms of eigenstates of the momentum operators of the field. The analogy with classical fields leads to the idea that a description in terms of plane waves propagating in a direction determined by the momentum and with corresponding wavelengths is appropriate. Since these states are eigenstates of the momentum operator, one could also speak intuitively of a picture in which the field states describe a certain number of particles with a particular momentum. A typical process is then that, as a result of the interaction, the directions of the incoming particles can be changed. That is seen in the Fock space representation by the "annihilation" of a state with a certain momentum from the initial state, while in the final state, a new state is "created" with its momentum in a different direction. Initial and final states thus differ in their occupation numbers for the states of the relevant momenta.

The computation of the scattering-matrix elements $\langle a \mid S \mid b\rangle$ is rather tedious and leads to a large number of individual contributions (cf. Peskin and Schroeder 1995, Chap. 4). One obtains these elements as sums of expressions in which combinations of creation and annihilation operators occur and act on the vacuum state $|0\rangle$. The individual expressions can be associated with numerical values or integrals,

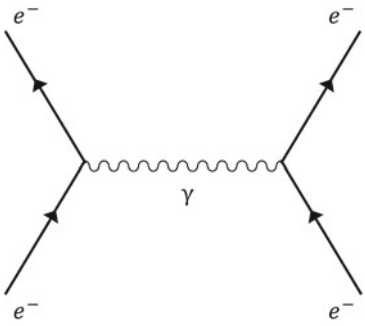


Fig. 6.4 Feynman diagram of a contribution to electron–electron scattering

whose sum then yields the value of the scattering-matrix element. An important and exceedingly helpful method here are the so-called Feynman diagrams. The individual components of a diagram (see e.g. Fig. 6.4) correspond to mathematical expressions, so that the intuitively clear diagrams aid in maintaining an overview while carrying out the computations.

Although Feynman diagrams are frequently used to help visualize elementary processes, within the theory they do not have the task of imaging spatiotemporal processes. They cannot play that role for reasons of principle. They are merely a graphic aid to the perturbation-theoretical evaluation of the scattering matrix.[17]

In Fig. 6.4, the lines representing electrons have open ends. They stand for mathematical expressions which can be associated with incoming and with outgoing particles. In contrast, the wavy line in between, which represents operators for the electromagnetic field, has a starting point and an endpoint within the diagram. Intuitively speaking, one often says that it describes a photon which is created at one point (vertex) and annihilated at another. Feynman already (in 1949) called such particles "virtual", because they do not appear in the initial or the final states. In fact, this still common manner of speaking is misleading. As mentioned, the diagrams are only illustrations of mathematical expressions, so that the internal lines do not describe the trajectories of any sorts of particles, whether virtual or real. Thus, there is no particular type of "virtual" particles, which exhibit a special mode of existence.

Figures such as Fig. 6.4 have also led to another popular misunderstanding, because it looks as though a photon is forcing the two electrons apart. It thus seems appropriate to say that the exchange of photons mediates the electromagnetic interaction between charged particles. However, a more careful look at the mathematical formulation of QFT shows that electrons, protons and photons are all represented in a similar manner by quantum fields. Interactions between such fields are described by their own terms within the Lagrangian, whose action can then be seen in the elements of the scattering matrix. The intuitively attractive picture of "particle exchange" occurs only in the Feynman graphs for perturbation calculations, and must not be understood as the description of a spatial process. Even when one interprets QFT as

[17] For a discussion of the significance of Feynman diagrams, see Kuhlmann (2010), Sects. 10.3–10.4, and Wüthrich (2012).

a theory of particles, the image of particle exchange remains a metaphor. And even on the metaphorical level, it is not intuitively clear how the exchange of particles could lead to attractive and repulsive forces.

Since the field states $|a\rangle$ and $|b\rangle$ are generated from the vacuum state $|0\rangle$ by the action of momentum operators, one can trace back the scattering-matrix elements to vacuum expectation values $\langle 0 \mid A \mid 0\rangle$, where the operator A is composed of field operators or of creation and annihilation operators. The field states are no longer explicitly visible in this expression. Nevertheless, they affect which creation and annihilation operators are chosen to characterize $|a\rangle$ and $|b\rangle$, with which one in turn determines the directions where outgoing particles can be detected. Already from this rough sketch, it is clear that in applying QFT, a scattering process cannot be completely reconstructed as a spatiotemporal event. Since we are interested only in transition matrix elements of the type $S_{ab} = \langle a \mid S \mid b\rangle$, the answer to the question as to whether it is actually the field *operators* or the field *states* which evolve over time must remain open. Both variants lead to the same transition probabilities. The whole scattering process is treated as a "stationary" problem. The dynamics of the interaction is mathematically analysed only with reference to a current of incoming and outgoing free quantum objects. Since only transition probabilities follow from the theory in such an application, it also remains open which of the possible measurement outcomes will in fact occur. Thus in QFT, too, the problem of the measurement process remains unclarified, which itself also has an aspect of spatiotemporal discontinuity.

6.3.5 Problems of Conventional Quantum Field Theory

This section treats some of the difficulties with the established formalism of QFT (as introduced here). These problems form the background of various attempts to reformulate QFT axiomatically. Here, the general assumption is that the deficiencies of conventional QFT are due to its evolved theoretical structure, and that they could be eliminated by a conceptually clearer and mathematically more precise new formulation. At the end of this section, we will briefly introduce the approach that is most fruitful for the interpretation of QFT, namely "algebraic QFT" (AQFT).

As we have already briefly mentioned in Sects. 1.2 and 1.2.4, for quantum mechanics there is a certain freedom in the choice of formalism, especially in the way in which observables (and states) are represented mathematically. As we have seen, within quantum mechanics the fundamental canonical commutation relations, e.g. for position and momentum, can be fulfilled only by operators. In general, observables are represented by linear operators on the vector space of states (the Hilbert space). The canonical commutation relations determine the corresponding operators to a considerable extent, but there remains a certain freedom in the choice of the vector space of the states, which determines the mathematical form of the operators. This fixes a representation of the algebraic structure which is determined by the commutation relations. All the representations in the various spaces of states have a common feature: The corresponding operators exhibit the same algebraic relations to each

other. For instance, the same commutation relations hold. This fact makes them into different representations of the same commutation relations.

In the Schrödinger version of quantum mechanics—i.e. wave mechanics—which works with a position basis, the position observables are represented by a multiplication operator, and the momentum by a differential operator. In Sect. 1.2.4, we wrote these operators out explicitly.[18] Another possibility is given by making the transition to a momentum representation. A further, still different alternative is matrix mechanics, due in particular to Heisenberg (see also Chap. 7)—also called the "energy representation"—in which observables are represented as matrices. As pointed out by Stone and von Neumann already in the early 1930s, wave mechanics and matrix mechanics are equivalent, because they can be converted into one another by certain ("unitary") transformations.[19] Stated more precisely, they are two equivalent representations of the more abstract Hilbert-space formalism, which at that time was just being formulated by von Neumann. Physically, (unitary) equivalence of different representations means that observables are represented differently while leading to the same measurable expectation values.

With which version of quantum mechanics—i.e. in which representation—one works usually depends on the pragmatic goal of making computations as easy as possible. Since all (irreducible)[20] representations of the commutation relations by operators on the respective Hilbert space are equivalent, the choice of representation has no physical relevance at all. This is fundamentally different in QFT, however, since there the commutation relations permit *inequivalent* representations. The reason for this difference is that fields are systems with an infinite number of degrees of freedom, which has the consequence that there are also infinitely many commutation relations (6.3), namely a set of relations for every point in spacetime. Usually, in QFT a certain Hilbert-space representation of the commutation relations is chosen without further justification. This pragmatic attitude finds a certain vindication in the fact that most of the inequivalent representations which were ignored have no physical meaning (because they violate basic physical requirements). Nevertheless, there are still various physically reasonable inequivalent representations left over.[21] In particular for the interpretation of QFT, it must therefore be emphasized that— in

[18] Since most operators are differential operators in Schrödinger's version of quantum mechanics, we called this approach the "calculus approach".

[19] In the following, some still-unexplained concepts will occasionally appear in the text, and they will be placed in quotation marks or parentheses, either to ensure that the statements made are not incorrect, and/or to guarantee to readers who wish to pursue their interests further in the literature that they will be able to follow the concepts employed there. In a first reading of this chapter, these concepts can be ignored, however.

[20] See Sect. 3.1.4.

[21] One notorious example for the significance of inequivalent representations is the so-called Unruh effect, according to which what appears to be a vacuum for one observer takes the form of a thermal bath of particles for another, accelerated observer. The deeper reason for this apparent paradox is that different inequivalent representations are associated with the two observers, which means among other things that they experience different vacuum states. In fact, different inequivalent representations are even systematically related to different vacuum states. This is the basis of the so-called GNS construction, which plays an important role in the relation between AQFT and

contrast to quantum mechanics—the choice of the representation is not a harmless, merely pragmatic matter.[22]

From the existence of different physically reasonable inequivalent representations of the commutation relations two questions arise: (i) Is it possible to either avoid choosing a particular representation or to justify the choice? (ii) Does this have an effect on the interpretation of QFT, in particular concerning ontological questions? Question (ii) will be treated in the following Sect. 6.4. An answer to question (i) is attempted within the framework of a reformulation of QFT, to which we now turn.

As we have seen, quantum field theory was first developed with the aid of analogies to classical field theory and to single-particle quantum mechanics. Due to the infinite number of degrees of freedom dealt with by QFT, particular mathematical difficulties arose again and again. Since the 1950's, there have been attempts to reformulate QFT in a systematic, preferably axiomatic manner so that these deficits could be avoided from the beginning. Arguably the most successful approach turned out to be *algebraic quantum field theory*[23] (AQFT). The central idea of AQFT is, in contrast to conventional QFT, not to choose a particular representation of the canonical commutation relations in terms of operators on a Hilbert space. Instead, the starting point for the formulation of AQFT is the level of the algebraic relations expressed by the commutation relations. Thus, abstract observable algebras such as those determined by the commutation relations stand at the centre of the theory. At this level, then—at least this was the original idea—various physical requirements are imposed in an axiomatic manner, for instance referring to the relativistic nature of the theory. It is thus required, e.g., that observables associated with spacelike-related regions must commute with each other, since measurements in causally separated spacetime regions must not influence each other, according to special relativity theory.[24]

Unfortunately, it has proven to be unexpectedly difficult to find realistic models for AQFT, especially for quantum field theories with *interactions*. This means that one cannot use AQFT in high-energy physics for computing scattering cross sections, since it cannot (yet?) deliver a satisfactory connection to experiments. In spite of this substantial limitation, it is nevertheless possible to obtain some fundamental results within the framework of AQFT, which are highly significant for the interpretation of QFT. On this background, it is not surprising that this approach forms the basis of many, possibly even most, philosophical investigations of QFT since the 1990s.[25] One prominent example are insights concerning the non-localizability of quantum objects, which yield decisive arguments against a particle interpretation (see Sect. 6.4.2).

---

conventional QFT—because, beginning with observable algebras, different operator representations lead to different vacuum states.

[22] Ruetsche (2003) discusses the philosophical consequences of this in detail.

[23] Haag (1996) offers a complete treatment of AQFT. Buchholz (2000) emphasizes the basic features of AQFT. Halvorson and Müger (2007) give a medium-long introduction, which is directed especially to philosophers of physics.

[24] This requirement does not contradict the possibility of non-local EPR correlations (see Chap. 4).

[25] See e.g. Redhead (1995), Halvorson and Clifton (2002), Earman und Fraser (2006), Baker (2009), Kuhlmann (2010), and Ruetsche (2011).

Furthermore, AQFT is crucial for some of the most recent interpretative approaches (see Sect. 6.5).

Conventional QFT has an additional problem, which explains the desire to reformulate it in a mathematically rigorous way. In Sect. 6.3.4, we sketched how in conventional QFT the connection to experiments is obtained through stepwise calculation of the scattering-matrix elements S_{ab}. In carrying out these approximate calculations, one encounters mathematical problems that cause doubts as to whether conventional QFT is even a consistent theory in a strict mathematical sense. Nearly all contributions that should approach a particular ("correct") value on summing are in fact infinite. That is, the calculation of the associated elements in the scattering matrix leads to infinite values (the corresponding integrals diverge). Computations of transition probabilities in perturbation theory thus result in series expansions which in the first few orders give excellent agreement with experimental results, but at higher orders (i.e. on adding more correction terms) diverge again.

Various pragmatic methods have been developed to avoid these difficulties. One possibility is to add additional terms to the Lagrangian, so that the scattering-matrix elements become finite again. In certain quantum field theories (i.e. with Lagrangians leading to so-called renormalizable theories), it is possible to obtain a finite result for all the relevant matrix elements by adding a limited number of additional terms. These pragmatic procedures can be justified in a certain sense, by re-interpreting the parameters which occur in the Lagrangian, such as mass and charge. One can assume, for example, that the observed mass of an electron consists of its "bare" mass plus a contribution that can be attributed to the interaction of the electron with other fields. The value of the observable mass cannot then be calculated by the theory; it must be determined experimentally.

This renormalization programme, which we have only briefly sketched here, thus makes it possible to avoid the divergences (infinite quantities) and to predict experimental effects with remarkable precision. The mathematical procedures applied for the renormalization of mass and charge suggest that the effects of processes occurring at very high energies can be taken into account by adjusting parameters such as mass and charge. This idea was extended during the 1970s into the concept of *effective field theories* (see Kuhlmann 2012, Sect. 2.4). Effective field theories describe only those interactions which are characteristic of a certain energy range. The precise form of the Lagrangian function thus depends on the energy range being investigated (e.g. on the energies of the particles that can be produced in a particular accelerator). Certain types of particles can be produced only at correspondingly high energies. The processes in the various energy ranges are to a large extent decoupled from each other. The effects of processes at very high energies can be taken into account by adjusting a few parameters, which are themselves energy dependent. While Newton's theory of gravitation holds equally well for the Moon and for the apple falling from a tree, in high-energy physics, according to this view, one is dealing with a whole series of quantum field theories (which differ in their Lagrangians), each of which is valid only within a certain energy range, and evidently must not be understood as an "ultimate" theory. If this picture is correct, there are no pressing reasons to assume that a fundamental theory in the strict sense exists at all. Therefore, ontological con-

siderations should begin now with the general theoretical framework of QFT and should not wait for a fundamental theory to be developed.

A further deficit of conventional QFT is that the concept of a field operator at a spacetime point is not mathematically well-defined.[26] These various problems, which we have sketched only briefly here, certainly make the desire for an improved mathematical reformulation of QFT understandable. However, since such approaches have been of little practical value so far, the only recourse remains to continue to work within the framework of conventional QFT for the present. As long as there is no improved version, "dirty physics" is unavoidable. Such a pragmatic approach is not unusual in modern physics. Philosophy of science, however, has dealt mainly with mature theories up to now. Nevertheless, there are good reasons to analyse the methods of "preliminary physics" more carefully and thus to contribute to an improved understanding of the current theories (compare Audretsch 1989).

The pragmatic approach within the framework of a preliminary physics is certainly quite reasonable for research in practice, in order to find a practically useful theoretical formulation under the given circumstances, even if it is not ideal. However, from a philosophical point of view, one could say that such "unfinished" theories should rather not be interpreted realistically. This would recommend deferring the clarification of the ontology of the theory until mathematically more stable versions are available. Nevertheless, in the next section, we will turn to such questions and will also explain why it is reasonable to do so.

6.4 Interpretations of Quantum Field Theory

6.4.1 Preliminary Remark

Perhaps the philosophical concerns regarding a realistic interpretation of preliminary physics are the reason why it took until about the year 2000 for interpretations of QFT to develop into an important subfield within the philosophy of physics (for an overview, see Kuhlmann 2012). There are a number of reasons for not waiting with semantic and ontological investigations until QFT has attained the degree of maturity of, for example, classical electrodynamics: For one thing, working with provisional theories is a rewarding topic in itself for the philosophy of science. Furthermore, one cannot know how long QFT will remain in its current state, and perhaps there are already some indications of the ontology of the microscopic world to be found within its present formulations. In addition, the heuristic role of ontological intuitions for the further development of theories is not to be underestimated. In this case, it would be good if the intuitions could be focused by the results of current analytical ontology.

[26] A compact treatment of the various deficits of conventional QFT can be found in Kuhlmann (2012), Sect. 4.1.

In classical physics, it is possible to interpret theories in a straightforward spatiotemporal manner. That is considerably more difficult for QFT, however, with its much richer mathematical structure. Asking for the ontology of QFT presupposes that the theory is understood not only as an instrument for predicting experimental results, but rather, at least to some extent, that it can be interpreted realistically. Under this precondition, one must state how QFT relates to the world also in those areas which are not accessible to experiment. In this chapter, we investigate in particular whether and how QFT describes processes and events in space and time.

There is no simple rule for determining the ontology of a theory, that is, for finding out with which objects a theory deals (in the broadest sense). For theories about our solar system or about the range of Darwin finches, the matter would seem to be relatively simple, since one already knows quite a lot about their objects, independently of the particular theory. In the search for an ontology of QFT, the following methods have been utilized: One refers to an intuitive understanding of the phenomena which become apparent in relevant experiments (i.e. in a certain sense, to pre-theoretical knowledge). One continues to rely upon analogies to classical theories and to interpretations of the quantum theory of "single-particle systems" and thus evaluates the heuristics which led to the newer theory. Above all, one analyses the various possibilities of giving a physical significance to the mathematical formalism of QFT. A famous philosophical view is that one can find the elements of the theory which carry its meaning, and thereby the entities about which it speaks, by looking at the quantities over which the theory quantizes in axiomatic formulations, i.e. by looking at the quantities on which the axioms impose their requirements. This route has proved to be impassable, however, at least in the current situation, due to the complicated mathematical structure of QFT.

The discussions around various interpretations of quantum mechanics suggest that quantum objects can neither be classical particles nor classical fields. Nevertheless, even the newer contributions to the interpretation of QFT usually begin with an examination of the classical particle and field concepts. This may be due to the fact that many physicists believe that one may find a way back to a particle picture, perhaps somewhat modified, via QFT. Or they are seeking new possibilities within the framework of QFT for connecting classical concepts in some way or another, since alternative models for the spatial embedding of quantum objects are not available. We will therefore likewise investigate whether or not QFT deals with particles (Sect. 6.4.2) or with fields (Sect. 6.4.3), and then we will analyse alternative suggestions for an ontology of QFT.[27]

6.4.2 The Particle Interpretation

Let us first consider in more detail the advantages and the difficulties of a *particle interpretation* of QFT. Experimental particle physics appears to favour a particle

[27] An overview of the various interpretations of QFT can be found in Kuhlmann 2012, Sect. 5.1.2.

ontology. "Particle accelerators" are constructed, detectors register "particle tracks" or make it possible to reconstruct "particle trajectories" in complicated counters, and in the end, there are Nobel prizes for the discovery of new "elementary particles". However, one must look more closely at what exactly is understood in such formulations by the term "particle". Evidently, objects are *not* meant that behave in every respect like classical particles. The problems of a particle interpretation of QFT become obvious when one investigates which features of the classical particle concept would have to be dispensed with in QFT.

The Classical Particle Concept

There is no canonical definition of what exactly a classical particle is.[28] In order to make the following discussion more transparent, we will work with the suggestion that classical particles are discrete, sharply localized, massive[29] objects with synchronic and diachronic identities. *Discreteness* means that one can specify their number. This is not the case when we specify the amount of a continuous quantity such as the field strength of a classical electric field. *Localization* distinguishes particles from fields, which extend over large regions of space. *Synchronic identity* means that particles are individuals at a given moment in time. This distinguishes them, for instance, from 100 Euros in a bank account. Discrete entities with synchronic identities are not only cardinally, but also ordinally countable.[30]

We can thus say "this is the first and that is the second particle", and this statement has an ontological significance. It refers to a real difference in the world. *Diachronic identity*, finally, means that particles as individuals can be followed along their time evolution. In classical mechanics, the diachronic identity of particles is guaranteed by the existence of their *trajectories*, i.e. their spatiotemporal paths, which cannot cross each other. Therefore, classical particles are also mutually impenetrable. This is *not* the case with two wave crests, for example, which can approach each other, meet, superpose, and finally move apart again.

As we shall see in the following, all of the characteristics of particles listed above are lost in QFT, depending upon the context. In part, this is already the case in quantum mechanics. QFT gives additional reasons for the inapplicability of the classical particle concept, however.

Discreteness

In Sect. 6.3.2, we saw that there are some features in the formalism of QFT which support the claim that QFT deals with particles. The application of the creation operator $a^\dagger(\mathbf{p})$ to the vacuum vector generates states with some properties that would

[28] Wigner's group-theoretical classification of the elementary particles (Wigner 1939) also gives no definition of the particle concept, in contrast to what is often assumed. What Wigner defines instead is "elementarity" (see Kuhlmann 2010, Sect. 8.1.2). This can be seen already by the fact that spatial localizability plays no role in Wigner's definition.

[29] *Relativistic* classical particles must obey the energy condition expressed by Eq. (6.4), owing to the equivalence of mass and energy.

[30] Instead of "ordinally countable" sometimes one simply says "countable" and contrasts it with "aggregable" (here: "cardinally countable"). Teller (1995) in particular uses alternative terminology.

also be exhibited by particles, for example integer eigenvalues of the particle-number operator. Thus, we are dealing with something that occurs in the form of discrete portions, where the summed energy and momentum values stand in the correct ratio for relativistic particles.[31] The latter result is surely less surprising than the integer eigenvalues of the particle number operator, since we obtained Eq. (6.5) on the basis of the relativistic energy-momentum relation (6.4). However, we used this latter relation to "derive" a *single-particle theory*, and then interpreted the result as a classical field. Therefore, it is not self-evident that this connection can also be transferred to the resulting theory, which deals with (possibly) *many particles*. All together, we have found two non-trivial and for relativistic particles essential properties: their discreteness and the validity of the relativistic energy condition. This is non-trivial because we imposed the commutation relations in Sect. 6.3.2 as referring to a classical relativistic field theory and not a particle theory. The resulting discreteness as well as the fact that these objects, which occur in discrete portions, fulfil the relativistic energy condition, are thus remarkable results.

Nevertheless, even the most clear-cut particle properties in QFT, i.e. discreteness and cardinal countability, are not exhibited without limitations. The first of these consists in the fact that the eigenstates of the particle-number operator are only a certain subset of the state vectors. As we saw above, the Fock space in Eq. (6.12) or in Eq. (6.19) written as a direct sum of n-particle state spaces also contains linear combinations of n-particle states with different particle numbers. Thus, there are also states with indeterminate particle numbers. This distinguishes QFT essentially from quantum mechanics, where we always have a fixed particle number. If particles are supposed to be the fundamental objects of the ontology of QFT, it would seem hardly acceptable, however, that we cannot even determine how many fundamental objects there are at a given moment in time.[32]

A second problem for the discreteness, or cardinal countability, of our potential particles may be the *Unruh effect*. An observer who is undergoing constant acceleration sees a thermal bath of particles, the so-called Rindler quanta, where another observer ('at rest') sees simply the vacuum (cf. also Footnote 21). A mere change of reference frame should not lead to the creation of many new particles, however, if particles are the fundamental entities of our ontology. Both points appear to be incompatible with an interpretation that takes the particle concept seriously.[33]

Often, however, a classical particle interpretation is not defended at all, but instead only a weaker position. Teller's so-called quanta interpretation is the approach which

[31] Fraser (2008) expresses this as follows: The countability and energy conditions are fulfilled. However, since countability is often understood in the ordinal sense—-for example by Teller—but in the present connection we are concerned with the cardinal sense, we speak instead of "discreteness".

[32] Baker (2013, p. 267) argues conversely that the situation could be similar to that of atoms in superposition states, which nevertheless does not cause us to doubt the existence of atoms.

[33] Teller (1995) argues in opposition to this view that probabilistic statements are a generic feature of quantum physics, and that both problems can be resolved with a propensity interpretation of quantum-mechanical probabilities (see Sect. 2.2.2). Teller (1995) discusses the first problem on pp. 31–33 and the second one on pp. 110–112. An accessible treatment and critical discussion of Teller's arguments is given by Huggett and Weingard (1996).

has been most discussed. The main difference is that quanta—in contrast to (classical) particles—are not individuals, which brings us to their next characteristic.

Synchronic Identity

As we saw in Chap. 3 on many-particle quantum mechanics, according to the standard view, quantum objects can violate Leibniz's principle: There are many-particle systems with quantum objects of the same type (namely various kinds of bosons or fermions) that differ in none of their permanent properties and in none of their time-dependent properties, without being numerically identical, as required by Leibniz's principle. If one subscribes to this standard view and argues that quantum objects are thus not individuals (in the sense of Leibniz's principle), then one can continue to argue – as does Teller (1995)—that it is unfavourable to work with a formalism which (apparently) enumerates particles and thus creates the impression that one is dealing with different individuals.[34] Exactly this is the case for the labelled tensor-product many-particle formalism of quantum mechanics, which we encountered in Chap. 3. Teller (1995) argues that this formalism contains so-called surplus structure, and that it would be a positive step to have a formalism in which states like that in Eq. (6.13), which correspond to nothing in nature, would no longer occur at all. Precisely such a formalism without surplus structure, says Teller, is the Fock space formalism in symmetrized form that is used in QFT (see Sect. 6.3.3). This is because with its description of states in the occupation-number representation, as for example in (6.15), it reflects in a quite natural way what is closest to being a particle in the quantum world: "quanta" as Teller calls them. They can indeed aggregate, but cannot be enumerated like the individual particles of classical physics.

Furthermore, the existence of the Fock space representation demonstrates, says Teller, that it is possible (and indeed is the only possibility) to describe infinitely many degrees of freedom and countable entities within the same framework—where there is also a relativistically invariant state containing precisely zero such entities, i.e. a vacuum. A field theory (with infinitely many degrees of freedom) thus appears to be compatible with the existence of countable particles. And, indeed, this is an essential feature of the formalism that we have encountered in Sect. 6.3.1: the equivalence of many-particle and field descriptions.

Teller has been strongly criticized for his viewpoint. Huggett and Weingard (1996) argue that the Fock space formalism is equivalent to the labelled tensor-product many-particle formalism. In particular, it was for just this reason that we described in such detail in Sect. 6.3.3 that the Fock space representation of the commutation relations for the Klein–Gordon field is equivalent to the corresponding symmetrized representation in the labelled tensor-product many-particle formalism. If these two formalisms are in fact equivalent, it is not comprehensible why the one formalism should be more relevant to ontological questions than the other. There are two more still "harder" arguments against Teller's choice of the Fock space representation as the

[34] Note that Teller presumes that it is not reasonable to assume quantum objects to be primitively individuated. However, as we have seen from the discussion on the possibility of weak distinguishability in Sect. 3.2.3, this presumption has been frequently criticized, especially in recent years.

basis of his ontological investigation of QFT. As we saw in Sect. 6.3.5, in quantum *field* theory, due to the infinite number of degrees of freedom, there are infinitely many different, *inequivalent* representations, of which the Fock space representation is only one example.[35] Therefore, Huggett and Weingard (1996) reason:

> ...the quantum field is richer than any single Fock space description, but this point is obscured by presenting the field in terms of a particular Fock space [p. 306]. […] Thus Teller's attempt to establish the quanta representation as the appropriate way to view QFT obscures some of the most crucial and startling aspects of that theory [p. 307].

Now one could object that it is questionable whether those infinitely many inequivalent representations have any physical significance at all. And indeed, this is the case for the vast majority of them. In any case, however, the Fock space representation is only valid for the *free* theory, according to Haag's theorem (see Fraser 2008). It thus cannot be the correct state space for theories which include interactions.[36] Thus, there is no unitary operator which leads from the free to the interacting representation. However, since the free theory is an idealization, it is highly problematic to draw ontological conclusions from a special representation which exists only for the free theory (see Fraser 2008).[37]

From the preceding considerations, we can extract two consequences. First, a particle interpretation is untenable, even in the sense of a "weak quanta" version ("weak" because it dispenses with synchronic identity). This is because the countability of discrete entities depends on the Fock space representation, which—as we have just seen—itself has only a very limited validity. A second consequence, based on the first, could be drawn in reference to our debate over Leibniz's principle. If QFT permits neither a particle nor a quanta interpretation, the symmetrization requirements of quantum physics would not refer to permutations of actual *objects* of any kind, but instead merely to the order in which creation operators are applied to field states.[38] We will treat this idea in more detail below.

As was shown in Sect. 3.1.2 (on many-particle tensor products), the indistinguishability postulate is fulfilled when the exchange of two of the indices (which enumerate the n factors of the tensor product of n single-particle Hilbert spaces) leads to states which are are physically indistinguishable; see Eqs. (3.3) and (3.4). This symmetry requirement is obeyed for bosonic and fermionic fields as a result of the commutation relations of the corresponding field operators.

The question of whether and in which sense QFT deals with indistinguishable quantum objects, and whether Leibniz's principle is fulfilled within quantum physics, is more difficult. The fundamental problem here is to identify mathematical elements

---

[35] A simple example can be found in Huggett and Weingard (1996 p. 306). Furthermore, in QFT on curved spacetimes, there are also infinitely many different inequivalent Fock space representations (Baker 2013).

[36] This holds in spite of the fact that the Fock space is used in perturbation-theoretical calculations within conventional quantum mechanics.

[37] Bain (2011) has formulated an alternative quanta interpretation for the asymptotically free theory.

[38] According to Baker (2013), the permutations refer to the order in which the charges are added (to "algebraic states").

of the theory which represent quantum objects. For two-particle systems such as those that we encountered in Chap. 4 (EPR), there are good reasons for speaking of "two photons" or of "two protons" as objects. These reasons are related among other things to the corresponding experimental setups for the production and detection of such systems. In experimental particle physics, also, one can speak of the fact that a detector has registered a certain particle, since a state has been measured which emerged from the interaction zone with a particular momentum and a particular charge. Here, again, one initially makes statements only about the results of a measurement process.

It thus remains an open question as to how quantum objects for which one could formulate a Leibniz principle can occur within the formalism of QFT. On closer examination, one must even ask whether the so-called quanta, whose existence is suggested by the occupation-number representation, can themselves in fact be seen as possible candidates for quantum objects. The manner of speaking that a certain state "is n-fold occupied" makes a statement about the eigenvalues of eigenstates of the so-called particle-number operator, but it fails to clarify what it is that "occupies" the state n times, and whether the search for it is meaningful at all. The idea that creation operators describe the production of particles or of quanta is suggestive, but it could lead us astray. A competing idea is that they describe transitions between different excitation states of a field, excitation states which are commonly not associated with any object-like characteristics. The state description of QFT does not associate objects in the world with states in any simple manner. "Instead, states simply characterize propensities for what will be manifested with what probability under various activating conditions". So says Teller (1995, p. 105), who then continues, "Among the items for which there can be propensities for manifestations is the occurrence of various numbers of quanta exhibiting various properties". One can certainly follow Teller insofar as there are experimental situations in which states occur from which it follows that a certain number of outgoing particles with a certain momentum can be measured. If one considers the mathematical structure of QFT in general, however, one simply finds no elements that could be considered to represent objects in any reasonable sense. And therefore, one cannot reasonably ask whether they fulfil Leibniz's principle.

In a similar vein, Baker (2013) thus draws the conclusion:

> …there is no analogue of the existing debate in interacting or curved-space-time QFTs. So puzzles about the statistical behaviour of quantum particles would seem not to bear on the question of whether the actual world is made up of individuals. According to the QFTs that offer the best available approximation to reality, there are no quantum particles, and we have no particular reason to expect that they will be reintroduced by some later more fundamental theory. […] Since QFT is probably best understood as describing the assignment of fundamental quantities to regions of space-time […], it is plausible that the best candidates for the "individuals" posited by the theory are space-time points, or space-time regions. For this reason […] space-time theories, and not basic QM, should be the locus of philosophical debate about the nature of identity and individuality in modern physics (Baker 2013, p. 284).

Diachronic Identity and Localizability

Whoever still finds a particle interpretation of QFT attractive should begin to brood at the latest on reading the following results. The arguments against a particle inter-

pretation which are generally held to be strongest are namely those related to the non-localizability of "quantum objects", which turns out to be still more radical in QFT than in quantum mechanics. Even in the heuristic derivation of the operators $a^\dagger(\mathbf{p})$ (see Eq. 6.8), one can already see that the states generated by their application are eigenstates of the momentum operator; that is, they are more similar to plane waves than to particles, and even simply due to Heisenberg's uncertainty relation, they cannot occupy a precisely-defined location. Moreover, in the light of QFT, it becomes clear that in the explanation of the photoelectric effect it is the conservation of momentum and energy that matters and not the idea of localized photons.[39]

In addition to these considerations, which relate to the conventional formulation of QFT, there are quite fundamental arguments which show that in the framework of a relativistic QFT no localized objects can occur. These mathematical indications demonstrate that no theory which fulfils certain general principles can yield localized states, that is states in which putative particles can be found with certainty within a limited region of space. In a much-noted theorem, for example, David Malament (1996) showed, making use in particular of relativistic considerations, that particle theories which fulfil a small number of not-too-limiting and plausible conditions lead to the result that the probability of finding a particle in some finite spatial region is equal to zero. That is an absurd result, however, so that a *reductio ad absurdum* points to the conclusion that quantum objects cannot be particles if their localizability is considered to be an indispensable characteristic.[40] Correspondingly, in QFT one cannot even meaningfully define a position operator.

Conclusions for the Particle Interpretation

The assumption that QFT deals with classical particles or with entities which are similar to classical particles thus runs up against great difficulties. If "particles" are often mentioned in textbooks and in everyday research practice, then this is due to the fact that the concept of "particles" has been almost arbitrarily extended, and the different meanings are related only through a family resemblance. There is no characteristic left that all the various usages of "particle" would have in common (cf. Falkenburg 2012).

6.4.3 The Field Interpretation

If the objects of QFT are not particles, then from the viewpoint of classical physics, the only option remaining is the assumption that QFT deals with *fields*. It seems to fit this interpretation that the quantum fields $\Phi(\mathbf{x}, t)$ have the spacetime manifold

[39] This was formulated somewhat more cautiously than in many other textbooks by Peskin and Schroeder (1995, p. 22): "It is quite natural to call these excitations *particles*, since they are discrete entities that have the proper relativistic energy-momentum relation. (By a *particle*, we do not mean something that must be localized in space; $a^\dagger_{\mathbf{p}}$ creates particles in momentum eigenstates)".

[40] For this topic and other proofs, cf. Halvorson and Clifton (2002), Kuhlmann (2010), Chap. 8, and Kuhlmann (2012), Sect. 5.3.

as their argument, so that to every spacetime point a quantity is assigned, thereby exhibiting the central characteristic of a field.[41] However, $\Phi(\mathbf{x}, t)$ is an *operator*, so that the spacetime points are not directly associated with any definite physical properties, in contrast to the case of the electromagnetic field for example. The field operators are important for the dynamics of the states, but only together with the states can they be associated with a spatiotemporal interpretation of definite properties.

Only when one combines the field operators $\Phi(\mathbf{x}, t)$ with the states $|\psi\rangle$ of the system on which the field operators or combinations $f(\Phi(\mathbf{x}, t))$ of them act, can one assign concrete values of physical quantities to spacetime points via expectation values of the form

$$\langle\psi| f(\Phi(\mathbf{x}, t))|\psi\rangle \ .$$

This approach has a series of difficulties, however. First, it is not at all clear just what one knows as a result of calculating the expectation values $\langle\psi| f(\Phi(\mathbf{x}, t))|\psi\rangle$. If we analyse the role of the field operators in their application, we find that they characterize possible types of interactions and observations, but not particular systems. The field operators appear, when one tries to make use of them for explanations, to belong rather to the level of laws than to that of boundary conditions which contain the changing properties of the system. In practical applications, the expectation values mentioned above play only an indirect role. The field operators are important when one asks the question of how large the probability is that a field state $|a\rangle$ will make a transition into another field state $|b\rangle$. This function of the field operators has itself no field-like character, however, and it provides no argument in favour of a field interpretation of QFT.

A similar difficulty with the practical application of QFT arises for an interpretation in which the field operators lead to the probabilities for finding classical field configurations. Huggett (2000) calls this the "wave–functional interpretation": Analogously to the way in which one can understand the wavefunction in quantum mechanics as assigning probabilities for detecting a quantum object in given regions of space, one could claim that in QFT, field configurations (which are themselves functions) are associated with probabilities for observing the corresponding configuration. (While functions map numbers onto numbers, *functionals* map functions onto numbers).

Nevertheless, a field interpretation would also have to make it clear just how it happens that, for instance, the complete charge and energy of an extended electron field can be localized within a detector "at one point". The field operators ("quantum fields") thus appear not to be interpretable as physical fields, but also the state vectors $|\psi\rangle$ are not fields in the classical sense, and furthermore can provide no information without the operators. Moreover, no other mathematical structures are apparent which

[41] In the mathematically correct algebraic formulation of QFT (see Sect. 6.3.5), the quantities are not associated with points, but rather with finite regions of spacetime; furthermore, they are not associated with individual operators, but with algebras of operators. For the following arguments, this makes no essential difference, however.

could be interpreted unproblematically as fields.[42] Thus, the proposal to interpret QFT in terms of a field ontology, too, does not seem to promise much success.

Conclusions for the Field Interpretation

The current, very lively discussions about the proper ontology for QFT have thus far not led to a result which could find wide acceptance. The least controversial point is the assessment that a particle interpretation of QFT is not tenable, in view of the multitude of problems that it encounters. The situation with regard to a field interpretation looks somewhat different. Since the discussion so far has been focused on particle interpretations, while a field interpretation was often viewed without further discussion as the only remaining alternative, there are too few careful considerations about just what a field interpretation should actually look like. It is clear that one cannot expect a classical field theory here. It needs to be spelled out in which way it would be a *field theory* in a physical sense, and not simply the assignment of certain mathematical expressions to certain points in spacetime. Furthermore, it has been investigated too little how far the arguments against a particle interpretation also count against a field interpretation. In summary, one can say that the overwhelming arguments against a particle interpretation, along with the open question of just how QFT should be understood to describe physical fields, suggest looking for new possibilities of interpreting QFT. Two alternative interpretive approaches will be discussed in the following section.

6.5 New Paths to an Interpretation

6.5.1 Ontic Structural Realism

Some philosophers believe that not things such as electrons, but rather *structures* or *relations* are the fundamental elements of the world. This is *ontic structural realism* in its strongest variant, so-called "eliminative" structural realism (Ladyman 1998). In the non-eliminative variant of ontic structural realism, which is prevalent at present, structures or relations are ontologically on at least the same level as things. This is true in the sense that not only must things exist, in order for the relations to be realized, but also that the corresponding things themselves are determined only through certain structures. The assertion is thus that there are structures or relations which enter the world not just through an arrangement of the previously existing things, but are themselves constitutive for those things. Concerning physics, the ontologically fundamental structures are mostly claimed to be *symmetries* (Lyre 2012). In QFT, these are primarily symmetries that are decisive for the classification of elementary systems (in particular elementary particles). In Sect. 3.1.4, we have seen, for example, that elementary particles can be divided into two main groups

[42] Baker (2009) has analysed some additional difficulties of a field interpretation.

according to their symmetry behaviour under permutations: bosons (symmetric) and fermions (antisymmetric).

In QFT, ontic structural realism obtains at least a certain initial plausibility through the fact that symmetry considerations were crucial when the theory was first formulated. Moreover, the fundamental symmetry structures were sometimes established before the elementary particles which obey them were discovered (Cao 2010, esp. Chaps. 1 and 9). In any case, symmetry considerations play such a fundamental role in modern physics that the question is justified whether this should not be reflected in its ontology.

A physical theory is said to be invariant under a particular symmetry transformation if its laws do not change their form when the corresponding transformation is carried out. This can be either a spatiotemporal transformation (e.g. a rotation, or time reversal) or a non-spatiotemporal transformation (e.g. a permutation, or a "gauge transformation" of potential fields). According to a pathbreaking article by Eugene Wigner (1939), the analysis of spatiotemporal symmetry groups yields a classification of elementary physical systems, such as elementary particles. Supporters of ontic structural realism interpret this as an indication that symmetries characterize elementary particles and in this sense are constitutive for them.

One important background for ontic structural realism is the ontological debate around many-particle systems with indistinguishable particles (see Chap. 3). In quantum mechanics, it seems possible that "particles" can match exactly in all their properties, both their permanent properties (such as charge and mass) as well as their time-dependent properties (such as their probability-density distributions), and nevertheless not be numerically identical. But then what is the basis for saying that we are dealing with two objects and not with one, when both of these objects match exactly in every property? The structural realist has the following answer (Esfeld and Lam 2011): The two electrons indeed have the same *monadic* properties (i.e. properties which refer only to themselves alone). However, the decisive *relational* properties with respect to their entanglement in many-particle systems were completely ignored up to now. These can be of the form that, for example, spin measurements on two electrons always show spin components in opposite directions, *"spin up"* and *"spin down"* (see Chap. 4). This irreflexive relation of properties can exist only for two *different* electrons, however. *Ergo*, the numerical diversity of the electrons is guaranteed by relational and not by monadic properties. The entanglement structure of the two-electron system is constitutive for the two electrons as distinct entities. Therefore, this structure is not produced by the "arrangement" of already existing electrons, but has at least the same ontological significance as the electrons themselves.

Building on the considerations of permutation invariance of many-particle wavefunctions, there are some further arguments in favour of ontic structural realism. Stachel (2002) suggests the following generalization: If permutations of putative individuals, whether they be spacetime points or particles, have no kind of observable effects, then Ockham's razor advises us that individual spacetime points or particles should not be regarded as fundamental, but instead the relevant symmetry structures—for example the metric structure of spacetime—or the symmetry-group

structure for the description of matter. In Sect. 3.1.4, we encountered several examples of such symmetry groups, which can be directly extended to QFT.

Whether or not there are arguments in favour of ontic structural realism which are specific to QFT and do not already hold for quantum mechanics, or even quite generally, is thus far not clear. Independently of this question, however, it could be the case that ontic structural realism can solve some problems which have arisen in the interpretation of QFT. An especially obstinate problem was that of the occurrence of different inequivalent representations in theories of systems with infinitely many degrees of freedom, such as are treated by QFT (see Sect. 6.3.5). The problem of choosing one of these inequivalent representations would be eliminated, however, if we had a justification for classifying the level of the algebraic structure of the commutation relations as ontologically fundamental. Ontic structural realism could deliver the general basis for precisely this justification.

Various objections have been raised against ontic structural realism. The principal argument against the strong eliminative variant of ontic structural realism is that assuming relations without *relata* is a contradiction in terms. For the weaker non-eliminative variant of ontic structural realism, which we have just discussed, the main problem is to spell out how exactly objects should be understood as being "structurally characterized". More traditional ontologies do not deny that there are relations or structures, and neither do they deny that they play a decisive role in the formulation of theories, in particular of QFT. The question is, however, whether or not it is reasonable to say that structures are ontologically primary, or at least stand on the same level as objects. If the advocate of ontic structural realism merely asserts that there are structures and that they are important, then no new ontology has been established.

6.5.2 A Trope-Ontological Interpretation

The fundamental idea of ontic structural realism consists in countering the problems of traditional interpretations by asserting that "things" (or "substances") such as electrons are not to be taken as fundamental elements of the ontology, but that something else should rather be considered to be primary. This characterization applies also to the so-called trope-ontological interpretation of QFT.[43] While ontic structural realism sees *relations* as fundamental, in trope ontology, *properties* are the basic elements of the ontology—where properties are understood in a non-standard way and are called "tropes" to mark this new understanding.

Trope ontology developed independently of the considerations of modern physics and has been the subject of active discussions in recent years.[44] In the philosophical discipline ontology, the term "trope" refers to single occurrences of properties. With recourse to Aristotle, we can offer the following analogy: An individual occurrence

[43] Wayne (2008), Morganti (2009), and Kuhlmann (2010).

[44] Maurin (2013) gives an up-to-date overview.

of white is related to the universal property of being white in the same way as a single human is related to the natural species of humanity. According to the point of view of "trope ontology" (in its standard form), tropes are in fact *the* fundamental category of being, to which everything else can be reduced. This standard version of trope ontology is therefore a single-category theory. "Things" (or "substances") such as the eraser which is lying in front of me on my desk are analysed as bundles of tropes, i.e. this individual occurrence of white, this rubbery consistence, this blunted cuboid shape, etc.

The tropes or properties of a bundle are evidently not to be regarded as spatiotemporal parts, but insofar as things are composed of them, they are indeed *parts*. The decisive point is now that tropes or properties are not individuated via the things whose properties they are, but their particularity is instead seen as primitive. The particularity of the tropes is the basis of the individuality of the object which they constitute. Consequently, the properties to which the objects are reduced in trope ontology cannot be *universalia*, i.e. multiply realizable entities. Otherwise, particular substantive objects—being nothing but bundles of tropes—could occur more than once, which would be a contradiction in terms.

The examples of tropes mentioned so far are only usable as illustrations of the basic idea, however. Precisely speaking, the things around us are themselves bundles of bundles, since otherwise there would be numerous problems.[45] Genuine tropes are to be found only on the fundamental level, and precisely for that reason, it is very important for philosophical trope ontology to consider what modern physics has to say about the basic elements of our material world.

How can we now spell out trope ontology in the case of quantum physics? The first suggestion is due to Simons (1994), where the background is the debate over many-particle systems once again (see Chap. 3 and Sect. 6.5.1). If properties are not considered to be *universalia*, but rather particular things, i.e. tropes, then the conflict with Leibniz's principle does not even arise: Two electrons which have precisely similar charge tropes, i.e. both carry a negative elementary charge, are two different things, since they are *not* identical in their properties after all—in the sense of tropes. According to trope ontology, properties are particulars[46] and as a result, things are also particulars, since they are nothing but bundles of tropes.

Now, what would a trope-ontological interpretation of QFT look like? Morganti (2009) connects his approach to the "Standard Model" of particle physics, in which the particles are classified according to various properties. Another variant, called "dispositional trope ontology"[47], is supported in particular by the algebraic formulation of QFT, i.e. AQFT (cf. Sect. 6.3.5). An essential point here is also the problem of inequivalent representations. Similarly to ontic structural realism, dispositional trope

[45] One concern is the *boundary problem* (Campbell 1990): Consider a blue sheet of paper that you tear apart. Are there suddenly two blue tropes now? Or were they "in" the original sheet already. If yes, then we have acquired a general problem: where does one colour trope end and another one begin on a macroscopic level?.

[46] The term "particular" is less misleading here than "individual thing", since single tropes are precisely *not* things.

[47] Kuhlmann (2010), Chaps. 11–15.

ontology also holds algebraic structures to be ontologically fundamental and avoids deriving fundamental ontologies (prematurely) from special representations. In contrast to ontic structural realism, however, dispositional trope ontology maintains that it is not possible to capture the ontology of the material world with (abstract) structures. For this reason, concrete representations which refer to the empirical world play a primary role in dispositional trope ontology.[48] A further important point is that in AQFT, it is not individual observables which are at the basis of the formulation, but rather nets of observable algebras. In other words, the central items are neither individual observables nor single observable algebras, but rather the manner in which observable algebras are related to each other with respect to different spacetime regions. According to dispositional trope ontology, this aspect is best captured by a bundle theory of properties.

There are a number of other reasons why it is advantageous not to consider either (quantum) fields or elementary particles as fundamental, but instead dispositional tropes, which in bundles correspond to the objects we know, e.g. electrons. In the following, we sketch three arguments in favour of a trope-ontological interpretation of QFT. Here is the first one: There are conservation laws for diverse physical quantities, in particular for various types of charges. There is, however, no conservation law for the particle number. Such a conservation law would contradict empirical results of high-energy physics, where millions of particles are annihilated and other types of particles are formed in scattering experiments. The trope-bundle theory can describe this fact very naturally, since particles have no fundamental status here, but rather are formed and can decay continually through new bundlings, in particular of charge tropes. A second topic where dispositional trope ontology is found to be advantageous is the vacuum of QFT. In a particle interpretation, it is incomprehensible that detectors produce signals even in a vacuum, although the vacuum is the "zero-particle" state. In contrast, dispositional trope ontology can make sense of the seething QFT vacuum: Dispositional tropes are also present in the vacuum and can lead to particle-like detection events. Finally, the non-localizability problems, which haunt the particle interpretation, disappear in a trope-ontological reading of QFT, since elementary particles are no longer fundamental objects. Therefore, it raises no difficulties that particle-like aspects can be observed only under certain conditions.

6.5.3 Conclusions for the Ontology of Quantum Field Theory

Philosophical discussions about the ontology of QFT are still relatively new. Correspondingly, especially the more recent alternative approaches have not yet been sufficiently worked out. A particular challenge for any interpretation that takes QFT seriously is the question of how to deal with the existence of different inequivalent

[48] Rossanese (2013) argues that the idea that the Unruh effect undermines the particle interpretation is not only transferable to the field interpretation (as asserted by Baker 2009), but may also be raised against a trope-ontological interpretation.

representations. How should we ontologically account for the fact that free and interacting theories, resting and accelerated observers, observers in a flat and in a curved spacetime experience such radically different pictures?[49] The open questions with respect to the ontology of QFT lead to the conclusion that we can say little about the spatiotemporal embedding of quantum objects. As we have seen, in conventional QFT there are procedures that permit at least spatial statements in connection with experiments. We can thus work successfully with QFT as a physical theory. Concerning the philosophical questions about QFT, however, there is still much to be done.

Exercises

1. Gather information about theories of light from the history of physics. Why was Newton's theory of light not considered satisfactory? Compare the mathematical descriptions of particles and of waves. What is, in your opinion, the principal difference?
2. Is there anything in classical physics which is neither a particle nor a field?
3. State two arguments which in your opinion provide the strongest support for considering quantum field theory to be a theory of particles. What objections can be raised against these arguments?
4. Would it be helpful in your opinion to extend the concepts of "particles" and "fields", i.e. by referring not only to those entities as "particles" or "fields" which exhibit all the characteristics of classical particles or classical fields?

References

Audretsch, Jürgen. 1989. Vorläufige Physik und andere pragmatische Elemente physikalischer Naturerkenntnis. In *Pragmatik. Handbuch des pragmatischen Denkens*, vol. III, ed. H. Stachowiak, 373–392. Hamburg: Meiner.

Bain, Jonathan. 2011. Quantum field theories in classical spacetimes and particles. *Studies in History and Philosophy of Modern Physics* 42: 98–106.

Baker, David J. 2009. Against field interpretations of quantum field theory. *British Journal for the Philosophy of Science* 60: 585–609.

Baker, David J. 2013. Identity, superselection theory, and the statistical properties of quantum fields. *Philosophy of Science* 80: 262–285.

Bohr, Niels. 1961. The quantum postulate and the recent development of atomic theory. In *Atomic Theory and the Description of Nature*, ed. Niels Bohr, 52–91. Cambridge: Cambridge University Press.

Brown, Harvey R., and Rom Harré (eds.). 1988. *Philosophical Foundations of Quantum Field Theory*. Oxford: Clarendon Press.

[49] One proposed solution is the "Swiss army approach" by Ruetsche (2011). However, it is not clear whether the problem is actually solved by such a multi-perspective approach, rather than merely explicitly formulated.

Buchholz, Detlev. 2000. Current trends in axiomatic qantum field theory. In *Quantum Field Theory. Proceedings of the Ringberg Workshop Held at Tegernsee, Germany, 21–24 June 1998 On the Occasion of Wolfhart Zimmermann's 70th Birthday*, vol. 558, ed. P. Breitenlohner, and D. Maison, 43–64, Lecture Notes in Physics Berlin, Heidelberg: Springer.

Butterfield, Jeremy, and Hans Halvorson (eds.). 2004. *Quantum Entanglements - Selected Papers of Rob Clifton*. Oxford: Oxford University Press.

Campbell, Keith. 1990. *Abstract Particulars*. Oxford: Blackwell.

Cao, Tian Y. 2010. *From Current Algebra to Quantum Chromodynamics: A Case for Structural Realism*. Cambridge: Cambridge University Press.

Earman, John, and Doreen Fraser. 2006. Haag's theorem and its implications for the foundations of quantum field theory. *Erkenntnis* 64: 305–344.

Einstein, Albert. 1905. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. [On a heuristic point of view about the creation and conversion of light]. *Annalen der Physik* 17: 132–148.

Esfeld, Michael. 2011. Ontic structural realism as a metaphysics of objects. In *Scientific Structuralism*, ed. A. and P. Bukolich, Chap. 8. Dordrecht: Springer.

Falkenburg, Brigitte. 2012. Was sind subatomare Teilchen? In *Philosophie der Physik*, ed. M. Esfeld, 158–184. Frankfurt: Suhrkamp.

Fraser, Doreen. 2008. The fate of "particles" in quantum field theories with interactions. *Studies in History and Philosophy of Modern Physics* 39: 841–59.

Haag, Rudolf. 1996. *Local Quantum Physics: Fields, Particles, Algebras*, 2nd ed. Berlin: Springer.

Halvorson, Hans and Rob Clifton (2002). No place for particles in relativistic quantum theories? *Philosophy of Science* 69, 1–28; See also in Butterfield and Halvorson (2004) and in Kuhlmann *et al.* (2002).

Halvorson, Hans, and Michael Müger. 2007. Algebraic quantum field theory (with an appendix by Michael Müger). In *Handbook of the Philosophy of Physics - Part A*, ed. Jeremy Butterfield, and John Earman, 731–922. Amsterdam: Elsevier.

Huggett, Nick. 2000. Philosophical foundations of quantum field theory. *The British Journal for the Philosophy Science* 51: 617–637.

Huggett, Nick, and Robert Weingard. 1996. Critical review: Paul Teller's interpretive introduction to quantum field theory. *Philosophy of Science* 63: 302–314.

Jammer, Max. 1966. *The Conceptual Development of Quantum Mechanics*. New York: McGraw Hill.

Kuhlmann, Meinard. 2010. *The Ultimate Constituents of the Material World - In Search of an Ontology for Fundamental Physics*. Frankfurt: ontos Verlag.

Kuhlmann, Meinard (2012). Quantum Field Theory. In: E.N. Zalta (ed.). *The Stanford Encyclopedia of Philosophy* (Winter 2012).

Kuhlmann, Meinard, Holger Lyre, and Andrew Wayne (eds.). 2002. *Ontological Aspects of Quantum Field Theory*. New Jersey: World Scientific.

Ladyman, James. 1998. What is structural realism? *Studies in History and Philosophy of Science* 29: 409–424.

Lyre, Holger. 2012. Symmetrien, Strukturen, Realismus. In *Philosophie der Physik*, ed. M. Esfeld, 368–389. Frankfurt: Suhrkamp.

Malament, David. 1996. In defense of dogma: Why there cannot be a relativistic quantum mechanics of (localizable) particles. In *Perspectives on Quantum Reality: Non-Relativistic, Relativistic, and Field-Theoretic*, ed. R. Clifton, 1–10. The University of Western Ontario Series in Philosophy of Science. Dordrecht: Kluwer Academic Publishers.

Mandl, Franz, and Graham Shaw. 2010. *Quantum Field Theory*, 2nd ed. Chichester: Wiley.

Maurin, Anna-Sofia (2013). Tropes. In: E. N. Zalta (ed.). *The Stanford Encyclopedia of Philosophy* (Fall 2013).

Morganti, Matteo. 2009. Tropes and physics. *Grazer Philosophische Studien* 78: 185–205.

Peskin, Michael E., and Daniel V. Schroeder. 1995. *Introduction to Quantum Field Theory*. Cambridge MA: Perseus Books.

Redhead, Michael L.G. 1988. A philosopher looks at quantum field theory. In *Brown and Harré* 1988: 9–23.

Redhead, Michael L.G. 1995. More ado about nothing. *Foundations of Physics* 25: 123–137.

Reichenbach, Hans (1955). Die philosophische Bedeutung der Wellen-Korpuskel-Dualität. In *Louis de Broglie und die Physiker*, 79–94. Hamburg: Claassen Verlag. [Engl. translation: The philosophical significance of wave-particle dualism. In *Selected Writings: 1909–1953*, 279-289 (ed. by M. Reichenbach and R.S. Cohen). Dordrecht-Boston 1978: Reidel.]

Rossanese, Emanuele. 2013. Trope ontology and algebraic quantum field theory: An evaluation of Kuhlmann's proposal. *Studies In History and Philosophy of Modern Physics* 44: 417–423.

Ruetsche, Laura. 2003. A matter of degree: Putting unitary inequivalence to work. *Philosophy of Science* 70: 1329–1342.

Ruetsche, Laura. 2011. *Interpreting Quantum theories: The Art of the Possible*. Oxford: Oxford University Press.

Ryder, Lewis H. 1996. *Quantum Field Theory*, 2nd ed. Cambridge: Cambridge University Press.

Simons, Peter M. 1994. Particulars in particular clothing: Three trope theories of substance. *Philosophy and Phenomenological Research* LIV (3): 553–575.

Srednicki, Mark. 2007. *Quantum Field Theory*. Cambridge: Cambridge University Press.

Stachel, John. 2002. 'The relations between things' versus 'the things between relations': The deeper meaning of the hole argument. In *Reading Natural Philosophy: Essays in the History and Philosophy of Science and Mathematics*, ed. D.B. Malament, 231–266. La Salle: Open Court.

Teller, Paul. 1995. *An Interpretive Introduction to Quantum Field Theory*. Princeton: Princeton University Press.

Wayne, Andrew. 2008. A trope-bundle ontology for field theory. In *The Ontology of Spacetime II. Philosophy and Foundations of Physics 4,* ed. D. Dieks, 1–15. Amsterdam: Elsevier.

Wigner, Eugene P. 1939. On unitary representations of the inhomogeneous Lorentz group. *Annals of Mathematics* 40: 149–204.

Wüthrich, Adrian. 2012. Zur Anwendung und Interpretation der Feynman-Diagramme. In *Philosophie der Physik*, ed. M. Esfeld, 227–246. Frankfurt: Suhrkamp.

Chapter 7
Chronology and Outlook

Cord Friebe, Meinard Kuhlmann and Holger Lyre

The following chronology places emphasis on the basics and on interpretations of quantum physics; it should not be considered to be a history of quantum physics as a whole. In particular, the special developments within quantum field theory, and the advances in particle physics which grew out of them, are not included. However, along with the interpretations which were given a detailed treatment in earlier chapters (Copenhagen, GRW, Everett, Bohm), a number of other approaches are briefly mentioned here. They could not be treated in detail in the rest of the book. These include

- Consciousness-Causes-Collapse Interpretation (from 1939)
- Modal Interpretations (from 1972)
- Consistent Histories (from 1984)
- Transactional Interpretation (1986)
- Relational Quantum Mechanics (from 1994)
- Epistemic Interpretations and the Quantum-Bayesian Interpretation (from 2000)
- Information-based interpretations (from 2003).

All together the development of quantum physics can be roughly divided into three phases: In the early phase, no complete theory was available. Instead, there was a mixture of models which combined new elements with classical results. This early phase began in 1900 with Planck's quantum hypothesis and ended around 1925. Thereafter, a phase of breakthroughs and the establishment of a consistent new theory began. This was quantum mechanics in the modern sense, and a corresponding mathematical formalism was developed in parallel. This phase ended around 1935 with Bohr's answer to the challenge presented by the EPR thought experiment. This marked also the tentative end of the philosophical interpretation debate around the newly developed theory, which had continued since the middle of the 1920s, after the *Copenhagen interpretation*, propagated by Bohr, Heisenberg, and other leading quantum physicists, was accepted by the majority of physicists—at least for the time

C. Friebe et al., *The Philosophy of Quantum Physics*,
https://doi.org/10.1007/978-3-319-78356-7_7

being, but in fact continuing up to the present. Since the end of the 1930s, or at the latest in the course of the 1940s, the evolution of quantum mechanics began a third phase. It is characterized, on the one hand, by newer theoretical developments such as relativistic quantum mechanics, quantum field theory and quantum gravity, but on the other, also by important experimental verifications of the basic assumptions of quantum mechanics, by the development of innovative modern fields such as quantum information theory, and finally by the establishment of alternative approaches to the interpretation of quantum mechanics.

7.1 The Early Period of Quantum Physics

1900: Planck's Quantum Hypothesis. In order to describe the thermal radiation spectrum of a black body, Max Planck suggests a formula which is based—as we now understand it—on the assumption that a body emits radiation energy only in small portions, as so-called quanta. The full implications of this assumption were recognized only some years later, by Einstein. Planck's work however led to the introduction of the energy quantum $h\nu$, containing the new universal constant which he had suggested already in 1899, "Planck's constant" or "Planck's quantum of action" h.

1905: Einstein's Light-Quantum Hypothesis. During his *annus mirabilis*, Albert Einstein established not only his theory of special relativity and his microscopic theory of diffusion. Based on the assumption that light energy (i. e. the energy of electromagnetic radiation) can be transported only in small portions (quanta), he was able to explain the experimental findings of the photoelectric effect. In 1921, he received the Nobel Prize in physics for this work. It would be misleading, however, to identify the "light quanta" of 1905 with the photons of quantum electrodynamics (QED), which was developed only decades later.

1911–13: Bohr's Model of the Atom. In 1911, Ernest Rutherford discovered experimentally that atoms consist of a positively charged nucleus (small but massive) and outer "shells" of electrons. According to the laws of classical physics, the electrons circling the nucleus would emit radiation, and as a result, atoms would necessarily be unstable. Niels Bohr "solved" this problem in 1913 with his *ad hoc* assumption of quantization of the energies of the electrons and the postulate of the corresponding "allowed" orbits on which they can circle the nucleus without emitting radiation. When the electrons release or take up radiant energy, they can jump between the allowed orbits, whose energy differences correspond to the observed spectral lines of the emitted or absorbed radiation. The model can explain much of the observed experimental data for hydrogen, but it contains an unexplained conflict with classical electrodynamics, and it cannot be applied to more complex atoms.

1916: The Need for Quantum Corrections to the Theory of Gravitation. One year after the publication of his General Relativity Theory (GRT) Einstein points out

the need to modify it in order to take quantum effects into account. Up to the present, gravitation is the only one of the four fundamental forces for which there is no accepted quantum theory. The unification of General Relativity Theory and quantum physics, still not accomplished today, is the greatest gap remaining in contemporary physics (compare the entries for 1949–57 and 1967).

1922–23: The Compton Effect and its Explanation. In 1922, the American physicist Arthur H. Compton discovers that X-rays which have scattered from electrons suffer a change in their wavelengths. In the following year, he explains this effect using energy and momentum conservation in a collision between the radiation quanta, acting as particles, and the electrons. This was received as a confirmation of Einstein's light-quantum hypothesis.

1922–25: The Discovery of the Electron's Spin. In 1922, Otto Stern and Walther Gerlach pass silver atoms through an inhomogeneous magnetic field and observe a splitting of the atomic beam. In order to explain this effect (and other spectroscopic data as well as the Einstein–de Haas effect), Samuel Goudsmit and George Uhlenbeck postulate, in 1925, a completely new property of electrons, which has no classical analogue: the (electronic) *spin*.

1923: De Broglie's Matter Waves. The radiationless orbital motion of the electrons within atoms remains mysterious. In general, it is questionable whether quantum particles can move on classical orbits at all. In his dissertation, Louis de Broglie suggests that (bound) electrons and other particles can be understood as standing waves with discrete frequencies. Thus, not only do waves have particle character, but also particles have wave character.

1924: Bose–Einstein Statistics. Satyendranath Bose and Albert Einstein introduce a probability distribution for the quantum statistics of bosons which differs from classical particle statistics (cf. Sect. 3.1.3).

7.2 Establishing Standard Quantum Mechanics

1925: Heisenberg's Matrix Mechanics. Werner Heisenberg finds an algebraic approach to quantum mechanics in which measurable quantities such as position and momentum can be represented as matrices (mathematical objects whose multiplication is non-commutative). Because of this transition to a new "kinematics", classical concepts such as "position" and "particle trajectory" become obsolete.

1925: Pauli's Exclusion Principle. Wolfgang Pauli states the requirement that no two electrons within an atom can have the same values of all four of their quantum numbers. The structure of the periodic table becomes much more transparent due to this rule (cf. Sect. 3.1.3).

1926: Schrödinger's Wave Mechanics. Erwin Schrödinger establishes the fundamental equation of quantum mechanics: a linear, deterministic and time-reversible

wave equation for the wavefunction $\Psi(x,t)$. The interpretation of this equation as the time evolution of a delocalized matter field is unsuccessful, however, since the function for N particles is defined in a $3N$-dimensional configuration space and not in position space.

1926: Quantum Hydrodynamics. Erwin Madelung reformulates Schrödinger's wave mechanics in a hydrodynamic form, putting a pair of quantum Euler equations at its centre instead of the Schrödinger equation. Bohm's quantum theory, developed later, has some similarities to this approach (cf. Sect. 5.1).

1926: Born's Probability Interpretation. Max Born interprets Schrödinger's wavefunction in a new, apparently anti-realistic manner. According to his interpretation, $\Psi(x,t)$ does not represent any material quantity, but rather it simply allows probability statements to be made. In the generalization suggested by Pauli in the same year, the absolute square $\int_X dx|\Psi(x,t)|^2$ expresses the *probability* of finding the particle within a particular region X. "Born's rule" becomes a basic element for the practical application of quantum theory (cf. Sect. 2.1).

1927: Heisenberg's Uncertainty Relations. Heisenberg demonstrates in the framework of his uncertainty or indeterminacy relations that, simultaneously, canonically conjugate quantities (i. e. position and momentum) are not sharply measurable in quantum mechanics.

1927: The Solvay Conference of 1927. 1927 is considered to be a key year for the establishment of the Copenhagen interpretation of quantum mechanics. With knowledge of the equivalence of wave and matrix mechanics—and thus a theory which had been freed of inconsistencies—a confrontation between Einstein and Bohr takes place during the 5th Solvay Conference, and is settled in favour of Bohr, in the opinion of the majority of physicists.

From 1927: Beginnings of Quantum Field Theory (QFT). In the late 1920s and early 1930s, Jordan, Heisenberg and Pauli as well as Dirac carry out pioneering work on the emerging question of the quantization of fields.

1928: Dirac's Relativistic Equation. Paul A.M. Dirac formulates a relativistic version of the Schrödinger equation for spin-$\frac{1}{2}$ particles: a milestone along the way to the generalization of quantum mechanics which also leads to the successful prediction of antimatter (positrons).

From 1928: Statistical or Ensemble Interpretations. Here, we are dealing with a whole cluster of interpretative approaches, whose precise authorship and chronology are difficult to reconstruct. At its centre stands the viewpoint that the quantum-mechanical formalism should not be applied to single systems, but only to ensembles in the sense of a statistical theory (which obeys Born's rule). Einstein and Schrödinger are early proponents of this direction. Karl Popper's Propensity Approach, interpreting probabilities as "realization tendencies", can be roughly attributed to this group of interpretations, as can the school of collectivistic interpretations of quantum mechanics inspired by Marxist thought and represented by Dimitri Blochinzev starting in the

late 1940s. Beginning in the 1950s, works by Günther Ludwig, and from the 1970s by Leslie E. Ballentine, each led to their own ensemble-theoretical programmes.

1929: Weyl's Gauge Principle. In 1929, Hermann Weyl shows in an article, in which he introduces tetrads and Weyl spinors, that the freedom in the choice of the local phase of the wavefunction, which he called "gauge freedom", leads to a term in the free Schrödinger equation which can be seen as a coupling term to the electromagnetic field. Invariance under local gauge symmetry thus leads to a minimal coupling.

1932: von Neumann's Standard Formalism. Johann von Neumann publishes his book *Mathematische Grundlagen der Quantummechanik* (Mathematical Foundations of Quantum Mechanics). It is generally recognized that the development of the new quantum mechanics attained, for the time being, its theoretical and mathematical completion with this publication. Von Neumann's book also contains the remark that there are two temporal processes for the wavefunction in quantum mechanics, namely, on the one hand, a continuous and deterministic evolution as a result of the Schrödinger dynamics and, on the other, a discontinuous collapse. He also derives a no-go-theorem which is intended to identify quantum mechanics as a complete theory and thus denies the possibility of a theory with "hidden variables". Later, John Bell in particular was able to show that the preconditions for this theorem are too specific for such a general conclusion.

1935: Schrödinger's Cat. The new paradigm has stubborn opponents: In 1935, Schrödinger publishes an essay in which he attacks the Copenhagen interpretation. If the decay of a single radioactive atom is coupled to a macroscopic trigger, then the microscopic superposition is transferred to the macroscopic realm. A cat could then for example be in a superposition state of "dead" and 'alive', a scenario that Schrödinger called "burlesque".

1935: The EPR Argument. Albert Einstein also remains an opponent of quantum mechanics. Together with Boris Podolsky and Nathan Rosen, he publishes a thought experiment which directs attention to compound systems and is intended to cast doubt on the completeness of quantum mechanics (Compare Sect. 4.2).

7.3 Confirmation and New Challenges

From 1936: Quantum Logic. Garrett Birkhoff and John von Neumann show that the logical structure of quantum mechanics does not correspond to that of classical logic. While the set of statements from classical two-valued logic forms a (distributive) Boolean lattice, the set of projectors or subspaces of the Hilbert space is a non-distributive lattice which is in particular characterized by the property that within it, *tertium non datur* is not valid. The work of Birkhoff and von Neumann is deemed to be the birthplace of quantum logic.

From 1936: Algebraic Quantum Mechanics (AQM). Von Neumann and, in the 1940s, also Gelfand, Neumark and Segal develop AQM. It represents a third mathematical formulation of quantum mechanics, in addition to the standard formulation in Hilbert space and quantum logic. It is not very useful in practice, however. Its basic idea is that a quantum system can be characterized in principle by the set of its observables and thus by an observable algebra. Physical states can be defined as linear functionals on the algebra. The so-called GNS construction guarantees that each state functional permits a representation of the algebra in the form of operators on the Hilbert space. Later, AQFT is based upon these fundamental ideas (compare the entry for 1956: AQFT).

From 1939: Consciousness-Causes-Collapse Interpretation. In von Neumann's theory of the measurement process (compare the entry for 1932), the measurement apparatus, too, is described quantum mechanically. This gives rise to the problem that the appearance of a unique measured value requires a discontinuous change of state (the "collapse" of the wavefunction). Von Neumann leaves open the possibility that this change of state occurs only through the perception of the observer, i.e. through the observer's consciousness. Fritz London and Edmund Bauer develop this concept into a theory of measurements, according to which the collapse occurs via the interaction of the physical system with the observer who is gifted with consciousness, since he or she has the special ability to uniquely determine his/her own state. Their position is influenced by philosophical tenets which originate with a substantive dualism of mind and matter. In a similar manner, Eugene Wigner also attributes a special role to the consciousness of the observer in the collapse of the state during a measurement (compare the entry for 1961). In more recent times, Henry Stapp has suggested a role for mental states (brain states). These approaches have received little resonance from the scientific and philosophical communities, however. One reason is that their basic assumptions on the relation of matter and consciousness are philosophically problematic, and the characterizations of the measurement process based on them have little explanatory value and appear to be rather *ad hoc* in nature.

From 1940: Quantum Electrodynamics (QED). In the 1940s, Richard Feynman, Julian Schwinger and Shinichiro Tomonaga develop quantum electrodynamics (QED), a quantum field theory which can be formulated gauge-theoretically and relativistically (compare the entries for 1964: QCD; and 1967: GSW).

1948: Feynman's Path-Integral Method. In the canonical formulation of QFT (compare Sect. 6.3.1), fields are described quite analogously to the usual quantum mechanics by non-commuting operators and are "quantized" in this manner. Richard Feynman finds a different transition from classical physics to quantum theory, and thus a formulation which dispenses with operators and the Hilbert space. In his formulation, integral expressions occupy a central position; they can be understood as integrals over all the possible particle trajectories. This path-integral method is also called the *sum over paths* or the *sum over histories* approach. Although classical particle trajectories may have played an heuristic role in the construction of the theory, Feynman's path-integral method cannot be considered a return to a particle

ontology, however. Feynman himself understood his approach to be primarily a mathematical instrument. It yields the same results for the transition matrix elements (e.g. in scattering theory) as the standard formulation. The two variants of the theory are thus equivalent on the level of observations. Path integrals have advantages for taking relativistic considerations into account explicitly (compare the entry for 1949–57) and for semi-classical approximations. In many applications of the path-integral method, Feynman diagrams play an important role (compare the entry for 1949) in the calculation of transition probabilities. In principle, however, these represent two separate methods.

1949: Feynman Diagrams. Feynman introduces within the framework of his path-integral approach (compare the entry for 1948) a highly effective method for ordering the probabilities with which particular initial states in scattering processes transform into particular final states: *Feynman diagrams* are a type of "graphical stenography" which relates individual elements to expressions that can be computed in a perturbation calculation. They are a graphical aid for finding all the relevant perturbation-theoretical contributions and for their calculation. Their function is not, however, to make fundamental processes intuitively understandable, as asserted by a widespread misunderstanding. Although Feynman diagrams were introduced in connection with the path-integral approach, they are not limited to this original application. They can serve the same purpose also in scattering theory within the standard formulation of QFT. (Compare Sect. 6.3.4.)

1949–57: First Theories of Quantum Gravity. Following diverse preliminary work in the 1930's and 1940's, approaches for a quantum theory of gravitation were formulated that are still the most important ones. The challenge for theories of quantum gravitation consists in the fact that gravitation within General Relativity Theory (GRT) is not a force *within* space and time, but rather the curvature *of* space and time themselves. A quantization of gravitation might thus imply a quantization of space and time, and it is not clear just what that would mean. In principle, there are four different methods of combining GRT and QFT: (i) GRT is quantized; (ii) QFT is "generally relativized"; (iii) one of the theories is a limiting case of the other; or (iv) both GRT and QFT are limiting cases of a fundamentally new theory. It is usual to refer to (i) as the canonical approach (due to the *canonical* quantization (compare Sect. 6.3.1), and to (ii) as the covariant approach (since the *covariance* of GRT is its starting point). These terms are somewhat misleading, however, and the categories are not always clearly distinguished. A further important approach, also formulated during this period, is the application of Feynman's path-integral method of quantization to gravitation; it is also known as the *sum over histories* approach (compare the entry for 1948). Today, the most important canonical approach is quantum loop gravity (compare the entry for 1986). The covariant approach converged for the most part with string theory in the end (compare the entry for 1987).

1952: Bohm's Deterministic Quantum Mechanics (de Broglie–Bohm Theory or Bohm's Mechanics). David Bohm develops a deterministic theory, competing with standard quantum mechanics, in which particles move along trajectories again,

"led" by the wavefunction Ψ which forms a guidance field. The French physicist de Broglie had formulated a similar theory already in the 1920's (which was not known to Bohm, however, when he developed his theory) (compare Sect. 5.1).

1954: Yang–Mills Theories. Chen Ning Yang and Robert L. Mills extend the concept of gauge theories (compare the entry for 1929: Weyl) to non-Abelian unitary symmetry groups. The later electro-weak unification and QCD (quantum chromodynamics) fall in this category (compare the entries for 1964: QCD; and 1967: GSW).

From 1956: Axiomatic and Algebraic Quantum Field Theory (AQFT). Partly due to the difficulties concerning infinities in the conventional formulation of QFT, various programmes start for reformulating QFT axiomatically in a mathematically rigorous and transparent manner. Following Wightman's axiomatics from the early 1950s, Daniel Kastler, Huzihiro Araki and, most enduringly, Rudolf Haag introduce the most successful axiomatically oriented approach, known as "AQFT". From the 1980s onwards, it was elaborated by, among others, Klaus Fredenhagen and Detlev Buchholz. While it proved that the original intention of a purely axiomatic formulation could not be carried out, AQFT remains particularly useful when one is considering fundamental questions such as the compatibility of localizability and relativistic invariance, or the significance of non-equivalent representations of systems with infinitely many degrees of freedom (see Chap. 6).

1957: Everett's Many-Worlds Interpretation. Hugh Everett III introduces his *Relative-state* formulation of quantum mechanics, which is popularized as the "Many-Worlds Interpretation". The latter considers a pure Schrödinger dynamics and dispenses with any sort of collapse (compare Sect. 5.2).

1959: The Aharonov–Bohm Effect. Yakir Aharonov and David Bohm discover a non-local effect which implies that the electromagnetic vector potential is more than just an ancillary mathematical construction. More precisely: The loop integral of the potential yields the experimentally observable shift of an interference pattern.

1961: "Wigner's Friend". In a thought experiment, Eugene Wigner attempts to show that the introduction of a second observer ("*Wigner's friend*") demonstrates that the standard formulation cannot specify where the "Heisenberg cut" between the observer and the object of the measurement is located, and thus when and where the collapse occurs. For example, his friend could have carried out a Schrödinger's cat experiment in the laboratory and have already found a particular result. Outside the laboratory, however, Wigner must consider the quantum-mechanical superposition of two states of the combined system "friend + cat". Since Wigner holds it to be obvious that his friend has already found the cat to be either dead or alive, he argues that this difficulty can be solved only if we assume, in accordance with a consciousness-causes-collapse interpretation (compare the entry for 1939), that the human consciousness gives rise to the collapse.

1964: Bell's Theorem and Bell's Inequalities. John S. Bell demonstrates the non-local character of quantum theory through a general consideration of the correlations observed in EPR-style experiments. Under the precondition that a local explanation

is possible, the probability distributions for this measurement must fulfil a certain inequality. Within quantum theory, however, this inequality is violated, and this could be confirmed experimentally (compare the entry for 1982; see also Sect. 4.3).

1964: Quarks and Quantum Chromodynamics (QCD). Murray Gell-Mann introduces the Quark Model in 1964; it makes an essential contribution to the development of QCD, a gauge-theoretical quantum field theory of the strong nuclear force.

1967: The Glashow–Salam–Weinberg Theory (GSW). The electro-weak theory of GSW accomplishes the unification of the electromagnetic and the weak interactions within the framework of a gauge-theoretical quantum field theory (compare the entry for 1954: Yang–Mills).

1967: The Wheeler–DeWitt Equation. A milestone for attempts to achieve a canonical quantization of General Relativity Theory (compare the entry for 1949–57) is the Wheeler–DeWitt equation, which was originally called the "Einstein–Schrödinger equation". It expresses a condition that must be fulfilled by all wavefunctions, namely the equivalence of all possible coordinate systems.

1967: The Kochen–Specker Theorem. Simon Kochen and Ernst Specker present an argument in favour of the completeness of standard quantum mechanics—thus, a no-go-theorem for hidden variables. According to this result, quantum systems must possess certain properties such as spin in a manner which is not context independent, i.e. not independent of *how* a measurement is carried out.

1970: Decoherence. H. Dieter Zeh gives the impetus to a project which is successful only in the 1990s, however. The basic idea of decoherence is that taking its environment into account, the state of a system which is comprised of a measurement apparatus and a microscopic object will pass from the unwanted superposition to a mixed state that no longer contains interference terms (and is thus "decoherent") (compare Sect. 2.3.2).

From 1972: Modal Interpretations. Bas van Fraassen suggests an interpretation of quantum mechanics which he continues to develop up to 1991. The principal goal of this approach is to solve the problems that occur in particular in connection with the quantum-mechanical measurement process. Modal interpretations, like Bohm's quantum mechanics and the Many-Worlds interpretation, assume that there is no collapse of the wavefunction. The projection postulate, according to which the state of the measured object collapses suddenly into the (eigen-) state corresponding to the measured value, is thus rejected. A crucial idea of van Fraassen's modal approach is distinguishing between the "dynamic state" and the "value state". While the dynamic state describes what *could* be measured, the value state determines what is *in fact* the case, that is: which physical properties are obtained sharply. The dynamic state is the usual Hilbert-space state of quantum mechanics, which evolves always in agreement with the Schrödinger equation. A decisive idea of this modal interpretation is now that the physical properties of a system can have sharp values *without* the dynamic state being an eigenstate of the corresponding observables. In the following decades,

a series of alternative modal interpretations were formulated, in particular by Simon Kochen, Dennis Dieks and Richard Healey.

1974 and 1976: Hawking Radiation and the Unruh Effect. In 1974, Stephen Hawking calculates the radiation emitted by black holes in the framework of QFT in curved spaces. Since, due to the extremely high mass density of black holes, spacetime attains an infinite curvature (a singularity), his work also has a strong influence on the development of theories of quantum gravitation, where there is hope that the singularities of GRT can be avoided or removed. Hawking further develops Feynman's "sum over histories" approach and uses it for his calculations (compare the entry for 1948). At the same time, with his work he opens up a rather fruitful new field: the thermodynamics of black holes. Within this framework, in 1976 William Unruh discovers what will later be called the "Unruh effect", and which seems to undermine a particle interpretation of QFT (cf. Sects. 6.3.5, 6.4.2 and 6.5.2).

1977: The Quantum Zeno Effect. Misra and Sudarshan predict an effect in which the decay rate of a quantum system can be drastically reduced through continuous measurements. Thus, the system can be almost "frozen" (in analogy to Zeno's motion paradoxes).

1978: Delayed Choice Experiments. John A. Wheeler sharpens the double-slit paradox by "delayed choice", in which the experimenter can decide only *after* a particle has passed through the double slit (in a classical sense) whether and what will be measured. In 1984, a corresponding experiment is carried out, and it confirms the quantum-mechanical predictions.

From 1980: Quantum Computing. Driven by the early theoretical publications of Yuri Manin, Richard P. Feynman, Charles H. Bennett, Paul A. Benioff and David Deutsch, during the 1980 s the concepts of quantum information and quantum computing become mature, along with the resulting possibilities for specialized quantum computations. Beginning in the 1990's, quantum information theory and quantum computer science continue to develop, increasingly in terms of experiments as well as theory.

1982: The Aspect Experiment. Alain Aspect *et al.* for the first time show experimentally that Bell's inequality is violated (compare the entry for 1964).

1982: The No-cloning Theorem. Wootters and Zurek, as well as Dieks, are able to show that quantum states cannot be copied or "cloned".

From 1984: Consistent-Histories Interpretation. An interpretation which is often considered to be similar to the Copenhagen interpretation of quantum mechanics is the *consistent-histories* interpretation. It was introduced in 1984 by Robert Griffiths and in the following years was developed further by him as well as by Roland Omnés, Murray Gell-Mann and Jim Hartle. The basic idea is somewhat similar to Feynman's "sum over histories" approach (compare the entry for 1948), in that the dynamics of physical systems is traced back to "consistent histories". Here, histories are temporally ordered series of events which are associated in a consistent

manner to certain probabilities, i. e. in particular without coming into conflict with the Schrödinger equation. The requirement of (essentially) interference-free histories leads to the so-called decoherent histories. The consistent-histories interpretation dispenses both with a collapse of the state and with any kind of description of the quantum-mechanical measurement process. Today, this interpretation plays a role especially in quantum cosmology.

1984: Berry's Geometrical Phase. Michael Berry discovers a geometrical aspect of the quantum-mechanical wavefunction. It has the effect that a quantum system which experiences a cyclic evolution in the state space due to its dynamics also exhibits an observable holonomy in the phase of its wavefunction.

1984: Quantum Cryptography. Charles H. Bennett and Gilles Brassard introduce a protocol for the exchange of quantum keys. According to them, quantum theory leads to encoding techniques which are in principle tap-proof.

From 1986: GRW: Spontaneous Collapse Theory. Giancarlo Ghirardi, Alberto Rimini and Tullio Weber replace the Schrödinger equation by a nonlinear, non-deterministic and temporally irreversible equation, which implies a real, spontaneous collapse mechanism for the wavefunction (compare Sect. 2.4).

1986: Transaction Interpretation. According to John Cramer, the quantum-mechanical wavefunction is a real, physical wave which obeys relativistic quantum mechanics and propagates both in the form of outbound, "retarded" waves towards the future, and in the form of incident, "advanced" waves towards the past. This brings about transactions ("handshakes") which then constitute quantum events. Cramer's interpretation is based on the Wheeler–Feynmann absorber theory of electromagnetic radiation (1945).

From 1986: Loop Quantum Gravity. Loop quantum gravity becomes the most important canonical approach to a theory of quantum gravitation (compare the entry for 1949–57) and remains today the principal competitor of string theory (compare the entry for 1987). The Indian physicist Abhay Ashtekar makes an important contribution towards establishing loop quantum gravity in 1986 through a new choice of variables. With this method, Lee Smolin and others soon discover the so-called Wilson loops as exact solutions to the Wheeler–DeWitt equation (compare the entry for 1967). An essential assumption of loop quantum gravity is the so-called background independence, which asserts that spacetime is not simply a background, but rather, on the microscopic scale, itself a dynamic entity. Spacetime is thus not presupposed, but instead it emerges in a certain manner. In contrast, string theory is not (manifestly) background independent, since it works within a given spacetime.

From 1987: String Theory. String theory, already proposed in the 1960 s in connection with QCD, experiences a strongly increased interest in the late 1980 s as a candidate for a unified theory of QFT and gravitation. String theory becomes the most important covariant approach to a theory of quantum gravitation (compare the entry for 1949–57), whereby among others perturbation calculations of certain scattering amplitudes lead to a breakthrough. The basic idea of string theory is that not particles,

but rather small vibrating one-dimensional "strings" or filaments are assumed to be the most fundamental objects. A decisive advantage of this approach is that strings do not have point-like interactions with each other, so that certain infinities can be avoided, which cause problems already for conventional QFT and even more so for the attempted unification with gravitation. In 1995, Edward Witten suggests a more comprehensive theory, the so-called M theory, to which the existing string theories are only approximations.

1993: Quantum Teleportation. Bennett *et al.* show in 1993 that quantum states can be transported to distant locations. Along with a quantum channel based on an entangled system (e.g. an EPR pair), a classical channel of information transmission is also required. In 1997, Zeilinger *et al.* implement the corresponding experiments using photons for the first time.

From 1994: Relational Quantum Mechanics. According to Carlo Rovelli's relational interpretation, the states of a quantum system have no absolute significance; they are important only in relation to another system (analogous to the frame dependence in relativity theory).

1994: The Shor Algorithm. Peter Shor develops a quantum algorithm for decomposing prime factors in polynomial time (compare the entry for 1980). His work has an enormously stimulating effect on further developments, especially on the experimental aspects of the newly established field of quantum computing, beginning around the year 2000.

From 2000: Epistemic Interpretations and Quantum-Bayesian Interpretations. The Quantum-Bayesian interpretation (most energetically advanced by Christopher Fuchs) formulates an explicitly subjectivist or epistemological interpretation, according to which the quantum-mechanical probabilities are due not to an objective indeterminacy in the world, but instead to our limited knowledge. The associated probabilities are not arbitrary, but rather they fulfil the minimal conditions of the subjectivist concept of probability developed by Thomas Bayes. According to the quantum-Bayesian interpretation, the collapse of the wavefunction does not describe a physical process of the object considered. Instead, it indicates the updating of our knowledge. Further important epistemological interpretations of quantum mechanics are due to David Mermin and Richard Healey, among others.

From 2003: The CBH Theorem and the Information-Based Interpretation. Robert **C**lifton, Jeffrey **B**ub and Hans **H**alvorson show in the now so-called CBH theorem that the algebraic structure of quantum mechanics (compare the entry for 1936: AQM) can be derived from three information-theoretical assumptions concerning the impossibility of superluminal transmission of information, the copying of an unknown state and a special transfer protocol. In the opinion of Bub, one should therefore interpret quantum mechanics as a theory of the manipulation of *information*, which is considered as the physically fundamental quantity.

Appendix A
Example Solutions for the Exercises

Chapter 1

1. *Niels Bohr introduced the concept of "complementarity" into the interpretation of quantum mechanics. Distinguish two readings of how it is to be understood.*

Concepts which are mutually exclusive in classical physics complement each other in quantum mechanics (wave/particle dualism) versus quantities which complement each other classically but are mutually exclusive in a precise sense in quantum mechanics (non-commuting observables such as position/momentum or spins along different directions).

2. *In sequential spin measurements, we apparently distinguished finally between two effects which could be reversed by mixing the particles. Describe these two ostensible effects and explain why they are in fact only a single effect. What can we deduce from this?*

In Fig. 1.5, particles with a well-defined spin value along the y direction pass into the apparatus. When they go through the (first) $\hat{S}_x$ filter (a Stern–Gerlach magnet), on the one hand, a definite spin value along the x direction is evidently *produced*; on the other hand, the definite spin value along the y direction is *destroyed*. Both effects could be verified by spin measurements along the paths of the particles, if one were to carry them out. The apparently *two* effects, however, cannot be separately reversed: Surprisingly, all the particles that leave the apparatus again have their original spin values. From this, one deduces a principle which states that (referring to the present example) "The state of a particle with a definite spin value along a particular direction is none other than the *superposition* of opposite spin values along other directions".

3. *Consider the expectation values of operators in regard to whether the physical system is represented by an eigenvector of the given operator or not. Compute the expectation values of the spin operators discussed earlier, relative to the various vectors described there. Explain the results by referring to the figures showing a repeated and a destructive measurement.*

Example: Spin operator (1.8): Let the state correspond initially to one of its eigenvectors (1.11). The expectation value (1.18) of a repeated measurement (compare Fig. 1.2) is then identical to the associated eigenvalue, and the vari-

C. Friebe et al., *The Philosophy of Quantum Physics*,
https://doi.org/10.1007/978-3-319-78356-7

ance around that value is 0. Here is the calculation (it is analogous for the other eigenvector):

$$\langle EV_1|\hat{S}|EV_1\rangle = \begin{pmatrix}1/\sqrt{2} & -i/\sqrt{2}\end{pmatrix}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix}1/\sqrt{2}\\ i/\sqrt{2}\end{pmatrix}$$
$$= \begin{pmatrix}1/\sqrt{2} & -i/\sqrt{2}\end{pmatrix}\begin{pmatrix}1/\sqrt{2}\\ i/\sqrt{2}\end{pmatrix} = 1\,,$$

$$\triangle\hat{S} = \langle EV_1|\hat{S^2}|EV_1\rangle - |\langle EV_1|\hat{S}|EV_1\rangle|^2 = 1 - 1 = 0\,.$$

Now, we represent the state in terms of one of the basis vectors from Fig. 1.11. The expectation value is now no longer identical to one of the eigenvalues, and the variance is also no longer 0. This process or measurement from which the new state emerged has destroyed the previously definite spin value (cf. Fig. 1.3). Calculation (analogously for the other basis vector):

$$\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} = \begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}0\\ i\end{pmatrix} = 0\,,$$

$$\triangle\hat{S} = 1 - 0 = 1\,.$$

4. *What does von Neumann's projection postulate state? Explain in particular to what extent this postulate goes beyond what we have considered to be well established in connection with expectation values.*

It is uncontested that when the quantum-mechanical system is in a state which is represented by a particular eigenvector, then the associated eigenvalue will be measured with certainty. The converse, however, is controversial: When a particular eigenvalue has been measured, is the quantum-mechanical system (immediately after the measurement) in a state which is represented by the (or a) corresponding eigenvector? This eigenvalue–eigenvector connection is demanded by von Neumann's projection postulate; it implies the problem of wavefunction collapse.

5. *In contrast to the general opinion of many philosophical schools of thought and also of some alternative physical interpretations (e. g. GRW, Bohm), (intuitive or physical) space is not at the centre of standard quantum mechanics. Discuss this hypothesis, first informally and then by referring to the particular mathematical characteristics of the position operator.*

This is an open question related to Sect. 1.2.4.

Chapter 2

1. *Distinguish between two readings of Born's rule, depending upon whether the reference to a measurement in it is essential or not.*

If the reference to a measurement is *relevant*, the properties which correspond to the measured values are not already present before the measurement. Standard quantum mechanics would then be complete, and we would have to explain

the problem of what in fact happens during such a measurement. If the reference to the measurement is however *irrelevant*, the standard formalism would be incomplete, since then properties would be objectively present which could not be predicted with certainty. We must, however, take into account that due to EPR/Bell (cf. Chap. 4), such local properties are in no case present and that the de Broglie–Bohm theory (cf. Chap. 5) adds only *positional* properties.

2. *According to the Copenhagen interpretation (in Heisenberg's version), there are two temporal dynamics of the state vector. Describe them in your own words. How is the second dynamics related to Born's rule and to von Neumann's projection postulate? What is problematic about it?*

The first dynamic obeys the linear Schrödinger equation and is continuous, deterministic and reversible. It occurs between successive measurements. During a measurement, according to this interpretation, a second dynamic occurs, and it is discontinuous, non-deterministic and irreversible. Correspondingly, for Born's rule, the essential point is the reference to a measurement; the second dynamic is required by von Neumann's projection postulate. This is particularly problematic since there is no physical criterion for deciding where to place the boundary ("Heisenberg cut") between the observer and the observed.

3. *The interpretation problem in quantum mechanics may be regarded as a trilemma. Explain the three statements and show that they are inconsistent when taken together. What is the advantage of this description as compared to the conventional one, which is guided by Born's rule?*

The three assertions are:

(a) All objectively present properties are determined by the state vector (completeness of the standard formalism).
(b) The temporal evolution of the state vector conforms to the linear dynamics (Schrödinger equation).
(c) Measurements always yield definite measured values.

The conjunction of these three assertions is inconsistent, which was already shown by the spin measurements described in the introductory chapter. The advantage of this representation of the interpretation problem consists in the fact that in this way, all the available options (plus their variants) lie open on the table: Bohm's mechanics as the negation of the first assertion; "Copenhagen" and GRW as the negation of the second assertion; and Everett as the negation of the third.

4. *The decoherence programme is an essential step forward. Highlight the ways in which all the interpretation variants could profit from it. Why, however, can the programme not solve the measurement problem in the end?*

The decisive progress made by the decoherence programme lies in the fact that a basis—the pointer basis (as a rule position space)—is distinguished physically dynamically; this solves one part of the interpretation problem, which is advantageous to all the options. Furthermore, it can be shown that the interference terms vanish *locally*, which corresponds to our perceptions. *Globally*, however, the interference is even amplified. Even if one would wish to accept this, decoherence produces at best a (classically understandable) *either-or* and thus a continuing form of indeterminacy rather than a definite measured value.

5. *Formulate in your own words what is accomplished by the GRW theory in the opinion of its supporters. Defend the standard view of physics against it.*

The GRW theory supports a unified (nonlinear) dynamics without mysterious measurements. It prefers spatiotemporal descriptions, which corresponds to our experience. (Its objective indeterminacy could offer further advantages, for example in the discussion about the arrow of time.) The new universal constants, however, are chosen in an extremely arbitrary way, precisely such that the results of measurements are reproduced, without an independent empirical confirmation. Ordinary quantum mechanics can furthermore be modified in a well-established manner so that it is compatible with the special theory of relativity (compare Chap. 6)—in contrast, a (generally accepted) Lorentz-invariant GRW theory is not (yet) known.

Chapter 3

1. *Construct the fully symmetrized basis functions in the state space of three similar objects a, b, and c (Hint: As presented in* Sect. 3.1.4*, these are the one-dimensional irreducible representations of* S_3*).*

S_3 has the dimensionality $3! = 6$. The 6 permutations of three objects a, b, c are abc, bac, cba, acb, cab and bca. They are generated from abc by application of the 6 permutation operators (elements of S_3): $\mathbb{1}$, $\hat{P}_{12}$, $\hat{P}_{13}$, $\hat{P}_{23}$, $\hat{P}_{13}\hat{P}_{12}$ and $\hat{P}_{12}\hat{P}_{13}$. Just as the fully symmetric state (3.9) of two objects a and b results from application of the symmetrization operator $\hat{S}_2 = \mathbb{1} + \hat{P}_{12}$ of the two elements of S_2, the symmetrization operator of S_3 is:

$$\hat{S}_3 = \mathbb{1} + \hat{P}_{12} + \hat{P}_{13} + \hat{P}_{23} + \hat{P}_{13}\hat{P}_{12} + \hat{P}_{12}\hat{P}_{13}.$$

The fully symmetric basis function of S_3 follows from it and is given by

$$\Psi_S = \frac{1}{\sqrt{6}}\Big(\psi_a\psi_b\psi_c + \psi_b\psi_a\psi_c + \psi_c\psi_b\psi_a + \psi_a\psi_c\psi_b + \psi_c\psi_a\psi_b + \psi_b\psi_c\psi_a\Big).$$

In a fully antisymmetric state, the permutation or exchange of two objects causes a change of sign of the wavefunction. In the case of S_2, the antisymmetrization operator $\hat{A}_2 = \mathbb{1} - \hat{P}_{12}$ gives the function (3.7). To construct the fully antisymmetric basis function of S_3, one requires an analogous antisymmetrization operator $\hat{A}_3$, which is obtained from $\hat{S}_3$ by reversing the signs of all the $\hat{P}_{ij}$; it leads finally to the fully antisymmetric basis function

$$\Psi_A = \frac{1}{\sqrt{6}}\Big(\psi_a\psi_b\psi_c - \psi_b\psi_a\psi_c - \psi_c\psi_b\psi_a - \psi_a\psi_c\psi_b + \psi_c\psi_a\psi_b + \psi_b\psi_c\psi_a\Big).$$

2. *What is the content of the Leibniz PII in the contrapositive formulation?*

No two objects share all their properties.

3. *To what extent is a bundle ontology naturally associated with PII?*

Leibniz individuation is individuation via properties. In the case that objects

are bundles of properties, there is nothing available for the individuation of objects besides properties; therefore, bundle ontology is most naturally associated with Leibniz's PII.

4. *What do the concepts of "synchronic" and "diachronic identity" refer to?*

"Synchronic identity" refers to the identity (or distinguishability) of objects at a particular moment in time, while "diachronic identity" refers to persistence through time.

5. *Define the three types of discernability according to Quine.*

Difference in terms of at least one . . .

- monadic property ("absolute discernability");
- ordering relation ("relative discernability");
- irreflexive relation ("weak discernability").

6. *Discuss to what extent both the weak PII and structural realism lead to related object conceptions.*

This is an open question related to Sect. 3.2.4.

Chapter 4

1. *In the EPR article, there is an assumption of major significance which we have called "locality assumption": "Since at the time of measurement, the two systems no longer interact, no real change can take place in the second system in consequence of anything that may be done to the first system". How does this assumption relate to the other concepts of locality which we introduced in Sect. 4.4.5: Does it imply global and causal Einstein locality as well as spatiotemporal separability?*

The ability to distinguish between the first system and the second system indicates that Einstein requires separability. The exclusion of interactions implies causal Einstein locality. From separability and causal Einstein locality, global Einstein locality then follows.

2. *Assume that an EPR/B experiment was correctly described by a local causal structure with hidden variables λ (see Fig. 4.5). One can then show that the existence of perfect correlations implies that measurements must proceed deterministically. Try to formulate a suitable argument.*

Perfect correlations denote the fact that when the measurement direction is the same on both sides, the measurement results agree *with certainty* (50% $\alpha = +, \beta = +$ and 50% $\alpha = -, \beta = -$). A local structure claims that the measurement outcome on one side depends only on the state of the local photon and the local measurement setting.

Assume that a measurement of a photon at A with the setting $a = 0°$ gave the result $\alpha = +$ and at B the measurement direction is likewise $b = 0°$. Then, the photon at B *must* also yield $\beta = +$. Since, according to the local causal structure, the outcome at B depends only on the local setting and the state of the photon at B, according to the assumed causal structure, these two factors must determine the result. This then holds even if in fact a different setting is chosen at A (i. e. one that is not parallel to that at B), or if the measurement result is different. An analogous argument holds in reverse for the photon at A when the result at B is known.

Therefore, for each photon, only its state together with the given measurement setting *determines* the outcome of the measurement, and these outcomes must agree when the settings do. (See also Table 4.2, where possible determination schemes are given, corresponding in each case to different photon states.)

3. *List the minimal set of assumptions which are required to derive a Bell inequality and sketch out what they state.*

An overview of the required assumptions from Sect. 4.3.4:

- **Global Einstein locality**: There are no causal processes faster than light (see Sect. 4.3.1).
- **No backwards causation**: Effects cannot occur earlier than their (timelike-related) causes (see Sect. 4.5.3).
- **Intervention assumption**: The experimenter (or a device constructed by the experimenter) can control the setting of a macroscopic apparatus through an intervention (see Sect. 4.5.2).
- **Causal Markov condition**: Given its direct causes Z, a variable X is statistically independent of all variables Y which are not effects of X: $P(X|YZ) = P(X|Z)$ (see Sect. 4.3.4).

4. *Apply the causal Markov condition to the local causal structure in Fig. 4.5 and note the resulting statistical independencies.*

The Markov condition states that given its direct causes, a variable becomes independent of its non-effects. That is, if x is the variable under consideration, then $P(x|\text{direct causes of } x, \text{non-effects of } x) = P(x|\text{direct causes of } x)$. If we apply this rule to every variable in the graph, we obtain:

$$P(\lambda|ab\psi) = P(\lambda)\,, \qquad P(\psi|ab\lambda) = P(\psi)\,, \qquad P(a|b\psi\lambda) = P(a)\,,$$
$$P(b|a\psi\lambda) = P(b)\,, \quad P(\alpha|\beta ab\psi\lambda) = P(\alpha|a\psi\lambda)\,, \quad P(\beta|\alpha ab\psi\lambda) = P(\beta|b\psi\lambda)\,.$$

5. *Outline the four fields of conflict between a non-local theory and the theory of relativity.*

See Sect. 4.4.2. (and the following sections for details): In a relativistic spacetime ...

- ... transport of matter or energy faster than the speed of light is not possible.
- ... signals transmitted at superluminal velocities could lead to paradoxical loops.
- ... fundamental spacelike causal connections contradict the temporal asymmetry of causality.
- ... non-local connections prefer frames of reference and thereby violate the principle of relativity.

6. *Discuss the two known kinds of non-locality which are not in conflict with the principle of relativity. Take into account both physical and ontological consequences.*

See Sect. 4.4.7:
Hyperplane dependence:

– A completely Lorentz-invariant theory (including interactions);
– States are defined only relative to hyperplanes; i. e., at every point in spacetime, there are infinitely many states for a quantum system, namely one for each hyperplane:
  * ontologically extravagant
  * highly counterintuitive
  * What is it supposed to mean that an object can have different states at a single point in spacetime?

GRW Flash:

– At present, a Lorentz-invariant theory is available only for particles without interactions.
  * It is possible that a theory with interactions will not be Lorentz invariant; then, this would not be an appropriate solution.
– Objects are not continuously present within spacetime, but instead are represented by a rapid series of "flashes".
  * This is likewise counterintuitive!
  * Objects in spacetime are epiphenomena: A flash at time 1 does *not* cause the flash at time 2. (The quantum state in Hilbert space at time 1 causes both the flash in spacetime at time 1 and also the quantum state at time 2; the latter causes the flash at time 2.)
– The fundamental space is not four-dimensional spacetime, but instead the (possibly) infinite-dimensional Hilbert space of a quantum system.
  * How do spacetime and the Hilbert space behave? How might one conceive of the fact that objects from the latter can act within the former (but not *vice versa*)?
  * Does spacetime supervene on the Hilbert space? If so, then how exactly?

Chapter 5

1. *Describe an argument between an advocate of the de Broglie–Bohm theory and a supporter of the Copenhagen interpretation about the status of hidden variables.*

The advocate of the DBB theory could point out that it is the *wavefunction* (and not the position) that is not directly observable. Only the *statistical predictions* which can be derived from it are possible subjects of experimental tests. The wavefunction of a *single* quantum object can *in principle* not be determined. In contrast, the positions of the particles are the quantities that can be measured (even in single-particle systems)—consider, for example, the experiments with electrons and a double slit. The supporter of the Copenhagen interpretation can interject here that the quantum equilibrium condition (Sect. 5.1.2) limits our *knowledge* of the positions, again in principle—in this sense, they are therefore "hidden". At this point, the DBB supporter can reply that the status of probability statements is still problematic, even within the Copenhagen interpretation. According to this

view, they in fact do *not* refer to the properties themselves, but rather only to their measurement.

2. *Explain why within the de Broglie–Bohm theory, the uncertainty relation* $\Delta x \cdot \Delta p \geq \frac{\hbar}{2}$ *is not violated!*

Here, one can argue on several levels: (i) Fundamentally, the quantum equilibrium hypothesis guarantees that the de Broglie–Bohm theory fulfils *all* of the (as a rule statistical) predictions of quantum theory; compare Sect. 5.1.2. The Heisenberg uncertainty relations provide a connection between the variances of the measured values of these predictions, and they are therefore also valid within the DBB theory. (ii) The momentum p (in contrast to the position x) indeed does *not* have a definite value at all times within the DBB theory; instead, it is a "contextual property" (cf. Sect. 5.1.5). In this sense, there is no reason at all to assume that the position *and* the momentum should have well-defined values at all times within the DBB theory. See also Bricmont (2016, 159–161).

3. *Compare the solutions to the measurement problem in the de Broglie–Bohm and the Everett interpretations. Give examples of structural similarities and differences between them.*

Both dispense with the discontinuous changes of state of the wavefunction ("wavefunction collapse"). In both cases, all the branches of a superposition state of the wavefunction are thus present at all times. In order to nevertheless guarantee that measurements yield a definite result, the de Broglie–Bohm theory introduces additional determining quantities which are necessary to define a state uniquely: the positions of the particles. They characterize that part of the wavefunction which corresponds to the factual result. In the Everett interpretation, only the appearance of definite measurement results is explained. All the possible results in fact *do* occur—to be sure in "different worlds". In both interpretations, effects of the so-called decoherence (cf. Sect. 5.2.4) play a central role. They guarantee that the branches, once they have separated, can no longer interfere with one another. Without this property, the "empty" branches of the wavefunction could not be effectively neglected in the DBB theory. Within the Everett interpretation, a lack of decoherence would in principle endanger the closure of the different "worlds" as well as the existence of a "preferred basis" (Sect. 5.2.4).

Chapter 6

1. *Gather information about theories of light from the history of physics. Why was Newton's theory of light not considered satisfactory? Compare the mathematical descriptions of particles and of waves. What is, in your opinion, the principal difference?*

Newton, among others, advocated a corpuscular theory of the nature of light. Field theories could, however, offer a better explanation of the various diffraction and refraction effects observed, by attributing them to the interference of waves.

The state of a particle is defined by its position and its momentum; it follows a trajectory as a function of time. Waves occur in fields. To characterize a field, one must specify, e. g., its field strength at every point in spacetime.

2. *Is there anything in classical physics which is neither a particle nor a field?*

In classical physics, mechanical processes are described in terms of particles and electromagnetic processes in terms of fields. There are thus no objects within classical physics which are neither particles nor fields. However, classical physics involves numbers, which are neither particles nor waves and thus are not objects embedded in spacetime at all.

3. *State two arguments which in your opinion provide the strongest support for considering quantum field theory to be a theory of particles. What objections can be raised against these arguments?*

In experiments, mainly particle properties are detected; this can be seen, e. g., in the name "particle physics" (cf. Sect. 6.3.4). However, for the calculation of scattering cross sections, for example, the field aspects of quantum field theory are also needed. A particle interpretation seems to be supported in particular by the occupation-number representation (Sect. 6.3.2). But we should keep in mind that what is created by creation operators (or whose creation is described by them) does not have the properties of classical particles. For example, there are states in quantum field theory with an undefined particle number, which could not occur in a purely particle theory.

4. *Would it be helpful in your opinion to extend the concepts of "particles" and "fields", i. e. by referring not only to those entities as "particles" or "fields" which exhibit all the characteristics of classical particles or classical fields?*

There is nothing to prevent the objects which are described by quantum field theory from being called "particles" or "fields" in everyday laboratory jargon, especially when they have the *properties* of particles or of fields in the given situation. In considering the questions of what is the *nature* of the fundamental objects of the world, and how they are embedded in spacetime, one should, however, keep in mind that quantum field theory does not refer to classical particles nor to classical fields. What we choose to denote, e. g., as "particles" is in the end a matter of language conventions. This, however, sheds no light on the problem of how quantum objects are embedded in spacetime.

Index

C. Friebe et al., *The Philosophy of Quantum Physics*,
https://doi.org/10.1007/978-3-319-78356-7

G

H

I

J