Philosophy of Quantum Probability

An empiricist study of its formalism and logic

Ronnie Hermens

Philosophy of Quantum Probability

An empiricist study of its formalism and logic

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. E. Sterken
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

maandag 25 januari 2016 om 16.15 uur

door

Ronnie Hermens

geboren op 4 november 1981
te Nijmegen

FILOSOFIE

VAN DE

QUANTUMKANSTHEORIE

EEN EMPIRISTISCHE ONDERZOEKING VAN HAAR
FORMALISME EN LOGICA

Ronnie Hermens

This page intentionally defiled with these words

Preface

If this dissertation were in Dutch, I would call this preface "Verantwoording". But since I lack the skills to ascertain whether any of the dictionary translations of this word (Responsibility/Justification/Amenability) is an apt one, I stick to the more traditional "Preface". On the other hand, the Dutch meaning of the phrase also appears to be suffering some damage. Nowadays, when some big chief makes a mess he can just take responsibility (verantwoording afleggen) by quiting his job with a bonus pay check, leaving everybody else with the mess. It is not my intention to leave the reader with a mess though, and thus I aim to give a "verantwoording" in a more traditional sense. So what follows (among other things) is a short explanation/defense of why this dissertation is the way it is, and not something else.

It is commonplace that a dissertation is the culmination of scientific work performed over the period of four years that comprise the 'promotion'. This dissertation is not of this kind. Such a culmination may already be found in the papers I published in the past four years. Instead of stapling those papers together, I took the opportunity to try to write a story that ties together some of these papers, even though at first sight they may look somewhat disjoint. I simply think it is nicer if a dissertation at least tries to be a book with a story. A peculiar consequence is that the overarching topic of this dissertation (quantum probability) isn't the specific topic of any of the papers that present the 'main research results'. Nonetheless, this topic has been in the back of my mind a lot during much of the research I did. Perhaps it is because it was a course on quantum probability that I followed as a master student that triggered my interest in the foundations of quantum mechanics in the first place. It could also be because officially my research topic was the philosophy of probability. Either way, I took this dissertation as a chance to put some of those background ideas to the front.

The plan to write a dissertation that makes up a single coherent story even though the papers upon which it is based are not directly linked to one another has both positive and negative consequences. Let me start with some of the nega-

tive ones. First, this dissertation most likely does not answer to the expectations one might have for a dissertation on the philosophy of quantum probability. For example, there is no discussion of the literature on generalized probability spaces. Second, I wasn't able to incorporate all aspects of all my papers in this dissertation. But the upside is that this dissertation provides both a look at quantum probability from a refreshing angle, as well as a new perspective on the papers I wrote. A fortunate consequence is that this new framing itself has amounted to some interesting thoughts, which is why part I of this dissertation is not directly related to any of the papers I have written. It may be considered to be the icing on the cake.

Apart from giving a "verantwoording" for the content of this dissertation, its style also deserves to be mentioned. Although this is a philosophy dissertation, and I am now allowed to call myself a philosopher, I have been trained as a mathematician. This is reflected in my writing style, in particular when adopting the "definition-theorem-proof" format. As a consequence, on some occasions, my arguments can be concise and dense to the point of becoming opaque or more formal than strictly necessary. Over the past few years I have tried to learn to expand my arguments. But the problem is that I have a short attention span. When an argument becomes long I loose track and get confused, even when I am myself making the argument. So the mathematical style is a way to reduce the amount of nonsense. I was hoping to compensate by writing only a short (less than a hundred pages) dissertation. The idea would be that readers would then have enough time to read the compact stuff twice. I apologize for not having been able to reach this goal.

Ronnie Hermens
Groningen, August 2015

Acknowledgments

Although writing this dissertation was mostly an isolating experience I couldn't have done it without many positive influences from the outside. It is impossible to acknowledge all of them because influences are often unconscious, but also because I am a forgetful person when it comes to such things. So in order to include as many as possible, and before going into the specifics, let me start with thanking my parents, family, friends, and colleagues for whatever their contribution may have been to this dissertation.

While it is commonplace to save the last thanks for your life partner, I want to start with thanking Femke Reijnen before all readers have dozed off. I think that is only fair given the amount of support she's given me, and through how much of my frustrations she had to suffer. Luckily, the dissertation is finished now.

Now let's continue in more or less chronological order. It is only in retrospect, by meeting many people from universities over the world, that I recognize how much my own thinking is influenced by the education I received at the Radboud University Nijmegen. But in particular I want to mention Klaas Landsman and Michiel Seevinck for stimulating me to pursue a career in the philosophy of physics. I also want to thank Hans Maassen for introducing me to quantum probability theory and for helping me with some technical issues I encountered when writing this dissertation. Needless to say, all mistakes that remain are to blame on me.

A bit less directly, I'd like to thank the people at Sealed Air for providing a working environment in which I could come up with many of the early versions of ideas that went into this dissertation. Also, they helped me to develop a good working attitude that has made my life as a PhD student a lot less stressful than it could have been. And then Jan-Willem came into my life, who was willing and able to offer me a job in which I could further develop these ideas and to allow a mathematical physicist into the philosophy department of Groningen. As a supervisor his help does not point to any specific part of this dissertation. Instead his influence is woven much deeper, and has shaped my way of thinking as a philosopher.

During the four years in Groningen I have of course benefited from the expertise of the people around me. Especially the PCCP (Poor aCronym Codifying (a particular group of) Philosophers (in Groningen)) meetings have been wonderful (although I do think somebody should bring back the beer during the afternoon sessions). For connections with the philosophy of physics community though I often had to look outside the city borders. The reading group in Utrecht with F. A. Muller, Gijs Leegwater, Tim de Haan and Lara van Zuilen was a nice place of refuge.

I found even more refuge outside country borders. The summer schools in Saig provided a tremendous blurring of the divide between holiday and work. On no other occasion have I ever learned so much with so little effort, except perhaps during my stay in Oxford. I'd especially like to thank Chris Timpson and Harvey Brown for making me feel welcome there, and of course the Oxford Graduate Philosophy of Physics group.

The chronological order is now closing in on the present and my incomplete list is also coming to an end. To tie up some loose ends I should thank Ariel Fernandez for suggesting to me that my manuscript on weak intuitionistic quantum logic could be transformed into a publishable paper. And I'd like to thank Jos Uffink who has supported me both directly and indirectly since I graduated in Nijmegen in 2009. I am very grateful that he agreed to be my second promotor. This dissertation has benefited a lot from his critical thinking. Similarly, I am grateful to the members of the reading committee J. Bub (Maryland), N. P. Landsman (Nijmegen) and F. A. Muller (Rotterdam-Utrecht) for their devotion.

Finally I want to thank you, dear reader. Thank you for your sharing interest in the philosophy of quantum probability, and for ensuring that writing this dissertation has not been an exhausting waste of resources.

Contents

x

1

Introduction

Why philosophy of quantum probability? As this question is likely to be in the top ten of questions first asked when someone picks up this dissertation, it seems a good idea to break the tension and come up with an answer. Quantum probability theory, like classical[1] probability theory, is a subfield of mathematics. It is the study of quantum probability spaces conceived as abstract mathematical objects. But like classical probability theory, it falls under the header of applied mathematics. The main applications are of course quantum mechanics and derived sciences like quantum computation. But there are also attempts to apply it within other fields such as decision theory and cognitive science (Busemeyer and Bruza, 2012; Pothos and Busemeyer, 2009) and biology (Aerts et al., 2014). The topic of this dissertation, then, lies at the intersection of philosophy of probability and philosophy of quantum mechanics. Although these two field have influenced each other, the overlap itself hasn't been studied that much. This dissertation, then, is an attempt to fill up this lacunae.

A more elaborate answer to the why of philosophy of quantum quantum probability is obtained by delving into the question "what is philosophy of quantum probability?". In the remainder of this introduction I work towards an answer of this question. By doing so, I motivate both the specific topic on which I focus in this dissertation, as well as the methodology I adopt to investigate it. But in short, this dissertation is an investigation of the formalism of quantum probability in the light of its empirical role in quantum mechanics. What this program holds in more detail is spelled out in the next two sections. First, by adopting a perspective from the philosophy of probability, and then by adopting a perspective from the philosophy of quantum mechanics. Finally, in section 1.3, an outline of the remainder of this dissertation is presented.

---

[1]I use the term 'classical' to indicate non-quantum probability. In most occasions this refers to probability theory as based upon the axioms of Kolmogorov (1933). But sometimes it may also refer to something broader, such as the axioms of Kolmogorov minus the assumption of σ-additivity.

1.1   Philosophy of probability

A useful starting point is to try to think of philosophy of quantum probability as a subfield of philosophy of probability. Within this field Lyon (2010) distinguishes two central questions. The first concerns the interpretation of probability and the second the formalism of probability. The interpretation of probability has spawned a lot of philosophical debate (Hájek, 2012a), revolving around several central questions. Are there probabilities 'out there in the world' independent of any observers, that is, are there probabilities (often called chances) that are ontological? Or can one only make sense of the concept in the context of a form of ignorance about the actual state of affairs, that is, are probabilities only epistemic? Or perhaps probabilities may be something in between (Lyon, 2011; Myrvold, 2012). Furthermore, answers to these questions may differ from case to case: there need not be a unique universally applicable interpretation of probability.

As an example consider probabilities for coin flips. These have a tendency to generate stable relative frequencies for the outcomes. This stability is then often taken to suggest a connection between these relative frequencies and some objective notion of probability. A further connection between these frequencies and symmetrical properties of the coin and the stability of the physical properties of the coin over time then suggests that there may be something ontological about these probabilities as well. Although it is a non-trivial matter to make these ideas precise, they suffice to argue that probabilities for coin tosses may be of a different kind than, for example, the probabilities at a race track. Two horses participating in a race certainly show less resemblance to each other than the two sides of a coin. Furthermore, the physical conditions of a horse change significantly over time, and the relative frequency for it winning a race is not very stable. Somewhat intuitively, then, probabilities for coin tosses appear 'more ontological' than those adopted at the race track.[2] But making such intuitions precise, or showing that they are misleading or just flat out wrong, has proven to be a difficult endeavor.

The question of interpretation doesn't change much when shifting attention from classical probability to quantum probability. The answer may differ of course, but at first glance, not much is changed by adding the adjective 'quantum'. Quite the opposite is the case though when turning to the second central question posed by Lyon (2010): that of the (correct) formalism of probability. The prime example of a formalism for probability are the axioms of Kolmogorov (1933). But, as noted

---

[2]One may also have intuitions yielding precisely the opposite conclusion. There are deterministic models for coin tosses (Keller, 1986; Strzałko et al., 2008) which may even be exploited to create biases in the outcomes of tosses (Diaconis, Holmes, and Montgomery, 2007). This then can be taken to suggest that an epistemic (ignorance) interpretation for these probabilities is more apt. Horses, on the other hand, are considerably more complex systems, and possibly there is no way their behavior can be accurately described in a deterministic way, making the probabilities in some sense irreducible.

above, quantum mechanics introduces a different formalism. In fact, apart from these two examples, many other formalisms are possible and may even be preferable depending on how one conceives of probability. For example, von Mises (1931) builds his formalism entirely from a frequentist interpretation of probability relying on the notion of so-called Kollektivs. The difference between the formalisms of Kolmogorov and von Mises may be explained in the light of background intuitions about the interpretation of probability (Van Lambalgen, 1996). In the case of quantum probability, however, the formalism arises more or less by accident as part of the axiomatization of quantum mechanics.

This 'accidental' character of the formalism of quantum probability may be seen as a consequence of the fact that 'probability' occurs as a primitive concept in quantum mechanics: the axiomatization of this theory isn't guided by strong intuitions on the nature of probability. Furthermore, the use of probabilities even seems to be an indispensable and fundamental aspect of the theory. This suggests the (common) view that the theory describes a world of which probability is an intrinsic aspect. In fact, quantum mechanics has sometimes even served as a motivation for developing theories of objective chance. For example, Loewer (2004) opens his paper on objective chance with the statement that "The most important theories in fundamental physics, quantum mechanics and statistical mechanics, posit objective probabilities or chances. As important as chance is there is little agreement about what it is." Also, the phenomenon of radioactive decay has served as the paradigm example of an inherently 'chancy' event in the philosophical literature.[3] However, as intuitive as the introduction of objective chance in quantum mechanics may seem, and despite the considerable support of physicists for this idea, there is nothing in the formalism of quantum mechanics that appears to force this notion of chance to be a part of the theory. Moreover, since Loewer is correct about the little agreement concerning the meaning of objective chance, a different interpretation of quantum probability, if it fits the theory, may even be preferable.

Now while a theory of objective chance may not be necessary to make sense of probability in quantum mechanics, such a theory may well be useful. However, this requires that at some point this theory has to be connected to the formalism of quantum mechanics. Unfortunately, this isn't done very often. As a telling historical example one can consider the work of Popper on probability in quantum mechanics. In his 1957 paper "The propensity interpretation of the calculus of probability, and the quantum theory" he proposes an interpretation that "takes the mystery out of quantum theory, while leaving probability and indeterminism in it" (p. 68). However, Popper nowhere recognizes that the classical calculus of probability, which he is interpreting, is manifestly distinct from the way probability

---

[3] See for example (Skyrms, 1977; Lewis, 1981; Pollock, 1984; Hall, 1994; Noordhof, 1999; Kment, 2012).

occurs in quantum mechanics. Without bridging this gap it is inevitably unclear how an interpretation of the former can play any role in resolving conceptual issues in the latter. Presumably, Popper's interpretation was supposed to fit one of his own axiomatic frameworks (Popper, 1938; Popper, 1959, Appendix *iv). But this is even less helpful as they seem far removed from the quantum formalism.[4]

Now at this point one may have several worries. First, one may try to argue for the requirement of objective chance in quantum mechanics not on the basis of the formalism of quantum mechanics alone, but also on the basis of no-go theorems for hidden variables like the Kochen-Specker theorem or Bell's theorem. Connections between such theorems and the role of probability are not uncommon in the literature, especially in popular science. For example, Wolchover (2014) writes: "An experimentally tested theorem by the Northern Irish physicist John Bell says there is no "true" state of the particle; the probabilities are the only reality that can be ascribed to it." Second, one may wonder how important the emphasis I put on formalism really is: can't a well-developed interpretation of probability just work for any decent formalism of probability? This last point can be dealt with directly. If an interpretation of probability is as flexible as suggested, then the connection with the formalism must be very weak. However, without a solid connection, it is not clear if the given interpretation is actually an interpretation of probability as it is used, or of something else entirely.

The first point relies on a misrepresentation of what no-go theorems actually prove. The setup of such theorems is that, first, certain plausible assumptions for physical theories are formulated. Then, it is shown that no theory that satisfies these assumptions can reproduce the predictions of quantum mechanics. Consequently, if quantum mechanics is taken to provide the correct predictions, one of the assumptions has to be rejected. However, these assumptions do not incorporate restrictions on the interpretation of probability. Consequently, simply positing the existence of objective chance cannot provide an explanation of these theorems. On the other hand, the quantum mechanical predictions that play a role in no-go theorems do rely on the use of the formalism of quantum probability. This suggests that an explanation of the formalism can help in explaining these theorems. It is for this reason that in this dissertation the focus is on the formalism of quantum probability, rather than its interpretation. The adopted method for this investigation is best explained from the perspective of the philosophy of quantum mechanics, to which I turn now.

---

[4]Bub (1975) provides a more extensive critique of Popper's alleged resolution of the foundational problems in quantum mechanics. Although I agree with his critique, I also think it should be recognized that when Popper introduced his ideas, formal results against their success such as Gleason's theorem (1957) and the Kochen-Specker theorem (1967) were just fresh on the market, or yet to be discovered.

1.2 Philosophy of quantum mechanics

In comparison to preceding theories in physics, quantum mechanics is odd, in the following way. Newtonian mechanics is a theory about forces acting between rigid bodies, and their influence on how these bodies move. Maxwell's theory of electromagnetism is about electromagnetic fields and their influence on the movement of bodies with electrical charge. Thermodynamics is about thermal properties of all physical processes, as well as the behavior of particular substances in processes. What these theories have in common, is that they are about *something*. Of course, there is the question of the nature of the relationship between the abstract things in the theory, and things that actually exist in the world. But they sketch a metaphysical picture of the world that is at least clear enough for the student of physics to be able to understand the theory, and to apply it.

In contrast to the theories just mentioned, quantum mechanics does not pose a clear candidate for a metaphysical picture of the world. Or, as argued by Maudlin (2014), the theory cannot be taken literally. In its minimalist form, quantum mechanics is about measurements on physical systems and their possible outcomes. But it is obnoxiously silent about what may be going on when no measurement is performed, or what precisely constitutes a physical system. To try to fill in these gaps, or even contemplating whether they can be filled at all, is to delve into the philosophy of quantum mechanics. To be sure, there is a natural candidate in the theory for spelling out what may be going on: the quantum state a.k.a. the wave function. However, this state appears to behave differently depending on whether or not a measurement is performed. This leads one to question how or why measurements are physical processes to be distinguished from other physical processes. But as measurements themselves dictate in part how the state evolves, there is little hope of explaining what kind of physical processes measurements are in terms of evolutions of the quantum state. In short, for a theory that is almost solely about measurements, it gives very little indication of what measurements are. As Bell (1981, p. 117) noted: "The concept of 'measurement' becomes so fuzzy on reflection that it is quite surprising to have it appearing in physical theory *at the most fundamental level.*"

What I am stumbling upon is, of course, the notorious measurement problem. This problem is so deeply woven into the theory, that there isn't even agreement about what the problem is.[5] But one useful way to think about it is in terms of a question. Although quantum mechanics is about measurements and their outcomes, in general, given a specified measurement, the theory does not prescribe which outcome will occur. On the other hand, actual measurements are usually seen to have a single relatively well-defined outcome. How can one account for the

---

[5]For example, Maudlin (1995) distinguishes three measurement problems, while both Pitowsky (2006) and Griffiths (2013) distinguish two. All of these problems differ from each other.

occurrence of definite measurement outcomes when the theory only gives sets of possible outcomes? Where do definite outcomes come from? The measurement problem can then be thought of as the culmination of all possible problems one encounters when trying to answer these questions.

But although the measurement problem may be the biggest problem in the philosophy of quantum mechanics, it certainly isn't the only one. Here I want to consider a problem that occurs right at the same level when looking at the theory. Apart from providing the possible outcomes for measurements, quantum mechanics also attaches probabilities to these possible outcomes. Like the notion of 'measurement', the notion of 'probability' can also be used in practical situations. Scientists usually have a good idea of what kind of processes to count as measurements, and similarly, they are good in adopting a working definition of probability as reflecting the relative frequencies of outcomes of a long series of similar experiments. But as is well known, frequentist interpretations of probability are fraught with problems (Hájek, 1996 ; 2009). More generally, as seen above, the interpretation of probability is a difficult topic. The following paraphrasing of Bell then seems apt: "The concept of 'probability' becomes so fuzzy on reflection that it is quite surprising to have it appearing in physical theory *at the most fundamental level.*"

The philosophy of probability in quantum mechanics (and also in other fields of physics) has regained interest over the past twenty years or so. Thus in the relatively recent book edited by Beisbart and Hartmann (2011) one reads in the introduction that it "is the first to provide a detailed philosophical appraisal of the status of probabilities in all of physics." What is typical of these investigations of the status of quantum probability, is that they are often performed with respect to a certain interpretation of quantum mechanics. This is obviously useful research. Indeed, adopting an interpretation of quantum mechanics is an intuitive method to explain the formalism of quantum probability. Once one goes beyond the minimalist version of the theory and starts sketching what the theory is actually about, one may also find ways to get a grip on quantum probability.

Historically, finding a solution to the measurement problem has been a priority in the development of an interpretation of quantum mechanics. As a consequence, one now finds a plethora of interpretations. Apart from the Copenhagen-like interpretations as developed mainly by Bohr and Heisenberg (Faye, 2014), one also has (among others) Bohmian mechanics (Bohm, 1952; Dürr, Goldstein, and Zanghì, 2013), spontaneous collapse theories (Ghirardi, Rimini, and Weber, 1985 ; 1986; Pearle, 1989), many-worlds interpretations (Everett III, 1957; Wallace, 2012), modal interpretations (Van Fraassen, 1972; Dieks and Vermaas, 1998) and consistent histories (Omnès, 1999; Griffiths, 2003). When the minimal statistical interpretation of Born is abandoned and probabilities are assumed to play a less primitive role, the question arises how these probabilities precisely emerge from the postulated ontology, and why they numerically coincide with the values given by the Born postulate.

Thus one finds results, for example, on how Born probabilities arise in Bohmian mechanics (Callender, 2007), in the many worlds interpretation (Wallace, 2012, Ch. 4-6), or in quantum Bayesianism (QBism) (Fuchs and Schack, 2013). A particularly friendly paper on probability in several realist interpretations of quantum mechanics is (Timpson, 2011).

Evaluating quantum probability against the background of the adopted interpretation may be considered to be a bottom-up approach. First one establishes the ontology, and from there one reconsiders the specific aspects of quantum mechanics. To complement these kinds of investigations, I aim to provide a more top-down approach: what constraints does quantum probability impose on a possible ontology? For such an approach a more abstract investigation of quantum probability is required, in which the measurement problem is (at least temporarily) set aside. The hope, then, is that by using a top-down approach the results obtained in this dissertation are in a sense independent of the interpretation of quantum mechanics. This can only be achieved by adopting a sort of ignorance concerning the ontology of quantum mechanics, and instead focusing on the empirical content of the theory. Roughly, it means that I will act as an operationalist concerning the concept of measurement, but not concerning the concept of probability.

A sharp distinction between empirical and ontological content of a theory is of course not to be expected. Instead, I will do my best to highlight whenever a particular assumption is being made and whether it relies on an ontological or empirical intuition. This is by no means trivial. As an example of such assumptions, consider what Bub and Pitowsky (2010) call the "two dogmas about quantum mechanics". The first of these is the idea that measurement is not allowed to play a fundamental role in the theory. That is, the theory should be able to provide (in principle) a dynamical account of measurement processes as regular physical processes. The second is the view that the quantum state provides a (possibly incomplete) representation of some aspect of physical reality. The information-theoretical approach to quantum mechanics of Bub and Pitowsky[6] relies on the rejection of these two dogmas. The methodology I adopt, on the other hand, requires the recognition of these dogmas and adopting a strict agnosticism about their truth. The truth or falsity of these assumptions is, after all, an aspect of an interpretation of quantum mechanics and is not determined directly by the empirical content of the theory.

Possibly, the sketched mathematical empiricist methodology still sounds sketchy. However, it lies in the nature of methodologies that they are best understood by seeing them being applied. Particular illustrative applications found in this dissertation are the discussion of the postulates of quantum mechanics, the Kochen-Specker theorem and its finite precision loophole, and the development of an empiricist quantum logic. As a final note, I wish to emphasize that my adoption of an empiricist

---

[6]See also (Pitowsky, 2006; Bub, 2011b).

methodology is not intended as an endorsement of an empiricist interpretation of quantum mechanics. On the other hand, the explicit agnosticism adopted does often reflect my own cluelessness concerning what quantum mechanics is about.

1.3   Outline

This outline is mainly intended as a guide for the reader who only wishes to read at most a few parts of this dissertation. It gives a quick sketch of what can be found where, in a slightly more comprehensive way than in the table of contents. Most chapters are not self-contained. On the other hand, the reader wishing to skip to a specific chapter most likely already knows a lot of the relevant background that may have been covered in preceding chapters. If, however, one does find a concept that sounds unfamiliar, try the index, which may guide directly to the relevant page. Finally, most of the mathematical concepts used are concisely introduced in the appendix.

1.3.1   Part I

The first part of this dissertation is mainly concerned with preliminary considerations. To quote Peres (2003, p. 459): "As it often happens in science, these things were well known to those who know things well." If you are one of those readers well familiar with both the foundations of quantum mechanics and the mathematics of quantum probability, you will not find a lot of surprises here. But I suspect many readers will not fall in this category. Furthermore, while the topics discussed in part I are well-known to some people, I do not know of any texts that neatly link them all together. The value of part I, then, is to make a short and smooth connection from the foundations of quantum mechanics to the mathematics of quantum probability, highlighting the philosophical aspects of the steps made along the way.

In chapter 2, I give a short historical introduction of how the notion of probability in quantum mechanics was received in the early days of quantum mechanics (1925–1935) and continued to shape the ideas of people like Born, Einstein and Bohr well into the period 1935–1950. Opposing views on the role of probability in quantum mechanics that emerge from this discussion are then linked to the programs in parts II and III. The postulates of quantum mechanics in the style of von Neumann are presented in chapter 3. These are analyzed for their empirical relevance, and this analysis at the same time will be useful for distilling the Hilbert space formalism of quantum probability from the axioms of quantum mechanics, which is done in chapter 4. In the same chapter, the more general algebraic formulation of quantum probability is also introduced. A comparison between the two formalisms, as well as with classical probability is made. The upshot is that classical probability

can be conceived as a special case of algebraic quantum probability, restricted to commutative algebras.

1.3.2 Part II

The fact that classical probability can be represented as a special case of quantum probability does not in itself imply that quantum probability is more general than classical probability. Generality can be a matter of perspective. Possibly, quantum probability spaces can also be represented as classical probability spaces. If so, the philosophy of the formalism of quantum probability may be a red herring, since there may be nothing new beyond classical probability. This possibility is studied in part II of this dissertation.

The possibility of classical representations of quantum probability spaces is technically the same question as the possibility of hidden variable interpretations of quantum mechanics. Conceptually, though, there is a difference. For example, hidden variables may be required to be *physically* explanatory, or to be constrained by further metaphysical assumptions. A classical representation of a quantum probability space, in turn, may be required to be *mathematically* explanatory. For example, such a representation may be required to capture all relevant aspects of the quantum formulation, i.e., a classical representation may be required to be *faithful*. In chapter 5 these considerations are spelled out in more detail.

The remainder of part II is devoted to evaluating the theorem of Kochen and Specker and its implications for both the possibility of hidden variables and classical representations. In chapter 6 the Kochen-Specker theorem is discussed, together with the well-known possibility of dropping the assumption of non-contextuality. A less well-known possibility of circumventing the Kochen-Specker theorem, which goes by the name of the finite precision loophole, is investigated in chapter 7. Formally, both these loopholes may be exploited to construct classical representations of quantum probability. However, none of the classical representations are found to be entirely faithful in the sense of providing a clear uncontroversial reformulation of quantum probability. This motivates the more direct investigation of the formalism of quantum probability in part III.

1.3.3 Part III

In chapter 8, I explain how the formalism of quantum probability is tightly connected to that of orthodox quantum logic.[7] An investigation of the latter is therefore useful for understanding the former. However, I reach the conclusion that a direct interpretation of orthodox quantum logic (specifically in terms of natural language)

---

[7]I use the adjective "orthodox" to refer to the quantum logic of Birkhoff and von Neumann (1936) specifically, thereby allowing a broader reading of the term "quantum logic".

is next to impossible. This then is taken as a motivation to try to reformulate quantum logic in a way such that it helps explaining orthodox quantum logic and quantum probability.

A first attempt is made in chapter 9. The method applied here relies on the work of Coecke (2002), which suggests a direct expansion of the orthodox quantum logic. However, as the method used relies on an abstract construction, the interpretation of the obtained structure is not immediately clear. Some interesting properties are discussed, including the idea that as a logic of propositions about quantum states it is more appropriate than orthodox quantum logic. In the end, though, the conclusion is that for the interpretation of the formalism of quantum probability a different route may be more fruitful.

In chapter 10, a new quantum logic is constructed from scratch. The starting point is the introduction of a set of elementary experimental propositions that are considered to be indispensable for the formulation of quantum probability in terms of natural language. The emphasis on natural language is adopted to prevent the difficulties encountered with the interpretation of the quantum logics discussed earlier. Logical connections between these propositions are then made based upon the empirical non-probabilistic content of quantum mechanics. These constructions lead first to the introduction of an intuitionistic quantum logic, i.e., a Heyting algebra of experimental propositions is constructed. This logic is then further expanded to obtain a Boolean algebra. In short, on the basis of empirical considerations concerning quantum measurements, a classical quantum logic is constructed.

In chapter 11, the possibility of introducing probability functions on the Boolean algebra of chapter 10 is investigated. It is shown that every quantum probability function can indeed be represented as a probability function on this new lattice. Thus quantum probability functions can then be understood as functions assigning probabilities to experimental propositions in a classical logic. It is shown that the reformulation of quantum probability also allows probability functions that do not obey the trace-rule that is typical of quantum probability. Sections 11.3 and 11.4 are devoted to investigating the possibility of characterizing the probability functions that do satisfy the trace-rule. In section 11.3, it is shown that these probability functions can be identified with those that may be called non-contextual. A short investigation is made of how non-contextuality can be understood as a natural criterion for probability functions by looking at three interpretations in which contextual probability functions are (in principle) also a possibility, to wit: Everettian quantum mechanics, Bub and Pitowsky's information-theoretical interpretation, and QBism. In section 11.4 I investigate two other possibilities to characterize the non-contextual probability functions. The first focuses on the possible role for a continuity assumption, and the second focuses on a link with the quantum logic that was obtained in chapter 9. The chapter ends with a short evaluation of the results of part III.

Part I

Introductory considerations

<div style="text-align: right; font-size: 3em;">*2*</div>

Probability in the early quantum theory

In this chapter I sketch a short history of how probability was introduced into quantum mechanics, and how this introduction was received at the time. It is not my aim to trace a full historical development. For extensive works on the history of quantum mechanics I refer the reader to the books of Jammer (1966) and Mehra and Rechenberg (1982). Instead, the aim is to show that, using a historical context as a method, quantum mechanics does not dictate how one should think of the probabilities in the theory. Although it is still commonly heard that quantum mechanics dictates the existence of objective chances or indeterminism, this is quite far from the truth. There is a lot of room for differing views on probability in quantum mechanics, and this was already apparent in the days when probabilities got first introduced into the theory. Different views on probability in quantum mechanics may pull into opposite directions for how one wishes to (re)formulate quantum probability. The chapter ends with a section that links the possible views on quantum probability that emerged in this chapter to two separate programs for reformulating quantum probability. These programs in turn are studied in parts II and III of this dissertation.

2.1 Probability in early quantum mechanics

The pioneering period of quantum mechanics (circa 1900-1925) is often referred to as "the old quantum theory" (Ter Haar, 1967). It encompasses seminal results such as Planck's law for black body radiation (1900), Einstein's description of the photoelectric effect (1905), and Bohr's model of the atom (1913). The following decade saw the culmination of these and other results into a single theory, starting with the matrix mechanics of Heisenberg (1925) and the wave mechanics of Schrödinger (1926), and leading to the formalization of quantum mechanics by Dirac (1930) and von Neumann (1932). It is within this period that Born introduced his probability interpretation, and the early philosophy of quantum mechanics developed.

Heisenberg's matrix mechanics was not heartily welcomed by everyone. As one of the first theories of atomic particles, it was quite radical in eliminating the classical concept of particle trajectories from the physical picture of the world. This was accompanied by a philosophy (often called the Copenhagen interpretation,[1] mainly developed by Bohr and Heisenberg) that supported the idea that this elimination was somehow necessary. What remained notoriously vague, however, was what physical picture was to replace the old one, if any. A direct reading of the statement, ascribed to Bohr,[2] that "there is no quantum world", would suggest that a purely instrumentalist point of view underlies the Copenhagen interpretation. But things aren't that simple, and Folse (1985) for example, attributes a more realist interpretation to Bohr. A balanced discussion is given by Landsman (2006 ; 2007).

Half a year later, Schrödinger (1926) presented his wave mechanics as a rival theory. In contrast to matrix mechanics, this theory allegedly *did* provide an ontology which would ground the empirical aspects. Not long before, de Broglie (1924) had introduced the wave aspect of particles: trajectories of particles were supposed to be constrained by waves. Schrödinger took the wave picture a step further by letting the waves represent the distribution of matter in space itself. He further demonstrated that his theory was empirically equivalent with Heisenberg's, thus saving all the merits of that theory while getting rid of the accompanying philosophy he disliked.

However, wave mechanics and its intended interpretation were not without problems either. One problem is that dynamically the waves do not typically evolve into peaked Gaussian-like functions that would represent the localization of a particle. The way matter is described by the theory then is rather different from how it is usually observed in experimental settings.[3] Another problem is that, for n-particle systems, these waves live in $3n$-dimensional space and thus do not obviously describe the distribution of anything in the 3-dimensional space we experience around us. So even if the waves in wave mechanics could provide an ontology, their relation to observed phenomena is not clear.

It deserves to be noted that the earlier mentioned empirical equivalence of the two theories also isn't entirely without problems. Both theories require reformulations and modifications to be put on a level in which they can be compared formally. For a discussion of the details see (Muller, 1997a,b). However, it appeared to be commonly accepted at the time that both theories were aiming towards the same

---

[1]There are many, not necessarily compatible, philosophical ideas that go under the header of 'Copenhagen interpretation'. This makes it, generally, a dangerous term to use without further elaboration of what is meant (Howard, 2004). Here, I use it as a place holder for the common ground between the views of Heisenberg and Bohr.

[2]See for example (Petersen, 1963, p. 12).

[3]This may be considered a preliminary version of the measurement problem. And a recent attempt at solving the measurement problem by Landsman and Reuvers (2013) starts precisely with this older issue.

new theory that was to be developed. Thus the physical pictures illustrated by Heisenberg and Schrödinger shifted from being independent theories, to being possible interpretations of the same theory.

Both the interpretation of Heisenberg and that of Schrödinger faced difficulties. Born noted these difficulties and proposed a solution: "Neither of these two views seem satisfactory to me. I would like to try to give a third interpretation here, and to test its applicability to collisions."[4] (1926b, p. 803). The solution was to keep the wave functions of Schrödinger, but to replace their interpretation as matter densities with probability densities.[5] Thus the probabilistic, or statistical, interpretation of quantum mechanics was born. In its minimalist form, it coincides with a common popular science account of quantum mechanics in which the probabilities themselves are the only elements of reality.

Not everybody was persuaded by this view. And even though Born's rule for calculating probabilities became part of the axioms of quantum mechanics, there was (and still is) neither consensus on the correct interpretation of these probabilities, nor on the question whether these probabilities reflect all that the theory has to say about nature. The probabilistic interpretation was most famously opposed by Einstein. However, the famous relevant quote has too often been misrepresented, and in an effort to do my part on countering this misrepresentation I give here the relevant passage from Einstein's letter to Born in full.

> Quantum mechanics is certainly imposing. But an inner voice tells me that it is not yet the real thing. The theory says a lot, but does not bring us any closer to the secret of the 'old one'. I, at any rate, am convinced *He* is not playing at dice. Waves in 3-dimensional space, whose velocity is regulated by potential energy (for example, rubber bands)...I am working very hard at deducing the equations of motion of material points regarded as singularities, given the differential equation of general relativity. (Einstein 1926, published in (Einstein, Born, and Born, 1971, p. 91))

Now surely Einstein rejects Born's interpretation here. The reason, however, doesn't seem to be that this interpretation implies indeterminism. Rather, it seems that Einstein is expressing doubt about whether the shift to a probabilistic reading of quantum waves takes away the main interpretational problems of the theory.

---

[4]My translation; original: "Keine dieser beiden Auffassungen scheint mir befriedigend. Ich möchte versuchen, hier eine dritte Interpretation zu geben und ihre Brauchbarkeit an den Stoßvorgängen zu erproben."

[5]Actually, instead of directly taking the wave function to represent a probability density, one has to take the square of its absolute value. This was something Born had to correct twice. First as a footnote in (Born, 1926a) in which the function was replaced by its square, and then again (recognizing that wave functions are complex valued) by the square of its absolute value in (Born, 1926b).

Simultaneously, he appears to be expressing doubt about whether Schrödinger's approach makes sense. This last point is clearer when noting that the phrase '3-dimensional space' has been wrongly translated from the original '$3n$-dimensionalen Raum' (Einstein, Born, and Born, 1969, p. 98). Instead, Einstein seems to hold the view that a new theory has to replace quantum mechanics before one can understand what the objects in the quantum formalism represent. His own theory of general relativity seemed a good starting point to him.

It has been stressed many times in the literature that Einstein's main objection to quantum mechanics was not that it was probabilistic or indeterministic. The most famous defense of this reading of Einstein is perhaps found in the letter of Pauli to Born.

> Einstein gave me your manuscript to read; he was not at all annoyed with you, but only said you were a person who will not listen. This agrees with the impression I have formed myself insofar as I was unable to recognise Einstein whenever you talked about him in either your letter or your manuscript. It seemed to me as if you had erected some dummy Einstein for yourself, which you then knocked down with great pomp. In particular, Einstein does not consider the concept of 'determinism' to be as fundamental as it is frequently held to be (as he told me emphatically many times) [. . .] Einstein's point of departure is 'realistic' rather than 'deterministic', which means that his philosophical prejudice is a different one. (Pauli 1954, published in (Einstein, Born, and Born, 1971, p. 221))

The misinterpretation of Einstein's criticisms is not surprising when noting (using Pauli's terminology) how intertwined the two departures 'realistic' and 'deterministic' become within the context of quantum mechanics. This twining has to be taken into careful consideration when doing philosophy of probability. A first aid for the required unraveling comes from a look at the status of Heisenberg's uncertainty principle.

2.2   Probabilities in the uncertainty principle

One of the first criticisms against matrix mechanics (which still applies to quantum mechanics), is that there are no states in the theory that simultaneously assign a definite value to both the position and momentum of a particle. In fact, most states assign a definite value to neither. If one supposes that these values in fact do exist, one can only conclude that the formalism of Heisenberg is incomplete. However, for Heisenberg, the view was that particles only have a definite position or momentum if these properties are actually observed/measured. Since in these cases the formalism can give these definite values, it seems perfectly complete.

Nevertheless, matrix mechanics was found to be difficult to interpret, and its abstractness seemed ill-motivated. Moreover, it was unclear why certain properties (like position and momentum) were never ascribed a definite value simultaneously. The uncertainty principle was meant to shed light on these matters. That is, the principle is used to defend a certain interpretation of matrix mechanics.[6]

The original uncertainty principle (Heisenberg, 1927) demonstrates that, for particular experimental setups, there is a bound on the accuracy with which pairs of conjugate variables (such as position and momentum) can be measured. The most famous example concerns the determination of the position of an electron using a microscope. This determination requires that the electron at least interacts with a photon. The accuracy of the measurement is then limited by the wave length of the light. However, the shorter the wave length, the more energetic the photon is. Consequently, the interaction will establish a disturbance in the momentum of the electron proportional to the inverse of the wave length.

These setups are then considered to be generic, thereby suggesting that the obtained bounds are empirically universally valid. A special trait of these considerations is that they are independent of the formalism of matrix mechanics. Thus the uncertainty principle in this approach has a flavor similar to that of empirical principles such as they appear in thermodynamics (the impossibility of certain kinds of perpetual motion). However, the step from the inaccessibility of values of certain quantities to the claim that therefore these definite values should play no role in a theory is not straightforward, and requires an additional philosophical principle. Born, when reflecting in his Nobel lecture on the development of quantum mechanics, phrased it as follows:

> The principle asserts that concepts and pictures that do not correspond to physically observable facts should not be used in theoretical description. When Einstein, in setting up his theory of relativity, eliminated the concepts of the absolute velocity of a body and of the absolute simultaneity of two events at different places, he was making use of the same principle. Heisenberg banished the picture of electron orbits with definite radii and periods of rotation, because these quantities are not observable; he demanded that the theory should be built up by means of quadratic arrays of the kind suggested in a preceding paragraph. Instead of describing the motion by giving a coordinate as a function of time $x(t)$, one ought to determine an array of transition probabilities x_{mn}. (Born, 1955, p. 676)

A similar idea can be found in the following quote of Heisenberg in which he recollects a conversation with Einstein:

---

[6]Or, more closer to the original debate at the time, to show that the theory is in a sense 'anschaulich'. For more details see (Hilgevoord and Uffink, 2014).

> "[. . .] since a good theory must be based on directly observable magnitudes, I thought it more fitting to restrict myself to these [. . .]" "But you don't seriously believe," Einstein protested, "that none but observable magnitudes must go into a physical theory?" "Isn't that precisely what you have done with relativity?" I asked in some surprise. [. . .] "Possibly I did use this kind of reasoning," Einstein admitted, "but it is nonsense all the same. [. . .] on principle, it is quite wrong to try founding a theory on observable magnitudes alone. In reality the very opposite happens. It is the theory which decides what we can observe. [. . .]" (Heisenberg, 1971, p. 63)

The analogy between Einstein's disposal of absolute simultaneity and Heisenberg's disposal of definite positions and momenta is certainly interesting. If the analogy is indeed apt, it is understandable why Born thought Einstein's objections had to be directed against indeterminism rather than against the lack of realism; from the realistic point of view quantum mechanics is not more eliminative than special relativity. In both cases, concepts were disposed that lacked direct empirical access, which need not harm a realist point of view. There are however important distinctions. The notion of simultaneity plays no role in the application of special relativity to describing the relevant phenomena. The notions of position and momentum, on the other hand, play an indispensable role in the application of quantum mechanics. (In fact, this is even part of Bohr's correspondence principle.)

The dis-analogy with the use of Born's principle in special relativity reveals that there appears to be some double standard at play when applying the principle to quantum mechanics. In deriving the uncertainty principle, Heisenberg (1927) explicitly refers to both the value of the position and the momentum of the electron, as if these concepts are simultaneously applicable: "At the instant when the position is determined [. . .] the electron undergoes a discontinuous change in momentum." The uncertainty principle appears to hold because the definite values for position and momentum are disturbed by the act of measurement in an uncontrollable way. Then, the motivation to remove simultaneous definite values for position and momentum from the theory, seems to presuppose the existence of these simultaneous definite values. This is a peculiarity that requires an explanation. This explanation may be sought in the philosophy of Bohr, to which I will turn now.

2.3 Uncertainty and complementarity

Bohr explicitly rejected the idea that measurements could be seen as disturbing certain definite properties of the system. Instead of viewing the uncertainty principle as reflecting an uncontrollable aspect of the measurement process, he saw it as expressing a limit to the *definability* of concepts such as position and momentum:

the uncertainty in question is not simply a consequence of a discontinuous change of energy and momentum say during an interaction between radiation and material particles employed in measuring the space-time coordinates of the individuals [. . .] the question is rather that of the impossibility of defining rigorously such a change when the space-time coordination of the individuals is also considered. (Bohr, 1985, p. 93)

The background of this view rests on yet another principle, namely, that of complementarity. According to this principle, 'position of a particle' and 'momentum of a particle' refer to mutually incompatible modes of description of a system. Furthermore, neither mode of description provides an exhaustive description: both are needed to obtain a full characterization of the system.

The fact that Bohr was not entirely satisfied with Heisenberg's exposition of the uncertainty principle was brought to Heisenberg's attention just before his paper was published. The clash resulted in a small "addition in proof" in the paper. In this addition it is not entirely clear what Heisenberg's position on Bohr's complaint was, but in a later work he reflects on the period between finishing his paper and publishing it as one in which they reached agreement:

After several weeks of discussion, which were not devoid of stress, we soon concluded, not least thanks to Oskar Klein's participation, that we really meant the same, and that the uncertainty relations were just a special case of the more general complementarity principle. Thus, I sent my improved paper to the printer and Bohr prepared a detailed publication on complementarity. (Heisenberg, 1967, p. 105)

It is, however, not entirely clear how the uncertainty principle can be viewed as a special case of the complementarity principle. After all, the uncertainty principle links two mutually exclusive modes of description into a single relation. The complementarity principle, if taken seriously, should preclude a meaningful reading of this relation. Such a view can indeed be found in the work of Bohr (1949, p. 211): "It must here be remembered that even in the indeterminacy relation we are dealing with an implication of the formalism which defies unambiguous expression in words suited to describe classical pictures." On this reading, the uncertainty principle only makes sense in the limit cases where either position ($\Delta x = 0$) or momentum ($\Delta p = 0$) is well-defined.

This view is often rejected though, and the uncertainty relation is asserted to be a meaningful mathematical expression. The reason for this may possibly be traced back to the idea, upheld by many physicists, that the wave function or quantum state expresses something real. For example, Born writes

We regard waves on a lake as real, though they are nothing material but only a certain shape of the surface of the water. The justification is that

> they can be characterized by certain invariant quantities, like frequency and wavelength, or a spectrum of these. Now the same holds for light waves; why then should we withhold the epithet 'real', even if the waves represent in quantum theory only a distribution of probability? (Born, 1953, p. 149)

To see why this view on quantum waves is relevant, it needs to be noted that the uncertainty principle not only occurs as an empirical principle, but also comes in a version that is a theorem of quantum mechanics itself. In fact, this version of the principle (due to Kennard (1927)) is more widely known, and discussed in any textbook. Instead of considering possible experimental setups of position and momentum measurements, this principle is directly derived from the theory. To be precise, it states that for every quantum state ψ the variances in position X and momentum P satisfy the inequality

$$\mathbb{Var}_\psi(X)\,\mathbb{Var}_\psi(P) \geq \hbar/2. \tag{2.1}$$

Here X and P are now operators on a Hilbert space, $\hbar$ denotes Planck's constant, and the variance is given by the Born rule.

Besides the shift from a pretheoretical principle to a theorem of a theory, there is another important distinction between this relation and the one obtained by Heisenberg. In Heisenberg's relation the uncertainties refer to the precision of actual measurements. The inequality (2.1), on the other hand, uses Born's interpretation of the wave function and identifies the uncertainties with spreads in the probability distributions associated with ψ. If this wave function is real in some sense, then it seems at least plausible that the associated probability distributions are also real. The uncertainty principle then is not only meaningful in the limit cases, but also expresses a definite feature of the world in any other case. This is, of course, in strong contrast with the view on the quantum formalism held by Bohr:[7]

> The entire formalism is to be considered as a tool for deriving predictions, of definite or statistical character, as regards information obtainable under experimental conditions described in classical terms [...] These symbols themselves [...] are not susceptible to pictorial interpretation; and even derived real functions like densities and currents are only to be regarded as expressing the probabilities for the occurrence of individual events observable under well-defined experimental conditions. (Bohr, 1948, p. 314)

---

[7]It deserves to be noted though that Born himself at many points also sided with Bohr's philosophy. Even though for him the quantum state was perfectly real, the associated probability distributions only have meaning with respect to an experimental setup: "The prediction made by the theory [...] has a meaning only in relation to the whole experimental arrangement" (Born, 1953, p. 146).

There are several things that may be taken away from these considerations. First, Born's rule for calculating probabilities from the quantum state was widely accepted. But second, there wasn't much consensus on the ontological status of these states or their associated probabilities. With Born one sees the strongest tendency to view quantum probabilities as inherent aspects of reality. In the writing of Bohr one finds a more modest interpretation, associating quantum probabilities with descriptions of experiments rather than that they define the mechanisms behind these experiments. And finally, in the writing of Einstein, one finds the view that quantum mechanics is at best a preliminary theory. The ontological status of quantum probabilities can then only be understood after a more complete theory has been devised.

2.4 Reflection

The early history of probability in quantum mechanics given in this chapter illustrates that, although the introduction of probabilities by Born was welcomed as part of the theory, there was no consensus on the foundational role of these probabilities. This situation still holds today, and this may be seen to be reflected in the way in which probability appears in the theory (see also chapter 3). Born's rule for calculating probabilities has an operationalist flavor. This may be alleviated by giving a reformulation of quantum probability so as to give a clearer view on its physical role, which is the topic of this dissertation. This may seem a bit odd at first since there is a well established mathematical formalism of quantum probability. This formalism is, of course, adequate for the application of quantum mechanics. But it may not be the most adequate formalism for illustrating the role of probability in quantum mechanics.

The investigations in this dissertation are divided into two parts (parts II and III). In the light of the discussion in this chapter, part II may be seen to latch on to Einstein's view, while part III is closer connected to Bohr's views. This claim deserves some elaboration. The mainstream view on quantum probability is that, although the theory provides probabilities for particular measurement outcomes, these outcomes themselves do not relate to any meaningful properties of the system outside the context of the measurement. This idea is reflected in the formalism of quantum probability by the fact that there is no set of which the elements determine outcomes for all possible measurements. This should be contrasted to classical probability in which the probability space can be understood as representing all possibilities.

Einstein's plea for a completion of quantum mechanics can now loosely be thought of as a plea for returning to classical probability.[8] After all, if one can recast

---

[8]This is a somewhat a-historic way of thinking about it. The axiomatization of classical prob-

quantum probability into the language of classical probability, then the points in the classical probability space associated with a quantum system are natural candidates for more complete descriptions of the system than the quantum states. The question of whether quantum probability can be recast into the language of classical probability thus parallels the question of the possibility of a hidden variable theory. It is for this reason that part II is to a large extent devoted to the discussion of a particular no-go theorem for hidden variables, namely, the Kochen-Specker theorem. This theorem is re-interpreted as one posing constraints on possible representations of quantum probability within classical probability.

It is not too surprising that the shift to a theory that completes quantum mechanics would demand a reformulation of quantum probability. But also if one remains agnostic concerning the completeness of quantum mechanics, or even accepts its completeness, there is good reason to ask for a reformulation of quantum probability. I will argue for this in more detail in part III, but a short motivation can already be given at this stage. The quantum state (as a possibly complete description) gives rise to a probability function. But being a function implies that there are two additional ingredients: the range and the domain. In quantum probability the range is the familiar interval $[0, 1]$ just like in classical probability. The domain is a bit more complicated. Mathematically, it is well-defined as the set of projection operators on a Hilbert space. But now consider, for example, the perspective of Bohr. On his view elements in the domain of the probability function should be "individual events observable under well-defined experimental conditions." The link between individual events and projection operators is not trivial (see chapter 8). A reformulation of quantum probability then can help to clarify what this link is.

It deserves to be noted that I do not believe that a philosophical investigation of the formalism of quantum probability can magically lead to a solution of the foundational problems in quantum mechanics. The interest in doing philosophy of quantum probability is more basic. It finds its roots in the recognition that the framework of probability that naturally arises in quantum mechanics, is manifestly distinct from the classical formalism of probability. To evaluate the significance of this change in formalism both for the foundations of quantum mechanics and for the foundations of probability theory, a philosophical investigation of quantum probability itself is helpful and perhaps even necessary. This of course requires first to have a firm grip on this formalism and its relation to quantum mechanics. The next two chapters are devoted to obtaining this grip.

---

ability was only given by Kolmogorov in 1933.

3

Postulates of quantum mechanics

The purpose of this chapter is to introduce the formalism of quantum mechanics. For this introduction I make use of a set of postulates that go back to the work of von Neumann (1932). Although they don't appear literally in this form in his book, it is generally accepted that these postulates form the backbone of his formalization of quantum mechanics. The particular approach of axiomatizing a physical theory adopted by von Neumann is not uncontroversial though. It mainly consists of providing a mathematical framework together with a manual of how to apply this formalism to obtain experimental predictions. Isham (1995) for example, refers to the framework as providing the rules of quantum mechanics rather than the postulates. This gives it a somewhat instrumentalist flavor. The precise physical meaning of the mathematical symbols in the theory, as well as that of basic terms in the manual such as 'measurement' or 'system', remain largely unexplained.

One may then argue, as for example Bell (1987) did, that the formalism of quantum mechanics given by von Neumann does not deliver what is to be expected of a serious physical theory. To obtain an interpretation of the theory then, one has to build on, or even adjust the postulates to obtain a new and enriched formalism (Muller, 2014). However, here I am not concerned with the interpretation of quantum mechanics. Instead, the aim is to focus on the empirical aspects of the theory, so as to get a grip on necessary constraints for a reformulation of quantum probability. An axiomatic skeleton, then, is a convenient starting point.

Now one may suspect that if one is just interested in the empirical content of the theory, the formalism of von Neumann is just fine. However, this is not the case either. The empirical content and mathematical formalism are tightly intertwined. This implies that, obeying the constraint that possible reformulations of quantum probability respect the quantum mechanical predictions, is not entirely trivial. Therefore, the empirical value of each of the postulates presented in this chapter will be discussed. Possibly this sounds somewhat vague now, but hopefully a clearer picture emerges once discussing these postulates. Before going on there is a final remark though. The postulates presented here do not capture the entirety of

quantum mechanics, nor are all postulates generally accepted. Still, they provide a solid ground for now, and further postulates will be introduced in this dissertation when the discussion requires it.

3.1   States and observables

The first postulate sets the mathematical stage for quantum mechanics:

StaP (State Postulate) Every physical system S is associated with a Hilbert space $\mathcal{H}$. The state of this system is given by a non-zero vector $\psi \in \mathcal{H}$ (pure state) or a density operator ρ acting on $\mathcal{H}$ (mixed state). When two systems S_1 and S_2 are associated with the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, then the joint system is associated with the Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Often, when two vectors ψ_1, ψ_2 give rise to the same projection operator $P_{\psi_1} = P_{\psi_2}$ they are considered to represent the same pure state. In this case the pure states form a proper subset of the set of mixed states. The term mixed state is then usually reserved for density operators that are not projection operators.

Although this first postulate provides much structure, it is still quite silent on what the role of this structure will be for experimental statements. Indeed, it does not state what it means that "ψ is the state of the system", or what it entails. Part of this will be cleared up by further postulates, but a definite answer to these questions is lacking in the formalism. So the answer is open to philosophical debate.

Main themes in this debate are the question whether the state provides a complete characterization of the system, or merely a partial description. And the unpacking of this distinction often involves taking a stance on the question whether the state pertains to a definite property of the system, or concerns merely an epistemic description of the system. In recent years this has become known as the ψ-ontic/ψ-epistemic distinction (Leifer, 2014). The relevance of this distinction to the question of completeness lies in the two senses in which a characterization may be complete. An ontologically complete characterization is something from which all the facts pertaining to a system may be derived, whereas an epistemologically complete characterization is something from which everything that can be known about the system may be derived. Then there are also people who believe that this is the wrong way of framing the debate, as they understand "state of a system" as a short-hand for "a description of an *ensemble* of identically prepared systems". According to this view, the state ψ does not give a characterization of a single system at all.

These observations serve to show that a reserved attitude towards the notion of states in quantum mechanics is advisable. Furthermore, it will become clear later on that taking a stance on their status is tightly connected with one's views on the

meaning of probability in quantum mechanics. And reservation with respect to the latter is precisely what I adopt here as part of the methodology for investigating the formalism of quantum probability.

Besides the primitive use of the notion of a state, it is also noteworthy that StaP does not state how the association of a Hilbert space with a physical system works. It is left to the user of quantum mechanics to find the appropriate Hilbert space for describing the system at hand. There are of course widely accepted conventions. For example, a single particle in three-dimensional space is associated with the Hilbert space $L^2(\mathbb{R}^3)$ of Lebesgue-square integrable functions. The explicit examples that play a role in this dissertation though will usually be simpler. In particular, the prime examples are the Hilbert spaces $\mathbb{C}^2$ associated with a spin-$\frac{1}{2}$ particle, and $\mathbb{C}^3$ associated with a spin-1 particle, both endowed with the standard inner product. More generally, I shall mostly be concerned with finite dimensional Hilbert spaces in this dissertation.

The association of a system with a Hilbert space is of course quite hollow if no further connection is made between possible experiments on the system and the associated Hilbert spaces. The second postulate is a first step towards making this connection.

OP (Observable Postulate) Every observable $\mathcal{A}$ for the system S is associated with a self-adjoint operator A acting on $\mathcal{H}$. If the observable $\mathcal{A}$ for a system S is associated with the operator A and the system is coupled to another system S', then the operator associated with $\mathcal{A}$ for the joint system is $A \otimes \mathbb{1}$, where $\mathbb{1}$ is the unit operator on the Hilbert space $\mathcal{H}'$ associated with S'.

Like "state", the term "observable" occurs as a primitive concept in the theory. The intended reading is that an observable signifies something that is measurable. But how the observable is discerned from the unobservable, or what process leads to something being a measurement is not determined by the theory. Neither does the theory dictate explicitly which operator is to be associated with a certain measurement procedure. However, in practice, physicists are very well capable of envisaging experimental procedures and relating them to self-adjoint operators. There are also several (heuristic) rules for making these relations dating back to the work of Dirac (1925).

In this dissertation I will take the existence of such rules for granted. To be specific, OP will be understood in the following way. For any system there is a set of observables $\mathcal{Obs}$. For each $\mathcal{A} \in \mathcal{Obs}$ it is assumed that there is an experimental procedure that can be understood as a measurement of $\mathcal{A}$. There further is a rule $f : \mathcal{Obs} \rightarrow \mathcal{O}_{\mathrm{sa}}$ which assigns to each observable $\mathcal{A}$ a self-adjoint operator A. It is assumed that this rule is in some sense known. In short, the problems mentioned above play no role in this dissertation. In the case of concrete examples

this approach is also legitimate, for then it is well-established what the relevant observables are, and with which operators they are to be associated. Two of such examples are the following.

Example 3.1. For a spin-$\frac{1}{2}$ particle the operators associated with spin measurements along the z, y and x axis are

$$\sigma_z := \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \ \sigma_y := \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_x := \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.1}$$

If $r = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi)$ is an arbitrary axis, then the observable associated with the spin measurement along that axis is given by

$$\sigma_r := \frac{1}{2} \begin{pmatrix} \cos\phi & e^{-i\theta}\sin\phi \\ e^{i\theta}\sin\phi & -\cos\phi \end{pmatrix}. \tag{3.2}$$

$\triangle$

Example 3.2. For the case of spin-1 particles a similar set of observables is relevant. Here the spin operators are given by

$$S_z := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \ S_y := \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \ S_x := \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{3.3}$$

These are again associated with spin measurements along the z, y and x axis respectively. And similarly one has for an arbitrary axis r that

$$S_r := \begin{pmatrix} \cos\phi & \frac{e^{-i\theta}\sin\phi}{\sqrt{2}} & 0 \\ \frac{e^{i\theta}\sin\phi}{\sqrt{2}} & 0 & \frac{e^{-i\theta}\sin\phi}{\sqrt{2}} \\ 0 & \frac{e^{i\theta}\sin\phi}{\sqrt{2}} & -\cos\phi \end{pmatrix}. \tag{3.4}$$

$\triangle$

The association of observables with self-adjoint operators only goes in one direction. Let $\mathcal{O}_{\text{sa}}^{Obs}$ denote the set of all self-adjoint operators for which there is an observable with which they are associated, i.e., $\mathcal{O}_{\text{sa}}^{Obs} = f(Obs)$. It is not inconceivable that this will be a proper subset of all the self-adjoint operators $\mathcal{O}_{\text{sa}}$. Specifically, what may be questioned is the general validity of the following principle:

IP (Identification Principle) For every self-adjoint operator A there exists an observable $\mathcal{A}$ such that A is associated with $\mathcal{A}$ via OP.

Although IP is often implicitly assumed, it has also been recognized that at best this assumption should be qualified as an idealization.[1] As phrased by Wigner (1963, p. 14): "For some observables, in fact for the majority of them (such as xyp_z), nobody seriously believes that a measuring apparatus exists".[2] Now, as unlikely as it may be that IP is true, it is even more unlikely that one would find a definite precise way to determine what the set $\mathcal{O}_{\text{sa}}^{Obs}$ actually should be. How could one ever declare the definite impossibility of finding an observable to be associated with a certain self-adjoint operator?

There are, however, constructive ways to downsize the set $\mathcal{O}_{\text{sa}}^{Obs}$. One way is by the introduction of so-called superselection rules. These rules were first introduced by Wick, Wightman, and Wigner (1952) following an address presented by Wigner in 1951. The idea is, roughly, that the set of self-adjoint operators associated with observables is trimmed down by simultaneously trimming down the set of vectors in the Hilbert space that can be associated with a state. Superselection rules will be discussed in more detail in the next section.

Wigner seemed to have been quite concerned with the validity of IP. Besides the two contributions mentioned above, he expressed a third critique in (Wigner, 1952) (translation available as (Wigner, 2010)). In this paper the compatibility of IP with von Neumann's theory of measurement processes (von Neumann, 1932, Ch. 6) was investigated. To measure an observable $\mathcal{A}$ on a system S, the system has to be coupled to a measurement apparatus M. The joint system is associated with the Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_M$, and the operator on this space associated with measuring the observable $\mathcal{A}$ on S is $A \otimes \mathbb{1}$. Taking into account the dynamics of the measurement process, it was proven (see also Araki and Yanase, 1960) that a precise measurement of the operator $A \otimes \mathbb{1}$ is only possible if it commutes with the conserved quantities for the total system $S + M$.

These considerations later led to the introduction of a generalized notion of observables in terms of positive operator valued measures (POVMs). The mathematical background of POVMs dates back to (Naimark, 1943), but the earliest application to quantum mechanics I found is by Holevo (1982). POVMs will play no significant role in this dissertation. My claim is that the restriction to self-adjoint operators is not conceptually damaging, but does keep some things simpler. In some possibly confusing cases qualifying remarks will be made though.

---

[1]The assumption is well known, although I have not found a name for it in the literature. I started calling it IP in (Hermens, 2011) for lack of a better name.

[2]Here xyp_z denotes the 'observable' which would correspond to measuring the x and y coordinates of the position of a particle and the z component of its momentum. Classically, one would just make a joint measurement of the position and momentum in this case. But in quantum mechanics, because of the uncertainty relations, this is not possible. It is unclear what other kind of experimental setup then would establish a measurement of this 'observable'.

3.2   Measurement outcomes and their probabilities

The postulates introduced in the previous section determine the mathematical structure of quantum mechanics. But although the concepts of states and observables have also been introduced, no connection with empirical statements has been made thus far. The two postulates introduced in this section make this connection.

VaP (Value Postulate) For an observable $\mathcal{A}$ associated with the self-adjoint operator A via OP, the set of values of possible measurement outcomes is given by the spectrum $\sigma(A)$ of A.

Example 3.3. For the observables in examples 3.1 and 3.2 one may show that the possible measurement outcomes for spin measurements for spin-$\frac{1}{2}$ and spin-1 particles are given by

$$\sigma(\sigma_r) = \{-\tfrac{1}{2}, \tfrac{1}{2}\}, \ \ \sigma(S_r) = \{-1, 0, 1\}. \tag{3.5}$$

Further, the possible outcomes for projection operators are

$$\sigma(P) = \{0, 1\}, \tag{3.6}$$

with exception of the trivial cases $P = \mathbb{0}$ or $P = \mathbb{1}$ where $\sigma(\mathbb{0}) = \{0\}$ and $\sigma(\mathbb{1}) = \{1\}$.   △

  VaP only gives the possible outcomes of a measurement, but does not state which of these outcomes will occur. Famously, quantum mechanics only posits probabilities for these outcomes, and these are given by the Born rule. The formulation of this rule makes use of the spectral theorem. This theorem associates with every self-adjoint operator A a unique projection valued measure (PVM) μ_A. That is, a map that takes measurable subsets of $\sigma(A)$ to projection operators on $\mathcal{H}$ such that $\mu_A(\sigma(A)) = \mathbb{1}$ and $\mu_A(\bigcup_{n \in \mathbb{N}} \Delta_n) = \sum_{n \in \mathbb{N}} \mu_A(\Delta_n)$ for any sequence of pairwise disjoint subsets.

BoP (Born Postulate) If the observable $\mathcal{A}$ is associated with the self-adjoint operator A, and ψ is the state of the system, then the probability to find a value in $\Delta \subset \sigma(A)$ upon a measurement of $\mathcal{A}$ is given by the Born rule[3]

$$\mathbb{P}_\psi(\mathcal{A} \in \Delta) = \frac{\langle \psi, \mu_A(\Delta)\psi \rangle}{\langle \psi, \psi \rangle} = \mathrm{Tr}\left(P_\psi \mu_A(\Delta)\right). \tag{3.7}$$

Alternatively, if the state is given by the density operator ρ, the probability is given by

$$\mathbb{P}_\rho(\mathcal{A} \in \Delta) = \mathrm{Tr}\left(\rho \mu_A(\Delta)\right). \tag{3.8}$$

---

[3]Here, Tr denotes the trace operation.

For two jointly measurable observables $\mathcal{A}_1$ and $\mathcal{A}_2$, corresponding to commuting operators A_1 and A_2, the probability of finding values in $\Delta_1 \times \Delta_2$ for a joint measurement is given by

$$\mathbb{P}_\rho(\mathcal{A}_1 \in \Delta_1, \mathcal{A}_2 \in \Delta_2) = \mathrm{Tr}\left(\rho \mu_{A_1}(\Delta_1)\mu_{A_2}(\Delta_2)\right). \tag{3.9}$$

It may be noted that BoP implies a probabilistic version of VaP: the outcome of a measurement will lie within the spectrum with probability one. Therefore, VaP is often not stated as a separate postulate. For the investigations in this dissertation however, it is useful to keep them separated. VaP is a non-probabilistic postulate, and therefore characterizes aspects of quantum mechanics that are independent of how probability in quantum mechanics is conceived. The constraints it imposes are to be taken seriously, and the non-triviality of this will become apparent, for example, from its role in the Kochen-Specker theorem in chapter 6, and in the development of an empiricist quantum logic in chapter 10. There is also a somewhat historical justification for separating the two postulates. The reproduction of the Rydberg formula by Bohr's atomic model (1913) can be seen as an instantiation of VaP. But BoP played no part in this model and had to wait for more than ten more years before being introduced.

The only role of the quantum state introduced by StaP is the determination of the probabilities in BoP. Without any further role for the quantum state, BoP posits Born's interpretation of the state (sections 2.1–2.3) as a postulate. That is, the quantum state is nothing but a probability function. This then also introduces an equivalence relation on the set of states.

Definition 3.1. Two states ψ_1 and ψ_2 are said to be *statistically equivalent* (with respect to the set of observables *Obs*) if

$$\mathrm{Tr}\left(P_{\psi_1}\mu_A(\Delta)\right) = \mathrm{Tr}\left(P_{\psi_2}\mu_A(\Delta)\right) \tag{3.10}$$

for all $A \in \mathcal{O}_{\mathrm{sa}}^{Obs}$ and all measurable subsets $\Delta \subset \sigma(A)$.

In the case that IP holds, or even if $\mathcal{O}_{\mathrm{sa}}^{Obs}$ is only a finite set of operators short of coinciding with $\mathcal{O}_{sa}$, statistical equivalence is determined by the equivalence relation

$$\psi_1 \sim \psi_2 \text{ iff } \exists \lambda \in \mathbb{C} \text{ s.t. } \psi_1 = \lambda \psi_2 \tag{3.11}$$

on $\mathcal{H}\backslash\{0\}$. The equivalence class generated by ψ is the line spanned by ψ:

$$[\psi] = \{\lambda\psi \,;\, \lambda \in \mathbb{C}, \lambda \neq 0\} \tag{3.12}$$

also called a *ray*. The set of all rays is called the *ray space* and is denoted $\mathcal{R}(\mathcal{H})$.

This construction suggests that when $\mathcal{O}_{\mathrm{sa}}^{Obs}$ becomes smaller, classes of statistically equivalent states become larger. It is in this sense that superselection rules

both deny IP and the idea that every vector corresponds to a pure state. However, this is done in an indirect way, namely, by imposing a certain property on the relation of statistical equivalence. The remainder of this section is devoted to explaining this in more detail. The reason for doing this is twofold. First, it gives an explicit example of how IP may be relaxed. Second, it introduces the language of algebras of operators which will be used throughout this dissertation.

Definition 3.2. An *orthogonal resolution of the identity* is a finite or countable set $\{P_k \mid k \in K\}$ of pairwise orthogonal projection operators such that $\sum_{k \in K} P_k = \mathbb{1}$. An orthogonal resolution of the identity is said to be a *superselection rule* if for all $\psi_1, \psi_2 \in \mathcal{H}$, ψ_1 and ψ_2 are statistically equivalent if and only if there exist numbers $\varphi_k \in [0, 2\pi)$ and $\lambda \in \mathbb{C} \backslash \{0\}$ such that

$$P_k \psi_1 = e^{i\varphi_k} \lambda P_k \psi_2 \ \forall k \in K. \tag{3.13}$$

The vectors ψ_1 and ψ_2 are said to differ by a *relative phase* (with respect to $\{P_k \mid k \in K\}$). The subspaces $P_k \mathcal{H} = \{P_k \psi \mid \psi \in \mathcal{H}\}$ are called *superselection sectors*.

The trivial superselection rule is given by the resolution $\{\mathbb{1}\}$. This corresponds to the ordinary case where two states ψ_1 and ψ_2 are statistically equivalent if and only if $[\psi_1] = [\psi_2]$. In this case it is clear that one can take $\mathcal{O}_{\text{sa}}^{Obs} = \mathcal{O}_{\text{sa}}$. In other cases a superselection rule only implicitly defines what $\mathcal{O}_{\text{sa}}^{Obs}$ is. A useful characterization is available in terms of algebras of operators. For the sake of definiteness, it is useful to first give a fully worked out concrete example.

Example 3.4. Consider the Hilbert space $\mathcal{H} = \mathbb{C}^2$. The set of self-adjoint operators is given by

$$\mathcal{O}_{\text{sa}} = \left\{ \begin{pmatrix} a & c+di \\ c-di & b \end{pmatrix} \ \middle| \ a, b, c, d \in \mathbb{R} \right\}. \tag{3.14}$$

Let $\{P_+, P_-\}$ be the orthogonal resolution of the identity with P_+ the projection on $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and P_- the projection on $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Then, for any $\psi \in \mathbb{C}^2$, the class of vectors that are statistically equivalent to ψ is given by

$$[\psi] = \{\lambda \psi_\varphi \mid \lambda \in \mathbb{C} \backslash \{0\}, \varphi \in [0, 2\pi)\}, \ \ \psi_\varphi := P_+ \psi + e^{i\varphi} P_- \psi. \tag{3.15}$$

Indeed, the vectors of the form ψ_φ are precisely those that only differ by a relative phase from ψ.

Now, for $A \in \mathcal{O}_{\text{sa}}$ to be an observable (with respect to the given superselection rule), it should satisfy

$$\langle \psi, A\psi \rangle = \langle \psi_\varphi, A\psi_\varphi \rangle, \ \forall \psi \in \mathcal{H}, \varphi \in [0, 2\pi). \tag{3.16}$$

Writing $\psi = (x, y)$, this amounts to the constraint

$$c\mathrm{Re}(x\bar{y}) + d\mathrm{Im}(x\bar{y}) = c\mathrm{Re}(e^{-i\varphi}x\bar{y}) + d\mathrm{Im}(e^{-i\varphi}x\bar{y}), \ \forall x, y, \varphi, \tag{3.17}$$

which only holds if $c = d = 0$. Then, the biggest possible set of self-adjoint operators associated with observables is

$$\mathcal{O}_{\mathrm{sa}}^{Obs} = \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \ \middle| \ a, b \in \mathbb{R} \right\}. \tag{3.18}$$

$$\triangle$$

In the above example the set of observables corresponds with the set of all self-adjoint operators in the algebra generated by the projections in the orthogonal resolution of the identity. Specifically, one has

$$\mathcal{A}\!\mathit{lg}(P_+, P_-) = \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \ \middle| \ a, b \in \mathbb{C} \right\} \ \text{and} \ \mathcal{O}_{\mathrm{sa}}^{Obs} = \mathcal{O}_{\mathrm{sa}} \cap \mathcal{A}\!\mathit{lg}(P_+, P_-). \tag{3.19}$$

This is always the case when all the projections in the superselection rule are 1-dimensional. More generally, for a superselection rule $\{P_k \mid k \in K\}$, the set of self-adjoint operators corresponding to observables is determined by the commutant of the set of projections as follows:

$$\mathcal{O}_{\mathrm{sa}}^{Obs} = \mathcal{O}_{\mathrm{sa}} \cap \{P_k \mid k \in K\}' = \{A \in \mathcal{O}_{\mathrm{sa}} \mid [P_k, A] = 0 \ \forall k \in K\}. \tag{3.20}$$

Thus a superselection rule reduces the set of 'relevant' operators on a Hilbert space to $\{P_k \mid k \in K\}'$ which generally is a proper sub-*-algebra of all operators.

Thus far I have discussed superselection rules as restrictions on the set of observables. But as noted before, another common way of introducing them is not as constraints on the observables, but as constraints on the states, or as a constraint on the superposition principle[4]. At first sight this seems to hint at a totally different construction because in the above approach every state still gives rise to a legitimate probability function. As it turns out, though, the two notions are mathematically equivalent (although the one given above is more interpretation neutral). This deserves some clarification.

Let $\{P_k \mid k \in K\}$ be an orthogonal resolution of the identity. Then a superselection rule (for states) for this resolution is the restriction that an element of $\mathcal{H}$ is a state if and only if it is an element of one of the superselection sectors. Thus, the only pure states are those from the set

$$\mathcal{H}_{\mathrm{states}} = \bigcup_{k \in K} P_k \mathcal{H}. \tag{3.21}$$

---

[4]This is the principle that states that for every pair of (pure) states any linear combination of the two is again a (pure) state.

This will generally be a proper subset of the entire Hilbert space $\mathcal{H}$. So what justifies getting rid of all the other states? The answer is that one still allows mixed states of the states in $\mathcal{H}_{\text{states}}$, and these are able to generate all the probability distributions that can be given by the pure states that are thrown out.

An explicit construction elucidates this point. Let $\psi \in \mathcal{H}$ be an arbitrary unit vector and set $\psi_k := P_k \psi$. One can associate ψ with the density operator

$$\rho_\psi := \sum_{k \in K} \|\psi_k\|^2 P_{\psi_k}. \tag{3.22}$$

One may now show that $\psi \in \mathcal{H} \backslash \mathcal{H}_{\text{states}}$ if and only if for every $A \in \mathcal{O}_{\text{sa}}^{Obs}$ (with the set of observables given by (3.20)) and for every measurable set $\Delta \subset \sigma(A)$

$$\text{Tr}(P_\psi \mu_A(\Delta)) = \text{Tr}(\rho_\psi \mu_A(\Delta)). \tag{3.23}$$

Thus the observables cannot be used to distinguish between pure and mixed states. In other words, ψ and ρ_ψ are statistically equivalent (stretching up Definition 3.1 here to also include mixed states).

By allowing mixed states, the set of probability functions remains the same, and from this probabilistic perspective the set of states is not really made smaller by the superselection rule. The accompanying view though, is that all the mixed states of the form ρ_ψ are given an ignorance interpretation, and that the 'actual' pure state is an element of $\mathcal{H}_{\text{states}}$. It is this interpretation of mixed states that fits best with this view on superselection rules. However, I shall not adopt it here as it complicates matters. It requires that pure states are the kinds of objects for which it makes sense to assign probabilities to. But the pure states themselves in turn are also associated with probability distributions. Adopting an epistemic interpretation with respect to the first kind of probability then suggests a more ontological interpretation of probability of the second kind.

This view is not forced upon us. In fact, BoP does not state what is meant with probability in quantum mechanics. Thus the ontological status of both quantum probability and the quantum state is still open for debate. Which position one adopts in this debate is tightly related to how one conceives of the postulates introduced in the next section.

3.3 The status of states

Often the occurrence of probability as a primitive concept in a theory is taken to suggest that it should be given an objective or ontological status. This suggestion at least was strong enough for Popper (1957) to develop a new interpretation of probability. But as seen in chapter 2, the conclusion was not that straightforward for the physicists who worked on the theory at the time. As de Broglie (1960, p. 29)

reflects: "The more the formalism of employing the Ψ wave became apparent, the more it appeared as a kind of formal and subjective representation making possible the evaluation of the probabilities of certain results of measurement." Traits of this epistemic view[5] were present in the writing of both endorsers and opponents of the Copenhagen school (Shimony, 1983). This is not entirely unsurprising, as an epistemic stance helps in understanding why traditionally there have been two postulates to describe the dynamics of quantum systems:

SchröP (Schrödinger Postulate) When no measurement is performed between time points t_1 and t_2 the state evolves unitarily, that is,

$$\psi(t) = U_{t-t_1}\psi(t_1),\ t \in [t_1, t_2] \tag{3.24}$$

for some (strongly continuous) group homomorphism $t \mapsto U_t$ from the reals to the unitary operators on $\mathcal{H}$.[6]

ProP (Projection Postulate) When a measurement of $\mathcal{A}$ is performed and the result is maximally specified by the set $\Delta \subset \sigma(A)$, then the state changes discontinuously following the rule

$$\psi \mapsto \mu_A(\Delta)\psi. \tag{3.25}$$

If the state is given by the density operator ρ, then the state changes discontinuously following the rule

$$\rho \mapsto \frac{\mu_A(\Delta)\rho\mu_A(\Delta)}{\mathrm{Tr}(\rho\mu_A(\Delta))}. \tag{3.26}$$

Adopting an epistemic reading of the quantum state, SchröP can be seen to take into account that the system changes over time, while ProP takes into account that one should update (collapse) the state in accordance with newly obtained information. But the epistemic view of quantum states is not endorsed by everyone. One argument against the view is that on this view it is not clear what the information is about, other than measurements and their outcomes. If one aspires a realist interpretation, this situation may be unsettling. On the other hand, on an ontological view of quantum states there is at least something that is physical within the theory (other than measurements and outcomes). However, on this view the existence

---

[5] I take it that where de Broglie uses 'subjective' this should be understood as 'epistemic'. On other occasions he speaks of Ψ as representing a state of knowledge. The interchangeable use of the epistemic/ontological distinction and the subjective/objective distinction is ubiquitous in the philosophy of physics and probability. Every now and then this leads to confusion, as noted for example by Jaynes (1985).

[6] By Stone's theorem (Stone, 1930 ; 1932), this statement is equivalent to the statement that $\psi(t)$ is a solution to the time-dependent Schrödinger equation.

of two dynamical postulates is quite mysterious. It introduces a sharp distinction between physical processes worthy of the name 'measurement' and 'ordinary' processes. An ontological view then often goes hand in hand with a rejection of ProP. These are the so-called no-collapse interpretations such as Bohmian mechanics and Everettian quantum mechanics. A notable exception is the proposal of Wigner (1961) who adopts Cartesian dualism and suggests that the discontinuous change in ProP is caused by the interaction of matter with the mind. Another option is to replace both postulates with a single time-evolution such as in the approach of Ghirardi, Rimini, and Weber (1985 ; 1986).

The growing number of interpretations of quantum mechanics in which epistemic readings of quantum states are rejected in itself does not explain the growing support for this rejection. From a philosophical point of view, much support may be drawn from the work of Bell (1987), who pressed the importance of coming up with solutions to the measurement problem and introducing an ontology for quantum mechanics. From the physical point of view, the rise of quantum cosmology can be seen as an important influence to the growing unease with epistemic interpretations (Wheeler, 1977). After all, an epistemic reading of quantum states requires the separation between subject and object: something which is not easily done when the object is to be the entire universe. Both Bohmian mechanics and Everettian quantum mechanics, on the other hand, have no problem with the notion of a quantum state of the universe.

More recently, the distinction between epistemic and ontological interpretations of quantum states has become the topic of formal investigations. In the important work by Pusey, Barrett, and Rudolph (2012) it is shown that "any model in which a quantum state represents mere information about an underlying physical state of the system, and in which systems that are prepared independently have independent physical states, must make predictions that contradict those of quantum theory." And, as is tradition, experimental tests are found to be in line with quantum predictions (Nigg et al., 2012). The paper by Pusey, Barrett, and Rudolph has spawned a lot of follow-up research, as well as philosophical debate.[7] Noteworthy contributions are the works of Colbeck and Renner (2011 ; 2012a), Hardy (2013), Barrett et al. (2014), and the overview paper by Leifer (2014). But it should be noted that none of the results thus far entirely rule out epistemic interpretations of quantum states. Nor could they. An interpretation alone cannot alter the predictions of a theory, and auxiliary assumptions will always be needed.

It is tempting to think that an ontological view of quantum states goes hand in hand with an interpretation of quantum probability as objective chance. This could be argued for by using Born's identification of quantum states with probability

---

[7]On April 16th 2015 Google scholar listed 240 entries citing the paper by Pusey, Barrett, and Rudolph (2012).

functions. This identification, however, is not maintained in all interpretations of quantum mechanics. In Bohmian mechanics, for example, the state also plays a role in determining the path of the particles.[8] Also, in some approaches to the Everett interpretation, the Born rule does not explicate what quantum states are, but rather is a normative rule for how rational agents should set their credences provided they know the state.[9] In both these examples an epistemic interpretation of quantum states is rejected while an epistemic interpretation of quantum probabilities may still be tenable.

As seen, what counts as the postulates of quantum mechanics may depend on one's interpretation of quantum states and probabilities. An empiricist interpretation-neutral investigation of quantum probability based solely on the formalism of quantum mechanics then may seem hopeless, as the formalism itself is not neutral. However, in this dissertation the focus is on the static aspect of quantum mechanics, as presented in the first two sections. This helps in maintaining neutrality concerning the ontological status of quantum states. Furthermore, like with classical probability, many aspects of quantum probability can be investigated without delving into the dynamics of the system under investigation. There is, however, room for discussing dynamics for probability functions like classical conditionalization and its relation to ProP. But in these cases judgment is suspended about whether the change in the probability function corresponds to a physical process. Now then let us turn to this formalism of quantum probability.

---

[8]This role has often been taken to suggest that the quantum state should have the status of a physical field. This view has grown somewhat out of fashion among proponents of the interpretation, and the quantum state is now often given a nomological status (Esfeld et al., 2013). At any rate, it *is* being deprived of its purely epistemic status.

[9]The status of probability in Everettian quantum mechanics is a subtle matter and cannot be portrayed justly in one sentence. For details see (Wallace, 2012) and references therein, and for a critical assessment see (Kent, 2010).

4

Formalism of quantum probability

Classical probability theory can be thought of as a field in mathematics, sometimes characterized as "measure theory + independence". One can work through a course on measure-theoretic probability without ever having to think about probability as a notion that corresponds to something outside the realm of mathematics. And the same holds true for quantum probability. The aim of this chapter is to give a short introduction to the mathematical formalism of quantum probability. For thorough introductions into the field of quantum probability the reader may consult (Parthasarathy, 1992), (Meyer, 1993), and (Cuculescu and Oprea, 1994). A compact introduction is provided by Maassen (2010).

Instead of just starting with the formal definition of a quantum probability space, I take some time to distill it from the postulates presented in the previous chapter so as to make the relation to quantum mechanics more apparent. In section 4.1 a definition of quantum probability spaces is given that focuses on the use of classical probability spaces in quantum mechanics. Despite the importance of the use of classical probability in quantum mechanics, the two notions of classical and quantum probability are quite distinct. This is explained by means of an example based on the violation of Bell-type inequalities in quantum mechanics. A definition of quantum probability spaces that is more common, and that highlights the formal analogy between classical and quantum probability, is given in section 4.2. It is shown that the two definitions are equivalent. In this section Gleason's theorem, which provides a relation between quantum states and quantum probability functions, is discussed as well. Finally, in section 4.3, a more general definition of quantum probability spaces is given that makes use of the language of algebras. The upshot of this generalization is that both classical and (Hilbert space) quantum probability can be viewed as special instances of this more general formalism. A theorem that establishes this embedding of classical probability into quantum probability is presented and discussed.

4.1 Classical and quantum probability

A good introduction to the formalism of quantum probability is not possible without some background in classical probability. This is because classical probability spaces play an important role in the characterization of quantum probability spaces. Instead then of directly giving the definition of a quantum probability space, I start with an elaboration that is helpful for understanding why quantum probability spaces are defined in a particular way. The methodology adopted is one that is standard in mathematics. I start with a few intuitive properties one can plausible desire to hold for a quantum probability space, and then take these properties as the definition. Thus, for the sake of definiteness, I start with the definition of classical probability spaces and random variables.

Definition 4.1. A *(classical) probability space* is a triple $(\Omega, \mathcal{F}, \mathbb{P})$ with Ω a set, $\mathcal{F}$ a σ-algebra of subsets of Ω called the set of *events*, and $\mathbb{P} : \mathcal{F} \to [0, 1]$ a probability measure, i.e., a function such that $\mathbb{P}(\Omega) = 1$ and

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} \Delta_n\right) = \sum_{n \in \mathbb{N}} \mathbb{P}(\Delta_n) \tag{4.1}$$

for every countable sequence $(\Delta_n)_{n \in \mathbb{N}}$ of pairwise disjoint sets in $\mathcal{F}$. If $(\Omega_X, \mathcal{F}_X)$ is a measurable space and $X : \Omega \to \Omega_X$ is a measurable function, then X is called an *(classical) Ω_X-valued random variable.*

The formalism of quantum mechanics gives rise to classical classical probability spaces in a natural way. Let A be a self-adjoint operator. With this operator one can associate the measurable space $(\Omega_A, \mathcal{F}_A)$ with $\Omega_A := \sigma(A)$ the spectrum of A, and $\mathcal{F}_A$ the Borel σ-algebra of subsets of Ω_A. Now every quantum state ρ gives rise to a probability measure $\mathbb{P}_\rho^A$ on $(\Omega_A, \mathcal{F}_A)$ via the rule

$$\mathbb{P}_\rho^A(\Delta) := \text{Tr}(\rho \mu_A(\Delta)). \tag{4.2}$$

There is a sense in which one can think of a self-adjoint operator as somehow representing a quantum version of a random variable. The analogy becomes more apt when noting that a classical random variable $X : \Omega \to \Omega_X$ can also be characterized by the function $X^{-1} : \mathcal{F}_X \to \mathcal{F}$, where $X^{-1}(\Delta) := \{\omega \in \Omega \mid X(\omega) \in \Delta\}$. Similarly, the self-adjoint operator A can be characterized by the PVM $\mu_A : \mathcal{F}_A \to L(\mathcal{H})$. This suggests that $L(\mathcal{H})$ may be taken as the domain of what is to be a quantum probability function.

A second motivation for this particular domain is drawn from the following observation. The numerical value of the probability of a measurement outcome assigned by the Born rule depends only on the projection operator $\mu_A(\Delta)$, and

doesn't further depend on the observable $\mathcal{A}$ or the specific numbers in the spectrum $\sigma(A)$. Specifically, if $\mathcal{A}$ is an observable associated with a self-adjoint operator A with spectrum $\sigma(A)$ and $f : \sigma(A) \to \mathbb{R}$ is an injective function, then an observable $\mathcal{A}'$ associated with the self-adjoint operator $A' := f(A)$ has the same probability distribution associated with it as $\mathcal{A}$:

$$\mathbb{P}_\rho(\mathcal{A} \in \Delta) = \mathbb{P}_\rho(\mathcal{A}' \in f(\Delta)) \ \forall \rho, \forall \Delta \subset \sigma(A). \tag{4.3}$$

Thus the numerical values of probabilities only depend on the projection operators P_ρ and $\mu_A(\Delta)$, i.e.,

$$\mathbb{P}_\rho(\mathcal{A} \in \Delta) = \mathrm{Tr}(P_\rho \mu_A(\Delta)). \tag{4.4}$$

These considerations suggest that a quantum probability function is to be a function on $L(\mathcal{H})$ that provides a classical probability space for every self-adjoint operator on $\mathcal{H}$. This is indeed the definition I shall work with for now. Later on, I shall show that it is equivalent to a definition that is more common in the literature.

Definition 4.2. A *quantum probability space* is a triple $(\mathcal{H}, L(\mathcal{H}), \mathbb{P})$ with $\mathcal{H}$ a Hilbert space, $L(\mathcal{H})$ the set of projection operators on $\mathcal{H}$, called the set of *events*, and $\mathbb{P} : L(\mathcal{H}) \to [0,1]$ a function such that for every self-adjoint operator A the function

$$\mathbb{P}^A(\Delta) := \mathbb{P}(\mu_A(\Delta)) \tag{4.5}$$

turns $(\Omega_A, \mathcal{F}_A, \mathbb{P}^A)$ into a classical probability space. An Ω_X-*valued quantum random variable* on a quantum probability space $(\mathcal{H}, L(\mathcal{H}), \mathbb{P})$ is a measurable space $(\Omega_X, \mathcal{F}_X)$ together with a PVM $\mu_X : \mathcal{F}_X \to L(\mathcal{H})$.

Note that for every quantum state ρ the map $\mathbb{P}_\rho : L(\mathcal{H}) \to [0,1]$, given by

$$\mathbb{P}_\rho(P) := \mathrm{Tr}(\rho P), \tag{4.6}$$

indeed turns $(\mathcal{H}, L(\mathcal{H}))$ into a quantum probability space. The analogy between classical and quantum probability thus far can be summed up by the following table:

	Events	Probability function	Random variable	States
Classical	$\mathcal{F}$	$\mathbb{P} : \mathcal{F} \to [0,1]$	$X^{-1} : \mathcal{F}_X \to \mathcal{F}$	Ω
Quantum	$L(\mathcal{H})$	$\mathbb{P} : L(\mathcal{H}) \to [0,1]$	$\mu_X : \mathcal{F}_X \to L(\mathcal{H})$	?

The analogy breaks down when trying to point out a quantum version of the set Ω. The candidate for a quantum analog would of course be the Hilbert space $\mathcal{H}$. However, there is no function $X : \mathcal{H} \to \Omega$ of which the PVM μ_X can be considered to be the inverse. This is because quantum states do not determine the values of observables. In this sense a quantum probability space is like a classical

probability space without a set of points Ω (or without a "set of possible worlds"). The omission of such a set in the formalism of course doesn't imply the impossibility of introducing such a set. Constraints on such possibilities are considered in part II of this dissertation. But for now it is useful to consider one particular example that illustrates that the question of the existence of such a set is not trivial.[1] First consider the following theorem in classical probability.

Theorem 4.1. *Let $(\Omega, \mathcal{F})$ be a measurable space and A_1, A_2, B_1, B_2 be four 0,1-valued random variables. Then every probability measure $\mathbb{P}$ satisfies the Bell-type inequality*

$$\mathbb{P}(A_1 = B_1) \leq \mathbb{P}(A_1 = B_2) + \mathbb{P}(A_2 = B_1) + \mathbb{P}(A_2 = B_2). \tag{4.7}$$

Proof. Although the proof can be done easily by writing out, it is insightful to take a pictorial approach. The probabilities for the four random variables for any probability function can be characterized using the Karnaugh map[2]

$$
\begin{array}{c}
\quad\quad A_1 \\
B_1 \left|
\begin{array}{|c|c|c|c|}
\hline
p_1 & p_2 & p_3 & p_4 \\
\hline
p_5 & p_6 & p_7 & p_8 \\
\hline
p_9 & p_{10} & p_{11} & p_{12} \\
\hline
p_{13} & p_{14} & p_{15} & p_{16} \\
\hline
\end{array}
\right| B_2 \\
\quad\quad A_2
\end{array}
\tag{4.8}
$$

with $p_i \geq 0$ for all i and $\sum_i p_i = 1$. Every square fixes the values of all the random variables and the value of p_i denotes the probability of these variables having these values. For example

$$p_3 = \mathbb{P}(B_1 = A_2 = 1, A_1 = B_2 = 0). \tag{4.9}$$

---

[1] The example is a mathematical view on the violation of Bell inequalities in quantum mechanics. This analysis is based on (Kümmerer and Maassen, 1998), but the pictorial proof of Theorem 4.1 is my own.

[2] These graphical tools were first introduced by Karnaugh (1953) for their convenience with studying logic circuits. The key difference with Venn diagrams is that while the latter focuses on atomic propositions, Karnaugh maps put more emphasis on the atoms in the Boolean lattice of propositions (i.e., the 'maximal' conjunctions). The advantage of using Karnaugh maps in connection with probability theory, is that a specification of the probabilities of all the atoms (corresponding to the 16 smallest squares in (4.8)) fully determines the probability function on the Boolean lattice, whereas the same does not hold for a specification of the probabilities of all atomic propositions (corresponding to the 4 rectangles specified by A_1, A_2, B_1, B_2 in (4.8)).

The events $A_i = B_j$ can then be identified with colorings of the Karnaugh map, where their probabilities correspond to the sum of all the colored p_i's. The inequality (4.7) is then rewritten as

This shows that the inequality holds because every square that is colored on the left is at least colored once on the right. $\qquad\square$

The crux is that for random variables on a quantum probability space the inequality (4.7) need not hold. The following example shows this.

Example 4.1. Consider the Hilbert space $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$. Four 0,1-valued random variables are introduced based on the projection operators

$$P_\varphi := \begin{pmatrix} \cos^2 \varphi & \cos \varphi \sin \varphi \\ \cos \varphi \sin \varphi & \sin^2 \varphi \end{pmatrix} \tag{4.10}$$

on $\mathbb{C}^2$ with $\varphi \in [0, 2\pi)$. Define the projection operators

$$A_1 := P_{\alpha_1} \otimes \mathbb{1}, \ A_2 := P_{\alpha_2} \otimes \mathbb{1}, \ B_1 := \mathbb{1} \otimes P_{\beta_1}, \ B_2 := \mathbb{1} \otimes P_{\beta_2}, \tag{4.11}$$

where $\mathbb{1}$ denotes the unit operator on $\mathbb{C}^2$. For any pair i, j, the pair of operators A_i, B_j commute and so BoP defines a joint probability distribution for them. Now consider the quantum state

$$\psi := \frac{1}{2}\sqrt{2} \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right). \tag{4.12}$$

One then finds that

$$\begin{aligned}
\mathbb{P}_\psi(A_i &= B_j) \\
&= \left\langle \psi, (P_{\alpha_i} \otimes P_{\beta_j})\psi \right\rangle + \left\langle \psi, \left((\mathbb{1} - P_{\alpha_i}) \otimes (\mathbb{1} - P_{\beta_j})\right) \psi \right\rangle \\
&= \frac{1}{2} \left(\cos^2 \alpha_i \sin^2 \beta_j + \sin^2 \alpha_i \cos^2 \beta_j - 2 \cos \alpha_i \sin \alpha_i \cos \beta_j \sin \beta_j \right) \\
&\quad + \frac{1}{2} \left(\cos^2 \alpha_i \sin^2 \beta_j + \sin^2 \alpha_i \cos^2 \beta_j - 2 \cos \alpha_i \sin \alpha_i \cos \beta_j \sin \beta_j \right) \\
&= (\cos \alpha_i \sin \beta_j - \sin \alpha_i \cos \beta_j)^2 = \sin^2(\alpha_i - \beta_j).
\end{aligned} \tag{4.13}$$

When setting $\alpha_1 = 0$, $\alpha_2 = \frac{1}{3}\pi$, $\beta_1 = \frac{1}{2}\pi$, $\beta_2 = \frac{1}{6}\pi$ one obtains

$$\mathbb{P}_\psi(A_1 = B_1) = 1, \ \mathbb{P}_\psi(A_1 = B_2) = \mathbb{P}_\psi(A_2 = B_1) = \mathbb{P}_\psi(A_2 = B_2) = \frac{1}{4}. \quad (4.14)$$

This establishes a contradiction with (4.7). $\triangle$

The discrepancy between classical and quantum probability can be seen to be a consequence of the existence of the space Ω in the classical case. The existence of this set implies that the joint probabilities for each pair A_i, B_j are marginals of a distribution over the possible values for A_1, A_2, B_1, B_2. In the quantum case only the joint probabilities for commuting pairs are defined. The violation of Bell-type inequalities can be associated with the non-existence of a total probability function (Fine, 1982b,a), which in turn can be associated with the non-existence of an underlying space Ω for quantum probability.

It should be noted that this example does not really demonstrate the incompatibility of classical probability with quantum probability. For example, the non-existence of a joint distribution for the three variables A_1, A_2, B_1 may be seen as evidence that it is a wrong-headed idea that B_1 should be modeled with the same classical random variable when taken in consideration together with A_1 as when taken in consideration together with A_2. Indeed, if one uses eight classical random variables

$$A_1^{11}, A_1^{12}, A_2^{21}, A_2^{22}, B_1^{11}, B_1^{21}, B_2^{12}, B_2^{22}, \quad (4.15)$$

with the pair (A_i^{ij}, B_j^{ij}) as the classical representation of the quantum pair (A_i, B_j), one can violate (4.7). The possibility of classical representations will be investigated more deeply in part II. But before that it is good to first delve a bit deeper into the formalism of quantum probability.

4.2   Quantum probability and Gleason's theorem

As noted in the previous section, Definition 4.2 is a non-standard definition of quantum probability. In this section it is shown to be equivalent to a more common definition such as found, for example, in (Gudder, 1979; Parthasarathy, 1992). This more common definition is also the one that will be used the most throughout this dissertation. It has the benefit that it shows a strong structural resemblance to the definition of a classical probability space. On the other hand, in comparison to Definition 4.2, its connection with actual classical probability spaces is less clear. The benefit of having both definitions, and establishing their equivalence, is that it helps to get a firmer grip on the topic. The theorem by Gleason, presented in this section as well, of course also aids in getting a firmer grip on the topic. But this fact alone does not do justice to its importance. As noted in the previous section,

$(\mathcal{H}, L(\mathcal{H}), \mathbb{P})$ is a quantum probability space whenever $\mathbb{P}$ is given by (4.6) for some density operator ρ. Gleason's theorem establishes that this is the generic case.[3] But first the theorem that relates Definition 4.2 to the standard conception of quantum probability spaces.

Theorem 4.2. *A triple $(\mathcal{H}, L(\mathcal{H}), \mathbb{P})$ is a quantum probability space if and only if $\mathbb{P} : L(\mathcal{H}) \to [0,1]$ satisfies*

(i) $\mathbb{P}(\mathbb{1}) = 1$.

(ii) $\mathbb{P}(P_1 + P_2 + \ldots) = \mathbb{P}(P_1) + \mathbb{P}(P_2) + \ldots$ *for any countable sequence of pairwise orthogonal projection operators.*

Proof. Suppose $(\mathcal{H}, L(\mathcal{H}), \mathbb{P})$ is a quantum probability space. I shall first show that it satisfies criterion (ii), and then derive that (i) also holds. Let $P_1, P_2, \ldots$ be a countable sequence of pairwise orthogonal projections. Define $P_0 := \mathbb{1} - (P_1 + P_2 + \ldots)$ and let

$$A := \sum_{n=0}^{\infty} \frac{1}{n+1} P_n. \tag{4.16}$$

Then A is a self-adjoint operator with

$$\sigma(A) = \left\{ \frac{1}{n+1} \,\middle|\, n \in \mathbb{N} \right\} \tag{4.17}$$

and the PVM μ_A given by

$$\mu_A(\Delta) = \sum_{\left\{ n \in \mathbb{N} \,\middle|\, \frac{1}{n+1} \in \Delta \right\}} P_n. \tag{4.18}$$

Let $\Delta_0 := \sigma(A) \backslash \{1\}$. Then

$$\mathbb{P}(P_1 + P_2 + \ldots) = \mathbb{P}(\mu_A(\Delta_0)) = \mathbb{P}^A(\Delta_0)$$
$$= \sum_{n=1}^{\infty} \mathbb{P}^A(\{\tfrac{1}{n+1}\}) = \sum_{n=1}^{\infty} \mathbb{P}(\mu_A(\{\tfrac{1}{n+1}\})) = \sum_{n=1}^{\infty} \mathbb{P}(P_n). \tag{4.19}$$

Finally,

$$\mathbb{P}(\mathbb{1}) = \mathbb{P}(P_0 + P_1 + \ldots) = \mathbb{P}(\mu_A(\sigma(A))) = \mathbb{P}^A(\sigma(A)) = 1. \tag{4.20}$$

For the converse, suppose $\mathbb{P}$ satisfies (i) and (ii). Let A be any self-adjoint operator with spectrum $\Omega_A := \sigma(A)$ and associated PVM μ_A. Then $\mathbb{P}^A$ satisfies

$$\mathbb{P}^A(\Omega_A) = \mathbb{P}(\mu_A(\sigma(A))) = \mathbb{P}(\mathbb{1}) = 1. \tag{4.21}$$

---

[3]There is a peculiar exception for the case when $\mathcal{H}$ is 2-dimensional.

For any countable sequence $(\Delta_n)_{n \in \mathbb{N}}$ of mutually disjoint measurable sets, one has

$$\mathbb{P}^A \left(\bigcup_{n=0}^{\infty} \Delta_n \right) = \mathbb{P} \left(\mu_A \left(\bigcup_{n=0}^{\infty} \Delta_n \right) \right) = \mathbb{P} \left(\sum_{n=0}^{\infty} \mu_A (\Delta_n) \right)$$

$$= \sum_{n=0}^{\infty} \mathbb{P} \left(\mu_A (\Delta_n) \right) = \sum_{n=0}^{\infty} \mathbb{P}^A (\Delta_n). \tag{4.22}$$

Hence $(\Omega_A, \mathcal{F}_A, \mathbb{P}^A)$ is a classical probability space. $\qquad \square$

This formulation of a quantum probability space shows a resemblance with the classical Definition 4.1. It may also be used to give a characterization of all quantum probability spaces in terms of so-called frame functions. This is useful, as frame functions will also play a significant role in later chapters. So without further ado, here's the definition.

Definition 4.3. Let $\mathcal{H}$ be a Hilbert space and let $L_1(\mathcal{H})$ denote the set of one-dimensional projection operators. A *frame* is a sequence (P_i) of pairwise orthogonal projections in $L_1(\mathcal{H})$ with $\sum_i P_i = \mathbb{1}$. A *frame function* is a function $\lambda : L_1(\mathcal{H}) \to \mathbb{R}$ such that[4]

$$\sum_i \lambda(P_i) = 1, \tag{4.23}$$

for every frame (P_i).

Equivalently, a frame function can be taken to be a function on the unit vectors of a Hilbert space such that its value sums to 1 for every orthonormal basis. It is clear that every quantum probability function $\mathbb{P}$ determines a non-negative frame function simply by restricting $\mathbb{P}$ to $L_1(\mathcal{H})$. The converse is true as well, although this is less trivial. When the dimension of the Hilbert space is 1 or 2 this is easy to see. If $\dim(\mathcal{H}) = 1$, then $L_1(\mathcal{H}) = \{\mathbb{1}\}$ and hence there is only one frame function, and only one probability function. In the two-dimensional case $L_1(\mathcal{H})$ coincides with $L(\mathcal{H})$ except for the elements $\mathbb{0}$ and $\mathbb{1}$, i.e., $L(\mathcal{H}) = L_1(\mathcal{H}) \cup \{\mathbb{0}, \mathbb{1}\}$. For a frame function λ, the probability function is then given by

$$\mathbb{P}_\lambda(\mathbb{0}) := 0, \ \ \mathbb{P}_\lambda(\mathbb{1}) := 1 \ \text{ and } \ \mathbb{P}_\lambda(P) := \lambda(P) \ \text{ otherwise.} \tag{4.24}$$

In other cases the proof is non-trivial and relies on the following lemma by Gleason (1957).

---

[4]Often, this is called a frame function of weight 1. More generally one can speak of frame functions of weight W with $W \in \mathbb{R}$. These are defined to have the property that every sum in (4.23) adds up to W.

Quantum probability and Gleason's theorem **4.2**

Lemma 4.1 (Gleason 1957). *Let $\mathcal{H}$ be a Hilbert space with $\dim \mathcal{H} > 2$. Then for every non-negative frame function λ there exists a density operator ρ_λ such that*

$$\lambda(P) = \mathrm{Tr}\,(\rho_\lambda P) \ \ \forall P \in L_1(\mathcal{H}). \tag{4.25}$$

From this lemma it is clear that every non-negative frame function λ gives rise to a function $\mathbb{P}_\lambda$ on $L(\mathcal{H})$ by

$$\mathbb{P}_\lambda(P) := \mathrm{Tr}\,(\rho_\lambda P). \tag{4.26}$$

The fact that this defines a quantum probability function then follows from the additivity of the trace. The culmination of these considerations[5] is Gleason's theorem.

Theorem 4.3 (Gleason 1957). *Let $\mathcal{H}$ be a Hilbert space with dimension unequal to 2. Then $(\mathcal{H}, L(\mathcal{H}), \mathbb{P})$ is a quantum probability space if and only if there exists a density operator ρ such that*

$$\mathbb{P}(P) = \mathrm{Tr}(\rho P) \ \forall P \in L(\mathcal{H}). \tag{4.27}$$

Two formal remarks are in order. First, this theorem only holds when the Hilbert space is either finite-dimensional or separable. A generalization of Gleason's theorem that applies to non-separable Hilbert spaces was given by Eilers and Horst (1975). In this approach the notion of a quantum probability space is further constrained by replacing (ii) in Theorem 4.2 by a condition that also allows arbitrary (non-denumerable) sums. Second, as noted in section 3.1, the notion of PVMs can be generalized to POVMs. If done so, quantum probability functions are also generalized to stretch their domain to incorporate a class of positive operators known as effects.[6] A version of Gleason's theorem for this case has been given by Busch (2003). The upshot of that result for present purposes is that the class of quantum probability spaces remains the same, i.e., they are still characterized by density operators. A version of the theorem that holds for arbitrary von Neumann algebras is presented in the next section.

Now let me shortly recapitulate what has been shown thus far in this section. It follows from Theorem 4.2 that every quantum probability function $\mathbb{P}$ gives rise to a frame function λ. Conversely, Lemma 4.1 establishes that, if $\dim(\mathcal{H}) > 2$, every frame function gives rise to a quantum probability function through the introduction of a density operator. Theorem 4.3 shows that this correspondence is bijective. It is not hard to find counter examples for the case that $\dim(\mathcal{H}) = 2$. As noted above, every frame function determines a probability function by (4.24). In particular, λ

---

[5]Together with some details such as the fact that two quantum probability functions are identical when they give rise to the same frame function.

[6]An effect is a positive operator E with $\sigma(E) \subset [0,1]$.

need not be continuous.[7] However, every frame function induced by a probability function of the form (4.27) is continuous. So the discontinuity of λ is a sufficient condition for $\mathbb{P}_\lambda$ to be a counter example to Gleason's theorem.

Because the representation of quantum probability functions by density operators only fails for the case that $\mathcal{H} = \mathbb{C}^2$, one also often finds quantum probability spaces simply defined through Gleason's theorem. That is, $\mathbb{P}$ is considered to be a quantum probability function if and only if there exists a density operator ρ such that $\mathbb{P}$ is of the form of (4.27). This redefinition atomically takes place when one switches to the algebraic formulation (see section 4.3). I will implicitly follow this convention. There are several justifications for this. First, the probability functions excluded in this way do not play a role in quantum mechanics. As such, particular properties of these functions have no direct implications for the foundations of quantum mechanics. And second, when turning to POVMs, Gleason's theorem also holds in the two-dimensional case (Busch, 2003).

Like the set of all probability measures on a measurable space, the set of all density operators on a Hilbert space is convex. That is, for any pair of density operators ρ_1, ρ_2 and $\lambda \in [0, 1]$ the operator $\rho = \lambda \rho_1 + (1 - \lambda)\rho_2$ is again a density operator. The extreme points of this convex set are precisely the one-dimensional projections. This explains the terminology that one-dimensional projections are called pure states and the density operators mixed states. This can be compared to the classical case where the extreme points coincide with Dirac measures on $(\Omega, \mathcal{F})$. These extreme points can be identified with the points of the space Ω. Often then, these are interpreted as denoting the actual state of affairs. For example, Ω may be the set of possible worlds, one of which corresponds to the actual world. Any nonextreme point in the set of probability distributions can then be seen as expressing a form of ignorance concerning what the actual world is.

A similar reading is sometimes found with respect to the density operators in quantum mechanics. The density operator then expresses an uncertainty concerning the state of the system. In fact, this is, more or less, the view von Neumann (1932) adopted when introducing density operators. But such a view is not entirely unproblematic. One technical issue is that, unlike in the classical case, there are many distinct convex combinations of pure states that give rise to the same mixed state. Thus, while in the classical case the mixed state dictates the set of pure states of which it is a mixture, the same cannot be done in quantum mechanics. Consequently, there are distinct situations in which uncertainty about the pure state is represented by the same mixed state. This in itself is not an unsurmountable problem, but it is still peculiar, as is the following conceptual issue.

Adopting an uncertainty interpretation of density operators suggests that there

---

[7]A natural metric on $L_1(\mathcal{H})$ for defining continuity of frame functions is the one induced by the operator norm.

is an actual state (the 'true' state of the system) to be uncertain about. But since pure states themselves are (generators of) probability functions, these probabilities are then implicitly attributed an ontological status as well. This is not typically quantum, since attributing probabilities to probability functions also happens for example in Bayesian statistics, where similar conundrums arise (see for example Uffink (2011, p.33–34)). But it does show that taking a stance on the nature of the distinction between pure and mixed states has interpretative consequences for what quantum probabilities are. Therefore, a reserved stance towards the distinction is adopted in this dissertation.

Thus far I have introduced the Hilbert space formulation of quantum probability and drawn some analogies between classical and quantum probability spaces. But much more can be done to compare the two. The next step is the reformulation of quantum probability theory in terms of *-algebras in the next two sections. The upshot of this reformulation is that both quantum and classical probability spaces can be viewed as instances of these (algebraic) quantum probability spaces.

4.3 Algebraic formulation

There is a peculiar tension between BoP and the definition of a quantum probability space. The Born rule takes as arguments observables and sets of possible measurement outcomes, while a quantum probability function in turn takes as arguments projection operators. There is a good reason for this discrepancy: observables give rise to self-adjoint operators (OP) which in turn give rise to projection operators through their associated PVMs. However, this association goes only in one direction and, as discussed in section 3.2, the converse association (embodied by IP) need not be universally valid. In that same section it was demonstrated that one can constructively make the set of self-adjoint operators that correspond to observables smaller by introducing superselection rules. Such smaller sets were characterized as corresponding to sub-algebras of operators on the Hilbert space. This abstraction effectively allows one to 'forget' about the Hilbert space, and take a purely algebraic approach instead. To prepare for the abstractness of this step it is useful to first consider a classical analogy.

In classical probability theory, probability measures can be generalized to expectation value functions. Every probability measure $\mathbb{P} : \mathcal{F} \to [0,1]$ gives rise to such a functional by the definition

$$\mathbb{E} : L^1(\Omega, \mathbb{P}) \to \mathbb{R}, \ \mathbb{E}(X) := \int_\Omega X(\omega) \mathrm{d} \, \mathbb{P}(\omega), \tag{4.28}$$

with[8]

$$L^1(\Omega, \mathbb{P}) := \left\{ X : \Omega \to \mathbb{R} \,\middle|\, \int_\Omega |X(\omega)| \mathrm{d}\,\mathbb{P}(\omega) < \infty \right\}. \tag{4.29}$$

Then the probability of a set is recovered by taking the expectation value of the characteristic function of that set, i.e.,

$$\mathbb{P}(\Delta) = \mathbb{E}(1_\Delta), \ 1_\Delta(\omega) := \begin{cases} 1 & \omega \in \Delta, \\ 0 & \omega \notin \Delta. \end{cases} \tag{4.30}$$

Thus instead of probability being defined as a function on events, it is now a functional defined on random variables.

One point on which the analogy breaks down, is that in the quantum case, instead of considering the set of self-adjoint operators (representing random variables), one considers an entire algebra of operators of which the relevant self-adjoint operators are a subset. Explicitly, the construction of an algebraic quantum probability space from an 'ordinary' quantum probability space goes as follows. Let $(\mathcal{H}, L(\mathcal{H}), \mathbb{P})$ be a quantum probability space. Let $\mathcal{B}(\mathcal{H})$ denote the algebra of bounded operators on $\mathcal{H}$. Now assume, using Gleason's theorem, that $\mathbb{P}$ is characterized by the density operator ρ. Then $\mathbb{P}$ gives rise to a functional ϕ_ρ on $\mathcal{B}(\mathcal{H})$ via the definition

$$\phi_\rho : \mathcal{B}(\mathcal{H}) \to \mathbb{C}, \ \phi_\rho(A) := \mathrm{Tr}(\rho A). \tag{4.31}$$

The restriction to bounded operators ensures that the trace is well-defined. This may be compared to the classical restriction to integrable functions in (4.28).

When restricting ϕ_ρ to the self-adjoint operators in $\mathcal{B}(\mathcal{H})$, one indeed obtains the expectation value with respect to the probability measure on the spectrum of A:

$$\phi_\rho(A) = \int_{\sigma(A)} a \, \mathrm{d}\,\mathbb{P}_\rho^A(a) = \mathbb{E}_\rho^A(\mathrm{id}_{\sigma(A)}), \tag{4.32}$$

with $\mathrm{id}_{\sigma(A)} : \sigma(A) \to \sigma(A)$ the classical random variable given by $\mathrm{id}_{\sigma(A)}(a) = a$. Note further that ϕ_ρ is an extension of $\mathbb{P}$ since $L(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$. For projection operators, (4.32) is the quantum version of (4.30). The functional ϕ_ρ is a state in the algebraic sense (see Definition A.21) and $\mathcal{B}(\mathcal{H})$ is a *-algebra. These features then are taken as construing the definition of an algebraic quantum probability space (see (Meyer, 1993)).

Definition 4.4. An *(algebraic) quantum probability space* is a pair $(\mathcal{C}, \phi)$ with $\mathcal{C}$ a *-algebra and $\phi : \mathcal{C} \to \mathbb{C}$ a state.

---

[8]Formally, $L^1(\Omega, \mathbb{P})$ is the set of equivalence classes of functions, where two functions X_1, X_2 are considered to be equivalent if and only if $\int_\Omega |X_1(\omega) - X_2(\omega)| \mathrm{d}\,\mathbb{P}(\omega) = 0$.

This definition is much more general than what is needed in this dissertation. But also in mathematical practice one often imposes further restrictions. For example, one may require that $\mathcal{C}$ is a C*-algebra or a von Neumann algebra. Another noteworthy feature is that a state need only be finitely additive. The prototypical state ϕ_ρ from (4.31) on the other hand is countably additive. When restricting to von Neumann algebras, this feature is captured by the notion of normality (see Definition A.22). The following theorem ties these idea together with the existence of appropriate density operators.[9]

Theorem 4.4. *Let $\mathcal{H}$ be a Hilbert space and $\phi : \mathcal{C} \to \mathbb{C}$ a state on a von Neumann algebra $\mathcal{C} \subset \mathcal{B}(\mathcal{H})$, then the following conditions are equivalent:*

(a) *ϕ is normal;*

(b) *there exists a density operator ρ on $\mathcal{H}$ such that $\phi(A) = \mathrm{Tr}(\rho A)$ for all $A \in \mathcal{C}$;*

(c) *for every set $\{P_i\}_{i \in I}$ of pairwise orthogonal projections*

$$\sum_{i \in I} \phi(P_i) = \phi \left(\sum_{i \in I} P_i \right). \tag{4.33}$$

In section 3.2 it was shown that von Neumann algebras arise in a natural way when superselection rules are introduced. These rules can be interpreted as a restriction on the set of observables (by a rejection of IP). For such a smaller set of observables, an algebraic quantum probability space may then be found to be more fitting. After all, it allows one to ignore the self-adjoint operators on the Hilbert space that do not correspond to observables. But more interesting for the present discussion, is that the algebraic formulation can be used to view classical probability spaces as a special case of quantum probability spaces. Specifically, every classical probability space can be represented using an algebraic quantum probability space. The construction is not entirely trivial though, and to make the general structure clear it is useful to first consider the special case of finite classical probability spaces as an example.

Example 4.2. Let $(\Omega, \mathcal{F})$ be a measurable space with $\#\Omega < \infty$ and $\mathcal{F} = \mathcal{P}(\Omega)$. With the set Ω one can associate the set of all complex-valued functions on Ω

$$C(\Omega) := \{f : \Omega \to \mathbb{C}\}. \tag{4.34}$$

---

[9]A proof for this theorem can be obtained by taking together Theorem 3.6.4 in (Pedersen, 1979) and Theorem 7.1.12 in (Kadison and Ringrose, 1986). Also, it deserves to be noted that for the validity of the theorem $\mathcal{H}$ need not be separable.

This is turned into a *-algebra by introducing the rules

$$
\begin{aligned}
(\lambda_1 f_1 + \lambda_2 f_2)(\omega) &:= \lambda_1 f_1(\omega) + \lambda_2 f_2(\omega), \\
(f_1 f_2)(\omega) &:= f_1(\omega) f_2(\omega), \\
f^*(\omega) &:= \overline{f(\omega)},
\end{aligned}
\tag{4.35}
$$

where $f, f_1, f_2 \in C(\Omega)$, $\lambda_1, \lambda_2 \in \mathbb{C}$ and $\lambda \mapsto \overline{\lambda}$ denotes complex conjugation on $\mathbb{C}$. A norm is introduced with the definition

$$
\|f\|_\infty := \max_{\omega \in \Omega} |f(\omega)|.
\tag{4.36}
$$

This turns $C(\Omega)$ into a C*-algebra. The self-adjoint elements of $C(\Omega)$ are the real-valued functions and the projections are precisely the characteristic functions 1_Δ for $\Delta \subset \Omega$.

Every probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ gives rise to a state $\phi_\mathbb{P}$ on $C(\Omega)$ defined as

$$
\phi_\mathbb{P} : C(\Omega) \to \mathbb{C}, \ \ \phi_\mathbb{P}(f) := \sum_{\omega \in \Omega} f(\omega)\, \mathbb{P}(\{\omega\}).
\tag{4.37}
$$

The classical probabilities are regained by the action of the state on the projections:

$$
\phi_\mathbb{P}(1_\Delta) = \mathbb{P}(\Delta).
\tag{4.38}
$$

Conversely, every state ϕ on $C(\Omega)$ can be used to define a probability measure given by

$$
\mathbb{P}_\phi : \mathcal{F} \to [0,1], \ \ \mathbb{P}_\phi(\Delta) := \phi(1_\Delta).
\tag{4.39}
$$

$$\triangle$$

The example establishes a bijective correspondence between probability measures on a finite measurable space and states on the associated C*-algebra. What deserves further special attention is that $C(\Omega)$ is an Abelian algebra. This is an aspect which is typical for the connection with classical probability space. For example, if $(\mathcal{A}, \phi)$ is a quantum probability space with $\mathcal{A}$ a finite-dimensional Abelian *-algebra, then one may show that it is isomorphic to a space of the form $(C(\Omega), \phi_\mathbb{P})$ derived from a classical probability space. In the case that Ω is infinite things become slightly more complicated, but the main idea remains the same. To highlight these main ideas, I only provide a sketch of the proof for the following theorem.[10]

Theorem 4.5. *For every classical probability space $(\Omega, \mathcal{F}, \mathbb{P})$ there exists a quantum probability space $(\mathcal{A}, \phi)$ with $\mathcal{A}$ an Abelian von Neumann algebra that represents the*

---

[10]Proofs of the relevant details can be found in (Douglas, 1998) (in particular Propositions 4.50 and 4.51) and section 1 in (Maassen, 2003).

classical probability space. Specifically, there is a surjective PVM $\mu : \mathcal{F} \to L(\mathcal{A})$ from the measurable sets to the set of projections in $\mathcal{A}$, such that

$$\phi(\mu(\Delta)) = \mathbb{P}(\Delta) \ \forall \Delta \in \mathcal{F}. \tag{4.40}$$

Proof sketch. Like in the finite case an algebraic structure is introduced by taking a particular set of complex-valued functions on Ω. Let $\mathcal{M}(\Omega)$ denote the set of all measurable complex-valued functions. Scalar multiplication, addition, multiplication and a *-operation are then introduced by the rules in (4.35). These rules turn $\mathcal{M}(\Omega)$ into a *-algebra. Note that every subset of $\mathcal{M}(\Omega)$ that is closed under the operations (4.35) is automatically also a *-algebra. The aim is to have a subset that is both big enough and is mathematically well-behaved.

To say that the sub-algebra of $\mathcal{M}(\Omega)$ is big enough, is to say that it can at least model the measurable sets in $\mathcal{F}$. So what is needed, at least, is that the projections are included. These again correspond to the elements 1_Δ with Δ any measurable subset of Ω. The candidate for a PVM $\mu : \mathcal{F} \to L(\mathcal{A})$ is then $\mu(\Delta) = 1_\Delta$. This sets the inspiration drawn from the finite case, and this is where one bumps into some technicalities.

The technical details come in the form of several desiderata. First, the algebra of all complex-valued functions does not have a norm to turn it into a C*-algebra. Furthermore, for the algebra to be a von Neumann algebra, it has to be embedded into the algebra of bounded operators on some Hilbert space $\mathcal{H}$, and be shown to be equal to its double commutant. Finally, one has to find a state on the algebra with the desired properties. Let's start with this problem. A first constraint for a state is that

$$\phi_\mathbb{P}(1_\Delta) = \mathbb{P}(\Delta) = \int_\Omega 1_\Delta(\omega) \mathrm{d}\, \mathbb{P}(\omega), \ \forall \Delta \in \mathcal{F}. \tag{4.41}$$

A natural extension then is that for other functions

$$\phi_\mathbb{P}(f) = \int_\Omega f(\omega) \mathrm{d}\, \mathbb{P}(\omega). \tag{4.42}$$

This requires restricting attention to the $\mathbb{P}$-measurable functions for which the integral converges, i.e., the elements of $\mathcal{L}^1(\Omega, \mathbb{P})$.

However, $\mathcal{L}^1(\Omega, \mathbb{P})$ is not an algebra as it is not closed under the operations from (4.35). For example, one may have that $f \in \mathcal{L}^1(\Omega, \mathbb{P})$ while $f^2 \notin \mathcal{L}^1(\Omega, \mathbb{P})$. As one can take arbitrary many powers of functions, one should switch to the set $\mathcal{L}^\infty(\Omega, \mathbb{P})$ which is the set of complex-valued functions for which the essential supremum of the absolute value is finite, i.e.,

$$\mathcal{L}^\infty(\Omega, \mathbb{P}) := \{f \in \mathcal{M}(\Omega) \,|\, \mathrm{ess\,sup}\,|f| < \infty\}\,,$$
$$\mathrm{ess\,sup}\,|f| := \inf\{x \in \mathbb{R} \,|\, \mathbb{P}(\{\omega \in \Omega \,|\, |f(\omega)| > x\}) = 0\}\,. \tag{4.43}$$

The essential supremum is the proper generalization of the supremum in a measure-theoretic context. Roughly, it selects the supremum of a function while neglecting the values the function may have on some set of measure zero.

Thus far $\mathcal{L}^{\infty}(\Omega, \mathbb{P})$ is an Abelian *-algebra, and $\phi_{\mathbb{P}}(f)$ is now defined for every $f \in \mathcal{L}^{\infty}(\Omega, \mathbb{P})$ by (4.42). However, a norm is still missing. The natural candidate is

$$\|f\|_{\infty} := \operatorname{ess\,sup} |f|. \tag{4.44}$$

However, this is not a norm since $\|f_1 - f_2\|_{\infty} = 0$ does not imply $f_1 = f_2$ because the two functions may differ on a set of $\mathbb{P}$-measure zero. This is solved by instead looking at equivalence classes of functions, and one switches to the set $L^{\infty}(\Omega, \mathbb{P})$. This is indeed a C*-algebra with the norm given by (4.44) and $\phi_{\mathbb{P}}$ as given by (4.42) defines a state on it. It can be viewed as a subalgebra of the bounded operators on the Hilbert space $L^2(\Omega, \mathbb{P})$. Indeed, every $f \in L^{\infty}(\Omega, \mathbb{P})$ gives rise to a bounded operator $M_f : L^2(\Omega, \mathbb{P}) \to L^2(\Omega, \mathbb{P})$ via

$$(M_f \psi)(\omega) := f(\omega)\psi(\omega). \tag{4.45}$$

One may show that in this context $L^{\infty}(\Omega, \mathbb{P})$ is indeed a von Neumann algebra. $\boxtimes$

Theorem 4.5 establishes a firm argument for the view that classical probability is a special case of quantum probability. There are, however, two small caveats that deserve to be mentioned. A peculiar difference between the construction in Example 4.2 and that in Theorem 4.5, is that in the first case, the algebra $C(\Omega)$ depends only on the measurable space $(\Omega, \mathcal{F})$, while in the second case the algebra $L^{\infty}(\Omega, \mathbb{P})$ also depends on the measure. This suggests that there may be difficulties with modeling the dynamics of classical probability functions in the quantum framework: a change in the measure $\mathbb{P}$ may require one to not only change the state $\phi_{\mathbb{P}}$, but also its domain. The second caveat is that measurable subsets of Ω are not uniquely represented in the quantum model: when the difference between two sets is of measure zero, they are identified with the same projection in $L^{\infty}(\Omega, \mathbb{P})$. Thus measure zero sets have no place in the quantum formalism.

The situation isn't very grim though. The usual dynamics of probability, given by conditionalization, is easily modeled in the quantum case and doesn't require updating the algebra. Specifically, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathbb{P}_{\Delta}$ denote the measure obtained after conditionalizing $\mathbb{P}$ on the set $\Delta \in \mathcal{F}$. Then there is a state $\phi_{\mathbb{P}_{\Delta}}$ on $L^{\infty}(\Omega, \mathbb{P})$ that represents $\mathbb{P}_{\Delta}$, namely,

$$\phi_{\mathbb{P}_{\Delta}}(f) := \frac{1}{\mathbb{P}(\Delta)} \int 1_{\Delta}(\omega) f(\omega) \mathrm{d}\, \mathbb{P}(\omega). \tag{4.46}$$

More generally, any probability measure $\mathbb{P}'$ can be modeled on $L^{\infty}(\Omega, \mathbb{P})$ as long

as $\mathbb{P}' \ll \mathbb{P}$.[11] In that case, it follows from the Radon-Nikodym theorem, together with the observation that $\mathcal{L}^\infty(\Omega, \mathbb{P}) \subset \mathcal{L}^\infty(\Omega, \mathbb{P}')$, that

$$\phi_{\mathbb{P}'}(f) := \int f(\omega) \frac{\mathrm{d}\,\mathbb{P}'(\omega)}{\mathrm{d}\,\mathbb{P}(\omega)} \mathrm{d}\,\mathbb{P}(\omega) \tag{4.47}$$

is well-defined and equals $\int f(\omega)\mathrm{d}\,\mathbb{P}'(\omega)$.

In cases where $\mathbb{P}'$ assigns a positive probability to a set Δ of $\mathbb{P}$-measure zero the situation is more difficult. In fact, here the first caveat and the second come together. The problem is that $L^\infty(\Omega, \mathbb{P})$ does not recognize Δ as a viable possibility. The significance of this fact is perhaps best understood within the setting of a concrete example. Consider the Hilbert space $L^2(\mathbb{R})$ used to describe a single particle in one spatial dimension. The event of finding the particle in one particular region Δ is represented by the projection 1_Δ viewed as a multiplication operator using (4.45).

In the case that Δ has measure zero, the corresponding operator 1_Δ is identified with the zero operator. This is in particular the case for all singleton sets $\{x\}$ with $x \in \mathbb{R}$. The foundational significance of this peculiarity is debatable. Arntzenius (2003) takes this aspect of quantum probability as support for the idea that space itself has no points. But I think it is worth recognizing two possibly underlying assumptions to this claim. The first is that it is impossible to recover the points of a space within the quantum formalism. The second is the idea that this formalism is in some sense necessarily the correct one for describing these matters. Both these assumptions may be questioned. Doing so does of course not imply that one can't make a case for going 'pointless', but it does indicate that the support from quantum probability for such a metaphysical claim is limited.

A counter example to the first assumption can be found in the paper by Halvorson (2001). He defends the idea that there are (non-normal) states in quantum probability which can be thought of as representing a definite value for the position of a particle. To this end the following definition is introduced. A state ϕ on the algebra $\mathcal{B}(L^2(\mathbb{R}))$ is said to converge to the point $x \in \mathbb{R}$ if $\phi(M_{1_\Delta}) = 1$ for every open set Δ containing x. Such states exist for every $x \in \mathbb{R}$ and may be interpreted as expressing that the particle is indeed in the point x. A counter example to the second assumption is provided by Bohmian mechanics. In this theory, all particles have a definite position at all times. The fact that probabilities for position measurements cannot distinguish regions that differ a set of measure zero then expresses a form of empirical indistinguishability. The situation then isn't that different from the one in classical probability. Thus constructions like that of Halvorson (ibid.) or in Bohmian mechanics help smoothing the discrepancies surrounding Theorem 4.5.

---

[11]A measure μ is said to be absolutely continuous with respect to another measure ν (denoted $\mu \ll \nu$) if for every measurable set Δ, if $\nu(\Delta) = 0$, then $\mu(\Delta) = 0$. In other words, the set of sets with ν-measure zero is a subset of the set of sets with μ-measure zero.

The emerging view is that quantum probability is a proper generalization of classical probability. However, this embedding of classical probability is not entirely trivial. One may then wonder if, as indicated in the discussion around Theorem 4.1, one can't conversely embed quantum probability within the classical framework with a bit of creativity. This question, then, is the topic of the next part of this dissertation.

Part II

Hidden variables and classical representations

5

Introduction

5.1 Defining the question of classical representations

The formalism of quantum probability was introduced in chapter 4, and it was shown there that classical probability can be viewed as a special case of this formalism (Theorem 4.5). To my knowledge, this result is the strongest defense for the idea that quantum probability is a proper generalization of classical probability.[1] However, Theorem 4.5 does not show that the generalization is in fact proper. That is, while the theorem shows that it is possible to represent every classical probability space as a quantum probability space, it is silent about whether the converse isn't true as well: can a quantum probability space be represented as a classical probability space? This, roughly, is the topic of this part of the dissertation.

The question is by no means trivial, as it goes hand in hand with evaluating what counts as a positive answer. Apart from being able to represent quantum probability functions in a classical way, it is also desirable that the representation is in some sense faithful. That is, if a classical representation is to be helpful for the philosophy of quantum probability, it also has to be explanatory. A solid definition of faithfulness will not be given, nor do I think that aspiring for such a definition is very useful. Rather, I take it to be more fruitful to reevaluate the question of what counts as a satisfactory representation along the way, while discussing formal results concerning the (im)possibility of these representations. But to get a feeling of what is hinted at when I use the notion of faithfulness, a metaphorical example may be useful.

---

[1]Another, and perhaps more popular argument, is that classical probability is unable to produce certain correlations typical for quantum probability, as demonstrated by the violation of Bell inequalities (see Theorem 4.1). However, derivations of such inequalities require more ingredients than just 'classical probability', which are typically of a physical/philosophical nature (e.g. locality). These assumptions are used to argue that random variables like A_1^{11} and A_1^{12} in (4.15) should actually be identified with each other. As such, from a mathematical perspective, they do not evidently show quantum probability is more general than classical probability.

It has long been believed that trajectories of celestial bodies should in some sense be harmonious and perfect. Circular motion is the most natural candidate for such a view. However, even a shift from a geocentric to a heliocentric model of the planetary motions in our solar system turned out to have discrepancies with observations. There always remained the need to introduce epicycles to the circular motions. That is, until Kepler abandoned the idea of using circles, and replaced them with ellipses.

Now one can think of circular motion being analogous to classical probability, and elliptical motion as being analogous to quantum probability. It is clear that elliptical motions form a generalizations of circular motions. However, it is also possible to account for elliptical motions by only making use of circles. By adding enough epicycles, circular motions can approach elliptical motions up to any desired degree of precision. So it is possible to represent ellipses by circles. But the representation would not be considered faithful by anyone. It is a result from Fourier analysis that *any* periodic orbit can be approximated with the use of epicycles. An illustrative internet movie by Ginnobili (2008) establishes with the use of epicycles an orbit that draws Homer Simpson (🦶). This example shows that, although circles can be used to describe ellipses, this description fosters no insight in the importance of ellipses. In fact, ellipses are rather arbitrary motions from the epicycle point of view.

To bring the metaphor back to the topic of this dissertation: if a classical representation of quantum probability is like the epicycles, the representation would not be considered to be faithful. But this only sketches the question of classical representations. The remainder of this chapter is devoted to further charting the contours of the possibilities and impossibilities of classical representations. In section 5.2 a comparison is made between the formalisms of classical and quantum probability that highlights the main distinctions that would have to be overcome for classical representations. In section 5.3 the connection is made with the question of hidden variable interpretations of quantum mechanics. The contours for a definition of classical representations are illustrated in section 5.4 by means of two theorems. First, a possibility theorem is taken to determine an outer boundary for the definition. After that, an impossibility theorem (based on von Neumann's no-go theorem for hidden variables) is taken to determine an inner boundary. Finally, in section 5.5, an outline is given for part II of this dissertation.

5.2   Formalisms of probability

The main topic of this part of the dissertation is the distinction between Kolmogorov's axiomatization of probability and the quantum formalism. The primary interest in an investigation of the differences between the two is relatively straight-

forward. The classical formulation of probability has been the topic of many philosophical studies. A particularly interesting part of these studies has been the justification of a specific axiomatization of probability, namely, that of Kolmogorov or a slight variation thereof (Eells, 1983). It is somewhat peculiar then to realize that scientific investigations lead to a formulation of probability that is manifestly distinct from this classical structure. Have philosophers then been trying to find justifications for the wrong formalism?

The situation is not that grim of course. The classical formalism may be adequate for most, if not all, of the use of probability outside quantum mechanics. This idea is only strengthened by the philosophical justifications for the formalism. But for the case where the quantum structure is applicable a new investigation seems required. That is, unless the distinction is not proper at all. If every quantum probability space can be represented on a classical probability space, then one may take the justification for the classical structure as indirectly providing a justification for the quantum formalism. However, it will probably come as no surprise to the reader that in this part of the dissertation I work towards the conclusion that satisfactory classical representations of quantum probability are not to be had. After all, after part II comes part III, and that part is still on quantum probability, and not on classical probability.

An appropriate way to start is with a coarse comparison of the two formalisms of probability. This then gives a first suggestion of what is required of a classical representation, as well as what is at stake if such a representation does not succeed. For the sake of definiteness, consider the two definitions of probability spaces alongside each other.

Definition 5.1.

A *classical probability space* is a triplet $(\Omega, \mathcal{F}, \mathbb{P})$ such that
1. Ω is a set,
2. $\mathcal{F}$ is a σ-algebra of subsets of Ω containing Ω,
3. $\mathbb{P}$ is a function from $\mathcal{F}$ to the interval $[0, 1]$,
4. $\mathbb{P}(\Omega) = 1$,
5. $\mathbb{P}(\bigcup_n \Delta_n) = \sum_n \mathbb{P}(\Delta_n)$ for any countable sequence of mutually disjoint sets.

A *quantum probability space* is a triplet $(\mathcal{H}, L(\mathcal{H}), \mathbb{P})$ such that
1. $\mathcal{H}$ is a Hilbert space,
2. $L(\mathcal{H})$ is the collection of all projection operators on $\mathcal{H}$,
3. $\mathbb{P}$ is a function from $L(\mathcal{H})$ to the interval $[0, 1]$,
4. $\mathbb{P}(\mathbb{1}) = 1$,
5. $\mathbb{P}(\sum_n P_n) = \sum_n \mathbb{P}(P_n)$ for any countable sequence of mutually orthogonal projections.

I choose to work here with the definition from Theorem 4.2 as it resembles the classical axiomatization as much as possible. It is also sufficiently general to stick to the Hilbert space formalism. One reason is that every C*-algebra can be represented

as a sub-algebra of $\mathcal{B}(\mathcal{H})$ for some Hilbert space.[2] But more importantly, it is because the theorems discussed in this part of the dissertation place all philosophical considerations within the framework of finite-dimensional Hilbert spaces.

When it comes to formalizing probability it is useful to separate three specific ingredients. Namely, probability is thought of as a function with (a) a domain and (b) a range and that (c) satisfies certain calculus rules. When comparing classical and quantum probability, the most important distinction is in the structure of the domain as indicated by points 1 and 2 in Definition 5.1. Aside from these differences concerning the domain, quantum probabilities are quite similar to their classical counterparts. The items 3, 4 and 5 for quantum probabilities only differ from the classical axioms to the extent that the domain in the quantum case has a different structure. In other words, given the domain $L(\mathcal{H})$, the range of the quantum probability function $\mathbb{P}$ and the rules it satisfies are as similar to the classical case as possible.

This observation is quite intriguing in light of the literature on the philosophy of probability. There is a (tacit) consensus that the domain specified by the axioms of Kolmogorov is roughly correct. The focus is instead on the ingredients (b) and (c). Dutch book arguments, for example, pertain to the rules for the calculus. Rationality requires an agent to set her degrees of belief to satisfy $\mathbb{P}(A) + \mathbb{P}(\neg A) = 1$, while taking for granted that A can be represented as a set. There is also some discussion on the range of probability functions. One may ask whether real numbers are either too fine-grained (can one have epistemic access to non-rational numbers?) or too coarse-grained (should one introduce hyperreals to be able to separate impossible from improbable events?). To be sure, there is also discussion on the domain of probability functions. Twined with the discussion of whether $\mathbb{P}$ should be σ-additive is the discussion whether $\mathcal{F}$ should be closed under countable unions. A possibly more popular issue is whether probability should be a single-placed function, such that conditional probabilities are derived from unconditional probabilities, or a two-placed function, taking conditional probability as primitive. But what remains the same in these considerations is the assumption that the arguments of the probability function are taken from some (Boolean) algebra of subsets.

The set-theoretic approach is relatively uncontroversial and applied widely (for example, it is the cornerstone of possible world semantics). Therefore it deserves to be emphasized that it is precisely this kind of structure that is absent in the formalism of quantum probability. Thus it is worthwhile to investigate under which conditions this structure can be recovered. But the question is also interesting from the viewpoint of quantum mechanics, as it relates to the question of hidden

---

[2]This is done using the construction of Gelfand, Naimark and Segal (Kadison and Ringrose, 1986, Theorem 4.5.2).

variables. I turn to this relation in the next section.

5.3 Hidden variables

From the viewpoint of philosophy of quantum mechanics, the question of classical representations of quantum probability touches on the subject of hidden variable interpretations of quantum mechanics. The question of hidden variables is as old as quantum mechanics itself (see also chapter 2). It ties in with the issue of whether the quantum state provides a complete description of a system, or only a partial description. The idea that the description may be only partial is an intuitive one. Quantum mechanics postulates the existence of many observables, gives lists of possible measurement outcomes for these observables, but refuses to commit itself to the idea that the actual measurement outcomes can be attributed to these observables independently of the act of measurement. This reluctance points to an idea often found in the more popular accounts of this theory, namely, that acts of measurement *create* the outcome of a measurement instead of *revealing* it. A hidden variable account of quantum mechanics may seek to alleviate this confusing state of affairs in the most obvious way. That is, a hidden variable state can be thought of as a function that assigns values to observables, which are to be the values revealed upon measurement. Specifically, the following definition is an appropriate working definition to start with:

Definition 5.2. A *hidden variable state* for a physical system S associated with a Hilbert space $\mathcal{H}$ is a function $\lambda : \mathcal{O}bs \to \mathbb{R}$ such that $\lambda(\mathcal{A}) \in \sigma(A)$ for every $\mathcal{A} \in \mathcal{O}bs$, where A is the self-adjoint operator acting on $\mathcal{H}$ associated with $\mathcal{A}$.

Here $\mathcal{O}bs$ denotes the set of all observables and A is the self-adjoint operator quantum mechanics associates with the observable $\mathcal{A}$.

A connection between the existence of hidden variable states and classical representations of quantum probability spaces can now be made in the following way. In a classical probability space, every $\omega \in \Omega$ determines the value of every random variable. That is, if $\mathcal{RV}$ denotes the set of (real-valued) random variables on Ω, then every $\omega \in \Omega$ determines a functional

$$\hat{\omega} : \mathcal{RV} \to \mathbb{R}, \ \hat{\omega}(F) := F(\omega). \tag{5.1}$$

Now it is a reasonable demand that, for a classical representation $(\Omega, \mathcal{F})$ of a quantum probability space $(\mathcal{H}, L(\mathcal{H}))$, it is possible to associate a classical random variable on Ω with every quantum random variable. In particular, for every self-adjoint operator A there should be a random variable $F_A : \Omega \to \sigma(A)$. Consequently, every $\omega \in \Omega$ gives rise to a hidden variable state λ_ω by the rule

$$\lambda_\omega(\mathcal{A}) := \hat{\omega}(F_A), \tag{5.2}$$

where A is the self-adjoint operator associated with $\mathcal{A}$.

Thus every classical representation implies the existence of hidden variable states. It is then tempting to conclude that every no-go theorem for hidden variables provides a no-go theorem for classical representations. But this is where the questions of classical representations and hidden variables come apart. No-go theorems for hidden variable theories rely on assumptions that may not be directly relevant for the question of classical representations. This is because hidden variables come equipped with metaphysical connotations. As such, additional constraints (such as locality) that may be relevant for the question of hidden variables, may be less interesting from the formal perspective of classical representations. It is for this reason that this part of the dissertation is almost solely devoted to the Kochen-Specker theorem, while discussions of Bell inequalities are omitted entirely. What is left open is the question of what constraints *are* interesting for a classical representation. The next section is devoted to answering this question.

5.4   What is a classical representation?

Let $(\mathcal{H}, L(\mathcal{H}), \mathbb{P})$ be a quantum probability space and suppose $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}}, \mathbb{P}_{\mathcal{H}})$ is a classical probability space that represents it. What could one reasonably expect this to mean? A definite answer to this question will not be given here. In fact, the main message to take away from this section is that the answer depends on how faithful one wants the representation to be. Which aspects of the quantum formalism should be reproduced in the classical representation? This is a philosophical question and, as such, it is susceptible to multiple defensible answers.

Every possible answer has to make up a balance with on the one side the explanatory merit a classical representation can give, and on the other side the elegance and constraints posited by the quantum formulation. To clarify this idea, this section is devoted to explaining two formal results on the (im)possibility of classical representations that pull in opposite directions. The first result is a positive one, and assumes a very liberal reading of classical representations. The second is a negative one. It assumes a very restrictive reading, and requires the classical representation to adopt so much of the quantum structure, that one can show that such a representation does not exist. The gain of considering these two results, is that the first may be seen as giving a minimal requirement for a classical representation, while the second shows that it is a non-trivial matter to investigate what more requirements can be added. As such, they set the boundaries for the discussion in the remainder of this part.

A natural minimal requirement for a classical representation is that it is able to reproduce the empirical content of quantum probability. Tracking this empirical content is of course not a trivial matter. In fact, precisely assessing what empirical

ingredients underly the Kochen-Specker theorem is one of the points on which my presentation of the Kochen-Specker theorem differs from common presentations.

Whatever empirical content probability theory has, it is channeled through random variables. After all, these represent the possible experiments as well as their corresponding outcomes. Thus a minimal requirement is that for every quantum random variable there is a classical random variable that represents it. I restrict attention to the quantum random variables that are given by self-adjoint operators. The following theorem shows that it is possible to identify each one of them with a random variable on a single classical probability space.

Theorem 5.1. *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{O}_{\mathrm{sa}}$ the set of self-adjoint operators acting on it. Then there is a measurable space $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}})$ together with a collection of random variables $\{F_A \mid A \in \mathcal{O}_{\mathrm{sa}}\}$ such that for every quantum probability function $\mathbb{P} : L(\mathcal{H}) \to [0,1]$ there is a probability measure $\mathbb{P}_{\mathcal{H}}$ on $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}})$ that satisfies*

$$\mathbb{P}(\mu_A(\Delta)) = \mathbb{P}_{\mathcal{H}}(F_A^{-1}(\Delta)) \tag{5.3}$$

for all measurable subsets $\Delta \subset \sigma(A)$, where μ_A denotes the PVM associated with A.

Proof. For every self-adjoint operator the associated measurable space $(\Omega_A, \mathcal{F}_A)$ is given by the spectrum of A and the algebra of Borel subsets. The idea of the proof is to heap all the measurable spaces $(\Omega_A, \mathcal{F}_A)$ into a single big one. The probability measure $\mathbb{P}_{\mathcal{H}}$ is then constructed as the product measure of probability measures on the spaces $(\Omega_A, \mathcal{F}_A)$.

The set $\Omega_{\mathcal{H}}$ is constructed as the Cartesian product of all the Ω_A:

$$\Omega_{\mathcal{H}} := \{\omega : \mathcal{O}_{\mathrm{sa}} \to \mathbb{R} \mid \omega(A) \in \sigma(A)\} \simeq \prod_{A \in \mathcal{O}_{\mathrm{sa}}(\mathcal{H})} \Omega_A. \tag{5.4}$$

The σ-algebra $\mathcal{F}_{\mathcal{H}}$ will be the one generated by the cylinder sets. That is, $\mathcal{F}_{\mathcal{H}}$ is defined as the smallest σ-algebra containing the subsets

$$C_{(A_1, \Delta_1), \dots, (A_n, \Delta_n)} := \{\omega \in \Omega_{\mathcal{H}} \mid \omega(A_i) \in \Delta_i, \ i = 1, \dots, n\} \tag{5.5}$$

for any finite sequence $(A_1, \Delta_1), \dots, (A_n, \Delta_n)$ with $A_i \in \mathcal{O}_{\mathrm{sa}}$ and $\Delta_i \in \mathcal{F}_{A_i}$ for all i.

For each self-adjoint operator A the random variable F_A is defined as

$$F_A(\omega) := \omega(A). \tag{5.6}$$

One may check that this is indeed a measurable function. Now let $\mathbb{P}$ be a quantum probability function. The corresponding probability measure $\mathbb{P}_{\mathcal{H}}$ is now defined by first specifying its action on the cylinder sets:

$$\mathbb{P}_{\mathcal{H}}\left(C_{(A_1, \Delta_1), \dots, (A_n, \Delta_n)}\right) := \prod_{i=1}^{n} \mathbb{P}(\mu_{A_i}(\Delta_i)). \tag{5.7}$$

This already warrants that (5.3) holds. The only puzzle left is to show that this function can be extended to a measure on the whole of $\mathcal{F}_{\mathcal{H}}$. For this result, I refer to the short proof by Saeki (1996), which moreover shows that the extension is unique. □

On a liberal reading of what classical representations are to be, the above theorem establishes the possibility of such representations. Every quantum random variable is identified with a classical random variable on the measurable space, and for every quantum probability function there is a probability measure that reproduces the distributions for all quantum random variables. But there are two ways in which this representation may be found to be unsatisfactory or unfaithful. The first problem is similar to the one with the epicycle representation of elliptical orbits discussed in section 5.1. The probability measures on $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}})$ that derive from quantum probability functions are a rather special subset of all the probability measures. Indeed, these measures may be seen to satisfy a form of non-contextuality. Namely, for any pair of operators A_1, A_2 and sets Δ_1, Δ_2 the conditional

$$\mu_{A_1}(\Delta_1) = \mu_{A_2}(\Delta_2) \Rightarrow \mathbb{P}_{\mathcal{H}}(F_{A_1} \in \Delta_1) = \mathbb{P}_{\mathcal{H}}(F_{A_2} \in \Delta_2) \tag{5.8}$$

holds, and it is easy to see that it need not hold for arbitrary probability measures on $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}})$. Without at least an additional explanation of why (5.8) should hold, the classical representation of Theorem 5.1 is of little explanatory value when it comes to quantum probability.

The second problem is that the classical representation does not respect the algebraic structure of quantum probability. Then, whenever this structure is relevant for the application of quantum probability theory, the classical representation will not be adequate. Let me first explain by example what is formally the problem. In the quantum formalism, for any two self-adjoint operators A_1 and A_2, their sum is again a self adjoint operator. And whenever the two operators commute, their product is also again a self-adjoint operator. But this structure is not respected by the associated random variables $F_{A_1}, F_{A_2}, F_{A_1+A_2}$ and $F_{A_1 A_2}$. In fact, by the construction of the product measure (5.7), these random variables are all stochastically independent. Now, if the algebraic structure is respected in actual measurements, then the classical representation is no longer adequate. For example, if the outcomes a_1, a_2, a_3 of a joint measurement of $A_1, A_2, A_3 = A_1 + A_2$ always satisfy the relation $a_3 = a_1 + a_2$, then the variables F_{A_1}, F_{A_2} and $F_{A_1+A_2}$ should not be stochastically independent.

Quantum probability itself suggests that the algebraic structure actually is relevant. For every quantum probability function and for all pairs of self-adjoint operators A_1, A_2 the expectation values satisfy the relation

$$\mathbb{E}(A_1 + A_2) = \mathbb{E}(A_1) + \mathbb{E}(A_2). \tag{5.9}$$

Von Neumann (1932) took this to suggest that for a classical representation this relation should hold for all probability measures. In particular, it should hold for the Dirac measures $\mathbb{P}_\omega$ for all ω. This boils down to the assumption that

$$F_{A_1 + A_2} = F_{A_1} + F_{A_2} \tag{5.10}$$

for all $A_1, A_2 \in \mathcal{O}_{\mathrm{sa}}$. However, this assumption is one that cannot hold, as is shown by the following simplified version of von Neumann's argument against hidden variables (ibid., §IV.2).

Theorem 5.2. *Let $(\mathcal{H}, L(\mathcal{H}), \mathbb{P})$ be a quantum probability space. Then there is no classical representation $(\Omega, \mathcal{F}, \mathbb{P})$ such that for every self-adjoint operator A there is a random variable $F_A : \Omega \to \sigma(A)$ and (5.10) holds for all pairs of self-adjoint operators A_1 and A_2.*

Let me start on a positive note by pointing out that the theorem establishes that not all of the algebraic structure of quantum probability can be recovered in a classical representation. However, it is not clear why all this structure is something that should be preserved. As a definite argument against the possibility of hidden variables the above result is quite poor, and it has been criticized for having been presented as such by von Neumann. In particular, the constraint (5.10) has become known as von Neumann's "silly assumption" (Bell, 1988; Mermin, 1993).

The argument against the constraint runs as follows. The range of the function $F_{A_1} + F_{A_2}$ is given by the set $\{a_1 + a_2 \mid a_1 \in \sigma(A_1), a_2 \in \sigma(A_2)\}$. But in general this does not coincide with the spectrum of $A_1 + A_2$. In fact, these two sets can even be disjoint. Then, by assuming (5.10), values could be attributed to the observable corresponding to $A_1 + A_2$ that are not even among the possible measurement outcomes, and these possible outcomes in turn are not available as possible values for the observables. Incidentally, this gives a proof of Theorem 5.2 as well as a proof of the fact that the empirical motivation for the constraint given just above (5.9) is flawed.

It is clear from this analysis that von Neumann's theorem proved something what was deemed irrelevant by Bell and his followers. It is not immediately clear from this though, that his proof is thereby silly. This depends more on what von Neumann aimed to prove, rather than on what others expected from it. A compelling defense for the thesis that the theorem is indeed significant is given by Bub (2011a). For the present discussion this issue is slightly less relevant and would also require going into the details of the original theorem instead of focusing on the simplified version given here. What is relevant to note, is that it follows that a notion of classical representation that adopts (5.10) as a constraint is not only one that cannot be realized, but would also deliver more than needed. The question of the definition and possibility of faithful representations thus remains open at this

point. But the contours have been sketched, and it is now clear that the possibility of classical representations ties in with the precise relevance of the algebraic structure of quantum probability spaces for the application of quantum probability theory. This highlights the main topic of study in the remainder of this part of the dissertation.

5.5   Outline

As noted in section 5.3, the remainder of this part of the dissertation solely focuses on the Kochen-Specker theorem. The theorem is first presented, proven, and investigated in chapter 6. In more detail, sections 6.1 and 6.2 work towards a formulation of the theorem that highlights and disentangles the philosophical and technical assumptions that underlie it. In the first of these two sections a more or less traditional formulation is presented, while in the second a more modern formulation is given that has some of my own ingredients.

Although the theorem presents severe constraints on the possibility of classical representation, it also relies on two assumptions that are good candidates for rejection: non-contextuality (NC) and the identification principle (IP). Sections 6.3 and 6.4 are dedicated to the possibilities of classical representations and hidden variables when NC is rejected. The rejection of IP is more involved and controversial, and requires a chapter on its own. In chapter 7 the possibility of non-contextual classical representations and hidden variables is discussed by means of a discussion of the models of Meyer, Kent and Clifton (MKC). First, in section 7.2 a proof of the Kochen-Specker theorem is presented that is close to the original proof, and that allows a significant weakening of IP. Section 7.3 is dedicated to investigating the possibility and justification of rejecting IP. Specifically, the so-called finite precision loophole which allows the rejection of IP is explained. In section 7.4 the possibility for reproducing quantum probabilities is discussed, and in section 7.5 some peculiar aspects of the MKC models are highlighted. A short review of the results is then given in section 7.6.

6
The Kochen-Specker theorem

6.1 The Kochen-Specker theorem

In their own words, Kochen and Specker state that their theorem aims "to give a proof of the nonexistense of hidden variables", and they continue by noting that "This requires that we give at least a precise necessary condition for their existence." The accepted existence of hidden variable theories proves that the necessary condition posed by Kochen and Specker can be contested. On the other hand, there is some consensus that the condition is not obviously too restrictive. The theorem hasn't met the same kind of criticism as von Neumann's theorem did. In short, the theorem provides valuable insight in the kind of constraints that play a role in the possibility of hidden variables and classical representations.

The import of the Kochen-Specker theorem is often phrased with the credo that it excludes the possibility of so-called non-contextual hidden variables. Some have even gone as far as to conclude from the result that nature itself is contextual. These are claims that deserve to be made more precise. But I will not yet go deeply into the impact of the theorem in this section. That is reserved for the later sections and chapter 7. Instead, this section is devoted to picking out problematic aspects of the classical representation introduced by Theorem 5.1, and to use the Kochen-Specker theorem as a way to demonstrate that there is no easy way to patch up these problems.

Now let's pick up on the issue I discussed in section 5.4. Which algebraic structures of quantum probability should be respected by a classical representation in order for the representation to be called faithful? A definite answer is not given here, but a minimal requirement is that a structure should be respected when it is reflected in empirical predictions. As noted in section 5.4, such a reflection is possible within the setting of joint measurements. Whenever algebraic relations are reflected in outcomes of joint measurements, this introduces statistical correlations among the observables. A probability measure in a classical representation should of course be able to account for such correlations.

   The formalism of quantum mechanics as presented in chapter 3 mostly dealt with single measurements at a time. However, it is quite generally accepted that (at least as an idealization) certain joint measurements are also possible. More precisely, the following postulate is often (implicitly) accepted.

CoP (Comeasurability Postulate) If $\{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$ is a finite set of observables and $\{A_1, \ldots, A_n\}$ is a corresponding set of self-adjoint operators that satisfies $[A_i, A_j] = 0$ for all i, j, then it is possible to perform a joint measurement of $\mathcal{A}_1, \ldots, \mathcal{A}_n$.

This postulate in itself poses no direct empirical content, but in conjunction with the Born Postulate (BoP) it does. Quantum mechanics also adopts joint probabilities for observables whose corresponding operators commute. For these joint probabilities, these observables need not be stochastically independent. This poses possible difficulties for the classical probability space of Theorem 5.1, for which all variables are independent by construction.

   One may envision that in light of these possible correlations, one can still be able to represent quantum probability functions on the space $(\Omega_\mathcal{H}, \mathcal{F}_\mathcal{H})$ constructed in the proof of Theorem 5.1. After all, the constructed probability measures from (5.7) need not be the only probability functions that reproduce the quantum statistics for single measurements. There are others that do allow for correlations. However, the Kochen-Specker theorem poses difficulties for this view, not by showing that the set of probability measures on $(\Omega_\mathcal{H}, \mathcal{F}_\mathcal{H})$ is too small to account for all quantum probability functions (when joint measurements are also considered), but by showing that the set $\Omega_\mathcal{H}$ *itself* is unable to account for the predictions of quantum mechanics. As a bonus, one does not have to go into the precise empirical aspects of stochastic dependence to investigate the import of the theorem. What is needed now is a further postulate that poses relations between measurement outcomes in the case of joint measurements. The following one is commonly accepted.

EFR (Experimental Functional Relations) For every finite set of observables $\{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$ for which a joint measurement is possible, and for every real-valued Borel function f that satisfies $f(A_1, \ldots, A_n) = \mathbb{1}$ (where $A_1, \ldots, A_n$ are the corresponding self-adjoint operators), the outcomes $a_1, \ldots, a_n$ of a joint measurement satisfy $f(a_1, \ldots, a_n) = 1$.

Example 6.1. Consider a single observable $\mathcal{A}$ corresponding to the operator A and consider the characteristic function $1_{\sigma(A)}$. Then $1_{\sigma(A)}(A) = \mathbb{1}$, and so EFR implies the value postulate (VaP), i.e., the outcome of the measurement of $\mathcal{A}$ should lie in the spectrum of A. $\qquad\qquad\triangle$

Example 6.2. Consider again the spin operators for a spin-1 particle from Example 3.2. If one takes the squares of these operators, then one can show that for any triplet of orthogonal axes r_1, r_2, r_3

$$[S_{r_i}^2, S_{r_j}^2] = 0. \tag{6.1}$$

Thus, according to CoP, it is possible to make a joint measurement of $S_{r_1}^2, S_{r_2}^2, S_{r_3}^2$. Furthermore, it follows from EFR that the outcome should be one of the triplets $(1, 1, 0)$, $(1, 0, 1)$ or $(0, 1, 1)$ because

$$\sigma\left(S_{r_i}^2\right) = \{0, 1\} \text{ and } \frac{1}{2}S_{r_1}^2 + \frac{1}{2}S_{r_2}^2 + \frac{1}{2}S_{r_3}^2 = \mathbb{1}. \tag{6.2}$$

$\triangle$

Example 6.3. For any self-adjoint operator A and real-valued Borel function f the operator $f(A)$ is well-defined (see equation (A.49)) and is again self-adjoint. Furthermore, $[A, f(A)] = 0$. If $\mathcal{A}_1$ and $\mathcal{A}_2$ are two observables associated with the operators A and $f(A)$ respectively, then it follows from CoP that they are jointly measurable and from EFR that the measurement outcomes for a joint measurement satisfy

$$a_2 = f(a_1). \tag{6.3}$$

$\triangle$

The first of these examples shows that VaP is a consequence of EFR. In fact, like VaP, EFR may be seen to be a non-probabilistic version of a consequence of BoP for joint measurements. Indeed, this postulate implies that for measurement outcomes $a_1, \ldots, a_n$ of a joint measurement, the relation $f(a_1, \ldots, a_n) = 1$ will hold with probability one for all quantum states.

The classical representation of Theorem 5.1 is not in accordance with the two assumptions CoP and EFR. This can be demonstrating by making use of the last example. The relation

$$F_{f(A)}(\omega) = f\left(F_A(\omega)\right) \tag{6.4}$$

does not hold for all $\omega \in \Omega_{\mathcal{H}}$ whenever f is not the identity. Therefore, a joint measurement of observables associated with the operators A and $f(A)$ cannot be taken to reveal the values of F_A and $F_{f(A)}$. It is worthwhile to delve a bit more into the tension at hand. For single measurements, the statistics of the classical representation *do* obey the functional relationships such as those in Example 6.3. Specifically, if $A_2 = f(A_1)$ then for every measurable set Δ

$$\mathbb{P}_{\mathcal{H}}\left(F_{A_1}^{-1}(\Delta)\right) = \mathbb{P}_{\mathcal{H}}\left(F_{A_2}^{-1}(f(\Delta))\right). \tag{6.5}$$

However, the joint probability distribution does not adhere to the functional relation since

$$\mathbb{P}_{\mathcal{H}}\left(F_{A_1}^{-1}(\Delta) \cap \left(F_{A_2}^{-1}(f(\Delta))\right)^c\right) = \mathbb{P}_{\mathcal{H}}\left(F_{A_1}^{-1}(\Delta)\right) \mathbb{P}_{\mathcal{H}}\left(\left(F_{A_2}^{-1}(f(\Delta))\right)^c\right), \quad (6.6)$$

while EFR on the other hand requires that

$$\mathbb{P}_{\mathcal{H}}\left(F_{A_1}^{-1}(\Delta) \cap \left(F_{A_2}^{-1}(f(\Delta))\right)^c\right) = 0. \quad (6.7)$$

It is tempting to try to accommodate for EFR in the classical representation by adjusting the probability measure $\mathbb{P}_{\mathcal{H}}$ such that stochastic variables corresponding to commuting observables are no longer independent. The idea would be that for every $A \in \mathcal{O}_{\mathrm{sa}}$ and function f the set

$$\left\{\omega \in \Omega_{\mathcal{H}} \mid F_{f(A)}(\omega) \neq f\left(F_A(\omega)\right)\right\} \quad (6.8)$$

would get probability zero. However, it will be a consequence of the Kochen-Specker theorem that this is not possible.

Possibly the most intuitive way to incorporate the assumptions CoP and EFR as constraints for a classical representation, is to demand that relations between commuting operators are reflected in relations between their corresponding random variables. This is precisely the key assumption adopted by Kochen and Specker. Using the terminology of Redhead (1987), the assumption is

FUNC (Functional Composition Principle) For every self-adjoint operator A and real-valued Borel function f, the classical random variables F_A and $F_{f(A)}$ (associated with the operators A and $f(A)$) should satisfy the relation

$$F_{f(A)} = f \circ F_A. \quad (6.9)$$

Note that, within the context of hidden variables, FUNC is indeed a sufficient (though not necessary) condition to ensure that EFR holds. This follows because for any finite set of commuting self-adjoint operators $A_1, \ldots, A_n$ there is an operator A_0 together with functions $f_1, \ldots, f_n$ such that $A_i = f_i(A_0)$ for each i. Then for any function f that satisfies $f(A_1, \ldots, A_n) = \mathbb{1}$ one can introduce the function $h_f(x) := f(f_1(x), \ldots, f_n(x))$. It follows that $h_f(A_0) = \mathbb{1}$ and from FUNC it then follows that

$$f(a_1, \ldots, a_n) = f(f_1(a_0), \ldots, f_n(a_0)) = h(a_0) = 1. \quad (6.10)$$

It also deserves to be noted that a particular consequence of FUNC is that

$$F_{A_1+A_2} = F_{A_1} + F_{A_2} \quad (6.11)$$

whenever A_1 and A_2 commute. The equation need not hold when the operators do not commute, and in this sense it weakens the constraint (5.10) used for Theorem 5.2.

The relation between FUNC and the above attempt to fix the classical representation of Theorem 5.1 by tweaking the probability functions, is that the admissible probability functions would assign probability zero to the set of ω that don't satisfy FUNC (6.8). This implies assigning probability 1 to all the ω that do satisfy FUNC. The Kochen-Specker theorem now shows that this cannot be done because this set is in fact empty.

Theorem 6.1 (Kochen-Specker 1967)**.** *For any Hilbert space $\mathcal{H}$ with dimension greater than 2 there is no non-empty set Ω such that for every self-adjoint operator A on $\mathcal{H}$ there is a function $F_A : \Omega \to \sigma(A)$ and such that FUNC is satisfied.*

Proof. Suppose Ω is not empty. Every $\omega \in \Omega$ determines a function $\lambda_\omega : \mathcal{O}_{\mathrm{sa}} \to \mathbb{R}$ through

$$\lambda_\omega(A) := F_A(\omega). \tag{6.12}$$

This defines a frame function by restricting to the set of 1-dimensional projections. I prove this here for finite-dimensional $\mathcal{H}$ only.

Let $(P_k)_{k=1}^n$ be a frame and define $A := \sum_k k P_k$. Then for every $j \in \{1, \ldots, n\}$ the function

$$f_j(x) := \prod_{k \neq j} \frac{x - k}{j - k} \tag{6.13}$$

satisfies $f_j(A) = \sum_k f_j(k) P_k = P_j$. Using FUNC it follows that

$$\sum_k \lambda_\omega(P_k) = \sum_k F_{P_k}(\omega) = \sum_k F_{f_k(A)}(\omega) = \sum_k f_k\left(F_A(\omega)\right). \tag{6.14}$$

Because $f_k(j) = \delta_{jk}$ for $j \in \{1, \ldots, n\}$ and $F_A(\omega)$ is one of these integers, it follows that

$$\sum_k \lambda_\omega(P_k) = 1. \tag{6.15}$$

Furthermore, since $\lambda_\omega(P) \in \{0, 1\}$ for every $P \in L_1(\mathcal{H})$ it is a positive frame function. Because both values 0 and 1 must be obtained and none of the values in between, λ_ω must be discontinuous. The existence of such a frame function, however, is excluded by Gleason's theorem whenever $\dim(\mathcal{H}) > 2$, because $P \mapsto \mathrm{Tr}(\rho P)$ is continuous for every density operator ρ.

The proof for infinite-dimensional $\mathcal{H}$ goes along similar lines, though it requires some more care in choosing A and the f_k in such a way that they are well defined. $\quad\square$

It is worth noting that it is only with the addition of CoP and EFR that the classical representation of Theorem 5.1 is found to be inadequate. Although these assumptions are taken to be a part of quantum mechanics, it is not entirely evident that they should hold in all possible applications of quantum probability. This indicates a possible separation between quantum probability as a formal framework and its use in quantum mechanics. Admittedly, this is quite a technical point, and CoP and EFR seem plausible candidates for any application of quantum probability. Furthermore, the proof of Theorem 6.1 establishes that EFR may also be weakened to hold only for polynomial functions.[1] The point made here then isn't that these assumptions are somehow of deep significance. Rather, the point is simply that it does matter whether such assumptions are taken to hold or not and recognizing such assumptions is a non-trivial enterprise.

More interesting assumptions will be uncovered in the next section. Some of them relate to something I passed over with a bit of hand-waving. Namely, FUNC was shown here to be a sufficient condition to incorporate EFR, but not a necessary one. Then, there may be some wiggle room between EFR and FUNC. The first is an empirical constraint, while the latter is a formal one, and even though FUNC has the appearance of innocence, it turns out to have tremendous consequences. This gives me the opportunity to let Bell (1966, p. 451) make the bridge to the next section: "That so much follows from such apparently innocent assumptions leads us to question their innocence."

6.2   Reformulating the Kochen-Specker theorem

Although the Kochen-Specker theorem purports to make a contribution to the philosophy of quantum mechanics, its precise importance is not immediately clear. This is a necessary consequence for any formal proof in philosophy; at some point philosophical intuitions have to be translated into mathematically well-defined assumptions, and derived consequences then have to be translated back. These translations are always somewhat muddy. However, in the present situation there is room for improvement. In the previous section it was not always clear where the philosophy stopped and the mathematics began. In this section I will make the separation more precise.

Since the mathematics is the easiest part I start there. Specifically, I start with a purely mathematical reformulation of the theorem. The mathematical assumptions are then scrutinized and motivated/derived from more conceptual/philosophical assumptions. This then allows the reformulation of the theorem in a more conceptual

---

[1]Actually, in the finite-dimensional case every function can be shown to behave like a polynomial. That is, for every function f and every self-adjoint operator A, there is a polynomial $p_{f,A}$ such that $f(A) = p_{f,A}(A)$.

language. In further sections then, the conceptual assumptions are subjected to a more intensive investigation. Without further ado, here is the Kochen-Specker theorem (again).

Theorem 6.2. *Let $\mathcal{H}$ be a Hilbert space of dimension at least 3, and let $\mathcal{O}_{sa}$ be the set of self-adjoint operators on $\mathcal{H}$. Then there is no function $\lambda : \mathcal{O}_{sa} \to \mathbb{R}$ such that*

$$\lambda(A) \in \sigma(A), \tag{6.16a}$$
$$\lambda(f(A)) = f(\lambda(A)) \tag{6.16b}$$

for every $A \in \mathcal{O}_{sa}$ and real-valued Borel function f.

This presentation of the theorem can be understood as a corollary of Gleason's theorem, which implies the non-existence of 0-1 valued quantum probability functions. On the other hand, any λ satisfying the constraints (6.16) gives rise to a 0-1 valued probability function by restriction to the projection operators. The important distinction is conceptual in nature.[2] Gleason proved his theorem as part of a purely mathematical investigation (i.e., the classification of all probability functions). Kochen and Specker, on the other hand, intended their theorem to be philosophically relevant by relating it to the question of the possibility of hidden variables. Such relevance can only be had by establishing relations between the Hilbert space structure, the empirical content of quantum mechanics, and further philosophical desiderata. So let's draw some relations.

In Theorem 6.2 the function λ is intended to represent a hidden variable state. The interest in the existence of such states comes from a desire to find a description of quantum mechanics that is more complete than the standard formalism. The standard formalism is silent about the origin of measurement outcomes; it is not clear whether they can be thought of as revelations of quantities existing independent of the performance of a measurement. Hidden variables are supposed to fill in this gap.[3] Formally, postulating their existence can be seen to be embodied by the conjunction of the following two assumptions.

VD (Value Definiteness) Every observable possesses a unique definite value at all times.

FM (Faithful Measurement) Given VD, a measurement of an observable reveals the value it possesses at the time of the measurement.

---

[2]To be fair, the original Kochen-Specker theorem is not a corollary of Gleason's theorem as they do not assume that $\lambda(A)$ is defined for all $A \in \mathcal{O}_{sa}$. So apart from the conceptual distinction between the two theorems there is also a mathematical one, which is discussed in chapter 7. However, to clarify the conceptual distinction the formulation in Theorem 6.2 suffices.

[3]There may be other reasons for wishing to look for hidden variable formulations of quantum mechanics, such as a desire for determinism. Here I intend to make their motivation more modest.

One may note a discrepancy between these assumptions and the role of the function λ in Theorem 6.2. VD and FM refer to a notion of hidden variable states as given by Definition 5.2. Such a hidden variable state acts on the set of observables. The function λ on the other hand is defined on the set of self-adjoint operators. The connection between the two is made by the observable postulate.

OP (Observable Postulate) Every observable $\mathcal{A}$ for the system S is associated with a self-adjoint operator A acting on $\mathcal{H}$.

While it is clear that every function λ satisfying (6.16) induces a hidden variable state by assigning to each observable $\mathcal{A}$ the value $\lambda(A)$ (where A is the operator associated with $\mathcal{A}$ by OP), it does not follow from this same postulate that every hidden variable state should be of this form. It *would* follow if one could take the set of observables to be *identical* to the set of self-adjoint operators. Two further assumptions can be recognized that together establish this.

NC (Non-Contexuality) Every self-adjoint operator is associated with at most one observable.

IP (Identification Principle) For every self-adjoint operator A there exists an observable $\mathcal{A}$ such that A is associated with $\mathcal{A}$ via OP.

To illustrate how these assumptions all tie together, note that OP implies the existence of a function $f : \mathcal{Obs} \to \mathcal{O}_{\mathrm{sa}}$ that assigns to each observable a self-adjoint operator. NC then states that f is injective, and IP states that it is surjective. Consequently, any hidden variable state $\kappa : \mathcal{Obs} \to \mathbb{R}$ can be translated to a function $\lambda : \mathcal{O}_{\mathrm{sa}} \to \mathbb{R}$ without loss of generality.

Thus far I have only discussed the domain of hidden variable states, but not their range or possible constraints they should satisfy. The first of the two constraints (6.16a) (together with the earlier assumptions) ensures that λ is a hidden variable state as in Definition 5.2. But the motivation for this constraint comes of course from quantum mechanics itself. In fact, it can be seen to follow from FM and EFR by making use of Example 6.1.

The second constraint (6.16b) is the FUNC rule from the previous section. The empirical motivation for this constraint derives from EFR. But EFR cannot do any work without a further assumption on the comeasurability of observables. CoP suffices to make this final step, but a slightly weaker version also works. Although weakening assumptions is often a good idea to obtain stronger results, the conceptual motivation for this particular weakening is lacking at this point. The explanation will be given in the next section, where it turns out to be the case that this slight alteration is quite crucial.

WCoP (Weak Comeasurability Postulate) For every finite set of observables $\{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$ with corresponding self-adjoint operators $A_1, \ldots, A_n$, if $[A_i, A_j] = 0$ for all i, j, then there exists a set of observables $\{\mathcal{A}'_1, \ldots, \mathcal{A}'_n\}$, corresponding to the same operators, such that it is possible to perform a joint measurement of $\mathcal{A}'_1, \ldots, \mathcal{A}'_n$.

The FUNC rule can now be seen to follow from WCoP and EFR together with the other assumptions. This establishes a solid empirical underpinning of one of the more technical ingredients of the Kochen-Specker theorem. To sum up, the theorem can be formulated as follows.

Theorem 6.3. *For any system described by a Hilbert space with dimension greater than 2 the assumptions OP, WCoP, EFR, VD, FM, IP and NC taken together lead to a contradiction.*

The implications of the theorem are straightforward: at least one of the assumptions has to be rejected.[4] The assumptions OP, WCoP, EFR are all commonly accepted facets of quantum mechanics and are to be kept in place to keep the discussion on track.[5] For further exposition, it is then useful to define

$$\mathrm{QM}_{\mathrm{KS}} := \mathrm{OP} \wedge \mathrm{WCoP} \wedge \mathrm{EFR} \tag{6.17}$$

as the part of quantum mechanics used in the Kochen-Specker theorem. VD and FM are reasonable assumptions for a hidden variable theory, and it is likewise useful to set

$$\mathrm{HV} := \mathrm{VD} \wedge \mathrm{FM}. \tag{6.18}$$

Also, these are requirements for the possibility of constructing a classical probability space for quantum mechanics (see section 5.2). Theorem 6.3 can then be formulated as the formula

$$\mathrm{QM}_{\mathrm{KS}} \wedge \mathrm{HV} \wedge \mathrm{IP} \wedge \mathrm{NC} \Rightarrow \perp. \tag{6.19}$$

In this formulation it is clear that the Kochen-Specker theorem relies on two assumptions that lack a good motivation: IP and NC. When adopting a purely formal stance with respect to the question of classical representations of quantum probability these assumptions have a certain natural appeal. Rejecting IP would imply that not every quantum random variable (self-adjoint operator) will be represented by at least one classical random variable. Not all facets of quantum probability would then be captured by the classical representation. Rejecting NC would imply that the quantum random variables do not fully take into account all relevant random variables, as some of them refer to multiple distinct classical random variables. If

---

[4]Unless one is willing to resort to paraconsistent logic.

[5]Remember that, in particular, EFR can be understood as a consequence of BoP.

this is the case, it raises the question of how the quantum formalism was found to be adequate in the first place. One would expect that in some way a distinction between random variables is able to represent itself. These considerations of course do not pose a definitive argument against the view that quantum probability cannot have a satisfactory classical representation. However, they do indicate that there are difficulties for a 'natural' classical representation.

For the question of hidden variables the costs of rejecting either IP or NC may appear to be less high. This would especially be the case when the rewards can be as big as solving the measurement problem. But there are benefits closer to home. Rejecting NC may be quite an attractive option. The incompleteness of the quantum formalism implied by denying it ($\mathcal{O}_{sa}$ does not adequately describe all observables) can even be seen as a bonus. Indeed, the incompleteness of quantum mechanics was the main motivation for going into the question of hidden variables in the first place. The Kochen-Specker theorem then only confirms the idea that something is missing.

Rejecting IP is more complicated. To be sure, there are good reasons to reject its validity. A nice example is given by Nielsen (1997) who constructs a self-adjoint operator that contradicts the Church-Turing thesis if it were to correspond to an observable. And as noted in chapter 3 there are many operators for which it is hard to imagine any experimental setup that would correspond to it. But it is difficult to find a satisfactory argument to ban a specific set of operators from the observables, and to show that by doing so the Kochen-Specker theorem can be circumvented. This is particularly difficult when realizing that the original proof of Kochen and Specker uses a much weaker version of IP, requiring only that a particular *finite* set of operators correspond to observables. In short, trying to circumvent their theorem by wiggling with IP is quite non-trivial and thus deserves to be discussed in a separate chapter.

Here I have only summed up some intuitions one may have considering the implications of the Kochen-Specker theorem. But real insight can only be gained by having a closer look at the assumptions IP and NC. The rejection of NC is discussed in the next section and the rejection of IP is the topic of chapter 7.

6.3 The role of non-contextuality

Traditionally, the Kochen-Specker theorem is interpreted as establishing the impossibility of non-contextual hidden variable theories. In other words, the rejection of NC is the commonly accepted response for saving hidden variables. Notably, this option was suggested by Bell (1966) even before the theorem had actually been proven. In a review of Gleason's theorem and its consequences for the hidden variables program he states that (p. 451) "[i]t was tacitly assumed that measurement

of an observable must yield the same value independently of what other measurements may be made simultaneously." This refers to the fact that although condition (4.23) in the definition of a frame function only presents a constraint for commuting projection operators, it also has consequences for incompatible measurements. This is because frame functions assign values to 1-dimensional projections *independently* of the orthonormal basis they are associated with.

As an example, consider two non-commuting self-adjoint operators A_1, A_2 and two possible measurement outcomes a_1, a_2 such that $\mu_{A_1}(\{a_1\}) = \mu_{A_2}(\{a_2\})$. Then, if A_1 gets assigned the value a_1 by a hidden variable state, then necessarily A_2 gets the value a_2. Thus the values attributed to one observable can have consequences for the possible values to be attributed to another observable even if the two cannot be measured together. As Bell remarks earlier on (p. 447), such assumptions[6] "are seen to be quite unreasonable when one remembers with Bohr "the impossibility of any sharp distinction between the behavior of atomic objects and the interaction with the measuring instruments which serve to define the conditions under which the phenomena appear.""

On a historical note it deserves to be said that the assumption called out by Bell was hardly a tacit one in Gleason's paper. In fact, as a mathematician he was quite explicit about his assumptions. Furthermore, the assumption is not unreasonable when put in the context of Gleason's investigation, namely to "Determine all measures on the closed subspaces of a Hilbert space" (Hooker, 1975, p. 123).

In terms of the assumptions formulated in the previous section, Bell's objection can be read as the rejection of NC: although one may associate the same self-adjoint operator to a particular measurement irrespective of what other measurements are performed simultaneously, the observable associated with the measurement may be distinct in these cases. Thus there may be many distinct observables associated with the same self-adjoint operator, and it is not clear why it is a reasonable assumption that these observables should be attributed the same value by a hidden variable state. A second way of reading Bell is as a rejection of VD. An observable may then be attributed a multitude of definite values, namely, one for every possible measuring context in which the observable can be measured.

To a certain extent the distinction between these two options is a matter of semantics. On the first view, a hidden variable state is a function $\kappa : \mathcal{O}bs \to \mathbb{R}$ where at least in some cases $\kappa(\mathcal{A}_1) \neq \kappa(\mathcal{A}_2)$ even though $\mathcal{A}_1$ and $\mathcal{A}_2$ are associated with the same operator. On the second reading, a hidden variable state is a (partial) function $\lambda : \mathcal{O}_{\mathrm{sa}}(\mathcal{H}) \times \mathfrak{C} \to \mathbb{R}$ where $\mathfrak{C}$ symbolizes the set of all possible measurement contexts, and $\lambda(A, C)$ denotes the value of A in the context C. The reason for assuming that λ is only a partial function, is that it only needs to assign a value to (A, C) if C denotes a context in which A can be measured. The two kinds of states

---

[6]Here Bell cites Bohr (1949).

can be translated into one another. Given a λ one can define a κ by taking

$$\mathit{Obs} := \{(A, C) \in \mathcal{O}_{\mathrm{sa}}(\mathcal{H}) \times \mathfrak{C} \mid A \text{ can be measured in context } C\}, \qquad (6.20)$$

and setting $\kappa = \lambda$. Conversely, for every κ one can define a λ by taking $\mathfrak{C} = \mathit{Obs}$ and setting $\lambda(A, \mathcal{A})$ equal to $\kappa(\mathcal{A})$ if A is the operator associated with $\mathcal{A}$, and leaving it undefined otherwise. Thus for convenience I consider both constructions to be versions of contextual hidden variables.

The clarity of Bell's criticism leads one to wonder how the assumption of non-contextuality crept into Kochen and Specker's argument. In fact, the assumption NC never explicitly enters into their paper: it is only used implicitly. It is not trivial to tease out how NC arises in the original Kochen-Specker paper. Naturally, it enters into any argument when no sharp distinction is made between observables and self-adjoint operators. Now, even though in their words Kochen and Specker make the distinction with care, no notation is developed to make the distinction sharp. NC then creeps into the argument because a single operator is (implicitly) associated with two possibly distinct observables. This, in turn, is a consequence of conflation of an assumption of comeasurability and a definition of comeasurability, which I will now explain.

The first assumption that enters the original proof is the following. For any observable $\mathcal{A}$ and function f one can *define* the observable $f(\mathcal{A})$ as the observable consisting of a measurement of $\mathcal{A}$ and then applying f to the outcome. The assumption is then that if A is the operator associated with $\mathcal{A}$, the observable $f(\mathcal{A})$ is associated with the operator $f(A)$. Next, two observables $\mathcal{A}_1$ and $\mathcal{A}_2$ are said to be *comeasurable* if there is an observable $\mathcal{A}_3$ and functions f_1, f_2 such that $A_1 = f_1(A_3)$ and $A_2 = f(A_3)$. Note that this is used by Kochen and Specker as the *definition* of comeasurability. The intended reading, however, is that comeasurability implies that $\mathcal{A}_1$ and $\mathcal{A}_2$ can be measured simultaneously. Furthermore, it is not entirely evident what the role of $\mathcal{A}_3$ is here. But it turns out that implicitly there is the assumption that a joint measurement of $\mathcal{A}_1$ and $\mathcal{A}_2$ is the same as a single measurement of $\mathcal{A}_3$. More specifically, they implicitly assume that in this case $\mathcal{A}_1 = f_1(\mathcal{A}_3)$, where now the left-hand side and the right-hand side refer to observables that have been introduced independently. Indeed, the observable $f_1(\mathcal{A}_3)$ has been defined earlier on as the measurement of $\mathcal{A}_3$ with f_1 applied to the outcome. It is however not evident that this is the same as a measurement of $\mathcal{A}_1$, even though both observables are associated with the same operator $A_1 = f_1(A_3)$.

The line of reasoning adopted by Kochen and Specker can be made more transparent using the terminology introduced in the previous sections. Their definition of comeasurability can be seen to allude to CoP. This is because the existence of the operator A_3 and functions f_1, f_2 is equivalent to the condition that $[A_1, A_2] = 0$. The equality $\mathcal{A}_1 = f_1(\mathcal{A}_3)$ can then be further motivated by using EFR within the context of hidden variables. In first instance $\mathcal{A}_1$ and $f_1(\mathcal{A}_3)$ may be two distinct

observables corresponding to the same operator. But it follows from CoP that they can be jointly measured, and EFR implies that in that case the measurement outcomes are identical. This in turn implies that the hidden variable state κ must assign the same value to both observables. More generally, $\kappa(\mathcal{A}_1) = \kappa(\mathcal{A}_2)$ whenever the two observables $\mathcal{A}_1, \mathcal{A}_2$ are identified with the same operator. In short, one has[7]

$$\text{HV} \wedge \text{CoP} \wedge \text{EFR} \Rightarrow \text{NC}. \tag{6.21}$$

Note that on the other hand one has

$$\text{WCoP} \wedge \text{NC} \Rightarrow \text{CoP} \tag{6.22}$$

and

$$\text{HV} \wedge \text{WCoP} \wedge \text{EFR} \nRightarrow \text{NC}. \tag{6.23}$$

In other words, CoP itself is almost a non-contextuality assumption. Its initial reasonableness presumably stems from the fact that NC holds in orthodox presentations of quantum mechanics. But when NC is contested, it is necessary to resort to something weaker like WCoP. A related point is that, in the context of HV, the assumption CoP makes EFR act like FUNC. Thus FUNC may be viewed as a non-contextual version of EFR.

Whether NC is adopted as a primitive assumption or derived from other assumptions, without it the Kochen-Specker theorem does not work. This, however, leaves open the question of whether rejecting it is sufficient to allow the construction of hidden variable theories. The question can be answered positively by way of giving an example: Bohmian mechanics is not only the prime example of a hidden variable theory, but also of a contextual hidden variable theory. However, this theory is formulated in a way that does not explicitly adopt a rejection of NC. For this reason it is not a useful example for gaining general insight in the nature of this assumption. Instead, I illustrate in the next section what a contextual hidden variable theory may look like, and I provide an example of a contextual classical representation of quantum probability.

6.4   A contextual classical representation

Although it is commonly believed that it is possible to make formal constructions of contextual hidden variables, explicit examples of such constructions are not so common. The earliest such construction is likely due to Gudder (1970). A possible reason for their unpopularity is that they are not very illuminating physically (Shimony, 1984). Here, however, I am interested in the formal comparison of quantum

---

[7]For convenience I assume here that, if two observables have the same value for all hidden variable states, then they are the same observable.

and classical probability. For such a program a formal construction of contextual hidden variables is useful.

Setting up a contextual theory requires the introduction of contexts. The quote from Bell suggests that the specification of a context may depend on the totality of observables that are to be measured or can be measured using the envisaged experimental setup. Thus specifying only the operator associated with the observable is insufficient for defining a context, but specifying the set of all operators associated with observables that can be measured using the experimental setup may well be enough. In fact, this is the notion of a context I adopt here, and I will show that it suffices for the construction of a contextual hidden variables theory.

It is natural to stipulate that the set of operators that specifies a context forms a von Neumann algebra. The rationale behind this is borrowed from Kochen and Specker. Suppose that $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are observables that can be measured within a certain context. Then there is also an observable $f(\mathcal{A}_1, \ldots, \mathcal{A}_n)$ associated with applying the function f to the outcomes of the joint measurement of $\mathcal{A}_1, \ldots, \mathcal{A}_n$. It is natural to associate this observable with the operator $f(A_1, \ldots, A_n)$.[8] When adding these operators to the set $\{A_1, \ldots, A_n\}$ for every real-valued Borel function f, one obtains a von Neumann algebra of operators (assuming the operators are bounded). To sum up, if $\mathcal{A}_1, \ldots, \mathcal{A}_n$ is a collection of observables associated with a context, then $\mathcal{A}lg(A_1, \ldots, A_n)$ is the algebra that will be used to specify the context.

It is natural to suppose that the algebra $\mathcal{A}lg(A_1, \ldots, A_n)$ is Abelian. However, this does not immediately follow from the above considerations, but in fact requires a further assumption.

CCoP (Converse Comeasurability Postulate) For every finite set of observables $\{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$, if it is possible to perform a joint measurement of $\mathcal{A}_1, \ldots, \mathcal{A}_n$, then their corresponding self-adjoint operators $A_1, \ldots, A_n$ commute.

With this assumption it follows that with every context one can associate an Abelian von Neumann algebra. This will be taken as the complete characterization of a context. This means that, for example, whenever two sets of observables $\{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$ and $\{\mathcal{A}'_1, \ldots \mathcal{A}'_m\}$ satisfy $\mathcal{A}lg(A_1, \ldots, A_n) = \mathcal{A}lg(A'_1, \ldots, A'_m)$, then they specify the same context. Now let $\mathfrak{A}$ denote the set of all Abelian von Neumann algebras, then, using the notation from (6.20), the set of observables is given by

$$Obs := \{(A, \mathcal{A}) \in \mathcal{O}_{\mathrm{sa}} \times \mathfrak{A} \mid A \in \mathcal{A}\}. \tag{6.24}$$

I will now show that this set of observables is rich enough to allow the formulation of a hidden variable theory, or classical representation for quantum probability. I

---

[8]Note that, unlike in the argument of Kochen and Specker, this association is not problematic, since now NC is rejected and there is no worry that the operator $f(A_1, \ldots, A_n)$ may have to be reserved for another observable.

restrict attention to the cases where $\mathcal{H}$ is finite-dimensional to avoid technicalities that may blur the discussion.

The assumptions QM$_{\text{KS}}$ and HV together require the existence of a function $\lambda : \mathit{Obs} \to \mathbb{R}$ such that

$$\lambda(A, \mathcal{A}) \in \sigma(A),$$
$$\lambda(f(A), \mathcal{A}) = f(\lambda(A, \mathcal{A})) \tag{6.25}$$

for all $(A, \mathcal{A}) \in \mathit{Obs}$ and real-valued Borel functions f. To demonstrate the existence of such functions I first give a sufficient criterion for their existence, and then show that this criterion can be met.

For every context $\mathcal{A} \in \mathfrak{A}$ let $\mathcal{O}_{\text{sa}}(\mathcal{A})$ denote the set of self-adjoint operators in it. Now, if for every $\mathcal{A}$ there exists a function $\lambda_{\mathcal{A}} : \mathcal{O}_{\text{sa}}(\mathcal{A}) \to \mathbb{R}$ that satisfies

$$\lambda_{\mathcal{A}}(A) \in \sigma(A),$$
$$\lambda_{\mathcal{A}}(f(A)) = f(\lambda_{\mathcal{A}}(A)), \tag{6.26}$$

for all A and f, then the function

$$\lambda : \mathit{Obs} \to \mathbb{R}, \ \ \lambda(A, \mathcal{A}) := \lambda_{\mathcal{A}}(A) \tag{6.27}$$

satisfies (6.25). So all that has to be shown now is that for every $\mathcal{A}$ there exist functions of the type $\lambda_{\mathcal{A}}$ that satisfy (6.26).

Let $\mathcal{A} \in \mathfrak{A}$ and let $L(\mathcal{A}) = \mathcal{A} \cap L(\mathcal{H})$ denote the set of projection operators in $\mathcal{A}$. This set is big enough to generate the algebra it derives from, i.e., $\mathcal{A} = \mathit{Alg}(L(\mathcal{A}))$. Many subsets of $L(\mathcal{A})$ share this property, and one of particular interest is the set $L_{\text{a}}(\mathcal{A})$ defined as[9]

$$L_{\text{a}}(\mathcal{A}) := \left\{ P \in L(\mathcal{A}) \,\middle|\, P \neq \mathbb{0}, \ PP' \in \{\mathbb{0}, P\} \forall P' \in L(\mathcal{A}) \right\}. \tag{6.28}$$

To see that this set indeed generates the algebra, it suffices to note that it recovers $L(\mathcal{A})$. This is the case since for every $P' \in L(\mathcal{A})$

$$P' = \sum_{P \in L_{\text{a}}(\mathcal{A})} PP'. \tag{6.29}$$

For every $P \in L_{\text{a}}(\mathcal{A})$ one can now define

$$\lambda_{(\mathcal{A}, P)} : \mathcal{O}_{\text{sa}}(\mathcal{A}) \to \mathbb{R}, \ \ \lambda_{(\mathcal{A}, P)}(A) := \frac{\text{Tr}(AP)}{\text{Tr}(P)}. \tag{6.30}$$

---

[9]Using the language of lattice theory, one may show that $L_{\text{a}}(\mathcal{A})$ corresponds to the set of atoms in $L(\mathcal{A})$ with the lattice structure given by the partial order $P_1 \leq P_2$ if and only if $P_1 P_2 = P_1$ and the meet and join are $P_1 \wedge P_2 = P_1 P_2$ and $P_1 \vee P_2 = P_1 + P_2 - P_1 P_2$ respectively (see also chapter 8). One may also note that at this point it is used that the Hilbert space is finite-dimensional. In the infinite-dimensional case one need not have that $\mathcal{A} = \mathit{Alg}(L_{\text{a}}(\mathcal{A}))$.

It is a consequence of the spectral theorem that $\lambda_{(\mathcal{A},P)}(A) \in \sigma(A)$. Specifically, $\lambda_{(\mathcal{A},P)}$ picks out the eigenvalue of A corresponding to the eigenstates in the space $P\mathcal{H}$. The second condition of (6.26) follows by making use of the Borel functional calculus (see (A.50)).

This construction establishes that QM_{KS} and HV can be satisfied. By construction of *Obs* it also satisfies IP. Thus it has been shown that rejecting NC is sufficient to circumvent the Kochen-Specker theorem. However, the existence of hidden variable states alone is not a sufficient condition for a classical representation of quantum probability. There also need to be enough states such that the quantum probability distributions can in fact be modeled. I now show that the above construction is rich enough to allow this.

It will be useful to introduce the following set that will be used to characterize the hidden variable states:

$$\Omega := \{\omega : \mathfrak{A} \to L(\mathcal{H}) \mid \omega(\mathcal{A}) \in L_{\text{a}}(\mathcal{A})\} \simeq \prod_{\mathcal{A} \in \mathfrak{A}} L_{\text{a}}(\mathcal{A}). \tag{6.31}$$

Each of the sets $L_{\text{a}}(\mathcal{A})$ in the product can be viewed as a state space for the context $\mathcal{A}$, which are then all glued together to form the set Ω. The set Ω determines the set of hidden variable states through the definition

$$\Lambda := \{\lambda_\omega \mid \omega \in \Omega\}, \quad \lambda_\omega(A, \mathcal{A}) := \lambda_{(\mathcal{A},\omega(\mathcal{A}))}(A) = \frac{\text{Tr}(A\omega(\mathcal{A}))}{\text{Tr}(\omega(\mathcal{A}))}. \tag{6.32}$$

One may even show that all functions that satisfy (6.25) are of this form.[10]

The construction of a σ-algebra and probability measures that simulate the quantum probabilities is similar to that in the proof of Theorem 5.1. It is convenient to define both within the language of the set Ω. Specifically, $\mathcal{F}$ is the σ-algebra generated by the cylinder sets

$$C_{(\mathcal{A}_1,\mathcal{D}_1),\dots,(\mathcal{A}_n,\mathcal{D}_n)} := \{\omega \in \Omega \mid \omega(\mathcal{A}_i) \in \mathcal{D}_i, \ i = 1,\dots,n\}, \tag{6.33}$$

where $(\mathcal{A}_1,\mathcal{D}_1),\dots,(\mathcal{A}_n,\mathcal{D}_n)$ can be any finite sequence of pairs $(\mathcal{A}_i,\mathcal{D}_i)$, where $\mathcal{A}_i$ is an Abelian von Neumann algebra and $\mathcal{D}_i \subset L_{\text{a}}(\mathcal{A}_i)$. For every context $\mathcal{A} \in \mathfrak{A}$ and for every self-adjoint operator $A \in \mathcal{A}$ the observable $(A, \mathcal{A})$ is identified with the measurable function

$$F_{(A,\mathcal{A})}(\omega) := \frac{\text{Tr}(A\omega(\mathcal{A}))}{\text{Tr}(\omega(\mathcal{A}))} = \lambda_\omega(A, \mathcal{A}). \tag{6.34}$$

---

[10]A proof can be given by using that $\mathfrak{A} = \{\mathcal{A}lg(A) \mid A \in \mathcal{O}_{\text{sa}}\}$, i.e., every Abelian von Neumann algebra can be generated by a single self-adjoint operator. For a given λ one then defines $\omega_\lambda(\mathcal{A}lg(A)) := \mu_A(\{\lambda(A, \mathcal{A}lg(A))\})$. One then only has to check that this definition is independent of the choice of the operator A that generates the algebra $\mathcal{A}lg(A)$.

This just takes the observable to the value assigned to it by the hidden variable state λ_ω.

Now let ρ be a density operator that specifies a quantum probability function. A corresponding probability measure $\mathbb{P}_\rho$ on $(\Omega, \mathcal{F})$ is defined by its action on the cylinder sets.

$$\mathbb{P}_\rho\left(C_{(\mathcal{A}_1, \mathcal{D}_1), \dots, (\mathcal{A}_n, \mathcal{D}_n)}\right) := \prod_{i=1}^{n} \sum_{P \in \mathcal{D}_i} \mathrm{Tr}(\rho P). \tag{6.35}$$

For every observable $(A, \mathcal{A})$ this probability measure reproduces the Born rule. To see this let $\Delta \subset \sigma(A)$ and associate with it the set $\mathcal{D}_\Delta \subset L_{\mathrm{a}}(\mathcal{A})$ given by

$$\mathcal{D}_\Delta := \{P \in L_{\mathrm{a}}(\mathcal{A}) \mid P \mu_A(\Delta) = P\}. \tag{6.36}$$

Then

$$\begin{aligned}
\mathbb{P}_\rho\left(F_{(A, \mathcal{A})} \in \Delta\right) = \mathbb{P}_\rho\left(C_{(\mathcal{A}, \mathcal{D}_\Delta)}\right) &= \sum_{P \in \mathcal{D}_\Delta} \mathrm{Tr}(\rho P) \\
&= \sum_{a \in \Delta} \mathrm{Tr}\left(\rho \mu_A(\{a\})\right) = \mathrm{Tr}\left(\rho \mu_A(\Delta)\right).
\end{aligned} \tag{6.37}$$

This construction reproduces the Born rule for every possible measurement context, and then sort of glues them together. This provides a nice way of thinking about quantum probability: a patchwork of classical probability spaces with every patch corresponding to a possible measurement context. Furthermore, within every context this model satisfies the functional relationships of EFR because the states satisfy (6.25). It thus respects the algebraic structure of quantum probability when it matters. In this sense it is a huge improvement on the classical representation constructed in Theorem 5.1. But be that as it may, there are also some aspects to the construction that can be used to argue that this representation doesn't deserve the label of being faithful.

Like the construction of Theorem 5.1, the measurable space $(\Omega, \mathcal{F})$ allows many more probability measures than those given by the Born rule. From the point of view of the contextual hidden variables, the probability functions that are quantum form a somewhat peculiar subset of all the probability measures. The functions $\mathbb{P}_\rho$ may be seen to satisfy a non-contextuality property. Specifically, consider a self-adjoint operator A together with to distinct contexts $\mathcal{A}_1, \mathcal{A}_2$ such that $A \in \mathcal{A}_1 \cap \mathcal{A}_2$. Then for every set $\Delta \subset \sigma(A)$ and every $\mathbb{P}_\rho$ it holds that

$$\mathbb{P}_\rho\left(F_{(\mathcal{A}_1, A)} \in \Delta\right) = \mathbb{P}_\rho\left(F_{(\mathcal{A}_2, A)} \in \Delta\right). \tag{6.38}$$

This is of course a direct consequence of the Born rule. According to quantum mechanical predictions, there is no way to discern the observables $(\mathcal{A}_1, A)$ and $(\mathcal{A}_2, A)$. The peculiarity is that the contextual hidden variable model on the other

hand explicitly requires that these observables *are* distinct, which causes a tension. While there have to be cases where

$$F_{(\mathcal{A}_1, A)}(\omega) \neq F_{(\mathcal{A}_2, A)}(\omega), \tag{6.39}$$

when it comes to the quantum probability functions, the models are not allowed to display this contextuality. Thus, although adopting a contextual point of view allows one to construct a classical representation, this point of view itself seems to be lacking support with the special role played by the quantum mechanical probability functions. Then there is little hope that the key assumption alone (rejecting NC) can provide an explanation of the structure of quantum probability.

The specific sense in which quantum probability functions are non-contextual is a topic I return to in chapter 11. This turns out to be an aspect that is lacking an explanation even outside the framework of (contextual) hidden variables. Within the framework, there are of course attempts to explain the Born rule. Bohmian mechanics in particular, as the main contender for a hidden variable interpretation of quantum mechanics, has to be able to account for the special role of quantum mechanical probability functions. The common solution, developed first by Valentini (1991), is the adoption of a dynamical approach. The quantum probability functions are taken to be equilibrium distributions that arise due to the evolution of the hidden variable states governed by the Schrödinger and guiding equations. The approach is reminiscent of Boltzmann's H-theorem explanation of the special role of the Maxwell-Boltzmann distribution in statistical mechanics (Brown, Myrvold, and Uffink, 2009). Unsurprisingly, it shares similar difficulties (Callender, 2007; Timpson, 2011).

Going deeper into the understanding of probability in Bohmian mechanics takes us too far adrift, because this theory does not easily relate to the formalism of quantum probability as it is studied here. For example, although Bohmian mechanics is considered a contextual hidden variable theory, it does not share the set of observables as defined in (6.24). To the extent that these observables play a role in the theory, they are translated to position measurements in a non-trivial way. Instead, it is time to move forward and to investigate in what way classical representations can be constructed when relaxing the other ingredient of the Kochen-Specker theorem: IP.

The finite-precision loophole

7.1 Introduction

From a logical perspective, Theorem 6.3 (see also (6.19)) provides a very tight result: the assumptions NC and IP appear as natural constraints considering the formalism of quantum probability. But from an empiricist perspective, these assumptions are less satisfactory. In section 6.3, NC was criticized for setting constraints that are empirically counterfactual. After all, the defining property of incompatible measurement contexts is that properties pertaining to these contexts cannot be investigated simultaneously. It is therefore unreasonable to assume that observables that occur in more than one context should be attributed the same value independent of this context.

Finding a critique of IP that can be used to circumvent the theorem is a lot harder. On the other hand, it is plausible that the assumption is unsatisfactory. For one, the assumption postulates an uncountable infinity of observables. Now such an amount of observables is not uncommon in scientific theories, but usually it is obtained in a rather trivial way. In classical mechanics, for example, length can be measured in meters, or in inches, or in any other of the uncountably many rescalings of the standard unit of length. As soon as one accepts that any observable can be rescaled by any real number, one obtains uncountably many observables. But the uncountable set postulated by IP is not of this form: there is an uncountable subset of $\mathcal{O}_{sa}$ such that neither pair of observables in it is related by a scaling or other transformation. For some of the elements it is even completely unclear what a corresponding measurement procedure could be (see also the discussion in section 3.1).

This situation is unsettling. On the one hand the assumption IP just seems absurd, as it is inconceivable that it would be possible even in principle to construct a measurement procedure for every self-adjoint operator. But on the other hand, denying IP constructively is also problematic. It requires pointing out specific self-

adjoint operators which are to be denied the status of observable. The situation would improve much if either the assumption IP can be relaxed to a more compelling one, or if one could find a constructive way to deny IP and provide a classical representation that exploits this. This representation would of course be required to be non-trivial, in the sense that the rejection of IP should not result in singling out a single specific Abelian algebra of operators (i.e., measurement context), making the probability space effectively classical. Such a result would be unsatisfactory since, for example, position and momentum are both considered observables and they correspond to non-commuting operators.

The peculiarity of the situation has it that one can actually improve in both ways. It is possible to weaken IP to a more satisfactory assumption, and it is possible to construct non-trivial hidden variable models that reject IP and are nonetheless empirically equivalent to quantum mechanics. The envisioned weakening of IP was already provided by Kochen and Specker themselves. Instead of requiring that every self-adjoint operator corresponds to an observable, it suffices to select a specific finite subset of $\mathcal{O}_{\mathrm{sa}}$. This alternative proof to the Kochen-Specker theorem is explained in section 7.2.

In section 7.3 an argument is explained that establishes that even the weakened version of IP can be rejected while respecting the empirical predictions of quantum mechanics. Roughly speaking, the argument rests on the idea that, due to the finite precision of measurements, elements of $\mathcal{O}_{\mathrm{sa}}$ that are close to each other cannot be distinguished empirically. This is then used to deny certain operators the status of observable. This is the so-called finite precision loophole for the Kochen-Specker theorem. In section 7.4 it is shown that this loophole can be exploited to reconstruct the quantum probability functions on a space that respects NC.

The classical representations are further investigated in section 7.5 where a positive and a negative aspect are highlighted. The positive aspect is that one can establish a formal argument that recovers the Born rule in these classical representations. The negative aspect is a tension within the models themselves reminiscent of the tension found in the models that reject NC. As discussed in section 6.4, these latter models require a simultaneous acceptance and rejection of contextuality: acceptance on the level of the hidden variables, but rejection on the level of probability distributions. The models that reject IP in turn require a similar simultaneous acceptance and rejection of a particular kind of continuity. Finally, in section 7.6, a short evaluation is given concerning the results obtained on classical representations.

7.2   Weakening the identification principle

The role of IP in the proof of Theorem 6.1 is to allow the application of Gleason's

theorem. It ensures that the hidden variable state λ_ω in (6.12) defines a frame function on $L_1(\mathcal{H})$. Thus weakening IP requires a new proof that is independent of Gleason's theorem. Such a proof was given by Kochen and Specker, who also recognized the unsatisfactory nature of IP. Instead of proving the non-existence of classical random variables for all self-adjoint operators, they focused on a specific finite set of self-adjoint operators.

Although the proof of Kochen and Specker applies to all Hilbert spaces with $\dim(\mathcal{H}) > 2$, I focus here on the three-dimensional case. The entire philosophical discussion can be held within this context without loss of generality. The Hilbert space $\mathbb{C}^3$ is the space used to describe spin-1 particles, and it suffices to restrict attention to the operators associated with the squared spin of spin-1 particles (see Example 3.2). For these operators it is uncontroversial that they correspond to observables.

As noted in Example 6.2, for any triple of orthogonal axes r_1, r_2, r_3 in $\mathbb{R}^3$ the corresponding squared spin operators $S_{r_1}^2, S_{r_2}^2, S_{r_3}^2$ sum up to $2\mathbb{1}$. Accordingly, the values assigned to the corresponding observables by a hidden variable state should also sum up to two. In their original proof, Kochen and Specker constructed a set of 117 axes for which such an assignment is not possible. Here I make use of the set constructed by Peres (2002, p. 198) which only requires 33 axes, namely, those given by the vectors in Table 7.1. But the reader may keep in mind that which specific set one uses is not relevant for the philosophical argument though, as long as it is a finite subset of $\mathcal{O}_{\text{sa}}$ of which the elements can be thought of as observables. What it boils down to, is that the assumption IP in Theorem 6.3 can be replaced by the following weaker assumption.

IP$_{\text{KS}}$ (Kochen-Specker Identification Principle) For every self-adjoint operator A in a specific finite subset $\mathcal{O}_{\text{KS}} \subset \mathcal{O}_{\text{sa}}$ there exists an observable $\mathcal{A}$ such that A is associated with $\mathcal{A}$ via OP.

Theorem 7.1 (Kochen-Specker 1967). *For a system described by a Hilbert space of at least dimension 3, the assumptions QM$_{\text{KS}}$, HV, IP$_{\text{KS}}$, NC taken together lead to a contradiction.*

The proof presented here is restricted to the 3-dimensional case where

$$\mathcal{O}_{\text{KS}} = \left\{ S_r^2 \mid r \in D_{\text{KS}} \right\} \tag{7.1}$$

with D_{KS} given by Table 7.1. The proof further relies on the following definition and lemma.

Definition 7.1. Let $\mathbb{S}^2$ denote the 2-sphere, i.e., the set of unit vectors in $\mathbb{R}^3$. A function $c : \mathcal{D} \subset \mathbb{S}^2 \to \{0, 1\}$ is called a *coloring function* if

1. $c(a) = c(-a)$ whenever $a, -a \in \mathcal{D}$,

2. $c(a_1) + c(a_2) + c(a_3) = 2$ whenever $(P_{a_i})_{i=1}^{3}$ is a frame and $a_1, a_2, a_3 \in \mathcal{D}$,

3. $c(a_1) + c(a_2) \geq 1$ whenever $a_1 \perp a_2$ and $a_1, a_2 \in \mathcal{D}$.

Table 7.1: The 33 vectors of the set D_{KS}.

$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$	$e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$	$e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$
$f_1^1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad f_1^2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$	$f_2^1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad f_2^2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$	$f_3^1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad f_3^2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$
$g_1^1 = \begin{pmatrix} 0 \\ 1 \\ \sqrt{2} \end{pmatrix}, \quad g_1^2 = \begin{pmatrix} 0 \\ -\sqrt{2} \\ 1 \end{pmatrix}$	$g_2^1 = \begin{pmatrix} \sqrt{2} \\ 0 \\ 1 \end{pmatrix}, \quad g_2^2 = \begin{pmatrix} 1 \\ 0 \\ -\sqrt{2} \end{pmatrix}$	$g_3^1 = \begin{pmatrix} 1 \\ \sqrt{2} \\ 0 \end{pmatrix}, \quad g_3^2 = \begin{pmatrix} -\sqrt{2} \\ 1 \\ 0 \end{pmatrix}$
$g_1^3 = \begin{pmatrix} 0 \\ \sqrt{2} \\ 1 \end{pmatrix}, \quad g_1^4 = \begin{pmatrix} 0 \\ -1 \\ \sqrt{2} \end{pmatrix}$	$g_2^3 = \begin{pmatrix} 1 \\ 0 \\ \sqrt{2} \end{pmatrix}, \quad g_2^4 = \begin{pmatrix} \sqrt{2} \\ 0 \\ -1 \end{pmatrix}$	$g_3^3 = \begin{pmatrix} \sqrt{2} \\ 1 \\ 0 \end{pmatrix}, \quad g_3^4 = \begin{pmatrix} -1 \\ \sqrt{2} \\ 0 \end{pmatrix}$
$h_1^1 = \begin{pmatrix} \sqrt{2} \\ -1 \\ 1 \end{pmatrix}, \quad h_1^2 = \begin{pmatrix} \sqrt{2} \\ 1 \\ -1 \end{pmatrix}$	$h_2^1 = \begin{pmatrix} 1 \\ \sqrt{2} \\ -1 \end{pmatrix}, \quad h_2^2 = \begin{pmatrix} -1 \\ \sqrt{2} \\ 1 \end{pmatrix}$	$h_3^1 = \begin{pmatrix} -1 \\ 1 \\ \sqrt{2} \end{pmatrix}, \quad h_3^2 = \begin{pmatrix} 1 \\ -1 \\ \sqrt{2} \end{pmatrix}$
$h_1^3 = \begin{pmatrix} \sqrt{2} \\ -1 \\ -1 \end{pmatrix}, \quad h_1^4 = \begin{pmatrix} \sqrt{2} \\ 1 \\ 1 \end{pmatrix}$	$h_2^3 = \begin{pmatrix} -1 \\ \sqrt{2} \\ -1 \end{pmatrix}, \quad h_2^4 = \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}$	$h_3^3 = \begin{pmatrix} -1 \\ -1 \\ \sqrt{2} \end{pmatrix}, \quad h_3^4 = \begin{pmatrix} 1 \\ 1 \\ \sqrt{2} \end{pmatrix}$

Lemma 7.1. *Let $\mathcal{D}_{\mathrm{KS}} \subset \mathbb{S}^2$ be given by*

$$\mathcal{D}_{\mathrm{KS}} = \left\{ \frac{r}{\|r\|} \,\middle|\, r \in D_{\mathrm{KS}} \right\} \tag{7.2}$$

where D_{KS} is given by table 7.1. Then there is no coloring function $c : \mathcal{D}_{\mathrm{KS}} \to \{0,1\}$.

Proof. The impossibility of a coloring function $c : \mathcal{D}_{\mathrm{KS}} \to \{0,1\}$ is demonstrated with the use of the graph in Figure 7.1.[1] In this graph the nodes represent the elements of $\mathcal{D}_{\mathrm{KS}}$ and a line connecting two nodes indicates that the corresponding vectors are orthogonal. A coloring function c has to assign exactly two nodes of every triangle the value 1, and of every pair of nodes connected with a line at least one has to be assigned the value 1.

Since (e_1, e_2, e_3) forms a triangle, one of the nodes gets the value 0. Because of the symmetry of the diagram, one may assume that $c(e_3) = 0$ without loss of generality. Consequently, e_1 is assigned 1. Now note that (e_1, g_1^2, g_1^4) and (e_1, g_1^3, g_1^1) both form a triangle. This gives rise to several options. Either both g_1^1 and g_1^4 are

---

[1]Each of the three large triangles in this graph is an example of the 13 rays used by Yu and Oh (2012) in their derivation of a Kochen-Specker inequality. In (Blanchfield, 2012) one finds the graph that pastes the three triangles together upon which the graph used here is based.

assigned 0, or at least one of them is assigned the value 1. In this last case, at least one of g_1^2 or g_1^3 is assigned 0. Each of these three options will be shown to lead to a contradiction.
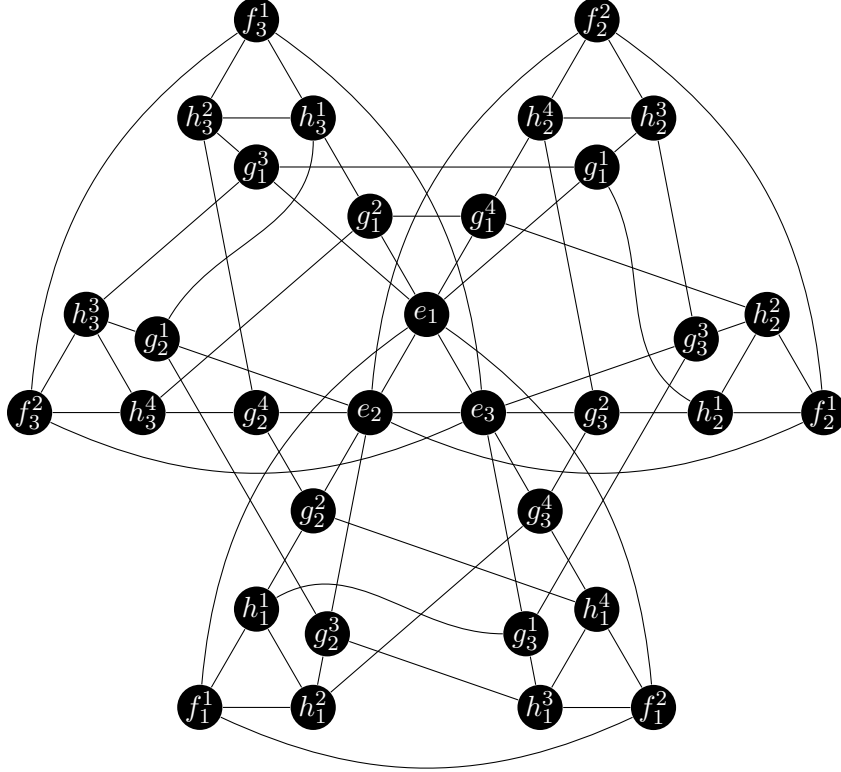

Figure 7.1: Graph demonstrating the orthogonality relations between the axes in D_{KS}.

Suppose $c(g_1^1) = c(g_1^4) = 0$. Then h_2^1, h_2^2, h_2^3 and h_2^4 are assigned the value 1. Consequently, f_2^1 and f_2^2 are assigned 0, which leads to a contradiction because they form a triangle together with e_2. Thus there is no coloring function with $c(e_3) = c(g_1^1) = c(g_1^4) = 0$.

Suppose $c(g_1^2) = 0$. Then h_3^1 and h_3^4 become 1. Also, because $c(e_3) = 0$, one has $c(f_3^1) = c(f_3^2) = 1$. This together implies that both h_3^2 and h_3^3 are assigned 0. The first implies that $c(g_2^4) = 1$ and the second that $c(g_2^1) = 1$. Note that these occur in the triangles (e_2, g_2^1, g_2^3) and (e_2, g_2^2, g_2^4). Because $c(e_2) = 1$, both g_2^2 and g_2^3 are assigned 0. This implies h_1^1, h_1^2, h_1^3 and h_1^4 are all assigned the value 1. Consequently, $c(f_1^1) = c(f_1^2) = 0$, showing there is no coloring function with both $c(e_3) = 0$ and $c(g_1^2) = 0$.

The assumption $c(g_1^3) = 0$ leads to a contradiction in an analogous way. This completes the proof. $\qquad\square$

Proof of Theorem 7.1. The idea of the proof is to show that the assumptions QM_{KS}, HV, IP_{KS}, NC together imply the existence of a coloring function on $\mathcal{D}_{KS}$. A contradiction is then obtained by invoking Lemma 7.1.

Because of NC and IP_{KS}, every operator in the set $\left\{ S_r^2 \mid r \in D_{KS} \right\}$ corresponds to precisely one observable. HV requires that each of these observables is assigned a definite value, and because of QM_{KS} this value has to lie in $\{0, 1\}$. Taken together, this implies the existence of a function $c : \mathcal{D}_{KS} \to \{0, 1\}$ via

$$c \left(\tfrac{r}{\|r\|} \right) = \lambda(S_r^2). \tag{7.3}$$

It follows from QM_{KS} that c satisfies the further properties of a coloring function. This proves the theorem. □

This result shows that Kochen and Specker took great care to weaken IP enough to make its use plausibly legitimate. A rejection of IP_{KS} now requires an argument to deny that one of the operators S_r^2 for $r \in D_{KS}$ may be attributed the status of observable. Such an argument (to be explained in the next section) was first given by Meyer (1999) accompanied with the construction of a non-contextual hidden variable model for spin-1 particles that led him to conclude the 'nullification' of the theorem. This result caused some controversy, and discussions on what the Kochen-Specker theorem 'actually' proves followed (Appleby, 2000; 2001; 2002; 2005; Barrett and Kent, 2004; Cabello, 1999; 2002; Mermin, 1999). To some extent this controversy also seemed to revolve around the meaning and use of the term 'nullification'. This latter controversy seems to me resolvable without much fuzz, the former deserves its own section though.

If 'nullifying' is taken to mean 'to strip it of all its foundational importance', then the phrasing seems just wrong. This follows from the discussions in the previous sections which show that, although rejecting IP is a logical option, doing so is no trivial matter. A more modest reading of 'nullifying' I think is 'to show that it doesn't prove what it intended to prove'. On this reading my take would be that if the Kochen-Specker theorem was supposed to prove the "nonexistence of hidden variables", as claimed by the authors themselves, then its nullification was already proven by Bohm's hidden variable theory in 1952, and the explanation for this nullification was provided by Bell (1966). But the most interesting reading of Meyer's claim I think is that it would show that the received interpretation of the theorem is flawed. Namely, the theorem does not show the impossibility of non-contextual hidden variable theories. Logically this is a truism. What has been established in this section is that the conditional (6.19) can be replaced by

$$QM_{KS} \wedge HV \wedge IP_{KS} \wedge NC \Rightarrow \bot. \tag{7.4}$$

But from this one cannot infer the conditional

$$QM_{KS} \wedge HV \Rightarrow \neg NC. \tag{7.5}$$

But similarly, it neither follows that the negation of the conditional (7.5) is true. More work is to be done for that, which I provide in the next section using the formal argument known as the finite precision argument. An explicit model, based on the work of Clifton and Kent (2001), in which $\mathrm{QM_{KS}}$, HV and NC hold is presented. As such, the received view of the theorem is 'nullified', but only by replacing it with a deeper understanding of the theorem.

7.3   The finite-precision argument

The MKC models are a specific kind of attempt to demonstrate the logical possibility of non-contextual hidden variable theories. Noticeable other attempts are the work of Pitowsky (1983 ; 1985) and La Cour (2009). The first one will not be discussed here because it adopts a peculiar kind of non-classical probability in which, for example, the conjunction of two probability-1 events can have probability zero. Such a model can hardly be considered a *classical* representation. The second manages to uphold both NC and IP by rejecting FM. Violations with quantum mechanical predictions (such as EFR) are circumvented by adopting a peculiar dynamics. This departure from static considerations takes us too far adrift from the discussion of classical representations and the theory will therefore not be discussed.

Before going into the finite precision argument, it is good to briefly recapitulate what is formally at stake. The assumptions NC and IP both concern the relation between operators and observables. Both presuppose the validity of the observable postulate (OP) of quantum mechanics, which states that observables can be associated with operators. Specifically, it postulates the existence of a function $f : \mathit{Obs} \to \mathcal{O}_\mathrm{sa}$. This function establishes that formal proofs concerning operators can have conceptual implications. However, the postulate does not specify what the function is, and this provides a gap for evaluating the philosophical implications of the formal proofs.

Formal proofs have conceptual implications only to the extent that they also apply to the relevant subset $f(\mathit{Obs}) \subset \mathcal{O}_\mathrm{sa}$. The difficulty of this construction is that there is no definite way to determine precisely what this subset is and thus one requires further assumptions. Kochen and Specker showed that there is a specific finite set $\mathcal{O}_\mathrm{KS} \subset \mathcal{O}_\mathrm{sa}$ such that if the function f satisfies

$$\mathcal{O}_\mathrm{KS} \subset f(\mathit{Obs}) \subset \mathcal{O}_\mathrm{sa}, \tag{7.6}$$

then at least one of the assumptions $\mathrm{QM_{KS}}$, HV, or NC has to fail. Many results have been presented since the original proof that aim to make $\mathcal{O}_\mathrm{KS}$ as small as possible.[2] What is left open in this competition is the converse question. How big

---

[2]Depending on the exact notion of 'small' there are several contests running on this topic. The

can a set of operators $\mathcal{O}_{\mathrm{MKC}}$ be such that if f satisfies

$$f(\mathit{Obs}) \subset \mathcal{O}_{\mathrm{MKC}} \subset \mathcal{O}_{\mathrm{sa}}, \tag{7.7}$$

then $\mathrm{QM_{KS}}$, HV and NC can be satisfied? This is the question Meyer (1999), Kent (1999) and Clifton and Kent (2001) (MKC) focused on, and it turns out that, at least for finite-dimensional Hilbert spaces, this set can be surprisingly big.[3]

From a purely logical perspective one seems to arrive at a stalemate here. If f satisfies $\mathcal{O}_{\mathrm{KS}} \subset f(\mathit{Obs})$, then non-contextual hidden variables are not possible, as shown by the Kochen-Specker theorem. But if on the other hand f satisfies $f(\mathit{Obs}) \subset \mathcal{O}_{\mathrm{MKC}}$, then they are possible, as shown by MKC. The tie-breaker is provided by the finite precision argument, which favors the second option. Before explaining this argument in its general form, I first return to the spin-1 particles that figure prominently in the Kochen-Specker theorem.

It is useful to think of measurements of spin-1 particles as being analogous to measurements of spin-$\frac{1}{2}$ particles.[4] A measurement of σ_r (see Example 3.1) is usually done with a Stern-Gerlach device which is then aligned to the axis r. The idea that every operator in the set $\{\sigma_r \mid r \in \mathbb{R}^3, \|r\| = 1\}$ corresponds to an observable is thus motivated by the idea that a macroscopic measurement device can be oriented in all possible ways in space. The same view may be used to motivate that all the operators in the set

$$\mathcal{O}_{\mathrm{spin}} := \left\{ S_r^2 \mid r \in \mathbb{R}^3, \ \|r\| = 1 \right\} \tag{7.8}$$

correspond to observables. Then any finite subset, like $\mathcal{O}_{\mathrm{KS}}$ in (7.1), does so as well.

The finite precision argument counters this conclusion. To determine whether one measures the squared spin along some axis r_1 or r_2, one has to determine whether some macroscopic device is aligned along either axis. But this direction can, of course, only be determined up to some finite precision. Thus for any direction r there is always some $\epsilon > 0$ for which one cannot determine which of the directions

---

size also depends on the dimension of the Hilbert space under investigation. In three dimensions Conway and Kochen hold the record for the smallest number of vectors (31) (Peres, 2002, p. 114). The constructions of Peres (ibid., p. 198) and Bub (1996) both use 33 vectors of which the latter requires the least number of frames of all these results. If instead of vectors the number of frames is to be minimized, then the record is held by Lisoněk et al. (2014) using only 7 frames for the Hilbert space $\mathbb{C}^6$. For more discussion on comparing sizes of Kochen-Specker sets see (Bengtsson, 2012; Pavičić et al., 2005) and references therein.

[3]It deserves to be noted that this question was also investigated by Bub and Clifton (1996), albeit from a different perspective.

[4]The reason for this little side-track is that I do not know of any actual experiments on spin-1 particles. Actual spin-1 particles such as the W and Z bosons or Higgs bosons are hard to manipulate, and experiments on similarly behaving systems (e.g. qutrits composed of photon pairs such as in (Lanyon et al., 2008)) also have a long way to go in comparison to the manipulation of qubit systems. On the other hand, the analogy with spin-$\frac{1}{2}$ measurements may be quite apt (Swift and Wright, 1980).

r' with $\|r - r'\| < \epsilon$ actually applies to the observable under investigation. The recognition of the practical underdetermination of directions was vividly exploited by Meyer (1999), who showed that it is possible to assign definite values to the operators in the set

$$\mathcal{O}'_{\text{spin}} := \left\{ S^2_r \mid r \in \mathbb{Q}^3, \ \|r\| = 1 \right\} \tag{7.9}$$

in a non-contextual way. Then, to determine whether a certain observable S^2_r corresponds to an operator in $\mathcal{O}'_{\text{spin}}$ or in $\mathcal{O}_{\text{spin}} \backslash \mathcal{O}'_{\text{spin}}$ requires empirically distinguishing rational numbers from irrational numbers. As this is generally considered to be impossible, there is no empirical reason to assert that there are operators in $\mathcal{O}_{\text{spin}} \backslash \mathcal{O}'_{\text{spin}}$ that *do* correspond to observables. The finite precision argument then is that, because one cannot determine with infinite precision which operator precisely should correspond to the observable (experimental setup) under investigation, one may as well assume that the operator is an element of a particular dense subset ($\mathcal{O}'_{\text{spin}}$) of all the operator candidates ($\mathcal{O}_{\text{spin}}$).

The argument may come off as being contrived. The idea that $\mathbb{R}^3$ is a good (approximate) model of our spatial surroundings is deeply entrenched in our scientific thinking. To replace it with $\mathbb{Q}^3$ then seems a desperate step just to save NC. There are several ways to respond to this. The most neutral one is that we are interested here in logical possibilities that need not be physically attractive. But somewhat more boldly, one can argue that holding on to $\mathbb{R}^3$ is itself a desperate step to save the standard reading of the Kochen-Specker theorem. And finally, one can deny the step going from the idea that $\mathbb{R}^3$ correctly models space, to the conclusion that the squared spin along the axis r is an observable for every spatial direction. Perhaps macroscopic measurement devices make an infinitesimal shift in alignment just to line up with some direction in $\mathbb{Q}^3$ in the case of a measurement.

The finite precision argument as it is used here works for spin observables specifically. In this argument the fact is used that the observables $S^2_{r_1}$ and $S^2_{r_2}$ will resemble each other when r_1 and r_2 start to approach each other. That is, whenever the operators $S^2_{r_1}$ and $S^2_{r_2}$ are close to each other, it is used that the experimental setups used to measure $S^2_{r_1}$ and $S^2_{r_2}$ are also close to each other in some sense. What is required to make the finite precision argument work in arbitrary cases, is that something similar applies in general. That is, the metric on *Obs* defined by[5]

$$d(\mathcal{A}_1, \mathcal{A}_2) := \|A_1 - A_2\| \tag{7.10}$$

has to be relevant in the following way:

CoO (Continuity of Observables) For any observable $\mathcal{A} \in \textit{Obs}$ and sequence of observables $\mathcal{A}_1, \mathcal{A}_2, \ldots$, if $d(\mathcal{A}_n, \mathcal{A})$ tends to zero as $n \to \infty$, then there is a

---

[5]There is nothing stringent concerning the choice of this particular metric because the finite precision argument itself rests on fuzzy terms concerning the resemblance of experimental setups.

sense in which the sequence of experimental setups associated with the $\mathcal{A}_n$ will resemble an experimental setup suitable for the measurement of $\mathcal{A}$ as $n \to \infty$.

The assumption is one that I think is tacitly accepted in quantum mechanics. It is part of the application of OP to select/construct the function $f : \mathit{Obs} \to \mathcal{O}_{\mathrm{sa}}$ in such a way that CoO comes out true. A closely related fact is that in quantum mechanics, when one tries to pinpoint the exact operator to be identified with a particular observable, there is some wiggle room due to the finite precision with which experimental setups can be specified. This is a consequence of the fact that, for two operators A_1, A_2 that are close to each other, the quantum mechanical predictions also resemble each other. For example, the following theorem holds.

Theorem 7.2. *Let A be a self-adjoint operator acting on a finite-dimensional Hilbert space, and let $(A_n)_{n \in \mathbb{N}}$ be a sequence of self-adjoint operators such that $\sigma(A) = \sigma(A_n)$ for all n and $\lim_{n \to \infty} \|A - A_n\| = 0$. Then for every $\Delta \subset \sigma(A)$ and density operator ρ one has*

$$\lim_{n \to \infty} |\mathrm{Tr}(\rho\mu_A(\Delta)) - \mathrm{Tr}(\rho\mu_{A_n}(\Delta))| = 0. \tag{7.11}$$

Proof. Because

$$|\mathrm{Tr}(\rho\mu_A(\Delta)) - \mathrm{Tr}(\rho\mu_{A_n}(\Delta))| \leq \sum_{a \in \Delta} |\mathrm{Tr}(\rho\mu_A(\{a\})) - \mathrm{Tr}(\rho\mu_{A_n}(\{a\}))|, \tag{7.12}$$

it suffices to show that (7.11) holds for singleton sets. By making use of Lemma A.5, it follows that

$$\begin{aligned} |\mathrm{Tr}(\rho\mu_A(\{a\})) - \mathrm{Tr}(\rho\mu_{A_n}(\{a\}))| &= |\mathrm{Tr}\left(\rho(\mu_A(\{a\}) - \mu_{A_n}(\{a\}))\right)| \\ &\leq \mathrm{Tr}(|\rho|)\|\mu_A(\{a\}) - \mu_{A_n}(\{a\})\| \qquad (7.13) \\ &= \|\mu_A(\{a\}) - \mu_{A_n}(\{a\})\|. \end{aligned}$$

Now for every $a \in \sigma(A)$ define the polynomial

$$p_a(x) := \prod_{\substack{a' \in \sigma(A) \\ a' \neq a}} \frac{x - a'}{a - a'}. \tag{7.14}$$

Restricted to $\sigma(A)$, this function acts as a Kronecker delta function: $p_a(a') = \delta_{aa'}$. Therefore

$$\|\mu_A(\{a\}) - \mu_{A_n}(\{a\})\| = \|p_a(A) - p_a(A_n)\|. \tag{7.15}$$

It now follows from Lemma A.6 that (7.11) holds. This completes the proof. $\qquad \square$

The restriction that $\sigma(A) = \sigma(A_n)$ for all n is a bit severe, and it is possible to weaken it. However, doing so makes the theorem a lot messier[6] and the formulation given here suffices for the present discussion. The problem at hand is that, given an experimental procedure, one has to select a corresponding self-adjoint operator. Since the set of possible outcomes is determined by the experimental procedure, the choice is effectively between operators with the same spectrum. Now one may object that, because of the finite precision of measurements, one is not able to determine exactly what this spectrum should be. This is true of course, but this can be viewed as part of a separate selection process, where one can choose between different scalings of the spectrum on a collection of operators with the same spectrum. The selection of an operator is then split into the selection of a spectrum, and of a set of projection operators. This splitting can be justified by noting that scaling the spectrum does not alter the relevant probabilities (see also the discussion on page 38).

To make the finite precision argument work one has to show that for every observable $\mathcal{A}$ and every possible candidate $A \in \mathcal{O}_{\mathrm{sa}}$ for representing $\mathcal{A}$, there exists an operator $A' \in \mathcal{O}_{\mathrm{MKC}}$ that is just as good a candidate as A given the finite precision with which the properties of $\mathcal{A}$ can be measured. Here $\mathcal{O}_{\mathrm{MKC}}$ is the set of operators corresponding to observables that can be assigned definite values by a hidden variable state in a non-contextual way. That is, for every possible function $f : \mathcal{O}bs \to \mathcal{O}_{\mathrm{sa}}$ there should be a function $f' : \mathcal{O}bs \to \mathcal{O}_{\mathrm{MKC}}$ such that $\|f(\mathcal{A}) - f'(\mathcal{A})\|$ is small for all $\mathcal{A} \in \mathcal{O}bs$. It is a corollary of a theorem by Kent (1999) that such a set can indeed be constructed. Specifically, the following was shown.[7]

Theorem 7.3. *For every finite-dimensional Hilbert space there exists a countable set of frames* $\mathcal{F}_{\mathrm{MKC}} := \{(P_i^n) \mid n \in \mathbb{N}\}$ *that is dense in the set of all frames, i.e., for every frame* (P_i) *and every* $\epsilon > 0$, *there exists an* $n \in \mathbb{N}$ *such that*

$$\max_i \min_j \|P_i - P_j^n\| < \epsilon. \tag{7.16}$$

Furthermore, there exist functions

$$\lambda : L_1^{\mathrm{MKC}}(\mathcal{H}) \to \{0, 1\} \tag{7.17}$$

that satisfy

$$\sum_i \lambda(P_i^n) = 1 \ \forall n \in \mathbb{N}, \tag{7.18}$$

---

[6]For example, it requires introducing a sequence of maps that link subsets of $\sigma(A)$ to subsets of $\sigma(A_n)$ and adjust (7.11) accordingly. Without such an adjustment the theorem would no longer hold as one can construct sequences $(A_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} \|A - A_n\| = 0$ while $\sigma(A) \cap \sigma(A_n) = 0$ for all n.

[7]Although it is implicitly assumed here that measurements can be represented by PVMs (as a consequence of OP), nothing hinges on this assumption. A similar theorem can be proven for POVMs.

where $L_1^{\mathrm{MKC}}(\mathcal{H})$ is the set of all 1-dimensional projections that occur in one of the frames in $\mathcal{F}_{\mathrm{MKC}}$.

Corollary 7.1. *For every finite-dimensional Hilbert space there is a set of self-adjoint operators $\mathcal{O}_{\mathrm{MKC}} \subset \mathcal{O}_{\mathrm{sa}}$ such that*

(i) *$\mathcal{O}_{\mathrm{MKC}}$ is closed under functional operations: for every real-valued Borel function f, if $A \in \mathcal{O}_{\mathrm{MKC}}$, then $f(A) \in \mathcal{O}_{\mathrm{MKC}}$,*

(ii) *$\mathcal{O}_{\mathrm{MKC}}$ is dense in $\mathcal{O}_{\mathrm{sa}}$: for every $\epsilon > 0$ and every $A' \in \mathcal{O}_{\mathrm{sa}}$ there is an $A \in \mathcal{O}_{\mathrm{MKC}}$ such that $\|A - A'\| < \epsilon$,*

and there exist functions $\lambda : \mathcal{O}_{\mathrm{MKC}} \to \mathbb{R}$ such that

$$\begin{aligned} \lambda(A) &\in \sigma(A), \\ \lambda(f(A)) &= f(\lambda(A)) \end{aligned} \tag{7.19}$$

for every $A \in \mathcal{O}_{\mathrm{MKC}}$ and real-valued Borel function f.

Proof. Let $\{(P_i^n) \mid n \in \mathbb{N}\}$ be a set frames as in Theorem 7.3. For every $n \in \mathbb{N}$ let $\mathcal{A}_n$ be the Abelian von Neumann algebra generated by the frame (P_i^n). Now define

$$\mathcal{O}_{\mathrm{MKC}} := \bigcup_{n \in \mathbb{N}} (\mathcal{O}_{\mathrm{sa}} \cap \mathcal{A}_n). \tag{7.20}$$

It immediately follows that criterion (i) is satisfied because the algebras are closed under the functional operations.

Now suppose $A' \in \mathcal{O}_{\mathrm{sa}}$ and $\epsilon > 0$ are given. Let (P_i) be a frame such that $A' = \sum_i a_i' P_i$. Now choose a frame (P_i^n) such that $\|P_i - P_i^n\| < \frac{\epsilon}{d\|A\|}$ for all i, where d is the dimension of the Hilbert space, and define $A := a_i' P_i^n$. It follows that

$$\|A - A'\| = \left\| \sum_i a_i'(P_i^n - P_i) \right\| \leq \sum_i |a_i'| \|P_i^n - P_i\| < \epsilon. \tag{7.21}$$

This establishes that (ii) holds.

Finally, every function λ that satisfies the criteria in Theorem 7.3 can be extended to a function $\overline{\lambda}$ on $\mathcal{O}_{\mathrm{MKC}}$ that satisfies (7.19). To see this note that for every $A \in \mathcal{O}_{\mathrm{MKC}}$ there exists a frame (P_i^n) such that $A = \sum_i a_i P_i^n$, with $a_i \in \sigma(A)$. Now define

$$\overline{\lambda}(A) := \sum_i a_i \lambda(P_i^n). \tag{7.22}$$

Then (7.19) follows from the spectral theorem together with the Borel functional calculus. □

The finite precision argument, together with the above theorem and corollary, motivate the following weakening of IP.

IP$_{\text{MKC}}$ (Meyer-Kent-Clifton Identification Principle) For every self-adjoint operator $A \in \mathcal{O}_{\text{MKC}}$ there exists an observable $\mathcal{A}$ such that A is associated with $\mathcal{A}$ via OP.

It follows from Corollary 7.1 that QM$_{\text{KS}}$, HV, NC and IP$_{\text{MKC}}$ can be jointly satisfied. In short, the result explicitly disproves the credo that the Kochen-Specker theorem establishes the impossibility of non-contextual hidden variable theories. On the more cautious side though, what has been established thus far is the possibility of non-contextual hidden variable states. What is still an open question at this point, is whether these states can be used to construct a classical representation of quantum probability. The answer is yes, but showing this requires a bit more work.

7.4 Recovering probabilities

The question at hand is if every quantum probability function (represented by a density operator ρ) can be modeled by a classical probability function $\mathbb{P}_\rho$ on the set

$$\Lambda_{\text{MKC}} := \left\{ \lambda : \mathcal{O}_{\text{MKC}} \to \mathbb{R} \, \middle| \, \forall f, A : {}^{\lambda(A) \in \sigma(A),}_{\lambda(f(A)) = f(\lambda(A))} \right\}. \tag{7.23}$$

That is, if for every ρ there exists a $\mathbb{P}_\rho$ such that

$$\mathbb{P}_\rho \left(F_A \in \Delta \right) = \text{Tr} \left(\rho \mu_A(\Delta) \right), \tag{7.24}$$

for all $A \in \mathcal{O}_{\text{MKC}}$ and $\Delta \subset \sigma(A)$ where F_A is the random variable defined as

$$F_A(\lambda) := \lambda(A). \tag{7.25}$$

Clifton and Kent (2001) showed that the answer to this question is yes. But much more interesting than the answer itself, is the way the answer is obtained. After all, it is only then that one can evaluate the classical representation of quantum probability provided in this approach.

The way quantum probabilities can be reproduced in MKC models is best understood by backing away from the Kochen-Specker theorem, and focusing attention to the general idea of reconstructing quantum probabilities. Let $\mathcal{H}$ be a finite-dimensional Hilbert space and let $\mathcal{P} := \{P_1, \ldots, P_n\}$ be a frame and $\mathcal{O}_{\text{sa}}(\mathcal{P})$ be the set of self-adjoint operators that commute with every element of $\mathcal{P}$. This means that $\mathcal{O}_{\text{sa}}(\mathcal{P})$ is the set of self-adjoint operators in the Abelian von Neumann algebra generated by $\mathcal{P}$. There is a straightforward way of finding a classical representation of any quantum probability function for this set of operators. In this case one can set

$$\Lambda_{\mathcal{P}} := \left\{ \lambda : \mathcal{O}_{\text{sa}}(\mathcal{P}) \to \mathbb{R} \, \middle| \, \forall f, A : {}^{\lambda(A) \in \sigma(A),}_{\lambda(f(A)) = f(\lambda(A))} \right\}. \tag{7.26}$$

This set contains exactly $n = \dim(\mathcal{H})$ elements, namely, one for each P_i:

$$\lambda_{P_i}(A) := \operatorname{Tr}(AP_i). \tag{7.27}$$

A density operator ρ now gives rise to a probability distribution over $\Lambda_{\mathcal{P}}$ by setting $\mathbb{P}_\rho(\lambda_{P_i}) := \operatorname{Tr}(\rho P_i)$. The reader may check that this probability distribution satisfies (7.24) for the operators in $\mathcal{O}_{\mathrm{sa}}(\mathcal{P})$.

The next step is to consider two frames $\mathcal{P}_1 := \{P_{1,1}, \ldots, P_{1,n}\}$ and $\mathcal{P}_2 := \{P_{2,1}, \ldots, P_{2,n}\}$ with the property that $[P_{1,i}, P_{2,j}] \neq 0$ for all $i, j \in \{1, \ldots, n\}$. Such frames are called *totally incompatible*. It follows that the only elements that $\mathcal{O}_{\mathrm{sa}}(\mathcal{P}_1)$ and $\mathcal{O}_{\mathrm{sa}}(\mathcal{P}_2)$ have in common are those of the form $c\mathbb{1}$ for $c \in \mathbb{R}$. Consequently, any pair $(\lambda_{P_{1,i}}, \lambda_{P_{2,j}})$ in $\Lambda_{\mathcal{P}_1} \times \Lambda_{\mathcal{P}_2}$ unambiguously defines an element $\lambda_{P_{1,i}, P_{2,j}}$ of

$$\Lambda_{\mathcal{P}_1, \mathcal{P}_2} := \left\{ \lambda : \mathcal{O}_{\mathrm{sa}}(\mathcal{P}_1) \cup \mathcal{O}_{\mathrm{sa}}(\mathcal{P}_2) \to \mathbb{R} \,\middle|\, \forall f, A : \begin{smallmatrix} \lambda(A) \in \sigma(A), \\ \lambda(f(A)) = f(\lambda(A)) \end{smallmatrix} \right\} \tag{7.28}$$

with the rule

$$\lambda_{P_{1,i}, P_{2,j}}(A) := \begin{cases} \lambda_{P_{1,i}}(A) & A \in \mathcal{O}_{\mathrm{sa}}(\mathcal{P}_1), \\ \lambda_{P_{2,j}}(A) & A \in \mathcal{O}_{\mathrm{sa}}(\mathcal{P}_2). \end{cases} \tag{7.29}$$

Conversely, every element of $\Lambda_{\mathcal{P}_1, \mathcal{P}_2}$ may be shown to be of this form and thus one has

$$\Lambda_{\mathcal{P}_1, \mathcal{P}_2} \simeq \Lambda_{\mathcal{P}_1} \times \Lambda_{\mathcal{P}_2}. \tag{7.30}$$

A probability distribution is now obtained by taking the product measure, resulting in

$$\mathbb{P}_\rho\left(\lambda_{P_{1,i}, P_{2,j}}\right) = \operatorname{Tr}(\rho P_{1,i}) \operatorname{Tr}(\rho P_{2,j}). \tag{7.31}$$

For a countable sequence $(\mathcal{P}_i)$ of totally incompatible frames the same procedure can be followed. The set of hidden variables can be taken to be

$$\Lambda_{(\mathcal{P}_i)} := \left\{ \lambda : \bigcup_i \mathcal{O}_{\mathrm{sa}}(\mathcal{P}_i) \to \mathbb{R} \,\middle|\, \forall f, A : \begin{smallmatrix} \lambda(A) \in \sigma(A), \\ \lambda(f(A)) = f(\lambda(A)) \end{smallmatrix} \right\}, \tag{7.32}$$

and the probability measure that reproduces the quantum probability rule is again the product measure.

The upshot is that for any countable sequence of totally incompatible frames, the assumptions $\mathrm{QM_{KS}}$, HV and NC can be jointly satisfied, provided that the set of operators associated with an observable is a subset of $\bigcup_i \mathcal{O}_{\mathrm{sa}}(\mathcal{P}_i)$. What Clifton and Kent showed is that a sequence of totally incompatible frames can be selected that satisfies the constraints from Theorem 7.3. Consequently, the set of operators $\bigcup_i \mathcal{O}_{\mathrm{sa}}(\mathcal{P}_i)$ satisfies the constraints of Corollary 7.1. That is, for every finite-dimensional Hilbert space, there is a countable sequence of totally incompatible frames such that $\bigcup_i \mathcal{O}_{\mathrm{sa}}(\mathcal{P}_i)$ is a dense subset of $\mathcal{O}_{\mathrm{sa}}$. With the adoption of

CoO, then, this set is big enough to accommodate any observable in quantum mechanics, and the above construction shows that every quantum probability function has a classical representation on the set $\Lambda_{(\mathcal{P}_i)}$. In short, the MKC models both provide an example of a non-contextual hidden variable theory, as well as a classical representation of quantum probability.

7.5 The role of continuity

Let me start this section with saying that the MKC models are highly artificial. As a candidate for an interpretation of quantum mechanics or a route to 'new physics', they perform a terrible job. Their value lies purely in their capacity to provide insight in the notion of contextuality in quantum mechanics, and the relation between quantum and classical probability. It is not my intention here to perform a re-hash of the pros and cons of these models, nor on the precise metaphysical meaning of contextuality in quantum mechanics. For this I refer the reader to the discussions in (Barrett and Kent, 2004; Appleby, 2005; Hermens, 2011) and references therein. Instead, I focus on some particular aspects of the way they perform as classical representations of quantum probability.

The finite precision argument may be seen to rest on two ingredients. The first is the uncontroversial idea that it is impossible to determine with infinite precision the parameters that fix the experimental setup associated with an observable. Indeed, if this were possible, the MKC models would not be empirically equivalent to quantum mechanics. One could 'just' check for any particular direction r if there is an observable corresponding to S_r^2. The scare quotes around 'just' indicate how implausible such a test is, making the finite precision assumption a very plausible one.

The second ingredient relates to the central theme of this section: continuity. To make the finite precision argument precise one needs a metric (or at least a topology) to specify when two experimental setups resemble each other. The assumption CoO states that this should be done by adopting the metric on the set of self-adjoint operators. This metric is plausible for spin-measurements: the two operators $S_{r_1}^2$ and $S_{r_1}^2$ are close to each other just in case the axes spanned by r_1 and r_2 are near each other. But in other situations its relevance may be less clear. For example, the metric also provides a distance between position and momentum measurements, and it is not clear how this relates to the experimental setups. However, this need not pose a difficulty. For CoO the distance between the position and momentum operator is only relevant when adopting the further assumption that there is a whole range of observables in between. It is this further assumption that is troublesome, rather than CoO itself. In fact, CoO is a relatively innocent assumption because it is formulated as a conditional.

But also when the finite precision assumption and CoO are accepted, there are reasons to reject the MKC models as faithful classical representations. I discuss here two reasons. The first is the by now familiar complaint that not all probability functions on $\Lambda_{(\mathcal{P})_i}$ are of the form $\mathbb{P}_\rho$. However, in this case there is a formal argument available that selects the quantum probability functions as a natural subset of all probability functions. After giving this argument below I discuss the second reason, which explains why CoO is an unsatisfactory assumption from the point of view of the MKC models themselves. Thus the MKC models undermine the very assumption that allows their construction.

7.5.1   Deriving the Born rule

The introduction of CoO was motivated by the idea that a similar claim tends to hold in quantum mechanics itself. This was illustrated by Theorem 7.2. The explanation of the special role of quantum probability functions in the MKC models given in this section uses the converse fact that, if probability functions satisfy a certain continuity assumption, they satisfy the Born rule. Thus, roughly speaking, probability functions in the MKC models satisfy the Born rule if and only if they are continuous. This statement will be made more precise below, after giving a possible defense for singling out continuous probability functions.

A particular way to motivate a continuity assumption for probability functions is by adopting an epistemic interpretation of probability.[8] That is not to say that such an interpretation is necessary here, but at least it provides a useful way to think about the formal side of the story. On an epistemic reading, probabilities are thought of as epistemic judgments of a rational agent concerning outcomes of particular measurements. However, when considering the outcomes of a particular actual measurement the agent not only has to reflect on the uncertainty pertaining to the outcome of the measurement, but also (due to the finite precision argument) on the uncertainty pertaining to which measurement is actually being performed. For example, consider the measurement of an observable corresponding to S_r^2. The agent is required to assign probabilities to the possible outcomes 0 and 1 for every possible value of r. But apart from the uncertainty about the outcome, the agent is also necessarily uncertain about the value of r. In order for this notion of probability to be made operational, the probabilities assigned to the possible outcomes 0 and 1 have to vary continuously with r. For example, when thinking of probabilities as betting ratios, one has to be able to agree on what the payoff will be after a measurement of S_r^2 without requiring that one determines with infinite precision

---

[8]Inspiration for this kind of motivation for a continuity assumption is drawn from the work of Malament and Zabell (1980) in statistical mechanics.

what the value of r is. Thus one has that

$$\lim_{r' \to r} \mathbb{P}\left(F_{S^2_{r'}} = 1\right) = \mathbb{P}\left(F_{S^2_r} = 1\right). \tag{7.33}$$

When taking into account CoO, this leads to the demand that $\mathbb{P}$ should be continuous with respect to the metric on the self-adjoint operators. Here I propose the following definition to make this precise.

Definition 7.2. A probability measure $\mathbb{P}$ on Λ_{MKC} is said to *respect CoO*, if for every $A \in \mathcal{O}_{\mathrm{MKC}}$ and every $\epsilon > 0$ there exists a $\delta > 0$ such that for all $A' \in \mathcal{O}_{\mathrm{MKC}}$ with $\sigma(A') = \sigma(A)$,

$$\|A - A'\| < \delta \;\Rightarrow\; |\mathbb{P}(F_A = a) - \mathbb{P}(F_{A'} = a)| < \epsilon \tag{7.34}$$

for all $a \in \sigma(A)$.

One can now show that the Born rule precisely singles out these continuous probability distributions. Conversely, from the perspective of the MKC models, the Born rule can be seen as a consequence of the continuity assumption. Note that although this continuity assumption may be justified with an epistemic interpretation of probability, the formal result is independent of the interpretation of probability.

Theorem 7.4. *A probability measure $\mathbb{P}$ on Λ_{MKC} respects CoO if and only if there is a density operator ρ such that*

$$\mathbb{P}\left(F_A \in \Delta\right) = \mathrm{Tr}\left(\rho \mu_A(\Delta)\right) \tag{7.35}$$

for all $A \in \mathcal{O}_{MKC}$ and $\Delta \subset \sigma(A)$.

Proof. Note that if $\mathbb{P}$ is given by a density operator ρ, then it follows from Theorem 7.2 that CoO is respected. It thus only has to be shown that, if $\mathbb{P}$ respects CoO, then there exists a density operator ρ such that (7.35) holds. Let $\mathbb{P}$ be a probability function that respects CoO. This function can be used to define a function $\lambda_{\mathbb{P}} : L(\mathcal{H}) \to [0,1]$ in the following way. Let $P \in L(\mathcal{H})$. According to Theorem 7.3 there is a sequence (P_i) in $L(\mathcal{H}) \cap \mathcal{O}_{\mathrm{MKC}}$ such that $P_i \to P$ as $i \to \infty$. Now define

$$\lambda_{\mathbb{P}}(P) := \lim_{i \to \infty} \mathbb{P}\left(F_{P_i} = 1\right). \tag{7.36}$$

Because $\mathbb{P}$ respects CoO this value is independent of the choice of the sequence. Thus the function $\lambda_{\mathbb{P}}$ is well-defined. Further, again because of CoO, this function satisfies $\lambda_{\mathbb{P}}\left(\sum_{i \in \mathbb{N}} P_i\right) = \sum_{i \in \mathbb{N}} \lambda_{\mathbb{P}}(P_i)$ for every countable sequence of pairwise orthogonal projection operators. Since it is also the case that $\lambda_{\mathbb{P}}(\mathbb{1}) = 1$ this defines a quantum probability function (see Definition 5.1). Finally, it follows from Gleason's theorem that (7.35) holds.

$\square$

It deserves to be emphasized that Theorem 7.4 is quite a remarkable result. It establishes success on a point where previous classical representations have failed: to derive the Born rule from within the classical model. Moreover, it strengthens a recurring idea in results on reconstructing quantum mechanics such as (Hardy, 2001), namely, that continuity is an essential ingredient. But the result also puts additional weight on the role of the assumption CoO, and it deserves to be investigated how well this assumption fits with the construction of the MKC models. The conclusion will be that it isn't a very attractive assumption.

7.5.2 The incompatibility with continuity

CoO establishes that the metric on $\mathcal{O}_{\mathrm{sa}}$ also plays a role on the set of observables $\mathcal{O}_{\mathrm{MKC}}$ of the MKC models. This adds additional structure of the Hilbert space formalism to the MKC models, on top of the functional relationships EFR. But from the point of view of the MKC models, this additional structure on the set is somewhat alienating. In the construction of $\Lambda_{(\mathcal{P}_i)}$ all relations between frames are placed on an equal footing. A state λ is blind to the distance between two observables whenever their corresponding operators do not commute. To put it another way, the MKC states do naturally reproduce EFR, but there is nothing in the MKC models that could even lead to the formulation of CoO.

The tension would be relieved if the metric could in some way be reflected within the properties of the MKC states. This would be the case if, typically, for observables that are close to each other, the values assigned to them are close to each other as well. The provisional 'typically' is to indicate that a too stringent demand is doomed to fail from the start. For example, in the 3-dimensional case, when restricting an MKC state λ to the squared spin observables S_r^2, this function cannot be a continuous function of r. This is because the range of the function is the discrete set $\{0, 1\}$, and both values have to be attained. Some discontinuities are therefore necessary. But there is still the possibility that these can be 'special' and that 'generally' the value of $\lambda(S_r^2)$ is a good indication of the values of $\lambda(S_{r'}^2)$ for r' close to r. It was shown by Appleby (2005) however, that this is not possible: there have to be wild discontinuities for every MKC state.

The theorem proven by Appleby is formulated in terms of coloring functions. Let $\mathbb{S}^2$ denote the 2-sphere whose points represent the squared spin observables. Now set

$$\mathbb{S}^2_{\mathrm{MKC}} := \left\{ r \in \mathbb{S}^2 \mid S_r^2 \in \mathcal{O}_{\mathrm{MKC}} \right\}. \tag{7.37}$$

This is a dense colorable subset of $\mathbb{S}^2$. In fact, every $\lambda \in \Lambda_{\mathrm{MKC}}$ provides a coloring function c_λ via

$$c_\lambda(r) := \lambda(S_r^2). \tag{7.38}$$

Theorem 7.5 (Appleby 2005)**.** *Let $\mathcal{O}_{\mathrm{MKC}}$ be an MKC set of observables for the Hilbert space $\mathbb{C}^3$. Then for every $\lambda \in \Lambda_{\mathrm{MKC}}$ there exists an open set $U \subset \mathbb{S}^2$ such that c_λ is discontinuous at every point of $U \cap \mathbb{S}^2_{\mathrm{MKC}}$.*

Although the theorem pertains to spin-1 systems specifically, it is likely to be generalizable to higher dimensional systems. For Appleby, the importance of this result lies in the implication that, in general, measuring some observable A doesn't provide any information about the state λ of the system. It doesn't do so in principle, whenever the measurement is of an observable A in a region in which λ is densely discontinuous. But also doesn't do so pragmatically, since there is no way to have empirical access to the question whether A is in such a region or not. The broader lesson Appleby draws from the (Bell-)Kochen-Specker theorem then is not so much on the (im)possibility of certain kinds of hidden variables. Rather, the theorem poses a limit on the extent to which such hidden variables are empirically accessible. Thus Appleby reads in the Kochen-Specker theorem an early version of what more recently Colbeck and Renner (2012b) aimed to show: that quantum mechanics is 'maximally informative'. Here though I am more interested in the implications of Theorem 7.5 for the MKC models and the classical representations based on them.

As mentioned at the beginning of this section, the MKC models are unsatisfactory candidates for a physical theory. One peculiarity about them is that the observables do not form a continuum. This is a big break with scientific theories of the past and present. In particular, there is a tension with the common conception of space being adequately modeled by $\mathbb{R}^3$. This in itself need not be a reason to be discontent with the models. A departure from this conception already occurs of course in general relativity, and in quantum gravity one may even move to a discrete conception of space. But it is problematic that the departure is ill-motivated. The set of operators $\mathcal{O}_{\mathrm{MKC}}$ used in section 7.4 is selected in an ad hoc way, and there is no unique way to construct it. Consequently, the departure from a continuum is not as simple and clear as going from $\mathbb{R}^3$ to $\mathbb{Q}^3$.[9] But what is most disturbing is that Theorem 7.5 shows that this departure causes a particular tension within the models themselves. The finite precision argument requires that observables appear to form a continuum by CoO. But this apparent continuity is nowhere coded into the theory: it is neither there at the level of the random variables $\mathcal{O}_{\mathrm{MKC}}$, nor at the level of the MKC states $\Lambda_{(\mathcal{P}_i)}$. Theorem 7.5 shows this is necessarily so.

This tension in the MKC models (i.e., requiring a continuity assumption while at the same time not being able to respect this assumption) is not just a moot point. The assumption CoO is not only used in the finite precision argument, but also in

---

[9]This is in contrast with the paper of Meyer (1999) in which spin-directions are associated with directions in $\mathbb{Q}^3$. Strictly speaking though, this does not give an MKC model, as the hidden variables in this case cannot be used to reproduce the quantum statistics (Cabello, 2002).

the derivation of the Born rule in Theorem 7.4. While this result is interesting, the fact that $\Lambda_{(\mathcal{P}_i)}$ does not respect CoO indicates that the theorem is not as important as it pretends. Difficulties arise, in particular, when one considers update rules for probability functions in the MKC models. Suppose one measures A and finds the result a. The usual way to update the probability function is by conditioning on the set of all λ with $F_A(\lambda) = a$. Now, because F_A is independent of $F_{A'}$ for every A' with $[A, A'] \neq 0$, this update has a negligible impact on the probability distribution. It only changes the probability distribution for one of the countably many frames that make up Λ_{MKC} (see (7.32)). This may be seen as a vindication of Appleby's objection that one can't learn anything about the hidden variables from a measurement. On top of that, one may note that the probability function one obtains by conditionalization is radically different from what one would get by using the projection postulate.

There are also some other problems with conditionalization as an update rule. For one, even if one started with a probability function that respects CoO (and thus is quantum like), after conditionalization it no longer respects CoO. Namely, for all the A' with $[A, A'] \neq 0$, the updated probability function still satisfies the trace rule, while for A the probability is now peaked around the value a. Another problem is that, due to the finite precision argument, one cannot know exactly for which A to perform the update. A naive solution could be to say that, since one knows roughly which A has been measured, and one knows that a repeated measurement of A would yield the same result (assuming λ wasn't influenced too much by the measurement), one should assign high probability to the set of all λ with $F_{A'}(\lambda) = a$ for all A' close to A. Effectively, this would mean applying (something very much resembling) the projection postulate. But although one could hope that this process saves the Born rule, the proposed update is completely unwarranted from the perspective of the MKC models. In general, the value of $\lambda(A')$ need not resemble $\lambda(A)$, and the existence of densely discontinuous regions indicates that changing the values of $\lambda(A')$ will often lead to false predictions.

In brief, the statistical predictions for the MKC models for consecutive measurements are problematic to define and possibly contradict quantum mechanical predictions. This is not to say it is impossible to define such measurements without contradictions. In (Hermens, 2011) I proposed the following artificial solution to the problem. Take the MKC states to specify the state of the system at any point in time, but let these states evolve stochastically in time. Specifically, at any instant the MKC state is selected in accordance with the Born rule, and the quantum state guiding this stochastic process evolves according to the dynamics of quantum mechanics. On this approach, the Born rule is applied as a postulate, rather than being derived from the assumption CoO. In essence, this MKC model is just orthodox quantum mechanics with some decorative fluff that is neither attractive nor explanatory. This works fine for showing the logical possibility of non-contextual

hidden variables. However, as a classical representation of quantum probability, the models recover the old difficulty of being incapable of explaining the special role of the probability functions that satisfy the Born rule.

7.6   Taking stock

Let me take the time to briefly recapitulate the results of part II before moving on. Two complementary programs have been running alongside each other. One concerns the possibility of understanding quantum probability from a classical perspective by means of a classical representation of a quantum probability space. The other concerns the possibility of understanding quantum mechanics from a hidden variable perspective. The main difficulty encountered is that neither is possible in such a way that all relevant aspects of the quantum are captured in a faithful way.

The Kochen-Specker theorem provides a compact description of the problem: classical representations or hidden variables are impossible when one holds on to both NC and IP. Giving up either of these two assumptions was shown to be sufficient to turn the impossibility into a possibility. A striking aspect of these possibilities is that for the constructed measurable space the probability functions that obey the Born rule form a special subset of all admissible probability functions. So while one of the important messages of section 4.3 was that classical probability is a special case of quantum probability, here the opposite message prevails: quantum probability is a special case of classical probability. These are not two contradictory messages, nor does their simultaneous truth establish that classical and quantum probability are one and the same thing. The fact that every classical probability space can be represented by a quantum probability space and vice versa rather may be seen to be analogous to the claim that every circle can be fitted within a square and vice versa. This claim is obviously true, but it is also clear that its truth does not imply that circles are squares.[10]

From the hidden variable perspective the situation is not entirely surprising. One of the reasons for introducing hidden variables in the first place was to correct for the incompleteness of the quantum formalism. The hidden variables describe putative states of affairs that aren't accounted for by the quantum state. Necessarily then, probability functions that are peaked around these states cannot be described

---

[10]It deserves to be noted that a similar situation is well-known within the literature on Bell inequalities (Cirel'son, 1980; Popescu and Rohrlich, 1994; Brunner et al., 2014). Although under certain conditions these inequalities cannot be violated within a classical representation, relieving these conditions allows for stronger violations than that can be obtained with quantum probabilities. The tension there is often framed by a contrast between signaling interactions and general non-local interaction. Such a contrast however cannot explain the general mismatch between quantum probability spaces and their classical representatives. After all, as seen in this part, this mismatch already occurs for single particle systems.

by a quantum probability function. However, one still has to explain why in practice these non-quantum probability functions play no role. This is best understood by way of an example. Consider again Bohmian mechanics. As noted earlier, this theory does not neatly conform to the formalism of a contextual hidden variable theory as illustrated in section 6.3. Instead, the hidden variables have a physical story to tell, and this story is helpful in explaining the Born rule: it is used to show that this rule arises as an equilibrium distribution. Importantly, then, the Born rule is not fundamental in Bohmian mechanics. Other probability functions are not prohibited, but are just 'atypical', and this even allows, in principle, for empirical discrepancies with quantum mechanics (Valentini, 2010).

When it comes to understanding the formalism of quantum probability on its own, the hidden variable perspective doesn't suffice. Putting it somewhat boldly, it 'explains' the formalism of quantum probability by stating that it is essentially incorrect, or at least incomplete. What I am hinting at is an important distinction between the quantum representation of classical probability and the classical representation of quantum probability. In the first case there is a criterion that singles out the classical from the quantum: classical probability corresponds, roughly speaking, to commutative probability. But conversely it isn't easy to understand from a classical perspective, either formally or conceptually, what the special role of the quantum probability functions is. Without such an understanding it isn't entirely satisfactory to speak of a faithful representation. This is not to say that faithful representations are an impossibility. But a more direct investigation of quantum probability may be more worth our time, and this is what I do in part III of this dissertation.

Worries about the possibility of deriving the Born rule within hidden variable theories seldom play a role in discussions on the pros and cons of hidden variables. As noted above, there are good reasons for this: the Born rule may just be a contingent rule. More common are the objections that "the introduction of hidden variables does not lead to new experimentally testable predictions", or that "hidden variable theories cannot be generalized to incorporate quantum field theories". But the most traditional objection is of course that "hidden variable theories are necessarily non-local". Correspondingly, there are also generic replies to these objections. The first objection misses the point: the aim of hidden variables is to solve the measurement problem, or, more modestly, to simply give an interpretation of quantum mechanics, not to necessarily lead to new physics. The second is an undecided issue, and in the third one may be applying a double standard: it is not evident that a rejection of hidden variables automatically gives you a local interpretation of quantum mechanics.

It is not my aim here to take sides by scrutinizing these objections and replies. The main point is rather that there are more stringent objections against hidden variables than those given by the Kochen-Specker theorem. However, when it comes

to the more formal questions of classical representations of quantum probability, the Kochen-Specker theorem already provides enough restrictions to be pessimistic. The general aspect of all ways to circumvent the theorem, is that the formalism of quantum probability is no longer fundamental. Now while this is of course a logical possibility, its denial is a logical possibility as well. In that case, letting go of classical probability, and instead focusing directly on quantum probability, is the road towards a physical understanding of quantum probability. This then is the project in the next part.

Part III

Quantum logic and quantum probability

<div align="right">*8*</div>

Orthodox quantum logic

8.1 Introduction

The central theme of this dissertation is the investigation of quantum probability in order to obtain a better understanding of its mathematical formalism. Thus far, the investigation focused on a comparison with classical probability. In chapter 4 a first purely mathematical comparison of the two formalisms was made. In part II the comparison focused on the (im)possibility of embedding quantum probability within the classical framework. Both these investigations led to useful insights. In chapter 4 the view emerged that classical probability can be viewed as a special case of quantum probability. In part II, on the other hand, the opposite view emerged. For any classical representation of quantum probability it was found that the quantum probability functions form a strict subset of the classical probability functions.[1]

These two opposing views are complementary rather than contradictory: they highlight different aspects of quantum probability. On the first view classical probability identifies with the commutative side of quantum probability. The richer structure of non-commutative probability is an essential ingredient of the advantages of quantum computation over classical computation (Nielsen and Chuang, 2010). On the second view the quantum probability functions can be understood as incomplete specifications of the actual (hidden variable) state. The total set of probability functions then also contains complete specifications of the states.

In this part of the dissertation I undertake a more direct investigation of quantum probability. The aim is to obtain a reformulation of the formalism in which the mathematical symbols have a clear physical interpretation. In the remainder of this section I elaborate on what this aim is, but two remarks are already in order. First, the aim presupposes that in the current formulation of quantum probability a clear physical interpretation is absent. This is a claim I defend in this chapter. Second, in the investigation I adopt an empiricist stance. The aim is to obtain a

---

[1]Theorem 7.4 is a peculiar exception. However, as discussed, it is not without its own problems.

formulation of quantum probability that connects smoothly with the way quantum probability is used to formulate empirical statements. Little or no assumptions are made concerning the correct metaphysics for quantum probability, and in particular no attempt is made to provide a solution to the measurement problem.

The program in this part of the dissertation is somewhat unorthodox, and it is therefore useful to separate it from a program that looks much like it. The summary above sketches a picture of quantum probability spaces lying in between two classical probability spaces. A commutative subspace of the quantum probability space on the inside, and a hidden variable space enveloping the quantum space on the outside. A similar picture is well-known in the foundations of quantum mechanics. Although quantum mechanics violates Bell inequalities, and thus contains non-local correlations, these correlations are found to be non-maximal: they satisfy the so-called Tsirelson bound (Cirel'son, 1980), which may be violated by non-local hidden variable theories. Quantum mechanics thus occupies a special place within the space of all possible theories. The quest to understand this particular space was triggered by Wheeler's famous question: "How come the quantum?" (1990). In this program, the aim is to reconstruct the formalism of quantum mechanics from physical principles.

In this dissertation I aim for something more modest and specialized. Part of the investigation of quantum probability is to try to reconstruct its formalism. But in this reconstruction I do not shy from picking (non-probabilistic) ingredients from quantum mechanics. Thus while in the reconstruction of quantum mechanics one may seek an underpinning for the use of Hilbert space theory, I take this mathematical structure for granted. Then, given this structure, the aim is to provide a conceptual reading/reformulation of this structure on which one can introduce a natural formalism of probability.

To understand the strategy, it is useful to make a final comparison with classical probability. Consider again the definitions of classical and quantum probability spaces alongside each other.

Definition 8.1.

A *classical probability space* is a triplet $(\Omega, \mathcal{F}, \mathbb{P})$ such that
1. Ω is a set,
2. $\mathcal{F}$ is a σ-algebra of subsets of Ω containing Ω,
3. $\mathbb{P}$ is a function from $\mathcal{F}$ to the interval $[0, 1]$,
4. $\mathbb{P}(\Omega) = 1$,
5. $\mathbb{P}(\bigcup_n \Delta_n) = \sum_n \mathbb{P}(\Delta_n)$ for any countable sequence of mutually disjoint sets.

A *quantum probability space* is a triplet $(\mathcal{H}, L(\mathcal{H}), \mathbb{P})$ such that
1. $\mathcal{H}$ is a Hilbert space,
2. $L(\mathcal{H})$ is the collection of all projection operators on $\mathcal{H}$,
3. $\mathbb{P}$ is a function from $L(\mathcal{H})$ to the interval $[0, 1]$,
4. $\mathbb{P}(\mathbb{1}) = 1$,
5. $\mathbb{P}(\sum_n P_n) = \sum_n \mathbb{P}(P_n)$ for any countable sequence of mutually orthogonal projections.

The most striking difference is in the first two points, which mark the replacement of the familiar set theory with the Hilbert space formalism. This suggests that once one can determine why it makes sense to adopt a Hilbert space structure for the appropriate domain of probability functions,[2] one is already close to an understanding of the whole of quantum probability. That is, given a conceptual underpinning of 1 and 2, it is reasonable to assume that 3, 4 and 5 can be motivated by appealing to an analogy with classical probability. Thus explaining the role of $L(\mathcal{H})$ really can be considered the central task.

To get a feeling for what one may be looking for in an explanation of 1 and 2, it is useful to make a comparison with the classical case. Here, the adoption of 1 and 2 is so 'trivial' that it becomes hard to find a direct motivation for it in modern literature. The use of set theory is deeply entrenched in our mathematical thinking, and so one has to search for answers to the question of why a set-theoretical structure is the appropriate domain of probability functions. In most of the math textbooks on probability, the sets in a σ-algebra are identified with 'events', usually taken to be a primitive term. This is not surprising since, pragmatically, it is easy to work with this notion. For example, considering the roll of a die, the set of possible outcomes is $\{1, 2, 3, 4, 5, 6\}$. The set $\{2, 4, 6\}$ can then be identified with the event of rolling an even number. Thus it is plausible to model the probability of events with the values assigned to the corresponding sets by a function $\mathbb{P}$.

When accepting the notion of 'event' as a primitive concept, a similar understanding can be given for the use of projection operators. This link is given by the Born postulate. The event that the measurement of an observable $\mathcal{A}$ yields an outcome in the set Δ is identified with the projection operator $\mu_A(\Delta)$ (with A the operator associated with $\mathcal{A}$). This primitive use of the notion of 'event' is quite common and, pragmatically, it is of course as natural as it is in the classical case. But more care is required when this notion starts playing a role in philosophical discussions of quantum probability. A good example of such use is in the information-theoretical interpretation of Bub and Pitowsky:

> On this interpretation, the structure of Hilbert space, i.e., the non-Boolean algebra of Hilbert space subspaces, defines the structure of a quantum event space, just as, classically, a Boolean algebra, the subsets of a set (phase space), defines the structure of a classical event space. Gleason's theorem then determines all possible probability measures on this structure as given by quantum states (pure and mixed) according to the trace rule, where the probabilities are interpreted as degrees of

---

[2]To contrast with the program of reconstructing quantum mechanics: It is not required here that the use of Hilbert space theory comes out as the necessarily correct formalism. It suffices if it can be shown to be a possibly correct formalism.

belief or measures of uncertainty about events in the Bayesian sense.
(Bub, 2007, p. 239)

Here, the quantum event structure is taken for granted and the interpretation of quantum probability is a story about this formalism, rather than something that provides an explanation of the formalism. This is problematic because it is not clear whether the notion of 'event' can be unpacked in such a way that it is compatible with the given interpretation. In fact, there is even cause for serious doubt, as there are good reasons to believe that events should be modeled by sets rather than projection operators. An interpretation of quantum probability should thus come equipped with tools to counter these reasons.

Note that the problem is not specific to the adoption of a Bayesian interpretation. The same worry also applies for example to the consistent histories approach of Griffiths (2013), in which the quantum event structure plays a similar primitive role, but where probabilities are understood as objective chances. But for the sake of definiteness it is useful to unpack this problem further with an eye on Bayesian interpretations. According to the Bayesian, probabilities are degrees of belief of a rational agent. The usual elaboration goes as follows. The formula $\mathbb{P}(A) = x$ is taken to express the idea that a rational agent believes with degree x that the proposition (represented by) A is true. It is relatively straightforward to relate events to propositions. The probability of the event that the roll of a die yields an even number can be understood as the degree of belief that the proposition "the roll of the die yields an even number" is true. Thus any motivation for a particular mathematical modeling of propositions also provides a motivation for using the same model for events.

In philosophy and classical logic, propositions are regularly associated with sets. This idea at least dates back to the works of Boole and Venn in the 19th century. The standard way to make the identification with sets is to construct the Lindenbaum-Tarski algebra from some formal language L. If L obeys classical logic this algebra is Boolean. Stone's representation theorem then shows that this algebra can be identified with an algebra of sets. In short, if one accepts that ultimately probabilities are the kind of things to be assigned to propositions, and that classical logic is correct, then one ends up with a Boolean algebra of sets as the domain of probability functions.[3] Admittedly, this does not immediately give one the classical probability framework. The algebras obtained need not be a σ-algebras, and the sets come equipped with a special topology (i.e., the Stone topology). However, these are just details in comparison to the mismatch with the quantum framework.

---

[3]It seems then that Popper (1959, p. 320) was somewhat reckless when criticizing Kolmogorov's framework as being "less 'formal'" than his own because "He [Kolmogorov] interprets the arguments of the probability functor as *sets*" (emphasis in original) while priding himself for being completely silent about what kind of mathematical objects are in the domain of his own probability functions.

One therefore cannot just replace the classical event structure with a non-Boolean structure and keep the classical interpretations as if nothing has changed. This of course does not imply that $L(\mathcal{H})$ cannot play the role of an event-set. But it makes any approach that takes this role as a separate fundamental assumption suspect. The aim then of this part of this dissertation is to gain a better understanding of $L(\mathcal{H})$. In other words, the focus is on quantum logic.[4]

I start in this chapter and the next one with a direct approach. That is, I take the set $L(\mathcal{H})$ as given and investigate the possibility to view this set as an algebraic structure of propositions. The obvious starting point is a discussion of the orthodox quantum logic of Birkhoff and von Neumann (1936), which is given in section 8.2. For reasons discussed in section 8.3, this logic is found to be unsatisfactory for providing a direct conceptual foundation for quantum probability. Consequently, one has to make a departure from $L(\mathcal{H})$, possibly by expanding it into a larger set of propositions. The new logic of quantum propositions is then supposed to better fit with the correct way of reasoning about such propositions. In section 8.4 I motivate why an appeal to intuitionistic logic can be useful in the search for such an expansion of $L(\mathcal{H})$. This motivation bases itself on a parallel between an empiricist stance in science and a constructivist stance in mathematics. Chapter 9 is devoted to the study of a particular kind of expansion of $L(\mathcal{H})$ that builds on the work of Coecke (2002). Although the algebraic structure obtained (a weak Heyting algebra) provides some useful insights, it remains conceptually vague due to its mostly mathematically motivated construction.

In chapter 10 the strategy is turned around by adopting a more conceptual approach. That is, I start with a minimal set of propositions about quantum measurements, and then expand to obtain an algebraic structure that adheres to empirical constraints posed by quantum mechanics. This first leads to the construction of the Heyting algebra that has also been derived by Caspers et al. (2009) within a topos-theoretical approach, but that is now given a conceptual derivation rather than a mathematical one. The upshot is that the Heyting algebra automatically has a clear interpretation. The algebra is then further expanded so as to obtain a Boolean algebra of empirical quantum propositions. That is, an empiricist quantum logic is derived that is completely classical. Chapter 11 is devoted to investigating probability functions on this new algebraic structure. It is shown that every quantum probability function can be identified with a conditional probability function on the Boolean algebra. Thus an empiricist reformulation of quantum probability is obtained.

---

[4]I adopt a relaxed attitude towards the meaning of the term 'logic'. For me, a logic can be any algebraic structure that represents some aspect of a set of propositions as well as the way they hang together.

8.2   Orthodox quantum logic

In this section the formalism of orthodox quantum logic (OQL) is presented. Usually, this formalism is introduced by making an analogy with the way classical logic derives from the state space formalism of classical mechanics (Isham, 1995). Propositions are then identified with sets of states, and logical connectives are introduced by suggesting manipulations of these sets. The downside of this approach for the present discussion is that it starts from the idea that projections can be seen as propositions, which is the very idea that is being questioned here. Instead, I provide a short derivation of OQL from the formalism of quantum probability. The upshot of this approach is that OQL necessarily presents itself to us as part of the quantum formalism, in that there are no real philosophical choices involved in the adoption of OQL. Rather, the philosophical work lies in providing an understanding of the role of this structure. Minimally, it is just a mathematical oddity that has no conceptual role to play, and maximally it points us towards a reinvestigation of what kind of reasoning is correct in a world in which quantum mechanics is true. I take it that the truth lies somewhere in between these two extremes. But before one can investigate where precisely, one has to know what OQL is.

To understand how one can derive quantum logic from quantum probability it is useful to consider a classical analogy. In the classical case the probability function is a map $\mathbb{P} : \mathcal{F} \to [0,1]$, where $\mathcal{F}$ is some σ-algebra. The structural properties of $\mathbb{P}$ are reflections of the the structure of the σ-algebra $\mathcal{F}$. Specifically, as a set of subsets, $\mathcal{F}$ has the natural partial order

$$\Delta_1 \leq \Delta_2 \text{ iff } \Delta_1 \subset \Delta_2. \tag{8.1}$$

This partial order also has another possible characterization.

Proposition 8.1. *For a measurable space $(\Omega, \mathcal{F})$, for every Δ_1, Δ_2, $\Delta_1 \leq \Delta_2$ if and only if $\mathbb{P}(\Delta_1) \leq \mathbb{P}(\Delta_2)$ for every probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$.*

Proof. Suppose $\Delta_1 \leq \Delta_2$. Then for every probability measure $\mathbb{P}$ it holds that

$$\mathbb{P}(\Delta_1) = \mathbb{P}(\Delta_1 \cap \Delta_2) \leq \mathbb{P}(\Delta_1 \cap \Delta_2) + \mathbb{P}(\Delta_1^c \cap \Delta_2) = \mathbb{P}(\Delta_2). \tag{8.2}$$

Conversely, suppose Δ_1 is not a subset of Δ_2. Then $\Delta_3 := \Delta_1 \cap \Delta_2^c$ is not empty. Now choose $\mathbb{P}$ such that $\mathbb{P}(\Delta_3) = 1$. Then $\mathbb{P}(\Delta_1) = 1 \nleq 0 = \mathbb{P}(\Delta_2)$, which is a contradiction. $\qquad\square$

The partial order on $\mathcal{F}$ may thus also be understood as a structure imposed on it by the set of all probability functions. In an analogous way a relation may be introduced on $L(\mathcal{H})$ that turns it into a partial ordered set.

Theorem 8.1. *The relation*

$$P_1 \leq P_2 \ iff \ \forall \mathbb{P} : \mathbb{P}(P_1) \leq \mathbb{P}(P_2) \tag{8.3}$$

is a partial order on $L(\mathcal{H})$.

Proof. From the properties of probability functions it follows directly that this relation is a preorder (i.e., reflexive and transitive). With a little more effort it can be seen that it is also antisymmetric. What has to be shown for this is that if $\mathbb{P}(P_1) = \mathbb{P}(P_2)$ for all $\mathbb{P}$, then $P_1 = P_2$. The proof relies on the following three identities.

$$PP_{P\psi} = P_{P\psi}, \ PP_{P^\perp\psi} = \mathbb{0} \ \text{and} \ \|P\psi\| = \|\psi\|\sqrt{\text{Tr}(PP_\psi)}$$

for every projection operator P and every vector ψ, where

$$P^\perp := \mathbb{1} - P \tag{8.4}$$

denotes the orthocomplement of P, and $P_{P\psi}$ is the projection onto the line spanned by $P\psi$, i.e.,

$$P_{P\psi}\phi := \frac{\langle P\psi, \phi\rangle}{\langle P\psi, P\psi\rangle}P\psi. \tag{8.5}$$

Now suppose for two projections P_1, P_2 that $\mathbb{P}(P_1) = \mathbb{P}(P_2)$ for all $\mathbb{P}$, then

$$
\begin{aligned}
\|P_2\psi - P_1\psi\| &= \|(P_2 - P_1)P_1\psi + (P_2 - P_1)P_1^\perp\psi\| \\
&\leq \|(P_2 - P_1)P_1\psi\| + \|(P_2 - P_1)P_1^\perp\psi\| = \|P_2^\perp P_1\psi\| + \|P_2 P_1^\perp\psi\| \\
&= \|P_1\psi\|\sqrt{\text{Tr}(P_2^\perp P_{P_1\psi})} + \|P_1^\perp\psi\|\sqrt{\text{Tr}(P_2 P_{P_1^\perp\psi})} \\
&= \|P_1\psi\|\sqrt{\text{Tr}(P_1^\perp P_{P_1\psi})} + \|P_1^\perp\psi\|\sqrt{\text{Tr}(P_1 P_{P_1^\perp\psi})} = 0,
\end{aligned}
$$

where in the last line it was used that $P \mapsto \text{Tr}(PP_{P_1\psi})$ and $P \mapsto \text{Tr}(PP_{P_1^\perp\psi})$ are probability functions for all P_1 and ψ and that $\mathbb{P}(P_1) = \mathbb{P}(P_2)$ if and only if $\mathbb{P}(P_1^\perp) = \mathbb{P}(P_2^\perp)$. $\qquad\square$

This theorem establishes that quantum probability functions introduce a partial order on $L(\mathcal{H})$ analogous to the way classical probability functions introduce a partial order on $\mathcal{F}$. The partial order on $\mathcal{F}$ is the one associated with the usual lattice structure: the join of two sets corresponds to the union, and their meet with the intersection. Note that, given a partially ordered set, the meet and join of two elements (if they exist) are unique. Thus union and intersection are operations that derive from the partial order of set-inclusion. Similarly, the meet and join in the orthodox quantum logic of Birkhoff and von Neumann (1936) (see also (von

117

Neumann, 1932, §III.5)) derive from the partial order (8.3) on $L(\mathcal{H})$. To see this it is required to show that indeed every pair of projections has a meet and a join with respect to the partial order (8.3). This, in turn, is best shown by proving that the partial order is equivalent to the usual one.

Proposition 8.2. *Let $(L(\mathcal{H}), \leq)$ be the poset with $\leq$ given by (8.3), then*

$$P_1 \leq P_2 \text{ iff } P_1\mathcal{H} \subset P_2\mathcal{H} \text{ iff } P_1P_2 = P_2P_1 = P_1. \tag{8.6}$$

Proof. Start with the first left-to-right implication. Suppose $P_1 \leq P_2$ and let $\psi \in P_1\mathcal{H}$. By making use of the fact that $P \mapsto \text{Tr}(PP_\psi)$ is a probability function, it follows that

$$\begin{aligned}
\|\psi - P_2\psi\|^2 &= \|P_2^\perp\psi\|^2 = \|\psi\|^2 \, \text{Tr}(P_2^\perp P_\psi) \\
&= \|\psi\|^2 \left(1 - \text{Tr}(P_2 P_\psi)\right) \leq \|\psi\|^2 \left(1 - \text{Tr}(P_1 P_\psi)\right) = 0.
\end{aligned} \tag{8.7}$$

Thus $P_2\psi = \psi$ and $\psi \in P_2\mathcal{H}$.

Now consider the second left-to-right implication. Because P_2 is idempotent, it follows directly from $P_1\mathcal{H} \subset P_2\mathcal{H}$ that $P_2P_1 = P_1$. The second equation is obtained by taking adjoints:

$$P_1P_2 = ((P_1P_2)^*)^* = (P_2^* P_1^*)^* = (P_2P_1)^* = P_1^* = P_1. \tag{8.8}$$

The last step closes the loop. Suppose $P_1P_2 = P_2P_1 = P_1$, and define $P_3 := P_2 - P_1$. Then P_3 is a projection operator:

$$\begin{aligned}
P_3^* &= (P_2 - P_1)^* = P_2^* - P_1^* = P_2 - P_1 = P_3, \\
P_3^2 &= (P_2 - P_1)(P_2 - P_1) = P_2 - P_1P_2 - P_2P_1 + P_1 = P_2 - P_1 = P_3.
\end{aligned} \tag{8.9}$$

It also follows directly that P_3 and P_1 are orthogonal. Finally then, for every probability function $\mathbb{P}$, one has

$$\mathbb{P}(P_1) \leq \mathbb{P}(P_1) + \mathbb{P}(P_3) = \mathbb{P}(P_1 + P_3) = \mathbb{P}(P_2). \tag{8.10}$$

$$\square$$

Proposition 8.2 shows that the partial order on $L(\mathcal{H})$ can be identified with the convenient partial order of set inclusion on the set of subsets of $\mathcal{H}$ of the form $P\mathcal{H}$ for $P \in L(\mathcal{H})$, i.e., there is an isomorphism $(L(\mathcal{H}), \leq) \simeq (L'(\mathcal{H}), \subset)$ with

$$L'(\mathcal{H}) := \{P\mathcal{H} \mid P \in L(\mathcal{H})\}. \tag{8.11}$$

The upshot is that meets and joins of projection operators can be understood as operations on sets provided one restricts attention to the closed linear subspaces.

So $P_1 \wedge P_2$ identifies with the biggest subspace $P\mathcal{H}$ that satisfies $P\mathcal{H} \subset P_1\mathcal{H} \cap P_2\mathcal{H}$, and $P_1 \vee P_2$ with the smallest subspace $P\mathcal{H}$ that satisfies $P_1\mathcal{H} \cup P_2\mathcal{H} \subset P\mathcal{H}$. What is needed now is a way to characterize sets of the form $P\mathcal{H}$ that helps to see that these meets and joins indeed exist. This characterization is given by Theorem A.2 which states that a subset of $\mathcal{H}$ is of the form $P\mathcal{H}$ for some projection P if and only if it is a closed linear subspace.

Showing the existence of meets and joins for closed linear subspaces is relatively straightforward. For any two closed linear subspaces $\mathcal{K}_1, \mathcal{K}_2$, the intersection is again a closed linear subspace. This is also the largest linear subspace that is a subspace of both $\mathcal{K}_1$ and $\mathcal{K}_2$. Thus

$$\mathcal{K}_1 \wedge \mathcal{K}_2 = \mathcal{K}_1 \cap \mathcal{K}_2. \tag{8.12}$$

Similarly, for any pair of closed linear subspaces $\mathcal{K}_1, \mathcal{K}_2$ one can take the linear span of all the vectors in $\mathcal{K}_1 \cup \mathcal{K}_2$. Although this gives a linear subspace, in general it will not be closed. This problem is solved by taking the closure of the set.[5] Thus

$$\mathcal{K}_1 \vee \mathcal{K}_2 = \overline{\text{span}(\mathcal{K}_1 \cup \mathcal{K}_2)}. \tag{8.13}$$

Since $L(\mathcal{H})$ as a partially ordered set is isomorphic to the set of closed linear subspaces with set inclusion, lattice operations on $L(\mathcal{H})$ can be completely understood in terms of the lattice operations on closed linear subspaces. Specifically, if $P_\mathcal{K}$ denotes the projection such that $P_\mathcal{K}\mathcal{H} = \mathcal{K}$, then

$$P_1 \vee P_2 = P_{P_1\mathcal{H} \vee P_2\mathcal{H}}, \; P_1 \wedge P_2 = P_{P_1\mathcal{H} \wedge P_2\mathcal{H}}, \tag{8.14}$$

where the join and meet for closed linear subspaces are given by (8.13) and (8.12). Although this shows the existence of joins and meets of projection operators, it does not give a very insightful formulation of these operations. In some cases, a clearer formulation can be given purely in terms of projection operators, as seen in the following example.

Example 8.1. Suppose P_1 and P_2 are two compatible projection operators. Then

$$P_1 \wedge P_2 = P_1 P_2,$$
$$P_1 \vee P_2 = P_1 + P_2 - P_1 P_2. \tag{8.15}$$

$\triangle$

Thus far I have shown that quantum probability gives rise to a partial order on $L(\mathcal{H})$, which in turn introduces a join and meet that turn it into a lattice. In

---

[5]If $\mathcal{K}$ is a linear subspace, then its closure $\overline{\mathcal{K}}$ consists of all the limits of all converging sequences in $\mathcal{K}$.

the classical case the lattice structure on $\mathcal{F}$ gives rise to a unique complement. On $(L(\mathcal{H}), \leq)$, however, complements are not unique. For example, in $\mathbb{C}^2$ every pair of unequal one-dimensional projection operators are complements of each other. This is because any two vectors ψ_1, ψ_2 that aren't on the same line, together span the whole Hilbert space ($P_{\psi_1} \vee P_{\psi_2} = \mathbb{1}$). This non-uniqueness already implies that $(L(\mathcal{H}), \leq)$ is not Boolean. But despite this non-uniqueness, there is still a particular complement singled out by quantum probability.

Theorem 8.2. *For every $P \in L(\mathcal{H})$ the orthocomplement $P^\perp$ is the unique complement that satisfies*

$$\mathbb{P}(P) + \mathbb{P}(P^\perp) = 1 \tag{8.16}$$

for all quantum probability functions $\mathbb{P}$.

Proof. It is easy to check that the orthocomplement is a complement and satisfies (8.16), so the task is to prove uniqueness. Suppose $P \in L(\mathcal{H})$ and P^c is a complement that satisfies (8.16), then it has to be shown that $P^c = P^\perp$. If $\psi \in P\mathcal{H}$, then

$$\begin{aligned}
\|P^c\psi\|^2 &= \langle P^c\psi, P^c\psi \rangle = \langle \psi, P^c\psi \rangle \\
&= \text{Tr}(P_\psi P^c) = \mathbb{P}_\psi(P^c) = 1 - \mathbb{P}_\psi(P) = 0.
\end{aligned} \tag{8.17}$$

If $\psi \in P^\perp\mathcal{H}$, then

$$\begin{aligned}
\|\psi - P^c\psi\|^2 &= \|(\mathbb{1} - P^c)\psi\|^2 = \langle (\mathbb{1} - P^c)\psi, (\mathbb{1} - P^c)\psi \rangle = \langle \psi, (\mathbb{1} - P^c)\psi \rangle \\
&= \text{Tr}(P_\psi(\mathbb{1} - P^c)) = \text{Tr}(P_\psi) - \text{Tr}(P_\psi P^c) = 1 - (1 - \mathbb{P}_\psi(P)) = 0.
\end{aligned} \tag{8.18}$$

Thus for arbitrary $\psi \in \mathcal{H}$:

$$P^c\psi = P^c P\psi + P^c P^\perp\psi = 0 + P^\perp\psi. \tag{8.19}$$

$\square$

This theorem provides the final step to conclude that the formalism of quantum probability alone has all the ingredients to turn $L(\mathcal{H})$ in a complemented lattice. To conclude this section, let me state explicitly what is meant by orthodox quantum logic in the remainder of this dissertation.

Definition 8.2. The *orthodox quantum logic* for a system described by a Hilbert space $\mathcal{H}$ is the set $L(\mathcal{H})$ of projection operators understood as a complemented lattice $(L(\mathcal{H}), \leq, \wedge, \vee, \perp)$ with the lattice structure given by (8.3) and (8.4).

8.3   The interpretation of orthodox quantum logic

In the previous section it was shown that orthodox quantum logic is part of the package deal that is quantum probability. Also, in the introduction of this chapter it was argued that any explanation of why $L(\mathcal{H})$ would be an appropriate domain for probability functions, is already halfway a conceptual foundation for quantum probability. The tempting idea then is to put two and two together, and aim for a physical interpretation of orthodox quantum logic as the basis for a conceptual foundation for quantum probability. The aim of this section is to argue that a direct physical interpretation of OQL cannot provide this basis.

There has been a vast amount of research in quantum logic over the past century, and any evaluation is necessarily selective and incomplete. Most of this research though isn't tailored for the question at hand. Often it is purely mathematical or, when a connection with probability is made, a primitive notion of probability is presupposed. In other cases one usually adopts a primitive concept of events or properties when thinking of the elements of $L(\mathcal{H})$. That is of course fine for practical purposes, but when one aims to interpret quantum probabilities without explaining what the elements of $L(\mathcal{H})$ are, one just replaces one mystery with another.

Probably the most radical (and possibly most familiar among philosophers) attempt at interpreting orthodox quantum logic stems from Putnam (1968). In his proposal the elements of $L(\mathcal{H})$ are understood as expressing possible properties that a quantum system can have, or, more accurately, as propositions expressing that the system has a certain property. The projection operator $\mu_A(\Delta)$ is taken to represent the proposition expressed by "the observable A has a value in Δ", usually shortened with the notation $A \in \Delta$. The meet and join of the lattice are further understood as expressing conjunctions and disjunctions of these propositions. On this reading of quantum logic, particles always have a position and a momentum, as both $X \in \mathbb{R}^3$ and $P \in \mathbb{R}^3$ identify with the top-element of $L(\mathcal{H})$, which is now understood as expressing a tautology. More generally, every observable has a definite value by the same argument.

There are several conceptual problems with Putnam's account of quantum logic. A very thorough debunking has been given by Dummett (1976).[6] Instead of discussing these conceptual problems I just want to sketch one technical problem for the view adopted by Putnam (see also (Friedman and Glymour, 1972)). When it comes to properties, it is common to think that any system either has them or doesn't. If this is accepted, then there exists a function $f : L(\mathcal{H}) \to \{0, 1\}$ marking

---

[6]Much of the discussion has focused on whether in principle scientific discoveries can lead to a revision of logic, and what such a revision could possibly mean (see also (Bacciagaluppi, 2009)). Maudlin (2005) gives an overview of this historical episode marking the birth and death of quantum logic in this form. But see (Griffiths, 2014) for an attempt to reanimate the view with slight modifications.

for every property whether the system has it (value 1), or doesn't (value 0). Further assuming that the meets and joins on $L(\mathcal{H})$ express conjunctions and disjunctions, requires this function to be a lattice homomorphism. But it is a corollary of the Kochen-Specker theorem that (if $\dim(\mathcal{H}) > 2$) no such function exists. So one of the assumptions has to go. Either there aren't any truth valuations, or meets and joins do not (always) express conjunctions and disjunctions. Putnam hoped to keep both assumptions in place. It is not surprising, then, that he later rejected his own proposal (Putnam, 1994 ; 2005).

Putnam aimed for much more than what I'm after here. The identification of elements of $L(\mathcal{H})$ with possible properties of a system puts forward a metaphysical picture. If the picture were consistent this could point to a solution of the measurement problem when one thinks of measurements as revealing properties of a system. The quantum probabilities could then just be probabilities about actually possessed properties. But this approach (if it were to work) presupposes too much metaphysics and leads towards a narrow conception of quantum probability. I do not aim to solve the measurement problem though, and neither do I aim to provide an ontology for quantum mechanics.

A broader conception of quantum logic would be obtained by adopting an empiricist approach. In such an approach it is more fitting to try to think of the elements of $L(\mathcal{H})$ as experimental propositions: propositions about measurements and their outcomes of the kind one can ascribe probabilities to. This is also closer to the approach of Birkhoff and von Neumann (1936) themselves.[7] Unsurprisingly, their approach is more mathematical than philosophical, but it is nevertheless a good starting point for philosophical discussion. In good mathematical fashion, one of the first things one finds in their paper is a definition of experimental propositions.

Definition 8.3. For a set $\{A_1, \ldots, A_n\}$ of commuting observables the *observation space* is given by the Cartesian product of their spectra

$$\mathfrak{O}(A_1, \ldots, A_n) := \sigma(A_1) \times \cdots \times \sigma(A_n). \tag{8.20}$$

A subset[8] $\Delta \subset \mathfrak{O}(A_1, \ldots, A_n)$ is called an *experimental proposition* concerning $\mathfrak{O}(A_1, \ldots, A_n)$.

The notion of an observation space can easily be adjusted to fit other theories of physics. An experimental proposition is nothing more than a set of possible

---

[7]Although von Neumann (1932) smoothly intertwines use of the words "property", "proposition" and "event", indicating that he may not have cared much about their distinctions in meaning, in his collaboration with Birkhoff (1936) he was conspicuously more careful and consistent.

[8]In general, this will be a Borel set. In their paper, Birkhoff and von Neumann (ibid.) only consider operators with discrete spectra, so all their experimental propositions are automatically measurable sets. For the sake of simplicity and continuity I adopt the same assumption in this section.

measurement outcomes. There is nothing typically quantum about that. But an important distinction noted by Birkhoff and von Neumann is that, in classical theories, the connection between experimental propositions and the phase space (set of pure states) is direct. In classical mechanics the phase space itself (i.e., the set of all possible position and momentum configurations) is itself an observation space. In quantum mechanics the connection between observation spaces and phase spaces has to be made separately. This is done with another definition.

Definition 8.4. The *mathematical representative* of an experimental proposition $\Delta \subset \mathfrak{O}(A_1, \ldots, A_n)$ is the set of vectors ψ in the Hilbert space for which there exists an element $(a_1, \ldots, a_n) \in \Delta$ such that

$$A_1\psi = a_1\psi, \ldots, A_n\psi = a_n\psi. \tag{8.21}$$

For an observation space built from a single observable A and an experimental proposition $\Delta \subset \sigma(A)$, the mathematical representative is just the subspace $\mu_A(\Delta)\mathcal{H}$. More generally, it is the subspace[9]

$$\left(\bigvee_{a \in \Delta} \bigwedge_{i=1}^{n} \mu_{A_i}(a_i) \right) \mathcal{H}. \tag{8.22}$$

By the connection between projections and closed linear subspaces, every observation space selects a proper subspace of $L(\mathcal{H})$ for its mathematical representatives. When ranging over all possible observation spaces, eventually every element of $L(\mathcal{H})$ will be a mathematical representative of some experimental proposition. Thus $L(\mathcal{H})$ exhaustively provides the mathematical representatives of all experimental propositions.

Even if on rejects Putnam's idea, one may still hope that orthodox quantum logic provides a logic of experimental propositions in which the lattice operations can be understood as logical connectives. In such an approach the elements of $L(\mathcal{H})$ need not represent properties, but instead refer to experimental propositions.[10] The meets and joins in $L(\mathcal{H})$ could then represent conjunctions and disjunctions for these properties. One may distinguish two assumptions that would establish such a view.

OQL1 The mathematical representative of an experimental proposition captures all relevant aspects of this proposition.

OQL2 Disjunctions and conjunctions of experimental propositions are once again experimental propositions.

---

[9]I adopt here the short-hand notation $\mu(a) = \mu(\{a\})$ for singleton sets.

[10]Or one could go one step further and have the elements refer to dispositional properties that are only instantiated (assigned a truth value) in the event of a measurement.

The first assumption expresses the idea that, when going from the observation spaces $\mathfrak{O}(A_1, \ldots, A_n)$ to the set of mathematical representatives $L(\mathcal{H})$, nothing essential is "lost in translation". That is, it is meaningful to think of projection operators themselves as expressing experimental propositions. The second assumption establishes that the disjunction (or conjunction) of two experimental propositions can again be represented by an element of $L(\mathcal{H})$. Because the disjunction of two experimental propositions is the strongest proposition that is weaker than both of the disjuncts, the only available element in $L(\mathcal{H})$ to be associated with it is the join of the mathematical representatives of the disjuncts. I shall now argue that, given OQL1, the meet and join cannot both be identified with conjunction and disjunction respectively. This, then, will be taken to motivate that, to understand $L(\mathcal{H})$, one has to first go beyond orthodox quantum logic, which is done in the next chapter. This should not be taken as a defense of the idea that OQL1 is unproblematic. In fact, in chapter 10 OQL1 is no longer assumed to hold.

Although experimental propositions have been given a sound mathematical representative, physically, it is not entirely clear how these representatives should match to propositions. A demand for a clear physical interpretation may be seen to follow from Bohr's doctrine that "all well-defined experimental evidence, even if it cannot be analyzed in terms of classical physics, must be expressed in ordinary language making use of common logic" (1948, p. 317). What kind of expression in ordinary language would fit the experimental propositions of quantum mechanics? To answer this question, consider first a simple example of the measurement of the spin of a spin-$\frac{1}{2}$ particle in the r-direction. The mathematical representatives of the relevant experimental propositions satisfy the relations

$$
\begin{aligned}
\mu_{S_r}\left(\tfrac{1}{2}\right) \vee \mu_{S_r}\left(-\tfrac{1}{2}\right) &= \mu_{S_r}\left(\{-\tfrac{1}{2}, \tfrac{1}{2}\}\right) = \mathbb{1}, \\
\mu_{S_r}\left(\tfrac{1}{2}\right) \wedge \mu_{S_r}\left(-\tfrac{1}{2}\right) &= \mu_{S_r}(\varnothing) = \mathbb{0}.
\end{aligned}
\tag{8.23}
$$

Classical intuition suggests adopting Putnam's proposal and identifying these experimental propositions with the revelations of properties of the system. But, as noted above, this runs into problems with the Kochen-Specker theorem (at least when moving to spin-1 particles). The broader consequence of this theorem is that, assuming OQL1 and OQL2, there is no consistent way to assign truth values to the experimental propositions in $L(\mathcal{H})$. Then, if Bohr's "common logic" is construed as "classical propositional logic", we already run into problems. But the common logic of everyday language is much broader, and allows for constructs for which the existence of truth values may be contested. The right amount of wiggle room seems to be available when making use of conditionals.[11] Consider the following proposal:

---

[11]The idea that conditional sentences may not have truth values at least dates back to the work of Adams (1975, Ch. 1). Often this view is accompanied with the view that conditionals do not express propositions. Especially when having a truth value is considered a necessary condition for

$$\mu_A(\Delta) \hat{=} \text{"if } A \text{ is measured, then the outcome lies in } \Delta \text{"}. \tag{8.24}$$

Two not too controversial rules for conditionals are

$$(A \to B) \wedge (A \to C) = A \to (B \wedge C),$$
$$(A \to B) \vee (A \to C) = A \to (B \vee C). \tag{8.25}$$

If one assumes these rules to hold, and further makes the idealization that a measurement always yields an outcome, then the relations (8.23) come out true under the reading (8.24).

The natural language translation of experimental propositions given by (8.24) thus looks promising. Adopting the rules (8.25), orthodox quantum logic is easily seen to work also for observables A with a richer spectrum. That is, it makes the relations

$$\bigvee_{i=1}^{n} \mu_A(\Delta_i) = \mu_A \left(\bigcup_{i=1}^{n} \Delta_i \right),$$
$$\bigwedge_{i=1}^{n} \mu_A(\Delta_i) = \mu_A \left(\bigcap_{i=1}^{n} \Delta_i \right) \tag{8.26}$$

come out true. The Kochen-Specker theorem is furthermore circumvented by rejecting the existence of a global truth valuation $f : L(\mathcal{H}) \to \{0, 1\}$. It is after all unclear what the truth value of "if A is measured, then the outcome lies in Δ" would be when in fact A is not measured. However, even on this liberal view on conditionals as propositions without truth values, there are problems with this interpretation of OQL.

Consider now the measurements of spin on a spin-$\frac{1}{2}$ particle along two distinct directions r_1 and r_2. In OQL the following relations hold:

$$\mu_{S_{r_1}} \left(\tfrac{1}{2} \right) \vee \mu_{S_{r_2}} \left(\tfrac{1}{2} \right) = \mathbb{1},$$
$$\mu_{S_{r_1}} \left(\tfrac{1}{2} \right) \wedge \mu_{S_{r_2}} \left(\tfrac{1}{2} \right) = \mathbb{0} . \tag{8.27}$$

These relations do not follow from (8.24). On the other hand, they do not directly violate it either. One of the difficulties of assessing these propositions is that there are no situations in which the antecedents of the conditionals $\mu_{S_{r_1}}(\frac{1}{2})$ and $\mu_{S_{r_2}}(\frac{1}{2})$ are both true. In keeping with the earlier line of argument, one would then expect that conjunctions and disjunctions of these conditionals never have a truth value. But (8.27) suggest that they *should* be attributed a truth value. Now, although the assigned truth values (1 for the disjunction, and 0 for the conjunction) are not

---

a sentence to express a proposition. Here I permit myself some sloppiness to avoid falling too deep in the complicated philosophy of conditionals.

directly in conflict with (8.24), they lack an explanation, if only for why the the two values are different.

But the real problem comes in when looking at legitimate translations into natural language. In particular, consider the following false derivation:

$$\mu_{S_{r_1}}\left(\tfrac{1}{2}\right) = \mu_{S_{r_1}}\left(\tfrac{1}{2}\right) \wedge \left(\mu_{S_{r_2}}\left(\tfrac{1}{2}\right) \vee \mu_{S_{r_2}}\left(-\tfrac{1}{2}\right)\right), \tag{8.28a}$$

$$\mu_{S_{r_1}}\left(\tfrac{1}{2}\right) \wedge \left(\mu_{S_{r_2}}\left(\tfrac{1}{2}\right) \vee \mu_{S_{r_2}}\left(-\tfrac{1}{2}\right)\right) \\ = \left(\mu_{S_{r_1}}\left(\tfrac{1}{2}\right) \wedge \mu_{S_{r_2}}\left(\tfrac{1}{2}\right)\right) \vee \left(\mu_{S_{r_1}}\left(\tfrac{1}{2}\right) \wedge \mu_{S_{r_2}}\left(-\tfrac{1}{2}\right)\right), \tag{8.28b}$$

$$\left(\mu_{S_{r_1}}\left(\tfrac{1}{2}\right) \wedge \mu_{S_{r_2}}\left(\tfrac{1}{2}\right)\right) \vee \left(\mu_{S_{r_1}}\left(\tfrac{1}{2}\right) \wedge \mu_{S_{r_2}}\left(-\tfrac{1}{2}\right)\right) = \mathbb{0} \vee \mathbb{0} = \mathbb{0}. \tag{8.28c}$$

This derivation is both false in OQL and in the natural language translation given by (8.24). The problem is that in both cases it fails for different reasons. It fails in OQL due to the failure of distributivity, i.e., equation (8.28b) is false. On the other hand, this step seems fine when using the reading (8.24). The use of conditionals would rather point to the rejection of the third equality (8.28c). That is, while the conjunction of the antecedents of $\mu_{S_{r_1}}\left(\tfrac{1}{2}\right)$ and $\mu_{S_{r_2}}\left(\tfrac{1}{2}\right)$ may be a contradiction, there is no clear reason to believe that the conjunction of the two conditionals should therefore also be a contradiction. In fact, in natural language one can easily come up with examples in which the conjunction would seem to be true. For example, it seems fine to sometimes attribute the truth value 1 to the conjunction "If I were alone, I would cry. And if I were with you, I'd be home and dry." (Waters, 1970).

More generally, though, the rejection of the second step (8.28b) just seems wrong simply because distributivity holds in natural language. This, then, is the main reason that assumptions OQL1 and OQL2 cannot both be true. Thus the conclusion is that a direct implementation of OQL as a logic for experimental propositions in which joins and meets can function as disjunctions and conjunctions is not possible. But a more positive message is that $L(\mathcal{H})$ need not be too far off for this goal, and so one may opt to only reject OQL2.

This proposal also seems natural considering how far the introduction of conditionals has brought us. Assuming that this proposal is correct, it points to a particular failure of OQL2. For example, what seems to be missing are propositions that allow conjunctions of the form $\mu_{S_{r_1}}\left(\tfrac{1}{2}\right) \wedge \mu_{S_{r_2}}\left(\tfrac{1}{2}\right)$ that are distinct from contradictions. But more elementary, the simpler propositions that are used as antecedents and consequents for the conditionals are also missing. Thus assumption OQL2 has to be rejected in favor of an expansion of the collection of propositions. This option is investigated in chapter 9. In that chapter (and later ones) I make liberal use of ideas from intuitionistic logic. As some readers may feel some unease with such ideas, I end this chapter with a section with a short motivation for appealing to intuitionistic logic; both for why one could accept the use of this kind of

non-classical logic in general, and for why it could be useful for quantum mechanics in particular.

8.4   Motivating intuitionistic logic

In this section I explain why it is useful to have an open mind about intuitionistic logic when the aim is to find a suitable empiricist logic for quantum mechanics. What follows is not a defense of the necessity of intuitionistic logic, but rather of its possibility. The arguments are meant to persuade the reader of the idea that a change from orthodox quantum logic to intuitionistic quantum logic would be a step forward instead of just replacing one failure of classical logic (distributivity) with another one (the law of excluded middle). To start with, I give two examples of cases where quantum mechanics may be taken to suggest a failure of the law of excluded middle. After that, I give a more general discussion about the use of intuitionistic logic in constructive mathematics, and how this use could potentially be translated to an empiricist stance in physics.

The prime candidate for the role of negation in orthodox quantum logic is the orthocomplement. Assuming for the moment also that the join expresses disjunction in orthodox quantum logic, then the law of excluded middle may be seen to hold:

$$\forall P \in L(\mathcal{H}): \ P \vee P^{\perp} = \mathbb{1}\,. \tag{8.29}$$

Consequently, for any pair $P_1, P_2 \in L(\mathcal{H})$, P_1 is equivalent to $P_1 \wedge (P_2 \vee P_2^{\perp})$. On the other hand, there are cases for which both $P_1 \wedge P_2$ and $P_1 \wedge P_2^{\perp}$ are associated with a contradiction. An example is given by (8.28). In such a case, the proposition P_1 is seen to be incompatible both with P_2 and with $P_2^{\perp}$. It thus presents itself as a sort of excluded middle. So while orthodox quantum logic adopts a failure of the law of distributivity to explain that (8.28) is invalid (thus rejecting (8.28b)), the analysis just given shows that it is more natural to put the blame on a failure of the law of excluded middle, rejecting (8.28a). And such was the conclusion of Popper:

> the kind of change in classical logic which would fit what Birkhoff and von Neumann suggest [. . .] would be the rejection of the law of excluded middle [. . .], as proposed by Brouwer, but rejected by Birkhoff and von Neumann (Popper, 1968).

Although Popper presents an interesting idea here, there is more wrong with his paper than correct (Scheibe, 1974). For the present discussion it is useful to focus on just one problem. In his argument, Popper appears to presuppose that orthodox quantum logic was proposed by Birkhoff and von Neumann as a candidate logic for reasoning within quantum mechanics. That is, he appears to attribute Putnam's proposal to their work. By doing this, he doesn't seem to fully appreciate

the fact that OQL is derived from quantum mechanics in a natural mathematical way. There are no substantial philosophical choices involved in this derivation. What Birkhoff and von Neumann did most of all, was to investigate from a broader lattice-theoretical perspective the special properties of orthodox quantum logic (as opposed to just comparing it to Boolean algebras). But they were careful not to presuppose that the lattice operations identify with traditional logical connectives. Indeed, their paper ends with the following question:

> What experimental meaning can one attach to the meet and join of two given experimental propositions? (Birkhoff and von Neumann, 1936)

Thus there is no proposal in their paper that distributivity should be given up as a logical law to understand quantum mechanics.

What one has to take away from this is that, although Popper provides an argument for the use of intuitionistic logic, he doesn't recognize that adopting such a proposal requires more work. Orthodox quantum logic is *given* to us as part of quantum mechanics. If one finds intuitionistic logic conceptually more suitable, then one has to construct a framework that incorporates this. At the very least, one needs a Heyting algebra of propositions that in some way relates to the lattice $L(\mathcal{H})$. Specifically, as the proposal requires that the disjunction of P with $P^\perp$ does not equal unity, the join in $L(\mathcal{H})$ has to be replaced with something else (say $\vee_I$), and $L(\mathcal{H})$ as a set has to be extended to make room for the new elements $P \vee_I P^\perp$. A construction like this will in fact be given in the next chapter.[12]

The above argument for intuitionistic logic is quite mathematical in nature, and the reader may not yet be convinced that pursuing an intuitionistic quantum logic is very relevant from a physical point of view. One may suspect that conceptually there is still not much to be gained by going to intuitionistic logic. So for a second argument I turn to one of the masters in avoiding unnecessary mathematical jargon. According to Feynman, the double-slit experiment "has in it the heart of quantum mechanics. In reality, it contains the *only* mystery" (Feynman, Leighton, and Sands, 1963, p. 37-2). He admits that he cannot give a fully physical explanation of the phenomenon. Instead, he arrives at a Copenhagen-like compromise, concluding that, if done correctly, quantum mechanics does provide an accurate description of the phenomenon:

> What we must say (to avoid making wrong predictions) is the following. If one looks at the holes or, more accurately, if one has a piece of apparatus which is capable of determining whether the electrons go through hole 1 or hole 2, then one *can* say that it goes either through hole 1 or hole 2. *But*, when one does *not* try to tell which way the electron goes, when there is nothing in the experiment to disturb the

---

[12]Although the lattice constructed will not be a Heyting algebra, but a weak Heyting algebra.

electrons, then one may *not* say that an electron goes either through hole 1 or hole 2. If one does say that, and starts to make any deductions from the statement, he will make errors in the analysis. This is the logical tightrope on which we must walk if we wish to describe nature successfully. (Feynman, Leighton, and Sands, 1963, p. 37-9)

Intuitionistic logic may now be seen as a possible attempt to formalize Feynman's tightrope. When "going through hole 2" is thought of as the negation of "going through hole 1", a rejection of the universal validity of the law of excluded middle provides a way to avoid making errors in the analysis of the double slit experiment.

Although the two arguments just given point towards the idea that rejecting the law of excluded middle within the context of quantum mechanics may be natural, they do not yet indicate what it means to in fact adopt intuitionistic logic within a scientific theory. It deserves to be stressed that the use of intuitionistic logic does not help in understanding the physical processes behind phenomena like the double slit experiment. If anything, it distances us from the metaphysics behind quantum mechanics. To say that "going through hole 1 or going through hole 2" isn't a tautology, while at the same time accepting that the two disjunctions are each other's negation, is to admit that one doesn't know what is really going on. Instead, it is better to think of intuitionistic logic as a tool for correct reasoning about physical phenomena in the absence of a metaphysical picture and with an emphasis on the empirical content. To clarify this viewpoint, it is useful to draw a parallel with mathematics.

Constructive mathematics[13] is (in)famous for denying the truth of certain well-established theorems in classical mathematics. Formally this is the case whenever a proof requires the use of the law of excluded middle. But to understand why this use is considered to be troublesome, it is useful to discuss an explicit Brouwerian counterexample.[14] Consider the following mathematical statement:

$$\forall x \in \mathbb{R}: \ x \geq 0 \text{ or } x \leq 0. \tag{8.30}$$

In classical mathematics this statement can be proven, but in constructive mathematics this is not the case. For the constructivist, a proof for (8.30) requires providing a method that can be used to show for any given real number x whether $x \geq 0$ or $x \leq 0$. That is, for every x one has to be able (in principle) to decide which of the two disjuncts is true.

---

[13]Constructive mathematics is an approach in mathematics that rejects 'non-constructive' methods such as proof by contradiction, or unrestricted use of the axiom of choice. Intuitionistic mathematics is a sub-branch of constructive mathematics that strictly follows the methods of Brouwer. A helpful introduction to these distinctions is provided by Iemhoff (2013).

[14]Brouwer presented his first so-called weak counterexample in 1908. In (Brouwer, 1929) he gave an example of a number "floating around zero", of which the example given here is a version. For more on weak counterexamples see (Mandelkern, 1989; van Atten, 2011).

An explicit example can be given to demonstrate that this is not possible in general. A real number in Cauchy's sense is an (equivalence class represented by an) infinite sequence $(q_1, q_2, q_3, \ldots)$ of rational numbers. This means that the construction of a real number is never really finished. Consequently, the rule that specifies the real number need not be entirely definite either. It suffices if one has a rule that allows one to construct the finite sequence $(q_1, q_2, \ldots, q_n)$ for every n (and, of course, it should satisfy the Cauchy-criterion). This allows, for example, for the definition of the sum of two real numbers, which is done by point-wise summing the elements of the two sequences, without requiring that the sequence is "finished".

Now as an explicit example consider, for any natural number $k \in \mathbb{N}$, the real number x_k defined by the sequence

$$(x_k)_n = \begin{cases} \frac{(-1)^n}{n+1} & n \le k, \\ \frac{(-1)^k}{k+1} & n > k. \end{cases} \tag{8.31}$$

Given the value of k, it is easy to decide whether $x_k \le 0$ or $x_k \ge 0$. But instead of taking a natural number, one may also let k depend on the development of another mathematical program. Define k_{99} as the first place in the decimal expansion of π that is preceded by a sequence of 99 nines. The value of k_{99}, and even the question whether it exists or not, is an open problem in mathematics, possibly one that will never be solved. Nevertheless, (8.31) still defines a Cauchy sequence, because for every $n \in \mathbb{N}$ one can determine the value of $(x_{k_{99}})_n$ simply by going through the decimal expansion of π up to point n. Consequently, $x_{k_{99}}$ is a real number for which one cannot decide whether $x_{k_{99}} \le 0$ or $x_{k_{99}} \ge 0$.

A possible response to this argument may be that, irrespective of our knowledge about π, there is a fact of the matter determining whether k_{99} exists and, if it does, whether it is even or odd. Consequently, even if we do not know whether $x_{k_{99}} \le 0$ or $x_{k_{99}} \ge 0$, there is still some fact determining which one is the case. The assumption that there are such facts is known as Platonism. Roughly speaking, this is the mathematical equivalent of realism in philosophy of science, and similarly, it is not accepted by everyone.[15] Another response may be that (8.30) is simply true because it follows from the axioms together with the laws of logic. These laws themselves are then considered to be in no further need of justification: they simply form the basis for mathematical truth (a view known as logicism). A third option, known as formalism, is that (8.30) is true because the string of symbols that make up the formula can be arrived at by using the agreed upon rules of symbol manipulations that make up mathematics. The conundrum is then solved by deflating the notion of mathematical truth.

But one need not dwell on the philosophy of mathematics to find an appreciation for constructive mathematics. Also if one accepts that perhaps in some way

---

[15]A classical argument against Platonism is given by Benacerraf (1973).

statements like (8.30) are true, it is nevertheless the case that, epistemically, there are mathematical facts that we do not have access to. From such a point of view, while perhaps the statement "$x_{k_{99}} \leq 0$ or $x_{k_{99}} \geq 0$" may be considered to be true, pragmatically, it also verges on being vacuous, since one cannot decide which of the two options is the case. These kinds of mathematical statements may not be the ones we are interested in. This more pragmatic motivation for constructive mathematics has been colorfully defended by Bishop:

> Mathematics belongs to man, not to God. We are not interested in properties of the positive integers that have no descriptive meaning for finite man. When a man proves a positive integer to exist, he should show how to find it. If God has mathematics of his own that needs to be done, let him do it himself. (Bishop, 1967, p. 2)

This should strike a chord with any physicist who found that not all mathematics courses were as useful as (s)he hoped they would be. At some point the disappointment kicks in when learning, for example, that a mathematical proof for the existence of a solution to some integral doesn't always provide the solution of said integral, and it is the latter that is usually needed in practice. The reason is that such proofs are often done by contradiction, and thus they do not point towards an actual solution.

The use of intuitionistic logic in mathematics may thus be understood as a consequence of adopting an epistemic view. To say that some statement is true is to state that one has a proof for it: one knows it to be true. Similarly, to say that a number with certain properties exists means that one has a construction by which one knows which number has these properties.

In physics, more important than the aspect of knowing, is the aspect of experiencing. The scientific analog of the mathematical epistemic view, then, is an empirical one. To say that "the particle went through slit 1" is true, is to have the scientific equivalent of a proof for this statement: the performance of a measurement that shows that the particle went through slit 1. Measurement then plays the role of a scientific equivalent of proof in mathematics. Hence, the use of intuitionistic logic within physics, like its use within mathematics, may be given a pragmatic motivation. Such a pragmatic stance may be found in the words of Bohr (as quoted by Petersen (1963, p. 12)): "It is wrong to think that the task of physics is to find out how nature is. Physics concerns what we can say about nature", which is reminiscent of the words of Bishop. What is being denied is not the idea of an underlying reality, but, rather, the idea that one should give a description of this reality. It suffices to give a description of what we can experience. It is not surprising that intuitionistic logic can play a helpful role in such a description.

9

Weak intuitionistic quantum logic

9.1 Introduction

In the previous chapter, in section 8.3, it was suggested that the peculiarity of orthodox quantum logic (in particular, the failure of distributivity) may be a consequence of the incompleteness of $L(\mathcal{H})$ (in particular, of a failure of OQL2). However, it is quite a non-trivial task to determine what precisely is missing. In section 8.3 the assumption that projections represent conditional propositions suggested that there may be elements in $L(\mathcal{H})$ representing propositions, while the conjunction of these propositions has no mathematical representative in $L(\mathcal{H})$. In section 8.4, on the other hand, it was suggested that, when adopting an intuitionistic perspective, disjunctions of experimental propositions can no longer be identified with projections in general. But these suggestions alone are too sketchy to form the basis for the mathematical work required, which is to embed $L(\mathcal{H})$ into a propositional lattice that is at least distributive, and, preferably, a Heyting algebra (i.e., the intuitionistic equivalent of a Boolean algebra). Therefore, a formal approach is adopted in this chapter to find such an embedding. The hope is, then, that afterwards the result may also be found to be conceptually adequate.

   In this chapter I work with the idea suggested in section 8.4 rather than with the one suggested in section 8.3. Although the suggestion that projections correspond to conditional statements is an interesting and promising idea, it is also a difficult idea to work with. Apart from the fact that it motivates expanding $L(\mathcal{H})$ to incorporate elements that correspond to conjunctions of conditionals (by the rejection of (8.28c)), it also requires introducing elements that correspond to the antecedents and consequents of the conditionals (unless one is willing to except conditionals as atomic propositions). On top of that, it is also the case that, even outside the context of quantum mechanics, intuitions concerning conditionals are a poor guide for constructing formal theories for them (Edgington, 2014). In contrast, when adopting an intuitionistic perspective, there is hope that one can stay quite close to $L(\mathcal{H})$. The perspective only motivates the introduction of new disjunctions. In

fact, conjunctions may be seen to be quite satisfactory. To see this, consider two projections such that $P_1 \wedge P_2 = \mathbb{0}$. If this is the case, then P_1 and P_2 either refer to measurement outcomes that contradict each other (when $[P_1, P_2] = \mathbb{0}$), or they refer to measurements that are considered to be incompatible (when $[P_1, P_2] \neq \mathbb{0}$). In either case, there is no empirical fact that could be interpreted as a 'proof' for $P_1 \wedge P_2$ and so the identification with $\mathbb{0}$ is appropriate. Therefore, keeping the conjunctions of $L(\mathcal{H})$ seems a good assumption, at least on a first pass, to keep things simple.

The method adopted in this chapter builds on the work of Coecke (2002). This is a natural place to start, as his paper is entitled "quantum logic in intuitionistic perspective". Mathematically, it relies on the theory of injective hulls developed by Bruns and Lakser (1970). This provides a formal method of embedding meet-semilattices into Heyting algebras, where the embedding is given by a meet-homomorphism (i.e., a map that does preserve meets and the partial order, but that doesn't necessarily preserve joins). Section 9.2 is devoted to explaining this method when applied to $L(\mathcal{H})$. The result is an embedding of $L(\mathcal{H})$ into a Heyting algebra $\mathcal{DI}(L(\mathcal{H}))$ that doesn't preserve joins. Thus disjunctions of experimental propositions are no longer identified with projections.

It turns out that the obtained Heyting algebra is in fact Boolean. The result then doesn't square well with the motivation drawn from section 8.4. There, the idea that disjunctions do not correspond to elements of $L(\mathcal{H})$ was related to a rejection of the law of excluded middle. But this law does hold in $\mathcal{DI}(L(\mathcal{H}))$. In section 9.3 a proposal is made to relieve this tension. The orthocomplement of $L(\mathcal{H})$ is transformed into a new negation on $\mathcal{DI}(L(\mathcal{H}))$, which behaves intuitionistically. With this operation in place it is showed that $\mathcal{DI}(L(\mathcal{H}))$ can be turned into a so-called weak Heyting algebra. The chapter ends in section 9.4 with a conceptual reflection on the obtained results. It is concluded that, although the weak Heyting algebra has promising and interesting features, a conceptual understanding of the structure requires more work. In section 11.4 then, after developing an empiricist quantum logic in chapter 10, a new connection is made with this weak Heyting algebra.

9.2   Bruns-Lakser completion

The main formal slice of this section relies on work by Coecke (2002). It therefore deserves to be pointed out that his approach and motivation are somewhat different from the one adopted here. In his paper, Coecke works in the tradition of the Geneva-Brussels school of quantum logic (Jauch and Piron, 1969; Piron, 1976; Aerts, 1982; Moore, 1999). In this approach, elements of $L(\mathcal{H})$ are viewed as (testable) properties. They correspond to particular experiments for which the

possible outcomes are 0 and 1, dubbed yes-no questions. For example, when a measurement of $P \in L(\mathcal{H})$ yields the outcome 0 this is identified with the system not having the property P. Thus 'property' is given an operational rather than a metaphysical meaning here. There is also a role for conjunctions in this framework. The property $P_1 \wedge P_2$ corresponds to the experiment where an arbitrary choice is made between the two experiments P_1 and P_2, and then taking the outcome of that experiment as the value for $P_1 \wedge P_2$.

In the view of Coecke (2002), the elements of $L(\mathcal{H})$ identify with propositions about these properties. But, in the light of Birkhoff and von Neumann's question (see page 128) he recognizes that disjunctions of such propositions need not correspond to properties. The program can then be summed up as follows:

> if we want to describe a physical system by a "language" that is closed under all disjunctions of properties, we formally need to introduce those additional propositions that express disjunctions of properties and that do not correspond to a property in the property lattice. (Coecke, 2002, pp. 419–420)

As no assumptions are made on the precise meaning of the propositions in $L(\mathcal{H})$, the extension of the lattice takes an abstract route. In such an approach restraint is a useful guide. The idea is to add just enough disjunctions of propositions in $L(\mathcal{H})$ to make the new lattice a Heyting algebra. The formal approach adopted to establish this uses the theory of injective hulls of meet-semilattices as developed by Bruns and Lakser (1970).[1] This requires the introduction of some terminology.

Definition 9.1. A meet-semilattice L is called *injective* if for every pair of meet-semilattices M_1, M_2 with M_1 a sub-meet-semilattice of M_2 and every homomorphism $\phi : M_1 \to L$ there exists an extension $\overline{\phi} : M_2 \to L$. That is, for every M_1, M_2, ϕ there exists a homomorphism $\overline{\phi}$ that makes the diagram

$$\begin{array}{ccc} M_1 & \xrightarrow{\phi} & L \\ \downarrow & \nearrow & \\ M_2 & \overline{\phi} & \end{array} \qquad (9.1)$$

commute.

---

[1] Within the context of quantum logic, the Bruns-Lakser completion has also been studied by Marsden (2010) and Heunen, Landsman, and Spitters (2012). In these works the construction is compared to methods within topos-theoretical approaches to quantum theories. Since topos theory plays virtually no role in this dissertation, it is hard to relate these results to the present discussion. One caveat deserves to be mentioned though: in the next chapter a Heyting algebra is constructed that also occurs in the topos-theoretical framework discussed by Heunen, Landsman, and Spitters (ibid.). However, this fact seems to be somewhat coincidental.

The interest in injective meet-semilattices stems from the fact that these objects turn out to be complete Heyting algebras. More precisely, one can show that for an injective meet-semilattice arbitrary meets and joins always exist and that they satisfy the relation

$$a \wedge \bigvee_{b \in M} b = \bigvee_{b \in M} (a \wedge b) \tag{9.2}$$

for all a and M. This relation allows the definition of the relative pseudo-complement $a \rightarrow b := \bigvee \{c \mid c \wedge a \leq b\}$. Conversely, one can show that every complete Heyting algebra is an injective meet-semilattice (Bruns and Lakser, 1970).

The upshot is that instead of looking for the smallest complete Heyting algebra that embeds a given meet-semilattice ($L(\mathcal{H})$ in our case), one can look for the smallest injective meet-semilattice. This notion is captured by the concept of an *injective hull* of a meet-semilattice L. An injective meet-semilattice H is an injective hull of the meet-semilattice L if there exists an injective homomorphism $\phi : L \rightarrow H$ and if for any other injective meet-semilattice H' that satisfies this criterion there is also an embedding of H into H'. In the form of a diagram one thus has:

$$
\begin{array}{ccc}
L & \xrightarrow{\ \phi\ } & H \\
 & \searrow_{\phi'} & \downarrow \\
 & & H'
\end{array}
\tag{9.3}
$$

The injective hull of L may be shown to be unique up to isomorphism.[2] So what is needed now is a helpful characterization of injective hulls of meet-semilattices. This was also given by Bruns and Lakser, using the following terminology.

Definition 9.2. Let L be a meet-semilattice. A subset $A \subset L$ is called *admissible* if $\bigvee_{a \in A} a$ exists and for every $b \in L$ the join $\bigvee_{a \in A} (b \wedge a)$ exists as well and equals $b \wedge \bigvee_{a \in A} a$. A non-empty subset I of L is called an *ideal* if

$$\forall a \in I, \forall b \in L : \text{ if } b \leq a, \text{ then } b \in I. \tag{9.4}$$

An ideal I is called *distributive* if $\bigvee_{a \in A} \in I$ for every admissible set $A \subset I$.

Now let $\mathcal{DI}(L)$ denote the set of distributive ideals in a meet-semilattice L. Set-inclusion turns this into a partial ordered set. Furthermore, it is easy to see that the intersection of two distributive ideals is again a distributive ideal, thus $\mathcal{DI}(L)$ is a meet-semilattice. Bruns and Lakser proved the following result.

Theorem 9.1. *For every meet-semilattice L every injective hull H is isomorphic to $\mathcal{DI}(L)$.*

---

[2]Injective hulls are usually defined as maximal essential extensions. But that terminology is irrelevant for the present discussion. For details see (Coecke, 2002, p. 421).

The embedding $i : L \to \mathcal{DI}(L)$ is given by assigning to each element $a \in L$ the corresponding down set

$$i(a) := \downarrow a = \{b \in L \mid b \leq a\}. \tag{9.5}$$

The reader may check that this is indeed a meet-homomorphism. However, generally it does not preserve joins even if L has joins. In fact, one then has that

$$\downarrow a_1 \vee \downarrow a_2 \leq \downarrow(a_1 \vee a_2). \tag{9.6}$$

When equality fails the element $\downarrow a_1 \vee \downarrow a_2$ may be viewed as the introduced disjunction that does not correspond to any element of L.

Example 9.1. Consider the non-distributive lattice known as N_5 given by the Hasse diagram
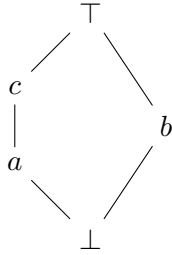
Distributivity can be seen to fail because

$$(a \vee b) \wedge c = \top \wedge c = c \neq a = a \vee \bot = (a \wedge c) \vee (b \wedge c). \tag{9.7}$$

The set of ideals is given by

$$\mathcal{I}(N_5) = \{\downarrow\bot, \downarrow a, \downarrow b, \downarrow c, \downarrow\top, I_1 := \{\bot, a, b\}, I_2 := \{\bot, a, b, c\}\}. \tag{9.8}$$

The subsets A of I_1 for which $\bigvee A \notin I_1$ are $\{a, b\}$ and $\{\bot, a, b\}$. Since both aren't admissible it follows that I_1 is distributive. The ideal I_2 on the other hand isn't distributive. The subset $\{b, c\}$ of I_2 is admissible since

$$(x \wedge b) \vee (x \wedge c) = x = x \wedge (b \vee c) \tag{9.9}$$

for all $x \in N_5$. But $\bigvee\{b, c\} \notin I_2$. The injective hull of N_5 can then be represented by the Hasse diagram

$\triangle$
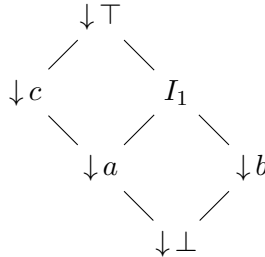
As noted before, $\mathcal{DI}(L)$ is an injective meet-semilattice and may therefore be understood as a complete Heyting algebra. Turning then to the case of $L(\mathcal{H})$, the new disjunctions are obtained by switching to $\mathcal{DI}(L(\mathcal{H}))$. But perhaps somewhat surprisingly, it turns out in this case that $\mathcal{DI}(L(\mathcal{H}))$ is not a proper Heyting algebra, but is in fact Boolean. Specifically, one has the following theorem.

Theorem 9.2. *The lattices $\mathcal{DI}(L(\mathcal{H}))$ and $\mathcal{P}(L_1(\mathcal{H}))$ are isomorphic (as complete bounded lattices).*

Here $\mathcal{P}(L_1(\mathcal{H}))$ denotes the power set of the 1-dimensional projections, which is clearly a Boolean algebra. The theorem is a special case of an example in (Coecke, 2002, p. 423). A direct proof of this version is given in (Hermens, 2013b, §2). Here I will not go into the details of the proof, but instead illustrate what is going on conceptually.

The result may be considered surprising against the background of the motivation for expanding the lattice $L(\mathcal{H})$ on the basis of Popper's analysis (see section 8.4). What was found unsatisfactory there, was the idea that in $L(\mathcal{H})$ the join of P and $P^\perp$ is identified with a tautology even though there exist projections P' that are incompatible with both P and $P^\perp$. The new disjunctions in $\mathcal{DI}(L(\mathcal{H}))$ were supposed to turn this around. That is, the idea was that the disjunction of P and $P^\perp$ would no longer correspond to a tautology by rejecting the law of excluded middle. But now we find that the law of excluded middle *does* hold here. So what has happened?

The answer to this question is that the introduction of a new disjunction forced the introduction of a new negation. This deserves to be spelled out. The negation in $\mathcal{DI}(L(\mathcal{H}))$ is defined as $\neg I := I \to \{\mathbb{0}\}$. This negation is much weaker than the orthocomplement in quantum logic, because one has

$$\downarrow P^\perp \leq \neg \downarrow P, \quad \forall P \in L(\mathcal{H}) \tag{9.10}$$

with equality if and only if $P = \mathbb{0}$ or $P = \mathbb{1}$.

This of course doesn't show yet how the negation on $\mathcal{DI}(L(\mathcal{H}))$ obeys the law of excluded middle. This is better understood by looking at $\mathcal{P}(L_1(\mathcal{H}))$. In this lattice it is obvious that negation satisfies the law of excluded middle. How this stands in relation to $\mathcal{DI}(L(\mathcal{H}))$ is best seen by explicitly defining the maps that connect the lattices. The isomorphism f from $\mathcal{P}(L_1(\mathcal{H}))$ to $\mathcal{DI}(L(\mathcal{H}))$ is defined as

$$f(\Delta) := \{P \in L(\mathcal{H}) \,|\, \downarrow P \cap L_1(\mathcal{H}) \subset \Delta\}, \tag{9.11}$$

and its inverse is given by

$$f^{-1}(I) = I \cap L_1(\mathcal{H}) = \bigcup_{P \in I} (\downarrow P \cap L_1(\mathcal{H})). \tag{9.12}$$

The existence of an isomorphism implies that there is also an embedding r of $L(\mathcal{H})$ into $\mathcal{P}(L_1(\mathcal{H}))$ such that $r = f^{-1} \circ i$. That is, there is a function r that makes the diagram

$$
\begin{array}{ccc}
L(\mathcal{H}) & \xrightarrow{\;\;i\;\;} & \mathcal{DI}(L(\mathcal{H})) \\
\downarrow{\scriptstyle r} & \nearrow{\scriptstyle f} & \\
\mathcal{P}(L_1(\mathcal{H})) & &
\end{array}
\tag{9.13}
$$

commute. One can show that this function is given by

$$
r(P) := \,\downarrow\! P \cap L_1(\mathcal{H}). \tag{9.14}
$$

As the equivalent of i, the function r also obeys

$$
r(P^\perp) \le \left(r(P) \right)^c, \tag{9.15}
$$

with equality if and only if $P = \mathbb{0}$ or $P = \mathbb{1}$. In fact, when equality doesn't hold, there is a huge gap between the two. Let $P' \in L_1(\mathcal{H})$ be any projection that doesn't commute with $P \in L(\mathcal{H})$. Then P' is neither an element of $r(P)$ nor of $r(P^\perp)$. Or, equivalently, $P' \in (r(P))^c \cap \left(r(P^\perp) \right)^c$. Thus while P' is incompatible with both P and $P^\perp$, this does not lead to a violation of the law of excluded middle because now P' has become compatible with both the negation of P and the negation of $P^\perp$. It is not entirely clear though what to make of this.

Thus, in going from $L(\mathcal{H})$ to $\mathcal{DI}(L(\mathcal{H}))$ not only the joins have changed, but also the orthocomplement no longer identifies with negation. In fact, the whole operation of orthocomplementation seems to have been lost. This is somewhat dissatisfying for an operation that was thought to be relevant for reasoning about experimental propositions. In the next section I show that this operation can be re-introduced on $\mathcal{DI}(L(\mathcal{H}))$ and can be used to turn the lattice into a weak Heyting algebra that is no longer Boolean.

9.3   An intuitionistic perspective

The objective of this section is to investigate how far one can get with introducing a new operation on $\mathcal{DI}(L(\mathcal{H}))$ that behaves as the orthocomplement does on $L(\mathcal{H})$. The motivation is that with the introduction of such an operation one may find an intuitionistic-like structure that fits with the original motivation for constructing $\mathcal{DI}(L(\mathcal{H}))$. As in the previous section, the method used is rather formal. As such, philosophical considerations play no fundamental role. It is only after the the formal work has been done that the results can be subjected to philosophical reflections. This is done in section 9.4.

The goal is to introduce and then investigate a unary operation $\sim\colon \mathcal{DI}(L(\mathcal{H})) \to \mathcal{DI}(L(\mathcal{H}))$ that mimics the orthocomplement on $L(\mathcal{H})$. That is, the operation is required to satisfy

$$\sim \downarrow P = \downarrow P^{\perp} \ \forall P \in L(\mathcal{H}). \tag{9.16}$$

The introduction of $\sim$ turns out to be quite straightforward when turning attention to $\mathcal{P}(L_1(\mathcal{H}))$. Note that

$$r(P^{\perp}) = \left\{ P' \in L_1(\mathcal{H}) \,\middle|\, P'P = \mathbb{0} \right\}. \tag{9.17}$$

This suggests the definition

$$\sim \Delta := \left\{ P' \in L_1(\mathcal{H}) \,\middle|\, P'P = \mathbb{0} \ \forall P \in \Delta \right\}. \tag{9.18}$$

It is easy to check that $\sim r(P) = r(P^{\perp})$ for all $P \in L(\mathcal{H})$, i.e., that (9.16) is satisfied. Let us then call $\sim$ the *pseudo-negation* on $\mathcal{P}(L_1(\mathcal{H}))$.

The pseudo-negation has many properties that are considered typical for the negation in intuitionistic logic. For example, the equalities

$$\Delta \vee \sim \Delta = \top, \tag{9.19a}$$

$$\sim \Delta \vee \sim\sim \Delta = \top \tag{9.19b}$$

are satisfied if and only if $\Delta = \bot$ or $\Delta = \top$. On the other hand,

$$\sim\sim (\Delta \vee \sim \Delta) = \top \tag{9.20}$$

holds for all Δ. Of the De Morgan laws, only

$$\sim \Delta_1 \wedge \sim \Delta_2 = \sim (\Delta_1 \vee \Delta_2) \tag{9.21}$$

holds for all Δ_1, Δ_2, while the other law only holds in one direction:

$$\sim \Delta_1 \vee \sim \Delta_2 \leq \sim (\Delta_1 \wedge \Delta_2), \tag{9.22}$$

for all Δ_1, Δ_2.

The pseudo-negation also provides a link between the join on $L(\mathcal{H})$ and the join on $\mathcal{P}(L_1(\mathcal{H}))$. For any subset $K \subset L(\mathcal{H})$, one has

$$\sim\sim \left(\bigvee_{P \in K} r(P) \right) = r\left(\bigvee_{P \in K} P \right). \tag{9.23}$$

This is reminiscent of the Gödel-Gentzen double-negation translation of classical logic into intuitionistic logic (Gödel, 1933; Gentzen, 1936). This translation uses the fact that, of every formula that may be proven in classical logic, the double

negation can be proven in intuitionistic logic. Similarly, (9.23) can be seen to state that, to translate the join in $L(\mathcal{H})$ into $\mathcal{P}(L_1(\mathcal{H}))$, one can't just take the joins in $\mathcal{P}(L_1(\mathcal{H}))$ of each of the terms in the join translated individually using r. Instead, one has to take the double-negation in $\mathcal{P}(L_1(\mathcal{H}))$ after this operation. So one may think of the join in $L(\mathcal{H})$ to be more classical-like, and the join in $\mathcal{P}(L_1(\mathcal{H}))$ to be more intuitionistic-like.

These results strengthen the idea that $\sim$ introduces an intuitionistic negation. On the other hand, it is also clear that it cannot really be an intuitionistic negation. This is because the negation is always defined in terms of the relative pseudo-complement ($\neg x := x \to \bot$), which is unique if it exists. And for $\mathcal{P}(L_1(\mathcal{H}))$ this just gives the complement. So if $\sim$ is to be the pseudo-complement on a Heyting algebra, this algebra, if it exists, has to be an extension of $\mathcal{P}(L_1(\mathcal{H}))$. But instead of searching for such an extension, one may wonder to what extent $\mathcal{P}(L_1(\mathcal{H}))$ is 'almost' a Heyting algebra. That is, is it possible to introduce an implication $\to$ that is 'almost' a relative pseudo-complement and satisfies

$$\sim \Delta = \Delta \to \bot \tag{9.24}$$

for all Δ? An appropriate formalism for investigating this question is that of weak Heyting algebras as introduced by Celani and Jansana (2005).

Definition 9.3. A *weak Heyting algebra* $(L, \vee, \wedge, \to)$ is a bounded distributive lattice with an implication relation that satisfies for all $a, b, c \in L$

(i) $a \to a = \top$,

(ii) $a \to (b \wedge c) = (a \to b) \wedge (a \to c)$,

(iii) $(a \vee b) \to c = (a \to c) \wedge (b \to c)$,

(iv) $(a \to b) \wedge (b \to c) \leq a \to c$.

This implication relation may be shown to have the following properties (proofs can be found in (ibid.)):

(a) If $a \leq b$, then for all c: $c \to a \leq c \to b$ and $b \to c \leq a \to c$.

(b) If $a \leq b$, then $a \to b = \top$.

(c) For all a, b, c: $(a \to b) \wedge (a \to c) \leq a \to (b \vee c)$.

These properties provide a useful guide for determining what an implication relation on $\mathcal{P}(L_1(\mathcal{H}))$ must look like if it is to satisfy (9.24). For ease of notation, let $\overline{P}$ denote the singleton set $\{P\}$ for $P \in L_1(\mathcal{H})$. As a special case of (b), note that whenever $P \in \Delta$

$$\overline{P} \to \Delta = \top. \tag{9.25}$$

On the other hand, when $P \notin \Delta$ one has

$$\overline{P} \to \Delta \stackrel{(i)}{=} \left(\overline{P} \to \Delta\right) \wedge \left(\overline{P} \to \overline{P}\right) \stackrel{(ii)}{=} \overline{P} \to \bot \stackrel{(9.24)}{=} \sim \overline{P}. \tag{9.26}$$

These two cases can be combined in a single relation using (9.20):

$$\overline{P} \to \Delta = \sim\sim \left(\sim \overline{P} \vee \left(\overline{P} \wedge \Delta\right)\right) = \begin{cases} \top & P \in \Delta, \\ \sim \overline{P} & P \notin \Delta. \end{cases} \tag{9.27}$$

Finally, if one assumes that (iii) also holds for arbitrary joins and meets, one finds that the implication relation should satisfy

$$\begin{aligned}
\Delta_1 \to \Delta_2 &= \left(\bigvee_{P \in \Delta_1} \overline{P}\right) \to \Delta_1 = \bigwedge_{P \in \Delta_1} \left(\overline{P} \to \Delta_2\right) \\
&= \bigwedge_{P \in \Delta_1} \sim\sim \left(\sim \overline{P} \vee \left(\overline{P} \wedge \Delta_2\right)\right) = \bigwedge_{P \in \Delta_1 \setminus \Delta_2} \sim \overline{P}
\end{aligned} \tag{9.28}$$

for all $\Delta_1, \Delta_2 \in \mathcal{P}(L_1(\mathcal{H}))$, where the empty meet is identified with $\top$. The next theorem show that this implication relation indeed turns $\mathcal{P}(L_1(\mathcal{H}))$ into a weak Heyting algebra that satisfies (9.24).

Theorem 9.3. *The map $(\Delta_1, \Delta_2) \mapsto \Delta_1 \to \Delta_2$, defined by*

$$\Delta_1 \to \Delta_2 := \bigwedge_{P \in \Delta_1 \setminus \Delta_2} \sim \overline{P}, \tag{9.29}$$

turns $\mathcal{P}(L_1(\mathcal{H}))$ into a weak Heyting algebra and satisfies (9.24).

Proof. All that is required is to show that (i)-(iv) and (9.24) are satisfied. The first one is trivial, and the others follow by writing out. For (ii):

$$\begin{aligned}
\Delta_1 \to (\Delta_2 \wedge \Delta_3) &= \bigwedge \left\{\sim \overline{P} \mid P \in \Delta_1 \setminus (\Delta_2 \wedge \Delta_3)\right\} \\
&= \left(\bigwedge_{P \in \Delta_1 \setminus \Delta_2} \sim \overline{P}\right) \wedge \left(\bigwedge_{P \in \Delta_1 \setminus \Delta_3} \sim \overline{P}\right) \\
&= (\Delta_1 \to \Delta_2) \wedge (\Delta_1 \to \Delta_3).
\end{aligned} \tag{9.30}$$

For (iii):

$$\begin{aligned}
(\Delta_1 \vee \Delta_2) \to \Delta_3 &= \bigwedge \left\{\sim \overline{P} \mid P \in (\Delta_1 \vee \Delta_2) \setminus \Delta_3\right\} = \bigwedge_{\substack{P \in \Delta_1 \setminus \Delta_3, \\ P \in \Delta_2 \setminus \Delta_3}} \sim \overline{P} \\
&= \left(\bigwedge_{P \in \Delta_1 \setminus \Delta_3} \sim \overline{P}\right) \wedge \left(\bigwedge_{P \in \Delta_2 \setminus \Delta_3} \sim \overline{P}\right) \\
&= (\Delta_1 \to \Delta_3) \wedge (\Delta_2 \to \Delta_3).
\end{aligned} \tag{9.31}$$

For (iv):

$$(\Delta_1 \to \Delta_2) \wedge (\Delta_2 \to \Delta_3) = \left(\bigwedge_{P \in \Delta_1 \backslash \Delta_2} \sim \overline{P} \right) \wedge \left(\bigwedge_{P \in \Delta_2 \backslash \Delta_3} \sim \overline{P} \right)$$

$$= \bigwedge_{\substack{P \in \Delta_1 \backslash \Delta_2, \\ P \in \Delta_2 \backslash \Delta_3}} \sim \overline{P} \leq \bigwedge_{P \in \Delta_1 \backslash \Delta_3} \sim \overline{P} \tag{9.32}$$

$$= \Delta_1 \to \Delta_3.$$

And finally (9.24):

$$\Delta \to \bot = \bigwedge_{P \in \Delta} \sim \overline{P} = \sim \Delta. \tag{9.33}$$

$\square$

With the above considerations that led up to the definition of the implication relation, it almost follows that the weak Heyting algebra is uniquely defined by the pseudo-negation $\sim$. The only additional assumption was that (iii) holds for arbitrary joins and meets. This assumption is not crucial though. In fact, for Boolean lattices, if a pseudo-negation arises from a weak Heyting algebra structure, then this structure is unique. Specifically, one has the following theorem:

Theorem 9.4. *Let $(L, \vee, \wedge)$ be a bounded Boolean lattice and suppose both $\to$ and $\to'$ turn it into a weak Heyting algebra. If*

$$a \to \bot = a \to' \bot \tag{9.34}$$

for all $a \in L$, then $\to = \to'$.

Proof. As noted by Celani and Jansana (2005), for bounded Boolean lattices there is a bijective correspondence between weak Heyting algebras $(L, \vee, \wedge, \to)$ and normal modal algebras $(L, \vee, \wedge, {}^c, \Diamond)$. Specifically, every weak Heyting algebra is turned into a normal modal algebra with the definition

$$\Diamond a := (\top \to a^c)^c, \tag{9.35}$$

where c denotes the complement. The converse mapping is given by $a_1 \to a_2 := (\Diamond(a_1 \wedge a_2^c))^c$.

Now let $\to$ and $\to'$ be two implication relations that turn $(L, \vee, \wedge)$ into a weak Heyting algebra, and let $\Diamond$ and $\Diamond'$ be their corresponding modal operations. Because $a \to \bot = a \to' \bot$ for all $a \in L$, it follows that

$$(\Diamond a)^c = (\Diamond(a \wedge \bot^c))^c = a \to \bot = a \to' \bot = (\Diamond'(a \wedge \bot^c))^c = (\Diamond' a)^c. \tag{9.36}$$

Thus $\Diamond = \Diamond'$ and consequently $\to = \to'$. $\square$

Incidentally, this proof shows that apart from a completion of $L(\mathcal{H})$ to a Heyting algebra (that is in fact Boolean), the Bruns-Lakser construction together with the introduction of the implication (9.29), also gives rise to a completion to a normal modal algebra. These latter kind of completions for orthodox quantum logic have also been studied by Kramer (2014). The normal modal algebra obtained here, however, does not appear to be related in a clear way to the constructions studied by Kramer (ibid.).

9.4   Evaluation

In the previous two sections I followed a formal approach to expand the orthodox quantum logic $L(\mathcal{H})$ to obtain a structure for which the label 'logic' is more appropriate. Specifically, while it was suggested that the elements of $L(\mathcal{H})$ do refer to propositions, the lattice operations do not refer to logical connectives. With the transition to the lattice $\mathcal{P}(L_1(\mathcal{H}))$ a plenitude of new operations have been introduced that may be candidates for logical connectives. Whether they can function as such depends on which propositions the elements of $L(\mathcal{H})$ were supposed to express. In the construction of $\mathcal{P}(L_1(\mathcal{H}))$ though, nothing was assumed about this. The benefit of this formal approach was that we could rely on a clean mathematical method to make the expansion as minimal as possible. The downside is that we are left with the question of interpretation. Bell's warning appears to be apt here:

> Now it is precisely in cleaning up intuitive ideas for mathematics that one is likely to throw out the baby with the bathwater. (Bell, 1990, p. 106)

Have we really thrown out the baby with the bathwater? This depends on the goal, with regard to which there is good news and bad news. Since the bad news leads to a new research task, let me start with the good news.

Perhaps the most straightforward interpretation of $\mathcal{P}(L_1(\mathcal{H}))$ arises from the identification of 1-dimensional projections and rays, together with a realist interpretation of rays as representing the states of quantum systems. A set $\Delta \subset L_1(\mathcal{H})$ then refers to the proposition that the state corresponds to one of the 1-dimensional projections in the set Δ. They are therefore propositions about the properties of a system. There is then a clear sense in which the elements of $\Delta \subset L_1(\mathcal{H})$ arising from $L(\mathcal{H})$ may be called 'testable' properties in line with the Geneva-Brussels interpretation, although the term falsifiable may be more apt: a single measurement can suffice to show that the proposition $r(P)$ for some $P \in L(\mathcal{H})$ does not hold, namely, when a measurement of P yields the outcome 0. The complement of $r(P)$ then corresponds to the proposition that the system does not have the property $r(P)$, while $\sim r(P)$ denotes the weakest falsifiable proposition that implies the negation

of $r(P)$. The modal operation introduced in the proof of Theorem 9.4 also suits this interpretation. Generally, this operation is given by

$$\Diamond \Delta = (\top \to \Delta^c)^c = \left(\bigwedge_{P \in \Delta} \sim \overline{P} \right)^c = (\sim \Delta)^c. \tag{9.37}$$

In particular, $\Diamond r(P)$ corresponds to all states for which it is possible that a measurement of P yields the value 1.

What I used implicitly in this analysis is the Born rule. For example, the sense of 'possibility' can be understood as meaning 'non-zero probability'. This means that on the given interpretation, the propositions in $\mathcal{P}(L_1(\mathcal{H}))$ reflect statements about probabilities. But this is in conflict with the initial motivation for constructing this lattice. The hope was that the elements of $\mathcal{P}(L_1(\mathcal{H}))$ could refer to measurements and their outcomes: that they would correspond to propositions to which one can assign probabilities, not propositions about probabilities. This does of course not mean that one can't define probability functions on $\mathcal{P}(L_1(\mathcal{H}))$. It just means that these probabilities do not in any obvious way refer to probabilities of measurement outcomes. Rather, they would be probabilities about probabilities.

The mixing up of probabilities of propositions and propositions about probabilities may be a trait that occurs due to the double role of the Hilbert space in quantum mechanics. The Hilbert space both makes up the domain of quantum probability functions (its closed linear subspaces), and the probability functions themselves (as rays). Then it appears that somewhere along the way when searching for a propositional lattice to clear up this domain role, we switched to the function role. It is my contention that the source of this mixing may be traced back to the assumptions OQL1 and OQL2 on page 123. Experimental propositions were introduced as subsets of observation spaces, and before even spelling out what these propositions were supposed to express, they were identified with closed linear subspaces. As a consequence, when interpreting $\mathcal{P}(L_1(\mathcal{H}))$ as a lattice of experimental propositions that expands $L(\mathcal{H})$, one now has a dual task of finding out what the correct formulation of these propositions is in natural language, and why it is justified to identify them with closed linear subspaces.

It does not follow from these considerations that the obtained weak Heyting algebra has no conceptual role to play for quantum probability. In fact, as a natural distributive extension of $L(\mathcal{H})$, and given the importance of $L(\mathcal{H})$ for the formulation of quantum probability, it seems that there should be a conceptual role for $\mathcal{P}(L_1(\mathcal{H}))$. It just is not clear at this point what its role can or should be. To avoid guess work, it is better to head back to the drawing board. It is time to let go of the work of Birkhoff and von Neumann, and re-invent quantum logic. What kind of propositions does one require to form the domain of quantum probability functions? How do these propositions generate a fitting propositional lattice? And

how does the end result relate to the lattices $L(\mathcal{H})$ and $\mathcal{P}(L_1(\mathcal{H}))$? This project is taken up in the next two chapters. As one says in Dutch: this tale will have a little tail. Specifically, the weak Heyting algebra will re-enter the discussion in section 11.4.

10

An empiricist quantum logic

The ongoing program is to obtain a reformulation of quantum probability in which the mathematical symbols have a clear physical meaning. In section 8.1 I explained that the main work is to be done on the domain $L(\mathcal{H})$ of quantum probability functions. The aim is to replace this lattice with a lattice of experimental propositions in which the lattice operations correspond to logical connectives. In the previous two chapters this program was followed by attempting to directly interpret the domain $L(\mathcal{H})$ as a lattice of propositions and possibly expand it on the basis of this interpretation. In the previous chapter this task has only been partially completed. A proper expansion of $L(\mathcal{H})$ had been obtained and it was argued that a clear physical interpretation for this lattice is available. However, the given interpretation was found to be unfitting for the ongoing program. Instead of dwelling on the possibility of different interpretations, the strategy will be reversed in this chapter. Here I start with propositions that already have an empirical interpretation, but are not yet tied together in a lattice. The goal is to build from these propositions a new quantum logic that is suitable to form a domain of quantum probability functions. To be sure, this is an unorthodox project, and that is why it requires an entire chapter to formulate this new logic. Furthermore, to keep things as simple as possible, attention is restricted to finite-dimensional Hilbert spaces throughout this chapter and the next one.

If the new quantum logic is to be part of a new formulation of quantum probability, it has to relate to the Born postulate, as well as respect the non-probabilistic predictions of quantum mechanics. A basic set of elementary experimental propositions that can play this role is introduced in section 10.1. But for such a set to become a logic, it has to be given a lattice structure that ideally is such that meets and joins can be interpreted as conjunctions and disjunctions. A lattice structure is imposed on the set of elementary experimental propositions by quantum mechanics. This structure is scrutinized in section 10.2. It is argued that this lattice is not rich enough to incorporate disjunctions of elementary experimental proposition. Thus, an extension is required, which is constructed in section 10.3. The obtained lat-

tice forms a Heyting algebra of experimental propositions. In the final section this intuitionistic quantum logic is extended one last time to obtain a Boolean algebra of experimental propositions. The motivation for this final step is somewhat more pragmatic: when it comes to evaluating possible probability functions, Boolean algebras are a more familiar domain than Heyting algebras. With the new empiricist quantum logic that is obtained in this chapter, we will be well prepared to undertake a new investigation of quantum probability in chapter 11.

10.1 Elementary experimental propositions

The aim is to find a lattice of propositions that can provide a formal conceptualization of quantum probability. What these propositions should be is open, but a reasonable minimal requirement is that they should play a role in formulating the Born rule. A logical first place to look for inspiration then is the Born postulate itself, as introduced in chapter 3.

BoP (Born Postulate) If the observable $\mathcal{A}$ is associated with the self-adjoint operator A, and ψ is the state of the system, then the probability to find a value in $\Delta \subset \sigma(A)$ upon a measurement of $\mathcal{A}$ is given by the Born rule

$$\mathbb{P}_\psi(\mathcal{A} \in \Delta) = \frac{\langle \psi, \mu_A(\Delta)\psi \rangle}{\langle \psi, \psi \rangle} = \mathrm{Tr}\left(P_\psi \mu_A(\Delta)\right). \tag{10.1}$$

Alternatively, if the state is given by the density operator ρ, the probability is given by

$$\mathbb{P}_\rho(\mathcal{A} \in \Delta) = \mathrm{Tr}\left(\rho \mu_A(\Delta)\right). \tag{10.2}$$

When trying to formalize the Born postulate in such a way that it connects to a lattice of propositions, there is one first obstacle that cannot be overcome in a non-controversial way: the postulate makes use of two conditionals. One is obvious because it starts with the word "if", and one more subtle (and troublesome) issued by the use of the word "upon". The reason to flag this use is that there is not much agreement on what the correct formal theory of conditionals is, or even on the possibility of a correct theory. A wide variety of problems emerge in the development of such theories. So much so that it led Hájek (2012b, p. 146) to conclude that "it's enough to make a philosopher turn to something easier, like solving the mind-body problem or the problem of free will." But delving into the vast amount of literature on conditionals also isn't likely to be very helpful for the present discussion. Though there is little agreement in this literature, many of the participants in the discussion presuppose possible world semantics, which is set against a Boolean background. However, the logical background is precisely what we are trying to shape here.

I have to be somewhat heuristic, then, in assessing the Born postulate. A useful start is to evaluate some short paraphrases of the postulate. Because I am interested specifically in the domain of quantum probability functions, the focus is on the specification of measurements and their outcomes. The state of the system specifies the probability function itself, rather than its domain. For this reason, the specification of the state is implicitly held constant in the background in the following three paraphrases:

BR1 If a measurement of $\mathcal{A}$ is performed, then the probability that the result lies in Δ equals $\mathrm{Tr}(P_\psi \mu_A(\Delta))$.

BR2 The probability that, if a measurement of $\mathcal{A}$ is performed, then the result lies in Δ, equals $\mathrm{Tr}(P_\psi \mu_A(\Delta))$.

BR3 The probability that a measurement outcome lies in Δ given that a measurement of $\mathcal{A}$ is performed equals $\mathrm{Tr}(P_\psi \mu_A(\Delta))$.

The first is a conditional statement in which probability occurs as a primitive term. Although formalization in this direction could be possible (e.g. with the use of probability logics (Demey, Kooi, and Sack, 2014)) such a project has little hope of fostering insight in quantum probability. The result would be a logic containing probabilistic propositions, and does not construct a domain for quantum probability functions. This is reminiscent of the situation at the end of the previous chapter. The second paraphrase concerns the probability of a conditional, and the third is a conditional probability. Generally, these are considered to be distinct things. This is because equating the two is known to lead to difficulties (see (Hájek, 2015) and references therein). So there is a real choice concerning which one to pursue.

The lazy option is to go for the easiest formulation: BR3. But a more important reason for this choice is that conditionals are not atomic propositions. Moreover, their formulation requires the introduction of at least all of the propositions that are also required when going for option BR3. While explaining this, I can already make some first steps towards the introduction of experimental propositions. There are two kinds of propositions that are useful to have:

$$M_{\mathcal{A}} \mathrel{\hat{=}} \text{``A measurement of } \mathcal{A} \text{ is performed''},$$
$$M_{\mathcal{A}}(\Delta) \mathrel{\hat{=}} \text{``}\mathcal{A} \text{ is measured and the outcome lies in } \Delta\text{''}. \tag{10.3}$$

These will be termed the *elementary experimental propositions*. It may be considered to be somewhat unsatisfactory to regard the second of these as elementary, since it is a conjunction. On the other hand, it is cumbersome, if not meaningless, to introduce propositions about measurement outcomes that do not presuppose the performance of a measurement. As long as "measurement" occurs as a primitive term, there seems to be no way around having such conjunctions as primitives as

well. With the introduction of these propositions, both options BR2 and BR3 can
be formulated as equations:

$$\begin{aligned}
\text{BR2 } & \mathbb{P}_\psi(M_\mathcal{A} \to M_\mathcal{A}(\Delta)) = \text{Tr}(P_\psi \mu_A(\Delta)), \\
\text{BR3 } & \mathbb{P}_\psi(M_\mathcal{A}(\Delta)|M_\mathcal{A}) = \text{Tr}(P_\psi \mu_A(\Delta)).
\end{aligned}$$

(10.4)

Note that the left-hand sides of these equations are not well-defined mathemat-
ically (yet). The task now is to formulate a logic in which at least the propositions
(10.3) are incorporated. This is obviously not a trivial task, and this entire chapter
is devoted to it. But it will be done in small steps, and the first step is to couple
these propositions to the formalism of quantum mechanics. This requires a classi-
fication of possible measurements and their outcomes. These are specified by the
observable and value postulate.

OP (Observable Postulate) Every observable $\mathcal{A}$ for the system S is associated
with a self-adjoint operator A acting on $\mathcal{H}$.

VaP (Value Postulate) For an observable $\mathcal{A}$ associated with the self-adjoint
operator A via OP, the set of values of possible measurement outcomes is
given by the spectrum $\sigma(A)$ of A.

These ingredients can hardly be called assumptions, as they are part and parcel
of quantum mechanics.[1] The next ingredient is a real assumption, though:

NC (Non-Contexuality) Every self-adjoint operator is associated with at most
one observable.

Although this assumption is controversial in the context of the possibility of hidden
variables, here it is more reasonable. It is an assumption of modesty. There is
nothing in quantum mechanics that indicates that the set of observables should be
more complex than the set of mathematical objects that represent them. Why then
should we not try to see if it is also sufficient for present purposes? Secondly, to go
beyond NC would similarly require assumptions about what the set of observables
should then be. No clear motivation for such assumptions can be given at this
point.[2]

---

[1]One may object that this characterization doesn't take into account the option to consider
POVMs to represent observables. This is a fair point. My response at the moment is that I believe
that the conceptual gain going from PVMs to POVMs is minimal. It would then seem masochistic
to make the situation more complicated than it already is.

[2]It deserves to be noted that adopting NC does not imply incompatibility with contextual hidden
variable models. As discussed in section 6.3, these can also be viewed as models that violate Value
Definiteness (VD) instead of NC. Contextual hidden variable models then aren't ruled out here
because VD is not assumed.

With these assumptions in place, the totality of all elementary experimental propositions can be given a set-theoretic structure. The benefit of NC for this formulation is that we no longer have to be careful to distinguish observables from operators. Thus I now switch the notation of (10.3) to M_A and $M_A(\Delta)$. The set of elementary experimental propositions is now

$$\mathcal{E}ep(\mathcal{H}) := \{M_A \mid A \in \mathcal{O}_{\mathrm{sa}}(\mathcal{H})\} \cup \left\{ M_A(\Delta) \;\middle|\; {}^{A \in \mathcal{O}_{\mathrm{sa}}(\mathcal{H}),}_{\Delta \subset \sigma(A)} \right\}. \qquad (10.5)$$

One may wonder whether the set of elementary experimental propositions should not also incorporate propositions of the form "the state of the system is ψ". However, assessing the extent to which this precisely is an *experimental* proposition is not something that can be done in a completely uncontroversial way. The main role of the quantum state is to determine probabilities for outcomes of measurements, and probabilities aren't the kinds of things that are easily and unambiguously empirically accessible.[3] Finally, as noted above, these kinds of propositions are not likely to play a part in the language that specifies the domain of quantum probability functions. Therefore, there is no reason to incorporate them in the empiricist quantum logic to be developed.

The next step is to impose more structure on the set of elementary experimental propositions. This is done by incorporating assumptions about the way these propositions logically relate to each other. The first is an idealization of the idea that measurements have outcomes.

IEA (Idealized Experimenter Assumption) Every experiment has an outcome: for every A, $M_A(\varnothing)$ denotes a contradiction, and $M_A(\sigma(A))$ is logically equivalent to M_A.

IEA introduces an equivalence relation on $\mathcal{E}ep(\mathcal{H})$. Explicitly, this relation is given by the rules

$$M_{A_1}(\varnothing) \overset{\text{IEA}}{\sim} M_{A_2}(\varnothing) \;\forall A_1, A_2,$$
$$M_A(\sigma(A)) \overset{\text{IEA}}{\sim} M_A \;\forall A, \qquad (10.6)$$

and for all other cases equivalence holds if and only if equality holds. The resulting set of equivalence classes is

$$\mathcal{E}ep(\mathcal{H})/\overset{\text{IEA}}{\sim} = \left\{ M_A(\Delta) \;\middle|\; {}^{A \in \mathcal{O}_{\mathrm{sa}}(\mathcal{H}),}_{\varnothing \subsetneq \Delta \subset \sigma(A)} \right\} \cup \{\perp\}, \qquad (10.7)$$

---

[3]Probabilistic statements are usually associated with observable relative frequencies. However, mismatches between probabilities and observed relative frequencies are allowed though 'unlikely' when they are large. It is not clear then in which sense observations of relative frequencies can be viewed as observations of probabilities. This issue concerning the empirical investigation of probabilities was emphasized by Popper (1959, §66). What the best way is to cope with this issue depends on ones interpretation of probability and philosophy of statistics (Gillies, 1990; Hájek, 2012a; Romeijn, 2014).

where $\perp$ denotes the equivalence class $\{M_A(\varnothing) \mid A \in \mathcal{O}_{sa}(\mathcal{H})\}$.

There are also less trivial relations between elementary experimental propositions. Indeed, quantum mechanics would be a rather dull theory if it didn't pose any structure on the experimental propositions. The relations can be used to impose a preorder. The precise preorder of course depends on which relations are assumed. A trivial relation is that for any A and $\Delta_1 \subset \Delta_2 \subset \sigma(A)$ the proposition $M_A(\Delta_1)$ implies the proposition $M_A(\Delta_2)$. But there are also relations among propositions concerning distinct observables and these are less trivial. A common assumption about joint measurements is the following.

EFR (Experimental Functional Relations) For every finite set of observables $\{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$ for which a joint measurement is possible, and for every real-valued Borel function f that satisfies $f(A_1, \ldots, A_n) = \mathbb{1}$ (where $A_1, \ldots, A_n$ are the corresponding self-adjoint operators), the outcomes $a_1, \ldots, a_n$ of a joint measurement satisfy $f(a_1, \ldots, a_n) = 1$.

EFR is an experimentally testable assumption that is taken to hold in quantum mechanics. The validity of this assumption is often extrapolated to cases where the observables are not jointly measured. Instead of directly measuring some observable, often another observable is measured and some suitable function is then applied to the outcome to obtain the outcome for the first observable. For example, if one is interested in the kinetic energy of a particle, it also suffices to measure its momentum, square the outcome, and rescale in accordance with the mass of the particle. Thus one may say that a measurement of the momentum of a particle also counts as a measurement of its kinetic energy. More generally, the following assumption seems legitimate.

LMR (Law-Measurement Relation) If A_1 and A_2 are two observables that can be jointly measured, and if f is a function such that for every joint measurement of A_1 and A_2 the outcomes a_1 and a_2 satisfy $a_1 = f(a_2)$ (f represents a law), then a measurement of A_2 alone also counts as a measurement of A_1 with outcome $f(a_2)$.

The assumption LMR imposes a preorder on $\mathcal{E}ep(\mathcal{H})/\overset{\text{IEA}}{\sim}$. Specifically, in terms of elementary experimental propositions, LMR has the following implications: for any pair of observables A_1, A_2 for which there is a function f such that $A_2 = f(A_1)$, and for any subset $\Delta_1 \subset \sigma(A_1)$, the proposition $M_{A_1}(\Delta_1)$ implies the proposition $M_{A_2}(f(\Delta_1))$. This notion of implication gives rise to the following preorder on $\mathcal{E}ep(\mathcal{H})/\overset{\text{IEA}}{\sim}$:

$$M_{A_1}(\Delta_1) \leq M_{A_2}(\Delta_2) \text{ iff } \exists f \text{ s.t. } A_2 = f(A_1), \Delta_2 \supset f(\Delta_1),$$
$$\perp \leq M_A(\Delta) \; \forall M_A(\Delta). \tag{10.8}$$

This preorder introduces the equivalence relation

$$M_{A_1}(\Delta_1) \overset{\text{LMR}}{\sim} M_{A_2}(\Delta_2) \text{ iff } M_{A_1}(\Delta_1) \leq M_{A_2}(\Delta_2) \text{ and } M_{A_2}(\Delta_2) \leq M_{A_1}(\Delta_1). \quad (10.9)$$

This is the case if and only if there are functions f and g such that

$$(A_2, \Delta_2) = (f(A_1), f(\Delta_1)) \text{ and } (A_1, \Delta_1) = (g(A_2), g(\Delta_2)). \quad (10.10)$$

Now, making use of the algebraic language, this in turn holds if and only if

$$\mathcal{Alg}(A_1) = \mathcal{Alg}(A_2) \text{ and } \mu_{A_1}(\Delta_1) = \mu_{A_2}(\Delta_2), \quad (10.11)$$

where $\mathcal{Alg}(A)$ denotes the Abelian von Neumann algebra generated by A.[4] Now let $\mathfrak{A}$ denote the set of all Abelian von Neumann algebras. The set of equivalence classes is then given by

$$\mathcal{EEP}(\mathcal{H}) := \{(\mathcal{A}, P) \mid \mathcal{A} \in \mathfrak{A}, \mathbb{0} \neq P \in L(\mathcal{A})\} \cup \{\perp\}, \quad (10.12)$$

i.e., $\mathcal{EEP}(\mathcal{H}) \simeq (\mathcal{Eep}(\mathcal{H})/\overset{\text{IEA}}{\sim})/\overset{\text{LMR}}{\sim}$.[5] Explicitly, the map that takes elements of $\mathcal{Eep}(\mathcal{H})$ to their representation is given by

$$M_A \mapsto \left(\mathcal{Alg}(A), \mathbb{1}\right),$$
$$M_A(\Delta) \mapsto \begin{cases} \left(\mathcal{Alg}(A), \mu_A(\Delta)\right) & \text{if } \mu_A(\Delta) \neq \mathbb{0}, \\ \perp & \text{otherwise,} \end{cases} \quad (10.13)$$

The preorder on $\mathcal{Eep}(\mathcal{H})/\overset{\text{IEA}}{\sim}$ automatically becomes a partial order on $\mathcal{EEP}(\mathcal{H})$. This partial order is given by

$$(\mathcal{A}_1, P_1) \leq (\mathcal{A}_2, P_2) \text{ iff } \mathcal{A}_1 \supset \mathcal{A}_2 \text{ and } P_1 \leq P_2,$$
$$\perp \leq (\mathcal{A}, P) \ \forall (\mathcal{A}, P). \quad (10.14)$$

The partially ordered set $\mathcal{EEP}(\mathcal{H})$ will play an important role throughout the remainder of this dissertation. Although formally every element $(\mathcal{A}, P)$ is an equivalence class of elementary experimental propositions, I will sometimes abuse language and refer to it as if it coincides with a particular element of this class. Specifically, $(\mathcal{A}, P)$ will sometimes be understood as the proposition "A is measured and the

---

[4]Note that this is well-defined because it follows from the assumption that $\mathcal{H}$ is finite-dimensional that A is bounded. Furthermore, this assumption ensures that there is a bijective correspondence between subsets Δ of the spectrum of A and the projection operators $\mu_A(\Delta)$.

[5]Equivalently, elements of $\mathcal{EEP}(\mathcal{H})$ may be viewed as pairs (R, P) where R is a resolution of the identity and P is a sum of projections in some subset of R. This may be a more fitting formulation if one wishes to generalize these results to include POVMs. Resolutions of the identity would then be replaced by sequences of positive operators that sum to the identity.

outcome lies in Δ". Furthermore, to make some of the mathematical derivations run smoother a particular notation is adopted. Although pairs of the form $(\mathcal{A}, \mathbb{0})$ are formally not elements of $\mathcal{EEP}(\mathcal{H})$, they will be used to denote the bottom element $\perp$, i.e.,

$$(\mathcal{A}, \mathbb{0}) := \perp \; \forall \mathcal{A} \in \mathfrak{A}. \tag{10.15}$$

Thus there are many ways to denote the same element $\perp$.

10.2   Lattice of elementary experimental propositions

In the previous section it was shown that the elementary experimental propositions for a system described by a Hilbert space $\mathcal{H}$ can be faithfully characterized by the partial ordered set $\mathcal{EEP}(\mathcal{H})$. As the title of this section suggests, this partial order also turns the set into a lattice: every pair of elements has a meet and a join. But just as with orthodox quantum logic, one has to take care when attempting to attach meaning to these operations. It is not obvious that they correspond to conjunctions and disjunctions of the elementary experimental propositions. In this section these lattice operations are scrutinized. There are some interesting properties, but in the end it is concluded that the join does not express a disjunction in all cases. In the next section these observations will be used to expand $\mathcal{EEP}(\mathcal{H})$ to a proper quantum logic of experimental propositions.

By making use of the notation introduced at the end of the previous section the meet on $\mathcal{EEP}(\mathcal{H})$ can be expressed as

$$(\mathcal{A}_1, P_1) \wedge (\mathcal{A}_2, P_2) = \begin{cases} \big(\mathcal{Alg}(\mathcal{A}_1, \mathcal{A}_2), P_1 \wedge P_2\big) & \text{if } \mathcal{Alg}(\mathcal{A}_1, \mathcal{A}_2) \in \mathfrak{A}, \\ \perp & \text{otherwise.} \end{cases} \tag{10.16}$$

The operation $P_1 \wedge P_2$ here denotes the meet on $L(\mathcal{H})$. Thus, whenever $P_1 \wedge P_2 = \mathbb{0}$, the meet in (10.16) corresponds to $\perp$. This meet gives a good representation of conjunction, but seeing this requires some argumentation. Suppose $M_{A_1}(\Delta_1)$ and $M_{A_2}(\Delta_2)$ are two experimental propositions whose representations correspond to $(\mathcal{A}_1, P_1)$ and $(\mathcal{A}_2, P_2)$ respectively. The conjunction of these two propositions can be paraphrased as "A_1 and A_2 are measured and the outcomes lie in Δ_1 and Δ_2 respectively." Assuming that such a joint measurement is possible only if the operators commute,[6] this proposition corresponds to a contradiction whenever the two operators do not commute. This in turn is the case if and only if the algebra $\mathcal{Alg}(\mathcal{A}_1, \mathcal{A}_2)$ is not Abelian. Thus the meet (10.16) does correspond to a conjunction whenever the propositions concern incompatible measurements, because in that case $(\mathcal{A}_1, P_1) \wedge (\mathcal{A}_2, P_2) = \perp$.

---

[6]This assumption follows from the conjunction of WCoP and CCoP from chapter 6.

When the operators A_1 and A_2 *do* commute, the algebra $\mathcal{A}lg(\mathcal{A}_1, \mathcal{A}_2)$ can be seen as the representation of an observable A_3 whose measurement counts as a measurement of both A_1 and A_2 (according to LMR). Conversely, a joint measurement of A_1 and A_2 counts as a measurement of A_3. That is, $\mathcal{A}lg(A_1, A_2) = \mathcal{A}lg(A_3)$ if and only if there exist functions f_1, f_2, f_3 such that $f_1(A_3) = A_1$, $f_2(A_3) = A_2$ and $f_3(A_1, A_2) = A_3$. In short, if a joint measurement of observables associated with the algebras $\mathcal{A}_1$ and $\mathcal{A}_2$ is possible, then it corresponds to a measurement of some observable associated with the algebra $\mathcal{A}lg(\mathcal{A}_1, \mathcal{A}_2)$. Thus with respect to the specification of such measurements the meet again behaves as a conjunction. What is left to check is that the meet also behaves correctly with respect to the specification of measurement outcomes. That this is indeed the case can be seen by considering again three observables A_1, A_2, A_3 as above. The conjunction of $M_{A_1}(\Delta_1)$ and $M_{A_2}(\Delta_2)$ precisely corresponds to the proposition $M_{A_3}(\Delta_3)$ if $\Delta_3 = f_1^{-1}(\Delta_1) \cap f_2^{-1}(\Delta_2)$, which is the case if and only if $\mu_{A_3}(\Delta_3) = \mu_{A_1}(\Delta_1) \wedge \mu_{A_2}(\Delta_2) = P_1 \wedge P_2$.

The meet on $\mathcal{EEP}(\mathcal{H})$ thus neatly coincides with what one expects for conjunctions of elementary experimental propositions. For the join, however, the situation is a lot more complicated. As a first step I restrict attention to equivalent measurements, i.e., $\mathcal{A}_1 = \mathcal{A}_2$. In this case the join actually behaves quite decently. One then has that

$$(\mathcal{A}, P_1) \vee (\mathcal{A}, P_2) = (\mathcal{A}, P_1 \vee P_2). \tag{10.17}$$

Correspondingly, the disjunction of $M_A(\Delta_1)$ and $M_A(\Delta_2)$ paraphrases neatly to $M_A(\Delta_1 \cup \Delta_2)$.

The situation becomes more difficult when considering distinct measurements. In general, the join is given by

$$(\mathcal{A}_1, P_1) \vee (\mathcal{A}_2, P_2) = \left(\mathcal{A}_1 \cap \mathcal{A}_2, \bigwedge \{ P \in L(\mathcal{A}_1 \cap \mathcal{A}_2) \mid P_1 \vee P_2 \leq P \} \right). \tag{10.18}$$

This is not a very insightful equation, and so it is better to look at a specific example. Consider two self-adjoint operators A_1, A_2 and their corresponding algebras $\mathcal{A}_1, \mathcal{A}_2$. Call these two operators (and their corresponding algebras) *totally incompatible*[7] when $\mathcal{A}_1 \cap \mathcal{A}_2 = \mathbb{C}\,\mathbb{1}$. For such a pair of operators one has

$$(\mathcal{A}_1, \mathbb{1}) \vee (\mathcal{A}_2, \mathbb{1}) = (\mathbb{C}\,\mathbb{1}, \mathbb{1}). \tag{10.19}$$

The right-hand side of this equation is the top element of $\mathcal{EEP}(\mathcal{H})$. This means that, if the join is interpreted as a disjunction, the proposition "A_1 is measured or A_2 is measured" is identified with a tautology. But of course one can consider many occasions when a measurement of neither A_1 nor A_2 is performed. Thus the join is unsatisfactory as a representation of disjunction.

---

[7]This definition is related to the notion of totally incompatible frames as used in chapter 7: two Abelian algebras are totally incompatible if and only if they are subalgebras of algebras generated by totally incompatible frames.

This example can further be used to demonstrate that $\mathcal{EEP}(\mathcal{H})$ is not distributive. Considered a third algebra $\mathcal{A}_3$ that is totally incompatible with both $\mathcal{A}_1$ and $\mathcal{A}_2$. Then

$$(\mathcal{A}_3, \mathbb{1}) \wedge ((\mathcal{A}_1, \mathbb{1}) \vee (\mathcal{A}_2, \mathbb{1})) = (\mathcal{A}_3, \mathbb{1}) \neq \bot \\ = ((\mathcal{A}_3, \mathbb{1}) \wedge (\mathcal{A}_1, \mathbb{1})) \vee ((\mathcal{A}_3, \mathbb{1}) \wedge (\mathcal{A}_2, \mathbb{1})). \quad (10.20)$$

So, much like orthodox quantum logic, the lattice $\mathcal{EEP}(\mathcal{H})$ does not provide a logic that fits well with natural language. There are no elements in $\mathcal{EEP}(\mathcal{H})$ that provide good candidates for disjunctions of elementary experimental propositions. These have to be added manually, which is done in the next section.

10.3 Intuitionistic quantum logic

The present situation is reminiscent of that in the beginning of chapter 9. We have a lattice of experimental propositions that isn't rich enough to also express disjunctions of these propositions. The key difference is that now we are working with an interpretation of these propositions in natural language, which will be the guide to expanding the lattice. So while it may be tempting to apply the theory of Bruns and Lakser (1970) to $\mathcal{EEP}(\mathcal{H})$, this is not what will be done here. The reason for this is that although this method is guaranteed to provide a Heyting algebra,[8] it is conceptually unclear what this method would establish. Indeed, there is no obvious interpretation of a distributive ideal in $\mathcal{EEP}(\mathcal{H})$. That being said, it is of course possible to construct the injective hull of $\mathcal{EEP}(\mathcal{H})$ and the reader interested in what this construction amounts to may consult (Hermens, 2014, p. 59).

The appropriate completion of $\mathcal{EEP}(\mathcal{H})$ is obtained almost immediately by formulating clearly what is desired. That is, specifying what precisely is meant with an expansion comes almost confusingly close to providing a proof of its existence. Thus it is required to carefully articulate what in fact is desired, as well as what has to be shown to achieve it. Let me start with setting out some demands. What is required in the end is the following:[9]

  (i) A distributive lattice $(\mathcal{IL}(\mathcal{H}), \preceq, \curlyvee, \curlywedge)$,

  (ii) An injective map $i_{IL} : \mathcal{EEP}(\mathcal{H}) \to \mathcal{IL}(\mathcal{H})$ that is a meet-semilattice homomorphism.

Injectivity of i_{IL} warrants that $\mathcal{IL}(\mathcal{H})$ is a proper extension of $\mathcal{EEP}(\mathcal{H})$, and the homomorphism assumption warrants that the meets in $\mathcal{EEP}(\mathcal{H})$ (which can properly

---

[8]In fact, it will be a Boolean algebra since $\mathcal{EEP}(\mathcal{H})$ is atomistic.

[9](Spoiler alert!) As the defense in section 8.4 makes clear, it would not be surprising to find that intuitionistic logic can play a role in an empiricist approach to quantum mechanics. The lattice $\mathcal{IL}(\mathcal{H})$ will in fact turn out to be a proper Heyting algebra, hence its name.

be understood as conjunctions) are preserved. Distributivity is a necessary condition for the demand that $\curlyvee$ and $\curlywedge$ can be interpreted as disjunction and conjunction. Because $\mathcal{EEP}(\mathcal{H})$ itself is not distributive, this implies that in general the equation[10]

$$i_{IL}\left((\mathcal{A}_1, P_1) \vee (\mathcal{A}_2, P_2)\right) = i_{IL}\left(\mathcal{A}_1, P_1\right) \curlyvee i_{IL}\left(\mathcal{A}_2, P_2\right), \tag{10.21}$$

will not hold. But for the special cases where $\mathcal{A}_1 = \mathcal{A}_2$ it was found that the join in $\mathcal{EEP}(\mathcal{H})$ does express a disjunction. So for these special cases i_{IL} is required to preserve joins. That is, it is required that

$$i_{IL}\left((\mathcal{A}, P_1) \vee (\mathcal{A}, P_2)\right) = i_{IL}\left(\mathcal{A}, P_1\right) \curlyvee i_{IL}\left(\mathcal{A}, P_2\right). \tag{10.22}$$

It deserves to be emphasized that these are just desiderata and that, for example, (10.22) is still meaningless at the moment because neither i_{IL} nor $\curlyvee$ has been defined yet. But these constraints will form the basis for the construction of $IL(\mathcal{H})$. Afterwards, when i_{IL} and $(IL(\mathcal{H}), \preceq, \curlyvee, \curlywedge)$ have been defined, it has to be shown that they satisfy the desiderata.

The next important ingredient makes use of the functions $S_{(\mathcal{A}, P)} : \mathfrak{A} \to L(\mathcal{H})$ that, for fixed $(\mathcal{A}, P)$, are defined as[11]

$$S_{(\mathcal{A},P)}(\mathcal{A}') := \begin{cases} P & \text{if } \mathcal{A} \subset \mathcal{A}', \\ \mathbb{0} & \text{otherwise.} \end{cases} \tag{10.23}$$

It has the property that $\left(\mathcal{A}', S_{(\mathcal{A},P)}(\mathcal{A}')\right) \leq (\mathcal{A}, P)$ for all $\mathcal{A}' \in \mathfrak{A}$, and equality holds in the case $\mathcal{A}' = \mathcal{A}$. These functions then allows us to rewrite every element $(\mathcal{A}, P) \in \mathcal{EEP}(\mathcal{H})$ as an infinite join over all Abelian algebras. Specifically, in $\mathcal{EEP}(\mathcal{H})$ the relation

$$(\mathcal{A}, P) = \left(\bigvee_{\substack{\mathcal{A}' \in \mathfrak{A}, \\ \mathcal{A} \subset \mathcal{A}'}} (\mathcal{A}', P) \right) \vee \left(\bigvee_{\substack{\mathcal{A}' \in \mathfrak{A}, \\ \mathcal{A} \not\subset \mathcal{A}'}} (\mathcal{A}', \mathbb{0}) \right) = \bigvee_{\mathcal{A}' \in \mathfrak{A}} \left(\mathcal{A}', S_{(\mathcal{A},P)}(\mathcal{A}')\right) \tag{10.24}$$

holds for all $(\mathcal{A}, P)$. By combining these observations one obtains

$$i_{IL}\left(\mathcal{A}, P\right) = i_{IL}\left(\mathcal{A}, S_{(\mathcal{A},P)}(\mathcal{A})\right) \leq \bigcurlyvee_{\mathcal{A}' \in \mathfrak{A}} i_{IL}\left(\mathcal{A}', S_{(\mathcal{A},P)}(\mathcal{A}')\right)$$

$$\leq i_{IL}\left(\bigvee_{\mathcal{A}' \in \mathfrak{A}} \left(\mathcal{A}', S_{(\mathcal{A},P)}(\mathcal{A}')\right) \right) = i_{IL}\left(\mathcal{A}, P\right). \tag{10.25}$$

---

[10]For notational clarity I write $i_{IL}\left(\mathcal{A}, P\right)$ instead of $i_{IL}\left((\mathcal{A}, P)\right)$ when there is no risk for confusion.

[11]Here, and in the remainder of this chapter, $\mathcal{A}'$ does not denote the commutant of $\mathcal{A}$. The prime is merely used as a notational convenience to avoid fiddling with subscripts.

This implies that the inequalities should actually be equalities. In particular

$$i_{IL}(\mathcal{A}, P) = \curlyvee_{\mathcal{A}' \in \mathfrak{A}} i_{IL}\left(\mathcal{A}', S_{(\mathcal{A},P)}(\mathcal{A}')\right). \tag{10.26}$$

Thus from the condition that i_{IL} preserves the order alone the constraint (10.26) is obtained. Combining this with the condition (10.22), one can now obtain an expression for joins of elementary experimental propositions in $IL(\mathcal{H})$:

$$
\begin{aligned}
&i_{IL}(\mathcal{A}_1, P_1) \curlyvee i_{IL}(\mathcal{A}_2, P_2) \\
&= \left(\curlyvee_{\mathcal{A}' \in \mathfrak{A}} i_{IL}\left(\mathcal{A}', S_{(\mathcal{A}_1,P_1)}(\mathcal{A}')\right)\right) \curlyvee \left(\curlyvee_{\mathcal{A}'' \in \mathfrak{A}} i_{IL}\left(\mathcal{A}'', S_{(\mathcal{A}_2,P_2)}(\mathcal{A}'')\right)\right) \\
&= \curlyvee_{\mathcal{A}' \in \mathfrak{A}} \left(i_{IL}\left(\mathcal{A}', S_{(\mathcal{A}_1,P_1)}(\mathcal{A}')\right) \curlyvee i_{IL}\left(\mathcal{A}', S_{(\mathcal{A}_2,P_2)}(\mathcal{A}')\right)\right) \\
&= \curlyvee_{\mathcal{A}' \in \mathfrak{A}} i_{IL}\left(\mathcal{A}', S_{(\mathcal{A}_1,P_1)}(\mathcal{A}') \vee S_{(\mathcal{A}_2,P_2)}(\mathcal{A}')\right).
\end{aligned}
\tag{10.27}
$$

The crucial thing to note is that both $i_{IL}(\mathcal{A}, P)$ and $i_{IL}(\mathcal{A}_1, P_1) \curlyvee i_{IL}(\mathcal{A}_2, P_2)$ have now been characterized by a function $S : \mathfrak{A} \to L(\mathcal{H})$. The first is characterized by the function $S_{(\mathcal{A},P)}$ and the second by the function defined by the action

$$S_{(\mathcal{A}_1,P_1,\mathcal{A}_2,P_2)}(\mathcal{A}) := S_{(\mathcal{A}_1,P_1)}(\mathcal{A}) \vee S_{(\mathcal{A}_2,P_2)}(\mathcal{A}) \; \forall \mathcal{A} \in \mathfrak{A}, \tag{10.28}$$

where the join on the right-hand side is the one from $L(\mathcal{H})$. This gives the inspiration to think of $IL(\mathcal{H})$ as a set of functions $S : \mathfrak{A} \to L(\mathcal{H})$ where the embedding i_{IL} is given by

$$i_{IL}(\mathcal{A}, P) := S_{(\mathcal{A},P)}. \tag{10.29}$$

The final details are given by the following theorem and its proof.

Theorem 10.1. *Formally adding disjunctions to the collection of elementary experimental propositions while respecting (10.22) leads to the complete distributive lattice*

$$IL(\mathcal{H}) := \left\{ S : \mathfrak{A} \to L(\mathcal{H}) \;\middle|\; \begin{array}{c} S(\mathcal{A}) \in L(\mathcal{A}), \\ S(\mathcal{A}_1) \le S(\mathcal{A}_2) \; whenever \; \mathcal{A}_1 \subset \mathcal{A}_2 \end{array} \right\} \tag{10.30}$$

with partial order defined as

$$S_1 \preceq S_2 \; iff \; S_1(\mathcal{A}) \le S_2(\mathcal{A}) \forall \mathcal{A} \in \mathfrak{A}, \tag{10.31}$$

giving rise to the join and meet

$$
\begin{aligned}
(S_1 \curlyvee S_2)(\mathcal{A}) &= S_1(\mathcal{A}) \vee S_2(\mathcal{A}), \\
(S_1 \curlywedge S_2)(\mathcal{A}) &= S_1(\mathcal{A}) \wedge S_2(\mathcal{A}).
\end{aligned}
\tag{10.32}
$$

The injective map $i_{IL} : \mathcal{EEP}(\mathcal{H}) \to IL(\mathcal{H})$ given by (10.29) is a meet-semilattice homomorphism.

Proof. The fact that $I\!L(\mathcal{H})$ is a complete distributive lattice follows from the fact that $L(\mathcal{A})$ is a complete distributive lattice for every $\mathcal{A} \in \mathfrak{A}$. The details of this observation are left to the reader.

The map i_{IL} is injective since $S_{(\mathcal{A}_1,P_1)} = S_{(\mathcal{A}_2,P_2)}$ if and only if $(\mathcal{A}_1, P_1) = (\mathcal{A}_2, P_2)$. To see that it preserves the partial order consider $(\mathcal{A}_1, P_1) \leq (\mathcal{A}_2, P_2)$, then for every $\mathcal{A} \in \mathfrak{A}$

$$S_{(\mathcal{A}_1,P_1)}(\mathcal{A}) = \begin{cases} P_1 \leq P_2 = S_{(\mathcal{A}_2,P_2)}(\mathcal{A}) & \mathcal{A}_1 \subset \mathcal{A}, \\ \mathbb{0} \leq S_{(\mathcal{A}_2,P_2)}(\mathcal{A}) & \mathcal{A}_1 \not\subset \mathcal{A}. \end{cases} \tag{10.33}$$

Furthermore, i_{IL} respects meets because

$$\bigwedge_{i \in I} S_{(\mathcal{A}_i,P_i)}(\mathcal{A}) = \begin{cases} \bigwedge_{i \in I} P_i & \text{if } \mathcal{A}_i \subset \mathcal{A} \; \forall i \in I, \\ \mathbb{0} & \text{otherwise} \end{cases} = S_{\bigwedge_{i \in I}(\mathcal{A}_i,P_i)}(\mathcal{A}). \tag{10.34}$$

The fact that the join respects (10.22) is basically a consequence of the construction of $I\!L(\mathcal{H})$. Explicitly, this now follows because for every $\mathcal{A}' \in \mathfrak{A}$

$$\left(S_{(\mathcal{A},P_1)} \curlyvee S_{(\mathcal{A},P_2)}\right)(\mathcal{A}') = S_{(\mathcal{A},P_1)}(\mathcal{A}') \vee S_{(\mathcal{A},P_2)}(\mathcal{A}') = S_{(\mathcal{A},P_1 \vee P_2)}(\mathcal{A}'). \tag{10.35}$$

Finally, it needs to be shown that $I\!L(\mathcal{H})$ contains nothing more than disjunctions and conjunctions of elementary experimental propositions. This is shown by proving that every $S \in I\!L(\mathcal{H})$ can be written as a join of elementary experimental propositions. Specifically, one has

$$S = \bigcurlyvee_{\mathcal{A} \in \mathfrak{A}} S_{(\mathcal{A},S(\mathcal{A}))}. \tag{10.36}$$

To see this, first note that

$$S_{(\mathcal{A},S(\mathcal{A}))}(\mathcal{A}') = \begin{cases} S(\mathcal{A}) & \mathcal{A} \subset \mathcal{A}', \\ \mathbb{0} & \mathcal{A} \not\subset \mathcal{A}'. \end{cases} \tag{10.37}$$

Now for any $\mathcal{A}' \in \mathfrak{A}$

$$\left(\bigcurlyvee_{\mathcal{A} \in \mathfrak{A}} S_{(\mathcal{A},S(\mathcal{A}))}\right)(\mathcal{A}') = \bigvee_{\mathcal{A} \in \mathfrak{A}} S_{(\mathcal{A},S(\mathcal{A}))}(\mathcal{A}') = \bigvee_{\substack{\mathcal{A} \in \mathfrak{A}, \\ \mathcal{A} \subset \mathcal{A}'}} S(\mathcal{A}), \tag{10.38}$$

Because $S(\mathcal{A}) \leq S(\mathcal{A}')$ whenever $\mathcal{A} \subset \mathcal{A}'$ (this follows from the definition of $I\!L(\mathcal{H})$), the final join in (10.38) is actually equal to $S(\mathcal{A}')$. This proves (10.36). $\qquad\square$

The lattice $\mathit{IL}(\mathcal{H})$ now consists of all possible disjunctions and conjunctions of elementary experimental propositions. This means that it is rich enough to allow a formulation of the Born rule of the type of option BR3 in (10.4), and allows the investigation of rules for a probability calculus. In fact, in chapter 11 it will become apparent that for a certain class of probability functions that may be defined on the lattice, this class is rich enough to model all quantum probability functions. But the philosophical investigation of this result of course runs much smoother when one has a better physical understanding of $\mathit{IL}(\mathcal{H})$. The remainder of this chapter is devoted to obtaining this understanding.

The first occurrence of the lattice $\mathit{IL}(\mathcal{H})$ in the literature that I know of is in the work of Caspers et al. (2009) and a further investigation of this logic was performed by Heunen, Landsman, and Spitters (2012). It is quite remarkable and reassuring that these authors arrived at the same logic by following an entirely different route, namely, a topos-theoretical reformulation of quantum mechanics. For me, this is useful as it provides a more sturdy mathematical basis. It was already noted by Caspers et al. (2009) that the lattice is turned into a Heyting algebra by obtaining the relative pseudo-complement in the usual way:

$$S_1 \to S_2 := \bigvee \left\{ S \in \mathit{IL}(\mathcal{H}) \, ; \, S \wedge S_1 \leq S_2 \right\}. \tag{10.39}$$

Furthermore, negation is introduced by $\neg S := S \to \bot$, where the bottom element is given by

$$\bot(\mathcal{A}) = \mathbb{0} \quad \forall \mathcal{A} \in \mathfrak{A}, \tag{10.40}$$

thus obtaining an intuitionistic logic for reasoning with experimental propositions. The logic $\mathit{IL}(\mathcal{H})$ is a proper Heyting algebra and even is radically non-classical: the law of excluded middle only holds for the top and bottom element. This is easily checked by observing that $S(\mathbb{C}\,\mathbb{1}) = \mathbb{1}$ if and only if $S = \top$, where the top element is given by

$$\top(\mathcal{A}) = \mathbb{1} \quad \forall \mathcal{A} \in \mathfrak{A}. \tag{10.41}$$

It then follows that $S \vee \neg S = \top$ if and only if either $S = \top$ or $\neg S = \top$. The interpretation of negation will be discussed later on in this section.

On a more philosophical level it is harder to make a comparison with this earlier occurrence of $\mathit{IL}(\mathcal{H})$. Because a purely formal route is followed, Caspers et al. (ibid.) find themselves in a situation that is reminiscent of the one I encountered at the end of chapter 9. Indeed, although it was claimed that $\mathit{IL}(\mathcal{H})$ is "the correct quantum logic" (ibid., p. 741), the physical motivation for this claim remained obscure. Also in (Heunen, Landsman, and Spitters, 2012) the logic is "proposed" (postulated?) rather than derived, and the interpretation remained vague:[12]

---

[12]The notation of symbols has been altered to coincide with the present discussion.

> we regard each $S \in I\!L(\mathcal{H})$ as a single proposition as far as logical struc-
> ture is concerned; physically, S breaks down into a family $\{S(\mathcal{A})\}_{\mathcal{A} \in \mathfrak{A}}$.
> This could either mean that one invents a question for each context $\mathcal{A}$
> separately (compatible with the monotonicity in (10.30)), or that one
> constructs such a family from a single proposition [projection operator]
> in the sense of von Neumann. (Heunen, Landsman, and Spitters, 2012,
> pp. 725–726)

The idea to view S as a family of propositions is linked to Bohr's ideas concerning the necessity for a specification of (classical) experimental contexts for unambiguous communication. However, for this link to work, another ingredient of Bohr's ideas is also mandatory: the demand for ordinary language. This requires that either S itself, or each of the elements $S(\mathcal{A})$ in the family can be given a translation into natural language. No such translation is provided. But what I find most confusing is that the suggestion for such a translation appears to presuppose orthodox quantum logic. After all, $\{S(\mathcal{A})\}_{\mathcal{A} \in \mathfrak{A}}$ is a family of projection operators, and each term in the family is supposed to be a proposition. In my view this approach turns the arrow of explanation in the wrong direction. One motivation for me to develop a new quantum logic came from an inability to interpret orthodox quantum logic. This logic that was found to be confusing and problematic. What's more, Heunen, Landsman, and Spitters (ibid.) appear to agree (at least in part) with this view as they speak of "the lure of [orthodox] quantum logic". It is strange then to see that precisely this logic is invoked to attempt an interpretation of $I\!L(\mathcal{H})$.

It deserves to be noted that the program of the topos-theoretical reformulation of quantum mechanics doesn't stand or fall with their reading of $I\!L(\mathcal{H})$. Nor is it clear whether this logic is as important to this program as the two cited papers could lead one to believe. But the critique of their interpretation is useful for posing a contrast with the one adopted here. In my own interpretation of $I\!L(\mathcal{H})$, each element S is a *disjunction* rather than a *family* of elementary experimental propositions. Moreover, this is not just a disjunction over the projections $\{S(\mathcal{A})\}_{\mathcal{A} \in \mathfrak{A}}$, but over the pairs $\{(\mathcal{A}, S(\mathcal{A}))\}_{\mathcal{A} \in \mathfrak{A}}$. In fact, von Neumann's idea to consider projections as representations of experimental propositions *independently* of the specification of experimental contexts seems to me precisely the deviation from Bohr's ideas that leads to the mess of orthodox quantum logic. It is only when considering a pair $(\mathcal{A}, P)$ that one can attach unambiguous meaning to the projection operator.[13]

This interpretation thus poses much weight on the requirement to specify a measurement. This observation can be used to understand the strong non-classical

---

[13]This subconscious forgetting about he measurement context is ubiquitous in quantum logic and occurs even among the greater thinkers in the field. As an example, consider the work of Hughes (1982) who takes great care to introduce experimental propositions as *pairs* (A, Δ) requiring both the specification of the observable and the outcome set, and then without blinking identifies this with the projection $\mu_A(\Delta)$, as if there were no loss in this translation.

behavior of $IL(\mathcal{H})$. The logic is entirely built up from propositions about *performed* measurements. The negation of $M_A(\Delta)$ is identified with other propositions about measurements that exclude the possibility of $M_A(\Delta)$. Specifically,

$$\neg S_{(\mathcal{A},P)} = \bigvee \left\{ S_{(\mathcal{A}',P')} \mid [\mathcal{A}',\mathcal{A}] \neq \mathbb{0} \text{ or } P' \wedge P = \mathbb{0} \right\}. \tag{10.42}$$

The propositions in this disjunction are either of the form $M_{A'}$ for some A' that is incompatible with A, or they are compatible with $M_A(\Delta')$ with $\Delta' \cap \Delta = \varnothing$ (i.e., the measurement outcomes are incompatible). But all these propositions leave open the option of no experiment having been performed. An unexcluded middle thus presents itself as the proposition $\neg M_A =$ "A is not measured", which does not correspond to any element of $IL(\mathcal{H})$. An explicit example illustrates these points further.

Example 10.1. Consider a spin-$\frac{1}{2}$ particle and restrict attention to two possible measurements, namely, its spin along the x-axis and around the z-axis. The sublattice of relevant projection operators and the partially ordered set of relevant algebras are given by the Hasse diagrams

$$\text{(10.43)}$$

where $\sigma_x = \frac{1}{2}P_x^+ - \frac{1}{2}P_x^-$, $\sigma_z = \frac{1}{2}P_z^+ - \frac{1}{2}P_z^-$ and $\mathcal{A}_x := \mathcal{Alg}(\sigma_x)$ and $\mathcal{A}_z := \mathcal{Alg}(\sigma_z)$. The function $S_{(\mathcal{A}_x,P_x^+)}$ represents the proposition expressing that σ_x is measured and the spin is found to be in the up-direction. One now has

$$\neg S_{(\mathcal{A}_x,P_x^+)} = S_{(\mathcal{A}_x,P_x^-)} \vee S_{(\mathcal{A}_z,\mathbb{1})}. \tag{10.44}$$

Explicitly, the actions of these functions are given by the following table.

	$\mathbb{C}\mathbb{1}$	$\mathcal{A}_x$	$\mathcal{A}_z$
$S_{(\mathcal{A}_x,P_x^+)}$	$\mathbb{0}$	P_x^+	$\mathbb{0}$
$S_{(\mathcal{A}_x,P_x^-)}$	$\mathbb{0}$	P_x^-	$\mathbb{0}$
$S_{(\mathcal{A}_z,\mathbb{1})}$	$\mathbb{0}$	$\mathbb{0}$	$\mathbb{1}$
$\neg S_{(\mathcal{A}_x,P_x^+)}$	$\mathbb{0}$	P_x^-	$\mathbb{1}$

$$\text{(10.45)}$$

The function $\neg S_{(\mathcal{A}_x,P_x^+)}$ expresses that either σ_x has been measured and the spin was found to be in the down-direction, or σ_z was measured. Both these options of course imply that $S_{(\mathcal{A}_x,P_x^+)}$ is not true, but both also imply that some measurement has in fact been performed. $\triangle$

I suggested in (Hermens, 2013a, §2) that, by incorporating propositions that explicitly mention that a particular measurement is not performed, one could obtain a classical logic. The motivation for not doing it there was that "many more propositions would have to be added to make the logic classical and most of them are rather dull." This is certainly true. I have never met a physicist enthusiastically telling everyone that no measurement has been performed. But in the current situation developing such a logic may actually be beneficial, as I explain below.

The hope is that in the end a quantum logic is constructed that can be used to obtain a better physical understanding of the formalism of quantum probability. Ideally, the situation would be that one has a general conception of what probability functions on a propositional lattice are, *independently of quantum mechanics.* One could then motivate what a probability function on $IL(\mathcal{H})$ should look like, and then compare it to the Born rule. A derivation of the Born rule would then follow from philosophical considerations that select out the quantum probability functions from all possible probability functions. Perhaps by a rationality argument, or perhaps just simply because every probability function on $IL(\mathcal{H})$ already satisfies the Born rule, or perhaps with the aid of an additional metaphysical assumption. In any case, one first has to know what a probability function on $IL(\mathcal{H})$ is. Now, although there have been attempts to characterize probability functions for intuitionistic logic,[14] a clear, easily acceptable, and widely accepted set of axioms for such functions on Heyting algebras is not available. When it comes to Boolean algebras, on the other hand, much more theoretical background is available and there is more agreement on what probability functions are. Then, if the incorporation of a few dull propositions leads to a Boolean algebra, this entire body of knowledge is at our disposal. This is enough reason to expand $IL(\mathcal{H})$, which is done in the next section.

10.4   Classical quantum logic

As suggested in the previous section, the Heyting algebra $IL(\mathcal{H})$ can be turned into a Boolean algebra $CL(\mathcal{H})$ by adding the propositions $\neg M_A$. One way to do this is to go back to the beginning of this chapter and start anew with expanding the set of elementary experimental propositions, now including these negations. Such an approach was roughly followed in (Hermens, 2014). But this requires taking some steps that are only justified in retrospect after $CL(\mathcal{H})$ has been defined. Here I follow a different route that I think is conceptually more appealing.[15] The idea is to stay close to the lattice $IL(\mathcal{H})$ and tweak it until the desired result obtains. That is, I start with the functions $S \in IL(\mathcal{H})$ (whose interpretation should now clear),

---

[14]See (Van Fraassen, 1981; Morgan and Leblanc, 1983a,b; Roeper and Leblanc, 1999; Weatherson, 2003).

[15]It may be interesting to note that the two routes also lead to distinct formulations of $CL(\mathcal{H})$. The fact that the two are actually equivalent is shown by Theorem 11.1.

and then tweak these functions till they may be taken to represent negations. This implies that the obtained lattice $CL(\mathcal{H})$ will properly contain $IL(\mathcal{H})$, and the task is to find out what constraints a function $S : \mathfrak{A} \to L(\mathcal{H})$ should satisfy in order to count as an element of $CL(\mathcal{H})$.

A good starting point is to consider the constraints that define the set $IL(\mathcal{H})$. The requirement of monotonicity is that $\mathcal{A}_1 \subset \mathcal{A}_2$ implies $S(\mathcal{A}_1) \leq S(\mathcal{A}_2)$. In the proof of Theorem 10.1 this feature arises from the demand that all propositions in $IL(\mathcal{H})$ should be obtained by forming disjunctions and conjunctions of elementary experimental propositions. Because the elementary experimental propositions already satisfy monotonicity, this feature is carried over to all other propositions. But what is the conceptual meaning of monotonicity for these elementary experimental propositions?

The answer to this question is found by taking a closer look at the functions that represent elementary experimental propositions. Consider an elementary experimental proposition $M_A(\Delta)$ and its representative $S_{(\mathcal{A},P)}$. This representative satisfies

$$S_{(\mathcal{A},P)} = \bigvee_{\substack{\mathcal{A}' \in \mathfrak{A}, \\ \mathcal{A} \subset \mathcal{A}'}} S_{(\mathcal{A}',P)}. \tag{10.46}$$

This embodies the idea that the proposition that A is measured does not exclude the proposition that an observable A' is measured for which there exists an f such that $A = f(A')$. Now consider the special case where $P = \mathbb{1}$. If one reads $S_{(\mathcal{A},\mathbb{1})}$ explicitly as a disjunction, it paraphrases to "A is measured or some more fine-grained measurement is performed." But when also allowing propositions that refer to the not-performing of measurements, it should also be possible to formulate propositions of the form[16]

$$M_A^!(\Delta) \mathrel{\hat{=}} \begin{array}{l} \text{``}A \text{ is measured and the outcome lies in } \Delta, \\ \quad \text{and no finer grained measurement is performed''}. \end{array} \tag{10.47}$$

This proposition effectively cancels out all disjuncts in (10.46) except for the one concerning $\mathcal{A}$. If there is to be a function $S : \mathfrak{A} \to L(\mathcal{H})$ that can represent this proposition, then the natural candidate is the function $S_{(\mathcal{A},P)}^! : \mathfrak{A} \to L(\mathcal{H})$ defined by

$$S_{(\mathcal{A},P)}^!(\mathcal{A}') = \begin{cases} P & \mathcal{A}' = \mathcal{A}, \\ \mathbb{0} & \mathcal{A}' \neq \mathcal{A}, \end{cases} \tag{10.48}$$

where again $\mathcal{A} = \mathcal{A}lg(A)$ and $P = \mu_A(\Delta)$.

Note that the function $S_{(\mathcal{A},P)}^!$ is not an element of $IL(\mathcal{H})$ precisely because it does not obey the monotonicity condition. The simplest suggestion one can then

---

[16]The use of the exclamation mark is a reference to its similar use in "∃!" meaning "there exists a unique". Here too the exclamation mark singles out one specific observable.

make for $\mathcal{CL}(\mathcal{H})$ is to just drop the monotonicity assumption. That is, I propose the lattice

$$\mathcal{CL}(\mathcal{H}) := \{S : \mathfrak{A} \to L(\mathcal{H}) \mid S(\mathcal{A}) \in L(\mathcal{A})\}, \tag{10.49}$$

with partial order precisely the one of $\mathcal{IL}(\mathcal{H})$:

$$S_1 \preceq S_2 \text{ iff } S_1(\mathcal{A}) \leq S_2(\mathcal{A}) \forall \mathcal{A} \in \mathfrak{A}. \tag{10.50}$$

I shall now demonstrate that this proposal is actually spot on.

It is not hard to see that the meet and join on $\mathcal{CL}(\mathcal{H})$ are again given by the equations

$$\begin{aligned}
(S_1 \curlyvee S_2)(\mathcal{A}) &= S_1(\mathcal{A}) \vee S_2(\mathcal{A}), \\
(S_1 \curlywedge S_2)(\mathcal{A}) &= S_1(\mathcal{A}) \wedge S_2(\mathcal{A}).
\end{aligned} \tag{10.51}$$

Consequently, $\mathcal{IL}(\mathcal{H})$ is a sublattice of $\mathcal{CL}(\mathcal{H})$. This shows the formal compatibility of elements of $\mathcal{IL}(\mathcal{H})$ with the newly added functions. But it is of course also required that the earlier interpretation of these elements is compatible with the interpretation of the new elements. This is not entirely trivial. The two subsets $\mathcal{IL}(\mathcal{H})$ and $\left\{ S^!_{(\mathcal{A},P)} \,\middle|\, (\mathcal{A}, P) \in \mathcal{EEP}(\mathcal{H}) \right\}$ have independently been given a translation in natural language, the first using the rules from (10.3) and the second with the rule (10.47). Furthermore, the meet and join are required to be read as conjunction and disjunction respectively, and this implies that the elements of the sublattice generated by $\left\{ S^!_{(\mathcal{A},P)} \,\middle|\, (\mathcal{A}, P) \in \mathcal{EEP}(\mathcal{H}) \right\}$ also have been given an interpretation using (10.47). In short, there are many elements of $\mathcal{CL}(\mathcal{H})$ that have been given a double interpretation, and it needs to be shown that these just refer to paraphrases in natural language.

It turns out that showing consistency of the two interpretations is a lot easier than it sounds, especially once one knows which steps to take. The first step is noticing that the new interpretation (10.47) amounts to a translation into natural language of *all* elements of $\mathcal{CL}(\mathcal{H})$. That is, the sublattice generated by $\left\{ S^!_{(\mathcal{A},P)} \,\middle|\, (\mathcal{A}, P) \in \mathcal{EEP}(\mathcal{H}) \right\}$ is in fact $\mathcal{CL}(\mathcal{H})$ itself. This is because for every $S \in \mathcal{CL}(\mathcal{H})$ one has

$$S = \curlyvee_{\mathcal{A} \in \mathfrak{A}} S^!_{(\mathcal{A}, S(\mathcal{A}))}. \tag{10.52}$$

Consequently, every element of $\mathcal{IL}(\mathcal{H})$ has been given an interpretation by (10.47). To show that this interpretation is consistent with the one of the previous sections, it suffices to restrict attention to the elementary experimental propositions corresponding to the functions of the form $S_{(\mathcal{A},P)}$. This is because $\mathcal{IL}(\mathcal{H})$ is the sublattice generated by the set $\left\{ S_{(\mathcal{A},P)} \,\middle|\, (\mathcal{A}, P) \in \mathcal{EEP}(\mathcal{H}) \right\}$. But for these elementary experimental propositions compatibility of the two interpretations follows immediately

from the fact that

$$S_{(\mathcal{A},P)} = \curlyvee_{\substack{\mathcal{A}' \in \mathfrak{A}, \\ \mathcal{A} \subset \mathcal{A}'}} S^!_{(\mathcal{A}',P)} \ \forall (\mathcal{A}, P) \in \mathcal{EEP}(\mathcal{H}). \tag{10.53}$$

The $\mathit{IL}(\mathcal{H})$-interpretation of the left-hand side is just the proposition $M_A(\Delta)$. This can be paraphrased as "some measurement that is at least as fine-grained as A is performed and the outcome is given by Δ." This in turn can be rewritten as a disjunction over all measurements A' that are at least as fine-grained as A like in (10.46). It follows from this that exactly one of these A' is measured and no other one. So one of the propositions represented by $S^!_{(\mathcal{A}',P)}$ must be the case. Since it is not specified which of these A' is the case, one must take the disjunction over all possibilities. This then gives the right-hand side of (10.53).

Thus far it has only been established that $\mathit{IL}(\mathcal{H})$ has been expanded so as to also incorporate propositions of the form (10.47). The new lattice of propositions $\mathcal{CL}(\mathcal{H})$ is closed under conjunctions and disjunctions, but I haven't said anything yet about negations. So what still has to be established is that $\mathcal{CL}(\mathcal{H})$ has complements, that these complements can be understood as expressing negations and that the lattice is in fact Boolean. The first and third task are quite easy to establish. The unary operation $S \mapsto S^c$ defined by

$$S^c(\mathcal{A}) := (S(\mathcal{A}))^\perp \ \forall \mathcal{A} \in \mathfrak{A}. \tag{10.54}$$

satisfies

$$S \curlyvee S^c = \top, \ S \curlywedge S^c = \bot \tag{10.55}$$

for all $S \in \mathcal{CL}(\mathcal{H})$ and thus turns the lattice into a Boolean algebra. So what is left to show is that this complement is consistent with the use of negations of the natural language interpretation of $\mathcal{CL}(\mathcal{H})$.

To understand the complement on $\mathcal{CL}(\mathcal{H})$ it is good to start with a simple case. The first thing to check is that the complement of a proposition M_A can be paraphrased as "A has not been measured". The proposition M_A is represented by the function $S_{(\mathcal{A},\mathbb{1})}$ and its complement satisfies

$$S^c_{(\mathcal{A},\mathbb{1})}(\mathcal{A}') = \begin{cases} \mathbb{0} & \mathcal{A} \subset \mathcal{A}', \\ \mathbb{1} & \mathcal{A} \not\subset \mathcal{A}' \end{cases} = \begin{cases} \mathbb{0} & \mathcal{A} \subset \mathcal{A}', \\ \mathbb{1} & \mathcal{A}' \subsetneq \mathcal{A}, \\ \mathbb{1} & [\mathcal{A}, \mathcal{A}'] \neq 0, \end{cases} \tag{10.56}$$

where $[\mathcal{A}, \mathcal{A}'] \neq 0$ indicates that the two algebras are generated by non-commuting observables. This implies that

$$S^c_{(\mathcal{A},\mathbb{1})} = \left(\curlyvee_{\substack{\mathcal{A}' \in \mathfrak{A}, \\ \mathcal{A}' \subsetneq \mathcal{A}}} S^!_{(\mathcal{A}',P)} \right) \curlyvee \left(\curlyvee_{\substack{\mathcal{A}'' \in \mathfrak{A}, \\ [\mathcal{A}'', \mathcal{A}] \neq 0}} S_{(\mathcal{A}'',P)} \right). \tag{10.57}$$

The right-hand side can be read as "either a measurement is performed that is compatible with A, but that is strictly coarser, or a measurement is performed that is incompatible with a measurement of A." The intended reading of the left-hand side is "A is not measured". Clearly the reading of the right-hand side implies the reading of the left-hand side. The converse is not entirely trivial though. At first sight one option appears to be missing on the right-hand side, namely, the case where no measurement is performed at all. The idea is, however, that this case has been included as 'the coarsest possible measurement': $S^!_{(\mathbb{C}\mathbb{1},\mathbb{1})}$. A measurement of the observable $\mathbb{1}$ necessarily yields the value 1 (remember IEA). It is a trivial measurement. Then $S^!_{(\mathbb{C}\mathbb{1},\mathbb{1})}$ expresses that no non-trivial measurement is performed. The only thing then left to argue for is that a trivial measurement is not really a measurement at all. Indeed, since the outcome is fixed, one doesn't have to do anything to obtain the it. The experimenter just needs to think of the number 1 and the deed is done.

Thus far it has been established that $S^c_{(\mathcal{A},\mathbb{1})}$ is a good representation of $\neg M_A$. For the next step, consider an arbitrary elementary experimental proposition. In this case one has

$$S^c_{(\mathcal{A},P)} = S^c_{(\mathcal{A},\mathbb{1})} \curlyvee S_{(\mathcal{A},P^\perp)}. \tag{10.58}$$

The right-hand side indicates the possible scenarios that either no measurement of A is performed at all, or that a measurement is performed whose outcome is in Δ^c instead of Δ. This is again consistent with the intended meaning of the left-hand side. For arbitrary $S \in I\!L(\mathcal{H})$, consistency follows from the previous cases together with the use of de Morgan's laws as

$$S^c = \left(\curlyvee_{\mathcal{A} \in \mathfrak{A}} S_{(\mathcal{A}, S(\mathcal{A}))} \right)^c = \curlywedge_{\mathcal{A} \in \mathfrak{A}} S^c_{(\mathcal{A}, S(\mathcal{A}))}. \tag{10.59}$$

This establishes that the complement can indeed be interpreted as a negation.

It should be clear that $C\!L(\mathcal{H})$ incorporates all propositions from $I\!L(\mathcal{H})$ as well as all negations of these propositions, and is closed under forming disjunctions, conjunctions and negations. As a final point, one may wonder whether it is also the smallest expansion that establishes this. After all, the starting ingredient was that $C\!L(\mathcal{H})$ also incorporates all the propositions of the form $M^!_A(\Delta)$. One can wonder if these would also have been part of the Boolean algebra if instead we had started with incorporating all propositions of the form $\neg M_A(\Delta)$ instead. The answer to this question is yes, and for the proof one only has to observe that

$$S^!_{(\mathcal{A},P)} = S_{(\mathcal{A},P)} \curlywedge \left(\curlywedge_{\substack{\mathcal{A}' \in \mathfrak{A}, \\ \mathcal{A} \subsetneq \mathcal{A}'}} S^c_{(\mathcal{A}',P)} \right). \tag{10.60}$$

This establishes that every proposition of the form $M_A^!(\Delta)$ can be written as a conjunction of propositions of the forms $M_A(\Delta)$ and $\neg M_A(\Delta)$. Thus by introducing (classical) negations to $I\!L(\mathcal{H})$ one automatically introduces the propositions of the form $M_A^!(\Delta)$ as well.

It may be clear that, although $I\!L(\mathcal{H})$ is a sub-lattice of $C\!L(\mathcal{H})$, it is not a sub-Heyting algebra. In particular, for any $S \in I\!L(\mathcal{H})$, the equation

$$\neg S = S^c \tag{10.61}$$

holds if and only if $S = \top$ or $S = \bot$. This is seen by noting that in all other cases $S^c \notin I\!L(\mathcal{H})$. Indeed, suppose $S \in I\!L(\mathcal{H})$ and $S \notin \{\bot, \top\}$, then $S(\mathbb{C}\mathbb{1}) = \mathbb{0}$ and so $S^c(\mathbb{C}\mathbb{1}) = \mathbb{1}$. Because $S \neq \bot$, there is some $\mathcal{A}$ such that $S(\mathcal{A}) = P \neq \mathbb{0}$. Consequently, $S^c(\mathcal{A}) = P^\perp \neq \mathbb{1}$, which shows that S^c does not satisfy the monotonicity requirement. Intuitively, the complement in $C\!L(\mathcal{H})$ is weaker than the one in $I\!L(\mathcal{H})$. The complement $\neg S$ lists all possible measurements and outcomes that ensure a conflict with S, while S^c in addition lists the options of not performing measurements. An example illustrates this point further.

Example 10.2. Consider again the spin-$\frac{1}{2}$ particle from Example 10.1 with the two possible measurements $\sigma_x = \frac{1}{2}P_x^+ - \frac{1}{2}P_x^-$, $\sigma_z = \frac{1}{2}P_z^+ - \frac{1}{2}P_z^-$. The function $S_{(\mathcal{A}_x, P_x^+)}$ represents the proposition expressing that σ_x is measured and the spin is found to be in the up-direction. The complement and the pseudo-complement for this proposition are listed in the following table:

	$\mathbb{C}\mathbb{1}$	$\mathcal{A}_x$	$\mathcal{A}_z$
$S_{(\mathcal{A}_x, P_x^+)}$	$\mathbb{0}$	P_x^+	$\mathbb{0}$
$\neg S_{(\mathcal{A}_x, P_x^+)}$	$\mathbb{0}$	P_x^-	$\mathbb{1}$
$S_{(\mathcal{A}_x, P_x^+)}^c$	$\mathbb{1}$	P_x^-	$\mathbb{1}$

$$\tag{10.62}$$

This shows that

$$S_{(\mathcal{A}_x, P_x^+)}^c = \neg S_{(\mathcal{A}_x, P_x^+)} \curlyvee S_{(\mathbb{C}\mathbb{1}, \mathbb{1})}^!, \tag{10.63}$$

where $S_{(\mathbb{C}\mathbb{1}, \mathbb{1})}^!$ again expresses the proposition that no measurement is performed.
$\triangle$

With the construction of $C\!L(\mathcal{H})$ we now have a classical quantum logic that is fit to express all experimental propositions one finds in quantum mechanics. This in itself may be considered a remarkable achievement, since one sometimes finds slogans expressing the incompatibility of classical logic with quantum mechanics.[17] It may be seen to be an explicit vindication of Bohr's claim that "all departures

---

[17]Feynman's "logical tightrope" may be considered such an example. On the other hand, I haven't come across any thorough defense for this alleged incompatibility.

from common language and ordinary logic are entirely avoided by reserving the word "phenomenon" solely for reference to unambiguously communicable information, in the account of which the word "measurement" is used in its plain meaning of standardized comparison" (1958). Apart from that, the construction of $\mathcal{CL}(\mathcal{H})$ also shows that quantum mechanics without probabilities already provides a non-trivial structure for experimental propositions. In the next chapter, then, I introduce probability functions acting on these propositions. It is shown that the structure is rich enough to incorporate all quantum probability functions, but also allows other probability functions. I end with an investigation of further constraints that may be used to single out the Born rule.

11

Quantum probability for empiricists

In this chapter the new quantum logic developed in the previous chapter is used to give a reformulation of quantum probability. In section 11.1 the logic is first reformulated so as to obtain a measurable space: the common structure for formalisms of probability. In section 11.2 the theory of conditional probability spaces is used to demonstrate that quantum probability functions can be modeled as conditional probability functions on the Boolean algebra

$$\mathcal{CL}(\mathcal{H}) = \{S : \mathfrak{A} \to L(\mathcal{H}) \mid S(\mathcal{A}) \in L(\mathcal{A})\} . \tag{11.1}$$

This means that this quantum logic provides a physical interpretation of the domain of quantum probability functions as experimental propositions.

But apart from a proper empirical reformulation of the formalism of quantum probability one could hope for more. In section 11.2 it is shown that not every conditional probability function on $\mathcal{CL}(\mathcal{H})$ gives rise to a quantum probability function. It would be interesting to see if there is a conceptual reason why quantum probability functions play a privileged role. In section 11.3 it is shown that the quantum probability functions precisely coincide with the so-called non-contextual probability functions. This characterization of quantum probability functions is common in the literature and in the remainder of the section I discuss three possible motivations for the importance of non-contextuality from the perspectives of Everettian quantum mechanics, Bub & Pitowsky's information-theoretical interpretation, and Quantum Bayesianism, respectively. None of these motivations are found to be entirely satisfactory though. In the final section I develop a new approach to characterizing the Born rule that focuses less on the importance of non-contextuality. Formally, this new approach delivers what it promises, but there are also some philosophical wrinkles in the carpet that will be highlighted. The chapter ends with a short selective afterthought on the accomplishments of part III in section 11.5.

11.1 Preparing for probabilities

Classical probability functions are often defined as functions acting on a Boolean algebra. Also, with $CL(\mathcal{H})$ we now have a Boolean algebra of propositions for quantum experiments. It then seems straightforward to put one and one together and investigate what probability functions on $CL(\mathcal{H})$ look like. But a little more patience goes a long way and it is useful to first do a bit more research on $CL(\mathcal{H})$ itself.

In many expositions of probability, the Boolean algebra that is the domain of the probability functions is an algebra of subsets of some set. The classical quantum logic $CL(\mathcal{H})$, on the other hand, is not of this kind. This is nothing to be frowned upon, as this is no indication that an underlying set would be 'missing' in some sense. After all, it follows from Stone's theorem that such an underlying set always exists. But although it may be tempting to use this theorem to construct this set, it is useful to take a less abstract approach. There is a 'natural' underlying set $\Omega_{\mathcal{H}}$ for $CL(\mathcal{H})$, yet Stone's theorem is not the way to find it. This is not to say that the result obtained by Stone's method (which is the set of all ultrafilters on $CL(\mathcal{H})$ endowed with the Stone topology) is not susceptible to interpretation, but just that such an interpretation doesn't arise naturally.

The construction of a particular underlying set for $CL(\mathcal{H})$ can be motivated by looking at the atoms in $CL(\mathcal{H})$. These are the maximally informative propositions in the lattice, and they are given by the set of functions

$$CL_{\mathrm{a}}(\mathcal{H}) := \left\{ S^!_{(\mathcal{A},P)} \in CL(\mathcal{H}) \,\middle|\, P \in L_{\mathrm{a}}(\mathcal{A}) \right\}. \tag{11.2}$$

Remember that $S^!_{(\mathcal{A},P)}$ is given by

$$S^!_{(\mathcal{A},P)}(\mathcal{A}') = \begin{cases} P & \mathcal{A} = \mathcal{A}', \\ \mathbb{0} & \text{otherwise}, \end{cases} \tag{11.3}$$

and expresses the proposition that A is measured (with $\mathcal{A} = \mathit{Alg}(A)$) and an outcome in Δ is obtained (with $P = \mu_A(\Delta)$) and that no finer grained measurement is performed. Furthermore, $L_{\mathrm{a}}(\mathcal{A})$ denotes the set of atoms in $L(\mathcal{A})$ and so the projections in $L_{\mathrm{a}}(\mathcal{A})$ correspond to the most precise possible specifications of measurement outcomes. When $\mathcal{H}$ is finite-dimensional, these atoms can be used to formulate any proposition in $CL(\mathcal{H})$ (i.e., the lattice is atomistic). This follows from the fact that (10.52) establishes that every element in $CL(\mathcal{H})$ can be written as a disjunction of propositions of the form $S^!_{(\mathcal{A},P)}$, while these in turn can be written as disjunctions of elements in $CL_{\mathrm{a}}(\mathcal{H})$. Specifically, for every proposition of the form $S^!_{(\mathcal{A},P)}$ one has

$$S^!_{(\mathcal{A},P)} = \bigvee \left\{ S^!_{(\mathcal{A},P')} \in CL_{\mathrm{a}}(\mathcal{H}) \,\middle|\, P' \le P \right\}. \tag{11.4}$$

The set of atoms $CL_a(\mathcal{H})$ may be seen to be characterized by the set

$$\Omega_\mathcal{H} := \{(\mathcal{A}, P) \in \mathcal{EEP}(\mathcal{H}) \mid P \in L_a(\mathcal{A})\}, \tag{11.5}$$

where $\mathcal{EEP}(\mathcal{H})$ is given by (10.12). The set $\Omega_\mathcal{H}$ may be taken to be the underlying set for $CL_a(\mathcal{H})$:[1]

Theorem 11.1. *Let $\mathcal{H}$ be a finite-dimensional Hilbert space and let $\mathcal{F}_\mathcal{H} := \mathcal{P}(\Omega_\mathcal{H})$ be the Boolean algebra of all subsets of $\Omega_\mathcal{H}$, where this set is given by (11.5). Then $\mathcal{F}_\mathcal{H}$ and $CL(\mathcal{H})$ are isomorphic as Boolean algebras.*

Proof. The proof consists of showing that

$$f : CL(\mathcal{H}) \to \mathcal{F}_\mathcal{H}, \;\; f(S) := \{(\mathcal{A}, P) \in \Omega_\mathcal{H} \mid P \leq S(\mathcal{A})\} \tag{11.6}$$

is an isomorphism of Boolean algebras.

First note that

$$\begin{aligned} f(\bot) &= \{(\mathcal{A}, P) \in \Omega_\mathcal{H} \mid P \leq \mathbb{0}\} = \varnothing, \\ f(\top) &= \{(\mathcal{A}, P) \in \Omega_\mathcal{H} \mid P \leq \mathbb{1}\} = \Omega_\mathcal{H}. \end{aligned} \tag{11.7}$$

The fact that the map respects arbitrary joins and meets is shown by focusing on the Boolean lattices $L(\mathcal{A})$.

$$\begin{aligned} f\left(\bigvee_{i \in I} S_i\right) &= \bigcup_{\mathcal{A} \in \mathfrak{A}} \left\{(\mathcal{A}, P) \in \Omega_\mathcal{H} \;\middle|\; P \leq \left(\bigvee_{i \in I} S_i\right)(\mathcal{A})\right\} \\ &= \bigcup_{\mathcal{A} \in \mathfrak{A}} \left\{(\mathcal{A}, P) \in \Omega_\mathcal{H} \;\middle|\; P \leq \bigvee_{i \in I} S_i(\mathcal{A})\right\} \\ &= \bigcup_{\mathcal{A} \in \mathfrak{A}} \bigcup_{i \in I} \{(\mathcal{A}, P) \in \Omega_\mathcal{H} \mid P \leq S_i(\mathcal{A})\} \\ &= \bigcup_{i \in I} \bigcup_{\mathcal{A} \in \mathfrak{A}} \{(\mathcal{A}, P) \in \Omega_\mathcal{H} \mid P \leq S_i(\mathcal{A})\} = \bigcup_{i \in I} f(S_i). \end{aligned} \tag{11.8}$$

---

[1] Note that $\Omega_\mathcal{H}$ corresponds to the set of principal ultrafilters on $CL(\mathcal{H})$. This would be the same result as provided by Stone's theorem if there are no free ultrafilters on $CL(\mathcal{H})$. I do not know, however, if such ultrafilters exist.

$$
\begin{aligned}
f\left(\bigwedge_{i\in I} S_i\right) &= \bigcup_{\mathcal{A}\in\mathfrak{A}}\left\{(\mathcal{A},P)\in\Omega_\mathcal{H}\,\middle|\,P\le\left(\bigwedge_{i\in I}S_i\right)(\mathcal{A})\right\}\\
&= \bigcup_{\mathcal{A}\in\mathfrak{A}}\left\{(\mathcal{A},P)\in\Omega_\mathcal{H}\,\middle|\,P\le\bigwedge_{i\in I}S_i(\mathcal{A})\right\} \\
&= \bigcup_{\mathcal{A}\in\mathfrak{A}}\bigcap_{i\in I}\{(\mathcal{A},P)\in\Omega_\mathcal{H}\mid P\le S_i(\mathcal{A})\}\\
&= \bigcap_{i\in I}\bigcup_{\mathcal{A}\in\mathfrak{A}}\{(\mathcal{A},P)\in\Omega_\mathcal{H}\mid P\le S_i(\mathcal{A})\}=\bigcap_{i\in I}f(S_i).
\end{aligned}
$$

(11.9)

Thus f is a homomorphism.

To see that f is a bijection I prove that the function

$$
g:\mathcal{F}_\mathcal{H}\to\mathcal{CL}(\mathcal{H}),\ (g(\Delta))(\mathcal{A}):=\bigvee\{P\in L_\mathrm{a}(\mathcal{A})\mid(\mathcal{A},P)\in\Delta\} \tag{11.10}
$$

provides an inverse for f. For any $S\in\mathcal{CL}(\mathcal{H})$ and $\Delta\subset\Omega_\mathcal{H}$ one has

$$
\begin{aligned}
(g(f(S)))(\mathcal{A}) &= \bigvee\{P\in L_\mathrm{a}(\mathcal{A})\mid(\mathcal{A},P)\in f(S)\}\\
&= \bigvee\{P\in L_\mathrm{a}(\mathcal{A})\mid P\le S(\mathcal{A})\}=S(\mathcal{A}),
\end{aligned}
$$

(11.11)

And conversely,

$$
\begin{aligned}
f(g(\Delta)) &= \{(\mathcal{A},P)\in\Omega_\mathcal{H}\mid P\le(g(\Delta))(\mathcal{A})\}\\
&= \left\{(\mathcal{A},P)\in\Omega_\mathcal{H}\,\middle|\,P\le\bigvee\{P\in L_\mathrm{a}(\mathcal{A})\mid(\mathcal{A},P)\in\Delta\}\right\}\\
&= \{(\mathcal{A},P)\in\Omega_\mathcal{H}\mid(\mathcal{A},P)\in\Delta\}=\Delta.
\end{aligned}
$$

(11.12)

$\square$

This theorem establishes that $\mathcal{CL}(\mathcal{H})$ can be identified with the measurable space $(\Omega_\mathcal{H},\mathcal{F}_\mathcal{H})$. Now one could evaluate all probability measures on this space. But doing so is quite unsatisfactory. These probability measures in general don't do justice to how quantum mechanics is applied or, more generally, how one reasons about measurements and their outcomes in any scientific theory. This is most effectively illustrated by an example.

Consider the element $(\mathbb{C}\mathbb{1},\mathbb{1})\in\Omega_\mathcal{H}$. The corresponding singleton set is identified with the element $S^!_{(\mathbb{C}\mathbb{1},\mathbb{1})}\in\mathcal{CL}(\mathcal{H})$. This element represents the proposition that no measurement is performed (see the discussion on page 167). A possible probability measure is the one that assigns probability one to this proposition, and probability zero to all the other elements of $\Omega_\mathcal{H}$. Clearly this probability function cannot

be used to assign probabilities to measurement outcomes for any of the possible measurements, since it is not possible to condition on a set of measure zero.

Probability functions of this kind are unsatisfactory even for the hardened instrumentalist. Although one may adhere to the idea that "unperformed experiments have no result" (Peres, 1978), a central idea in science is that, if these experiments were to be performed, they would have an outcome. Although one may reject the idea that these outcomes have a definite value when the experiment is not performed, it is a lot more radical to deny the possibility of assigning probabilities to the possible outcomes. When, for example, one thinks of probabilities as degrees of belief for a rational agent, it is perfectly natural that these are also defined for measurements that aren't performed. After all, one can contemplate what would happen if a measurement were to be performed. But also when one thinks of probabilities as propensities or dispositions that are properties of a system, it is perfectly natural to assume that these properties exist also for unperformed measurements precisely because these properties are dispositional. In each of these cases the existence of non-trivial conditional probabilities is presupposed and these aren't provided by many of the possible probability measures on $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}})$.

A way to get around these difficulties is to take conditional probability as primitive. On such an approach a probability distribution over possible measurement outcomes for each possible measurement is well-defined. Therefore, such a framework is adopted in the next section.

11.2 Recovering quantum probability

The idea that conditional probability is prior to regular probability is not uncommon in the philosophy of probability. For a defense of this view one may consult the work of Hájek (2003). Noticeably, he even presents quantum mechanics as an example of a theory in which conditional probability is primitive (p. 305): "Quantum mechanics apparently tells us that certain chances, conditional on free acts, are defined, and it even purports to tell us their values." Although it is easy to more or less understand what he means with this, it deserves to be noted that his construal of quantum probability does not reflect the way it is usually formulated. Quantum probability functions are usually defined as one-placed functions, and this is a tradition I have followed throughout this dissertation. To be sure, the paraphrasing of the Born postulate as a postulate on conditional probabilities I gave in section 10.1 is a natural one. But it does not pose in itself a mathematical formalism of conditional probability for quantum mechanics. Nor am I familiar with such formalizations other than the one given in this section. If this formulation catches on, then Hájek's argument will have a more solid foundation, although one would still have to do some work to justify his presupposition that quantum probabilities are chances.

The distinction between conditional and unconditional probability is mostly debated within the philosophy of probability. It is somewhat unfortunate that this leads to the two options as somehow being opposing views; a defense of conditional probability is often posed as a critique of unconditional probability, as if there were something fundamentally wrong with the standard axioms of probability in which Kolmogorov defines conditional probability in terms of unconditional probabilities. But when a mathematician defines x in terms of y, this almost never should be taken to imply that one concept is more fundamental than the other. With respect to this issue it deserves to be noted that Kolmogorov (1963) in a later text explicitly starts with a notion of probability that only makes sense conditional on an assumed set of conditions. He then arrives at the axioms he is famous for by keeping the set of conditions fixed. That being said, one of course still needs a formal definition of conditional probability that also works when the conditions have probability zero. Such a definition of conditional probability spaces was first given by Rényi (1955).[2]

Definition 11.1. Let $(\Omega, \mathcal{F})$ be a measurable space and let $\mathcal{C} \subset \mathcal{F}\backslash\{\varnothing\}$ be a non-empty set. Then $(\Omega, \mathcal{F}, \mathcal{C}, \mathbb{P})$ is called a *conditional probability space (CPS)* if $\mathbb{P} : \mathcal{F} \times \mathcal{C} \to [0, 1]$ satisfies

(i) for every $C \in \mathcal{C}$ the function $A \mapsto \mathbb{P}(A|C)$ is a probability measure on $(\Omega, \mathcal{F})$,

(ii) for all $A, B, C \in \mathcal{F}$ with $C, B \cap C \in \mathcal{C}$

$$\mathbb{P}(A \cap B|C) = \mathbb{P}(A|B \cap C)\,\mathbb{P}(B|C). \tag{11.13}$$

The set $\mathcal{C}$ is called the *set of conditions*. If this set is closed under finite unions the space is called an *additive CPS* and if $\mathcal{C} = \mathcal{F}\backslash\{\varnothing\}$ it is called a *full CPS*.

I take it that this definition is relatively uncontroversial so I shall adopt it for my investigation without defending it. In the previous section, the quantum logic $\mathcal{CL}(\mathcal{H})$ was related to the measurable space $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}})$. Although I adopt this space to evaluate possible conditional probability spaces, I shall refer to $\mathcal{F}_{\mathcal{H}}$ as if it were identical (and not just isomorphic) to $\mathcal{CL}(\mathcal{H})$. This means that the elements of $\mathcal{F}_{\mathcal{H}}$ are being thought of both as functions on $\mathfrak{A}$ and as subsets of $\Omega_{\mathcal{H}}$. As a modest set of conditions for quantum mechanics, it is appropriate to take the propositions that express that a measurement of the observable $\mathcal{A}$ is performed. These are the elements of the form

$$S_{\mathcal{A}} := S_{(\mathcal{A}, \mathbb{1})}, \ \mathcal{A} \in \mathfrak{A} \tag{11.14}$$

---

[2]Independently, Popper (1959, A *ii-*v) gave a similar axiomatization of conditional probability. For this reason, conditional probability spaces are also often called Rényi-Popper spaces.

where $\mathcal{A}$ is the algebra generated by the operator A that is associated with the observable $\mathcal{A}$. Remember that this function was defined by the rule

$$S_{\mathcal{A}}(\mathcal{A}') := \begin{cases} \mathbb{1} & \text{if } \mathcal{A}' \subset \mathcal{A}, \\ \mathbb{0} & \text{if } \mathcal{A}' \not\subset \mathcal{A}. \end{cases} \tag{11.15}$$

Then the set of conditions is taken to be

$$\mathcal{C}_{\text{QM}} := \{S_{\mathcal{A}} \mid \mathcal{A} \in \mathfrak{A}\}. \tag{11.16}$$

With this set of conditions, every CPS $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}}, \mathcal{C}_{\text{QM}}, \mathbb{P})$ provides a distribution over the outcomes for each possible measurement. To see this, let $\mathcal{A} \in \mathfrak{A}$ represent a possible measurement and let $\{P_1, \ldots, P_n\}$ be any set of pairwise orthogonal projections in $L(\mathcal{A})$ that sum to the identity. Then

$$\mathbb{P}\left(S_{(\mathcal{A}, P_1)} \middle| S_{\mathcal{A}}\right) + \ldots + \mathbb{P}\left(S_{(\mathcal{A}, P_n)} \middle| S_{\mathcal{A}}\right) = 1. \tag{11.17}$$

Joint measurements are also incorporated in this formulation whenever they are measurements of observables that are jointly measurable, since

$$S_{\mathcal{A}_1} \wedge S_{\mathcal{A}_2} = \begin{cases} S_{\mathcal{Alg}(\mathcal{A}_1, \mathcal{A}_2)} & \text{if } \mathcal{Alg}(\mathcal{A}_1, \mathcal{A}_2) \in \mathfrak{A}, \\ \bot & \text{otherwise.} \end{cases} \tag{11.18}$$

One may have the intuition that it would be more satisfactory to consider full CPSs for $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}})$. After all, what is the use of taking conditional probability as primitive if not all propositions can be used as conditions? The fact is that I do not know at the moment if there exists a full CPS for $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}})$. It can be shown, though, that in some cases a CPS $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}}, \mathcal{C}_{\text{QM}}, \mathbb{P})$ can be extended to an additive CPS (for a proof see (Hermens, 2014, pp. 61–62)). On the other hand, I am not sure what the advantage is of also allowing propositions like "A is measured, or no measurement at all is performed" as a condition. What *is* conceptually advantageous, is that we now have a plausible framework that introduces probabilities in quantum mechanics as a natural ingredient rather than as a brute postulate. What I mean is that, given the logic $\mathcal{CL}(\mathcal{H})$, it is a natural step to introduce functions that assign probabilities to its propositions. This may be contrasted to the orthodox quantum logic $L(\mathcal{H})$ for which the introduction of probability functions is less natural because it is unclear how to interpret its elements as propositions (see also chapter 8). The domain of ordinary quantum probability functions is the lattice of projection operators, and the only motivation for calling these operators "events" seems to be the Born postulate itself. These abstract events have now been replaced by a clear Boolean algebra of experimental propositions. It is now time to connect the two formalisms of quantum probability.

Definition 11.2. Let $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}}, \mathcal{C}_{\mathrm{QM}}, \mathbb{P})$ be a CPS. Then $\mathbb{P}$ is said to be a *Born measure* if there exists a density operator ρ such that for all $\mathcal{A} \in \mathfrak{A}$ and $P \in L(\mathcal{A})$

$$\mathbb{P}\left(S_{(\mathcal{A},P)}|S_{\mathcal{A}}\right) = \mathrm{Tr}(\rho P). \tag{11.19}$$

It is easy to see that there is a CPS $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}}, \mathcal{C}_{\mathrm{QM}}, \mathbb{P}_\rho)$ for every density operator ρ such that $\mathbb{P}_\rho$ is a Born measure that reproduces the Born rule for ρ, namely,

$$\mathbb{P}_\rho\left(S|S_{\mathcal{A}}\right) := \mathrm{Tr}(\rho S(\mathcal{A})). \tag{11.20}$$

This equation can be understood as follows. Remember that every function $S \in \mathcal{CL}(\mathcal{H})$ can be understood as expressing a disjunction over possible measurements because,

$$S = \bigYdown_{\mathcal{A} \in \mathfrak{A}} S^!_{(\mathcal{A}, S(\mathcal{A}))} \tag{11.21}$$

holds for every S with the right-hand side defined by (10.48). The condition $S_{\mathcal{A}}$ now specifies which of the measurements in the disjunction is in fact performed. The set of possible outcomes for this measurement specified by S is represented by the projection $S(\mathcal{A})$ since

$$S(\mathcal{A}) = \left(\bigYdown_{\mathcal{A}' \in \mathfrak{A}} S^!_{(\mathcal{A}', S(\mathcal{A}'))}\right)(\mathcal{A}) = \bigvee_{\mathcal{A}' \in \mathfrak{A}} S^!_{(\mathcal{A}', S(\mathcal{A}'))}(\mathcal{A}) = S^!_{(\mathcal{A}, S(\mathcal{A}))}(\mathcal{A}). \tag{11.22}$$

Thus the class of CPSs $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}}, \mathcal{C}_{\mathrm{QM}}, \mathbb{P})$ is at least as big as that of all quantum probability functions. This means that every quantum probability space can be reformulated as a conditional probability space in which it is at least conceptually clear what the elements in the domain of the function represent. The probability functions introduced as CPSs can unambiguously be seen as assigning probabilities to outcomes of measurements conditioned on the performing of said measurement. It deserves to be noted that this formalism has been achieved by adopting a purely empiricist investigation of the formalism of quantum mechanics. No additional metaphysical assumptions such as the existence of hidden variables were introduced. The conceptual underpinning of quantum probability provided may be seen to be independent of the interpretation of quantum mechanics, in the sense that it would fit with any interpretation that accepts the empirical reading of quantum mechanics adopted here.

Although this is already a very attractive result, it is worth investigating whether one can't obtain even more. The formulation of quantum probability as CPSs also allows many probability functions that are not Born measures. As an example, consider two distinct density operators ρ_1, ρ_2 such that ρ_1 is not a multiple of the unit operator and define a probability function $\mathbb{P}_{\rho_1,\rho_2}$ by the rule

$$\mathbb{P}_{\rho_1,\rho_2}(S|S_{\mathcal{A}}) := \begin{cases} \mathbb{P}_{\rho_1}(S|S_{\mathcal{A}}) & \rho_1 \in \mathcal{A}, \\ \mathbb{P}_{\rho_2}(S|S_{\mathcal{A}}) & \rho_1 \notin \mathcal{A}. \end{cases} \tag{11.23}$$

This probability function acts as the quantum probability function $\mathbb{P}_{\rho_2}$ in almost all cases except for measurement contexts that imply that a measurement of ρ_1 (now understood as an observable) is performed. In those cases the probability is calculated using the Born rule for the case that ρ_1 is the state. The constraint that $\rho_1 \neq \lambda \mathbb{1}$ ensures that there are cases for which $\rho_1 \notin \mathcal{A}$ and so $\mathbb{P}_{\rho_1,\rho_2}$ is not a Born measure.

It would be great if one could have a conceptual characterization of the special role of probability functions that follow the Born rule. The remainder of this chapter is devoted to investigating some attempts at finding such a characterization.

11.3 Non-contextuality and the Born rule

From the perspective of hidden variable interpretations it is commonly accepted that quantum mechanics is a contextual theory.[3] Somewhat counter-intuitively, at the level of probability distributions, quantum mechanics is non-contextual. The Born rule itself is non-contextual in the sense that probabilities are calculated using only the projection operator that specifies the measurement outcome, rather than the complete resolution of the identity specified by the measured observable. This was even taken as a guide to the formal definition of a quantum probability space in chapter 4. The fact has also been used to argue that quantum mechanics itself is non-contextual (Mermin, 1993, p. 811). In connection with more general frameworks of probability it was recognized already by Barnum et al. (2000) that non-contextuality is an important ingredient for singling out the quantum probability functions from all possible probability functions. These considerations can be made precise within the framework of the CPSs of the previous section. First off, a definition of non-contextual probability functions is required.

Definition 11.3. Let $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}}, \mathcal{C}_{\mathrm{QM}}, \mathbb{P})$ be a CPS, then $\mathbb{P}$ is called a *non-contextual probability function* if for all $\mathcal{A}_1, \mathcal{A}_2 \in \mathfrak{A}$ and all $P \in L(\mathcal{A}_1) \cap L(\mathcal{A}_2)$

$$\mathbb{P}\left(S_{(\mathcal{A}_1,P)}\big|S_{\mathcal{A}_1}\right) = \mathbb{P}\left(S_{(\mathcal{A}_2,P)}\big|S_{\mathcal{A}_2}\right). \tag{11.24}$$

The non-contextuality constraint for probability functions can be viewed as a generalization of the constraint of parameter independence adopted in derivations of Bell-type inequalities (Jarrett, 1984; Shimony, 1990). This constraint states that the probability of the outcome of a measurement is not allowed to depend on which measurement is performed at a space-like separated location. It is obtained from (11.24) by setting $\mathcal{A}_1 = \mathcal{Alg}(\mathcal{A}_A, \mathcal{A}_B)$ and $\mathcal{A}_2 = \mathcal{Alg}(\mathcal{A}_A, \mathcal{A}'_B)$ where $\mathcal{A}_A$ represents the measurement selected at one location, and $\mathcal{A}_B$ and $\mathcal{A}'_B$ represent two

---

[3]The only caveat is the finite precision loophole (see part II).

measurements that may be selected at a space-like separated location and letting $P \in L(\mathcal{A}_A)$.

The following theorem shows that the Born measures make up all the non-contextual probability functions.

Theorem 11.2. *Let $(\Omega_{\mathcal{H}}, \mathcal{F}_{\mathcal{H}}, \mathcal{C}_{\mathrm{QM}}, \mathbb{P})$ be a CPS with $\dim(\mathcal{H}) > 2$. Then $\mathbb{P}$ is a Born measure if and only if $\mathbb{P}$ is non-contextual.*

Proof. If $\mathbb{P}$ is a Born measure, then non-contextuality follows because, for all $\mathcal{A}_1, \mathcal{A}_2 \in \mathfrak{A}$ and all $P \in L(\mathcal{A}_1) \cap L(\mathcal{A}_2)$,

$$\mathbb{P}\left(S_{(\mathcal{A}_1, P)} \big| S_{\mathcal{A}_1}\right) = \mathrm{Tr}(\rho P) = \mathbb{P}\left(S_{(\mathcal{A}_2, P)} \big| S_{\mathcal{A}_2}\right). \tag{11.25}$$

The converse is slightly less trivial. Suppose $\mathbb{P}$ is a non-contextual probability function. Let $L_1(\mathcal{H})$ denote the set of 1-dimensional projections and define the function $\lambda : L_1(\mathcal{H}) \to \mathbb{R}$ as

$$\lambda(P) := \mathbb{P}\left(S_{(\mathcal{Alg}(P), P)} \big| S_{\mathcal{Alg}(P)}\right). \tag{11.26}$$

This defines a frame function. To see this let $(P_i)_{i \in I}$ be a frame. Then

$$\begin{aligned}
\sum_{i \in I} \lambda(P_i) &= \sum_{i \in I} \mathbb{P}\left(S_{(\mathcal{Alg}(P_i), P_i)} \big| S_{\mathcal{Alg}(P_i)}\right) \\
&= \sum_{i \in I} \mathbb{P}\left(S_{(\mathcal{Alg}(\{P_i \,|\, i \in I\}), P_i)} \big| S_{\mathcal{Alg}(\{P_i \,|\, i \in I\})}\right) = 1
\end{aligned} \tag{11.27}$$

According to Gleason's lemma (section 4.2) there exists a density operator ρ such that $\lambda(P) = \mathrm{Tr}(\rho P)$ for all $P \in L_1(\mathcal{H})$. Consequently,

$$\mathbb{P}\left(S_{(\mathcal{A}, P)} \big| S_{\mathcal{A}}\right) = \mathbb{P}\left(S_{(\mathcal{Alg}(P), P)} \big| S_{\mathcal{Alg}(P)}\right) = \lambda(P) = \mathrm{Tr}(\rho P), \tag{11.28}$$

for all $\mathcal{A} \in \mathfrak{A}$ and $P \in L(\mathcal{A})$. $\qquad\square$

This theorem gives a short and sharp characterization of the Born measures on $\mathcal{CL}(\mathcal{H})$. Conceptually, this brings us a step forward to understanding whether the Born rule should be adopted as a separate postulate in quantum mechanics, or whether it can be derived from more basic assumptions. The rule can now be seen to follow from any motivation of non-contextuality. This is not to say that non-contextuality is easily motivated, but at least the task has been reduced to showing why a certain constraint on probability functions should hold. Perhaps one can identify it as a rationality constraint, or motivate it by adopting a more metaphysically grounded story. The remainder of this section is devoted to a discussion of some of the motivations that exist in the literature.

As noted before, the concept of non-contextuality as applied to probability functions (instead of hidden variables) is not new. The first explicit occurrence of the idea that I know of is in the paper by Barnum et al. (2000). In a similar fashion they argue that this assumption (together with Gleason's theorem) is enough to derive the Born rule. They appear to be slightly more optimistic about the importance of this result than I am, as they conclude their paper with the statement that "It is hard to imagine a cleaner derivation of the probability rule than this." So whereas I see Theorem 11.2 as an important midway station to deriving the Born rule, the authors envision themselves to already be at the finish line.

This discrepancy likely stems from a disagreement concerning the plausibility of the non-contextuality assumption. In their own words, if a probability function is non-contextual, it "means that the probabilities are consistent with the Hilbert-space structure of the observables" (p. 1181). But I do not really know what to make of this. Apparently there is some sense in which the probability function (11.20) is consistent while (11.23) isn't, even though both are completely specified in terms of the Hilbert-space structure of observables. But which sense this should be is not found in the paper of Barnum et al. (ibid.). Somewhat ironically, the confusion may be cleared by adopting the approach of the very paper that the authors were criticizing: Deutsch's decision theoretical approach in Everettian quantum mechanics (Deutsch, 1999).

11.3.1 Everettian Quantum Mechanics

Connecting the present discussion to Everettian quantum mechanics (EQM) requires a bit more background. In this interpretation the state postulate is interpreted metaphysically. That is, the wave function is a complete representation of the actual state of the universe. This state always evolves according to the Schrödinger equation. The occurrence of actual measurement outcomes is explained by interpreting the corresponding terms in the superposition state as equally real. These terms are taken to correspond to parallel 'classical' worlds that emerge as a consequence of a decoherence process.

The precise details of this interpretation are not important here. What is important to note is that EQM claims to solve the measurement problem by arguing that in some sense all possible measurement outcomes obtain. This is often taken to lead to the problem of probability. The question is how to make sense of probabilities for different outcomes when all outcomes occur. When outcomes are assigned different probabilities, does this then mean that one occurs more than the other, or is more real?

Note that this problem is not the one I am concerned with here (why the Born rule?), but is much broader. However, the solution to the problem proposed by Deutsch, which was further developed by Wallace (2003 ; 2007 ; 2010 ; 2012), solves

both problems at the same time. The starting point of the solution is to decouple probability from the ontology of the Everettian universe, and instead attach it to the decisions made by rational agents within such a universe. The benefit of this approach is that, although at the fundamental level EQM proposes a deterministic universe that is completely described by the quantum state, rational agents live at the emergent level of semi-classical branches. It is argued that on such branches a rational agent can use decision theory just as if the universe were non-branching.[4]

This decision-theoretical turn (if it works) can be used to introduce probability into any branching universe. But the Everettian rational agent not only believes that the universe branches, but also that this occurs according to the laws of quantum mechanics. Deutsch and Wallace demonstrate that if a rational agent uses this knowledge to inform her decisions, she will act in accordance with the Born rule. In fact, if she knows the wave function pertaining to her branch, she will use this particular quantum state in the Born rule. This demonstration of course requires assumptions. The interesting thing to note for the present discussion is that non-contextuality is not among these assumptions, but is instead derived. The important distinction with the idea of Barnum et al. (2000), is that for the Everettian there is a metaphysical picture available that can be used to give meaning to a probability function being "consistent with the Hilbert-space structure". The Everettian can resort to the state postulate of quantum mechanics: there is a particular state of affairs which can inform a rational agent. This idea is adopted explicitly by Wallace (2010) by introducing the following rationality requirement.

StaS (State Supervenience) An agent's preferences between acts depend only on what physical state they actually leave his branch in.

Thus the only thing available for an agent to specify her preferences are the weights of the branches associated with the possible outcomes of an act, which are given by the Born rule. This is (of course) not the only rationality constraint that is needed to obtain the Born rule in EQM. But it is the one that lies behind non-contextuality, in the sense that violations of non-contextuality (such as (11.23)) specifically violate this constraint (Wallace, 2012, §5.8.8).

StaS seems natural enough in any interpretation of quantum mechanics in which the quantum state is assumed to give a complete ontological description. But what seems natural is not necessarily what is right, and from the point of view of decision theory StaS looks a lot less natural. It is not a kind of rationality constraint one usually finds. It may seem rational on some everyday use of the idea of 'rational', but

---

[4]What I am glossing over here is the fact that the branches in EQM are emergent structures. Thus there isn't really a spin-up branch and a spin-down branch in the case of a spin-measurements. Rather, these things are a bit woolly. So it is not entirely clear what kind of things an agent is deciding about in a universe that is governed by the quantum state. For a critical investigation the reader may consult the work of Kent (2010; 2014).

in decision theory one requires more. For a constraint to be a rationality constraint, it is often required that violating it has negative consequences (such as a sure loss on a bet). But it is not clear how an agent in an Everettian universe would be worse off when she violates StaS. This point is conceded by Deutsch, who opts for a wider understanding of rationality dubbed "scientific rationality", where your opinions should only depend on the actual state of affairs (Saunders et al., 2010, p. 403). But it seems to me that this only transforms the problem to motivating scientific rationality.

All in all though, EQM comes a long way in deriving the Born rule. But the way this is done does not excel in metaphysical parsimony. The most important assumptions used (that have played no role thus far in this part of the dissertation) are the state postulate together with an ontological interpretation of the quantum state. On top of this, an emerging structure of parallel worlds is assumed.[5] And of course there is the assumption of scientific rationality. Could these assumptions be necessary for the Born rule? Or are there other ways to flesh out the appeal to Hilbert-space structure invoked by Barnum et al. (2000)?

There are two main candidates for interpretations of quantum mechanics in which quantum states, in so far as they play a role, are given an epistemic interpretation. These are the quantum Bayesian approach (QBism) developed by Caves, Fuchs, and Schack (2002), (see also (Fuchs, 2010; Fuchs, Mermin, and Schack, 2013) for introductory approaches) and the information-theoretical interpretation of Bub and Pitowsky (2010) (see also (Bub, 2007 ; 2011b)). The main distinction between these two programs when focusing only on probability in quantum mechanics, is that the first adopts a subjective (personalist) Bayesian interpretation, whereas the second adopts an objective Bayesian interpretation. Typical for Bayesian interpretations of probability is that they are well accustomed to defending or deriving also the formalism of probability. In this case, a derivation of the Born rule would seem to be desirable. I shall now discuss these two options one at a time.

11.3.2   Bub & Pitowsky

In the pioneering work in Bayesian interpretations of quantum probability, non-contextuality is more or less taken to be a natural assumption. It is motivated by an appeal to the structure of the quantum formalism. This is the case in (Barnum et al., 2000), but also in the work of Pitowsky (2003). In his paper he attempts to give a derivation of the Born rule by making use of Dutch book arguments and rationality constraints. One of the rules (RULE 2) looks much like an instance of (11.24). It is admitted that the rule is not a rationality constraint. Instead, it is argued that "an agent who violates RULE 2 is failing to grasp the logic of the

---

[5]Although Everettians would argue that this structure is actually derived from the formalism of quantum mechanics.

gamble and wrongly assumes that she is playing a different game" (p. 400). The "logic of the gamble" refers to the claim that even though the contexts $S_{\mathcal{A}_1}$ and $S_{\mathcal{A}_2}$ are distinct, the propositions $S_{(\mathcal{A}_1, P)}$ and $S_{(\mathcal{A}_2, P)}$ refer to the same event. In other words, the correct logic is given by the lattice of projection operators $L(\mathcal{H})$. However, as seen in section 8.2, this lattice can be understood as a consequence of the Born postulate. To avoid circularity one thus needs a motivation independent of the Born postulate to argue that projection operators deserve to be called "events". But as the investigation of quantum logic in chapters 8 and 9 shows, finding such a motivation is no trivial matter.

A similar approach is found in the more recent work of Bub (2011b). Here the lattice of projections is also posited as being the non-classical event space for quantum mechanics. Why it has this form, or what is precisely meant by events is left unspecified. But a bit more is said about the interpretation of this structure. Specifically (p. 231): "the geometric structure of Hilbert space imposes objective probabilistic or information-theoretic constraints on correlations between events, just as in special relativity the geometric structure of Minkowski space imposes spatio-temporal kinematic constraints on events." This gives some more insight in why two events $M_{A_1}(\Delta_1)$ and $M_{A_2}(\Delta_2)$ that satisfy $\mu_{A_1}(\Delta_1) = \mu_{A_2}(\Delta_2)$ should be identified with each other. It is not that they actually signify the same event per se, but rather there is an objective probabilistic constraint on the correlation between these two events. There is some rule 'out there' that compels agents to assign the same credence to these two events.

It is hard to see this constraint as anything other than an invocation of the Born rule. A derivation of the Born rule postulating such a constraint again verges on circularity. But perhaps such a derivation isn't really what Bub is after. His goal rather appears to be to obtain an *objective* (or intersubjective) Bayesian interpretation of quantum probability. On such a view, two rational agents in a quantum world need not only both obey the Born rule, but they should also use the same density operator when both have the same information available. Thus there has to be some objective fact that makes one probability function more appropriate than the other. For this purpose, Bub makes an appeal to a Humean notion of chance to obtain a re-interpretation of Gleason's theorem:

> Rather, in the sense of Lewis's Principal Principle, Gleason's Theorem relates an objective feature of the world, the non-classical structure of objective chances, to the credence function of a rational agent. [...] On this analysis, the quantum state does not have an ontological significance analogous to the ontological significance of an extremal classical state as the 'truthmaker' for propositions about the occurrence and non-occurrence of events, i.e. as a representation of physical reality. Rather, the quantum state is a credence function, a book-keeping device for

keeping track of probabilities. (ibid., p. 255)

The non-contextual aspect of the Born rule then isn't deeply fundamental, metaphysically speaking. Adopting Lewis's best system analysis (Loewer, 2004), it is nothing but the best description of how events hang together. It is then a contingent fact that in this description the events $M_{A_1}(\Delta_1)$ and $M_{A_2}(\Delta_2)$ are treated identically, rather than a metaphysical necessity.

In this kind of interpretation, it is no longer fitting to ask for a derivation of the Born rule in strictly quantum mechanical terms. Instead, the problem has been shifted to a more general metaphysical issue, namely, the fleshing out of Lewis's best system analysis of physical laws. Thereby, the subject has shifted far outside the scope of this dissertation. Possibly, Bub's account of quantum probability could work well.[6] But without a fully worked out Lewisian account of the quantum formalism, it doesn't bring us much closer to an understanding of the Born rule in the present discussion. It is time, then, to turn attention to QBism.

11.3.3 QBism

The starting point of QBism is a personalist Bayesian interpretation of quantum probabilities. Quantum states are nothing more than reflections of the epistemic states of agents. But as an interpretation of quantum mechanics it adopts many more ingredients. For example, probability-1 statements are also said to have no other meaning than the reflection of credences. When a rational agent assigns probability one to a particular measurement outcome (e.g. by adopting a pure state), this does not refer to some fact in the world embodying that a measurement would in fact yield this outcome. There may not even be such facts. The world described by the QBist is indeterministic (Fuchs, 2010, p. 8) and possibly, at least on a certain level, lawless (Fuchs, 2014, p. 205).

Understandably, this interpretation is unattractive to many. But the high level of subjectivity it attributes to the formalism of quantum mechanics is also a cause for suspicion. For example, it casts doubt on the question whether quantum mechanics has any explanatory power left (Timpson, 2013, §10.2), or whether it reduces to a collection of predictions of a rational agent who just found it convenient to write them down using the Hilbert space formalism. Perhaps QBism takes its subjective view too far and thereby looses the merits of a subjective interpretation of quantum probability. Bacciagaluppi (2014), for example, wonders whether such an interpretation of probability wouldn't work better within other approaches to quantum mechanics that adopt a firmer ontology for quantum systems.

---

[6]That is not to say that the information-theoretic interpretation is without any problems. I just haven't discussed those. But see for example (Timpson, 2010) for a critical assessment of the program of Bub and Pitowsky, and (Timpson, 2013) for problems with information theoretical interpretations more generally.

Be that as it may, it is precisely the highly subjective aspect of QBism that has been put to work in justifying the Born rule. In several communications in (Fuchs, 2002b) (in particular between Caves and Fuchs) the question of the possibility of contextual probability functions in QBism was discussed. The emerging view is I think best explained in (Fuchs, 2002a, §4.1). The starting point is to note that the observable postulate does not dictate which operator is to be assigned to which observable. According to the QBist even this choice is subjective. The idea is now that the assignment of degrees of belief to the outcomes of experiments goes hand in hand with the attribution of operators to these experiments. When the agent assigns equal probabilities to two events $M_{A_1}(\Delta_1)$ and $M_{A_2}(\Delta_2)$ in all possible scenarios (notice that no operator is specified in the description of these events), then she will choose corresponding operators A_1 and A_2 such that these two events are identified with the same projection operator. The arrow of explanation is turned around, so to speak: two events are not ascribed the same probability because they are associated with the same projection, but they are associated with the same projection because they are given identical probabilities. As Fuchs (2002b, p. 20) concludes: "By this point of view, noncontextuality is a tautology–it is built in from the start. Asking why we have it is a waste of time."

Although it is indeed trivial that one obtains non-contextuality in this way, the argument is problematic. If one takes this line of reasoning further, the whole use of the theory of quantum mechanics occurs as a subjective choice. There is no obvious objective reason for a rational agent to adopt it. The only way the agent could obtain the theory is as a consequence of Bayesian conditionalization together with some further rule for updating. Such a further rule is required because conditionalization alone does not result in the stable identity of credences in distinct events. As long as the two events $M_{A_1}(\Delta_1)$ and $M_{A_2}(\Delta_2)$ have not been identified with each other, it is in general possible that future updates will lead to distinct probabilities for the events. So to introduce the equivalence of two events one needs a separate formal rule.

By making non-contextuality a tautology, the problem has just shifted up a notch. The question is no longer why a rational agent using quantum mechanics would assign non-contextual credences, but rather why a rational agent doing physics would use quantum mechanics in the first place. One would expect that an agent using quantum mechanics to describe (quantum) experiments would be better off in some sense than an agent using the outcomes of the local lottery to make predictions. Surely quantum mechanics describes *some* structure that is objectively relevant for the experiments? And wouldn't a rational agent be better off when taking this structure into consideration? The Qbist agrees. As Fuchs (ibid., p. 5) notes: "The one of us that ignores the structure of the world will be bitten by it!" But precisely what this structure is, and what kind of bite is considered here is unclear.

There is a parallel between the adherence to quantum structure by the QBist and the use of StaS in EQM. Both take it that it is in some sense better to respect it than not. Furthermore, in both cases it is unclear precisely what the penalty is for ignoring it. But there is also an important distinction. Much of the structure of quantum mechanics has been incorporated into the logic $CL(\mathcal{H})$. It dictates that an agent betting on the outcome of a spin measurement on a spin-$\frac{1}{2}$ particle should bet on one of the outcomes in $\{-\frac{1}{2}, \frac{1}{2}\}$. There is a clear penalty for not complying: one will simply lose the bet. This is true in EQM irrespective of whether one adopts StaS, but is not evident according to the QBist. After all, there is no fact of the matter about the outcome of the experiment.

What is odd about QBism is that by making non-contextuality (and the Born rule) tautologous, it throws away much of the objectivity of the Hilbert space structure. So much so, that one may wonder whether non-contextuality is really such a valuable thing to save. Why is it so important to use a density operator to set my credences for a measurement outcome when there is no fact of the matter concerning the correct self-adjoint operator to determine the possible measurement outcomes? Personally, I'd prefer to take $CL(\mathcal{H})$ as expressing the objectively correct structure of a particular system without being able to derive the Born rule, rather than reducing $CL(\mathcal{H})$ to a subjective choice just in order to have the Born rule.

This sentiment may even be taken up a notch. Why is it so important to recover the Born rule? How important is the rule for the descriptive and explanatory power for quantum mechanics? How many experiments actually distinguish between non-contextual and contextual probability assignments? These are all interesting questions to which I have no good answer. But in any case, an agent adopting a contextual probability function is coherent within each possible experimental context. The difficulty of accepting StaS as a rationality constraint indicates precisely that there is no obvious penalty for violating non-contextuality. Contextual probability functions may then be just fine for a subjective Bayesian, and it is not clear to me why QBists make such an effort to reject them. The broader lesson, I think, is that any argument that establishes that there is no objective sense in which using one quantum state rather than another is better (which is what the QBists claim), also establishes that no contextual probability function is worse than some non-contextual probability function (which is what the QBists deny).

The subjective Bayesian in quantum mechanics then could be satisfied with using any conditional probability function on $CL(\mathcal{H})$. The objective Bayesian (as well as the Everettian), on the other hand, is looking for much more than satisfaction of the Born rule, namely, the aim is to single out a unique quantum state. Once this state has been obtained, the Born rule drops out as a bonus. Thus either, on the subjective view, violating the Born rule is not problematic, or, on the objective view, it is ruled out as a consequence of some motivation for the objectivity of quantum states. These considerations indicate that failing to find a derivation

of the Born rule in an interpretation-neutral setting need not be problematic. It may be necessary that at some point interpretation-laden steps have to be made. Nevertheless, it would be interesting to see from an empiricist point of view what it means to adopt a Born measure instead of an arbitrary probability function. In the next section, then, I investigate some new attempts to characterize the Born measures on $\mathcal{CL}(\mathcal{H})$.

11.4 Logico-empirical characterization of the Born rule

In the previous section I have been looking at possible derivations of the Born rule without much success. There is no shame in this. As noted by Landsman (2009, p. 69): "no generally accepted derivation of the Born rule has been given to date". However, as indicated at the end of the previous section, it also isn't clear how important it is to have a separate derivation of the Born rule. It may just be a sub-problem concerning the metaphysical status of quantum states and their associated probabilities. The specific trace-rule then would just drop out as a byproduct in a solution of that problem.

Nevertheless, there is the intuition that the Born measures on $\mathcal{CL}(\mathcal{H})$ should play a special role. It is therefore interesting to see how these measures can be characterized. Theorem 11.2 established a connection with non-contextuality. However, the concept of non-contextuality didn't appear to be very illuminating in the context of $\mathcal{CL}(\mathcal{H})$. In this section, then, I investigate two entirely different possible characterizations. The first one is found to be unsuccessful, but the second is more promising and does lead to a particular characterization of the Born rule. However, since some technical steps are involved in obtaining this characterization, the precise physical significance of the characterization has to be the topic of future work.

11.4.1 Continuity

The first idea reverts to the discussion of the Kochen-Specker theorem in part II. Thus far, Theorem 7.4 is the closest I got to a derivation of the Born rule from conceptually motivated assumptions. But although non-contextuality at the level of hidden variables was presupposed in chapter 7, this alone turned out to be insufficient to derive the Born rule for the MKC models. Instead, the key assumption used in Theorem 7.4 was a continuity assumption. It is tempting then to believe that the Born measures on $\mathcal{CL}(\mathcal{H})$ can be characterized in a similar way. However, I argue in this section that the prospects for such a characterization are grim. Specifically, I construct an example of a probability measure that is not a Born measure, but that does appear to satisfy some proper continuity assumptions.

The motivation to use a continuity assumption stems from the idea that the norm on the set of self-adjoint operators is relevant at the level of actual experimental

setups. Whenever $\|A_1 - A_2\|$ is small enough, experimental setups that measure A_1 or A_2 will become practically indiscernible and the events $M_{A_1}(\Delta)$ and $M_{A_2}(\Delta)$ will resemble each other as well (assuming the spectra of the operators are fixed, see also Theorem 7.2). When adopting a Bayesian interpretation of probability this motivates the idea that credences for these events should resemble each other as well, as a betting agent will not know with infinite precision which event she will be betting on in practice. But on other interpretations of probability such a continuity assumption seems natural as well.

It is not trivial to give a formal definition of what it means for a conditional probability function on $\mathcal{CL}(\mathcal{H})$ to be continuous in this sense. A difficulty is that the continuity criterion would have to apply simultaneously to the two arguments of the probability function. After all, without such a constraint all probability functions are discontinuous. For example, the probability of $\mathbb{P}(S_{(\mathcal{A},P)}|S_{\mathcal{A}})$ need not be zero, but that of $\mathbb{P}(S_{(\mathcal{A},P)}|S_{\mathcal{A}'})$ is whenever $\mathcal{A}$ and $\mathcal{A}'$ are incompatible, which can be the case even when they are generated by operators A and A' that are close to each other. But irrespective of such details, there are strong reasons to believe that a continuity assumption alone will not suffice. The argument is based on the construction of a probability measure that poses as a counterexample. To understand this counterexample, some considerations have to be made.

The concept of continuity should be such that all Born measures are continuous. Thus, for every density operator ρ, the probability function $\mathbb{P}_\rho$ defined by (11.20) would be continuous. As the set of density operators is convex, for every $\lambda \in [0,1]$ and all pairs of density operator ρ_1, ρ_2

$$\lambda \, \mathbb{P}_{\rho_1} + (1 - \lambda) \, \mathbb{P}_{\rho_2} \tag{11.29}$$

is again a continuous probability function (as it is again a Born measure).

One can create non-Born measures by, instead of taking λ to be fixed, taking it to be a non-constant function on $\mathfrak{A}$ taking values in $[0,1]$. The probability function $\mathbb{P}_{\rho_1\rho_2}$ of (11.23) can be seen to be of this form. Specifically, define the function λ_{ρ_1} by

$$\lambda_{\rho_1}(\mathcal{A}) = \begin{cases} 1 & \rho_1 \in \mathcal{A}, \\ 0 & \rho_1 \notin \mathcal{A}. \end{cases} \tag{11.30}$$

Then

$$\mathbb{P}_{\rho_1\rho_2}(S|S_{\mathcal{A}}) = \lambda_{\rho_1}(\mathcal{A}) \, \mathbb{P}_{\rho_1}(S|S_{\mathcal{A}}) + (1 - \lambda_{\rho_1}(\mathcal{A})) \, \mathbb{P}_{\rho_2}(S|S_{\mathcal{A}}). \tag{11.31}$$

This is clearly not a Born measure. However, it is also a function that one wouldn't consider to be continuous because λ_{ρ_1} is not continuous. Now one could suspect that a continuity assumption would force the function λ to be constant, but this seems idle hope. Let $P \in L(\mathcal{H})$ and consider the function

$$\lambda_P(\mathcal{A}) := \sup \left\{ \mathrm{Tr}(PP') \mid P' \in L_1(\mathcal{H}) \cap \mathcal{A}' \right\}, \tag{11.32}$$

where $\mathcal{A}'$ denotes the commutant of $\mathcal{A}$. A small change in $\mathcal{A}$ ensures a small change in the set $L_1(\mathcal{H}) \cap \mathcal{A}'$. Specifically, consider operators A_1, A_2 with $\sigma(A_1) = \sigma(A_2)$, then, if $\|A_1 - A_2\|$ is small, the differences in corresponding projections in the resolutions of A_1 and A_2 are also small. A small change in $\mathcal{A}$ further ensures a small change in $\lambda_P(\mathcal{A})$ because the trace operation is continuous. The resulting probability function given by

$$\mathbb{P}_{(\rho_1,\rho_2,P)}(S|S_{\mathcal{A}}) = \lambda_P(\mathcal{A})\,\mathbb{P}_{\rho_1}(S|S_{\mathcal{A}}) + (1 - \lambda_P(\mathcal{A}))\,\mathbb{P}_{\rho_2}(S|S_{\mathcal{A}}), \qquad (11.33)$$

however, is not a Born measure in general. The following example explicitly shows this.

Example 11.1. Consider two orthonormal bases (x, y, z) and (x', y', z') in $\mathbb{C}^3$ with

$$x' = x, \;\; y' = \tfrac{1}{2}\sqrt{2}(y + z), \;\; z' = \tfrac{1}{2}\sqrt{2}(y - z). \qquad (11.34)$$

Let $\mathcal{A}_1 = \mathcal{A}lg(P_x, P_y, P_z)$ and $\mathcal{A}_2 = \mathcal{A}lg(P_{x'}, P_{y'}, P_{z'})$ and take λ_P as in (11.32) with $P = P_z$. Then

$$\lambda_{P_z}(\mathcal{A}_1) = 1 \text{ and } \lambda_{P_z}(\mathcal{A}_2) = \tfrac{1}{2}. \qquad (11.35)$$

Now take $\rho_1 = P_z$ and $\rho_2 = P_x$. Then

$$\begin{aligned}
\mathbb{P}_{(P_z,P_x,P_z)}\left(S_{(\mathcal{A}_1,P_x)}\big|S_{\mathcal{A}_1}\right) &= \mathrm{Tr}(P_x P_z) = 0 \\
&\neq \tfrac{1}{2} = \tfrac{1}{2}\,\mathrm{Tr}(P_x P_z) + \tfrac{1}{2}\,\mathrm{Tr}(P_x P_x) = \mathbb{P}_{(P_z,P_x,P_z)}\left(S_{(\mathcal{A}_2,P_x)}\big|S_{\mathcal{A}_2}\right).
\end{aligned} \qquad (11.36)$$

This is a violation of non-contextuality and so, by Theorem 11.2, the function $\mathbb{P}_{(P_z,P_x,P_z)}$ is not a Born measure. Nevertheless, it appears to be continuous as the probabilities vary continuously as one makes small changes in the event and the condition simultaneously. $\triangle$

This example suggests that a characterization of the Born rule cannot be given in the form of a continuity criterion alone. Or, at least, it shows that the criterion has to be rather non-trivial. Time, then, to try something totally different.

11.4.2   A logico-mathematical approach

If the Born rule is correct and a logico-empirical characterization for this rule is possible, then the fact that probability functions on $\mathcal{CL}(\mathcal{H})$ generally are contextual implies that in the derivation of this logic some important facts about quantum mechanics may have been overlooked. It is hard to figure out what these facts are when starting from scratch, but some clues may be found by backtracking. Suppose the method used in this chapter and the previous one is right, can we then find what is (conceptually) missing in the conditional probability spaces? An attempt will now be made by adopting a more abstract mathematical perspective to connect what we

already have (i.e., conditional probability functions on $\mathcal{CL}(\mathcal{H})$) to what is desired: a characterization of the Born measures that has a clear physical meaning.

There are two important observations that provide a starting point. First, the Born rule implies that probability functions can be completely defined as functions on $L(\mathcal{H})$. Second, non-contextuality implies that not all aspects of the algebra $\mathcal{Alg}(A)$ are relevant for the specification of the event $M_A(\Delta)$. It is now tempting to immediately bring these two together: the event $M_A(\Delta)$ is completely specified by the projection $\mu_A(\Delta)$ and this warrants that probabilities need only be assigned to projections. Theorem 11.2 established that this is sufficient to characterize the Born rule. At the level of empiricist quantum logic, it suggests that there is a natural map from $\mathcal{CL}(\mathcal{H})$ to $L(\mathcal{H})$. However, at this same level the two logics hardly seem compatible, if only because $L(\mathcal{H})$ is non-distributive. Instead, if non-contextuality leads to an equivalence relation on $\mathcal{CL}(\mathcal{H})$ at all, then the resulting logic would be more like the weak Heyting algebra $\mathcal{P}(L_1(\mathcal{H}))$ of chapter 9. On the other hand, probability functions on $L(\mathcal{H})$ do not trivially extend to probability functions on $\mathcal{P}(L_1(\mathcal{H}))$. Consequently, there also isn't an obvious way to relate the Born measures on $\mathcal{CL}(\mathcal{H})$ to probability functions on $\mathcal{P}(L_1(\mathcal{H}))$. Thus, there is an obvious gap between $L(\mathcal{H})$ and $\mathcal{CL}(\mathcal{H})$, and there is the intuition that $\mathcal{P}(L_1(\mathcal{H}))$ should provide a midway point in this gap. But this intuition introduces the additional problem of finding out what role there is for probabilities at the level of $\mathcal{P}(L_1(\mathcal{H}))$.

There is no obvious philosophical starting point for resolving the sketchy issues just raised. The proposal then is to take a more mathematical approach. Specifically, I propose to look for a map $\tau : \mathcal{CL}(\mathcal{H}) \to \mathcal{P}(L_1(\mathcal{H}))$ that satisfies the following criteria:

(i) It respects the interpretation of the propositions in $\mathcal{CL}(\mathcal{H})$.

(ii) Provides an interpretation of all the elements of $\mathcal{P}(L_1(\mathcal{H}))$ (τ is surjective).

(iii) Helps explain why it suffices to specify probabilities for elements of $L(\mathcal{H})$, now viewed as a subset of $\mathcal{P}(L_1(\mathcal{H}))$.

(iv) Implies that these probabilities define a quantum probability function on $L(\mathcal{H})$.

In short, the map should encode similarities in $\mathcal{CL}(\mathcal{H})$ that may have been overlooked in the derivation of this logic, but that are relevant for characterizing the Born rule.

Although I do not have a solid argument for what this map should be, I do have a candidate for it. Furthermore, apart from a few caveats, the map may be seen to satisfy the above desiderata. It is best introduced by first presenting a few conceptual considerations. Take the proposition M_A expressing that a measurement of A is performed. Letting $\mathcal{A} = \mathcal{Alg}(A)$, the corresponding element of $\mathcal{CL}(\mathcal{H})$ is $S_{\mathcal{A}}$. If one had to identify an element of $L(\mathcal{H})$ with this proposition, it would have to be

$\mathbb{1}$ or $\mathbb{0}$, as these are the only candidates that treat all projections in $\mathcal{A}$ on an equal footing. This is of course unsatisfactory, as the conjunction of M_{A_1} and M_{A_2} for two totally incompatible observables certainly isn't a tautology, and the conjunction of M_{A_1} and M_{A_2} for two compatible observables isn't an obvious contradiction either.

The lattice $\mathcal{P}(L_1(\mathcal{H}))$ has a structure that is much richer than $L(\mathcal{H})$, and it allows for a more satisfactory encoding of the propositions of the form $S_{\mathcal{A}}$. My proposal is that τ identifies $S_{\mathcal{A}}$ with the set $\Delta_{\mathcal{A}} := L_1(\mathcal{H}) \cap \mathcal{A}'$, where $\mathcal{A}'$ indicates the commutant of $\mathcal{A}$. Thus $\Delta_{\mathcal{A}}$ consists of all 1-dimensional projection operators that are compatible with $\mathcal{A}$. As such, the set $\Delta_{\mathcal{A}}$ gives a good characterization of the algebra $\mathcal{A}$. In fact, $\mathcal{A}$ can be recovered from it because it is the commutant of $\Delta_{\mathcal{A}}$. When the Hilbert space is finite-dimensional, $\Delta_{\mathcal{A}}$ can be rewritten as

$$\Delta_{\mathcal{A}} = L_1(\mathcal{H}) \cap \mathcal{A}' = \bigcup_{P \in L_{\mathrm{a}}(\mathcal{A})} (\downarrow P \cap L_1(\mathcal{H})). \tag{11.37}$$

The intended reading is that each component $\downarrow P$ in the union expresses that a measurement of A is performed, and that the result is one specific eigenvalue $a \in \sigma(A)$. This view is further justified by noting that, for every A with $\mathcal{Alg}(A) = \mathcal{A}$,

$$\Delta_{\mathcal{A}} = \bigcup_{a \in \sigma(A)} (\downarrow \mu_A(\{a\}) \cap L_1(\mathcal{H})). \tag{11.38}$$

The next step is to associate $S_{(\mathcal{A},P)}$ with

$$\Delta_{(\mathcal{A},P)} := \bigcup_{\substack{P' \in L_{\mathrm{a}}(\mathcal{A}), \\ P' \leq P}} (\downarrow P' \cap L_1(\mathcal{H})). \tag{11.39}$$

The intended reading is again that each component $\downarrow P'$ in the union expresses that a measurement of A is performed, and that the result is one specific eigenvalue $a \in \sigma(A)$. However, the union now only runs over the eigenvalues that are compatible with P, i.e., the eigenvalues that satisfy $\mu_A(\{a\}) \leq P$. The proposal for completely specifying the map τ is to make use of the identity

$$S = \bigvee_{\mathcal{A} \in \mathfrak{A}} S_{(\mathcal{A}, S(\mathcal{A}))} \tag{11.40}$$

on $\mathcal{CL}(\mathcal{H})$. The definition of τ then becomes

$$\tau(S) := \bigcup_{\mathcal{A} \in \mathfrak{A}} \bigcup_{\substack{P' \in L_{\mathrm{a}}(\mathcal{A}), \\ P' \leq S(\mathcal{A})}} (\downarrow P' \cap L_1(\mathcal{H})). \tag{11.41}$$

The reader may check that indeed

$$\tau\left(S_{(\mathcal{A},P)}\right) = \Delta_{(\mathcal{A},P)}. \tag{11.42}$$

The way in which τ is defined indicates that condition (i) is met in so far as this is possible. In the first step it was noted that $S_{\mathcal{A}}$ can be recovered from the set $\tau(S_{\mathcal{A}})$, and the other steps were straight forward generalizations. The intended reading of particular elements of $\mathcal{P}(L_1(\mathcal{H}))$ should then remain the same as the elements they originate from. Thus the intended reading of $\Delta_{(\mathcal{A},P)}$ still is that of (an equivalence class of sentences like) "A is measured and the result lies in Δ". On this reading, the pseudo-negation on $\mathcal{P}(L_1(\mathcal{H}))$ (as given in (9.18)) presents itself as a useful logical connective, because

$$\sim \Delta_{(\mathcal{A},P)} = \Delta_{(\mathcal{A}lg(P),P^{\perp})}. \tag{11.43}$$

This expresses a performance of the coarsest possible measurement with a result that contradicts $M_A(\Delta)$. Even more interestingly, the associated implication relation (9.28) has the property that

$$\Delta_{\mathcal{A}} \to \Delta_{(\mathcal{A},P)} = \downarrow P \cap L_1(\mathcal{H}). \tag{11.44}$$

Note that $\downarrow P \cap L_1(\mathcal{H})$ is precisely the translation of $P \in L(\mathcal{H})$ into $\mathcal{P}(L_1(\mathcal{H}))$ given by (9.14). Thus precisely the elements of $\mathcal{P}(L_1(\mathcal{H}))$ that correspond to projections express propositions of the form $M_A \to M_A(\Delta)$, i.e., "if A is measured, then the result lies in Δ". Note that this is the reading of projection operators as suggested in (8.24) in section 8.3.

Condition (iii) is then satisfied when one adopts the view that quantum probabilities are ultimately probabilities of conditionals. This is a view that was already presented in section 10.1 as one of the three paraphrases of the Born postulate. So while up till now I have worked with the reading BR3, we here have a hint that BR2 may be more suitable after all. An important distinction with earlier uses of conditionals, though, is that now these conditionals are not taken as atomic propositions, but as composite ones. Consequently, the non-distributivity of $L(\mathcal{H})$ can be viewed as a result of selecting a particular subset from the set of all propositions, namely, those corresponding to particular conditionals. The lattice $\mathcal{P}(L_1(\mathcal{H}))$ then explains why meets and joins in $L(\mathcal{H})$ in general do not express conjunctions and disjunctions for these propositions, but do so when the projections commute. This is because in that case the antecedents of the two corresponding conditionals can be considered to be compatible.

This same observation can be used to motivate that condition (iv) is met. A collection of pairwise orthogonal projection operators can be seen to refer to a collection of conditionals, all of which have the same measurement specified in the antecedent, and for which the corresponding measurement outcomes are pairwise incompatible. That is, a sequence $P_1, \ldots, P_n$ of pairwise orthogonal projections corresponds to a sequence of conditionals $M_A \to M_A(\Delta_1), \ldots, M_A \to M_A(\Delta_1)$, where $\Delta_1, \ldots, \Delta_n$ is a sequence of pairwise disjoint sets. This motivates that a

probability function on $L(\mathcal{H})$ (now understood as a set of (equivalence classes of) conditionals) should satisfy

$$\mathbb{P}\left(\bigvee_{i\in I} P_i\right) = \sum_{i\in I} \mathbb{P}(P_i) \tag{11.45}$$

for such a sequence. Gleason's theorem then shows that $\mathbb{P}$ satisfies the Born rule.

This is the hand-waving success story. In short, if one can defend that τ provides a proper translation of $\mathcal{CL}(\mathcal{H})$, and that the implication relation on $\mathcal{P}(L_1(\mathcal{H}))$ expresses a conditional, then one has a characterization of the Born measures on $\mathcal{CL}(\mathcal{H})$. Specifically, the Born measures are the probability measures that respect the translation of τ in the sense that they give rise to probabilities for the conditionals in $L(\mathcal{H})$ in an unambiguous way, i.e., the Born measures are precisely the measures for which

$$\mathbb{P}\left(\tau(S_{\mathcal{A}}) \to \tau(S_{(\mathcal{A},P)})\right) := \mathbb{P}\left(S_{(\mathcal{A},P)}\big|S_{\mathcal{A}}\right) \tag{11.46}$$

is well-defined.

But of course, the devil is in the details. Some of these details work out nicely. For example, one can show that condition (ii) is also satisfied. But arguing that (i) is indeed satisfied is more difficult, because it is a difficult matter to decide what it means exactly to respect the interpretation of $\mathcal{CL}(\mathcal{H})$. This can all be seen from the following theorem and some of its consequences.

Theorem 11.3. *The map $\tau : \mathcal{CL}(\mathcal{H}) \to \mathcal{P}(L_1(\mathcal{H}))$ given by (11.41) is a complete surjective join-homomorphism. This property remains true when τ is restricted to $\mathit{IL}(\mathcal{H})$.*

Proof. Let $\{S_i \,|\, i \in I\}$ be an arbitrary subset of $\mathcal{CL}(\mathcal{H})$. Then

$$\tau\left(\bigcurlyvee_{i\in I} S_i\right) = \bigcup_{\mathcal{A}\in\mathfrak{A}} \bigcup_{\substack{P\in L_{\mathrm{a}}(\mathcal{A}), \\ P\leq(\bigcurlyvee_{i\in I} S_i)(\mathcal{A})}} (\downarrow P \cap L_1(\mathcal{H})). \tag{11.47}$$

The second disjunction can be rewritten as

$$\left\{ P \in L_{\mathrm{a}}(\mathcal{A}) \,\bigg|\, P \leq \left(\bigcurlyvee_{i\in I} S_i\right)(\mathcal{A}) \right\} = \left\{ P \in L_{\mathrm{a}}(\mathcal{A}) \,\bigg|\, P \leq \bigvee_{i\in I} S_i(\mathcal{A}) \right\}$$
$$= \bigcup_{i\in I} \{P \in L_{\mathrm{a}}(\mathcal{A}) \,|\, P \leq S_i(\mathcal{A})\}. \tag{11.48}$$

Then

$$\tau\left(\bigvee_{i\in I} S_i\right) = \bigcup_{\mathcal{A}\in\mathfrak{A}}\bigcup_{i\in I}\bigcup_{\substack{P\in L_a(\mathcal{A}),\\ P\leq S_i(\mathcal{A})}} (\downarrow P \cap L_1(\mathcal{H})) = \bigcup_{i\in I}\tau\left(S_i\right). \tag{11.49}$$

This shows that τ is a complete join-homomorphism. To see that it is surjective first note that for every $P \in L_1(\mathcal{H})$

$$\tau\left(S_{(\mathcal{A}\mathit{lg}(P),P)}\right) = \{P\}. \tag{11.50}$$

Surjectivity then follows because for any $\Delta \subset L_1(\mathcal{H})$

$$\tau\left(\bigvee_{P\in\Delta} S_{(\mathcal{A}\mathit{lg}(P),P)}\right) = \Delta. \tag{11.51}$$

□

This theorem establishes that τ has some decent properties one would require of a translation of $\mathcal{CL}(\mathcal{H})$ into $\mathcal{P}(L_1(\mathcal{H}))$. However, while τ does preserve joins, it does not preserve meets. Furthermore, τ is surjective, but not injective. In fact, failure of injectivity already crept into the definition of τ in the second step. Although it is possible to recover $S_\mathcal{A}$ from the set $\Delta_\mathcal{A} = \tau(S_\mathcal{A})$, it is in general not possible to recover $S_{(\mathcal{A},P)}$ from the set $\Delta_{(\mathcal{A},P)} = \tau(S_{(\mathcal{A},P)})$ because one doesn't know which of the $P'' \in L_a(\mathcal{A})$ have been lost by the restriction $P' \leq P$. Specifically, for any two algebras $\mathcal{A}_1, \mathcal{A}_2$ with $P \in L(\mathcal{A}_1) \cap L(\mathcal{A}_2)$ one has

$$\begin{aligned} \{P' \in L_a(\mathcal{A}_1) \mid P' \leq P\} = \{P' \in L_a(\mathcal{A}_2) \mid P' \leq P\} \\ \Rightarrow\ \tau(S_{(\mathcal{A}_1,P)}) = \tau(S_{(\mathcal{A}_2,P)}). \end{aligned} \tag{11.52}$$

A demonstration of a failure of meet-preservation is obtained by considering $\mathcal{A}_1$ and $\mathcal{A}_2$ as in Example 11.1. In this case,

$$S_{(\mathcal{A}_1,P_x)} \curlywedge S_{(\mathcal{A}_2,P_x)} = \bot, \tag{11.53}$$

while on the other hand,

$$\tau\left(S_{(\mathcal{A}_1,P_x)}\right) \cap \tau\left(S_{(\mathcal{A}_2,P_x)}\right) = \{P_x\} \cap \{P_x\} = \{P_x\}. \tag{11.54}$$

Incidentally, this example also displays that τ isn't injective (even when restricted to $I\!L(\mathcal{H})$).

The failure of injectivity was to be expected. After all, it is precisely such a failure we were looking for to explain the special role of Born measures on $I\!L(\mathcal{H})$ (although one may wonder if injectivity fails in the 'right' way). More puzzling is

the failure of meet-preservation. As a more general example of this failure note that, whenever $\mathcal{A}_1$ and $\mathcal{A}_2$ are incompatible (i.e., $\mathcal{Alg}(\mathcal{A}_1, \mathcal{A}_2)$ is not Abelian), $S_{\mathcal{A}_1} \curlywedge S_{\mathcal{A}_2} = \bot$, while $\Delta_{\mathcal{A}_1} \cap \Delta_{\mathcal{A}_2}$ generally does not identify with the empty set. The rationale behind the first conjunction is that a joint measurement of $\mathcal{A}_1$ and $\mathcal{A}_2$ is not possible. To find a possible rationale behind the second it is useful to note that the intersection satisfies

$$\Delta_{\mathcal{A}_1} \cap \Delta_{\mathcal{A}_2} = \tau(S_{(\mathcal{A}_{1,2}, P_{1,2})}), \tag{11.55}$$

with

$$\begin{aligned} \mathcal{A}_{1,2} &= \mathcal{Alg}\left(\{P_1 \wedge P_2 \mid P_1 \in L_a(\mathcal{A}_1), P_2 \in L_a(\mathcal{A}_2)\}\right), \\ P_{1,2} &= \bigvee \{P_1 \wedge P_2 \mid P_1 \in L_a(\mathcal{A}_1), P_2 \in L_a(\mathcal{A}_2)\}. \end{aligned} \tag{11.56}$$

The algebra $\mathcal{A}_{1,2}$ is the largest Abelian algebra that is compatible both with $\mathcal{A}_1$ and $\mathcal{A}_2$. In a sense, it highlights what the two have in common. For example, if $\mathcal{A}_1$ and $\mathcal{A}_2$ are totally incompatible, then $\mathcal{A}_{1,2} = \mathbb{C}\mathbb{1}$ and $P_{1,2} = \mathbb{0}$.

This partial compatibility can be given meaning when considering consecutive measurements instead of joint measurements. Implicitly, consecutive measurements already played a role in defending the plausibility of LMR. A measurement of A_1 also counts as a measurement of $A_2 = f(A_1)$, not only because of the counterfactual statement that "if both had been measured, then their outcomes would satisfy $a_2 = f(a_1)$", but also because of the indicative "if a measurement of A_1 is directly followed by a measurement of A_2, then their outcomes satisfy $a_2 = f(a_1)$". In fact, it doesn't matter in which order A_1 and A_2 are measured.

Empirical rules of this kind led von Neumann (1932, §III.3) to the introduction of the projection postulate. But the importance of this postulate reaches beyond consecutive measurements of compatible observables.[7] For example, in the case of the two algebras of Example 11.1, if the statement $S_{(\mathcal{A}_1, P_x)}$ is true, then an immediately consecutive measurement of $\mathcal{A}_2$ would make $S_{(\mathcal{A}_2, P_x)}$ true. The algebra $\mathcal{A}_{1,2}$ from (11.56) precisely characterizes what kind of implications a measurement of $\mathcal{A}_1$ can have for the possible outcomes of a consecutive measurement of $\mathcal{A}_2$, and vice versa.

This is all just fine, but it doesn't *really* explain how one can think of $S_{(\mathcal{A}_{1,2}, P_{1,2})}$ as expressing the conjunction of $S_{\mathcal{A}_1}$ and $S_{\mathcal{A}_2}$. All that is made clear is that LMR may have been applied too strictly in the construction of $\mathcal{CL}(\mathcal{H})$, in the sense that when a measurement of A_1 doesn't imply a measurement of A_2, this doesn't mean that the measurement of A_1 has no implications whatsoever for the outcomes of a future measurement of A_2. Making this precise requires the introduction of a formal notion of consecutive measurements. The projection postulate can then play a role here.

---

[7]Although the postulate is formulated as dictating the change in the state of a system, and thus presupposes the state postulate, I am here referring to particular non-probabilistic empirical features encoded by this postulate.

It is beyond the scope of this dissertation to develop a new quantum logic that extends $\mathcal{CL}(\mathcal{H})$ to incorporate propositions of consecutive measurements. Nevertheless, some insight can be fostered by looking at the map τ. Consider an elementary experimental proposition represented by the function $S_{(\mathcal{A},P)}$. The starting point in chapter 10 is that these propositions are the idealized kind of propositions that encode relevant physical data. What are precisely the implications of such a proposition? First of all, it can be paraphrased as the disjunction

$$S_{(\mathcal{A},P)} = \bigvee_{\substack{P' \in L_{\mathrm{a}}(\mathcal{A}), \\ P' \leq P}} S_{(\mathcal{A},P')}. \tag{11.57}$$

When adopting the projection postulate, this implies that possible outcomes of (immediate) future measurements are constrained by the probability-1 statements implied by P' for one of the $P' \in L_{\mathrm{a}}(\mathcal{A})$ with $P' \leq P$. However the remainder of the algebra $\mathcal{A}$ is made up (i.e., the part specified by $\{P' \in L_{\mathrm{a}} \mid P' \leq P^{\perp}\}$) is irrelevant. It is also precisely this part that is forgotten by τ as expressed by (11.52). In this sense, τ may be viewed as extracting the information from the proposition $S_{(\mathcal{A},P)}$ that is relevant for future predictions. But again, to check if this all pans out, one has to make precise what the empirical quantum mechanical statements are concerning time-ordered events. This is something for future work.

To recapitulate, the translation of $\mathcal{CL}(\mathcal{H})$ into $\mathcal{P}(L_1(\mathcal{H}))$ provided by τ can be used to give a characterization of the Born measures on $\mathcal{CL}(\mathcal{H})$. They are the measures for which (11.46) is well-defined. It has also been shown that these Born measures then define a quantum probability function on $L(\mathcal{H})$ by making use of (11.44). The foundational importance of these results depends on whether one can give a proper physical interpretation of τ. I have indicated that the failures of injectivity and meet-preservation are in need of an explanation. This is not because these failures themselves are necessarily problematic, but precisely because explanations of these failures should foster an explanation for the special role of the Born measures. My suggestion is that τ incorporates relevant dynamical aspects that have to be taken into account when reasoning in quantum mechanics, which may have been lost in the construction of $\mathcal{CL}(\mathcal{H})$. But there are of course other possibilities. For one, it may be that τ just needs some tweaking. It is also possible that an empiricist characterization of the Born rule just isn't something to be had, and perhaps the rule has to be accepted as a contingent fact rather than as something that holds necessarily. At the moment, however, I do not have an argument in favor of either of these directions.

11.5   Afterthought

Unfortunately, the last two sections did not provide a proper ending for a success story on quantum probability. But even without a proper characterization of the Born rule I think a lot has been accomplished part III. On the basis of empiricist considerations, a clear reformulation of quantum probability has been obtained. Due to this empiricist approach the obtained formalism is compatible with most interpretations of quantum mechanics. Nonetheless, I think this formalism is most helpful for those who stay close to an empiricist stance, such as variations on the Copenhagen interpretation. It can be helpful, for example, in defending these views against the charges that they are vague or obscure. It shows that one doesn't require a hand-waving interpretation of $L(\mathcal{H})$ as something like a non-distributive event space to make sense of quantum probability. Neither does one need a story that is disjoint from the formalism to make fuzzy qualifications about what can and cannot be said in quantum mechanics. To elaborate on the point I'm trying to make, a sketchy attempt at an unlikely synthesis of the views of Bohr and Bell may be useful. Let me start with Bell.

Bell is well-known for his criticism of Copenhagen-like interpretations and 'standard textbook explanations' of quantum mechanics. Part (and possibly the basis) of his complaint is that these presentations of the theory do not provide a clear foundation for the theory. Or, in his own words, do they not provide "an exact formulation of some serious part of quantum mechanics" (Bell, 1990, p. 33). Bell elaborates that by an exact formulation he means that "the theory should be fully formulated in mathematical terms". The usual formalism does not meet this criterion. To be sure, quantum mechanics does have a solid mathematical foundation in Hilbert space theory. But the theory is not entirely formulated in these terms. The Born postulate provides an example. It combines the mathematical formalism with pre-theoretical concepts such as 'observable' and 'measurement' to obtain an empirical statement from quantum mechanics. These pre-theoretical concepts do not have a solid foundation in the mathematical formalism of the theory. To the extent that they are linked to the mathematical formalism, it is not entirely clear that their mathematical role fits well with the natural language interpretation of these concepts. The difficulty with interpreting the orthodox quantum logic discussed in chapter 8 illustrates this. Consequently, the precise role and meaning of these concepts within the context of quantum mechanics remains ambiguous.

The ambiguity of concepts like 'measurement' infects the whole of quantum mechanics, turning it into an ambiguous theory: it just isn't clear what the theory really is about. Bell advocates the view that ambiguity is best avoided by removing the pre-theoretical concepts from the theory. Rather, these concepts should be derivable from the theory. A measurement, for example, should be a physical dynamical process that can be given meaning using the mathematical formalism of the

theory. The formalism of von Neumann does not, and possibly cannot, provide such a dynamical account of measurements. This is, after all, the core of the measurement problem. A new formalism then seems to be required, and Bell proposes that Bohmian mechanics or the GRW collapse-theory are more promising candidates for an exact formulation of quantum mechanics.

To a large extent I am sympathetic to the criticisms of Bell. But I hesitate to accept that the only way to make quantum mechanics less ambiguous is by replacing it with a theory in which the notion of measurement is given a dynamical physical meaning. In essence, demanding such a replacement is a plea for an ontology for the theory, or, to use Bell's terminology, for a theory of (local) beables. The point where my view drifts apart from Bell's can be illustrated on the basis of the following quote.

> When one forgets the role of the apparatus, as the word 'measurement' makes all too likely, one despairs of ordinary logic – hence 'quantum logic'. When one remembers the role of the apparatus, ordinary logic is just fine. (ibid., p. 34)

As I argued in chapter 8, orthodox quantum logic provides an ambiguous formalism for correct reasoning within the context of quantum mechanics, and I agree with Bell that this ambiguity can be seen as a consequence of a reckless use of the concept of measurement. But providing a physical basis for this concept is only one way to prevent sloppy uses of it. Another way is the one I adopted: to reformulate quantum logic in such a way that the concept of measurement is taken seriously, even if it is still a primitive term. Indeed, the results of chapter 10 show that also on this approach ordinary logic is just fine.

There are of course other motivations (apart from avoiding unambiguity) for a physical account of measurement processes. One may simply hold the view that a (fundamental) physical theory should provide an ontology, and that it is the task of such a theory to describe what is really going on in the world instead of restricting attention to how we perceive the world. But that seems to me a much stronger demand than "an exact formulation" of a physical theory.

Now let me turn to Bohr. This is now only a small step. Above I related Bell's demand for an exact formulation of quantum mechanics to a demand for a formulation of quantum mechanics that is unambiguous with respect to what the theory has to say. A similar demand for an unambiguous formulation can be found in the writings of Bohr. But now this demand is used as part of a defense of the usual formulation of quantum mechanics, in which a strict separation occurs between the mathematical formalism and the description of actual experiments to which the formalism is applied. Specifically, the quantum-mechanical description is reserved for the system under investigation, while the apparatus used to perform a measurement on the system is given a classical description:

> Just the requirement that it be possible to communicate experimental findings in an unambiguous manner implies that the experimental arrangement and the results of observation must be expressed in the common language adapted to our orientation in the environment. Thus, the description of quantum phenomena requires a distinction in principle between the objects under investigation and the measuring apparatus by means of which the experimental conditions are defined. (Bohr, 1963, p. 78)

The reformulation of quantum probability developed in this part of the dissertation can be seen to respect the demand that experimental findings are communicable in common language. This was in fact one of the guiding ideas for the construction of the logics in chapter 10. But what is new is that this logic makes a clear connection with the mathematical formalism of quantum mechanics. The demand for the use of common language thus doesn't imply a strict separation from the Hilbert space formalism. There is of course still a distinction between the kinds of descriptions used for the system and the apparatus. This is necessarily so when 'measurement' is used as a primitive concept. But everything is now embedded in a single mathematical framework.

Now it may be seen that there is a synthesis of the views of Bell and Bohr in the following sense. The obtained reformulation of quantum probability adheres both to Bell's demand for an exact formulation and to Bohr's demand for the possibility of unambiguous communication of experimental findings. The synthesis can be illustrated further by returning to the discussion of non-contextuality in section 6.3. The assumption of non-contextuality for hidden variables resulted in the impossibility of defining hidden variable states as truth valuation functions on the set of projection operators. Bell (1966, p. 447) famously criticized this assumption of non-contextuality by quoting Bohr. Indeed, Bohr emphasized that no experimental finding can be unambiguously defined without a specification of the experimental context. This criterion can be used both against a non-contextuality assumption for hidden variables, as well as against orthodox quantum logic.

My reformulation of quantum probability respects the need for specifications of measurement contexts. In the underlying logic there is explicit attention for the specification of measurement contexts, and probability functions are explicitly introduced as two-placed functions that require both a measurement context (as a condition) and a specification of possible outcomes within this context. By being explicitly contextual, my reformulation leaves open the possibility of an underlying hidden variable theory. Thus by staying relatively close to the ideas of Bohr, the obtained framework of quantum probability is also compatible with the ideas of Bell. In which of these two directions (Bohr or Bell) one wants to go depends on what one expects/demands of a physical theory. Currently, I am still undecided with

respect to this issue. But in either case one can benefit from the disambiguation of the formalism of quantum probability provided here.

A

Mathematical background

The purpose of this appendix is to give a short overview of the main mathematical definitions and theorems that play a role in this dissertation. This text is not self-contained and does not pretend to give a thorough introduction. Instead, this appendix is meant to help the reader who is somewhat familiar with the mathematics used, but does not have all the relevant details immediately available in his/her mind. Instead then of having to look up the details in a separate book, I hope the reader can find what he/she is looking for here. The advice then is to not read this appendix from beginning to end. Instead, the index of the dissertation is a useful guide to points of interest in the appendix.

The theorems discussed here are presented without proof. Most of them can be found in the books by Burris and Sankappanavar (2012) (for lattice theory) and Conway (1990) (for Hilbert space theory). These books also provide thorough introductions to these subjects. The theory of modal algebras is less standard, and was first introduced by Lemmon (1966). For more on operator theory the books of Kadison and Ringrose (1986) are classics.

A.1 Basic algebraic structures

Much of mathematics can be viewed as part of set theory. Different areas of mathematics then differ in their focus on particular kinds of structures on sets. For example, group theory focuses on sets on which a multiplication rule is defined. Topology focuses on sets with particular kinds of sets of subsets. It is this kind of perspective on mathematics that sets a useful common ground for the main fields of mathematics used in this dissertation: Hilbert space theory and lattice theory. In this section I give the basic definitions that establish this common ground.

The structure on sets considered here is that of operations. A unary operation on a set X is a map from X to itself. A binary operation on X is a map from $X \times X$ to X. And so on for higher orders. A set with one or more operations is called an

algebraic structure.

A.1.1 Structures with a single operation

Definition A.1. A *semigroup* G is a pair $(X, +)$ with X a set and $+$ a binary operation $(x_1, x_2) \mapsto x_1 + x_2$ (often called addition or multiplication) that is associative, i.e.,

$$(x_1 + x_2) + x_3 = x_1 + (x_2 + x_3) \ \forall x_1, x_2, x_3 \in X. \tag{A.1}$$

If X in addition has a unit element, i.e., an element 0_X such that

$$x + 0_X = 0_X + x = x \ \forall x \in X, \tag{A.2}$$

then G is called a *monoid*. If for every $x \in X$ there is an element, denoted $-x$, such that

$$x + (-x) = (-x) + x = 0_X, \tag{A.3}$$

then G is called a *group*. The element $-x$ is called the *inverse* of x. If

$$x_1 + x_2 = x_2 + x_1 \ \forall x_1, x_2 \in X \tag{A.4}$$

then the operation is called *commutative* and G is called an *Abelian group*.

Often, when the binary operation is called multiplication, the operation is denoted $\cdot$, the unit element is denoted 1_X, and the inverse is denoted x^{-1}

Definition A.2. A *semilattice* $(S, \wedge)$ is a semigroup for which the operation $\wedge$ (usually called the meet or join) is commutative and idempotent, i.e.,

$$\forall s \in S : \ s \wedge s = s. \tag{A.5}$$

If in addition S has a unit element (i.e., $(S, \wedge)$ is a monoid), then the semilattice is called *bounded*. This element is usually denoted $\top$.

When the binary operation is called a join it is usually denoted $\vee$. In this case the unit element is denoted $\bot$.

Example A.1. Consider the set of two elements $\{0, 1\}$. The only possible binary operations that turn the set into a monoid are given by

+	0	1		×	0	1		∧	0	1		∨	0	1
0	0	1		0	1	0		0	0	0		0	0	1
1	1	0		1	0	1		1	0	1		1	1	1

$$\tag{A.6}$$

The first two operations are the only operations that turn the set into a group. The last two operations are the only operations that turn the set into a semilattice. The operations $+$ and $\times$ are identical up to a relabeling of the elements of the set. The same holds for $\wedge$ and $\vee$. △

Example A.2. Each of the number systems $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ is an Abelian group with respect to addition, with 0 the unit element. With respect to multiplication each of these sets is a monoid with unit element 1, and $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ are groups. The number system $\mathbb{N}$ is a monoid with respect to addition and multiplication. Each of these sets is turned into a semilattice by setting $x_1 \wedge x_2 := \min(x_1, x_2)$ or $x_1 \vee x_2 := \max(x_1, x_2)$. $\triangle$

As these examples illustrate, often more than one binary operation is given on a set. Several important definitions particularly refer to how some of these binary operations relate to each other.

A.1.2 Structures with multiple operations

Definition A.3. A *pseudo-ring* R is a triplet $(X, +, \cdot)$ with X a set and $+$ and $\cdot$ two binary operations (called addition and multiplication) such that $(X, +)$ is an Abelian group and $(X, \cdot)$ is a semigroup. The multiplication is further required to distribute over addition:

$$
\begin{aligned}
x_1 \cdot (x_2 + x_3) &= (x_1 \cdot x_2) + (x_1 \cdot x_3) \\
(x_1 + x_2) \cdot x_3 &= (x_1 \cdot x_3) + (x_2 \cdot x_3)
\end{aligned} \tag{A.7}
$$

If $(X, \cdot)$ is a monoid, then the structure is called a *ring*. If the multiplication is commutative, R is called a *commutative ring* and if the multiplication is idempotent then R is called a *Boolean ring*. If R is a commutative ring and $(X \backslash \{0_X\}, \cdot)$ is a group then the algebraic structure is called a *field*.

Lemma A.1. *If R is a Boolean ring, then R is a commutative ring that satisfies*

$$
\forall x \in R : \ x + x = 0. \tag{A.8}
$$

Commutativity can be seen to follow from considering that $(x_1 + x_2)^2 = x_1 + x_2$ and (A.8) from the identity $(x + x)^2 = x + x$.

Definition A.4. A *lattice* is a triplet $(L, \wedge, \vee)$ with two binary operations such that $(L, \wedge)$ and $(L, \vee)$ are semilattices. The operations are further required to satisfy the absorption laws

$$
a \wedge (a \vee b) = a \vee (a \wedge b) = a \tag{A.9}
$$

for all $a, b \in L$. If the operations further satisfy

$$
\begin{aligned}
a \wedge (b \vee c) &= (a \wedge b) \vee (a \wedge c) \\
a \vee (b \wedge c) &= (a \vee b) \wedge (a \vee c)
\end{aligned} \tag{A.10}
$$

for all $a, b, c \in L$, then the lattice is called *distributive*. The lattice is called *bounded* when both $(L, \wedge)$ and $(L, \vee)$ are bounded semilattices.

Example A.3. Consider again the set of two elements $\{0,1\}$ with binary operations as given in (A.6). Then $(\{0,1\}, +, \vee)$ is a field that is also a Boolean ring, and $(\{0,1\}, \wedge, \vee)$ is a bounded distributive lattice. $\triangle$

Example A.4. The sets $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ with addition and multiplication defined in the usual way are fields. The set of natural numbers with meet and join as defined in Example A.2 is a distributive lattice that is not bounded. $\triangle$

Example A.5. Let R be a Boolean ring. Then it follows from Lemma A.1 that the definitions
$$x_1 \wedge x_2 := x_1 x_2,$$
$$x_1 \vee x_2 := x_1 + x_2 + x_1 x_2 \tag{A.11}$$
turn R into a distributive lattice. $\triangle$

Apart from algebraic structures on their own, relations between them are at least as important.

Definition A.5. Let X_1, X_2 be two algebraic structures of the same kind. Then a *homomorphism* from X_1 to X_2 is a map $h : X_1 \to X_2$ that respects all operations on the structures. If the map is bijective, then it is called an *isomorphism*. The collection of all homomorphisms from X_1 to X_2 is denoted $\mathrm{Hom}(X_1, X_2)$.

Example A.6. Consider the two semilattices $S_1 = (\{0,1\}, \wedge)$ and $S_2 = (\{0,1\}, \vee)$ of Example A.1. Then the map $h : S_1 \to S_2$ given by $h(0) := 1$, $h(1) := 0$ is an isomorphism of semilattices since for all $a, b \in S_1$: $h(a \wedge b) = a \vee b$. $\triangle$

A.1.3 Linear Algebra

Definition A.6. A *vector space V* over a field F is an Abelian group $(V, +)$ together with an operation $F \times V \to V$, $(\lambda, v) \mapsto \lambda v$ called scalar multiplication that is compatible with the addition on V in the sense that for all $\lambda_1, \lambda_2 \in F$ and $v_1, v_2 \in V$
$$\lambda_1(\lambda_2 v_1) = (\lambda_1 \cdot \lambda_2) v_1, \ 1v_1 = v_1,$$
$$\lambda_1(v_1 + v_2) = \lambda_1 v_1 + \lambda_1 v_2, \ (\lambda_1 + \lambda_2) v_1 = \lambda_1 v_1 + \lambda_2 v_1. \tag{A.12}$$

Example A.7. For every field F, the set $F^n = \{(f_1, \ldots, f_n) \mid f_i \in F \text{ for all } i\}$ is turned into a vector space over F by defining
$$(f_1, \ldots, f_n) + (f'_1, \ldots, f'_n) := (f_1 + f'_1, \ldots, f_n + f'_n),$$
$$f(f_1, \ldots, f_n) := (f \cdot f_1, \ldots, f \cdot f_n). \tag{A.13}$$

$\triangle$

Definition A.7. Let V be a vector space over a field F and let $\{v_i \,|\, i \in I\}$ be a set of vectors in V. Then the *linear span* of this set is defined as

$$\mathrm{span}(\{v_i \,|\, i \in I\}) := \left\{ \lambda_1 v_{i_1} + \ldots + \lambda_n v_{i_n} \;\middle|\; \begin{smallmatrix} \lambda_1,\ldots,\lambda_n \in F, \\ i_1,\ldots,i_n \in I, \ n \in \mathbb{N} \end{smallmatrix} \right\}. \tag{A.14}$$

If $\mathrm{span}(\{v_i \,|\, i \in I\}) = V$ then $\{v_i \,|\, i \in I\}$ is called a *complete set of vectors*. If in addition $\mathrm{span}(\{v_i \,|\, i \in I\setminus\{j\}\}) \neq V$ for every $j \in I$, then $\{v_i \,|\, i \in I\}$ is called a *basis*.

Example A.8. Let F^n be as in Example A.7. Then the set

$$\{(f_1^1, \ldots, f_n^1), \ldots, (f_1^n, \ldots, f_n^n)\}, \tag{A.15}$$

with $f_j^i = \delta_{ij}$, is a basis for F^n. $\triangle$

Lemma A.2. *Let V be a vector space over a field F. If both $\{v_1, \ldots, v_n\}$ and $\{v_1', \ldots, v_m'\}$ are bases, then $n = m$. The number n is called the* dimension *of V.*

Definition A.8. A map $f : V_1 \to V_2$ between two vector spaces V_1, V_2 over F is called *linear* if

$$f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2) \tag{A.16}$$

for all $v_1, v_2 \in V_1$ and $\lambda_1, \lambda_2 \in F$. If $V_1 = V_2$ the map is called an *operator*, and its action is often denoted as Av with A the operator and $v \in V_1$. If $V_2 = F$ the map is called a *functional*. When $F = \mathbb{C}$ a map is called *anti-linear* if

$$f(\lambda_1 v_1 + \lambda_2 v_2) = \overline{\lambda_1} f(v_1) + \overline{\lambda_2} f(v_2) \tag{A.17}$$

for all $v_1, v_2 \in V_1$ and $\lambda_1, \lambda_2 \in \mathbb{C}$, where $\overline{x}$ denotes the complex conjugate of x.

Example A.9. Let V and W be two vector spaces over the same field F. Then the set $\mathrm{Hom}(V, W)$ of linear maps between V and W is itself turned into a vector space over F with the rules

$$\begin{aligned} (f_1 + f_2)(v) &:= f_1(v) + f_2(v), \\ (\lambda f)(v) &:= \lambda f(v). \end{aligned} \tag{A.18}$$

When $V = W$ a multiplication rule is also introduced with

$$(f_1 f_2)v := f_1(f_2(v)). \tag{A.19}$$

This turns $\mathrm{Hom}(V, V)$ into a monoid with the unit element given by the unit operator $\mathbb{1}$ which acts as $\mathbb{1} v = v$ for all v. Generally, it will not be a group as not all operators will have an inverse. Neither will it be commutative, and in fact the *commutator* defined as

$$[f_1, f_2](v) := (f_1 f_2)(v) - (f_2 f_1)(v) \tag{A.20}$$

will be a useful operation. $\triangle$

Definition A.9. Let V be a vector space over a field F. A map $B : V \times V \to F$ is called a *bilinear form* when for every $v \in V$ the maps $v' \mapsto B(v', v)$ and $v' \mapsto B(v, v')$ are functionals. When $F = \mathbb{C}$ a map $B : V \times V \to \mathbb{C}$ is called a *sesquilinear form* if it is linear with respect to one argument, and antilinear with respect to the other. An *inner product* is a bilinear or sesquilinear form that satisfies $B(v_1, v_2) = \overline{B(v_2, v_1)}$ for all v_1, v_2 and $B(v, v) \geq 0$ with equality if and only if $v = 0$.

Example A.10. If $\dim(V) = n$ and $\dim(W) = m$ then $\dim(\mathrm{Hom}(V, W)) = nm$. This can be seen by noting that every linear map f can be represented by a matrix A^f by choosing to bases $v_1, \ldots, v_n$ and $w_1, \ldots, w_n$ and setting

$$A^f_{ij} := \langle w_i, f(v_j) \rangle. \tag{A.21}$$

$\triangle$

A.2   Hilbert space theory

Hilbert space theory can be seen as a combination of linear algebra with analysis. Thus apart from the algebraic structure of vector spaces, also the notion of a norm is required to meaningfully address questions of convergence.

Definition A.10. A *norm* on a vector space V over $\mathbb{C}$ (or some subfield thereof) is a map from V to the positive reals denoted $v \mapsto \|v\|$ that satisfies
  (i) $\|\lambda v\| = |\lambda| \|v\|$,
  (ii) $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$,
  (iii) $\|v\| = 0$ if and only if $v = 0$.
A sequence $(v_n)_{n \in \mathbb{N}}$ is said to converge to a vector $v \in V$ with respect to the norm $\|.\|$ if for every $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $\|v_n - v\| < \epsilon$ for all $n \geq N$. The sequence is called a Cauchy sequence if for every $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $\|v_n - v_m\| < \epsilon$ for all $n, m \geq N$.

Definition A.11. A *Banach space* is a vector space X with a norm $\|.\|$ with respect to which every Cauchy sequence converges. A *Hilbert space* is a vector space $\mathcal{H}$ with an inner product $\langle ., . \rangle$ such that with respect to the norm generated by the inner product $\|\psi\| := \sqrt{\langle \psi, \psi \rangle}$ it is a Banach space. The space is called *separable* if it has a countable basis $\{\psi_n \,|\, n \in \mathbb{N}\}$.

Sometimes Hilbert spaces are explicitly defined as vector spaces over $\mathbb{R}$ or $\mathbb{C}$. Here I follow that tradition by implicitly assuming this to be the case unless stated otherwise.

Example A.11. The paradigm example of a separable Hilbert space is that associated with a single quantum particle in one dimension[1]:

$$\mathcal{H} = L^2(\mathbb{R}) := \left\{ \psi : \mathbb{R} \to \mathbb{C} \;\middle|\; \int_{-\infty}^{\infty} |\psi(x)|^2 \mathrm{d}x < \infty \right\}. \tag{A.22}$$

The inner product is given by

$$\langle \psi, \phi \rangle := \int_{\infty}^{\infty} \overline{\psi(x)}\phi(x)\mathrm{d}x. \tag{A.23}$$

△

While a subspace of a vector space is itself again a vector space, a subspace of a Hilbert space is not always itself a Hilbert space. The characterizing distinction is given by the following definition.

Definition A.12. A linear subspace $\mathcal{K} \subset \mathcal{H}$ is called *closed* if for every sequence $(\psi_n)_{n \in \mathbb{N}}$ in $\mathcal{K}$ for which $\psi = \lim_{n \to \infty} \psi_n$ exists in $\mathcal{H}$, $\psi \in \mathcal{K}$.

It follows from this definition that a linear subspace is closed if and only if it is itself a Hilbert space. Thus closed linear subspaces are the 'natural' subspaces to look at when doing Hilbert space theory. In the finite-dimensional case every linear subspace can be shown to be closed, but in the separable case this does not hold.

Example A.12. Consider the Hilbert space of all square-integrable functions on the interval $[0, 1]$, i.e.,

$$\mathcal{H} = L^2([0,1]) = \left\{ \psi : [0,1] \to \mathbb{C} \;\middle|\; \int_0^1 |\psi(x)|^2 \mathrm{d}x \text{ exists} \right\} \tag{A.24}$$

with inner product

$$\langle \psi_1, \psi_2 \rangle := \int_0^1 \overline{\psi_1(x)}\psi_2(x)\mathrm{d}x. \tag{A.25}$$

Then the set of all polynomials of finite order

$$\mathcal{K} = \left\{ \psi \in \mathcal{H} \;\middle|\; \begin{array}{c} \psi(x) = a_0 + a_1 x + \cdots + a_n x^n, \\ n < \infty, \; a_i \in \mathbb{C} \end{array} \right\} \tag{A.26}$$

is a linear subspace, but not closed. For example, the sequence $(\psi_n)_{n \in \mathbb{N}} \; \psi_n(x) = \sum_{k=0}^n \frac{1}{k!} x^k$ converges to $\psi(x) = e^x$ which is not an element of $\mathcal{K}$. △

---

[1]Actually one should take the set of equivalence classes here. That is, if $\mathcal{L}^2(\mathbb{R})$ denotes the set of all square (Lebesgue-)integrable complex functions on $\mathbb{R}$, then an equivalence relation is defined by $\psi \sim \phi$ iff $\int_{-\infty}^{\infty} |\psi(x) - \phi(x)|\mathrm{d}x = 0$. Then $L^2(\mathbb{R})$ denotes the set of all equivalence classes. The reason for this detour is that (A.23) does not define an inner product on $\mathcal{L}^2(\mathbb{R})$. Specifically, on $\mathcal{L}^2(\mathbb{R})$ the equality $\int_{-\infty}^{\infty} |\psi(x)|\mathrm{d}x = 0$ does not imply $\psi = 0$.

The norm on a Hilbert space can also be used to introduce a norm on the set of operators and functionals.

Definition A.13. Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{O}(\mathcal{H})$ denote the set of operators on $\mathcal{H}$. The *operator norm* on $\mathcal{O}(\mathcal{H})$ is defined as

$$\|A\| := \sup\{\|A\psi\| \mid \|\psi\| \leq 1\}. \tag{A.27}$$

An operator is called *bounded* if $\|A\| < \infty$. A bounded operator A is called *positive* if $\langle \psi, A\psi \rangle \geq 0$ for all $\psi \in \mathcal{H}$. The set of all bounded operators is denoted $\mathcal{B}(\mathcal{H})$ and the set of all positive operators is denoted $\mathcal{B}^+(\mathcal{H})$.

Let $\mathrm{Hom}(\mathcal{H}, F)$ be the set of all functionals on $\mathcal{H}$. The standard norm on this set is defined as

$$\|f\| := \sup\{|f(\psi)| \mid \|\psi\| \leq 1\}. \tag{A.28}$$

A functional is called *bounded* if there exists a positive constant $c \in \mathbb{R}$ such that

$$|f(\psi)| \leq c\|\psi\| \; \forall \psi \in \mathcal{H}. \tag{A.29}$$

The set of all bounded linear functionals on $\mathcal{H}$ is denoted $\mathcal{H}^*$ and is called the *dual space*.

Theorem A.1 (Riesz representation theorem). *For every bounded functional f on a Hilbert space $\mathcal{H}$ there exists a unique $\psi_f \in \mathcal{H}$ such that*

$$f(\psi) = \langle \psi_f, \psi \rangle \; \forall \psi \in \mathcal{H}. \tag{A.30}$$

For every bounded sesquilinear form $B : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ there exists a bounded operator A on $\mathcal{H}$ such that

$$B(\psi_1, \psi_2) = \langle A\psi_1, \psi_2 \rangle \; \forall \psi_1, \psi_2 \in \mathcal{H}. \tag{A.31}$$

For interesting unbounded operators it is often required to restrict the domain to a proper subset of the entire Hilbert space. For example, let $\mathcal{H} = L^2(\mathbb{R})$ and consider the position operator X defined by the action

$$(X\psi)(x) := x\psi(x), \tag{A.32}$$

then in general $X\psi$ is not a square-integrable function. Instead, one restricts attention to the domain

$$D(X) = \left\{ \psi \in L^2(\mathbb{R}) \mid X\psi \in L^2(\mathbb{R}) \right\}. \tag{A.33}$$

One may show that this domain lies dense in $\mathcal{H}$. In fact, only operators with dense domain will be considered here.

The additional structure of the inner product allows for the introduction of special kinds of operators.

Definition A.14. The *adjoint* A^* of an operator A with domain $D(A)$ is the unique operator such that

$$\langle A^*\psi, \phi \rangle = \langle \psi, A\phi \rangle \tag{A.34}$$

for all $\phi \in D(A)$ and $\psi \in D(A^*)$ with

$$D(A^*) = \{\phi \in \mathcal{H} \mid D(A) \ni \psi \mapsto \langle \phi, A\psi \rangle \text{ is a bounded functional}\} . \tag{A.35}$$

When $A = A^*$ (and thus $D(A) = D(A^*)$) the operator is called *self-adjoint operator* and when $AA^* = A^*A$ the operator is called *normal*. A normal operator that satisfies $AA^* = \mathbb{1}$ is called *unitary*. The set of all self-adjoint operators on $\mathcal{H}$ is denoted $\mathcal{O}_{\mathrm{sa}}(\mathcal{H})$, and the set of all bounded self-adjoint operators $\mathcal{B}_{\mathrm{sa}}(\mathcal{H})$. When a self-adjoint operator is idempotent, i.e., $A = A^2$, it is called a *projection operator* or simply *projection*. The set of all projection operators is denoted $L(\mathcal{H})$. Two projection operators P_1 and P_2 are called *compatible* if they commute (i.e., $[P_1, P_2] = \mathbb{0}$). If in addition $P_1 P_2 \psi = 0$ for all ψ they are called *orthogonal*.

Example A.13. Let $\mathcal{H}$ be a finite-dimensional Hilbert space. If M_A is the matrix associated with A (with respect to some basis), then the matrix M_{A^*} associated with A^* satisfies

$$M_{A^*} = \overline{M_A}^T, \tag{A.36}$$

where $\overline{M_A}^T$ the transposed complex conjugate of M_A, i.e., $(M_{A^*})_{ij} = \overline{(M_A)_{ji}}$. $\triangle$

Example A.14. Trivial examples of projections are the zero operator $\mathbb{0}$ (taking every vector to the zero vector) and the unit operator $\mathbb{1}$ (taking every vector to itself). A more interesting example is the projection on the line spanned by a non-zero vector ψ:

$$P_\psi \phi := \frac{\langle \psi, \phi \rangle}{\langle \psi, \psi \rangle} \psi. \tag{A.37}$$

$\triangle$

Constructing more interesting examples is easy once the following lemma is recognized.

Lemma A.3. *For any projection operator P the orthocomplement*

$$P^\perp := \mathbb{1} - P \tag{A.38}$$

is again a projection operator. For any pair of projection operators P_1, P_2

$$P_1 P_2 \text{ and } P_1 + P_2 - P_1 P_2 \tag{A.39}$$

are again projection operators if and only if P_1 and P_2 are compatible.

The example (A.37) also indicates where the name 'projection' comes from; a projection operator projects a vector on a linear subspace. In fact, one can unambiguously switch between talking about projection operators and closed linear subspaces, which is the content of the following theorem.

Theorem A.2. *Let $\mathcal{H}$ be a Hilbert space, then a subset $\mathcal{K} \subset \mathcal{H}$ is a closed linear subspace if and only if there exists a projection operator $P \in L(\mathcal{H})$ such that*

$$\mathcal{K} = P\mathcal{H} := \{P\psi \mid \psi \in \mathcal{H}\}. \tag{A.40}$$

Furthermore, this projection operator is unique.

Accordingly, a projection P is called *n-dimensional* just in case $P\mathcal{H}$ is an n-dimensional subspace.

Definition A.15. For an operator A the *spectrum* is defined as the set

$$\sigma(A) := \{a \in F \mid A - a\mathbb{1} \text{ is not invertible}\}. \tag{A.41}$$

A special subset are the *eigenvalues* given by

$$\{a \in F \mid \exists \text{ non-zero } \psi \in \mathcal{H} \text{ s.t. } A\psi = a\psi\}. \tag{A.42}$$

A non-zero vector for which there exists an $a \in F$ such that $A\psi = a\psi$ is called an *eigenvector* of A for the value a.

Lemma A.4. *If $\mathcal{H}$ is finite-dimensional then for every operator A the spectrum coincides with the set of eigenvalues.*

Definition A.16. Let $(\Omega, \mathcal{F})$ be a measurable space and $\mathcal{H}$ a Hilbert space. A *positive operator valued measure (POVM)* is a map $\mu : \mathcal{F} \to \mathcal{B}^+(\mathcal{H})$ that satisfies
  (i) $\mu(\Omega) = \mathbb{1}$,
  (ii) $\mu\left(\bigcup_{i=1}^{\infty} \Delta_i\right) = \sum_{i=1}^{\infty} \mu(\Delta_i)$
for every countable collection $\{\Delta_i \; ; \; i \in \mathbb{N}\}$ of mutually disjoint measurable sets. If the co-domain is $L(\mathcal{H}) \subset \mathcal{B}^+(\mathcal{H})$ then the map is called a *projection valued measure (PVM)*.

Projection valued measures can be used to introduce a particularly useful generalized notion of integration where the value of the integral is an operator on the Hilbert space instead of a complex number. The construction will be helpful in defining the function of an operator.

Consider a PVM μ on $(\mathbb{R}, \mathcal{F})$ with $\mathcal{F}$ the standard Borel-σ-algebra. For every pair of vectors $\psi, \phi \in \mathcal{H}$ define $\mu_{\psi,\phi} : \mathcal{F} \to \mathbb{C}$ as

$$\mu_{\psi,\phi}(\Delta) := \langle \psi, \mu(\Delta)\phi \rangle. \tag{A.43}$$

One may check that for every pair (ψ, ϕ) this is a complex Borel measure. Now let $\mathcal{B}(\mathbb{R})$ denote the set of bounded Borel functions. For every $f \in \mathcal{B}(\mathbb{R})$ the map

$$\mathcal{H} \times \mathcal{H} \ni (\psi, \phi) \mapsto [\psi, \phi]_{\mu,f} := \int f(x)d\mu_{\psi,\phi}(x) \tag{A.44}$$

is a bounded sesquilinear form, where the integral is well-defined because $\mu_{\psi,\phi}$ is a Borel measure. According to Riesz' representation theorem then there exists an operator $A_{\mu,f} \in \mathcal{B}(\mathcal{H})$ such that

$$[\psi, \phi]_{\mu,f} = \langle A_{\mu,f}\psi, \phi \rangle \; \forall \psi, \phi \in \mathcal{H}. \tag{A.45}$$

The map $f \mapsto A_{\mu,f}$ is now taken as the definition for the integral

$$A_{\mu,f} =: \int f(x)d\mu(x). \tag{A.46}$$

Theorem A.3 (Spectral theorem). *Let A be a self-adjoint operator on a Hilbert space $\mathcal{H}$, then there exists a unique PVM μ_A on the spectrum $\sigma(A)$ of A such that*

$$A = \int_{\sigma(A)} a d\mu_A(a). \tag{A.47}$$

If the spectrum is discrete this can be written as

$$A = \sum_{a \in \sigma(A)} a \mu_A(\{a\}). \tag{A.48}$$

This theorem shows that for a bounded self-adjoint operator A the map $\mathcal{B}(\mathbb{R}) \ni f \mapsto f(A)$ with

$$f(A) := \int_{\sigma(A)} f(a)d\mu_A(a) \tag{A.49}$$

is well-defined. This is the so-called Borel functional calculus. If the spectrum is discrete this definition can be written as

$$f(A) := \sum_{a \in \sigma(A)} f(a)\mu_A(\{a\}). \tag{A.50}$$

Consequently, $\sigma(f(A)) = \{f(a) \mid a \in \sigma(A)\}$.

The following definition defines the trace operation which plays a crucial role in the formulation of the Born rule.

Definition A.17. Let $\mathcal{H}$ be a Hilbert space with orthonormal basis $(e_i)_{i \in I}$. An operator $A \in \mathcal{B}(\mathcal{H})$ is said to of *trace class* if

$$\sum_{i \in I} \langle e_i, |A|e_i \rangle < \infty. \tag{A.51}$$

The set of all trace class operators is denoted $\mathcal{T}(\mathcal{H})$. For every $T \in \mathcal{T}(\mathcal{H})$ and every orthonormal basis $(e_i)_{i \in I}$ the sum

$$\text{Tr}(T) := \sum_{i \in I} \langle e_i, T e_i \rangle \tag{A.52}$$

converges and is independent of the choice of the basis. The outcome $\text{Tr}(T)$ is called the *trace* of T.

The trace operation has several elegant properties that are useful.

Lemma A.5. *Let $T \in \mathcal{T}(\mathcal{H})$ and $A \in \mathcal{B}(\mathcal{H})$, then both AT and TA are trace class operators for which the following relations hold*
   (i) $\text{Tr}(AT) = \text{Tr}(TA)$,
   (ii) $|\text{Tr}(AT)| \le \|A\| \text{Tr}(|T|)$.

The following lemma establishes a property of polynomials of operators that is used in the proof of Theorem 7.2.

Lemma A.6. *Let $A \in \mathcal{B}(\mathcal{H})$ and p be a polynomial. Then for every $\epsilon > 0$ there exists a $\delta > 0$ such that*
$$\|p(A) - p(A')\| < \epsilon \tag{A.53}$$
for all $A' \in \mathcal{B}(\mathcal{H})$ with $\|A - A'\| < \delta$.

Proof. First consider the case where $p_n(x) := x^n$. For $n = 1$ the assertion is trivial (choose $\delta = \epsilon$). Now suppose it is true for all p_m with $m \in \{1, \ldots, n\}$, then it also holds for p_{n+1}. This is shown using the estimate

$$
\begin{aligned}
\|p_{n+1}(A) - p_{n+1}(A')\| =& \|A^{n+1} - A'^{n+1}\| \\
=& \|(A^n - A'^n)(A' - A) + A^n(A - A') + (A^n - A'^n)A\| \\
\le& \|A^n - A'^n\|\|A' - A\| + \|A^n\|\|A - A'\| \\
& + \|A^n - A'^n\|\|A\|.
\end{aligned}
\tag{A.54}
$$

Because the lemma holds for p_1 one can choose δ_1 such that

$$\|A - A'\| < \min\left(\sqrt{\tfrac{\epsilon}{3}}, \tfrac{\epsilon}{3\|A^n\|}\right) \tag{A.55}$$

and because the lemma holds for p_n one can choose δ_n such that

$$\|A^n - A'^n\| < \min\left(\sqrt{\tfrac{\epsilon}{3}}, \tfrac{\epsilon}{3\|A\|}\right). \tag{A.56}$$

Then with the choice $\delta = \min(\delta_1, \delta_n)$ each of the three terms in (A.54) becomes smaller than $\epsilon/3$. This proves that the lemma also holds for p_{n+1}.

Now consider the general case where $p(x) := a_0 + a_1 x + \ldots + a_n x^n$. Then

$$
\begin{aligned}
\|p(A) - p(A')\| &= \left\| \sum_{m=1}^{n} a_m (A^m - A'^m) \right\| \\
&\leq \sum_{m=1}^{n} |a_m| \left\| A^m - A'^m \right\|.
\end{aligned}
\tag{A.57}
$$

Now set $a := \max(|a_1|, \ldots, |a_n|)$. For every $m \in \{1, \ldots, n\}$ there is a δ_m such that $\|A^m - A'^m\| < \frac{\epsilon}{n|a|}$. Choosing $\delta = \min(\delta_1, \ldots, \delta_n)$ ensures that every term in the sum in (A.57) is smaller than $\frac{\epsilon}{n}$ and so the total sum is smaller than ϵ. This completes the proof. $\qquad\square$

Definition A.18. An *algebra* is a vector space $\mathcal{C}$ with a multiplication operation $\mathcal{C} \times \mathcal{C} \ni (A, B) \mapsto AB \in \mathcal{C}$ such that $\mathcal{C}$ is turned into a ring[2] and for all scalars λ and $A, B \in \mathcal{C}$:

$$
\lambda(AB) = (\lambda A)B = A(\lambda B).
\tag{A.58}
$$

The algebra is called *Abelian* if $AB = BA$ for all $A, B \in \mathcal{C}$. A *Banach algebra* is an algebra $\mathcal{C}$ with a norm $\|.\|$ that turns $\mathcal{C}$ into a Banach space and such that

$$
\|AB\| \leq \|A\|\|B\| \quad \forall A, B \in \mathcal{C}.
\tag{A.59}
$$

An algebra is called a **-algebra* if it is a vector space over the complex numbers and has a map $\mathcal{C} \ni A \mapsto A^* \in \mathcal{C}$ such that for all $A, B \in \mathcal{C}$ and $\lambda \in \mathbb{C}$
  (i) $(A^*)^* = A$;
  (ii) $(AB)^* = B^* A^*$;
  (iii) $(\lambda A + B)^* = \overline{\lambda} A^* + B^*$.
A *-algebra that is also a Banach algebra and satisfies

$$
\|A^* A\| = \|A^*\|\|A\| = \|A\|^2
\tag{A.60}
$$

for all $A \in \mathcal{C}$ is called a *C*-algebra*.

Example A.15. The generic example of a C*-algebra is the set of all bounded operators $\mathcal{B}(\mathcal{H})$ on a Hilbert space $\mathcal{H}$ with the norm given by (A.27) and the *-operation is taking the adjoint. The subset

$$
\mathbb{C}\mathbb{1} := \{\lambda \mathbb{1} \mid \lambda \in \mathbb{C}\}
\tag{A.61}
$$

---

[2]I adopt the convention that a ring is assumed to have a unit element w.r.t. multiplication. That is, there is an element $\mathbb{1}$ such that $\mathbb{1}A = A\mathbb{1} = A$ for all $A \in \mathcal{C}$. Although this is often done in ring theory, the reader may be warned that in most of the work on *-algebras this convention is usually not made. If the reader has a background in this kind of work, (s)he is advised to read "unital algebra" whenever I write "algebra".

is an Abelian C*-(sub)algebra. In the special case that $\mathcal{H} = \mathbb{C}^n$, $\mathcal{B}(\mathcal{H})$ coincides with $M_n(\mathbb{C})$. △

Example A.16. The set $\mathcal{T}(\mathcal{H})$ of all trace class operators on a Hilbert space $\mathcal{H}$ is a *-algebra. Let $(e_i)_{i \in I}$ be an orthonormal basis for $\mathcal{H}$, then the norm

$$\|A\|_1 := \sum_{i \in I} \langle e_i, |A|e_i \rangle \tag{A.62}$$

turns $\mathcal{T}(\mathcal{H})$ into a Banach algebra. △

Definition A.19. Let $\mathcal{S}$ be any subset of $\mathcal{B}(\mathcal{H})$, then the *commutant* of $\mathcal{S}$ is

$$\mathcal{S}' := \{ A \in \mathcal{B}(\mathcal{H}) \mid [A, S] = \mathbb{0} \; \forall S \in \mathcal{S} \}. \tag{A.63}$$

A C*-subalgebra $\mathcal{A}$ of $\mathcal{B}(\mathcal{H})$ is called a *von Neumann algebra* if it is equal to its double commutant, i.e., $\mathcal{A} = (\mathcal{A}')' =: \mathcal{A}''$. For a set $\mathcal{S} \subset \mathcal{B}(\mathcal{H})$ the smallest von Neumann algebra containing $\mathcal{S}$ is denoted $\mathcal{Alg}(\mathcal{S})$.

Example A.17. Suppose $\mathcal{H} = \mathbb{C}^2$ and let P be a projection operator on $\mathcal{H}$. Then

$$\mathcal{Alg}(\{P\}) = \left\{ \lambda_1 P + \lambda_2 P^\perp \; \middle| \; \lambda_1, \lambda_2 \in \mathbb{C} \right\}. \tag{A.64}$$

If $P = \mathbb{0}$ or $P = \mathbb{1}$, this algebra is isomorphic to $\mathbb{C}$. In other cases it is isomorphic (as a vector space) to $\mathbb{C}^2$. It may be shown that every C*-subalgebra of $M_2(\mathbb{C})$ is of this form. △

A C*-algebra that is a subalgebra of $\mathcal{B}(\mathcal{H})$ for some Hilbert space is called a *concrete* C*-algebra. This indicates that these examples are really the paradigm examples. And in line with this paradigm, it is not surprising that much of the terminology for operators on a Hilbert space can be carried over to elements of a *-algebra.

Definition A.20. Let $\mathcal{C}$ be a *-algebra. An element $A \in \mathcal{C}$ is called *self-adjoint* if $A = A^*$ and a *projection* if in addition $A = A^2$. The element is called *invertible* if there is an $A^{-1} \in \mathcal{C}$ such that $A^{-1}A = AA^{-1} = \mathbb{1}$. The spectrum $\sigma(A)$ of A is the set of all complex numbers λ for which $A - \lambda \mathbb{1}$ is not invertible. A self-adjoint element A is called *positive* (denoted $A \geq 0$) if $\sigma(A) \subset [0, \infty)$.

Definition A.21. Let $\mathcal{C}$ be a *-algebra, then a functional $\phi : \mathcal{C} \to \mathbb{C}$ is called *positive* if $\phi(A) \geq 0$ for all $A \geq 0$. The functional is called a *state* if it is positive and $\phi(\mathbb{1}) = 1$. A functional is called *multiplicative* if for all $A_1, A_2 \in \mathcal{C}$ $\phi(A_1 A_2) = \phi(A_1)\phi(A_2)$.

Definition A.22. A state ϕ on a von Neumann algebra $\mathcal{C}$ is called *normal* if for any monotone increasing net of operators $\{A_i\}$ the net $\{\phi(A_i)\}$ converges to $\phi(A)$ with A the least upper bound of $\{A_i\}$.

A.3 Lattice theory

In section A.1 lattices have been defined as a particular kind of algebraic structures. Here, a different approach is used, which starts with the notion of a relation on a set.

Definition A.23. A *relation* R on a set X is a subset $\mathcal{R}$ of $X \times X$. Whenever an element $(x_1, x_2) \in X \times X$ is an element of $\mathcal{R}$ this is indicated with the notation $x_1 R x_2$. A relation is called

-) *reflexive* if $\forall x \in X$: xRx,
-) *irreflexive* if $\nexists x \in X$ such that xRx,
-) *transitive* if $\forall x_1, x_2, x_3 \in X$: if $x_1 R x_2$ and $x_2 R x_3$, then $x_1 R x_3$,
-) *symmetric* if $\forall x_1, x_2 \in X$: if $x_1 R x_2$, then $x_2 R x_1$,
-) *antisymmetric* if $\forall x_1, x_2 \in X$: if $x_1 R x_2$ and $x_2 R x_1$, then $x_1 = x_2$.

If a relation is both reflexive and transitive it is called a *preorder* and one often writes $\leq$ instead of R. If the preorder is also antisymmetric it is called a *partial order*. If a preorder is symmetric it is called an *equivalence relation* and one often writes $\sim$ or $=$ instead of R.

Example A.18. Let $\mathbb{N}$ be the natural numbers then

$$n \sim m \text{ iff } n + m \text{ is even} \tag{A.65}$$

is an equivalence relation. The usual order given by

$$n \leq m \text{ iff } \exists k \in \mathbb{N} \text{ s.t. } n + k = m \tag{A.66}$$

turns $\mathbb{N}$ into a partial ordered set. A less trivial example is the relation

$$n \leq m \text{ iff } \exists k \in \mathbb{N} \backslash \{0\} \text{ s.t. } nk = m. \tag{A.67}$$

$\triangle$

Example A.19. Let $(S, \wedge)$ be a semilattice. Then

$$s_1 \leq s_2 \text{ iff } s_1 \wedge s_2 = s_1 \tag{A.68}$$

is a partial order. $\triangle$

This example indicates that partial ordered sets form the backbone of lattice theory, as meets give rise to partial orders. Conversely, partial orders may give rise to meets as established by the following definition.

Definition A.24. Let $(X, \leq)$ be a partial ordered set. Two elements $x_1, x_2 \in X$ are said to have an *upper bound* if there exists an element $x_3 \in X$ such that $x_1 \leq x_3$ and $x_2 \leq x_3$. If the orders are reversed x_3 is called a *lower bound*. A partial ordered set in which all pairs have an upper bound is called a *directed set*. A partial ordered set $(X, \leq)$ is said to have meets if for all $x_1, x_2 \in X$ there is an element in X denoted $x_1 \wedge x_2$ that is a lower bound and satisfies

$$\forall x \in X: \text{ if } x \leq x_1 \text{ and } x \leq x_2, \text{ then } x \leq x_1 \wedge x_2. \tag{A.69}$$

The partial ordered set is said to have joins if for all $x_1, x_2 \in X$ there is an element in X denoted $x_1 \vee x_2$ that is an upper bound and satisfies

$$\forall x \in X: \text{ if } x_1 \leq x \text{ and } x_2 \leq x, \text{ then } x_1 \vee x_2 \leq x. \tag{A.70}$$

A partial ordered set that has both joins and meets is called a *lattice*. A lattice is called *bounded* if it has elements $\top, \bot$ (called top and bottom) that satisfy

$$\forall x \in X: \ \bot \leq x \leq \top. \tag{A.71}$$

One may show that if a partial ordered set has joins and meets, that these are uniquely determined by the partial order. Further, one can show that a partial ordered set that has meets or joins is a semilattice in the sense of Definition A.2. And if the it has both joins and meets then it is also a lattice in the sense of Definition A.4. To obtain a lattice structure on a given set, one may thus first look for a partial order. The following definition and theorem show how one obtains a partial ordered set from a preordered set.

Definition A.25. If $(X, \sim)$ is a set with an equivalence relation, then for every $x \in X$

$$[x] := \left\{ x' \in X \mid x \sim x' \right\} \tag{A.72}$$

is called the *equivalence class of* x. An element of $[x]$ is called a *representative* of the class $[x]$.

Theorem A.4. *If $(X, \sim)$ is a set with an equivalence relation, then $\{[x] \mid x \in X\}$ is a partition of X. For every preordered set $(X, \leq)$ the relation*

$$x_1 \sim x_2 \text{ iff } x_1 \leq x_2 \text{ and } x_2 \leq x_1 \tag{A.73}$$

is an equivalence relation. The preorder on X gives rise to a partial order on $\{[x] \mid x \in X\}$ via

$$[x_1] \leq [x_2] \text{ iff } x_1 \leq x_2. \tag{A.74}$$

Example A.20. Define the equivalence relation on $\mathbb{N}$ as

$$n \sim m \text{ iff } n + m \text{ is even.} \tag{A.75}$$

Then $[1]$ is the set of all odd numbers and $[2]$ is the set of all even numbers. $\triangle$

Definition A.26. Let $(X, \leq)$ be a partial ordered set with least element $\perp$. An element $a \neq \perp$ is called an *atom* if for every $x \in X$ it follows from $\perp \leq x \leq a$ that either $x = \perp$ or $x = a$. The set X is called *atomic* if for every $x \in X$ with $x \neq \perp$ there exists an atom a such that $a \leq x$. The set is called *atomistic* if every $x \neq \perp$ is equal to the join of all the atoms below it.

Example A.21. Let X be a set and $(\mathcal{P}(X), \leq)$ the lattice of subsets of X with partial order given by set inclusion. The atoms are the singleton sets. Since $\Delta = \bigcup_{x \in \Delta}\{x\}$ for every $\Delta \in \mathcal{P}(X)$ the lattice is atomistic. $\triangle$

Example A.22. Consider the set of natural numbers without zero and with the partial order given by

$$n \leq m \text{ iff } m - n \in \mathbb{N} \text{ and } \frac{m}{n} \in \mathbb{N}. \tag{A.76}$$

The bottom element is 1 and the atoms are given by the prime numbers. The set is atomic, but not atomistic. As a counterexample note that the atoms below 12 are 2 and 3, but that $2 \vee 3 = 6$. $\triangle$

Lattices can be further qualified in terms of additional operations that may be defined on them and how these operations relate to each other.

Definition A.27. A unary operation $a \mapsto a^c$ on a bounded lattice $(L, \leq, \wedge, \vee, \perp, \top)$ is called a *complement* if it satisfies

$$\forall a \in L : \ a \wedge a^c = \perp, \ a \vee a^c = \top. \tag{A.77}$$

In that case L is called a *complemented lattice*. A *Boolean algebra* is a distributive complemented lattice. A *modal algebra* is a Boolean algebra with a unary operation $a \mapsto \Diamond a$ that satisfies

$$\forall a, b \in L : \ \Diamond(a \vee b) = \Diamond a \vee \Diamond b. \tag{A.78}$$

If in addition

$$\Diamond \perp = \perp, \tag{A.79}$$

then the lattice is called a *normal modal algebra*. A binary operation $(a, b) \mapsto a \rightarrow b$ is called a *relative pseudo-complement* if it satisfies

(i) $a \wedge (a \to b) \le b$,

(ii) for all c such that $a \wedge c \le b$: $c \le a \to b$.

A *Heyting algebra* is a bounded lattice with a relative pseudo-complement. On a Heyting algebra the unary operation

$$\neg a := a \to \bot \tag{A.80}$$

is called the *pseudo-complement.*

Example A.23. Let R be a Boolean ring. Then the lattice operations defined in Example A.5 turn R into a Boolean algebra with the complement given by

$$x^c := 1 + x. \tag{A.81}$$

Conversely, every Boolean algebra $(B, \wedge, \vee, ^c, \bot, \top)$ is turned into a Boolean ring with the definitions

$$\begin{aligned} a + b &:= (a \wedge b^c) \vee (a^c \wedge b), \\ ab &:= a \wedge b. \end{aligned} \tag{A.82}$$

$\triangle$

Example A.24. The standard example of a non-distributive lattice is that of the projection operators on a Hilbert space $\mathcal{H}$ with the partial order given by

$$P_1 \le P_2 \text{ iff } P_1 \mathcal{H} \subset P_2 \mathcal{H}. \tag{A.83}$$

A non-standard example is obtained by defining the partial order on $\mathbb{N} \backslash \{0\}$ as

$$n \le m \text{ iff } \begin{cases} m - n \in \mathbb{N} \text{ and } \gcd(n, m) > 1 & \text{or} \\ n = 1 \end{cases} \tag{A.84}$$

where gcd denotes the greatest common divisor. An example of the failure of distributivity is

$$(2 \vee 3) \wedge 4 = 6 \wedge 4 = 4 \ne 2 = 2 \vee 1 = (2 \wedge 4) \vee (3 \wedge 4). \tag{A.85}$$

$\triangle$

Example A.25. Consider the following three Hasse diagrams.
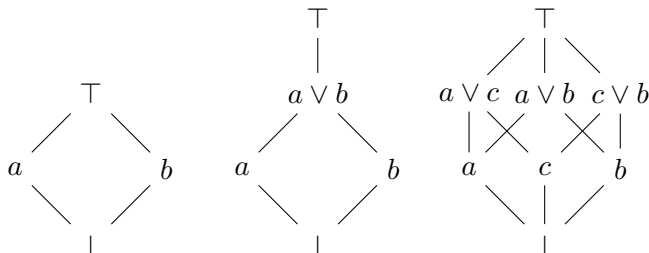
The first and last define a Boolean algebra, while the one in the middle is only a Heyting algebra. In the first b is the complement of a and in the second b is the pseudo-complement of a. That $a \vee b$ and $\top$ are not the same element in the second can be seen as an example of the failure of the law of excluded middle. The third may be thought of as a classical interpretation of the second. The element c then represents the 'third option'. $\triangle$

Example A.26. Let $(\Omega, \mathcal{F})$ be a measurable space, i.e., a set Ω with $\mathcal{F}$ a collection of subsets of Ω that satisfies
   (i) $\Omega \in \mathcal{F}$,
   (ii) if $\Delta \in \mathcal{F}$, then $\Omega \backslash \Delta \in \mathcal{F}$,
   (iii) if $\Delta_n \in \mathcal{F}$ for all $n \in \mathbb{N}$, then $\bigcup_{n \in \mathbb{N}} \Delta_n \in \mathcal{F}$.
Then $\mathcal{F}$ is a Boolean algebra with the partial order given by set inclusion. $\triangle$

Example A.27. Let (X, τ) be a topological space, i.e., a set X with τ a collection of subsets of X that satisfies
   (i) $X \in \tau$ and $\varnothing \in \tau$,
   (ii) if $U_1, U_2 \in \tau$, then $U_1 \cap U_2 \in \tau$,
   (iii) if $U_i \in \tau$ for all $i \in I$, then $\bigcup_{i \in I} U_i \in \tau$.
Then τ is a bounded lattice with the partial order given by set inclusion. The operation

$$U_1 \to U_2 := \text{int}(U_1^c \cup U_2) \tag{A.86}$$

is a relative pseudo-complement that turns τ into a Heyting algebra. The fact that in general $U \cup \in (U^c) \neq X$ represents the idea that the law of excluded middle does not hold in general. $\triangle$

A collection of subsets of some set that is closed under intersection, unions and complements is the prototype of a Boolean algebra. It was shown by Stone (1936) that every Boolean algebra is isomorphic to a Boolean algebra of this form. The final part of this appendix is devoted to explaining the construction behind this result. To identify a Boolean algebra with an algebra of subsets, one first has to find these sets. The starting point is the set of ultrafilters on the algebra.

Definition A.28. An *ultrafilter* $\mathcal{U}$ on a Boolean algebra B is a subset of B such that
   (i) $\top \in \mathcal{U}$,
   (ii) if $a_1 \leq a_2$ and $a_1 \in \mathcal{U}$, then $a_2 \in \mathcal{U}$,
   (iii) if $a_1, a_2 \in \mathcal{U}$, then $a_1 \wedge a_2 \in \mathcal{U}$,
   (iv) for every $a \in$ either $a \in \mathcal{U}$ or $a^c \in \mathcal{U}$.
An ultrafilter $\mathcal{U}$ on B is called a *principal ultrafilter* if there is an atom $a \in B$ such that $\mathcal{U}$ is given by the up set of a, i.e.,

$$\mathcal{U} = \uparrow a := \left\{ a' \in B \mid a \leq a' \right\}. \tag{A.87}$$

An ultrafilter $\mathcal{U}$ is called a *free ultrafilter* if it is not a principal ultrafilter. The set of all ultrafilters on B is denoted $S(B)$.

The set $S(B)$ is turned into a topological space by the introduction of the so-called *Stone topology*. This is the topology generated by the basis

$$\Delta_a := \{\mathcal{U} \in S(B) \,|\, a \in \mathcal{U}\} \tag{A.88}$$

for $a \in B$. The topological space $S(B)$ is called the *Stone space*. One can show that the set $\{\Delta_a \subset S(B) \,|\, a \in B\}$ precisely is the set of sets that are both open and closed. Furthermore, the following lemma holds.

Lemma A.7. *Let B be a Boolean algebra and $S(B)$ its associated Stone space. Then the map $B \ni a \mapsto \Delta_a \subset S(B)$ satisfies*

$$\begin{aligned} \Delta_{a_1 \vee a_2} &= \Delta_{a_1} \cup \Delta_{a_2}, \\ \Delta_{a_1 \wedge a_2} &= \Delta_{a_1} \cap \Delta_{a_2}, \\ \Delta_{a^c} &= \Delta_a^c, \end{aligned} \tag{A.89}$$

for all $a, a_1, a_2 \in B$.

This lemma thus establishes that the map $a \mapsto \Delta_a$ is a homomorphism. Stone's theorem now states that it is in fact an isomorphism. The technical culprit is to show that the map is injective. And this actually requires the Boolean prime ideal theorem (or something equivalent to it).

Samenvatting

I. Kansen: interpretatie en formalisme

Van alle wiskunde die we in ons dagelijks leven tegenkomen is, na de rekenkunde, de kansrekening waarschijnlijk de meest voorkomende. Daarbij kan men bijvoorbeeld denken aan kansspelen zoals loterijen, casinospelen en de WK-pool, maar ook aan verzekeringen, pensioenfondsen en aandelenkoersen. Ook is het tegenwoordig gebruikelijk dat bij het weerbericht kansen op zon en regen in getal worden genoemd. Behalve in de meteorologie, speelt kanstheorie ook een grote rol in andere wetenschappelijke theorieën zoals de genetica of de statistische thermodynamica. Verder is het gebruik van statistiek niet meer weg te denken bij de toetsing van wetenschappelijke theorieën.

Hoewel het kansbegrip veelvuldig voorkomt, is het bepalen van wat kansen zijn een lastige filosofische aangelegenheid. Hierbij doel ik niet op de vraag naar het bepalen van de waarde van een kans (wat is de kans dat een munt op kop terecht komt wanneer deze opgeworpen wordt?), maar de betekenis van het woord "kans" zelf (wat wordt er bedoeld met de uitspraak dat de kans op kop een half is?). Een reden hiervoor is dat kansen zelf niet eenduidig waarneembaar zijn. In het weerbericht kan men bijvoorbeeld stellen dat er morgen 90% kans op regen is, maar deze uitspraak wordt niet weerlegd als het in feite droog blijft. Immers, daar was ook 10% kans op.

Mede door deze onmogelijkheid om een directe empirische betekenis aan het kansbegrip te geven, is er een breed scala aan interpretaties. Een belangrijk onderscheid in interpretaties is of kansuitspraken alleen betekenis hebben in de context van een reeks vergelijkbare gebeurtenissen, of dat ze ook van toepassing zijn op eenmalige gebeurtenissen. In het eerste geval gaat de bewering dat een munt kans een half heeft om op kop te landen, over het (eventueel hypothetische) geval dat deze munt heel vaak (potentieel oneindig vaak) opgeworpen word. De bewering impliceert dan dat in (ongeveer) de helft van de gevallen de munt op kop land. Volgens frequentisten is dit de betekenis van kans: een kans is de relatieve frequentie

waarmee iets gebeurt. Er bestaan verschillende stromingen van deze interpretatie die op verschillende manieren de gegeven definitie precies proberen te maken. Eén van de stromingen die voortvloeit uit het frequentisme is de geneigdheidsinterpretatie. Volgens deze interpretatie zijn de relatieve frequenties enkel een manifestatie van kansen. Een munt heeft dan een kans om op kop te landen onafhankelijk of deze daadwerkelijk opgeworpen word. Ook in het geval van een enkele muntworp is het kansbegrip dan betekenisvol. Kansen zijn dan dispositionele eigenschappen (d.w.z, eigenschappen die een bepaalde geneigdheid uitdrukken) vergelijkbaar met dat breekbaarheid een eigenschap van een glas is ook als deze nooit zou breken.

De genoemde interpretaties hebben de eigenschap dat ze kansen plaatsen in de wereld. Volgens het frequentisme zijn het eigenschappen van een reeks gebeurtenissen en volgens de geneigdheidsinterpretatie zijn het eigenschappen van systemen zelf. Hiertegenover staan variaties van de Bayesiaanse interpretatie waarin kansuitspraken worden gekoppeld aan de overtuigingen van de persoon die de uitspraak doet. In deze context worden kansen ook wel graden van geloof genoemd: een kwantificering van de mate waarin men overtuigd is van de waarheid van een bepaalde uitspraak. Wanneer iemand de zin "morgen is er 90% kans op regen" uitspreekt, betekent dit dat deze persoon er veel vertrouwen in heeft dat het morgen gaat regenen, maar er zijn dan niet per se implicaties voor de feitelijke condities van het weer. Of die implicaties er zijn of niet hangt er van af hoe de persoon in kwestie aan zijn of haar vertrouwen is gekomen, bijvoorbeeld door het doen van meteorologisch onderzoek, of door het kijken in een glazen bol.[3]

Naast de genoemde interpretaties zijn er natuurlijk nog verscheidene anderen. Voor de verschillende voorkomens van kansen (in kansspelen, in wetenschappelijke theorieën) kunnen verschillende interpretaties beter of juist slechter passen. Er is echter één ding wat al deze kansbegrippen en hun toepassingen met elkaar verbindt: de kanstheorie. Of men nu kansen wilt berekenen voor de paardenrennen, voor het bepalen van de verzekeringspremie, of voor de luchtvochtigheid voor het komende uur, de rekenregels zijn (zeker sinds de axiomatisering van Kolmogorov uit 1933) steeds hetzelfde.

Op de toepasbaarheid van de axioma's van Kolmogorov (ook wel de klassieke kanstheorie genoemd) bestaat een eigenaardige uitzondering: de quantummechanica. Deze theorie heeft zelfs geleid tot de ontwikkeling van een aparte kanstheorie: de quantumkanstheorie. Hier dient men met enige verwondering bij stil te staan. Door de eeuwen heen is de ontwikkeling van de wiskundige theorie van kansen hand in hand gegaan met interpretationele overwegingen. Hoe deze theorie er ook uit zou

---

[3]In eerste instantie lijkt het misschien vreemd om de mate van geloof te verbinden aan getallen. Deze stap wordt vaak onderbouwd door een koppeling met (potentieel) gedrag van mensen te maken. De persoon die stelt dat er morgen 90% kans op regen is zou dan bijvoorbeeld bereid zijn om maximaal 90 cent te betalen voor een lot dat recht geeft op 1 euro wanneer het morgen daadwerkelijk regent en waardeloos is wanneer het droog blijft.

komen te zien, ze zou compatibel moeten zijn met de intuïties die we over kansen hebben. De quantummechanica aan de andere kant, is voor een geheel andere reden ontwikkeld, namelijk, voor de correcte beschrijving van het gedrag van atomaire deeltjes. Het zou nog een redelijke tijd duren na Born's introductie van kansen in de theorie in 1926 voordat duidelijk werd dat deze kansen niet op dezelfde manier samenhangen als de kansen in de klassieke kanstheorie.

De situatie leidt tot interessante vragen. Zijn de eerder genoemde interpretaties ook toepasbaar voor de kansen in de quantummechanica? Of vereist een nieuw formalisme een nieuw kansbegrip? Kunnen verschillende kanstheorieën zinvol naast elkaar bestaan? Deze vragen zijn alleen betekenisvol wanneer duidelijk is in welke zin de quantumkanstheorie een nieuwe kanstheorie is. Wiskundig is dit een heldere opgave: men onderzoekt dan de verschillen in definities van klassieke en quantum-kansruimten, en hun mogelijke instanties. In dit proefschrift ligt de nadruk meer op de conceptuele verschillen. Met name de vraag wat het betekent om de quantumkanstheorie te gebruiken voor de beschrijving van kansen komt aan bod.

Om dit onderzoek in goede banen te leiden is het eerste deel van het proefschrift volledig gewijd aan het in kaart brengen van de nodige achtergrondkennis. Er wordt een weg bewandeld van de introductie van kansen in de quantummechanica naar de axiomatisatie van de theorie door von Neumann in 1932. Het formalisme van de quantumkanstheorie wordt gerelateerd aan de axioma's van de quantummechanica, en enkele zinvolle karakteriseringen van quantumkansruimtes worden gegeven. Ook wordt een formulering van de quantumkanstheorie gepresenteerd waarbinnen de klassieke kanstheorie zich voordoet als een speciaal geval. Dit betekent dat de quantumkanstheorie opgevat kan worden als een generalisatie van de klassieke kanstheorie in plaats van als een rivaal. Het materiaal gepresenteerd in deel I is dus veelal niet nieuw en de rest van deze samenvatting gaat daarom over de delen II en III.

II. Klassieke representaties van de quantumkanstheorie

Een onderzoeking van de quantumkanstheorie kan niet geschieden zonder aandacht te besteden aan de grondslagen van haar belangrijkste toepassing: de quantum-mechanica. Het succes van deze theorie vormt immers een belangrijke rechtvaardiging voor het idee dat het formalisme van de quantumkanstheorie een serieuze kandidaat is voor een kanstheorie. Maar tegelijkertijd betekent dit ook dat deze rechtvaardiging gevoelig is voor problemen in de grondslagen van de quantummechanica. Wanneer fundamentele problemen leiden tot een herformulering van de quantummechanica kan dit gepaard gaan met een herformulering van de quantum-kanstheorie.

Het zou zomaar het geval kunnen zijn dat eigenaardige aspecten van het formalisme van de quantumkanstheorie een artefact zijn van een ongelukkige formulering

van de quantummechanica. In het bijzonder kan men hierbij denken aan de vraag of de quantummechanica volledig is. Sinds de ontwikkeling van de theorie zijn er argumenten naar voren gebracht (in eerste instantie met name door Einstein en Schrödinger) die verdedigen dat de beschrijving die deze theorie geeft wel correct kan zijn, maar niet volledig. Het bekendste voorbeeld is waarschijnlijk de kat van Schrödinger. Schrödinger beargumenteerde dat als we de quantummechanica beschouwen als een volledige universele theorie (dat wil zeggen, geldig voor alle fysische systemen) we moeten accepteren dat de onbepaaldheid van eigenschappen van microscopische systemen (in het algemeen legt bijvoorbeeld de toestand van een electron niet vast waar dit electron zich bevind) overslaat op eigenschappen van macroscopische systemen. De quantummechanica zou dan toelaten dat een kat niet altijd levend of dood is, maar zich soms in een zogenaamde superpositie van deze twee eigenschappen kan bevinden. Schrödinger beschouwde deze conclusie als absurd en concludeerde daarom dat er een vollediger theorie moet bestaan die wel ten allen tijde bepaald dat een kat levend of dood is.

Theorieën die een vollediger beschrijving geven dan de quantummechanica staan bekend als verborgen variabelen theorieën. De terminologie slaat op het idee dat een dergelijke theorie variabelen dient te introduceren die alle eigenschappen van systemen vastleggen. Omdat de quantummechanica zelf deze variabelen niet kent, zijn ze vanuit dat oogpunt verborgen. Dit betekent dus niet dat de variabelen noodzakelijk niet vatbaar zijn voor empirische onderzoekingen. Sterker, in de bekendste verborgen variabelen theorie, de Bohmiaanse mechanica, zijn de verborgen variabelen de positie-coördinaten van alle deeltjes en zijn het precies deze variabelen die worden waargenomen bij een positie-meting.

Het construeren van een verborgen variabelen theorie is niet bepaald een triviale activiteit en er bestaan allerlei resultaten die restricties leggen op zulke theorieën. Deze resultaten staan bekend als zogenaamde 'no-go stellingen' voor verborgen variabelen. De achterliggende gedachte is dat de restricties niet acceptabel zijn vanuit fysisch of filosofisch oogpunt. Uiteraard is niet iedereen het eens met deze lezing en er wordt nog voldoende gewerkt aan verborgen variabelen theorieën. De interesse voor deze stellingen in dit proefschrift komt niet zozeer voort uit de vraag of verborgen variabelen mogelijk zijn, maar uit de vraag van de noodzaak van een quantumkanstheorie. De relevantie van deze stellingen voor deze vraag berust op het feit dat er wiskundig een sterke overeenkomst is tussen de mogelijkheid van een verborgen variabelen theorie en de mogelijkheid om een klassieke representatie van een quantumkansruimte te formuleren.

Deze overeenkomst wordt geïllustreerd door het schema in Figuur S.1. Het bestaan van een verborgen variabelen theorie impliceert dat ieder quantummechanisch systeem evengoed beschreven kan worden door een verborgen variabelen systeem. De probabilistische aspecten van dit systeem kunnen op hun beurt beschreven worden met behulp van een klassieke kansruimte. Dit is een gevolg van de wiskundige

structuur die men doorgaans oplegt aan verborgen variabelen theorieën. Aan de andere kant kunnen de probabilistische eigenschappen van een quantummechanisch systeem beschreven worden met behulp van een quantumkansruimte. Omdat men eist dat de empirische voorspellingen van beide theorieën (quantummechanica en de verborgen variabelen theorie) met elkaar in overeenstemming zijn, geeft de klassieke kansruimte een representatie van de quantumkansruimte.
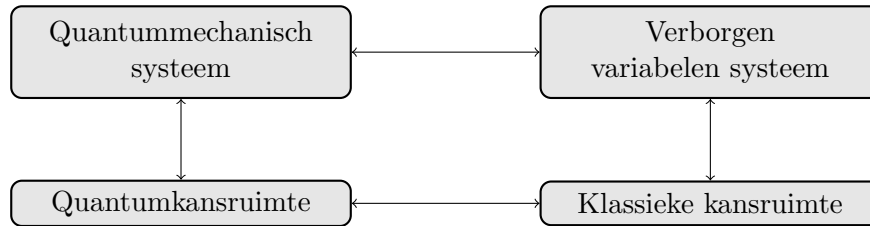
Figuur S.1: Schematische weergave van de relatie tussen een quantumechanische en verborgen variabelen beschrijving van een systeem en hun onderliggende kanstheorieën.

De mogelijkheid van verborgen variabelen theorieën schetst een beeld dat precies het tegenovergestelde is van wat we eerder zagen, namelijk, dat de quantumkanstheorie een speciaal geval is van de klassieke kanstheorie. De situatie, hoewel in eerste instantie mogelijk verwarrend, is niet strijdig. Als analogie voor de dubbele vertaling kan men denken aan vierkanten en cirkels. Het is mogelijk om ieder vierkant binnen een cirkel te plaatsen en het is ook mogelijk om het omgekeerde te doen, maar dat zorgt er natuurlijk niet voor dat cirkels en vierkanten hetzelfde zijn.

De interessantere vraag is of men een klassieke representatie kan construeren die ook inzicht geeft in het waarom van het formalisme van de quantumkansruimte. Om deze vraag te beantwoorden wordt in dit proefschrift de aandacht gevestigd op de stelling van Kochen en Specker uit 1967. In vergelijking met andere no-go stellingen legt deze stelling een grote nadruk op de wiskundige structuur van de quantummechanica en maakt ze nauwelijks gebruik van additionele metafysische aannames zoals localiteitsaannames. De strekking van de stelling is goed uit te leggen aan de hand van wat voorbeelden. Beschouw een experiment waarbij bepaald wordt of een gegeven bal blauw of rood is. De mogelijke uitkomsten van dit experiment zijn schematisch weergegeven in Figuur S.2a. Hierin geeft R de uitkomst "rood" aan en B de uitkomst "blauw". De symbolen $\perp$ en $\top$ representeren respectievelijk een onmogelijke uitkomst (bijvoorbeeld "rood en blauw tegelijkertijd") en een triviale uitkomst (bijvoorbeeld "rood of blauw"). Een lijn tussen twee blokken geeft aan dat de onderste van de twee de bovenste impliceert.

In zowel de klassieke als de quantumkansrekening vormt het diagram in Figuur S.2a de totale collectie van gebeurtenissen waaraan men kansen toekent. Het verschil tussen de twee kanstheorieën komt pas naar voren wanneer het over meer dan
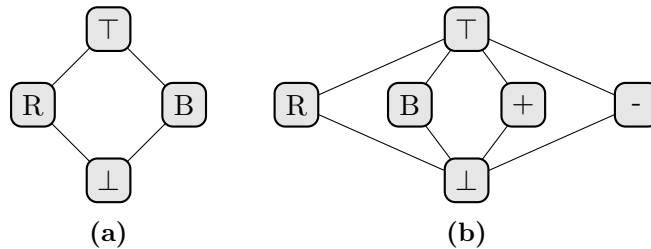
Figuur S.2: Schematische weergave van de tralies van gebeurtenissen voor twee quantumkansruimtes. Links voor een ruimte voor één mogelijk experiment met twee mogelijke uitkomsten, en rechts voor een ruimte voor twee mogelijke experimenten met ieder twee mogelijke uitkomsten.

één mogelijk experiment gaat. Beschouw daarom nu de situatie waarin we kunnen kiezen tussen twee experimenten. De eerste is dezelfde als voorheen, en bij het tweede experiment kunnen we bepalen of er een plus of een min op de bal staat. In de quantumkanstheorie kan zo'n situatie beschreven worden met behulp van het diagram in Figuur S.2b. Dit diagram is beknopt in het specificeren van gebeurtenissen: alleen de gebeurtenissen die zich daadwerkelijk kunnen voordoen in een experiment zijn gespecificeerd. Het is verkregen door twee versies van het eerste diagram aan elkaar te plakken door de gebeurtenissen "rood of blauw" en "plus of min" met elkaar te identificeren en de gebeurtenissen "rood en blauw" en "plus en min" met elkaar te identificeren.

In de klassieke kansrekening wordt meer verlangd van een specificatie. Ook volledige karakteriseringen van de bal dienen opgenomen te worden. In dit geval betekent dat bijvoorbeeld dat, wanneer alleen gekeken wordt naar de kleur van de bal, we aannemen dat er ook dan wel een plus, dan wel een min op de bal staat. Deze aanname is onafhankelijk van de vraag of het mogelijk is om experimenteel vast te stellen dat een bal voldoet aan één van de vier criteria "rood en plus", "rood en min", "blauw en min" of "blauw en plus". Voor klassieke representaties dienen dus mogelijke gebeurtenissen te worden gepostuleerd waarvan niet op voorhand duidelijk is of die experimenteel geverifieerd kunnen worden. In het huidige geval zou een klassieke representatie er uit zien zoals in Figuur S.3. In dit diagram kan bijvoorbeeld R geïnterpreteerd worden als een afkorting voor "rood met een plus of rood met een min". Verder is dan bijvoorbeeld X1 een afkorting voor "rood met een plus of blauw met een min" en X2 een afkorting voor "rood met een min of blauw met een plus".

Het is op voorhand niet duidelijk of een representatie zoals in Figuur S.3 met de gegeven interpretatie altijd mogelijk is voor een gegeven quantumkansruimte. Het voorbeeld geeft aan dat in het geval van twee mogelijke experimenten met ieder twee mogelijke meetuitkomsten het mogelijk is om een klassieke representatie van een quantumkansruimte te construeren. De stelling van Kochen en Specker laat
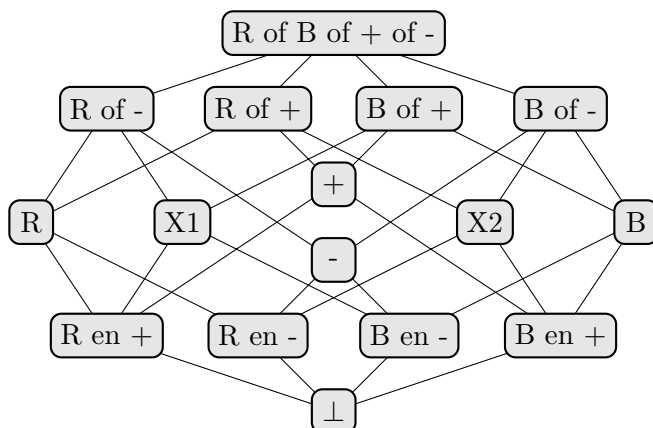
Figuur S.3: Schematische weergave van de tralie van gebeurtenissen voor een klassieke kansruimte voor twee mogelijke experimenten met ieder twee mogelijke meetuitkomsten. Formeel is dit hetzelfde tralie als voor een enkel experiment met vier mogelijke meetuitkomsten.

echter zien dat wanneer we kijken naar experimenten met drie of meer mogelijke meetuitkomsten het in het algemeen niet meer mogelijk is om een klassieke representatie te construeren. Een cruciale aanname daarbij is dat in sommige gevallen twee uitkomsten van verschillende experimenten als dezelfde gebeurtenis gezien worden. Deze aanname is gebaseerd op het feit dat in de quantumkanstheorie in deze gevallen de twee gebeurtenissen dezelfde wiskundige representatie hebben. Maar vanuit een klassiek oogpunt mist deze aanname rechtvaardiging.

In hoofdstuk 6 onderscheid ik twee manieren waarop deze aanname weerlegd kan worden. De eerste is welbekend en werd reeds door Bell in 1966 voorgedragen. Deze optie accepteert dat de wiskundige representatie die gebruikt wordt in de quantumkanstheorie correct is, maar stelt dat deze onvolledig is. Een volledige representatie zou ook de meetcontext (het gekozen experiment) moeten weergeven. Verborgen variabelen theorieën die deze optie volgen worden dan ook wel contextueel genoemd. De tweede optie is gebaseerd op het werk van Meyer, Kent en Clifton rond het jaar 2000. Volgens deze optie is de gebruikte wiskundige representatie niet altijd correct, maar slechts correct bij benadering. Twee verschillende gebeurtenissen hebben dan nooit precies dezelfde wiskundige representatie.

In beide opties is het resultaat dat gebeurtenissen die in eerste instantie met elkaar geïdentificeerd werden nu onderscheiden worden. In de hoofdstukken 6 en 7 laat ik expliciet zien dat beide opties de mogelijkheid geven tot het construeren van klassieke representaties (hoewel deze mogelijkheden bekend zijn, zijn expliciete constructies niet zo bekend). Echter, in beide gevallen hebben de klassieke representaties weinig verklarende kracht voor de quantumkanstheorie. Het probleem is

dat in allebei de gevallen er nieuwe kansfuncties geïntroduceerd worden die geen tegenhanger hebben in de quantumkanstheorie: de quantumkansfuncties vormen een bijzondere deelverzameling van alle kansfuncties in de klassieke representatie.

In beide gevallen is het mogelijk om een karakterisering van de quantumkansfuncties in de klassieke representaties te geven. In de optie waarin meetcontexten expliciet worden opgenomen is deze karakterisering relatief bekend: de quantumkansfuncties zijn precies diegenen die een bepaalde vorm van niet-contextualiteit respecteren. Deze functies voldoen dus precies aan een eigenschap waarvan de weerlegging de basis vormde voor de constructie van de representatie! In hoofdstuk 7 laat ik zien dat in de tweede optie zich een soortgelijke situatie voordoet. Ik bewijs dat in dit geval de quantumkansfunctie precies die functies zijn die aan een bepaalde continuïteitseis voldoen. Met behulp van een resultaat van Appleby uit 2004 beargumenteer ik vervolgens dat deze continuïteitseis zelf niet gerespecteerd wordt in de klassieke representatie op het beschrijvingsniveau van de gebeurtenissen.

In beide gevallen is het dus zo dat de karakterisering van de quantumkansfuncties op gespannen voet staat met de aannames die gebruikt werden voor het construeren van de klassieke representaties. Het accepteren van een voorkeurspositie voor de quantumkansfuncties ondermijnt de aanname die het toeliet om de klassieke representatie überhaupt te construeren. De conclusie is dat de quantumkanstheorie niet goed begrepen kan worden vanuit een klassiek oogpunt. Er zijn dan twee opties: of we moeten accepteren dat de quantumkanstheorie geen correcte volledige beschrijving geeft van alle mogelijke kansfuncties voor een bepaald systeem, of we moeten zoeken naar een andere manier om het formalisme van de quantumkanstheorie te begrijpen. Deze laatste optie wordt onderzocht in het derde deel van het proefschrift.

III. Quantumkanstheorie en de logica van de quantummechanica

Het doel van de onderzoekingen in het derde deel is om een beter begrip te krijgen van het formalisme van de quantumkanstheorie. De methode die hiervoor wordt gehanteerd is het maken van een conceptueel gefundeerde herformulering van het formalisme. Het is hierbij natuurlijk van belang dat de herformulering de originele formulering wel respecteert. Dit is nodig om er voor te zorgen dat de interpretatie van de herformulering ook toepasbaar is op het origineel. Een obstakel in het opnemen van deze eis is dat er weinig consensus is over waar de belangrijkste toepassing van de quantumkanstheorie (de quantummechanica) over gaat. Verschillende interpretaties van de quantummechanica suggereren verschillende herformuleringen van de quantumkanstheorie. Dit probleem wordt in dit proefschrift omzeild door de aandacht te richten op wat al deze interpretaties met elkaar gemeen hebben: de empirische voorspellingen.

In de zoektocht naar een herformulering is het belangrijk om drie componenten

van een kansfunctie te onderscheiden waarvan we ons kunnen afvragen waar die precies aan moeten voldoen. Ten eerste is er het bereik van de functie: de mogelijke waarden voor kansen. Het is redelijk onomstreden om hiervoor de getallen tussen 0 en 1 te kiezen. Ten tweede is er het domein van de functie: de dingen waaraan we kansen toeschrijven. En ten slotte zijn er de rekenkundige regels waaraan de functie dient te voldoen zoals de regel dat de som van kansen van elkaar uitsluitende gebeurtenissen niet groter is dan 1.

De component onder beschouwing hier is het domein van kansfuncties. Het is namelijk dit aspect waarin quantumkansfuncties het meest verschillen van hun klassieke tegenhangers. In de klassieke kansrekening is het doorgaans mogelijk om de wiskundige objecten in het domein (verzamelingen) op een eenduidige manier te koppelen aan mogelijke gebeurtenissen in de wereld. De gebeurtenis bijvoorbeeld dat een dobbelsteenworp resulteert in een even getal wordt gekoppeld aan de verzameling getallen $\{2, 4, 6\}$. Anderzijds kan de gebeurtenis ook geïdentificeerd worden met een propositie, bijvoorbeeld, de propositie uitgedrukt door de zin "de worp van de dobbelsteen resulteert in een even getal". Uit proposities kunnen we nieuwe proposities bouwen door ze aan elkaar te koppelen met de woorden "en" en "of". Aan de andere kant kan men verzamelingen met elkaar verenigingen of doorsnijden. Het mooie van deze soorten bewerkingen is dat ze met elkaar in harmonie zijn. Bijvoorbeeld, de zin "de worp van de dobbelsteen resulteert in een even getal en de worp van de dobbelsteen resulteert in een getal kleiner dan vier" kan geherformuleerd worden tot "de worp van de dobbelsteen resulteert in het getal twee". Anderzijds vinden we aan de kant van de verzamelingen dat

$$\{2, 4, 6\} \cap \{1, 2, 3\} = \{2\}.$$

Meer algemeen kan men bewijzen dat het gebruik van deze verzamelingstheoretische structuur consistent is met de (klassieke) propositielogica. Het is dus mogelijk om het domein van een klassieke kansfunctie te begrijpen als een verzameling proposities die (onder andere) gesloten is onder het vormen van disjuncties en conjuncties (samenstellingen met "en" en "of"). Dit geldt niet voor de quantumkanstheorie. Er is wel een duidelijke regel over hoe gebeurtenissen geïdentificeerd dienen te worden met projectie-operatoren (de wiskundige objecten in het domein van de quantumkansfunctie). Dit is noodzakelijk voor de theorie om toepasbaar te zijn. Maar de connectie tussen projectie-operatoren en eventuele proposities is niet evident.

In hoofdstuk 8 beargumenteer ik dat, als we aannemen dat er een eenduidige manier is om projectie-operatoren met proposities te identificeren, we gedwongen zijn tot twee opties: ofwel de verzameling van proposities die geassocieerd worden met wiskundige objecten is niet gesloten onder disjuncties en conjuncties, of ze voldoen niet aan de klassieke propositielogica (als gevolg van de stelling van Kochen en Specker). In het bijzonder worden we bij de laatste optie gedwongen tot het accepteren van de orthodoxe quantumlogica $L(\mathcal{H})$ als propositielogica. Deze optie

beschouw ik als onacceptabel omdat ze strijdig is met het gebruik van de woorden "en" en "of" in de alledaagse taal. In het bijzonder, de distributieregels

$$(A \text{ en } B) \text{ of } C = (A \text{ of } C) \text{ en } (B \text{ of } C),$$
$$(A \text{ of } B) \text{ en } C = (A \text{ en } C) \text{ of } (B \text{ en } C)$$

zijn niet geldig in de orthodoxe quantumlogica.

De eerste optie wordt onderzocht in hoofdstuk 9. Gebruikmakende van eerder werk van Coecke uit 2002 wordt een uitbreiding $\mathcal{DI}(L(\mathcal{H}))$ van de orthodoxe quantumlogica $L(\mathcal{H})$ geconstrueerd waarin de distributieregels wel geldig zijn. Verder worden er operaties geïntroduceerd die aanleiding geven tot twee mogelijke lezingen van de uitbreiding. De eerste lezing geeft een klassieke propositielogica met een modale operator. Deze logica kan een fysische interpretatie gegeven worden door de elementen te identificeren met proposities over de mogelijke quantumtoestanden waarin een fysisch systeem zich kan bevinden. Als een domein voor quantumkansfuncties is dit echter geen bevredigende interpretatie: quantumtoestanden zijn namelijk zelf objecten die quantumkansfuncties specificeren. Hoewel het zinvol kan zijn om kansen toe te kennen aan de mogelijke toestanden waarin een systeem zich kan bevinden, geeft dit geen direct inzicht in de kansen voor meetuitkomsten van mogelijke experimenten. De tweede lezing geeft een (bijna) intuïtitionistische logica. Echter, de interpretatie van proposities in deze logica blijft onduidelijk op dit punt.

In hoofdstuk 10 wordt de strategie omgegooid. In plaats van te zoeken naar mogelijke uitbreidingen of andere herformuleringen van de orthodoxe quantumlogica construeer ik een nieuwe quantumlogica aan de hand van enkele empirische regels die opgelegd worden door de quantummechanica. De basis is een verzameling van zogenaamde elementaire experimentele proposities die geacht worden noodzakelijk te zijn voor het construeren van een domein voor quantumkansfuncties, namelijk, proposities uitgedrukt door zinnen van de vorm "$\mathcal{A}$ is gemeten en de meetuitkomst ligt in de verzameling Δ". Deze proposities worden vervolgens wiskundige representaties gegeven op basis van voorspellingen van de quantummechanica.[4]

Dezelfde quantummechanische voorspellingen worden gebuikt om de verzameling van wiskundige representaties een partiële ordening te geven die van de verzameling een tralie maakt. Dit betekent dat er wiskundige operaties geïntroduceerd worden (infimum en supremum) die in principe kandidaten zijn voor de connectieven "en" en "of". Een analyse van deze connectieven wijst uit dat het infimum een goede kandidaat is voor de conjunctie, maar dat het supremum als kandidaat voor disjunctie niet consistent is met de gegeven interpretatie. Wederom blijkt de boosdoener het falen van distributiviteit. Gebruikmakende van interpretationele

---

[4] Specifiek, de bovengenoemde zin wordt geïdentificeerd met een paar $(\mathcal{A}, P)$. Hier is $\mathcal{A}$ de Abelse von Neumann algebra is die gegenereerd wordt door de zelf-geadjungeerde operator A geassocieerd met de observabele $\mathcal{A}$ en P is de projectie-operator $\mu_A(\Delta)$, waarbij μ_A de spectraalmaat is behorende bij A.

232

eisen voor disjunctie wordt het tralie dan uitgebreid tot een Heyting algebra $I\!L(\mathcal{H})$ (waarin distributiviteit automatisch wel geldt). Ten slotte wordt het tralie nog uitgebreid tot een Booleaanse algebra $C\!L(\mathcal{H})$ door het introduceren van proposities uitgedrukt door zinnen van de vorm "$\mathcal{A}$ is niet gemeten". De belangrijkste motivatie voor deze laatste stap is een pragmatische: het is eenvoudiger (en minder controversieel) om kansfuncties op Booleaanse algebra's te introduceren dan op Heyting algebra's. Aan de andere kant, conceptueel lijkt er niet veel te veranderen aangezien $I\!L(\mathcal{H})$ een sub-algebra is van $C\!L(\mathcal{H})$. Kansfuncties gedefiniëerd op de laatste kunnen dus altijd gebruikt worden om een kansfunctie op de eerste te definiëren door het domein in te perken.

In hoofdstuk 11 onderzoek ik de mogelijkheid om kansfuncties te introduceren op de empiristische quantumlogica $C\!L(\mathcal{H})$. Gebruikmakende van de theorie van conditionele kansruimtes van Rényi en Popper laat ik zien dat iedere quantumkansfunctie gerepresenteerd kan worden door een conditionele kansfunctie op $C\!L(\mathcal{H})$. Hiermee is een herformulering van de quantumkanstheorie bewerkstelligd waarin het domein van de kansfuncties een heldere fysische interpretatie heeft: het zijn proposities over mogelijke metingen en meetuitkomsten. Hoewel dit op zichzelf een mooi resultaat is, kan men hopen op meer. De verzameling van alle conditionele kansfuncties op $C\!L(\mathcal{H})$ is groter dan de verzameling verkregen uit de quantumkansfuncties. Het zou interessant zijn om een karakterisering van deze deelverzameling te hebben die gebruikt kan worden om de voorkeursrol van quantumkansfuncties te begrijpen.

De situatie doet denken aan die van de klassieke representaties in deel II. Er is echter een belangrijk verschil. In deel II bleken aannames die nodig zijn om de voorkeursrol van quantumkansfuncties te begrijpen op gespannen voet te staan met de aannames die nodig waren om de klassieke representaties te krijgen. In het huidige geval zijn er geen sterke aannames die nodig waren voor de constructie van $C\!L(\mathcal{H})$. De situatie lijkt daarom juist omgedraaid te zijn: de toelaatbaarheid van niet-quantumkansfuncties lijkt een gevolg van het over het hoofd zien van fysisch-relevante aannames.

De moeilijkheid is om grip te krijgen op wat deze fysisch-relevante aannames zouden kunnen zijn. Een belangrijke rol voor het vinden van deze aannames is weggelegd voor het karakteriseren van de quantumkansfuncties. Ik laat zien dat de quantumkansfuncties precies corresponderen met de conditionele kansfuncties die aan een niet-contextualiteits-eis voldoen. Dit resultaat toont sterke gelijkenis met een eerder resultaat van Barnum et al uit 2000. Niet-contextualiteit als eigenschap van quantumkansfuncties is dus welbekend, en ik bespreek drie interpretaties van de quantummechanica die een poging doen om het belang van niet-contextualiteit te duiden, namelijk, Deutsch en Wallace's versie van de vele werelden interpretatie, de informatietheoretische interpretatie van Bub en Pitowsky en de quantum Bayesiaanse interpretatie van Caves, Fuchs en Schack. Mijn conclusie is dat geen van de drie een bevredigende fundering voor een eis van niet-contextualiteit leveren.

In het laatste deel van hoofdstuk 11 ga ik op zoek naar een andere karakterisering van de quantumkansfuncties. Deze wordt verkregen door een link te leggen tussen $\mathcal{CL}(\mathcal{H})$ en de orthodoxe quantumlogica $L(\mathcal{H})$. Deze link loopt via de bijna intuïtionistische quantumlogica $\mathcal{DI}(L(\mathcal{H}))$ die werd gevonden in hoofdstuk 9. De elementen van $\mathcal{DI}(L(\mathcal{H}))$ worden nu geïnterpreteerd als disjuncties en conjuncties van (equivalentieklassen van) proposities van de vorm "als $\mathcal{A}$ gemeten wordt, dan ligt de meetuitkomst in Δ", waar het antecedent en het consequent opgevat kunnen worden als elementaire experimentele proposities uit $\mathcal{CL}(\mathcal{H})$. De conditionele zinnen zelf komen precies overeen met de projectie-operatoren uit $L(\mathcal{H})$, die nu worden opgevat als elementen van $\mathcal{DI}(L(\mathcal{H}))$. De quantumkansfuncties zijn dan precies de conditionele kansfuncties op $\mathcal{CL}(\mathcal{H})$ die op niet-ambigue wijze gebruikt kunnen worden om kansen aan de conditionele zinnen toe te kennen. Wiskundig geeft dit een degelijke karakterisering. Hoe goed deze vanuit een fysisch/filosofisch perspectief te motiveren is hangt af van de degelijkheid van de gepostuleerde interpretatie van de elementen van $\mathcal{DI}(L(\mathcal{H}))$. Dit is een punt voor toekomstig onderzoek.

Curriculum vitae

The author was born in Nijmegen, the Netherlands on the 4th of November in 1981. After finishing high school (Kandinsky college, Nijmegen, the Netherlands) in 2000 he found himself more interested in other things than education (in particular: music) and took one year of before deciding to study physics at the Radboud University in Nijmegen. In the meantime he financed his hobbies by working at Sealed Air producing packaging materials. During a seven year struggle for obtaining his Bachelor's degree (obtained in 2008) he decided his way of thinking was better fit for doing mathematics. This decision turned out to be a good one as in 2009 he obtained his Master's degree in mathematics cum laude (also at the Radboud University) with a thesis on the foundations of quantum mechanics.

Afterwards he started working again at Sealed Air in the expectation to soon find a PhD position. This search turned out to almost take two years, and in 2011 he was accepted for a position at the University of Groningen at the faculty of philosophy. The dissertation you are currently reading is a partial reflection of the work done there in the past four years and the list of publications below is another partial reflection. He is currently a postdoctoral researcher at the Department of Materials at the University of Oxford where he can combine his interests in physics, mathematics and philosophy.

List of publications:

- **-)** Hermens, R. (2014). Conway-Kochen and the finite precision loophole, *Foundations of Physics* 44(10):1038-1048.
- **-)** Hermens, R. (2014). Placing probabilities of conditionals in context, *The Review of Symbolic Logic* 7(3):415-438.
- **-)** Hermens, R. (2014). Speakable in quantum mechanics: Babbling on, *Electronic Proceedings in Theoretical Computer Science* 158:53-64. Eds. R. Duncan and P. Panangaden. Proceedings of the 9th Workshop on Quantum Physics and Logic 2012.

-) Hermens, R. (2014). The measurement problem is your problem too. Published in *New Directions in the Philosophy of Science*, Eds. M.C. Galavotti, D. Dieks, W.J. Gonzalez, S. Hartmann, Th. Uebel, M. Weber. Springer.

-) Hermens, R. (2013). Speakable in quantum mechanics, *Synthese* 190(15):3265-3286.

-) Hermens, R. (2013). Weakly intuitionistic quantum logic, *Studia Logica* 101(5): 901-913.

-) Hermens, R. (2011). The problem of contextuality and the impossibility of experimental metaphysics thereof, *Studies in History and Philosophy of Modern Physics* 42(4):214-225.

Index

References

Adams, E. W. (1975). *The logic of conditionals: An application of probability to deductive logic.* 86. Dordrecht: D. Reidel.

Aerts, D. (1982). "Description of many separated physical entities without the paradoxes encountered in quantum mechanics". In: *Foundations of Physics* 12.12, pp. 1131–1170.

Aerts, D., J. Broekaert, M. Czachor, M. Kuna, B. Sinervo, and S. Sozzo (2014). "Quantum structure in competing lizard communities". In: *Ecological Modelling* 281, pp. 38–51.

Appleby, D. M. (2000). "Contextuality of approximate measurements". In: arXiv:quant-ph/0005010v2.

— (2001). "Nullification of the nullification". In: arXiv:quant-ph/0109034.

— (2002). "Existential contextuality and the models of Meyer, Kent and Clifton". In: *Physical Review A* 65, p. 022105.

— (2005). "The Bell-Kochen-Specker theorem". In: *Studies in History and Philosophy of Modern Physics* 36, pp. 1–28.

Araki, H. and M. M. Yanase (1960). "Measurement of quantum mechanical operators". In: *Physical Review* 120.2, pp. 622–626.

Arntzenius, F. (2003). "Some problems for conditionalization and reflection". In: *The Journal of Philosophy* 100.7, pp. 356–370.

Atten, M. van (2011). "Weak counterexamples". In: *The Stanford Encyclopedia of Philosophy.* Ed. by E. N. Zalta. URL: http://plato.stanford.edu/entries/brouwer/weakcounterex.html.

Bacciagaluppi, G. (2009). "Is logic empirical?" In: *Handbook of Quantum Logic and Quantum Structures.* Ed. by K. Engesser, D. M. Gabbay, and D. Lehmann. Amsterdam: Elsevier, pp. 49–78.

— (2014). "A critic looks at QBism". In: *New Directions in the Philosophy of Science.* Heidelberg: Springer, pp. 403–416.

241

Barnum, H., C. M. Caves, J. Finkelstein, C. A. Fuchs, and R. Schack (2000). "Quantum probability from decision theory?" In: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 456.1997, pp. 1175–1182.

Barrett, J. and A. Kent (2004). "Non-contextuality, finite precision measurement and the Kochen-Specker theorem". In: *Studies in History and Philosophy of Modern Physics* 35, pp. 151–176.

Barrett, J., E. G. Cavalcanti, R. Lal, and O. J. E. Maroney (2014). "No ψ-epistemic model can fully explain the indistinguishability of quantum states". In: *Physical Review Letters* 112, p. 250403.

Beisbart, C. and S. Hartmann, eds. (2011). *Probabilities in Physics*. Oxford: Oxford University Press.

Bell, J. S. (1966). "On the problem of hidden variables in quantum mechanics". In: *Reviews of Modern Physics* 38.3, p. 447. Reprinted in (Bell, 1987).

— (1981). "Quantum mechanics for cosmologists". In: *Quantum Gravity 2*. Ed. by C. Isham, R. Penrose, and D. Sciama. Oxford: Clarendon Press, pp. 611–637. Page numbers refer to the reprinted version in (Bell, 1987).

— (1987). *Speakable and Unspeakable in Quantum Mechanics*. Cambridge: Cambridge University Press.

— (1988). In: *Omni*. Interview, pp. 85–92,121.

— (1990). "La nouvelle cuisine". In: *Between Science and Technology*. Ed. by A. Sarlemijn and P. Kroes. Amsterdam: Elsevier, pp. 97–115.

Benacerraf, P. (1973). "Mathematical truth". In: *The Journal of Philosophy* 70.19, pp. 661–679.

Bengtsson, I. (2012). "Gleason, Kochen-Specker, and a competition that never was". In: *AIP Conference Proceedings* 1508.1, pp. 125–135.

Birkhoff, G. and J. von Neumann (1936). "The logic of quantum mechanics". In: *Annals of Mathematics* 37.4, pp. 823–843.

Bishop, E. (1967). *Foundations of Constructive Analysis*. New York: McGraw-Hill.

Blanchfield, K. (2012). "Configurations in Quantum Information". Licentiate Thesis. Stockholm University.

Bohm, D. (1952). "A suggested interpretation of the quantum theory in terms of "hidden" variables I & II". In: *Physical Review* 85.2, pp. 166–193.

Bohr, N. (1913). "On the constitution of atoms and molecules, part I". In: *Philosophical Magazine* 26.151, pp. 1–25.

— (1948). "On the notions of causality and complementarity". In: *Dialectica* 2.3-4, pp. 312–319.

— (1949). "Discussion with Einstein on epistemological problems in atomic physics". In: *Albert Einstein: Philosopher-Scientist*. Ed. by P. A. Schilpp. Evanston: Library of Living Philosophers, pp. 199–242.

— (1958). "Quantum physics and philosophy – causality and complementarity". In: *Philosoph in the Mid-Century*. Ed. by R. Klibansky. reprinted as (Bohr, 1996, pp. 385–394). Firenze: La nuova Italia editrice, pp. 308–314.

— (1963). "The genesis of quantum mechanics". In: *Essays 1958-1962 on atomic physics and human knowledge*. New York: Wiley, pp. 74–78.

— (1985). *Niels Bohr Collected Works 6*. Ed. by J. Kalckar. Amsterdam: North-Holland Physics Publishing.

— (1996). *Niels Bohr Collected Works 7*. Ed. by J. Kalckar. Amsterdam: Elsevier.

Born, M. (1926a). "Zur Quantenmechanik der Stoßvorgänge". In: *Zeitschrift für Physik* 37.12, pp. 863–867. English translation by Wheeler and Zurek (1983, pp. 52–55).

— (1926b). "Quantenmechanik der Stoßvorgänge". In: *Zeitschrift für Physik* 38.11–12, pp. 803–827.

— (1953). "Physical reality". In: *The Philosophical Quarterly* 3.11, pp. 139–149.

— (1955). "Statistical interpretation of quantum mechanics". In: *Science* 122.3172, pp. 675–679.

Broglie, L. de (1924). "Recherches sur la théorie des quanta". PhD thesis.

— (1960). *Non-Linear Wave Mechanics*. Trans. by A. J. Knodel and J. C. Miller. Amsterdam: Elsevier.

Brouwer, L. E. J. (1908). "De onbetrouwbaarheid der logische principes". In: *Tijdschrift voor Wijsbegeerte* 2, pp. 152–158.

— (1929). "Mathematik, Wissenschaft und Sprache". In: *Monatshefte für Mathematik* 36.1, pp. 153–164.

Brown, H. R., W. Myrvold, and J. Uffink (2009). "Boltzmann's H-theorem, its discontents, and the birth of statistical mechanics". In: *Studies in History and Philosophy of Modern Physics* 40.2, pp. 174–191.

Brunner, N., D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner (2014). "Bell nonlocality". In: *Reviews of Modern Physics* 86.2, pp. 419–478.

Bruns, G. and H. Lakser (1970). "Injective hulls of semilattices". In: *Canadian Mathematical Bulletin* 13.1, pp. 115–118.

Bub, J. (1975). "Popper's propensity interpretation of probability and quantum mechanics". In: *Induction, Probability, and Confirmation*. Ed. by G. Maxwell and R. M. Anderson. Minneapolis: University of Minnesota Press, pp. 416–429.

— (1996). "Schütte's tautology and the Kochen-Specker theorem". In: *Foundations of Physics* 26.6, pp. 787–806.

— (2007). "Quantum probabilities as degrees of belief". In: *Studies in History and Philosophy of Modern Physics* 38, pp. 232–254.

— (2011a). "Is von Neumann's "no hidden variables" proof silly?" In: *Deep Beauty*. Ed. by H. Halvorson. Cambridge: Cambridge University Press, pp. 393–407.

Bub, J. (2011b). "Quantum probabilities: an information-theoretic interpretation". In: *Probabilities in Physics*. Ed. by C. Beisbart and S. Hartmann. Oxford: Oxford University Press, pp. 231–262.

Bub, J. and R. Clifton (1996). "A uniqueness theorem for 'no collapse' interpretations of quantum mechanics". In: *Studies in History and Philosophy of Modern Physics* 27.2, pp. 181–219.

Bub, J. and I. Pitowsky (2010). "Two dogmas about quantum mechanics". In: *Many Worlds?: Everett, Quantum Theory, and Reality*. Ed. by S. Saunders, J. Barrett, A. Kent, and D. Wallace. Oxford: Oxford University Press, pp. 433–459.

Burris, S. and H. P. Sankappanavar (2012). *A Course in Universal Algebra*. URL: `http : / / www . math . uwaterloo . ca / ~snburris / htdocs / UALG / univ - algebra2012.pdf`. Free online updated edition, original published by Springer-Verlag, New-York, 1981.

Busch, P. (2003). "Quantum states and generalized observables: a simple proof of Gleasons theorem". In: *Physical Review Letters* 91.12, p. 120403.

Busemeyer, J. R. and P. D. Bruza (2012). *Quantum models of cognition and decision*. Cambridge: Cambridge University Press.

Cabello, A. (1999). "Comment on "non-contextual hidden variables and physical measurements"". In: arXiv:quant-ph/9911024.

— (2002). "Finite-precision measurement does not nullify the Kochen-Specker theorem". In: *Physical Review A* 65, p. 052101.

Callender, C. (2007). "The emergence and interpretation of probability in bohmian mechanics". In: *Studies in History and Philosophy of Modern Physics* 38.2, pp. 351–370.

Caspers, M., C. Heunen, N. P. Landsman, and B. Spitters (2009). "Intuitionistic quantum logic of an n-level system". In: *Foundations of Physics* 39, pp. 731–759.

Caves, C. M., C. A. Fuchs, and R. Schack (2002). "Quantum probabilities as Bayesian probabilities". In: *Physical Review A* 65, p. 022305.

Celani, S. and R. Jansana (2005). "Bounded distributive lattices with strict implication". In: *Mathematical Logic Quarterly* 51.3, pp. 219–246.

Cirel'son, B. S. (1980). "Quantum generalizations of Bell's inequality". In: *Letters in Mathematical Physics* 4.2, pp. 93–100.

Clifton, R. and A. Kent (2001). "Simulating quantum mechanics by non-contextual hidden variables". In: *Proceedings: Mathematical, Physical and Engineering Sciences* 456.2001, pp. 2101–2114.

Coecke, B. (2002). "Quantum logic in intuitionistic perspective". In: *Studia Logica* 70.3, pp. 411–440.

Colbeck, R. and R. Renner (2011). "No extension of quantum theory can have improved predictive power". In: *Nature communications* 2, p. 411.

— (2012b). "The completeness of quantum theory for predicting measurement outcomes". In: arXiv1208.4123v2 [quant-ph].

References

— (2012a). "Is a system's wave function in one-to-one correspondence with its elements of reality?" In: *Physical Review Letters* 108 (15), p. 150402.

Conway, J. B. (1990). *A Course in Functional Analysis*. 2nd ed. New York: Springer-Verlag.

Cuculescu, I. and A. G. Oprea (1994). *Noncommutative probability*. Dordrecht: Springer.

Demey, L., B. Kooi, and J. Sack (2014). "Logic and probability". In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. URL: http://plato.stanford.edu/archives/fall2014/entries/logic-probability/.

Deutsch, D. (1999). "Quantum theory of probability and decisions". In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. Vol. 455. 1988. The Royal Society, pp. 3129–3137.

Diaconis, P., S. Holmes, and R. Montgomery (2007). "Dynamical bias in the coin toss". In: *SIAM review* 49.2, pp. 211–235.

Dieks, D. and P. E. Vermaas (1998). *The modal interpretation of quantum mechanics*. Dordrecht: Kluwer Academic Publishers.

Dirac, P. A. M. (1925). "The fundamental equations of quantum mechanics". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 109, pp. 642–653.

— (1930). *The principles of quantum mechanics*. Oxford: Clarendon Press.

Douglas, R. G. (1998). *Banach Algebra Techniques in Operator Theory*. Second edition. New York: Springer-Verlag.

Dummett, M. (1976). "Is logic empirical?" In: *Contemporary British Philosophy* 4. Reprinted in (Dummett, 1978) pp. 269–289, pp. 45–68.

— (1978). *Truth and Other Enigmas*. Harvard: Harvard University Press.

Dürr, D., S. Goldstein, and N. Zanghì (2013). *Quantum physics without quantum philosophy*. Springer.

Edgington, D. (2014). "Indicative conditionals". In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Winter 2014.

Eells, E. (1983). "Objective probability theory theory". In: *Synthese* 57, pp. 387–442.

Eilers, M. and E. Horst (1975). "The theorem of Gleason for nonseparable Hilbert spaces". In: *International Journal of Theoretical Physics* 13 (6), pp. 419–424.

Einstein, A. (1905). "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt". In: *Annalen der Physik* 322.6, pp. 132–148.

Einstein, A., M. Born, and H. Born (1969). *Briefwechsel 1916–1955*. Germany: Rowohlt.

— (1971). *The Born-Einstein Letters*. Trans. by I. Born. London and Basingstoke: MacMillan.

Esfeld, M., M. Hubert, D. Lazarovici, and D. Dürr (2013). "The ontology of Bohmian mechanics". In: *The British Journal for the Philosophy of Science*, axt019.

Everett III, H. (1957). ""Relative state" formulation of quantum mechanics". In: *Reviews of modern physics* 29.3, pp. 454–462.

Faye, J. (2014). "Copenhagen interpretation of quantum mechanics". In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Fall 2014. URL: http://plato.stanford.edu/archives/fall2014/entries/qm-copenhagen/.

Feynman, R. P., R. B. Leighton, and M. Sands (1963). *The Feynman Lectures on Physics*. Reading: Addison-Wesley Publishing Company.

Fine, A. (1982b). "Joint distributions, quantum correlations, and commuting observables". In: *Journal of Mathematical Physics* 23, pp. 1306–1310.

— (1982a). "Hidden variables, joint probability, and the Bell inequalities". In: *Physical Review Letters* 48 (5), pp. 291–295.

Folse, H. J. (1985). *The Philosophy of Niels Bohr*. New York: Elsevier Science Pub Co.

Fraassen, B. van (1972). *A formal approach to the philosophy of science*. Pittsburgh: University of Pittsburgh Press.

— (1981). "Probabilistic semantics objectified: I. postulates and logics". In: *Journal of Philosophical Logic* 10.3, pp. 371–394.

Friedman, M. and C. Glymour (1972). "If quanta had logic". In: *Journal of Philosophical Logic* 1.1, pp. 16–28.

Fuchs, C. A. (2002a). "Quantum mechanics as quantum information (and only a little more)". In: arXiv:quant-ph/0205039.

— (2002b). "The anti-Växjö interpretation of quantum mechanics". In: arXiv:quant-ph/0204146.

— (2010). "QBism, the perimeter of quantum Bayesianism". In: arXiv:1003.5209 [quant-ph].

— (2014). "My struggles with the block universe". In: arXiv:1405.2390 [quant-ph].

Fuchs, C. A., N. D. Mermin, and R. Schack (2013). "An introduction to QBism with an application to the locality of quantum mechanics". In: arXiv:1311.5253 [quant-ph].

Fuchs, C. A. and R. Schack (2013). "Quantum-Bayesian coherence: the no-nonsense version". In: arXiv:1301.3274 [quant-ph].

Gentzen, G. (1936). "Die Widerspruchsfreiheit der reinen Zahlentheorie". In: *Mathematische Annalen* 112.1, pp. 493–565.

Ghirardi, G. C., A. Rimini, and T. Weber (1985). "A model for a unified quantum description of macroscopic and microscopic systems". In: *Quantum Probability and Applications II*. Springer, pp. 223–232.

— (1986). "Unified dynamics for microscopic and macroscopic systems". In: *Physical Review D* 34.2, pp. 470–491.

References

Gillies, D. (1990). "Bayesianism versus falsifigationism". In: *Ratio* 3, pp. 82–98.

Ginnobili, S. (2008). *Ptolemy and Homer (Simpson)*. `https://www.youtube.com/watch?v=QVuU2YCwHjw`. Accessed: 2015-04-05.

Gleason, A. (1957). "Measures on the closed subspaces of a Hilbert space". In: *Indiana University Mathematics Journal* 6 (4), pp. 885–893. Reprinted in (Hooker, 1975) pp. 123–134.

Gödel, K. (1933). "Zur intuitionistischen Arithmetik und Zahlentheorie". In: *Ergebnisse eines mathematischen Kolloquiums* 4.1933, pp. 34–38.

Griffiths, R. B. (2003). *Consistent quantum theory*. Cambridge: Cambridge University Press.

— (2013). "A consistent quantum ontology". In: *Studies in History and Philosophy of Modern Physics* 44.2, pp. 93–114.

— (2014). "The new quantum logic". In: *Foundations of Physics* 44.6, pp. 610–640.

Gudder, S. P. (1970). "On hidden variable theories". In: *Journal of Mathematical Physics* 11, pp. 431–436.

— (1979). *Stochastic Methods in Quantum Mechanics*. New York: Elsevier North Holland.

Haar, D. ter (1967). *The old quantum theory*. Oxford: Pergamon.

Hájek, A. (1996). ""Mises redux"–redux: fifteen arguments against finite frequentism". In: *Erkenntnis* 45.2-3, pp. 209–227.

— (2003). "What conditional probability could not be". In: *Synthese* 137.3, pp. 273–323.

— (2009). "Fifteen arguments against hypothetical frequentism". In: *Erkenntnis* 70.2, pp. 211–235.

— (2012a). "Interpretations of probability". In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Winter 2012. URL: `http://plato.stanford.edu/archives/win2012/entries/probability-interpret/`.

— (2012b). "The fall of "Adams' thesis"?" In: *Journal of Logic, Language and Information* 21.2, pp. 145–161.

— (2015). "On the plurality of Lewis's triviality results". In: *A Companion to David Lewis*. Ed. by B. Loewer and J. Schaffer. Chichester: John Wiley & Sons, pp. 425–445.

Hall, N. (1994). "Correcting the guide to objective chance". In: *Mind* 103.412, pp. 505–517.

Halvorson, H. (2001). "On the nature of continuous physical quantities in classical and quantum mechanics". In: *Journal of Philosophical Logic* 30.1, pp. 27–50. ISSN: 0022-3611.

Hardy, L. (2001). "Quantum mechanics from five reasonable axioms". In: arXiv:quant-ph/0101012.

— (2013). "Are quantum states real?" In: *International Journal of Modern Physics B* 27.01n03.

Heisenberg, W. (1925). "Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen." In: *Zeitschrift für Physik A Hadrons and Nuclei* 33.1, pp. 879–893.

— (1927). "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik". In: *Zeitschrift für Physik* 43.3–4, pp. 172–198. English translation by Wheeler and Zurek (1983, pp. 62–84).

— (1967). "Quantum theory and its interpretation". In: *Niels Bohr*. Ed. by S. Rozental. New York: Interscience Publishers.

— (1971). *Physics and Beyond*. Trans. by A. J. Pomerans. New York: Harper.

Hermens, R. (2011). "The problem of contextuality and the impossibility of experimental metaphysics thereof". In: *Studies in History and Philosophy of Modern Physics* 42.4, pp. 214–225.

— (2013b). "Weakly intuitionistic quantum logic". In: *Studia Logica* 101.5, pp. 901–913.

— (2013a). "Speakable in quantum mechanics". In: *Synthese* 190 (15), pp. 3265–3286.

— (2014). "Speakable in quantum mechanics: babbling on". In: *Electronic Proceedings in Theoretical Computer Science* 158, pp. 53–64. Proceedings of the QPL2012 conference.

Heunen, C., N. P. Landsman, and B. Spitters (2012). "Bohrification of operator algebras and quantum logic". In: *Synthese* 186.3, pp. 719–752.

Hilgevoord, J. and J. Uffink (2014). "The uncertainty principle". In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Spring 2014. URL: http://plato.stanford.edu/archives/spr2014/entries/qt-uncertainty/.

Holevo, A. S. (1982). *Probabilistic and Statistical Aspects of Quantum Theory*. Amsterdam: North-Holland. originally published in Russian (1980).

Hooker, C. A., ed. (1975). *The Logico-Algebraic Approach to Quantum Mechanics Volume I*. Dordrecht: D. Reidel.

— ed. (1979). *The Logico-Algebraic Approach to Quantum Mechanics Volume II*. Dordrecht: D. Reidel.

Howard, D. (2004). "Who invented the "Copenhagen interpretation"? A study in mythology". In: *Philosophy of Science* 71, pp. 669–682.

Hughes, R. I. G. (1982). "The logic of experimental questions". In: *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*. Chicago: The University of Chicago Press, pp. 243–256.

Iemhoff, R. (2013). "Intuitionism in the philosophy of mathematics". In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Spring 2015. URL: http://plato.stanford.edu/archives/spr2015/entries/intuitionism/.

Isham, C. J. (1995). *Lectures on Quantum Theory*. London: Imperial College Press.

Jammer, M. (1966). *Conceptual Developments of Quantum Mechanics*. New York: McGraw-Hill.

References

Jarrett, J. P. (1984). "On the physical significance of the locality conditions in the Bell arguments". In: *Noûs*, pp. 569–589.

Jauch, J. M. and C. Piron (1969). "On the structure of quantal proposition systems". In: *Helvetica Physica Acta* 42, pp. 842–848.

Jaynes, E. T. (1985). "Some random observations". In: *Synthese* 63, pp. 115–138.

Kadison, R. V. and J. R. Ringrose (1986). *Fundamentals of the Theory of Operator Algebras*. Orlando: Academic Press Inc.

Karnaugh, M. (1953). "The map method for synthesis of combinational logic circuits". In: *Transactions of the American Institute of Electrical Engineers part I* 72.9, pp. 593–599.

Keller, J. B. (1986). "The probability of heads". In: *The American Mathematical Monthly*, pp. 191–197.

Kennard, E. H. (1927). "Zur quantenmechanik einfacher Bewegungstypen". In: *Zeitschrift für Physik* 44.4-5, pp. 326–352.

Kent, A. (1999). "Noncontextual hidden variables and physical measurements". In: *Physical Review Letters* 83.19, pp. 3755–3757.

— (2010). "One world versus many". In: *Many Worlds?: Everett, Quantum Theory, and Reality*. Ed. by S. Saunders, J. Barrett, A. Kent, and D. Wallace. Oxford: Oxford University Press, pp. 307–354.

— (2014). "Does it make sense to speak of self-locating uncertainty in the universal wave function? remarks on Sebens and Carroll". In: *Foundations of Physics*, pp. 1–7.

Kment, B. (2012). "Haecceitism, chance, and counterfactuals". In: *Philosophical Review* 121.4, pp. 573–609.

Kochen, S. and E. P. Specker (1967). "The problem of hidden variables in quantum mechanics". In: *Journal of Mathematics and Mechanics* 17.1, pp. 59–67. Reprinted in (Hooker, 1975) pp. 293–328.

Kolmogorov, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin: Springer. English translation by N. Morrison, *Foundations of the Theory of Probability*, Chelsea Publishing Company, 1956.

— (1963). "The theory of probability". In: *Mathematics: Its Content, Methods, and Meaning*. Ed. by A. D. Aleksandrov, A. N. Kolmogorov, and M. A. Lavrent'ev. Vol. 2. Cambridge: M.I.T. Press, pp. 229–264.

Kramer, S. (2014). "Quantum logic as classical logic". In: arXiv:1406.3526v3 [quant-ph].

Kümmerer, B. and H. Maassen (1998). "Elements of quantum probability". In: *Quantum Probability Communications* 10, pp. 73–100.

La Cour, B. R. (2009). "Quantum contextuality in the Mermin-Peres square: a hidden-variable perspective". In: *Physical Review A* 79.1, p. 012102.

Lambalgen, M. van (1996). "Randomness and foundations of probability: von Mises' axiomatisation of random sequences". In: *Lecture Notes-Monograph Series*, pp. 347–367.

Landsman, N. P. (2006). "When champions meet: rethinking the Bohr-Einstein debate". In: *Studies in History and Philosophy of Modern Physics* 37.1, pp. 212–242.

— (2007). "Between classical and quantum". In: *Philosophy of Physics, Part B*. Ed. by J. Earman and J. Butterfield. Amsterdam: Elsevier, pp. 415–555.

— (2009). "Born rule and its interpretation". In: *Compendium of Quantum Physics*. Ed. by D. Greenberger, K. Hentschel, and F. Weinert. Berlin: Springer-Verlag, pp. 64–70.

Landsman, N. P. and R. Reuvers (2013). "A flea on Schrödingers cat". In: *Foundations of Physics* 43.3, pp. 373–407.

Lanyon, B. P., T. J. Weinhold, N. K. Langford, J. L. OBrien, K. J. Resch, A. Gilchrist, and A. G. White (2008). "Manipulating biphotonic qutrits". In: *Physical Review Letters* 100.6, p. 060504.

Leifer, M. (2014). "Is the quantum state real? an extended review of ψ-ontology theorems". In: *Quanta* 3.1, pp. 67–155.

Lemmon, E. J. (1966). "Algebraic semantics for modal logics I". In: *The Journal of Symbolic Logic* 31.1, pp. 46–65.

Lewis, D. (1981). "A subjectivists guide to objective chance". English. In: *IFS*. Ed. by W. L. Harper, R. Stalnaker, and G. Pearce. Vol. 15. The University of Western Ontario Series in Philosophy of Science. Dordrecht: D. Reidel, pp. 267–297.

Lisoněk, P., P. Badziąg, J. R. Portillo, and A. Cabello (2014). "Kochen-Specker set with seven contexts". In: *Physical Review A* 89 (4), p. 042101.

Loewer, B. (2004). "David Lewis's Humean theory of objective chance". In: *Philosophy of Science* 71.5, pp. 1115–1125.

Lyon, A. (2010). "Philosophy of probability". In: *Philosophies of the Sciences: A Guide*. Ed. by F. Allhoff. Wiley-Blackwell, pp. 92–125.

— (2011). "Deterministic probability: neither chance nor credence". In: *Synthese* 182, pp. 413–432.

Maassen, H. (2003). "Quantum probability applied to the damped harmonic oscillator". In: *Quantum Probability Communications*. Ed. by S. Attal and J. Lindsay. Vol. XII. World Scientific, pp. 23–58. Preprint available on arXiv:0411024.

— (2010). "Quantum probability and quantum information theory". In: *Quantum Information, Computation and Cryptography*. Ed. by F. Benatti, M. Fannes, R. Floreanini, and D. Petritis. Vol. 808. Lecture Notes in Physics. Berlin Heidelberg: Springer, pp. 65–108.

Malament, D. B. and S. L. Zabell (1980). "Why Gibbs phase averages work–the role of ergodic theory". In: *Philosophy of Science*, pp. 339–349.

References

Mandelkern, M. (1989). "Brouwerian counterexamples". In: *Mathematics Magazine* 62.1, pp. 3–27.

Marsden, D. (2010). "Comparing intuitionistic quantum logics". MA thesis. University of Oxford. URL: `http://www.cs.ox.ac.uk/people/bob.coecke/Marsden.pdf`.

Maudlin, T. (1995). "Three measurement problems". In: *Topoi* 14.1, pp. 7–15.

— (2005). "The tale of quantum logic". In: *Hilary Putnam.* Ed. by Y. Ben-Menahem. Cambridge: Cambridge University Press, pp. 156–187.

— (2014). "Critical study – David Wallace, the emergent multiverse: quantum theory according to the Everett interpretation". In: *Noûs* 48.4, pp. 794–808.

Mehra, J. and H. Rechenberg (1982). *The Historical Development of Quantum Theory.* New York: Springer-Verlag.

Mermin, N. D. (1993). "Hidden variables and the two theorerns of John Bell". In: *Reviews of Modern Physics* 65.3, pp. 803–815.

— (1999). "A Kochen-Specker theorem for imprecisely specified measurement". In: arXiv:quant-ph/9912081.

Meyer, D. A. (1999). "Finitie precision measurement nullifies the Kochen-Specker theorem". In: *Physical Review Letters* 83.19, pp. 3751–3754.

Meyer, P. A. (1993). *Quantum Probability for Probabilists.* Berlin Heidelberg: Springer-Verlag.

Mises, R. von (1931). *Wahrscheinlichkeitsrechnung und ihre Anwendungen in der Statistik und theoretischen Physik.* Leipzig und Wien: Deuticke.

Moore, D. J. (1999). "On state spaces and property lattices". In: *Studies in History and Philosophy of Modern Physics* 30.1, pp. 61–83.

Morgan, C. G. and H. Leblanc (1983a). "Probabilistic semantics for intuitionistic logic". In: *Notre Dame Journal of Formal Logic* 24.2, pp. 161–180.

— (1983b). "Probability theory, intuitionism, semantics, and the dutch book argument". In: *Notre Dame Journal of Formal Logic* 24.3, pp. 289–304.

Muller, F. A. (1997a). "The equivalence myth of quantum mechanics–part I". In: *Studies in History and Philosophy of Modern Physics* 28.1, pp. 35–61.

— (1997b). "The equivalence myth of quantum mechanics–part II". In: *Studies in History and Philosophy of Modern Physics* 28.2, pp. 219–247.

— (2014). "Circumveiloped by obscuritads. The nature of interpretation in quantum mechanics, hermeneutic circles and physical reality, with cameos of James Joyce and Jacques Derrida". In: arXiv:1406.6284 [quant-ph].

Myrvold, W. C. (2012). "Deterministic laws and epistemic chances". In: *Probability in Physics.* Ed. by Y. Ben-Menahem and M. Hemmo. The Frontiers Collection. Berlin Heidelberg: Springer, pp. 73–85.

Naimark, M. A. (1943). "On a representation of additive operator set functions". In: *Bulletin de l'Acad'emie des Sciences de l'URSS* 7, pp. 237–244.

Neumann, J. von (1932). *Mathematische Grundlagen der Quantenmechanik.* Berlin: Verlag Julius-Springer. English translation by R. T. Beyer, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, 1955.

Nielsen, M. A. (1997). "Computable functions, quantum measurements, and quantum dynamics". In: *Physical Review Letters* 79.15, p. 2915.

Nielsen, M. A. and I. L. Chuang (2010). *Quantum computation and quantum information.* Cambridge university press.

Nigg, D., T. Monz, P. Schindler, E. A. Martinez, M. Chwalla, M. Hennrich, R. Blatt, M. F. Pusey, T. Rudolph, and J. Barrett (2012). "Can different quantum state vectors correspond to the same physical state? an experimental test". In: arXiv:1211.0942 [quant-ph].

Noordhof, P. (1999). "Probabilistic causation, preemption and counterfactuals". In: *Mind* 108.429, pp. 95–125.

Omnès, R. (1999). *Understanding quantum mechanics.* Princeton: Princeton University Press.

Parthasarathy, K. R. (1992). *An introduction to quantum stochastic calculus.* Basel: Birkhäuser Verlag.

Pavičić, M., J. P. Merlet, B. McKay, and N. D. Megill (2005). "Kochen-Specker vectors". In: *Journal of Physics A: Mathematical and General* 38.7, pp. 1577–1592.

Pearle, P. (1989). "Combining stochastic dynamical state-vector reduction with spontaneous localization". In: *Physical Review A* 39.5, p. 2277.

Pedersen, G. K. (1979). *C*-Algebras and their Automorphism Groups.* London: Academic Press Inc.

Peres, A. (1978). "Unperformed experiments have no results". In: *American Journal of Physics* 46.7, pp. 745–747.

— (2002). *Quantum Theory: Concepts and Methods.* New York: Kluwer Academic.

— (2003). "How the no-cloning theorem got its name". In: *Fortschritte der Physik* 51.4-5, pp. 458–461.

Petersen, A. (1963). "The philosophy of Niels Bohr". In: *Bulletin of the Atomic Scientists* 19.7, pp. 8–14.

Piron, C. (1976). *Foundations of Quantum Physics.* Massachusetts: W. A. Benjamin Inc.

Pitowsky, I. (1983). "Deterministic model of spin and statistics". In: *Physical Review D* 27.10, p. 2316.

— (1985). "Quantum mechanics and value definiteness". In: *Philosophy of Science*, pp. 154–156.

— (2003). "Betting on the outcomes of measurements: a Bayesian theory of quantum probability". In: *Studies in History and Philosophy of Modern Physics* 34.3, pp. 395–414.

References

— (2006). "Quantum mechanics as a theory of probability". In: *Physical Theory and its Interpretation*. Ed. by W. Demopoulos and I. Pitowsky. Vol. 72. The Western Ontario Series in Philosophy of Science. Dordrecht: Springer Netherlands, pp. 213–240.

Planck, M. K. E. L. (1900). "Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum". In: *Verhandlungen der Deutschen Physikalischen Gesellschaft* 2, pp. 237–245.

Pollock, J. L. (1984). "A solution to the problem of induction". In: *Noûs* 18.3, pp. 423–461.

Popescu, S. and D. Rohrlich (1994). "Quantum nonlocality as an axiom". In: *Foundations of Physics* 24.3, pp. 379–385.

Popper, K. (1938). "A set of independent axioms for probability". English. In: *Mind* 47.186, pp. 275–277.

— (1957). "The propensity interpretation of the calculus of probability, and the quantum theory". In: *Observation and Interpretation*. Ed. by S. Körner. London: Buttersworth Scientific Publications, pp. 65–70.

— (1959). *The Logic of Scientific Discovery*. Page numbers in this dissertation refer to the 2005 electronic edition published by Taulor & Francis. Hutchinson & Co.

— (1968). "Birkhoff and von Neumann's interpretation of quantum mechanics". In: *Nature* 219, pp. 682–685.

Pothos, E. M. and J. R. Busemeyer (2009). "A quantum probability explanation for violations of 'rational' decision theory". In: *Proceedings of the Royal Society B: Biological Sciences*.

Pusey, M. F., J. Barrett, and T. Rudolph (2012). "On the reality of the quantum state". In: *Nature Physics* 8.6, pp. 475–478.

Putnam, H. (1968). "Is logic empirical?" In: *Boston Studies in the Philosophy of Science, vol. 5*. Ed. by R. S. Cohen and M. W. Wartofsky. Reprinted in (Hooker, 1979), pp. 181–208. Dordrecht: D. Reidel, pp. 216–241.

— (1994). "Michael Redhead on quantum logic". In: *Reading Putnam*. Ed. by P. Clark and B. Hale. Oxford: Blackwell, pp. 265–280.

— (2005). "A philosopher looks at quantum mechanics (again)". In: *The British journal for the philosophy of science* 56.4, pp. 615–634.

Redhead, M. (1987). *Incompleteness, nonlocality, and realism: a prolegomenon to the philosophy of quantum mechanics*. Oxford: Oxford University Press.

Rényi, A. (1955). "On a new axiomatic theory of probability". In: *Acta Mathematica Academiae Scientiarum Hungarica* 6.3-4, pp. 285–335.

Roeper, P. and H. Leblanc (1999). "Absolute probability functions for intuitionistic propositional logic". In: *Journal of philosophical logic* 28.3, pp. 223–234.

Romeijn, J. W. (2014). "Philosophy of statistics". In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Winter 2014. URL: http://plato.stanford.edu/archives/win2014/entries/statistics/.

Saeki, S. (1996). "A proof of the existence of infinite product probability measures". In: *The American Mathematical Monthly* 103.8, pp. 682–683.

Saunders, S., J. Barrett, A. Kent, and D. Wallace, eds. (2010). *Many worlds?: Everett, Quantum Theory, & Reality.* Oxford: Oxford University Press, pp. 227–263.

Scheibe, E. (1974). "Popper and quantum logic". In: *British Journal for the Philosophy of Science* 25, pp. 319–342.

Schrödinger, E. (1926). "An undulatory theory of the mechanics of atoms and molecules". In: *The Physical Review* 28.6, pp. 1049–1070.

Shimony, A. (1983). "Reflections on the philosophy of Bohr, Heisenberg, and Schrödinger". In: *Physics, philosophy and psychoanalysis.* Ed. by R. S. Cohen and L. Laudan. Springer, pp. 209–221.

— (1984). "Contextual hidden variables theories and Bell's inequalities". In: *The British Journal for the Philosophy of Science* 35.1, pp. 25–45.

— (1990). "An exposition of Bell's theorem". In: *Sixty-two years of uncertainty.* Ed. by A. I. Miller. New York and London: Plenum Press, pp. 33–43.

Skyrms, B. (1977). "Resiliency, propensities, and causal necessity". In: *The Journal of Philosophy* 74.11, pp. 704–713.

Stone, M. H. (1930). "Linear transformations in Hilbert space: III. Operational methods and group theory". In: *Proceedings of the National Academy of Sciences of the United States of America* 16.2, p. 172.

— (1932). "On one-parameter unitary groups in Hilbert space". In: *Annals of Mathematics*, pp. 643–648.

— (1936). "The theory of representations of Boolean algebras". In: *Transactions of the American Mathematical Society* 40, pp. 37–111.

Strzałko, J., J. Grabski, A. Stefański, P. Perlikowski, and T. Kapitaniak (2008). "Dynamics of coin tossing is predictable". In: *Physics reports* 469.2, pp. 59–92.

Swift, A. R. and R. Wright (1980). "Generalized Stern-Gerlach experiments and the observability of arbitrary spin operators". In: *Journal of Mathematical Physics* 21.1, pp. 77–82.

Timpson, C. G. (2010). "Rabid dogma? comments on Bub and Pitowsky". In: *Many Worlds?: Everett, Quantum Theory, and Reality.* Ed. by S. Saunders, J. Barrett, A. Kent, and D. Wallace. Oxford: Oxford University Press, pp. 460–466.

— (2011). "Probabilities in realist views of quantum mechanics". In: *Probabilities in Physics.* Ed. by C. Beisbart and S. Hartmann. Oxford: Oxford University Press, pp. 201–229.

— (2013). *Quantum Information Theory & the Foundations of Quantum Mechanics.* Oxford: Oxford University Press.

Uffink, J. (2011). "Subjective probability and statistical physics". In: *Probabilities in Physics.* Ed. by C. Beisbart and S. Hartmann. Oxford: Oxford University Press, pp. 25–49.

Valentini, A. (1991). "Signal-locality, uncertainty, and the subquantum H-theorem. I". In: *Physics Letters A* 156.1, pp. 5–11.

— (2010). "Inflationary cosmology as a probe of primordial quantum mechanics". In: *Physical Review D* 82.6, p. 063513.

Wallace, D. (2003). "Everettian rationality: defending Deutsch's approach to probability in the Everett interpretation". In: *Studies in History and Philosophy of Modern Physics* 34.3, pp. 415–439.

— (2007). "Quantum probability from subjective likelihood: improving on Deutsch's proof of the probability rule". In: *Studies in History and Philosophy of Modern Physics* 38.2, pp. 311–332.

— (2010). "How to prove the Born rule". In: *Many worlds?: Everett, Quantum Theory, & Reality.* Ed. by S. Saunders, J. Barrett, A. Kent, and D. Wallace, pp. 227–263.

— (2012). *The emergent multiverse: Quantum theory according to the Everett interpretation.* Oxford: Oxford University Press.

Waters, R. (1970). *If.* From: *Atom Heart Mother* by Pink Floyd.

Weatherson, B. (2003). "From classical to intuitionistic probability". In: *Notre Dame Journal of Formal Logic* 44.2, pp. 111–123.

Wheeler, J. A. (1977). "Include the observer in the wave function?" In: *Quantum mechanics, a half century later.* Ed. by J. L. Lopes and M. Paty. Dordrecht: D. Reidel, pp. 1–18.

— (1990). "Information, physics, quantum: the search for links". In: *Complexity, Entropy and the Physics of Information.* Ed. by W. H. Zurek. Redwood City: Addison-Wesley, pp. 3–28.

Wheeler, J. A. and W. H. Zurek (1983). *Quantum Theory and Measurement.* Princeton: Princeton University Press.

Wick, G. C., A. S. Wightman, and E. P. Wigner (1952). "The intrinsic parity of elementary particles". In: *Physcial Review* 88.1, pp. 101–105.

Wigner, E. P. (1952). "Die Messung quantenmechanischer Operatoren". In: *Zeitschrift fü Physik* 133, pp. 101–108.

— (1961). "Remarks on the mind-body question". In: *The Scientist Speculates.* Ed. by I. J. Good. London: Heineman, pp. 284–302.

— (1963). "The problem of measurement". In: *American Journal of Physics* 31.6, pp. 6–15.

— (2010). "Measurement of quantum mechanical operators". Trans. by P. Busch. In: arXiv:1012.4372 [quant-ph].

Wolchover, N. (2014). "Time's arrow traced to quantum source". In: *Quanta Magazine.* URL: https://www.quantamagazine.org/20140416-times-arrow-traced-to-quantum-source/.

Yu, S. and C. H. Oh (2012). "State-independent proof of Kochen-Specker theorem with 13 rays". In: *Physical Review Letters* 108.3, p. 030402.