

Quantum Measurement as Theory: Its Structure and Problems

Marco Del Seta

LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Thesis submitted for the degree of Doctor in Philosophy
of the University of London

Date: July 1998.

UMI Number: U615434

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U615434

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

THESES

F

7584

701789

ABSTRACT. This thesis deals with the set of issues commonly known as the ‘measurement problem’ in quantum mechanics. The main thesis is that the problems are best understood as typically theoretical problems, in the sense that they are not problems directly concerned with the ability of the quantum theory to account for, or represent, actual measurements. This is contrary to the standard view that the quantum measurement problem is in fact about how to fit theory to experiment.

I explain how I characterise a theoretical problem and argue against claims that quantum measurement theory is unrealistic or ineffective because it bears so little relation to actual measurement practice: I argue that the quantum theory’s analysis of measurement need not be committed to doing for the experimenter what Henry Margenau and other critics think it should do. Its principal aim is to answer two questions. First, it tells us what properties are to be associated to quantum states; secondly, it tells us what, in the theory, a measurement must be if these properties are to emerge.

I then discuss some of the specific aspects of the problem of measurement, in particular the results known as insolubility proofs of the quantum measurement problem and the characterisation of the quantum measurement interactions satisfying standard probabilistic constraints. I prove several results here, amongst them characterisations of all interactions jointly satisfying the conditions of unitarity and, first, objectification, then secondly, probability reproducibility conditions. These are the standard conditions which capture our intuitions about quantum measurement. I show how the results lead to negative consequences with respect to the interpretive questions in quantum mechanics.

The discussion of these specific aspects of quantum measurements does, on the other hand, suggest a particular strategy for solving the problems. This is found in Arthur Fine’s solution to the measurement problem, which is based on the idea of a selective interaction. The discussion of Fine’s solution emphasises in general how simply implementing technical strategies is not sufficient to solve the measurement problem in quantum mechanics: further arguments must be given for why the strategy is appropriate, rather than just mathematically satisfactory. I claim that the arguments given by Fine are far from sufficient.

The thesis concludes that, although the quantum theory of measurement is immune from Margenau’s critique, and retains a theoretical autonomy, it is still plagued by numerous problems: the thesis identifies clearly what some of these problems are and considers some solutions, most of which, however, raise serious philosophical questions about the interpretation of quantum mechanics.

Contents

Statement of Conjoint Work in Marco Del Seta's Thesis	5
Acknowledgements	6
Introduction	11
Chapter 1. The Relative Autonomy of Theory: Quantum Measurements and Real Measurements	20
1. Quantum measurement and the real world	20
2. How far do the criticisms go?	25
3. The 'relative autonomy' of the quantum theory of measurement	30
4. Conclusions	37
Chapter 2. Measurements and Insolubility Proofs	39
1. A statement of the problem	39
2. A 'proof'	48
3. Insolubility proofs and Real Unitary Evolution	63
4. Insolubility without probability	75
5. Insolubility proofs in the case of unsharp pointer observables	82
6. Conclusions	90
Chapter 3. Describing Measurement Interactions in Quantum Measurement Theory	95
1. Some introductory remarks and definitions	96
2. The contractive partially isometric evolution and its completion in system-apparatus measurement interactions	103
3. Partially isometric contractive mappings in the density operator formalism	112
4. The problem of completion for mappings on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$	118
5. Some informational results for mixed initial apparatus states	121

6. Conclusions	131
Chapter 4. Algebraic Approaches in Fine's Solution to the Measurement Problem	134
1. Fine's Solution to the Measurement Problem	135
2. Algebraic Quantum Theory and Fine's Solution	139
3. Arguing for the algebraic rationale	150
4. Problems for Fine and Superselection	156
5. Two further problems for Fine's account	159
6. Conclusions	169
Chapter 5. Conclusions	171
Appendix A. Proofs of Theorems	175
Bibliography	184

Statement of Conjoint Work in Marco Del Seta's Thesis

I certify that parts of Chapter 3 of my thesis, specifically sections 3.1 and 3.2, describe joint work with Gianpiero Cattaneo, of the University of Milan. My share of the work in this part of the thesis is 50%.

Marco Del Seta

I confirm that the previous statement, concerning Del Seta's collaboration with Gianpiero Cattaneo, is correct to the best of my knowledge.

Nancy Cartwright

Acknowledgements

When I look at the work that makes up this thesis, I can't help but feel how relatively little it reflects who I am now. It is probably common to change considerably over a period of four years, to learn new things. Most of the material in this thesis, Chapters 2, 3 and 4, was initially written (in much worse form) in 1994, and at least in the gist of the topics remains the same. It is perhaps normal, then, that I should feel this way. The only substantive part of the thesis that is newer is Chapter 1. This reflects how I've come to think about this work in more recent times, and reflects the serious influence that many people have had on me in the meantime. I want to take the occasion to thank them and many others.

My interest in philosophy of physics started in an almost incidental manner. Having spent my undergraduate years, and a year of masters study, in a mathematics department, doing pure and applied mathematics, with not a single person who even knew what philosophy of physics was, I could have been excused for missing out entirely. Nevertheless I couldn't even have dreamt of doing the work I have done without the training I received there. What brought me back to philosophy was philosophy of mathematics; what brought me to philosophy of physics was participation in a seminar organised by the Philosophy of Science Research Students Group (PSRSG) at LSE, where Mao Suárez discussed his Masters thesis on Quantum Logic. I am particularly grateful to him for this, obviously. He also was a constant source of conversations, and has been sorely missed since he has left, for no one else now really works on these topics here. The PSRSG taught me philosophy of science at a time when nobody else was really doing it; I have learned from the seminars by the members of this remarkable group about every manner of topic, and I should mention in particular Samet Bagce, Tim Childers, Robin Hendry and Stathis Psillos. It still is the best forum I can remember for the discussions, the quality of outside speakers when we invited them, and the beer sessions.

My greatest intellectual debt in this thesis is, in different ways, to three people: Nancy Cartwright, Gianpiero Cattaneo and Arthur Fine. Nancy for being a supervisor and friend, and extremely good at both. Thanks to her I'm just about getting the hang of this 'writing philosophy' lark. She has also been responsible for me becoming a jack of all trades, particularly encouraging me to move into economics. I'm still unsure whether this was a wise thing to do at this stage of my career, but it's been enormously fun anyway. Her work in reading and commenting the thesis has been really thorough. I wanted to make some jokes about this, but it really wouldn't be appropriate: any mistakes and faults that remain must surely be blamed on me. Gianpiero Cattaneo has taught me a fraction of what he knows (I'm still lagging behind with the rest), in his inestimable manner: as well as being co-responsible for one of the results in the thesis, he has taught me about such things as unsharp observables, fancy functional analysis and many other things, and has been a truly enjoyable visitor and friend in his many trips to London. Arthur Fine had a whirlwind trip through London in 1994, dropped a couple of papers, gave two seminars and that prompted me to write my first two papers; it was my first year as a PhD student, I was young and impressionable and that was that. Since then I have scarcely met him a couple of times, but I still feel that his general attitude to philosophy is one of the ones I respect the most.

More specific debts are to Nancy Cartwright and several people in the Philosophy Department, particularly Craig Callender, Colin Howson, Thomas Uebel and John Worrall, for chapter 1; James Cushing commented on a *very* early version of chapter 2, Abner Shimony helped with a more recent one, and Paul Busch was of great assistance over email for questions regarding the latter part of that chapter; chapter 3 originates from a result worked out in conjunction with Gianpiero Cattaneo, and Mauricio Suárez also commented extensively; chapter 4 has been revised with the help of R. I. G. Hughes and Nancy Cartwright.

This takes care of the most direct intellectual debts vis-a-vis my thesis. The list of "without whom's", on the other hand, is another question altogether.

It's impossible not to start with my family, whose support has been never ending, on the Italian side, but also on the English one; I can never thank my parents enough for all the support they have given me over the years. My girlfriend, Claudia, has borne my moods stoically, in a way that I am infinitely grateful for. Her constant presence, here, on the phone, on the email, any which way, has made it all much easier, both

to work and to forget about work. On which topic, friends of various non-academic kinds have been indispensable in getting away from it all: we've drunk, eaten, played bridge, danced in various of the London establishments and done all sorts of things unbecoming of a seriously minded academic. Most of all, we've not talked philosophy. For that, I thank them.

The Centre for Philosophy of Natural and Social Science has been my home for the past five years, first as a student, then as staff. I have enjoyed the academic opportunities offered by the place enormously, and it is fair to say that the Centre as a whole supervised my work. I talked physics at some time or other with many people. Some have helped in a decisive way with my thesis: Jim Cushing and Abner Shimony in particular spring to mind. Some have been great company both intellectual and social: Teddy Seidenfeld, who was in my office encouraging me during difficult times trying to prove a damned result, Lawrence Sklar on trips to Chinatown, Jeff Bub who was very helpful with other work I have been doing, and, on many different and pleasant occasions, Cristina Bicchieri and David Malament. Many other people have visited the centre, and it's been a pleasure to know them all. I'm very grateful to all of them, especially James Woodward who gave me a wonderful book which I would never have been able to find otherwise. I've benefited from the presence of many other people around the Centre: amongst the students I would like to mention Sebastiano Bavetta, Francesco Guala and Towfic Shomar in particular, who have been great friends and with whom I have worked a lot in very different fields. Amongst the academics, I should mention in particular Mary Morgan, who has been tremendously encouraging in my forays in economics, and Margie Morrison. Amongst the colleagues here, I would like to thank Rudy Fara for always being there in moments of need.

In the Philosophy Department I have to thank Craig Callender, John Worrall and Elie Zahar for reading my material at different times and being of great help to me in many other ways. Peter Urbach was instrumental in guiding me through what was not always an easy first year in London, during my Masters, and if I have gone on to do my PhD it is largely thanks to him. The administrative staff in the Department have always been a great help in sorting out the details, and I owe them a lot, too. I must not forget Colin Howson and Thomas Uebel, partners in crime, who have spent endless hours, good and bad, with me in the Beaver's Retreat, surely the greatest bar I have ever seen in an academic institution anywhere. The conversations there, more often

gossip and snide remarks than deep philosophy if truth be told, have been priceless and will, in all likelihood, be irreplaceable. I also want to thank R.I.G. Hughes, to whom I had the opportunity to talk about work both in this thesis and outside of it with great benefit, and who was also a regular member of the Beaver's gang during his permanence here.

Thanks are also due to several members of the Philosophy Department of the University of Florence, in particular Maria Luisa Dalla Chiara and Roberto Giuntini. In my frequent trips to Italy they have proved to be a valuable source of conversations of an intellectual nature, and have always been good company in many other ways, too.

The last academic thanks, while clearly not the least, must go to 'the historians'. The term denotes a somewhat heterogeneous group of people from London, Cambridge and other places, who have of course spent a lot of time with me in bars, but have also done a lot more. They have reminded me of what it is to think historically and taught me much more about this, so much so that I now find it very hard to stop doing it. It is in large part due to them that I now find it quite difficult to look at my thesis in the same way that I did four years ago. I would like to mention, in particular, Rob Iliffe and Andrew Warwick from Imperial College as the prime movers for this change. To these I should add many others, at least William Ashworth, Clinton Chaloner, David Edgerton, Simon Schaffer and undoubtedly many more whose names escape me now.

Finally, I want to thank the London School of Economics for providing me with a substantial scholarship during the first three years of my PhD and the Department of Philosophy, Logic and Scientific Method for topping that up with the Popper Scholarship. I also have to thank Pat Gardner, Theresa Hunt, Frances Townshend, Charlotte Savery and many others who have worked in the Philosophy office; amongst the many valuable services they have provided for us all, they have organised my job files and made sure that I could apply for jobs in the best possible way. The CPNSS also contributed research funds for me, and eventually took over as my prime funder two years ago when they offered me a job here. They got their pound of flesh! And it would not have been possible to even attempt to do my job here without the constant backup of Kate Workman, Dorota Rejman and all the other people who have worked in the administrative office for the past five years. I am also grateful to the Centre for History of Science, Technology and Medicine, and to the Humanities Section, at Imperial College. They have employed me to lecture the undergraduate Philosophy of Science

course there for the last two years, an extremely valuable experience. Again, the administrative staff made it all the more simple. My thanks go also to the many people who have invited me to contribute talks to seminars and conferences, amongst other places in Cambridge, Florence, Maastricht and Chicago. Sometimes I'm embarrassed at how messy my material has been when I have given talks on it. It has certainly improved enormously as a result of presenting it, and all these experiences have been fun and a great help.

Introduction

This thesis discusses the work on quantum measurement from a foundational and philosophical perspective. Its title reflects the status I associate with this work: I want to argue that it is a typically theoretical enterprise, concerned with establishing theoretical facts, as opposed to fitting experimental facts to theoretical ones. In order to do this, I will both put forward the philosophical argument for this claim, and explain and critically analyse quite what the theoretical work I refer to consists of. I will also analyse a case where the questions that quantum measurement raises are analysed from a point of view that is somewhat opposite to the one I endorse here, in that it seeks to rely on experimental facts. As well as raising a number of problems associated with the answers this case provides, I hope to show not only how quantum measurement is typically theoretical, but also that bringing it closer to experiment is fraught with difficulties.

I will begin by outlining the contents of the chapters which follow in the thesis, and then make some general comments on its contents.

Chapter 1 argues that the theoretical account of measurement that quantum mechanics provides is not about giving realistic accounts of experimental measurements, but rather about an attempt to answer questions which emerge in quantum *theory*. The argument is directed against criticisms of the formalism of quantum measurement which go back, in some form or other, at least to Margenau [64, 66], and which have been repeated in some form or other by, amongst others, Park [69] and Chang [22, 23].

The strategy of the arguments against the formalism of quantum measurement begins from the fact that projections (or ‘wave function collapses’) either happen, or appear to happen. By *projection* I indicate the well-known phenomenon in quantum theory that takes place upon measurement of a physical quantity on a system. In Dirac’s original statement [29, p. 5–6], the state of the system undergoes a discontinuous

transition when such a measurement takes place: it evolves from a *superposition* of states where the measured physical quantities have exact values, to one of these states.

In the standard Hilbert space representation of quantum mechanics, *states* are taken to be one-dimensional projectors on the Hilbert space which models the system (or indeed the unit vectors of the one-dimensional subspaces which are the image of the projectors). This is usually generalised to include convex combinations of the projectors, this whole set representing all the possible states. Elements of this set are normally called density operators, and the projections are also called pure states.

Observables, the physical magnitudes of the systems, are usually represented as self-adjoint operators on the Hilbert space; *eigenstates* of these operators are states where the physical quantity associated with the operator has a definite value, this value being the eigenvalue associated with the eigenstate. Projections over eigenstates of operators representing physical observables, together with their algebraic sum, capture *properties*: in the standard approach to quantum theory a state then has the property represented by such a projector (which can be a projector over a subspace of whatever dimension) if and only if it itself projects on an eigenstate of the ‘property’. Pure states which are not eigenstates of a given observables are said to be in a *superposition* with respect to that observable.

Von Neumann [82, Ch. V] generalises Dirac’s projection postulate in the following way. The transition should not necessarily be understood as one from a superposition to an eigenstate of a measured observable. It is best understood as a transition from a projector on a superposition to a non-pure density operator which is a convex combination of appropriate eigenstates; this is Von Neumann’s *type I* evolution.

The quantum theoretical accounts of measurement assume either that projections happen, or that it is as if they happened, so an account of measurement that denies that projections take place should nevertheless be able to account for measurement results *consistent* with projections happening.

Margenau first criticised the very idea that projections happened, or seemed to happen, drawing on examples from empirical practice. In doing this he would undermine the whole approach of the quantum formalism: what use would a theory trying to account for projections be, if projections didn’t happen, or happened very rarely, and if states changed in a manner that did not mirror what projection postulates claim?

His most famous example is that of destructive photon measurements: when we measure photon polarisation, photons are destroyed, they never end up in eigenstates. What does the projection postulate have to say about this?

Chang in particular [23] has recently re-emphasised and reinforced this critique, explicitly directing it to the formalism of quantum measurement in general, and providing a number of good examples to prove his point.

My argument claims that this criticism is wrong. Its mistake lies in assuming that the aim of the formal account of quantum measurement is to provide a description or a calculational tool for all real-life experiments. I claim that the quantum theoretical accounts of measurement are designed to account for two questions:

1. What properties can we assign to quantum states?
2. What kinds of evolutions are appropriate for quantum systems?

I think that in answering these two questions it is not crucial that we maintain a close connection to modelling experimental results, for these questions are focused on problems of a specifically theoretical nature. In making this claim, it is still necessary to explain why theoretical work need not depend on what happens in experiments. I will try to make sense of this by arguing for what Lakatos called the relative autonomy of theoretical science.

These questions have two traditional answers. The quantum theorist's answer to the first question has been captured by the so-called eigenvalue-eigenstate link, which was mentioned above: a state has the property represented by a projector if and only if this state projects on an eigenstate of the 'property'. The answer to the second question claims that unitary evolutions, identified by the quantum mechanical Schrödinger equation, are the appropriate evolutions for quantum systems. Von Neumann's type I evolutions are problematic, and it would be nice to be able to do without them. For one thing they violate historical determinism. As Earman has put it,

Letting \mathcal{W} stand for the collection of all physically possible worlds . . . [the] world $W \in \mathcal{W}$ is *Laplacian deterministic* just in case for any $W' \in \mathcal{W}$, if W and W' agree at any time, then they agree at all times. . . . This concept of determinism can be broken down into two subconcepts. A world $W \in \mathcal{W}$ is *futuristically* (respectively, *historically*) *Laplacian deterministic* just in

case for any $W' \in \mathcal{W}$, if W and W' agree at any time then they agree at all later (respectively, earlier) times. [30, p. 13]

If quantum mechanics admits of Von Neumann's type I evolutions, then it clearly is not historically deterministic. There will almost always be infinitely many states mapped by a type I evolution to the same final state. If they are to be typical of the measurement process, this would make quantum mechanics indeterministic in a way that is not only conceptually problematic, but also problematic for the relationship it implies with other theories which do not share this same feature. Can we do without these evolutions?

This question is at the root of the work in subsequent chapters of the thesis. The second chapter examines in detail why the two traditional answers are not co-existent, clarifying and extending the literature on the insolubility proofs of quantum measurement. The starting point here is that we want our measured object to look like it is undergoing a type I evolution, while arguing that this is not really what is going on. The answer to this is to enlarge the system under consideration to include the measuring apparatus. Can the evolution be deterministic? Can we assign the right properties at the end of the evolution in a way that the eigenvalue-eigenstate link is in some sense preserved, while giving rise to the *correct* properties?

This latter requirement, in the context of the description on the enlarged system, is demanded not necessarily of the object system, but at least of the apparatus, which records the measured values. In order to ensure that the eigenvalue-eigenstate link is applicable to the final state after a measurement has taken place, the unitary measurement operator U effecting the measurement has to satisfy the *objectification* requirement. This requirement demands that the final state of the combined system be a convex combination of eigenstates of the pointer observable for the combined system.

Is objectification sufficient to guarantee that the right properties are possessed by the apparatus subsystem? No; we need some sort of probabilistic condition, too. The standard condition that is assumed is the *probability reproducibility condition*: if the object state is such that, upon measurement and projection, it is transformed into a certain eigenstate of the measured observable with probability α , then the pointer should point to the value associated with that eigenstate also with probability α . The problem is that no unitary operator which satisfies objectification can satisfy either the

probability reproducibility or any reasonable generalisation of it. That is the core of traditional insolubility results.

Chapter 2 will discuss these results. It provides several generalisations of the results known so far. In particular, it proves a characterisation theorem for all unitary operators satisfying objectification, which allows for a ‘constructive’ proof of insolubility of the quantum measurement problem.

In the process of doing this the chapter will try to resolve a long standing conceptual controversy between Shimony and Stein on the one hand, and Fine and Brown on the other, as to the validity of the latter two’s proof of insolubility, which assumes an additional principle, Real Unitary Evolution (RUE). I will show how the technical content of RUE is in fact deducible from unitarity and objectification.

I will then provide a preliminary insolubility proof for the case when the pointer observable is unsharp. This is a rather delicate question: as I will show, the formal statement of an unsharp objectification is not trivial, and it is far from clear that insolubility results will continue to hold under all possible conditions of unsharp objectification. I will propose such a statement and go on to prove an insolubility result for a restricted set of cases.

It is important to discuss why it is necessary to consider the case of unsharp pointer observables. The framework in which unsharp observables are introduced is slightly different from the one outlined earlier for standard observables. Another way of looking at observables is as mappings from possible values of physical quantities to projection operators, that is as projection-valued measures (PV measures for short). It is possible to recover the self-adjoint operators by adding up the projection operators in the range of the measure, multiplied by the values associated with them, so that there is a one-to-one correspondence between PV measures and self-adjoint operators on a Hilbert space. PV measures can be generalised in much the same way that pure states are generalised to density operators. Namely, values can be mapped to convex combinations of projections, called effects. In this case the projections can be over subspaces of dimension greater than one, rather than one-dimensional projections only as is the case with states, and the resulting, extended set of observables is known as the set of positive operator-valued measures (POV measures for short). The set of PV measures is a subset of this, known as the set of *sharp observables*. POV measures which aren’t sharp are known as *unsharp*. This construction is consistent with states and observables yielding

probabilities, through the calculation of the trace of the product of the state (as density operator) by the appropriate property (as effects).

The generalisation is not only mathematical, but has important physical consequences and applications, too. It was introduced by, amongst others, Davies and Lewis [26] in the context of setting up an operational approach to quantum mechanics. Many of the physical applications have been extensively discussed by Busch, Grabowski and Lahti [13], who show for example how the analysis of the original Stern-Gerlach experiment necessarily involves POV measures. Another important physical aspect of the set of POV measures is that it contains unsharp joint position and momentum observables.

It will help to give an example of an experimental setup which is necessarily analysed in terms of unsharp observables. This I do in Figure 1, which is taken from a paper by Pessoa, which discusses some of the aspects of measurement and unsharp observables [70].

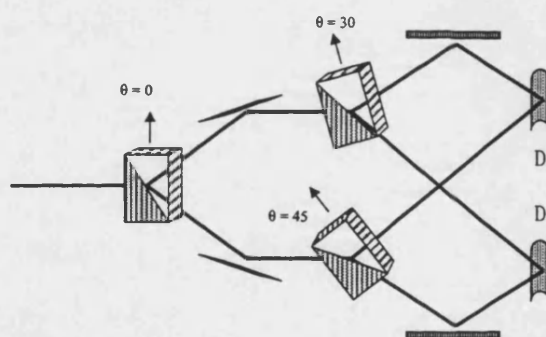


FIGURE 1. A measurement setup where the detection operators are unsharp

Here a beam of light is separated at the birefringent prism oriented at $\theta = 0$ degrees. The top and bottom components then fall on two further analysers, oriented at $\theta = 30$ and $\theta = 45$ degrees respectively. The observable representing the fact that the particle is detected at D_1 or D_2 is resolved as a positive operator valued measure which is unsharp: each one of the two detectors receives a fraction of the top and bottom components coming out of the first prism, rather than receiving definitely up or definitely down photons as would be the case if the observable was a PV measure.

In the case of the measurement problem, consideration of unsharp pointer observables is particularly important, for it is an open question whether unsharp objectification and unitarity will contradict probabilistic conditions, as happens for the sharp case. In fact it is pretty much an open question what unsharp objectification precisely

consists of. The importance of this analysis is that, were it possible to provide a model of measurement satisfying these conditions for unsharp observables, this would count as a solution of the measurement problem, at least from the formal point of view. While my conclusions are sceptical on this topic, it should be stressed that this does depend in a fundamental way on the definition that I propose, and defend, of unsharp objectification.

Chapter 3 addresses questions connected to probability and measurement in quantum mechanics. Its aim is to explore to the full the implications of the other assumption made in the previous chapter, namely the probability reproducibility condition. I discuss how this condition works with the demand of unitarity of the measurement process and more generally, for the case when states are taken to be density operators, bijectivity.

The probability reproducibility condition is not the only condition that has been proposed in order to capture the probabilistic aspect of a good measurement. While it dates back to at least Von Neumann, Fine [35] suggested other, more general conditions: these amount essentially to asking the pointer apparatus to discriminate between probabilistically different object states, without asking it to get the probabilities associated with the object absolutely right. These are an important conditions to bear in mind when discussing insolubility proofs, for they are considerably more general than the probability reproducibility condition. However they are problematic in several different ways, as again Fine, and also others, have pointed out, when used to characterise actual unitary operators effecting *premeasurements*. Premeasurements are unitary operators which map an initial state to a semi-final state, so to speak, namely to a state which just predates collapse of the wave function. In so-called no collapse interpretations, these premeasurements are actually all that there is to measurement.

In this chapter I discuss how to define premeasurement interactions that satisfy the unitarity and probability reproducibility conditions. I will present a detailed generalisation of a result by Beltrametti, Cassinelli and Lahti, which fixes the type of interactions for the Hilbert space case. I will then explore the question of how to define such mappings for the density operator formalism. The question is less straightforward in this case, though I will show that when measurements satisfying probability reproducibility and a suitable generalisation of unitarity exist, they must be actions on density operators of the form UTU^* , where T is a density operator and U is a unitary operator on the

composite object + apparatus system fixed by the result which generalises Beltrametti et al.'s theorem.

This result enables us to establish a number of consequences for the density operator case concerning the types of unitary premeasurements that will be possible. I will show that, if the initial apparatus state is a non-pure density operator, then no Von Neumann, repeatable or first kind measurements are ever possible. I will also show how there are cases where *no* unitary premeasurement satisfying probability reproducibility is possible at all. I discuss some of the consequences of these cases at the end of the chapter.

Chapters 2 and 3 both give characterisations of operators which satisfy objectification and the probability reproducibility condition, respectively, as well as being both unitary. If we wish to maintain all three of these conditions, then from a formal point of view an obvious thing to do is to look for subsets of the set of states where the two characterisations agree. This is effectively what superselection approaches to the measurement problem do. It is also what Fine does, in a series of recent papers proposing a solution to the measurement problem. I will be examining Fine's solution and comparing it to solutions which adopt superselection rules. The aim of Fine's solution is to consider what measurements actually do. This is of interest, for it is not the usual strategy that we find in work on the quantum measurement problem. On the other hand, as I argued in chapter 1, it is not entirely clear that this is the aim of the work in the measurement problem. This makes Fine's work doubly unusual, for he is changing the question somewhat, as well as offering a different answer.

His solution claims that measurement interactions interact not with the full physical details contained in the standard quantum states, but only with 'aspects' of these details, namely the probabilistic aspects related to the measured object observable. This leads him to rewrite the initial states in such a way that they encode only the information of these aspects. These states happen to be states for which the standard unitary operators satisfying probability reproducibility also satisfy objectification, so that the measurement problem can be solved for them.

I think there are basically two kinds of difficulties with Fine's solution to the measurement problem. The first kind enters into the picture when the formal similarity with superselection approaches is taken on board. The similarity becomes more than

formal, I argue, if we properly explore Fine's suggestion that we give an algebraic argument in support of his solution. I explore how the various kinds of difficulties implicit in the superselection approach re-emerge in Fine's approach, and argue that for the most part the difficulties are still there.

The second kind is concerned with how the technical device which gives a solution of the measurement problem is connected with what Fine claims real measurements do. I argue, amongst other things, that the notion of interaction with an aspect faces problems in certain kinds of measurement interactions, namely non-ideal ones, and that in general Fine is unsuccessful at persuading us that the measurement operators model interactions with an aspect, for there are far too many aspects that are captured by the standard measurement operators analysed in Chapter 3, which Fine assumes to be the appropriate ones to model measurement.

Summing up, Chapter 1 provides an overview of what I take the quantum theory's analysis of measurement to be about, and a defence of this project against criticisms that accuse it of falling short of the standards of accurately describing what happens in real measurements. Chapters 2 and 3 examine some of the standard conditions that are required of the quantum theoretical account of measurement, proving a number of new results, analysing some of the debates in the field and raising some new problems. Chapter 4 examines a particular solution that has been proposed by Fine and that naturally suggests itself in the light of the preceding analysis. I claim that the solution fails both if it is supported by algebraic techniques of quantum mechanics, as Fine, suggested, and in general, too.

CHAPTER 1

The Relative Autonomy of Theory: Quantum Measurements and Real Measurements

Is the quantum theoretical description of measurement empty? If it is, some serious consequences follow: a great deal of foundational work is barking up the wrong tree.

There is a long tradition arguing precisely this; the quantum theoretical description of measurement is not a theory of measurement when we think about the nature of actual physical measurement. In this chapter I want to show how the quantum theory of measurement is not empty and can be defended *in its current aims* from such criticisms, by relying on the idea of the relative autonomy of theoretical science. The phrase is due to Lakatos (see [60]), but such ideas are present in a variety of forms in a variety of writers.

I will begin by describing an important line of criticism of quantum theories of measurement, due essentially to Margenau [66, 65], further discussed by Park [69] and, in a more general form, by Chang [23]. I will then raise some problems for these criticisms and show how I think we should understand them. Finally I will claim that the critiques do not affect the work done in the quantum theoretical accounts of measurement and that it is no bad thing. I will explain this by placing work on quantum measurement squarely in the “theoretical” camp, and arguing for the relative autonomy of theory.

1. Quantum measurement and the real world

The criticism I want to address is that quantum mechanical accounts of measurement processes do not accurately describe (let alone explain) what happens in real world measurements, and that they should therefore be rejected, or that at least their role in foundational debates in quantum mechanics should be much reduced.

I will look at two versions of the argument. Margenau is concerned to establish the unreasonableness of the projection postulate in its original form (as stated in Dirac [29] and further expressed as Von Neumann’s type 1 evolution) as an account of what

happens in real measurements. More recent work, such as the paper by Chang, extends Margenau's type of argument to include the two-system interaction picture of measurement, again originally due to Von Neumann; the latter account is what is generally termed 'quantum theory of measurement' today.

Recall the projection postulate in its original form: we want to measure an observable which has a set of allowed values, represented by eigenvalues $\{e_i\}$ of the corresponding operator O , associated with eigenstates $\{\varphi_i\}$. The projection postulate tells us that at the end of the measurement each individual observed system will be in one of the states $\{\varphi_i\}$. The postulate guarantees, in its usual understanding, that at least by the time the measurement is over, if not before, the system has the value that we observe it to have upon measurement: it will yield just that value again if the measurement is immediately repeated. Von Neumann generalises the idea to that of a type 1 evolution of a system [82, p. 351]. Instead of discussing states as one-dimensional subspaces of a Hilbert space, think of the description of states as density operators, where a density operator is either a projection onto a one-dimensional subspace (conventionally a *pure* state), or a convex combination of such projections, that is a weighted sum with positive coefficients summing to one (conventionally termed a *mixed* state). Then a type 1 evolution is a mapping of a pure state to a mixed state, where projections decomposing the final state are projections onto eigenstates of a measured observable.

Extending this story to the interaction picture of measurements requires including in the description of the measurement evolution an apparatus system, and studying the possible interactions which will yield a final state, after measurement, compatible with our observation that the pointer points to a definite value. We assume that the object states are rays in a Hilbert space \mathcal{H}_S , while the apparatus is represented in a Hilbert space \mathcal{H}_M . The measured observable O must be linked up to a pointer observable M , so that a pointer being in an eigenstate ψ_i indicates that the value associated with the eigenstate of the measured system observable φ_i has been detected. The measurement will consist in the first instance of an interaction on the joint system $\mathcal{H}_S \otimes \mathcal{H}_M$ which transmits information from the object to the measuring apparatus. Then crucially, we will want the final state of the combined object + apparatus system to be an eigenstate (or perhaps a mixture of eigenstates) of the pointer observable $\mathbb{I} \otimes M$ on the joint system.

In *The Nature of Physical Reality* and subsequent work Margenau gives an argument against the projection postulate. The claim for which he is most known concerns the absurdity of applying the projection postulate to concrete measurement situations. His example is that of the measurement of photon polarisation, the very example that Dirac [29, p. 8] uses. To measure photon polarisation we place a photometer next to a polarisation prism to record the passage of a photon through the prism. The photon is *absorbed* by the photometer and leads to the excitation of atoms in an energy chain reaction which causes the counter to click. The key fact, for Margenau, is that the photon is absorbed. In a sense, then, it ceases to exist. In his famous phrase, “we may be pardoned for refusing to discuss the eigenstates of non-entities” [66, p. 377].

The argument is a little more subtle than this. The first part of the argument contends that a measurement must yield *numerical* values. Margenau takes this to be a necessary condition: unless an interaction yields a number it is not a measurement, but merely an ‘operation’. Margenau wants to distinguish the latter, ‘qualitative’ activity from quantitative ones which produce numbers. The term ‘operation’ for Margenau then designates, amongst other things, observations (the kinds of procedures where we observe an effect, like the superconductivity effect, without having made a precise numerical determination of anything), construction of equipment (like setting up the laboratory material for performing an experiment) and, crucially for Margenau, state preparation. A state preparation is any procedure which sets up a state in such a way that it *can* be subjected to a measurement, but does *not* actually produce a number. The preparation procedure can generate a statistical ensemble from which the data are collected, but it need not, and in general will not, be an act of measurement: we can, by filtering photons through a crystal, produce an ensemble of systems prepared in a certain polarisation state, but we cannot call this a measurement until the photons have been suitably recorded, and a numerical value has been produced.

According to the usual story, when a photon is filtered through a prism, if it emerges on the other side a measurement has been performed and the photon must be in a polarisation eigenstate, because if another prism with the same orientation is placed on its path the photon will still go through the prism: we can tell this, for instance, by putting a photometer on the exit path of the second prism. But according to Margenau this story won’t work: the first part of the experiment is a state preparation, and we do not record a numerical value in it. How can we know what state the system

actually is in at that stage? Well, we could place a photocounter between the first and second prism, but then the photon would be destroyed and certainly would not be in a polarisation eigenstate. So the projection postulate is an unrealistic assumption.

Chang has extended Margenau's criticism to the interaction picture of measurements introduced by Von Neumann [82, Ch. IV]. This picture, as mentioned before, generalises the analysis by describing a situation where the system upon which a measurement is to be performed (the object system) is placed in interaction with an apparatus, and the two evolve in such a way that the apparatus reproduces, according to some specific conditions, the properties of the system. Chang claims that measurements, as thus represented, correspond with almost no known real measurement. He calls the picture of measurement emerging from Von Neumann's theory the 'single-interaction' picture: one group of time-indexed unitary operators, all generated by the same Hamiltonian, describes a measurement.

Recall that, in quantum mechanics, the dynamical evolution of a system is characterised by the Schrödinger equation which is specified via an energy Hamiltonian, the sum of the kinetic and potential energies of the system. Given this Hamiltonian, which identifies the kind of interaction one is dealing with, one can calculate for each instant of time t an operator which maps the initial state of the system to the state that the system has at t . Such an operator is unitary, and all the operators for different times, when taken together, form a continuous group, so that the system changes in a continuous way. From this follows that we can think of measurements being realised in one single run, the different time-indexed unitary operators telling us what the stage of the measurement is at a given time.

According to Chang, the normal situation in the laboratory is that we have many more interactions going on in a measurement, and so the picture painted by the quantum theory of measurement is misleading if not downright mistaken. The argument is a generalisation of Margenau's in that it tries to maintain the charge that quantum theoretical descriptions of measurement are unrealistic, while bypassing some possible objections to the Margenau story. I will return to this briefly further on.

Chang argues the point through a variety of examples; the most persuasive one concerns Millikan's experiments which determine the kinetic energy of the electron through electrostatic retardation. The context of such experiments is Millikan's attempt to measure Planck's constant.

The structure of this experiment involves the production of electrons, and the attempt to make them pass through potential barriers of increasing intensity. When the electrons begin to pass the potential barrier in large number, it is possible to say that the kinetic energy has value equal to the corresponding threshold potential energy. Chang contends that it is not possible to make much sense of this experiment unless many runs are performed. This involves carrying through many interactions, with different potential barriers which the electron has to try to ‘overcome’, and these different potential barriers lead to different hamiltonians (expressed as a sum of kinetic and potential energy), lots of them, not just the single one that the Von Neumann interaction picture talks about. Note that we are discussing genuinely different kinds of interactions, since the potential energy is actually different for each one, not many repeated instances of the same interaction.

Chang further contends that there are grounds, given these examples, for appropriating a distinction between notions of measurements, due originally to Park [69]. According to Park we have a notion of measurement, M_1 , which is theory driven, perhaps typically quantum mechanical and is “used in stating the values of observables possessed by quantum states and objects. This concept would be employed even if there were never any laboratory measurements to be described or referred to” [23, p. 156]. This is what we do, for example, when we theoretically ascribe probabilities for certain values to be found *upon measurement*. This is to be distinguished from M_2 , the second sense of measurement, which is actual physical measurement, the act of obtaining certain pieces of information about certain phenomena. This distinction seems plausible at least in order to separate what the quantum theory analyses as measurement from the apparently very different practices that constitute the experimentalist’s measurement.

It is less clear why we should take M_1 seriously, if it fails so conspicuously to make contact with real measurements, which he designates by M_2 .

The key in answering that question [about the relationship between M_1 and M_2] lies in recognising that it is the final information gained from M_2 , rather than anything about the process of gaining that information, that should stand in some significant relationship to the requirements of M_1 . [23, p. 158].

Having argued that the distinction is there, Chang claims that the distinction is best interpreted as one between a mediating formalism M_1 that can be compared with the *numerical reports* of activities in M_2 . The purpose of the mediating formalism M_1 is to help bring out some kinds of predictions, or at least numbers, from the quantum theory. M_2 tells us what happens in real measurements, which in Margenau's sense produce numbers:

M_1 often serves as the conduit of predictions made by the formal theory, and in such cases there is a rather straightforward correspondence between the M_1 -results and some M_2 -results. For instance, the formal theory might make predictions about the energy distributions in certain groups of electrons, and M_2 -measurements of energy can verify those predictions. However, often the predictions of the formal theory are couched in other terms than M_1 results, and in such cases M_1 serves at best in some intermediate steps in making the predictions. [23, p. 158].

Some predictions of quantum mechanics (energies, for example) will be arrived at thanks to M_1 . The details of how M_1 should play this role are left unspecified. In short, according to Chang, M_1 , the kind of measurement studied in the quantum theory of measurement, is at best an occasional, and very peculiar, bridge principle, linking the theory with some kind of observation reports or, in the absence of a theory-observation distinction, a vocabulary that is 'antecedently understood'.

2. How far do the criticisms go?

I now want to point out some problems for these critiques, and explain how I understand them.

Must quantum mechanical descriptions of measurement be abandoned, as Margenau has claimed, or confined to providing occasional predictions, as Chang says?

I want to argue that quantum mechanical accounts of measurement cannot function as literal descriptions of all real measurement processes of physical quantities, nor should they be understood as bridge principles helping the quantum theory to yield its predictions. Insofar as I maintain the first claim I am conceding something to Margenau and Chang's argument. Unlike Margenau and Chang, however, I see this as a first step for a proper appreciation of the role of quantum theoretical accounts of measurement, rather than as a reason for rejecting these accounts.

The notion of ‘relative autonomy of theoretical science’ will then form an important part of my argument for what quantum theoretical accounts of measurement are and why they can legitimately be what they are.

Kronz [54] has pointed out that there are several problems with Margenau’s story. Margenau’s concept of non-entity fails to take into account properly what the absence of a particle means in the quantum domain. What is there, in particular, to stop us considering a state for the absence of a photon, a vacuum state?

Even only as a purely empiricist move, to save the phenomena, this would allow us to rescue the theory of quantum measurements from Margenau’s criticism, particularly when allied with an obvious generalisation of the projection postulate as a reduction not on one system, but on two.

Kronz’s argument goes as follows. Suppose we have an object plus apparatus system described by a state $\varphi \otimes \psi_0$ in the space $\mathcal{H}_S \otimes \mathcal{H}_M$ where $\varphi \in \mathcal{H}_S$ is the initial state of the object and $\psi_0 \in \mathcal{H}_M$ is the initial state of the apparatus. Let the system evolve in the first instance by an appropriate unitary evolution U , let’s call it a premeasurement, and then suppose that a discontinuous reduction, akin to the projection postulate, occurs such that the evolved state $U(\varphi \otimes \psi_0)$ is transformed into a state which is a mixed state for which we can claim that the pointer observable has definite values.

This ‘recipe’ will leave completely unspecified what the final object states might be. Roughly speaking, if the final states of the object are eigenstates of the measured object observable we can call the measurement ideal. This is the case originally studied by Von Neumann. But the possible final object states could all be vacuum states, reflecting the fact that the photons have been absorbed in the process of determining their polarisation; this does not affect the desideratum that the final state be a mixed state where the pointer observable $\mathbb{I} \otimes M$ has definite values, a desideratum which is guaranteed by the generalisation of the projection postulate in the shape of the discontinuous reduction.

Admittedly this opens up another problem: what does it mean to say that we have measured an object to have the value e_i for observable E in this case? In the standard case, we say that an object system has value e_i if, upon measurements being repeated, the system always displays e_i as a value. This implies that (a) the system is in the eigenstate φ_i associated with e_i (at least in the standard reading of quantum theory). Such measurements are called non-disturbing. Non-ideal measurements, on the other

hand, are measurements for which the final object state is not such an eigenstate. So when did the system ever have the value e_i ? Several answers can be given (see for example D’Espagnat’s reading of non-ideal measurements [27, Sec. 18.2]). They seem to presume an ensemble interpretation of quantum mechanics in D’Espagnat’s sense. We can also point out that, insofar as we think that what is important about quantum measurement is the recording of probabilistic frequencies of observables’ values in the object system, non-ideal measurements are perfectly compatible with this requirement (see Chapter 3 for a discussion of this). Furthermore, as we think that it is crucial that at least the pointers point to the right values, this is also compatible with non-ideal measurements. The crucial point here, nevertheless, is that Margenau claims that (a) does not hold in real measurements, and that this invalidates the projection story. The counterclaim, as expressed by Kronz, is that the fact that (a) does not hold does not invalidate the projection story in general: it’s just a matter of recognising at what level (whether at the level of an individual system or at that of a composite system) the projection postulate applies. What we should replace (a) with is quite another question, and one that is addressed by a number of different interpretations of quantum mechanics, with which I won’t be concerned here.

When we consider that in fact the vacuum state is much more than a fictional artifact, but is in fact a concept much used in Quantum Field Theory, Kronz’s argument provides a more than reasonable response to Margenau’s point that the projections are empirically undermined by most measurements: if we treat quantum measurements in the two-system formalism and adopt a suitable generalisation of the projection postulate, which guarantees that pointers behave in the right way, while allowing objects the freedom to disappear, then the examples that Margenau cites do not empirically undermine this picture.

Chang’s examples, however, function as a critique of the two-system formalism of quantum measurement as much as of the projection postulate. Kronz’s argument can work as a reply to Margenau, but Chang’s point still remains: if there are some measurements which do require multiple runs, and thus rely in a crucial way on different Hamiltonians to be described by quantum mechanics, then even the two-interaction picture, which relies on a single Hamiltonian to describe a measurement, will not work. There will be examples of real measurements which cannot be described by it, for it requires, as we have just seen in the discussion of Kronz’s argument, that there

be a unitary premeasurement evolution which initially evolves the state for object + apparatus, and this unitary evolution will be generated by a unique Hamiltonian.

There is something important in Chang's point. That some distinction must be drawn between quantum measurement theory and real measurements was apparent all the way back to Wigner, and good examples of measurements which are problematic in the quantum theoretical account of measurement, formally, methodologically or ontologically, might be actual measurements of mass, or Wigner's own example, that of the experimentalist's measurement of the cross-section. These are, I think, more instructive examples than Chang's own one if all he wants is to establish that there are real measurements for which the quantum theory of measurement has no story to tell.

In fact, the quantum theory of measurement is a theory of how we collect particular information about certain observables, where 'observable' is a technical term for a quantity that is represented by linear, Hermitean operators. But what is being measured when we measure cross-section or mass is a quantity which is not an observable in standard quantum mechanics in this sense, i.e. it has no operator representation. Yet we measure mass and cross-section all the time, and certainly they are real, routine measurements. Quantum measurement theory does not offer descriptions of all the possible measurements that one might perform, that much is clear.

Chang criticises M_1 , the theoretical notion of quantum measurement, on the grounds that it is defective in describing what happens in real measurements. His further claim, as is evident in the quote from his paper reproduced on page 24, is that the role M_1 plays is one of providing predictions.

But if the quantum account of measurement, denoted by Chang's M_1 , is so defective in its descriptions of real measurements, on what grounds should we want to use it at all as a source of empirical *predictions*?

We will not be able to use it to predict what states the measured system will end in. This is precisely one of the points that is being denied in Chang's and Margenau's critique of the projection postulate. The projection postulate claims that measured objects end up in eigenstates of the measured observable. Real experiments directly deny this, as is the case with photon polarisation experiments, or with the repeated measurements performed to calculate the electron's energy in Millikan's experiment: the latter seems to provide no definitive numerical output as to what the final state

of the electron will be after we have determined its energy, so there is no standard prediction for M_1 to make here.

The other possibility is that M_1 predicts probabilities and eigenstates that are measured on object systems. I think that for these cases quantum measurement interactions, as described by M_1 , do not lead to specific predictions for real systems, predictions that could not be arrived at otherwise. Therefore I claim that M_1 has no role to play over and above the standard quantum theory in making these predictions, so if we want to know why this kind of work is important we need to look for something different from this.

In order to see this, think of standard Hamiltonians in quantum theory generating interactions from which we derive predictions, like the Hamiltonian for the hydrogen atom. These are grounded in physical considerations which have a long and established tradition. We can use the Hamiltonians to calculate eigenvalues and eigenstates for a system like a hydrogen atom, and then calculate the probability of a state of the atom having a certain energy, thus generating predictions. We can, for example, calculate with this the probability of the direction of an electron scattering off (rebounding off) the atom, assumed to be in a definite energy eigenstate to begin with, and of the atom returning to a particular energy eigenstate after the collision.

This shows that we do not need to invoke projection postulates or their generalisations to predict energy eigenvalues or transition probabilities, it is just not part of standard practice in quantum mechanics. What we need to know is the Hamiltonian for the former kind of prediction, and the unitary evolution of the system under study for the latter. Born's paper on collisions in quantum mechanics [8] is a good example, amongst many others, for this.

Also, note that measurement interactions are never discussed in terms of the Hamiltonians that generate them, unlike the standard cases of interactions just described. We look directly for unitary operators which have little physical intuition behind them (again this will be evident in the discussion in Chapter 3, and will pose a problem for any discussions of the quantum measurement problem which claim to tell us about how measurements really work, as I will illustrate in chapter 4). Therefore, if we are to find a notion of prediction in the quantum measurement theory M_1 , it is not one that follows the standard patterns of quantum theory which are well exemplified by Born's discussion of hydrogen atom collisions.

It follows that, if prediction were to be the rationale for taking M_1 seriously, we would need to find out quite what prediction means with M_1 , as it doesn't seem to be anything that is standardly understood as prediction in quantum theory. However, I think that, whatever role quantum theoretical accounts of measurement might have, they do not include that of being a source of predictions in quantum mechanics. I will try to give some sense of what I see the quantum theory of measurement as doing in the next section.

3. The 'relative autonomy' of the quantum theory of measurement

Commitment to a precise descriptive account of quantum measurements is at the root of Margenau and Chang's criticism of the quantum measurement formalism. The criticism argues that the formalism does not yield such an account. Denying that the formalism need to subscribe to this commitment removes most of the force of the criticism. Outward commitment to a prediction role for the quantum measurement formalism is Chang's positive argument for the formalism. As I have just argued at the end of the previous section, I don't think that the quantum measurement formalism is about yielding predictions. We need to see what it does, then, before we can assess whether its role can be fulfilled without commitment to a precise description of real measurements.

So what is the role of the quantum theoretical accounts of measurement? I think there are two. The first follows from a question posed by Wigner:

[the quantum theorist] has appropriated the word 'measurement' and used it to characterise a special type of interaction by means of which information can be obtained on the state of a definite object. Thus, the measurement of a physical constant, such as cross section, does not fall into the category called 'measurement' by the theorist. His measurements answer only questions relating to the ephemeral state of a physical system, such as, '*What is the x component of momentum of this atom?*' [87, p. 329]

The quantum theoretical account of measurement seeks in the first instance to establish the necessary conditions for answering Wigner's question: "What is the x component of momentum of this atom?" Measurement, for Wigner, is about determining the *theoretical conditions* for properties to be assigned to physical systems and investigating the cases in which these conditions are fulfilled. A typical answer to this might require

that the system be in an appropriate eigenstate, and would then investigate how this can come about. Alternatively it might just postulate that properties are assigned upon completion of a unitary coupling with an apparatus and impose a number of conditions for this strategy. Note that Wigner distinguishes the theorist's notion of measurement from the experimentalist's simply because the former is more restricted and does not include a number of cases like cross-section. I want to say more: I want to claim that the distinction is such that the theorist does not have to worry about the conditions that the experimentalist has to deal with, and the practice of the laboratory. What makes the conditions theoretical in nature is clear: they are not concerned in the slightest with laboratory practices and conditions. So even though Wigner's question looks like one that an experimentalist might ask, there is no concern whatsoever with the kinds of things that Millikan, for example, mentions as essential conditions for the determination of the photon's energy: these conditions concern the repetition of the experiment several times with successively more refined equipment, in order to guarantee an appropriate vacuum in the machine, the least possible direct human intervention when shaving off copper from the detector and so on. What I am interested in, effectively, is whether Millikan's concerns should be explicitly reflected in Wigner's.

The second aim of the quantum theoretical accounts of measurement is to address the problems of a theory which, if Von Neumann's projection postulate, or some such variation, is accepted, is not, to use Earman's terminology, historically deterministic at the very least.

Again Wigner [87, p. 326] talks of the strange dualism implied by the assumption of two types of change of the state vector: evolution with respect to Schrödinger's equation and projection. The latter evolution introduces indeterminism in the theory. I understand historical indeterminism more or less in the sense that Earman gives. One of the physically possible worlds W is historically deterministic if, given another world W' , if W and W' agree at one point, they agree at all past points.

If quantum mechanics admits of projection evolutions, then it is not historically deterministic. There will always be infinitely many states mapped by the projection evolution to the same final state: for example any two distinct states of the form $\varphi = \alpha_1\varphi_1 + \alpha_2\varphi_2$ will be mapped onto φ_1 by a projection onto φ_1 . A similar point also holds for Von Neumann's type 1 evolutions.

The problems here are felt because determinism is thought to be a powerful heuristic, or a metaphysically preferable characteristic, of physical theories, and because of the problems that arise when quantum theory is to be used or fused with theories that are essentially deterministic, such as for instance general relativity. The problem is further compounded by the fact that quantum theory offers no internal recipe for deciding when the indeterministic evolutions should apply: renouncing these evolutions in favour of determinism would enable us to sidestep this problem, too.

I should stress here that the claim which I will make, that these problems can be investigated independently of direct experimental comparisons, does not mean that experiments never play a role in theoretical physics. Rather my claim will suggest that the role experimental practice plays for theory is not one of the constant guardian against the excesses of theory. It would be wrong, however, to deny that much of the original source of the debates in quantum mechanics on the nature of measurement draws on the results of experiments. Clearly, observation of quantization of energy led to the formulation of the idea of a quantum jump. Also our intuitive idea of what properties should be assigned to a system is motivated by experimental findings. The point is rather that the discussion to which we have been led is by now quite distinct from questions about what happens in real measurements of physical quantities.

There are two kinds of approaches for dealing with the second question, related to the problem of determinism, for example:

- We can deny that one of the two evolutions, usually the non-unitary, indeterministic projection, is an evolution that quantum states actually undergo. We can then attempt to formulate some interpretive algorithm that will provide what the rejected evolution provided: this is the strategy associated, for example, with relative state and modal approaches to the problem.
- We can try to show that in a more general picture, one evolution reduces to the other, or both reduce to an encompassing one: this is the kind of strategy associated with decoherence approaches, which also reject indeterministic evolutions, or with generalising the Schrödinger approach, where we treat unitary, deterministic evolutions as special cases of non-unitary, indeterministic ones.

There is nothing new in singling out the two questions of property assignment and the nature of quantum evolutions, and they do have a place in quantum theory: it is by tackling these two issues that we attempt to solve the measurement problem. I do

not think, however, that discussions of both of these questions requires commitment to a precise descriptive account of quantum measurements, nor to the theory of quantum measurement as a tool for deriving predictions.

In fact it is hard to see how concerns about the structure of real measurements relate to the questions that quantum theoretical treatments of measurement consider. For instance, any final decision about the question of unitary versus non-unitary evolutions would seem to offer little clarification as to whether such evolutions describe actual measurement situations. Suppose for example that we were to argue for a resolution of the question of determinism in favour of unitary evolutions, and therefore purely deterministic evolutions. The implications for the case that Chang discusses, for instance, would be minimal; we would still be left with a single interaction picture of quantum measurements, and we would still have little grounds for modelling measurements of mass entirely within this framework. We would be, in sum, no closer to modelling real measurements by opting for unitary evolutions.

Similarly, it is unclear whether careful modelling of actual measurement situations would enable us to make any headway with regards to the two questions suggested above. Grant that, in general, accounts of measurement must be modelled with reference to repeated interactions; with reference to the second question I have outlined, would the interactions between photons and detectors be unitary or not? It is hard to see anything in Millikan's measurements providing an answer to this question. After all it is an aim of different interpretations of the theoretical accounts of quantum measurement that they yield empirically equivalent accounts of data, so nothing in the outcome of the Millikan experiment would be able to adjudicate between different proposals on the determinism issue. The structure of the experiment itself will suggest what Hamiltonian(s) we should adopt to describe it, but this is not where the problem is taken to be: the problem with assigning properties, and whether these properties come to be in a deterministic fashion or not, arises at the point of observation of the final outcome. It is the final outcome, if anything, that suggests why we might need a projection postulate: photons click on detectors, and what might this mean for our theory? The complications of the experimental setup, aptly described by Millikan as "a machine shop in vacuo" [67, p. 361], don't seem to bear on the question of what happens when the detector clicks.

Far from establishing the worthiness of such questions, however, their independence from the structure of real measurements could be construed as good reason for abandoning them. Their legitimacy and relevance, in spite of their remoteness from actual measurement practice, can be established by accepting that the relationship between theory and experiment is neither dust-bowl empiricist nor so profoundly theory-ridden that the distinction between theory and experiment makes no sense, and accepting the relative autonomy of theoretical science (and of course of experimental, too).

It was Lakatos who spoke of the relative autonomy of scientific theories, in his paper on *The Methodology of Scientific Research Programmes*. He took it to be a historical fact that theoretical science had often developed independently of, or with relatively tenuous links to, experimental science, and wanted to provide an explanation for this within his methodological framework, something that the naive falsificationist view of science had failed to do. The idea can also be found in many authors in different guises. Recently for instance Galison [44, Ch. 9] has provided an account of how the two traditions come together in twentieth century physics, with his attempt to articulate the idea of a trading zone; many of his ideas are close to Hacking's description of representational and interventionist scientists [47], though we owe the first articulation of a historical project of this kind to early modern historians, along the lines sketched by Kuhn [58] in some of his later writings. It is also worth pointing out that it is not only historians or 'historical' philosophers who have recognised the usefulness of drawing such a distinction: Nagel [68, p. 87], for example, gave different characterisations of theoretical and experimental laws and argued that the latter have, so to speak, a life of their own.

For present purposes two features of the notion of relative autonomy of theoretical work are relevant:

1. the use in theoretical research of methods distinct from empirical methods,
2. the attempt to answer different questions.

When viewed from this perspective, criticisms of the kind that Margenau and others have made, of the way in which theoretical accounts of measurement miss the mark with respect to real measurements, lose their force; on the other hand Wigner's appropriation of the term 'measurement' is properly understood and appreciated, as is the substantial work devoted to questions to do with the nature of quantum evolutions.

A case of relative autonomy of theory in action in discussions of quantum measurement is provided by debates on the modal interpretation of quantum mechanics. In the modal interpretation of quantum mechanics we assume that the only evolutions that are acceptable are deterministic, unitary evolutions. The quantum theory of measurement characterises deterministic *premeasurement* evolutions (in a manner that I will discuss in detail in Chapter 3) satisfying the probability reproducibility condition, which guarantees that the probabilistic information contained in the object system is transferred to the apparatus system in an appropriate way; they can be taken to characterise, for instance, the evolutions in Kronz's account which precede the application of the projection postulate on the combined system. Modal interpretations can then rely on this characterisation as yielding the measurement evolutions themselves. What properties should we then assign to the system?

The answer to this question is given by the Kochen biorthogonal decomposition rule (originally proposed by Kochen [52], the present formulation is taken from Baccigaluppi [3]):

Basic Rule: Let the reduced state of a system with Hilbert space \mathcal{H} be ρ , and let the spectral resolution of ρ be $\rho = \sum_i \lambda_i^2 P_i$. Then the system possesses a property P_i with probability $\lambda_i^2 \dim(P_i)$.

Kochen Rule: if $\mathcal{H} = \mathcal{H}^1 \otimes \mathcal{H}^2$ is in a pure state Ψ , then the properties P_i^1 and P_i^2 of \mathcal{H}^1 and \mathcal{H}^2 given by the basic rule are perfectly correlated.

Assuming deterministic, unitary evolutions and the Kochen biorthogonal decomposition rule for assigning properties, we have an answer to both questions which constitute the measurement problem. Is the Kochen rule ad hoc or can we motivate it? The rule has a motivation, but not by reference to experimental results. Clifton [24] motivates this rule by showing that it follows from four 'natural' axioms. On top of them he adds two others, one which ensures compliance with the Kochen-Specker no-go theorem (condition 5) and one which guarantees that the question of property assignment is answered adequately in the case of measurement (condition 6):

1. For any W and any U satisfying $UWU^{-1} = W : U[\text{Def}(W)]U^{-1} = \text{Def}(W)$.
2. For any $W : \{P | PW = 0 \text{ or } W\} \subseteq \text{Def}(W)$.
3. For any W , and any $P, P' \in \text{Def}(W) : P \oplus P' \in \text{Def}(W), P \cap P' \in \text{Def}(W)$, and $P^\perp \in \text{Def}(W)$.

4. For any W , there should exist at least some (not necessarily orthogonal) set $\{P_k\}$ of pure states satisfying $W = \sum_k \lambda_k P_k$ with $\sum_k \lambda_k = 1$, such that for every $P_k \in \{P_k\} : \text{Def}(W) \subseteq \text{Def}(P_k)$.
5. For any W , there exist sufficiently many noncontextual value states $[\]_W$ on $\text{Def}(W)$, and a measure μ_W on the set of all such $[\]_W$'s, such that for any pairwise commuting subset $\{P_1, P_2, \dots\}$ of $\text{Def}(W)$

$$\text{Tr}(P_1 P_2 \dots W) = \mu_W\{[\]_W | [P_1]_W = [P_2]_W = \dots = 1\}.$$

6. For any W , $\text{SR}(W) \subseteq \text{Def}(W)$.

$\text{Def}(W)$ is the set of projection operators which have a definite value in state W and $\text{SR}(W)$ is the set of projections in the spectral resolution of W . Clifton's justification of the axioms is not based on experimental considerations, but rather relies on formal arguments that block other options for answering the question. Condition 1, for example, is motivated by the need to avoid "that [something] else in addition to W and the Hilbert space \mathcal{H} is 'smuggled in' to define the definite-valued observables" [24, p. 49]. On the other hand it is very difficult to imagine which kind of experimental setup could be suggestive of this approach to answering the two questions about quantum measurement.

The upshot of this example, one of many possible ones, is that quantum theoretical accounts of measurements really do address themselves to answering the two questions above, and really do so in a manner that is relatively independent of experimental considerations. One way of further characterising the difference I have outlined here has been given by Kuhn in two of his papers. In "The Function of Measurement in Modern Physical Science" [57], Kuhn argues, like Margenau, that measurement is concerned with producing numbers, and with indicating to us what constitutes reasonable agreement between numbers that the theory gives and numbers emerging from experiments (as well as having occasional roles of confirming and refuting instances). Millikan's paper is a good example of this: from the long discussion of what could and could not be done with the experimental apparatus we learn what constituted reasonable agreement between predicted and measured numbers.

Kuhn assigns quite a different role to an altogether different type of experiment in the paper "A Function for Thought Experiments" [56]. Here Kuhn talks of thought

experiments as exercises which explore the structure of a theory, to eliminate confusions in the scientist's theoretical apparatus:

[...] the function of the thought experiment is to assist in the elimination of prior confusion by forcing the scientist to recognise contradictions that had been inherent in his way of thinking from the start. Unlike the discovery of new knowledge, the elimination of existing confusion does not seem to demand additional empirical data. Nor need the imagined situation be one that actually exists in nature. On the contrary, the thought experiment whose sole aim is to eliminate confusion is subject to only one condition of verisimilitude. The imagined situation must be one to which the scientist can apply his concepts in the way he has normally employed before. [56, p.242]

This is then mitigated, later in the paper, when Kuhn concludes that a thought experiment cannot be directed always and solely to displaying and resolving logical conflicts:

Though the imagined situation need not even be potentially realizable in nature, the conflict deduced from it must be one that nature itself could present. The conflict that confronts the scientist in the experimental situation must be one that, however unclearly seen, has confronted him before. Unless he has already had that much experience, he is not yet prepared to learn from thought experiments alone. [56, p. 265]

I read this as saying that the worth of a thought experiment is that we can take some problem that has presented a scientist in a laboratory and analyse it as a conceptual conflict of perhaps quite a different kind, the answer to which will hopefully shed some light both on the theory and on the problematic experimental situation. While Kuhn himself admits that this is not a universal characterisation, it seems a fair one to apply to most of the work in the quantum theory of measurement.

4. Conclusions

What I have tried to do in this chapter is provide a rationale for quantum theoretic accounts of measurements which allows us to sidestep 'real measurement' critiques of it. Arguments such as Margenau's and Chang's purport to show that, on the basis of actual experimental practice, foundational debates on the projection postulate and on the structure of quantum measurements play a minor, if not nonexistent role, in

providing a quantum theoretical understanding of measurement. I have argued that these criticisms do not apply and that the conclusions they lead to for quantum measurements do not hold: this is because they overemphasise the need for the quantum theory of measurement to account for real measurements. Acceptance of a less than tight link between theory and experiment allows us to weaken the need for such an account and is particularly helpful in the task of providing such a rationale for quantum theoretic accounts of measurements. Nevertheless, the quantum theoretic account of measurement is not totally exempt from problems if we accept this weakened link between theory and experiment: Chapter 3 will present an example of such a problem and contrast it briefly, on page 130, with the examples of this chapter.

The work of people like Lakatos, Kuhn, Nagel and Galison not only draws the distinction between theoretical and experimental science, but also shows how this distinction can function as a positive factor in the development of science, particularly when the two come together. This is the position Hacking argues for, and it is well exemplified by Galison's study of Schwinger's contributions at the MIT Radiation Lab during the war. It is to be hoped perhaps that the quantum theory of measurement might come to play a more active role in conjunction with experimental physics (as has been suggested, for instance, by the discussions on quantum nondemolition measurements [9]), but its role should not be constrained a priori to providing accurate descriptions of measurements or to helping to yield predictions, which for instance in the case of transition probabilities the quantum theory seems quite capable of providing itself.

I have tried to show why we should look at the quantum theory of measurement on its own merit, so to speak; it does not mean that the question of what theoretical treatment we should give of an experiment like Millikan's is uninteresting or uninformative *tout court*. It will be informative, but such an analysis will not offer much to the questions that debates on the measurement problem address. It might, on the other hand, tell us a lot about how theory and experiment come together, but that's another story.

CHAPTER 2

Measurements and Insolubility Proofs

It has been known for many years that the traditional accounts of measurement would not deliver what was required of them: the first well known result on the subject can be found in a paper by Wigner [87]. Earman and Shimony [31], Fine [36], Shimony [77], Brown [11] and Stein [79] have proved successively more general results. A proof by Busch and Shimony [17] has recently shown that the measurement problem cannot be solved even when the measured observable is unsharp.

This chapter aims to do three things. It describes, first of all, a simple and intuitive insolubility result using facts about the unitary dynamics of non-relativistic quantum systems. Secondly, it looks at Fine's proof [36], subsequently revised by Brown [11] and considers whether the philosophical objections to it, concerned with the status of quantum mixtures, can be overcome. This discussion will lead to a more general result, relying on a weaker form of one of the standard conditions imposed on quantum measurement in the discussion of insolubility proofs, and eventually to a characterisation proof for unitary operators satisfying the objectification conditions for quantum mechanical measurements. Finally the chapter tackles the question of insolubility for the case of unsharp pointer observables. Some results are obtained for this case: in particular the preceding proofs straightforwardly generalise to cases where a particular commutativity condition holds. The chapter concludes by assessing the importance of these results for the debates on quantum measurement and the interpretation of quantum mechanics.

1. A statement of the problem

The quantum theory of measurement works with unitary operators such as, for example,

$$(1.1) \quad W(\cdot) := \sum_{i,j} \left\langle \varphi_I \otimes \psi_0^j \middle| (\cdot) \right\rangle (\tilde{\varphi}_I^j \otimes \psi_I),$$

operators satisfying requirements of probability transmission and unitarity in measurements. For the sake of convenience I ignore in what follows the time index that would normally be associated with these operators, so that W represents the operator which maps the initial state of an object + apparatus system directly to the final state. Specific discussion of the details about these measurement mappings and of how they are to be characterised is taken up in the next chapter, together with some problems that emerge from these characterisations; the problems dealt with here do not need a specific discussion of them, but just require a few remarks at this stage.

Operators like W , which act on the tensor space $\mathcal{H}_S \otimes \mathcal{H}_M$, by definition map pure states to pure states. The way to infer the states of systems S and M in this formalism is to apply the partial trace technique (the tracing away of one subsystem given the density matrix corresponding to the pure state at the end of the interaction).

The operator W is in fact part of a tradition that goes back to the origins of discussions of quantum measurement: it is a generalisation of an operator that can be found in the last chapter of Von Neumann's book on Quantum Mechanics [82]. Some refer to it as a *premeasurement* (for instance by Busch, Lahti and Mittelstaedt [14]) and all interpretations of quantum mechanics agree that (some part of a) measurement is captured by operators like it. Some interpretations of quantum mechanics, essentially based on Everett's idea, work with operators such as W and require nothing more, for a measurement to have taken place, than that this operator describes the interaction.

Traditionally, however, it was thought that such evolutions would be problematic. The original problem of measurement saw a difficulty with the final state of a measurement interaction being a superposition of correlated object-pointer states. The idea with describing a measurement as an evolution over a composite system $\mathcal{H}_S \otimes \mathcal{H}_M$ was to try to describe Von Neumann's type I evolution (the irreversible 'projective' evolution from pure states to mixed ones) on a smaller system as a unitary evolution over a larger system. It was felt, as we have seen in the previous chapter, that to have two different kinds of evolution of a quantum mechanical system was a difficulty, compounded by the fact that the projection evolution seemed to occur in accordance with criteria of a non-physical character: references to the unavoidably subjective character of observation, to the necessity of revising our notions of objective reality, to the role of the conscious observer, abound in the early literature on quantum measurement

theory (the *locus classicus* is of course Von Neumann [82], but see also Wigner [86] and London and Bauer [62]).

It is always possible to give a description of a linear contraction on a space \mathcal{H}' as the projection on \mathcal{H}' of a unitary evolution on an extension space \mathcal{H} , with $\mathcal{H}' \subseteq \mathcal{H}$ (see Riesz and Nagy, [72, p. 466]). The operation of mapping an initial pure state to a mixture by means of a Von Neumann type I evolution is indeed a contraction¹. Von Neumann himself gives a means of constructing unitary mappings on extensions of a Hilbert space for the object system, where the extension is interpreted as adding a space for the apparatus.

Consider, however, an initial state $\varphi \otimes \psi_0 \in \mathcal{H}_S \otimes \mathcal{H}_M$, φ a superposition of orthonormal states φ_I , the latter eigenstates of a self-adjoint operator H representing the observable to be measured. The adoption of operators such as W would yield, when applied to states $\varphi \otimes \psi_0$, a superposition of final object-pointer states like $\tilde{\varphi}_I \otimes \psi_I$, for example. This is unsatisfactory in the sense that it exhibits interference between pointer states in the composite system. This precludes an ignorance interpretation for the reduced matrix of the apparatus system so, even if such reduced state is a density operator defined over pointer states, as a Von Neumann type I evolution requires, it is not a proper mixture and cannot be interpreted as a classical probability distribution over possible observed final pointer states.

1.1. Objectification, Ignorance Interpretations and the Eigenvalue-Eigenstate Link. It is worth examining in some more detail why a pure object-apparatus final state is unsatisfactory. The problem lies with the acceptance amongst a number of quantum theorists of the assumption Arthur Fine [37] has termed the eigenvalue-eigenstate link. Von Neumann [82, pp. 200–201, p. 215ff] is once more an obvious reference for the early use of this assumption. Again Fine [38] remarks how this assumption is easily derivable from Von Neumann's more fundamental principle P .

The eigenvalue-eigenstate link is the assumption that a system can be assigned a definite value for a certain quantity if and only if that system is described by a state in a Hilbert space which is an eigenstate of the self-adjoint operator which represents the quantity. That is, if a state is an eigenstate of a certain quantity, then it has a value for that quantity given by the eigenvalue associated to the eigenstate, and conversely

¹For a proof, see the next chapter.

if it has a value, it must correspond to an eigenvalue of the operator, and the system must be in the respective eigenstate.

A measurement leads to a pointer pointing to a precise and definite value, so the pointer system must be in an eigenstate of the appropriate pointer observable. In order for this to be the case, the final state for the pointer must at least be decomposable into a mixture of pointer observable eigenstates, with the appropriate weights. This, however, is not sufficient if we wish to claim that, for the individual pointer eigenstates in the final, mixed apparatus state, the eigenvalue-eigenstate link can be upheld. We need to be able to say that, for this mixture, only one of the states is actual; in other words we want to impose an ignorance interpretation on the mixture. This is formally interpreted as the requirement that the probability weights of the mixture be classical.

This is carefully explained for example by Busch, Lahti and Mittelstaedt² [14]: they show how, for a single isolated system to possess one of a number of properties, it is necessary that the state of such system be a density operator defined over the projections representing these properties.

There are two ways in which a system may come to possess a mixed state. The first is that the system is prepared to be a mixture, for example by constructing an ensemble of individuals with different pure states. The second is by considering separately a system which is part of a composite system that has an entangled state. For a system which is a subsystem of a composite system, the condition that the state of this system be a density operator resolved by orthogonal projectors characterising an observable A is not a sufficient condition for the system having a property defined by one of these projectors: one also needs the composite system to be a density operator having a spectral resolution in terms of eigenstates of the property, represented by the operator $A \otimes \mathbb{I}$.

More formally, for any state $T \in \mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$, where $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ denotes the set of trace class one operators on $\mathcal{H}_S \otimes \mathcal{H}_M$, representing the set of physical states of the system, the reduced state in $\mathcal{T}_1^+(\mathcal{H}_M)$ admits of an ignorance interpretation for a mixture over pointer states $P_{[\psi_i]} \in \mathcal{T}_1^+(\mathcal{H}_M)$ decomposing the reduced state if and only

²Interesting and more detailed investigations on this issue can be found in Busch and Mittelstaedt [16]. The discussion is not unrelated to the distinction between proper and improper mixtures, as introduced by D’Espagnat [27]; a recent good discussion of this can be found in Ghirardi [46].

if

$$(1.2) \quad T \equiv \sum_i \mathbb{I} \otimes P_{[\psi_i]} T \mathbb{I} \otimes P_{[\psi_i]}.$$

Achieving all the conditions for an ignorance interpretation for the mixture over pointer states $P_{[\psi_i]} \in \mathcal{T}_1^+(\mathcal{H}_M)$ will ensure that the assumption of the eigenvalue-eigenstate link will yield the correct pointer values at the end of the interaction.

The question becomes the following: is it possible to find a unitary operator which evolves an initial object-apparatus state to a final state such that the final state will be a density operator decomposable into eigenstates of a pointer observable and which satisfies the appropriate probabilistic constraints?

1.2. Some Notation and Definitions. The first step is to spell out some notation and list the conditions assumed in the following analysis.

$\mathcal{T}(\mathcal{H})$ is the set of trace class operators on a Hilbert space. This is a Banach space, i.e. a complete vector space with a norm defined on it, the trace operation. The set of states of a quantum system is represented by $\mathcal{T}_1^+(\mathcal{H})$, the set of trace class operators with trace equal to one. Observables are usually represented by Hermitean operators on the Hilbert space \mathcal{H} .³ When such observables have a discrete spectrum, they admit of a spectral resolution in terms of a complete orthonormal system of eigenvectors. The dynamics of the system is normally given by a unitary operator on the Hilbert space \mathcal{H} ; we will consider also operations on the space $\mathcal{T}_1^+(\mathcal{H})$ and represent them with bold characters.

The conditions for answering the above question are then as follows. Consider an initial state $T = T_S \otimes T_M$ in the object-apparatus system $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$; an object observable H , a self-adjoint operator on \mathcal{H}_S , and a discrete apparatus observable A , a self-adjoint operator on the space \mathcal{H}_M with spectral resolution $\{\psi_i\}$. If H is itself discrete, and therefore with spectral resolution given by the orthonormal basis $\{\varphi_i\}$, then for any given i we take observation of the pointer position symbolised by the state ψ_i to mean that the state φ_i was measured. In the case when H is a self-adjoint, but not discrete, observable, the apparatus observable, which is always taken to be a discrete observable, will induce a partition on the spectrum of the observable H via a *pointer function* $f : \mathcal{P}(\Omega_M) \rightarrow \mathcal{P}(\Omega_S)$, where $\mathcal{P}(\Omega_M)$ and $\mathcal{P}(\Omega_S)$ are σ -algebras of subsets

³Later on we will consider a generalisation of this, in the context of the discussion of unsharp pointer observables

of the pointer and object observable spectra $\Omega_{\mathbf{S}}$ and $\Omega_{\mathbf{M}}$, under the assumption that $f(\{I\}) \cup f(\{j\}) = \emptyset$ and $\sum_i f(\{I\}) = \Omega_{\mathbf{S}}$. This defines a coarse-grained, discrete version of the initial object observable, which has a spectral resolution in terms of eigenstates, through a procedure originally described by Von Neumann [82, Ch. III.3, pp. 220–221]. Discreteness of the operator describing the pointer observable is required to make sense of the objectification condition. Discreteness of the operator describing the object observable is not important, except insofar as we can note that, given a pointer function, the pointer observable can be taken to measure a discrete, coarse-grained observable defined via the pointer function. Suppose then that \hat{H} is the discretised observable with a spectral set $\{\varphi_i\}$, we again take recording of the pointer value ψ_i to mean that φ_i has been observed.

Is there an operator $U : \mathcal{H}_{\mathbf{S}} \otimes \mathcal{H}_{\mathbf{M}} \rightarrow \mathcal{H}_{\mathbf{S}} \otimes \mathcal{H}_{\mathbf{M}}$ for which the following conditions are satisfied?

1. *Unitarity:* The operator U is unitary
2. *Quantum condition:* The two spaces $\mathcal{H}_{\mathbf{S}}$ and $\mathcal{H}_{\mathbf{M}}$ are proper quantum systems, in the sense that no superselection rules are present
3. *Objectification:* The final state $U T U^{-1}$ after the measurement interaction has taken place is a density operator over eigenstates of the operator $\mathbb{I} \otimes A_{\mathbf{M}}$ representing the pointer observable and acting on $\mathcal{H}_{\mathbf{S}} \otimes \mathcal{H}_{\mathbf{M}}$
4. *The probability reproducibility condition:*

$$(1.3) \quad p_{T_{\mathbf{S}}}^E(X) = p_{\mathcal{R}_{\mathcal{A}}(U(T_{\mathbf{S}} \otimes T_{\mathbf{M}})U^{-1})}^{P_{\mathcal{A}}}(f^{-1}(X)).$$

It is important to make three points about the probability reproducibility condition.

The *first* is that the condition just stated applies in general. For a sharp, discrete measured observable on the object system, on the other hand, it can be put in the following way:

$$(1.4) \quad \text{Tr}[P_{\{\varphi_i\}} T_{\mathbf{S}}] = \text{Tr}[\mathbb{I} \otimes P_{\{\psi_i\}} U T U^{-1}]$$

In (1.3) $p_T^E(X)$ is the probability measure over the set of possible values of the observable E , relative to the state $T \in \mathcal{T}_1^+(\mathcal{H}_{\mathbf{S}})$, and f is the pointer function described above. The state $\mathcal{R}_{\mathcal{A}}(U(T_{\mathbf{S}} \otimes T_{\mathbf{M}})U^{-1})$ is the reduced state of the apparatus system (obtained via the partial tracing operation $\mathcal{R}_{\mathcal{A}}$) after the measurement evolution has taken place. That is, for each set of possible values in the set X , $p_T^E(X)$ is the probability that the

system \mathcal{H}_S , when in state T , has value for the observable E in such a set X . One way to see the difference between the two formalisations of the probability reproducibility condition can be seen by the following example. Consider a continuous observable E on an infinite dimensional Hilbert space \mathcal{H} . For each interval $[a, b]$, $b > a > 0$, in the set of continuous possible values for E , there is a distinct probability measure associated with this interval, call it $p_{[a,b]}$, which maps states to probabilities. Consider now a discretised version of E , denoted by E_D . Suppose that it is discretised over, amongst countably many others, two intervals $[a-x, (a+b)/2]$ and $[(a+b)/2, b+x]$, to which the discretised observable associates projection operators $P_{[a-x, (a+b)/2]}$ and $P_{[(a+b)/2, b+x]}$. Then the probability that E has value in the interval $[a, b]$ when in state T is given by $p_{[a,b]}(T)$, while such probability for E_D is given by $\text{Tr}[T(P_{[a-x, (a+b)/2]} + P_{[(a+b)/2, b+x]})]$, which will be different from $p_{[a,b]}(T)$.

In the present context, given that the objectification requirement forces a measurement of a discretised observable, the simpler and more intuitive condition (1.4) is almost sufficient for the results we will discuss: the important proviso is that the following proofs apply also to the case when the measured object observable is an unsharp discrete observable (a POV measure as outlined in the introduction). More will be said about unsharp observables later on in the chapter, in section 5. For the time being it suffices to note that for this, more general case the probability reproducibility condition would assume the form

$$(1.5) \quad \text{Tr}[F_X T_S] = \text{Tr}[\mathbb{I} \otimes P_{[\psi_i]} U T U^{-1}]$$

where F is an effect associated by the POV measure to the set of possible values X .

The *second* point to note about the probability reproducibility condition is that, in cases such as that of unsharp or continuous object observables, it is best thought of as a two-part condition. The first condition captures the intuitive idea that if the initial state of the system to be measured yields non-zero probability for some eigenstate of the observable to be measured, the final state should yield non-zero probability for the corresponding pointer eigenstate in the object + apparatus space and viceversa. The second condition will determine the probabilistic relations between initial object state and final object + apparatus state. Formally the two conditions can be spelt out as follows:

5. Let T_S be an initial state of the object system, T_M an initial state of the apparatus system. Then

$$(1.6a) \quad p_{T_S}^E(X) \neq 0$$

if and only if

$$(1.6b) \quad p_{\mathcal{R}_A(U(T_S \otimes T_M)U^{-1})}^{P_A}(f^{-1}(X)) \neq 0.$$

6. If $p_{\mathcal{R}_A(U(T_S \otimes T_M)U^{-1})}^{P_A}(f^{-1}(X)) \neq 0$, then

$$(1.7) \quad p_{T_S}^E(X) = p_{\mathcal{R}_A(U(T_S \otimes T_M)U^{-1})}^{P_A}(f^{-1}(X))$$

It then becomes possible to hold condition 5, but change 6. Namely we can ask that ‘pointers point’ to the appropriate measurement results (so that all the appropriate measurement results appear), but relax the probabilistic condition so that the probabilities for the results in the final state might not be equal to those given in the initial state by the Born rule. It is worth noting that it is the probability reproducibility condition that does this work. In fact the objectification condition requires the final density operator to be decomposable in terms of pointer eigenstates, but gives no specification about which pointer eigenstates should be part of the decomposition: an initial object state which is an eigenstate of the measured observable might be evolved into a final density operator over several pointer eigenstates, and still it will satisfy the objectification condition, as it is by definition a density operator decomposable into pointer eigenstates, and that is all that objectification requires. In the case of a sharp and discrete observable, condition 5 implies condition 6 and they are equivalent to the probability reproducibility condition.

The *third* point is that the probability reproducibility condition is equivalent, in the case of a sharp and discrete observable, to the calibration condition (see Busch et al. [14, p. 34]):

- *Calibration*: for each initial state of the form $\varphi_i \otimes T_M$, W maps this state to a final state $\tilde{T}_S \otimes \psi_i$, where ψ_i is the pointer state for the i^{th} value of H and the $\tilde{\varphi}_i$ are arbitrary, and thus neither necessarily distinct nor orthogonal.

An initial state φ_i of the object system has probability 1 with respect to $P_{[\varphi_i]}$ and probability 0 with respect to $P_{[\varphi_j]}$ with $j \neq i$, so that the final state must yield probability 1 for the pointer observable eigenstate $\mathbb{I} \otimes P_{[\psi_i]}$ and zero for other pointer observable

eigenstates not measuring the value i . Therefore the final state must be of the form $\tilde{T}_S \otimes \psi_i$. Linearity then easily implies the probability reproducibility condition. Note that the calibration condition plays no direct role in the following proofs, save for a brief remark made in the context of the discussion of unsharp pointer observables; in fact most of the traditional results analyse the compatibility of 3 and 4.

In what follows we will sometimes make the distinction between these conditions explicit, and work with condition 5, which is a more general condition for the general case of POV measures. It will hardly be required to discuss POV measures explicitly for the object system. Nevertheless, proving insolubility results for 5 gives a genuinely more general proof.

Given the splitting of the probability reproducibility condition, there are at least two options open for generalising it in the context of trying to deal with the objectification condition. We can maintain 5, which guarantees that object observable eigenstates are correlated to pointer observable eigenstates, but relax the requirement that probabilities be exactly matched up in cases when the initial object state is a superposition of eigenstates of the measured observable. Alternatively we can follow Fine [35] and substitute a more general probabilistic condition than 4 which allows for violation of both 5 and 6. He asked whether, given an observable H to be measured on the object system, there existed a unitary measurement W which, rather than ‘reproducing’ the probabilities, might simply discriminate between two initial states T_S and T'_S so that if, for some i ,

$$p_{T_S}^E(X) \neq p_{T'_S}^E(X)$$

then

$$p_{\mathcal{R}_A(U(T_S \otimes T_M)U^{-1})}^{PA}(f^{-1}(X)) \neq p_{\mathcal{R}_A(U(T'_S \otimes T_M)U^{-1})}^{PA}(f^{-1}(X)).$$

In such a case we say that the respective initial states of the combined system, T and T' , are H -distinguishable. This does not help in trying to answer the highlighted question on page 44; the answer is simply ‘no’.

In the present framework a measurement has taken place when the apparatus system is in an eigenstate of a pointer observable defined over the apparatus system. In the next section, for the sake of simplicity, the pointer observable is taken to have no multiple eigenstates, and only one eigenstate indicates measurement of an eigenstate

of a (discretised) object observable. A remark at the end of the section will emphasise how this restriction is not essential to the argument.

The analysis of the next section will also make use of the following definitions:

DEFINITION 2.1. An induced notion of linearity for $\mathcal{T}_1^+(\mathcal{H})$ is defined by considering a linear mapping $U : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$ on the Banach space of trace class operators $\mathcal{T}(\mathcal{H})$; its restriction $U : \mathcal{T}_1^+(\mathcal{H}) \rightarrow \mathcal{T}_1^+(\mathcal{H})$ (when well-defined as a mapping from $\mathcal{T}_1^+(\mathcal{H})$ to $\mathcal{T}_1^+(\mathcal{H})$) satisfies the corresponding *convex-preserving* property, i.e. for all $a_1, \dots, a_n \in \mathbb{R}_+$, $a_1 + \dots + a_n = 1$ and P_1, \dots, P_n one-dimensional projections on \mathcal{H}

$$U(a_1 P_1 + \dots + a_n P_n) = a_1 U(P_1) + \dots + a_n U(P_n).$$

Such U is by definition called *convex*⁴.

DEFINITION 2.2. The *rank* of a density operator T defined for a Hilbert space \mathcal{H} (written $\text{Rk}(T)$) is the minimal number of one-dimensional projections of which it is a convex combination. This is an invariant for the density operator under convex transformations of the form $U^* T U$, with U unitary. The more formal definition of rank considers, for a given density operator T , the projection operator E^T on the range of T , such range being a subspace $\mathcal{H}' \subseteq \mathcal{H}$. Then the rank of T is equal to $\text{Tr}[E^T]$, the trace of the projection operator on the range of T ⁵.

2. A 'proof'

Geometric arguments can take us a long way towards answering the question defined in the previous section. By this I mean that most of the negative results in the measurement problem follow from simple considerations about the structure of the state

⁴For the case of $\dim(\mathcal{H}) = \infty$ we have that, for $\{a_1, \dots, a_n, \dots\}$ a (possibly infinite) sequence of positive real numbers such that $\lim_{n \rightarrow \infty} \sum_{i=1}^n a_i = 1$, and $\{P_1, \dots, P_n, \dots\}$ a (possibly infinite) family of one-dimensional projections on \mathcal{H}

$$\lim_{n \rightarrow \infty} U\left(\sum_{i=1}^n a_i P_i\right) = \lim_{n \rightarrow \infty} \sum_{i=1}^n a_i U(P_i).$$

⁵This is an immediate consequence of the discrete spectral representation of density operators as proved already in Von Neumann[82, Ch. 2, Sec. 11]. The rank is basis independent, as it is defined through a trace operation, and can be infinite in the case when E^T is a projection on an infinite dimensional subspace of \mathcal{H} . For a detailed analysis of the geometry of stratified convex sets of states, see [7]; Rockafellar [73] gives a detailed account of convex analysis .

space and the possible transformations on this structure given by unitary dynamics. This section does not give a general proof, but rather considers a number of (successively more general) options for answering the question positively and shows why these options cannot work. In the end this will lead to a result as general as any of the ones present in the literature, but it is hoped, in the first instance, that the problems encountered when trying to solve the quantum measurement problem by trying to find a mapping that will yield an appropriate final density operator will become clearer. It will successively become clear also that strategies employed in this section can be applied to other questions about insolubility proofs of the quantum measurement problem.

The approach taken involves assuming the objectification condition 3 and the first part of the probability reproducibility condition 5, then seeing whether the second part of the probability reproducibility condition 6 is satisfied.

The strategy of the argument used in this section is essentially a generalisation of an argument that Von Neumann has given in his book [82, pp. 437–439]; the mathematical structure is the same. It is also discussed by Scheibe [75, pp. 151–155]. Its generality is comparable to that of Busch and Shimony [17], in the sense that it can be shown to include the case of an unsharp object observable.

2.1. A Theorem by Davies and Some Initial Results. The following theorem by Davies [25, Theorem 3.1, p.21] is a good starting point:

THEOREM 2.1. *Consider a Hilbert space \mathcal{H} . Every pure positive linear map $\mathbf{T} : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$ is of one of the following three forms:*

$$(1) \quad \mathbf{T}(T) = BTB^*,$$

where $B : \mathcal{H} \rightarrow \mathcal{H}$ is bounded and linear;

$$(2) \quad \mathbf{T}(T) = BT^*B^*,$$

where $B : \mathcal{H} \rightarrow \mathcal{H}$ is bounded and conjugate linear;

$$(3) \quad \mathbf{T}(T) = \text{Tr}[TB] (|\psi\rangle\langle\psi|),$$

where $B : \mathcal{H} \rightarrow \mathcal{H}$ is bounded, linear and positive and $\psi \in \mathcal{H}$ ⁶.

⁶Note that ψ is constant.

Pure positive linear maps on $\mathcal{T}(\mathcal{H})$ are such that every element in the set of positive trace class operators is mapped into the same set, and furthermore *every pure state is mapped to a pure state*. Type 3 maps as characterised by the theorem are called *degenerate*: they map all states to a unique pure state, multiplied by a weight $\text{Tr}[TB]$. Therefore all such maps are pure. As for maps of Type 1 and 2, it is easy to see that all such maps are pure. In fact pure density operators are characterised by being one-dimensional projection operators, and therefore have rank 1. But it is a well-known result of linear algebra that $\text{Rk}(AB)$, with A and B two operators on \mathcal{H} , is equal to $\min[\text{Rk}(A), \text{Rk}(B)]$. It follows that whenever $\text{Rk}(A) = 1$, then $\text{Rk}(AB) = 1$. This shows that maps of Type 1 and 2 are always pure, so that with the Davies theorem it is possible to establish necessary and sufficient conditions for a map to be pure.

The standard dynamics for a state, given by a unitary operator W on $\mathcal{T}(\mathcal{H})$, can be represented as the conjugation action $WP_{[\varphi]}W^*$. It is well known and worth repeating that this is equivalent to the evolution given by $P_{[W(\varphi)]}$ ⁷, therefore establishing in a different way that this mapping is pure. The remark of the previous paragraph clearly establishes that the conjugation action of a unitary operator on the set of positive trace class one operators must be pure, that is it maps pure states to pure states. This immediately eliminates the possibility that any such conjugation might map an initial state of the form $P_{[\varphi]} \otimes P_{[\psi_0]} = P_{[\varphi \otimes \psi_0]}$ to a final state satisfying the objectification condition and condition 5: $P_{[\varphi \otimes \psi_0]}$ is a pure state, and therefore of rank one, while generally speaking the rank of the final state will have to be greater than one if objectification and 5 are to be satisfied.

Furthermore, Davies's Theorem states that *all* pure positive mappings have this form, so if any invertible mapping is to provide a solution to the measurement problem, given a pure initial object + apparatus state, it will not itself be a pure mapping. However the following result rules out this possibility.

THEOREM 2.2. *A convex mapping $U : \mathcal{T}_1^+(\mathcal{H}) \rightarrow \mathcal{T}_1^+(\mathcal{H})$ is bijective if and only if it is rank preserving for all $T \in \mathcal{T}_1^+(\mathcal{H})$.*

Any convex-preserving mapping can be a bijection (and therefore an invertible mapping) on the space of density operators if and only if the rank of the density matrices is a 'constant of the motion'. In particular any conjugation action of a unitary operator

⁷Von Neumann again shows this [82, Sec. V.1].

on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ will be rank preserving and it can be shown that such mappings are the only ones that are rank preserving, and thus bijective on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$. The proof of this result and the discussion of its other implications in the context of the quantum theory of measurement is left to the next chapter; for the time being it will be applied in a fairly basic way in the context of providing an insolubility proof.

This leads to the heart of the classical technical results on the measurement problem. We assume that, ‘physically’ speaking, there might be a degree of uncertainty as to what the initial state of the apparatus might be; as Wigner put it “the state vector of the apparatus, which is under the conditions now considered usually a macroscopic object, is hardly ever known”⁸ [87, p. 333]. This might be modelled by assuming that the initial state of the apparatus is a non-pure density operator; then the initial state of the combined system will not be a pure state, and perhaps it is possible to satisfy the objectification condition this way. The first consideration must be to decide what density operator is appropriate, and for which types of measurement.

Theorem (2.2) immediately rules out that a perfect correlation measurement might satisfy the objectification condition. A *perfect correlation measurement* is a measurement where the objectification requirement and condition 5 are satisfied and, moreover, the final state of the system is such that each final pointer state is correlated with a unique object state. In other words in a perfect correlation measurement, for a pointer state $P_{[\psi_i]}$, the component $T \otimes P_{[\psi_i]}$ is such that $T = P_{[\tilde{\varphi}_i]}$, where again note that $\tilde{\varphi}_i$ is an arbitrary vector in \mathcal{H}_S . Given that the final object + apparatus state admits of an ignorance interpretation, observing a specific pointer state allows inference of a pure final state for the system. The final state for the object + apparatus system would then be

$$W (P_{[\varphi]} \otimes T) W^{-1} = \sum_{i=1}^n w_i P_{[\tilde{\varphi}_i \otimes \psi_i]}$$

with T a mixed state for the apparatus system and $\sum w_i = 1$. Perfect correlation measurements such as the one just described are just a generalisation of so-called *strong*

⁸Wigner in fact makes this remark in one of his papers on measurement immediately after making the point of the previous paragraph. Ignore for the sake of argument the apparent contradiction in Wigner’s argument in that we are supposed to hardly ever know what the state of a macroscopic apparatus is, yet are supposed to learn about the value of an observable for a system precisely by discerning such a state.

state-correlation measurements. The latter measurements have been studied, for instance, by Busch and Lahti [15]. The former are defined by the property of the final object states associated with a given pointer state being pure states, with in addition the property of these final object states being pairwise orthogonal. While it is the case that strong state-correlation measurements have much more interesting informational properties than perfect correlation measurements, for the purpose of the present work the specifics of correlation properties are not important. The case of perfect correlation measurements is mainly illustrative of the kind of argument that will be presented in the rest of this section, as well as providing an insolubility result for a wide class of measurements which includes Busch and Lahti's strong state-correlation measurements.

The following condition is also normally assumed when characterising quantum measurements; as for instance Araki and Yanase have put it,

According to von Neumann, the measurement of the operator M in a state φ is accomplished by choosing an apparatus in a state ξ (fixed normalised state *independent of φ*) in \mathcal{H}_M ... (my italics) [2, p. 622]

meaning that such an initial state (which might of course be a density operator), must be chosen so as to not depend on what state we are trying to measure, as presumably we do not know what we are trying to measure before we have measured it.

In a perfect correlation measurement (again assuming objectification and 5) if the initial state of the object system is a φ superposition of two eigenstates of the observable A which is being measured, the final state will be a density operator over two eigenstates of the pointer observable M , each representing the pointer recording one of the two eigenvalues corresponding to the two eigenstates of A which resolve φ . This is because such a measurement should avoid interference between different pointer states in the final object + apparatus state after the evolution, while at the same time each pointer state should appear just once in the resolution of the final state. The final density operator will then have rank 2.

REMARK 2.1. Note that the observable $\mathbb{I}_S \otimes A_M$ has eigenvalues with degeneracies of the order of the dimension of the space \mathcal{H}_S . For example, if $\dim(\mathcal{H}_S) = \dim(\mathcal{H}_M) = 2$, suppose A_M has eigenstates $\{\psi_1, \psi_2\}$. Then $\mathbb{I}_S \otimes A_M$ has eigenvalues $\varphi \otimes \psi_i$ for $i = 1, 2$, where φ is any state in \mathcal{H}_S . Therefore it is

quite possible, for instance, that an eigenstate of the apparatus observable A_M might appear several times in the final state's resolution, coupled to different φ 's.

Similar reasoning though tells us that, if φ is a superposition of three such eigenstates, we should expect a final state with rank 3 for a perfect correlation measurement. But in both cases the rank of the final state is determined by the rank of the initial state, which is equivalent to the rank of the apparatus state chosen initially and independently of φ . If the initial state is chosen independently of any φ its rank is therefore constant. It follows that any measurement satisfying the objectification condition will not be a perfect correlation measurement, which requires the rank of the final state to vary with the number of eigenstates of the measured observable which resolve the initial state φ .

The argument excluding perfect correlation measurements is a simple, but instructive case. It can be generalised to the conclusion that the initial apparatus state should be a density operator of rank at least as high as the dimension of the Hilbert space \mathcal{H}_S characterising the object to be measured, at least if we want a proof that will cover all possible initial object states.

It may of course be the case, as Stein has claimed in his recent paper [79], that we don't want to prove insolubility just for the whole of a possible Hilbert space associated with a system on which a measurement is made, but also for certain given subspaces of it. The idea is that what doesn't work for the whole space might after all work if certain conditions are imposed on only a smaller part of it. That is, according to Stein, an insolubility proof for evolutions given by a certain unitary operator U , mapping all possible initial states of the object system \mathcal{H}_S , coupled to an apparatus initial state T_M , will show how certain conditions will fail given all these possible evolutions for all these different states. But we also need to show explicitly that the evolutions given by U for initial object states $\varphi \in \mathcal{H}_S'$ coupled to T_M , with \mathcal{H}_S' a proper subspace of \mathcal{H}_S , fail. Stein's idea seems to be that there might be some unitary U 's which run foul of insolubility proofs when applied to a whole Hilbert space, but might work if we only consider their action on a restricted part of the Hilbert space.

Here is a good physical example: suppose you want to measure the position of a particle which you know is somewhere along a line. In practice it is confined to a segment in front of you in the laboratory, and not outside. We can break up this segment,

within which we know the particle to be confined, into finitely many 'subsegments', and measure which one of them the particle is in. We would think it necessary to prove or disprove theorems about measurement for all of the (discrete) position observables on \mathcal{H}_S which represent the particle position along a line. According to Stein, we should make sure that, in proving such theorems, we don't rule out unitary evolutions which fall foul of an insolubility proof when we argue at the level of the observable on the whole space, but which might model perfectly good measurements when the measured observables are restricted to the ones corresponding to positions in my lab. Consider a finite set of eigenstates, which span a proper subspace $\mathcal{H}_S' \subseteq \mathcal{H}_S$, corresponding to the particle being somewhere in the laboratory. If we have a unitary operator that can model a measurement (in the sense of satisfying objectification and the probability reproducibility condition) for this finite set of eigenstates of a position observable, but fails when we consider all the eigenstates of the observable, then we should not rule this unitary operator out as a good model for measurements. It would account for one way of usually looking at measurements, namely that measurements are made in the strict confines of a laboratory. So insolubility proofs should hold not only for measurements on the whole space, but also for measurements on proper subspaces of these spaces.

The argument given in this chapter fulfils Stein's desideratum all the way: if one wants to provide a proof for a finite dimensional subspace of the Hilbert space associated with a particle whose position we want to measure, we can just think of the Hilbert subspace representing the particle's position as, nevertheless, a finite dimensional Hilbert space of possible position states, and make the same arguments, beginning from the claim that the initial density operator for the apparatus will need to have rank equal to the number of position eigenstates which we think it is realistic to measure.

In the case in which the measured observable has degeneracies, the rank of the initial density matrix for the apparatus space must be bounded from below by the cardinality of the set of eigenvalues; without loss of generality in the argument it is possible to make the assumption that there are no such multiplicities. In this case an equivalent claim is that the mixed state should have range of dimension at least equal to $\dim(\mathcal{H}_S)$ ⁹.

⁹More appropriately in the case of an infinite-dimensional Hilbert space the mixed state should have infinite-dimensional range. Note also that again results in the next chapter establish that if the rank of the density operator representing the initial apparatus state is greater than the dimension of the Hilbert

If the mixture has rank m strictly less than $\dim(\mathcal{H}_S) = n$, then a measurement of an initial state φ , a superposition of all the eigenstates of the measured operator, will yield a final state mixture of rank less than the number of eigenstates in the initial state φ . This cannot possibly be a final state of a measurement satisfying the objectification requirement and condition 5. Objectification and 5 would require that the final state be a density operator defined over at least n projectors, so that the pointer states in the final state do not display any interference between them: the pointer states must point to n different values, and if the final state is a density operator of rank m some of these different pointer states will necessarily have to be superposed.

The elimination of these possibilities leaves the case of an imperfect correlation measurement of an observable defined on an object system \mathcal{H}_S , with the initial state of the measuring apparatus system being a mixed state (over projectors on states forming a complete orthonormal system for the apparatus system \mathcal{H}_M of rank equal to $\dim(\mathcal{H}_S)$).

Before moving on to the general argument it is worth summarising what has been established in this section. We assume that the initial density operator for the apparatus state is given independently of the initial object state when measuring an observable H , also independent of the initial states. This initial apparatus density operator must, when coupled with initial object states, in all cases lead to a final state which satisfies the objectification requirement and condition 5 (at least). But this assumption is enough to rule out that there exist any *perfect correlation* measurements that do the trick: for example when the initial object state is a pure state the final rank is constant, while objectification, 5 and perfect correlation would demand a 'variable' rank for the final state.

This achieves two things. First, it rules out perfect correlation measurements as possible candidates for measurements satisfying objectification and 5 (because they would require the rank of the final density operator to 'vary' when it is in fact constant). Second, it requires the initial density matrix to have rank greater than or equal to the

space \mathcal{H}_S for the object system some very peculiar problems can arise, given the procedures we adopt for defining actual unitary premeasurements: it is possible to show that under certain circumstances no unitary premeasurement *at all* (that is, independently of whether objectification is satisfied) can be defined for an initial object-apparatus state where the apparatus density operator is of rank greater than the dimension of the Hilbert space for the measured system.

cardinality of the spectrum of the (discretised) measured object observable, so that measurements on all initial object states (including for example initial object pure states which are superpositions of all eigenstates associated with distinct eigenvalues of the measured observable) might have a chance of satisfying objectification and 5.

2.2. A first version of the general argument. We begin by considering first the case of $\dim(\mathcal{H}_S) = \dim(\mathcal{H}_M) = n < \infty$. Again without loss of generality assume that $\{\varphi_1, \dots, \varphi_n\}$ pick out an observable with a non-degenerate spectrum. We have that, for an initial pure state $P_{[\varphi]} = P_{[\alpha_1\varphi_1 + \dots + \alpha_n\varphi_n]}$ and an initial apparatus state $T_M = \sum_{i=1}^n w_i P_{[\psi'_i]}$, where the $P_{[\psi'_i]}$'s form a set of arbitrary orthogonal projections, the dynamics of a measurement process represented by a unitary operator U would yield

$$(2.1) \quad \begin{aligned} U \left(\sum_{i=1}^n w_i P_{[\varphi \otimes \psi'_i]} \right) U^{-1} &= \sum_{i=1}^n w_i U P_{[\varphi \otimes \psi'_i]} U^{-1} \\ &= \sum_{i=1}^n w_i P_{[U(\varphi \otimes \psi'_i)]} \end{aligned}$$

As the states ψ'_i are pairwise orthonormal, $\{\varphi \otimes \psi'_i\}_i$ is an orthonormal set in $\mathcal{H}_S \otimes \mathcal{H}_M$. Therefore $\{U(\varphi \otimes \psi'_i)\}_i$ is an orthonormal set, too, under the assumption of unitarity of U .

Suppose that, for $i \neq j$, $w_i \neq w_j$. If $\sum_{i=1}^n w_i P_{[U(\varphi \otimes \psi'_i)]}$ satisfies the objectification requirement of being a sum over eigenstates of the pointer operator $\mathbb{I} \otimes A_M$, then the states $U(\varphi \otimes \psi'_i)$ must be such eigenstates, that is $U(\varphi \otimes \psi'_i) = \tilde{\varphi}_i \otimes \psi_i$, with $\tilde{\varphi}_i$ arbitrary states. Recall from remark 2.1 that all states in $\mathcal{H}_S \otimes [\psi_i]$, for all ψ_i eigenstates of a pointer observable A_M , are eigenstates of $\mathbb{I} \otimes A_M$. Furthermore the eigenstates must be eigenstates of different pointer values (as we are measuring an initial superposition φ of n states) and therefore pairwise orthogonal. Under the assumption that the weights of the mixed state are different, such mixtures will admit of only one resolution in terms of orthogonal projections.

It then follows that, calculating the probabilities for $\mathbb{I} \otimes P_{[\psi_i]}$, we obtain always and only w_i 's as a result. Therefore the probabilistic information is determined solely through the weights of the initial apparatus mixture and completely independent of initial states such as φ , which are superpositions of n distinct eigenvalues of the observable to be measured. The probabilistic part 6 of the probability reproducibility condition will then be violated and we already have a proof of insolubility.

It is not difficult to see that a similar calculation will work for superpositions of less than n eigenstates. Discussion of an example will be helpful here.

EXAMPLE 2.1. Consider an initial state for the measured system $\varphi = \alpha_1\varphi_1 + \alpha_2\varphi_2$, $|\alpha_1|^2 + |\alpha_2|^2 = 1$, φ_1, φ_2 eigenstates of the measured observable. Also again consider an initial state for the apparatus system given by $T_M = \sum_{i=1}^n w_i P_{[\psi'_i]}$, where we assume the ψ'_i 's form an orthonormal set.

The initial state of the system is then

$$T = P_{[\varphi]} \otimes \sum_{i=1}^n w_i P_{[\psi'_i]} = \sum_{i=1}^n w_i P_{[\varphi \otimes \psi'_i]}.$$

A unitary evolution U will then yield the final state

$$\sum_{i=1}^n w_i P_{[U(\varphi \otimes \psi'_i)]},$$

where the $P_{[U(\varphi \otimes \psi'_i)]}$'s are orthogonal 1-dimensional projections. For objectification and condition 5 to be satisfied, this state must be equal to a state of the form

$$T_{\text{Final}} = \sum_{i=1}^k w'_i P_{[\bar{\varphi}_i \otimes \psi_1]} + \sum_{i=k+1}^n w'_i P_{[\bar{\varphi}_i \otimes \psi_2]},$$

with $1 \leq k < n$.

Without loss of generality we can assume that the states $\bar{\varphi}_i \otimes \psi_j$, for fixed j , form an orthonormal set. This is because

$$\begin{aligned} \sum_{i=1}^k w'_i P_{[\bar{\varphi}_i \otimes \psi_j]} &= \left(\sum_{i=1}^k w'_i P_{[\bar{\varphi}_i]} \right) \otimes P_{[\psi_j]} \\ &= T_k \otimes P_{[\psi_j]} \end{aligned}$$

and T_k must have a decomposition in terms of orthogonal projections. Furthermore, states $\bar{\varphi}_i \otimes \psi_1$ are orthogonal to states $\bar{\varphi}_j \otimes \psi_2$ for all i, j . Therefore T_{Final} can be written in terms of an orthonormal decomposition of factorised states of the form $\bar{\varphi}_i \otimes \psi_j$.

Under the assumption that $w_i \neq w_j$ in the initial apparatus state, $\sum_{i=1}^n w_i P_{[U(\varphi \otimes \psi'_i)]}$ is the unique decomposition of the final state in terms of orthogonal projections. It follows that $U(\varphi \otimes \psi'_i) = \bar{\varphi}_k \otimes \psi_j$ for one of k, j and that, more importantly, w'_i is equal to some w_i of the initial mixture.

Then calculating, for example, the probability for $\mathbb{I} \otimes P_{[\psi_1]}$ will yield $\sum_{i=1}^k w_i$, again dependent only on the probability weights of the initial mixture and not equal to $|\alpha_1|^2$. Again therefore the probabilistic condition 6 is violated.

The argument clearly generalises to all cases where the number of eigenvalues making up the superposition of the measured state is less than the rank of the initial density matrix for the apparatus system. It then follows that, given an initial pure state of the measured system, the probabilities of the final apparatus state depend only on the weights of the initial apparatus state.

A similar result holds for the case when the initial measured system state is a density operator. Again an example will clarify what happens.

EXAMPLE 2.2. Consider the case of a measurement on a 2-dimensional Hilbert space of an observable whose eigenstates are φ_1 and φ_2 . Let the initial state of the measured system be

$$T_S = v_1 P_{[\varphi^1]} + v_2 P_{[\varphi^2]}$$

where $\varphi^1 = \alpha_1 \varphi_1 + \alpha_2 \varphi_2$ and $\varphi^2 = \bar{\alpha}_2 \varphi_1 - \bar{\alpha}_1 \varphi_2$, $\alpha_i \in \mathbb{C}$ (where, for $\alpha \in \mathbb{C}$, $\bar{\alpha}$ is the complex conjugate of α). In other words we write the initial state of the measured system in its resolution in terms of orthogonal projections. We then have that the probability for φ_1 , given that the initial state is T_S , will be $v_1 |\alpha_1|^2 + v_2 |\alpha_2|^2$.

Now we consider an initial state

$$T_M = w_1 P_{[\psi_{01}]} + w_2 P_{[\psi_{02}]}$$

for the apparatus, and an observable A on the apparatus space with eigenstates ψ_1 and ψ_2 representing pointer positions for φ_1 and φ_2 respectively.

The initial object + apparatus state will then be

$$T = \sum_{i,j=1}^2 v_i w_j P_{[\varphi^i \otimes \psi_{0j}]},$$

where the $P_{[\varphi^i \otimes \psi_{0j}]}$ are pairwise orthogonal projections on the composite system.

The final state

$$T_{\text{Final}} = \sum_{i,j=1}^2 v_i w_j P_{[U(\varphi^i \otimes \psi_{0j})]}$$

is then resolved in terms of a set of pairwise orthogonal projections $P_{[U(\varphi^i \otimes \psi_{0j})]}$. Exactly the same argument employed from the previous example will show that these must be factorised states of the form $P_{[\tilde{\varphi}^i \otimes \psi_j]}$. Then the probability for $\mathbb{I} \otimes \psi_j$ will be equal to a combination of weights $v_i w_j$. At this point more could be said about how the weights combine. For the purpose of this example, however, it suffices to note that, in any case, such probabilities are independent of $|\alpha_1|^2$ and $|\alpha_2|^2$, the coefficients which appear in

the probabilistic calculations for the initial state T_S . This again contradicts condition 6 and yields an insolubility proof.

2.3. The argument for initial apparatus states with multiplicities in their weights. The situation is not much different if multiplicities in the weights of the initial apparatus density operator are allowed for. It is important, nevertheless, to show explicitly why this is the case.

Consider again the case of an initial pure object state $\varphi \in \mathcal{H}_S$, where $\dim(\mathcal{H}_S) = n$ and φ is a superposition of all eigenstates of the measured observable. Suppose that the initial apparatus state T_M has a set of equal weights w_i , with multiplicity m . When some w_i 's are equal, the set $\{U(\varphi \otimes \psi'_i)\}$ may well contain elements which are not eigenvalues of the operator $\mathbb{I} \otimes A$. The reason is that neither the initial state, nor the final state, have a unique resolution in terms of pairwise orthogonal projections. Therefore, even if the initial state is resolved in terms of such projections, there is no guarantee that the final state will be resolved in terms of the orthogonal projections corresponding to the apparatus eigenstates. This then raises the possibility that the probabilities will not depend on the weights of the initial density matrix, whereas they do in the case when no multiplicities are present.

The final state is itself a positive, self-adjoint operator on $\mathcal{H}_S \otimes \mathcal{H}_M$. As is well known, it must contain the appropriate apparatus observable eigenvectors in its range in order to be expressible as a convex combination of them¹⁰ and thus satisfy objectification. Furthermore the apparatus observable eigenvectors will form an orthonormal basis for the range of the final state. Then the objectification requirement and condition 5 can be satisfied. Two possibilities have to be considered here.

Any pure states in a resolution of the final state in terms of orthogonal projections whose weights do not occur more than once in the final state are fixed by the objectification condition, which demands that the final density operator admit of an orthogonal representation in terms of eigenstates of a pointer observable $\mathbb{I} \otimes A_M$. The orthogonal representation of the density operator is given by a set of eigenstates of the operator, the weights being its eigenvalues. For eigenvalues (weights) with multiplicity 1 there is no possibility of eliminating their corresponding eigenvectors from the orthogonal representation. These eigenvectors appearing in the final state of the object + apparatus

¹⁰It is a general feature of density operators that any projection onto a state in the range of these operators can appear in a decomposition of the density operators.

system must then be eigenvectors of the pointer observable, as we assume that the final state satisfies objectification and 5.

The problem case therefore remains when eigenvalues (weights) of the final (and initial) density operator have multiplicity greater than 1. Denote by ψ_i'' the states of the initial apparatus density operator for which the weights of the representation are repeated m times. The final states $\{U(\varphi \otimes \psi_i'')\}$ for $1 \leq i \leq m$ must span m eigenstates of $\mathbb{I} \otimes A$ in order for the final state to have a representation in terms of eigenstates of the pointer observable, satisfying in this way the objectification requirement and condition 5. Rewriting the final state in this way will not, however, affect the weights associated with the projections. In this case again calculation of the probabilities will yield always and only the weights w_i . The convex-preserving property implies that the weights, as determined by the initial density operator for the apparatus system, will themselves be 'constants of the motion' under any possible action of the conjugation operation as can be seen in (2.1). This is true regardless of whether the weights exhibit multiplicities or not, and the result previously established still holds.

REMARK 2.2. *Assuming that the pointer observable has multiple eigenstates associated with a given eigenvalue will not affect the argument given in this section. Suppose that the apparatus eigenspace \mathcal{H}_M is such that $\dim(\mathcal{H}_M) > \dim(\mathcal{H}_S)$, and suppose that for some eigenvalue a of the pointer observable A_M , indicating measurement of an eigenvalue of the object observable being measured, the pointer eigenvalue is degenerate. Assume that the eigenvalues associated with a are spanned by an orthonormal set $\{\psi_{1a}, \dots, \psi_{na}\}$. Then the appropriate projector to test for the probability associated with the pointer value a of the observable $\mathbb{I} \otimes A$ will be $\mathbb{I} \otimes P_{\{\psi_{1a}, \dots, \psi_{na}\}}$, where $P_{\{\psi_{1a}, \dots, \psi_{na}\}}$ is the projection on the subspace of \mathcal{H}_M spanned by the orthonormal system $\{\psi_{1a}, \dots, \psi_{na}\}$.*

It is easy to check that the probability condition will fail in exactly the same way if this is the case. Objectification requires that, if the initial state yields probability $|\alpha_a|^2$ for the eigenstate φ_a of the measured observable which is associated with the pointers ψ_a , then the final density operator $U(T)$ of the object + apparatus space must admit of a decomposition into one or more of the projections $P \otimes P_{[\psi_a]}$. The pure final states $\mathbb{I} \otimes P_{[\psi_a]}$ will be orthonormal to the projections associated to all other pointer eigenvalues. They will, therefore, have

been unitarily evolved from pure states spanned by a subset of the initial apparatus states orthonormal to the other pure states decomposing the initial apparatus mixture. Suppose the weights associated with such initial orthonormal states are w_{ia} ; their sum $\sum_i w_{ia}$ will then clearly be equal to $\text{Tr}[U(T)\mathbb{I} \otimes P_{[\psi_{1a}, \dots, \psi_{na}]}]$, irrespective of the initial object state.

2.4. Some conclusions. This section establishes that one cannot construct an operator satisfying conditions 3 and 4, by deriving a contradiction for assumptions 3 and 4, given the further assumption of unitarity of the measurement operator.

While the proof is carried out for the finite dimensional case only, it is clear that much the same reasoning can be applied to the infinite dimensional case and the same result (independence of probabilities) applies. This is because of the form of the convexity assumption for infinite dimensional Hilbert spaces, namely

$$\begin{aligned} U \left(\lim_{n \rightarrow \infty} \sum_{i=1}^n w_i P_{[\varphi \otimes \psi'_i]} \right) U^{-1} &= \lim_{n \rightarrow \infty} \sum_{i=1}^n w_i U \left(P_{[\varphi \otimes \psi'_i]} \right) U^{-1} \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n w_i P_{[U(\varphi \otimes \psi'_i)]} \end{aligned}$$

Again the states $P_{[U(\varphi \otimes \psi'_i)]}$ must be pointer states (or a group of them must span pointer states in the case of multiplicities in the weights, even infinite multiplicities) for objectification to be satisfied. Thus the probabilities will again be equal to (possibly infinite) sums of weights w_i .¹¹

The argument of this section seeks to give an intuitive idea of what can happen when we assume that objectification of the pointer observable holds. It can also easily be made rigorous for cases where multiplicities and degeneracies are present, following

¹¹Note in particular that unitary transformations on infinite dimensional Hilbert spaces must map the whole space onto itself as well as being isometric. This means that transformations which for instance map an orthonormal basis of the Hilbert space to a countably infinite subset of that basis, even though they preserve the inner product and are thus isometric, are not unitary. The argument however works both for unitary and isometric transformations. In fact the condition that the projection operators $P_{[\psi_i]}$, the set of all relevant pointer observables, span the whole of the infinite dimensional Hilbert space \mathcal{H}_M is not a necessary condition for the argument just given. Therefore the eigenspaces $\mathcal{H}_S \otimes [\psi_i]$ of the pointer observable need not add up, for all non-zero pointer values indexed by i , to the whole space $\mathcal{H}_S \otimes \mathcal{H}_M$, but will add up to $\mathcal{H}_S \otimes \{\psi_i\}_i$, a proper subspace of $\mathcal{H}_S \otimes \mathcal{H}_M$. In such case the measurement operator might well be an isometry from $\mathcal{H}_S \otimes \mathcal{H}_M$ to $\mathcal{H}_S \otimes \{\psi_i\}_i$.

the points made, for example, in subsection 2.3 and in Remark 2.2. I avoid this, though, for the discussion of this section will be generalised by the result given in section 4.

The present discussion also makes use of condition 5, which is independent of objectification and enables a more intuitive discussion of what is happening in terms of the dynamics: it requires not just that pointers point, but that they point to the right values. This is not, however, necessary for the results of this section to work. Assuming just objectification means that the final state (2.1), that is

$$\sum_{i=1}^n w_i P_{[U(\varphi \otimes \psi'_i)]},$$

must be decomposed into eigenstates of $\mathbb{I} \otimes A_M$. Such eigenstates, however, need not point to the right result. Nevertheless, calculating probabilities for $\mathbb{I} \otimes P_{[\psi_i]}$ for a pointer state $P_{[\psi_i]}$, the same line of reasoning used in this section will yield a value of 0 or a sum of weights w_i yet again, once more independent of the probability distribution associated with the initial object state $P_{[\varphi]}$. Therefore the present strategy also yields a contradiction with Fine's distinguishability condition.

The difference will be that, without assuming condition 5, the rank of the initial density operator for the apparatus space will not be constrained. This difference, however, straightforwardly leads to insolubility results if the rank is too small, with arguments similar to the one given for the case of perfect correlation measurements.

This section presents a result which is close to those of a recent paper by Stein [79]. Both reach the same conclusion, namely that the probabilistic information yielded by a measurement satisfying objectification will in the end depend on the initial apparatus state only: no transfer of probabilistic information can happen between systems \mathcal{H}_S and \mathcal{H}_M . The probability reproducibility condition asks that the probabilistic information of an initial state $\varphi \in \mathcal{H}_S$ with respect to a measured observable on \mathcal{H}_S be transferred to information of the final state of the apparatus with respect to a pointer observable: this cannot be done. Stein assumes that the measurement operator U be unitary and satisfy objectification. This latter condition is expressed by the requirement that, given an initial state $T_S \otimes T_M$, the final state of the measurement evolution, $U(T_S \otimes T_M)U^{-1}$ should commute with the pointer observable operator on $\mathcal{H}_S \otimes \mathcal{H}_M$, namely $\mathbb{I} \otimes A$. Rather than explicitly calculating the probabilities as is done in this chapter, he proves, using the commutativity condition, a general result which entails the conclusion about the probabilities. Both the present proof and Stein's effectively show that the probabilistic

requirement 6 is violated. Busch and Shimony [17, p. 401] explicitly construct an operator which is unitary and satisfies objectification, but for which the probability reproducibility condition fails in exactly the way shown by Stein and in this paper. This example will be discussed at some length in section 4.

Stein also motivates his proof by reference to a debate with Fine and Brown, which will be further discussed in the next section. Fine and Brown's proof considers the possible evolutions of particular, preferred decompositions of the initial object + apparatus state; this actually leads to an obvious impossibility proof for the quantum measurement problem according to Stein. I think Stein has in mind something like the proof presented in subsection 2.2; the 'obvious' is due to the fact that such a proof would simply consider unitary evolutions of the initial object + apparatus pure states in the preferred decomposition.

On the other hand, this does not rule out the possibility that for some decomposition other than the preferred one the problem might be solved. In particular, if we have a density operator which is not a proper mixture, or if we reject the possibility of drawing the distinction between proper and improper mixtures (as, for example, Krips [53] does) we are not in any position to select a preferred decomposition. One way to read the results of this section is that they establish that such possibilities are not really present: objectification and unitarity constrain them, as is shown by considering the evolution of the initial orthogonal decomposition of the object + apparatus state, and showing that multiplicities in this state do not make a difference.

3. Insolubility proofs and Real Unitary Evolution

This section discusses the insolubility proof given by Fine [36] and Brown [11]. The previous section has shown how, assuming objectification and condition 5 (indeed assuming objectification alone) makes it impossible for the probabilistic condition 6 to be satisfied. Fine's proof relied implicitly on a further assumption, Real Unitary Evolution (RUE), in order to derive the impossibility result. Brown's paper makes this assumption explicit and defends it as a reasonable assumption for quantum measurements.

Stein (in the paper [79] mentioned at the end of the previous section) and Shimony [77] have criticised this assumption as untenable on two different grounds. By relying on the analysis given in the previous section it is possible to show that the technical

claim that RUE makes is in fact not independent of the assumption of objectification, but follows from it under the interpretation of mixtures that Stein in particular thinks is appropriate.

This result relies on an apparently stronger version of the condition of objectification. In showing that the condition is not really stronger, I will prove another result which establishes that, for unitary evolutions of the object + apparatus system, the requirement of objectification and condition 5 are incompatible. The result is of particular interest as it involves no consideration whatsoever of conditions such as 6, Fine's distinguishability condition, or indeed any other probabilistic condition that connects probabilities for objects to probabilities for apparatus. It is therefore more general than any of the proofs given before, which assume such detailed probabilistic conditions for quantum measurement. This will be discussed in the next section.

3.1. Real Unitary Evolution and its critics. The principle of Real Unitary Evolution (RUE) assumes that the initial state of the object + apparatus system, call it $T = \sum w_i P_{[\Phi_i]}$, $\Phi_i = \varphi \otimes \psi_{0i} \in \mathcal{H}_S \otimes \mathcal{H}_M$ is a (proper) mixture of the pure states $P_{[\Phi_i]}$, and asserts that the final state is actually a mixture over the states $P_{[U(\Phi_i)]}$, where U is a candidate measurement operator. This implies that, if the final object + apparatus state $\sum_i w_i P_{[U(\Phi_i)]}$ is to satisfy objectification, the states $P_{[U(\Phi_i)]}$ must be the eigenstates of the pointer observable $\mathbb{I} \otimes A_M$, for they are the only states which can appear in the decomposition of the final mixture, given that it is a proper mixture.

Recall that a mixture is defined as a density operator which is a genuine mixed state, i.e. a density operator which either represents an ensemble of distinct systems truly in pure states, or represents a system which admits of an ignorance interpretation, so that the weights of the density operator are in fact classical probabilities, and the system is really in a pure state, we just don't know which one.

Unpacking the definition, we can see that it entails a number of things. The first is that the initial apparatus state is equally a proper mixture: this is because if the initial object state is pure, the combined state will be a proper mixture only if the apparatus state is a proper mixture, too. As the apparatus state is selected independently of the initial object state, it must always be a proper mixture. So, secondly, both for the combined state and for the apparatus state RUE selects a preferred decomposition,

whatever decomposition is forced upon the states by the fact that they are proper mixtures. The final apparatus eigenstates, if they emerge, must be the result of evolutions of the pure states appearing in the preferred decomposition of the initial combined state. These evolutions are called Real Unitary Evolutions because they are evolutions of clearly defined pure states making up a mixture.

The assumption that the initial state is a proper mixture is needed, according to Brown, because if the final density operator of the combined object-apparatus system is to allow for an ignorance interpretation we must be able to rule out all other possible resolutions of it (orthogonal or not, presumably), and this is straightforward if the initial state is a mixture, for then the final state will be one, too. Fine's result relies on considerations of states precisely of the form $P_{[U(\varphi_n)]}$; the original weakness of the result was in the lack of justification for why we should think of states like $P_{[U(\varphi_n)]}$ as the only plausible candidates for being eigenstates of the pointer observable. Brown's own justification of the reliance on states such as $P_{[U(\varphi_n)]}$, by assuming RUE, rescues the proof.

There are several problems with the RUE principle. First of all, why should we assume that the initial state must be a mixture in order to have a genuine measurement? There is no particularly good reason for this. A problematic aspect of this assumption concerns for example how we are to read the evolution: if the system is a proper mixture, it might be argued that it is as if it evolves as an ensemble of pure states, 'sectioned off' from one another according to whether they have the same pure state or not. Then there is a problem, because perhaps in this case we should characterise the measurement evolution as many different measurement evolutions for different individual pure states and treat the measurement problem on this basis. But it is by now well understood, for instance as a result of the discussion on page 50, that no satisfactory solution to the measurement problem in the present framework can be found if the evolution is defined over initial pure states.

There are two other reasons for dropping RUE, one to do with the generality of proofs that one wants to achieve and one concerning the interpretation of density operators in general. It seems unreasonable to disallow density operators which are not proper mixtures from being initial apparatus states, particularly if this might solve the measurement problem. This requires only objectification for the final apparatus system, which in turn requires only that the final object + apparatus system be an

appropriate density operator, and not that such density operator be a proper mixture. Initial apparatus density operators which are not proper mixtures might arise, for instance, by a decoupling of an apparatus from an environment if the apparatus + environment system happens to be in an entangled state. Several reasons suggest that this will in general be the case: the Coulomb interaction is sufficient to guarantee that everything will become entangled, even when it is effectively shielded. Shimony's own statement is that "... the energy levels of macroscopic systems are so densely spaced that no shielding can prevent the entanglement of [the apparatus system] with the environment" [77].

Stein's argument against assuming a principle such as RUE is based upon rejecting the possibility of interpreting the density operator as an assignment of probabilities to possible pure states. As he puts it, "... the full content of the statistical operator lies in its assignment of probabilities to the possible *values of observables*, and [...] it has nothing whatever to do with probabilities that the system is *in* a given pure state" [79]. This is, in several different senses, a point well made; many recent approaches to quantum mechanics stress the fact that states are derived from the demand that they define a probability measure for the sets of observables (this is how Gleason, for example, characterises states). From this point of view pure and mixed states have the same status: they both define probability measures, albeit different ones. But also we assign states of incoherent polarisation (non pure density operators) to photons without necessarily intending to interpret this assignment as stating that the photons are in one as opposed to another state of coherent polarisation. The algebraic approach to quantum mechanics also defines states in this way.

Stein's position seems to me to be the correct one on the issue of interpreting density operators, at least for the present discussion. If we assume that a decomposition of a density operator is to be privileged at all, this is in a sense equivalent to saying that quantum mechanics is incomplete. This is because nowhere does quantum mechanics tell us how to distinguish such a decomposition from any other (Van Fraassen makes a similar point when discussing the ignorance interpretation [80, p. 206]). In particular there are *no* observables which can draw this distinction, in the sense that no observables lead to different expectation values for different decompositions of a density operator. Therefore no physical setup *according to the quantum theory on its own*, no experiment, can draw the distinction: this is the essence, from a theoretical

point of view, of what differentiates quantum mechanics from ‘classical’ theories. Van Fraassen in fact explicitly speaks of quantum mechanics being incomplete insofar as we think we can tell which decomposition of a density operator is the correct one.

Henry Krips similarly rejects the idea that we can interpret density operators epistemically, as he puts it. He flatly rejects that it makes sense to talk about proper mixtures from a different point of view, both for technical and mainly philosophical reasons: if it makes sense to talk about proper mixtures in quantum mechanics, then in truth this so-called mixture is really a pure state, and Krips gives a number of technical and philosophical reasons for treating it as such [53, Ch. 4].

Nevertheless I think that Fine and Brown’s proof works without assuming RUE, and I will try to show this in what follows. What is it that Fine and Brown actually do? The proof works by considering three initial states φ_1 and φ_2 , eigenstates of an observable H , and $\varphi = \alpha_1\varphi_1 + \alpha_2\varphi_2$ for the object system, as well as an initial density operator for the apparatus system, $\sum_i w_i P_{[\gamma_i]}$ (the γ_i ’s not necessarily eigenstates of the apparatus observable). The proof then examines the three end states after the evolution, namely $\sum_i w_i P_{[U(\varphi_1 \otimes \gamma_i)]}$ and the similar states for φ_2 and φ . These must all be weighted sums of eigenstates of $\mathbb{I} \otimes A$ if they are to satisfy the objectification condition. Now let $\Phi_1 = \varphi_1 \otimes \gamma_n$ for some fixed n , $\Phi_2 = \varphi_2 \otimes \gamma_n$ and $\Phi = (\alpha_1\varphi_1 + \alpha_2\varphi_2) \otimes \gamma_n$. Then we have that $U(\Phi) = \alpha_1 U(\Phi_1) + \alpha_2 U(\Phi_2)$. But $U(\Phi)$, $U(\Phi_1)$ and $U(\Phi_2)$ are all supposed to be eigenstates of $\mathbb{I} \otimes A$ and for this to be true it must be that these eigenstates have the same eigenvalue. If this is so then Fine’s distinguishability criterion (as well as the probability reproducibility condition) will fail: φ_1 , φ_2 and φ are H -distinguishable, while the final states are not $\mathbb{I} \otimes A$ -distinguishable.

A natural question arises here: can this proof work without explicitly assuming, as RUE does, that, for an initial density operator $\sum_i w_i P_{[\varphi \otimes \psi_{0i}]}$ for the object + apparatus system, the final state, which by objectification must be a decomposition in terms of projections on eigenstates of pointer observables, will actually be so decomposed by projections of the form $P_{[U(\varphi \otimes \psi_{0i})]}$? Can the proof work without assuming that the ignorance interpretation of mixtures selects a preferred decomposition of the initial state of the combined system, as RUE suggests? The aim is to attempt a different ‘rescue’ operation of the proof just outlined in the previous paragraph from the one that Brown suggests. The motivations are also different from Brown’s. He sees a virtue in the RUE condition itself, as well as in saving the proof in view of its simplicity: the

addition of the RUE condition, while philosophically debatable, certainly does not add to the difficulty of the mathematical argument we have just seen. My rescue of this argument has a somewhat different motive. I want to show that it is on a par as a result with other insolubility proofs, so that it is not necessary to invoke *any* additional conditions in order to show that Fine and Brown's proof is valid. This will hardly keep matters simple; on the other hand it shows that the insolubility result that Fine and Brown propose is in no way less general than other results, *irrespective* of the attitudes one has towards interpreting quantum density operators.

3.2. Saving Fine and Brown's proof? In order to answer the question on which the previous section has closed, I have to show that we can dispense with RUE, assuming only the standard conditions invoked when proving insolubility results, while keeping the proof just outlined as a valid one. The crucial step in doing this is to show that the technical part of the RUE assumption is in fact derivable from the other conditions usually assumed when trying to solve the measurement problem.

This technical content amounts to the claim that there exists a preferred resolution of the initial apparatus state, independent of the object state. This resolution must in addition have the following property: the pure states decomposing the initial apparatus state, when coupled to the initial object state, are *necessarily* mapped to pointer eigenstates. This will establish that there is a way of writing down the initial object + apparatus state, independent of the initial object state, for example as $\sum_i w_i P_{[\varphi \otimes \psi_{0i}]}$, such that the final state $\sum_i w_i P_{[U(\varphi \otimes \psi_{0i})]}$, after the evolution U has taken place, is decomposed in terms of pointer eigenstates of precisely the form $P_{[U(\varphi \otimes \psi_{0i})]}$. The independence of the ψ_{0i} 's from the initial object state φ is again needed for the proof on page 67 to go through, as can be seen by inspecting the role of the γ_n 's in that proof.

If I can show that this technical content follows from the usual conditions in the quantum measurement problem, RUE will then become redundant. If this is established it *will* be possible to appeal to that resolution and give as a proof the argument of Fine and Brown presented on page 67, independently of the philosophical reasonableness of RUE and regardless of the approach we want to take in the interpretation of quantum density operators.

It is worth stressing that what I am considering here is the objection to the validity of the proof given on page 67 *per se*, and not the objections to RUE. If it were possible to show that the result of Brown and Fine follows without assuming RUE, then debates over the status of RUE and debates over the status of the proof will become separate issues: it might be thought that RUE is a tenable interpretation for quantum measurement evolutions, or it might not, but this would not affect the status of Fine and Brown's proof.

The final state of the object + apparatus system should be a density operator which admits of a decomposition over eigenstates of a pointer observable $\mathbb{I} \otimes A$ for the composite system in order for objectification to be satisfied. A subset of these eigenstates (bearing in mind remark 2.1) forms an orthonormal basis for the composite system. The initial density operator can also be decomposed over a set of orthonormal states, $P_{[\Phi_i]}$ say. As any unitary operator maps orthonormal systems to orthonormal systems, and a density operator has a unique resolution in terms of orthonormal states, assuming there are no multiplicities in its weights, then the final state must have its orthonormal resolution given by $P_{[U(\Phi_i)]}$, given the Φ_i are orthonormal states. If the decomposition of the final state in terms of pointer eigenstates is orthonormal, then the pointer eigenstates must be of the form $U(\Phi_i)$.

Suppose the unitary operator measures an observable with eigenstates $\varphi_1, \varphi_2, \dots$ and that the initial object state is a pure state which is a superposition of m such eigenstates. Then the final state must be a density operator decomposable into pointer observable eigenstates having m different pointer eigenstates of the observable A_M defined on the apparatus space \mathcal{H}_M . If two pointer eigenstates have different eigenstates of A_M for their \mathcal{H}_M component, they are orthogonal. Some, on the other hand, will have the same \mathcal{H}_M component. Consider now two different ones, $P_1 = P_{[\tilde{\varphi} \otimes \psi_i]}$ and $P_2 = P_{[\tilde{\varphi}' \otimes \psi_i]}$, i fixed. Suppose they are not orthogonal, therefore not of the form $U(\Phi_i)$; they do span a subspace of dimension 2, however, and any density operator which is decomposable into P_1 and P_2 is decomposable into any two projections on the image space of $P_1 + P_2$. In particular, given the assumption of no multiplicities in the initial density operator, there must be exactly *one* orthogonal decomposition of the subspace; this will, as is easily seen by looking at P_1 and P_2 , necessarily contain ψ_i as its \mathcal{H}_M component.

It follows that the projections which decompose the final object + apparatus state after a unitary measurement, in accordance with the objectification condition, are such that

either: only one contains a particular ψ_j as its \mathcal{H}_S component, in which case it will be a projection orthogonal to all other projections and therefore necessarily of the form $P_{[U(\Phi_i)]}$,

or: more than one contains ψ_j , in which case there exists a decomposition of the particular 'subpart' of the density operator in terms of orthogonal projections, both containing ψ_i , and both orthogonal to all other projections in the final state, in which case they will also be of the form $P_{[U(\Phi_i)]}$ ¹².

¹² It is easy to show that this is the case. First of all, note that the decomposition of a (multiplicity free) density operator T in terms of pointer eigenstates might not be made up of pairwise orthogonal pure states of a density operator, but will consist at the very least of a set of states, grouped into subsets which span orthogonal subspaces of the Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_M$, where the subsets are characterised by the fact that the eigenstates belonging to them share an eigenvalue. Clearly the elements of the subsets need not be pairwise orthogonal, but they are orthogonal to all other eigenstates in the decomposition.

Given one such subset S and its complement S' in the set of eigenstates of the given decomposition of the density operator T , the latter can be written as $T_S + T_{S'}$, where T_S is the operator obtained by summing together the elements of the subset S , with the appropriate weights, and $T_{S'}$ is simply $T - T_S$. T_S maps all vectors in a Hilbert space such as $\mathcal{H}_S \otimes \mathcal{H}_M$ to vectors spanned by S , and is null on vectors spanned by S' . $T_{S'}$ clearly does just the opposite. The claim just made is that the *orthogonal* decomposition of the density operator will then of necessity be given by vectors spanned by S and S' separately, and cannot belong to the span of the union of these two sets.

The orthogonal decomposition of T is given by its eigenvectors multiplied by its eigenvalues. Namely, for any φ such that P_φ appears in the orthogonal decomposition of T , we have that $T(\varphi) = \alpha\varphi$. Now suppose that φ is in the span of T , and that $\varphi = \alpha_1\varphi_1 + \alpha_2\varphi_2$, where φ_1 is in the span of T_S and φ_2 is in the span of $T_{S'}$. Then, if φ is an eigenstate of T , we must have that

$$T(\varphi) = T_S + T_{S'}(\alpha_1\varphi_1 + \alpha_2\varphi_2) = \alpha_1T_S(\varphi_1) + \alpha_2T_{S'}(\varphi_2)$$

For this to be equal to $\alpha\varphi$ both φ_1 and φ_2 must be eigenstates of T_S and $T_{S'}$ respectively, and therefore also of T , and both must have eigenvalue α . But this would imply that T has a multiple eigenvalue, which gives a contradiction. Therefore the eigenstates must belong to S and S' . The argument can clearly be repeated for each of the subsets which partition the decomposing pointer eigenstates, therefore making sure that all elements in the orthogonal decomposition of a multiplicity-free final density operator T must be pointer eigenstates. Note that this applies *regardless* of whether the pointer observable has multiplicities or not: if the pointer observable has multiple eigenstates associated with a pointer eigenvalue, it might be the case that in the final density operator after measurement there

So, regardless of whether we assume a condition such as RUE, it is a result that, in the absence of multiplicities in the initial apparatus state, the technical requirement of RUE (that the final state has the form $P_{[U(\Phi)]}$ with $U(\Phi)$ an eigenstate of the pointer observable signifying that a measurement has unearthed a certain value) holds. The technical content of Shimony and Stein's objection to the proof is not so easily defeated, though: the $U(\Phi_n)$ need not be eigenstates of the apparatus observable if there are multiplicities in the initial apparatus state, until otherwise proved.

As the possibility of writing the final states as $U(\Phi_n)$, where these states are pointer eigenstates, is crucial for Brown's proof to work, the disagreement voiced by Stein and Shimony still looks strong. Is it tenable? No, because, as I will show next, it is always possible to find a unique way of writing the initial density matrix, independently of what the initial object state is and independently of whether the initial state contains multiple weights, so that the final object + apparatus state is a sum of eigenstates of the apparatus pointer observable.

It looks as though, in the argument leading to proposition 2.3, it is not possible to get the eigenstates right: the procedure for finding the preferred way for writing the initial density operator will make use of a violation of condition 5. It is useful to explain this in some detail, as it will also lead to a very simple insolubility proof involving a contradiction between the objectification requirement 3 and condition 5, given unitarity of the measurement, rather than between 3 and 4, and eventually to a characterisation of unitary operators satisfying objectification. This makes it doubly worthwhile to study the question.

The following argument is close in structure to the one given in the previous section, specifically in subsection 2.3, when dealing with the question of how the result in section 2 is affected by the case when the initial apparatus state contains multiplicities. Suppose $1 \leq i \leq m \leq n$, where n is the rank of the density operator for the initial object + apparatus system and m is an index which tracks the multiplicity of one of the weights of the initial density operator. The states $P_{[\Phi_i]}$, $\Phi_i \in \mathcal{H}_S \otimes \mathcal{H}_M$, $1 \leq i \leq m$, which appear in the decomposition of the initial density operator, all have equal weight $w \neq 0$, so that the density operator (which is just a standard self-adjoint operator on the Hilbert

are many eigenstate associated with such eigenvalue. In this case they could well have different ψ_i 's in them, but again this is not a problem; the same argument just given applies, as if two eigenstates share the same eigenvalue, then so do all its linear combinations.

space $\mathcal{H}_S \otimes \mathcal{H}_M$) has an eigenspace of dimension m , with eigenvalue w . $P_{[U(\Phi_i)]}$ need not be an eigenstate of $\mathbb{I} \otimes A$, but if the final state is to have a spectral resolution in terms of such eigenstates, the states $U(\Phi_i)$ for $1 \leq i \leq m$ must span m of them.

More specifically, given an initial state of the object φ , and an initial apparatus state that is a density operator resolved by arbitrary states of the apparatus system ψ_{0i} , suppose the density operator has, for $1 \leq i \leq m$, weight w . Suppose also that for $\varphi \otimes \psi_{0i}$, $U(\varphi \otimes \psi_{0i})$ are not eigenstates of $\mathbb{I} \otimes A$. Still these eigenstates must be in the range of the operator $\sum_i P_{[U(\varphi \otimes \psi_{0i})]}$. U is unitary, and therefore invertible, so there must be an orthonormal system of states Φ_j in the subspace spanned by the $\varphi \otimes \psi_{0i}$'s, each of which will be mapped onto an eigenstate of the apparatus observable. But this subspace spanned by the states $\varphi \otimes \psi_{0i}$ consists of states of the form $\varphi \otimes \sum_i \alpha_i \psi_{0i}$, so that $\Phi_j = \varphi \otimes \sum_i \alpha_{ij} \psi_{0i}$. Let $\psi'_{0j} = \sum_i \alpha_{ij} \psi_{0i}$, then it follows that $U(\varphi \otimes \psi'_{0j})$ is an eigenstate of $\mathbb{I} \otimes A$.

This establishes that, given an initial object state φ there is a way to write the initial object-apparatus state as a sum of the projectors $P_{[\varphi \otimes \psi'_{0j}]}$ such that these will be mapped onto eigenstates of the apparatus observable. This is not enough for Brown's proof to work: the density operator must be the same for all initial object states, it cannot be dependent on the initial state of the object system. The reason for this is that RUE asserts that the evolution U maps an initial state with a specific decomposition in terms of states $\varphi \otimes \psi'_{0j}$, which results from the state being a proper mixture, to a final state where the pointer eigenstates are assumed to be the evolutes of the initial states in the decomposition, namely $U(\varphi \otimes \psi'_{0j})$. The proof works with such final states, and the ψ'_{0j} are selected independently of the initial object state by arguing that there is a preferred decomposition of the initial apparatus state. In order to show that Fine and Brown's proof can do without assuming RUE it is necessary then to show that *there is* a unique initial decomposition of the initial apparatus state so that the final state *is* a density operator resolved by pointer eigenstates. In other words, there must be a set of states ψ'_{0j} which we can point to, independently of the initial object states, as the ones which, when coupled to initial object states, are mapped by the measurement evolution to a final pointer eigenstate. Then it is not necessary to assume that the initial state is a mixture in order to claim that the final state consists of eigenstates of the form $U(\varphi \otimes \psi'_{0j})$.

To show that there can be a unique decomposition of the initial apparatus state when there are multiplicities in the density operator representing this state, we need to focus on the parts of the density operators which contain these multiplicities. These are, so to speak, the problem areas. Consider any two parts of final states which have multiple weights of order m , and denote them by $T = \sum_{i=1}^m w P_{[U(\varphi \otimes \psi_{0i})]}$ and $T' = \sum_{j=1}^m w P_{[U(\varphi' \otimes \psi'_{0j})]}$ for different initial states φ and φ' of the object system. T and T' are sums of projections over eigenstates of the observable $\mathbb{I} \otimes A$.

Assume that they are such that, for $i \neq j$, $\langle U(\varphi \otimes \psi_{0i}) | U(\varphi' \otimes \psi'_{0j}) \rangle = 0$ ¹³.

Then $\varphi \otimes \psi_{0i}$ and $\varphi' \otimes \psi'_{0j}$ must be sets of vectors in $\mathcal{H}_S \otimes \mathcal{H}_M$ satisfying

$$(3.1) \quad \langle \varphi \otimes \psi_{0i} | \varphi' \otimes \psi'_{0j} \rangle = 0 \text{ for } i \neq j.$$

I want to show that from this there follows that ψ_{0i} and ψ'_{0j} are the same set of orthonormal states. Two cases can arise:

- $\varphi \not\perp \varphi'$: In this case (3.1) reduces to

$$\langle \psi_{0i} | \psi'_{0j} \rangle = 0 \text{ for } i \neq j.,$$

which implies that ψ_{0i} and ψ'_{0j} are the same set of orthonormal states, considering that they must span the same (sub)space of \mathcal{H}_M .

- $\varphi \perp \varphi'$: Choose a vector φ'' such that it is perpendicular neither to φ nor to φ' . Then apply the argument of the above case for the pairs φ, φ'' and φ'', φ' to obtain that the density operators must have the same spectral resolution in this case, too.

The following result has been established:

PROPOSITION 2.3. *Consider any initial state $P_{[\varphi]} \in \mathcal{T}_1^+(\mathcal{H}_S)$ of the measured system, an initial apparatus state $\sum_{i=1}^n w_i P_{[\psi_{0i}]}$, a unitary operator U acting on $\mathcal{H}_S \otimes \mathcal{H}_M$ and a pointer observable for the apparatus system A_M (and corresponding pointer observable $\mathbb{I} \otimes A_M$ for the object + apparatus system). Assume further that objectification is satisfied in such a way that, for a final state $T = \sum_{i=1}^n w_i P_{[U(\varphi \otimes \psi_{0i})]}$, $U(\varphi \otimes \psi_{0i}) = \tilde{\varphi}_i \otimes \psi_i$ with $\psi_i \neq \psi_j$ for $i \neq j$.*

¹³This is the problematic assumption, apparently a stronger assumption than objectification. It will be analysed in the next section, where I will show how it in fact follows from the objectification assumption together with the assumption of unitarity of the measurement operator. Only then will it be possible to claim that Fine and Brown's proof is entirely equivalent to all other insolubility results.

Then there must be a resolution of the initial density operator in the object + apparatus system, independent of $\varphi \in \mathcal{H}_S$, in terms of orthogonal projections on $\mathcal{H}_S \otimes \mathcal{H}_M$ of the form

$$\sum_i w_i P_{[\varphi \otimes \psi'_{0i}]}$$

such that, for the final density operator

$$\sum_i w_i P_{[U(\varphi_i \otimes \psi'_{0i})]},$$

the $U(\varphi_i \otimes \psi'_{0i})$ must be eigenstates of the pointer observable $\mathbb{I} \otimes A_M$.

Clearly such a resolution will apply also in the case of an initial mixed state, as it applies for all the individual states of the decomposition of such a state.

The result says that, for any initial apparatus state of rank n there is a unique resolution of this state in terms of orthogonal projections such that the state of the apparatus after the measurement interaction will be a sum of n distinct and orthogonal projections on pointer eigenstates. Note that no assumption is made here about the specific value of n , nor indeed about degeneracies in the pointer observable.

We can then, without loss of generality, rewrite any initial state of an object-apparatus system in such a way that the final state will be a density operator over eigenstates of the apparatus observable, *given the additional assumption about objectification* made on page 73. In this case it is possible to say that, even if it is not possible to interpret the initial object + apparatus state as a proper mixture, as Brown wishes to, the argument of the proof that Fine and Brown give can be applied to yield an insolubility proof.

There remains the question of whether this apparent restriction of the objectification condition is actually a restriction, or whether it itself follows from the other assumptions used in the proof. We examine this in the next section, for it will be of interest in its own right to do so, as it will yield two further results, one a very simple insolubility proof, the other informative as to what the structure of the unitary evolution must be if objectification is assumed.

4. Insolubility without probability

This section, by further analysing the arguments of the previous section, will make clear in what sense insolubility proofs are independent of specific conditions tying probabilities of objects to probabilities of apparata in quantum measurement. In Section 2 of this chapter the argument has been that it is impossible to satisfy objectification and unitarity (with or without condition 5) on the one hand, and the probabilistic condition 6 on the other. Section 3 has analysed Fine's and Brown's proof, which establishes that objectification implies a violation of Fine's distinguishability condition, hence *a fortiori* of the probability reproducibility condition.

Stein, however, makes a stronger claim in his recent paper. He claims that, given the assumptions of unitarity and objectification, "the observation [the measurement process] conveys *no information whatever* about the antecedent condition of the object" [79], and he is right, of course. The implication of this is that there is no meaningful sense of measurement that can be attached to these processes, if we assume both unitarity and objectification. The first result of this section shows how, independently of which probabilistic condition is imposed, there is, given the unitarity of the measurement operator, a contradiction between objectification and condition 5. The second result is more general, and the closest to Stein's: it provides an explicit characterisation of any unitary operator U on a composite system $\mathcal{H}_S \otimes \mathcal{H}_M$ which satisfies objectification. Through yet another different route this establishes Stein's claim.

The crucial assumption in the discussion of Fine and Brown in the previous section is made on page 73, as I've already remarked several times. The requirement is that, for $i \neq j$, $\langle U(\varphi \otimes \psi_{0i}) | U(\varphi' \otimes \psi'_{0j}) \rangle = 0$. This can hold only if the final state is a sum of eigenstates of the pointer observable $\mathbb{I} \otimes A_M$ with distinct eigenvalues, and this is immediately in contradiction with 5, the first part of the probability reproducibility condition. Consider an initial state $P_{[\alpha_1\varphi_1+\alpha_2\varphi_2]} \in \mathcal{T}_1^+(\mathcal{H}_S)$, the φ_i 's eigenstates of observables to be measured on the object system, and an initial apparatus state $\sum_{i=1}^3 w_i P_{[\psi_{0i}]}$. Why should the final state be a sum of projections over three pointer eigenstates with distinct eigenvalues? If this were true, by condition 5 the initial state would be a superposition of three different observable eigenstates, which is clearly not the case. This remark suggests the following result, which demonstrates the contradiction between 3 and 5, independently of which probabilistic condition we adopt.

PROPOSITION 2.4. *Suppose an operation U acting on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ satisfies the objectification requirement 3 and also condition 5. Then U cannot be a unitary measurement action U^*TU .*

PROOF. Consider two initial states of the object, $P_{[\varphi]}$ and $P_{[\varphi']}$. Suppose they are such that $P_{[\varphi]}$ has non-zero probability with respect to two eigenstates of the discrete measured object observable (effects, if the observable is unsharp) and $P_{[\varphi']}$ has non-zero probability with respect to three eigenstates (or effects) of the same object observable, two of which are the same as for $P_{[\varphi]}$. Consider also an initial apparatus state $\sum_{i=1}^n w_i P_{[\psi_{0i}]}$, with n possibly infinite. Assuming objectification the final state for φ will be of the form

$$\sum_{i=1}^n w_i P_{[U(\varphi \otimes \psi_{0i})]} = \sum_{i=1}^n w_i P_{[\tilde{\varphi}_i \otimes \psi_i]},$$

and for φ'

$$\sum_{i=1}^n w_i P_{[\tilde{\varphi}'_i \otimes \psi_i]}.$$

5 implies that, in the case of φ , we will have that $\psi_i = \psi_1$ or ψ_2 , these being pointer observable eigenstates associated with different eigenvalues (not necessarily unique if the eigenvalues are multiple), indicating that φ_1 or φ_2 respectively have been measured. For φ' , on the other hand, $\psi_i = \psi_1, \psi_2$ or ψ_3 , again eigenstates associated with distinct eigenvalues. Consider in the second case a final pointer eigenstate $\tilde{\varphi}'_a \otimes \psi_3$. This is *necessarily* orthogonal to all the final eigenstates of the pointer observable appearing in the decomposition of the first final state, corresponding to the initial state φ , because it is a pointer eigenstate recording a value that is different from the eigenvalues corresponding to both eigenvalues associated with pointer eigenstates containing ψ_1 or ψ_2 .

If the operation U is generated by the conjugation action of a unitary operator U , then for some $i = a$, $\varphi' \otimes \psi_{0a}$ is orthogonal to all the states in the decomposition of the initial state for φ . This gives a contradiction: in order for $\langle \varphi' \otimes \psi_{0a} | \varphi \otimes \psi_{0i} \rangle = 0$ it is necessary that either $\langle \varphi | \varphi' \rangle = 0$, clearly not the case, or else that $\langle \psi_{0a} | \psi_{0i} \rangle = 0$ for all i . This is impossible, for there is no decomposition for the initial state $\sum_{i=1}^n w_i P_{[\psi_{0i}]}$ such that ψ_{0a} will be orthogonal to all the states in the decomposition. Such states must span the range of the self-adjoint operator $\sum_{i=1}^n w_i P_{[\psi_{0i}]}$, and therefore they must

span ψ_{0a} , so that the inner product of ψ_{0a} with them must yield at least one non-zero number. \square

The previous result is in itself an insolubility proof; it says that unitarity of the operator and objectification yield a contradiction with 5. It is worth exploring a bit further what restrictions on the form of the unitary operators are imposed by objectification alone.

The following corollaries follow straightforwardly from the arguments in the proof of Proposition 2.4.

COROLLARY 2.5. *Suppose the unitary operator U satisfies objectification. For any two states $P_{[\varphi]}, P_{[\varphi']} \in \mathcal{H}_S$, and given an initial state for the apparatus system T_M , the final states after the evolution of the object + apparatus system will be decomposable into a sum of projections over eigenstates of the pointer observable $\mathbb{I} \otimes A_M$. If one such pointer eigenstate in the resolution of the final state, evolved from the initial state $P_{[\varphi]} \otimes T_M$, has eigenvalue a_M , then there must be a pointer eigenstate in the final state evolved from $P_{[\varphi']} \otimes T_M$ with the same eigenvalue a_M .*

PROOF. See Appendix A. \square

COROLLARY 2.6. *For any initial state $P_{[\varphi]}$ of the object system $\mathcal{T}_1^+(\mathcal{H}_S)$, and an initial state for the apparatus $T = \sum_{i=1}^m w_i P_{[\psi_{0i}]}$, where the projections $P_{[\psi_{0i}]}$ are mutually orthogonal, if the final state after the unitary evolution satisfies objectification, then it must be decomposable into projections of the form $P_{[\tilde{\varphi}_i \otimes \psi_i]}$, with the ψ_i 's distinct eigenstates of the pointer observable A_M . Furthermore there must exist an orthonormal set of such ψ_i 's.*

PROOF. See Appendix A. \square

The previous corollaries need some comment. Suppose we have a final state of the form $\sum_i w_i P_{[\tilde{\varphi}_i]} \otimes P_{[\psi_i]}$, measuring an initial pure state $\varphi \in \mathcal{H}_S$, where the ψ_i are eigenstates of the pointer observable A_M . The first corollary tells us that the span of the reduced apparatus state $\sum_i w_i P_{[\psi_i]}$ must be the same for all initial object states φ . If not there must exist two object states φ and φ' for which the final apparatus states will have different span. In this case we can find, as in the proof of Proposition 2.4, a projection operator appearing in the decomposition of the final state associated to one of the two initial states which is orthogonal to all possible projections decomposing the

final state associated with the other state. This violates the assumed unitarity of the evolution.

The second corollary establishes that the rank of an initial apparatus state T_M must be equal to that of the final reduced apparatus state. Then the span of any final apparatus state is not only fixed, but also maximal relative to the rank of the initial apparatus state. Furthermore, there always exists an orthogonal decomposition of the final object + apparatus state in terms of eigenvectors of the apparatus observable $\mathbb{I} \otimes A_M$. Note that this does *not* mean that all vectors appearing in such a decomposition will be of the form $\tilde{\varphi} \otimes \psi_i$, with ψ_i an eigenstate of A_M . It is quite possible that projections over vectors such as $\sum_{i=1}^2 \alpha_i \tilde{\varphi}_i \otimes \psi_i$ might appear in the decomposition if ψ_1 and ψ_2 are eigenstates of A_M corresponding to the same eigenvalue, for then such an entangled vector will nevertheless be an eigenstate of $\mathbb{I} \otimes A_M$.

REMARK 2.3. *Note that in the previous proofs (see Appendix A) no assumption is made that the resolution of the final state in terms of pointer eigenstates must correspond to the evolutions of initial projections of the form $P_{\{\varphi \otimes \psi_{0i}\}}$, with the ψ_{0i} 's an orthogonal decomposition of the initial apparatus state. However at this point it is possible to invoke Proposition 2.3. Suppose we have a unitary evolution satisfying the objectification criterion, which we now know to entail, by Corollary 2.6, the extra assumption in proposition 2.3. This proposition then shows that there must exist an orthogonal resolution of the initial apparatus density operator, independent of the initial object state, such that projections in the resolution of the initial composite density operator, induced by the resolution of the initial apparatus density operator, will be mapped to projections which are eigenstates of the composite pointer observable.*

This finally and somewhat laboriously establishes, amongst other things, that Fine and Brown's proof outlined on page 67 can be used to provide an insolubility proof of the quantum measurement problem independently of the interpretation we impose on quantum mixtures.

This result also leads to the further theorem which characterises unitary measurement mappings with sharp pointer observables satisfying objectification of the pointer observable.

Consider a complete orthonormal set in $\mathcal{H}_S \otimes \mathcal{H}_M$ such as $\{\hat{\varphi}_1, \dots, \hat{\varphi}_n\}$, used in the proof of the previous Corollary, and the resolution in terms of orthonormal projections

onto ψ_{0i} of the initial apparatus state selected by Proposition 2.3. The states $U(\hat{\varphi}_j \otimes \psi_{0i})$ must be equal to $\sum_{k=1}^l \tilde{\varphi}_{ijk} \otimes \psi_{ijk}$, where ψ_{ijk} is an eigenstate of A_M , with possible multiplicities in the eigenvalues of A_M . When $k > 1$, the eigenstate of the pointer observable $\mathbb{I} \otimes A_M$ for the composite system is entangled. A vector (state) in a composite system such as $\mathcal{H}_S \otimes \mathcal{H}_M$ is called *entangled* if it is not of the form $\tilde{\varphi} \otimes \psi_i$ under any basis for $\mathcal{H}_S \otimes \mathcal{H}_M$; if it has this latter form it is called *factorisable*. An entangled state is always the sum of factorised vectors in $\mathcal{H}_S \otimes \mathcal{H}_M$. In the present case an eigenstate can be entangled if and only if the set $\{\psi_{ijk}\}_k$ all share the same eigenvalue of A_M .

If there are ψ_{ij} 's which are not orthogonal for different values of j (either because they are equal, or simply non-orthogonal eigenstates corresponding to the same eigenvalue), then in order to preserve the orthogonality of the initial states $\hat{\varphi}_j \otimes \psi_{0i}$ and $\hat{\varphi}_{j'} \otimes \psi_{0i}$ it will be necessary that

$$\langle \tilde{\varphi}_{ij} | \tilde{\varphi}_{ij'} \rangle = 0.$$

Note also the following definitions:

DEFINITION 2.3. A linear partial isometry (from now on simply *partial isometry*) on a Hilbert space \mathcal{H} is an operator W which is isometric on a subspace \mathcal{H}' of \mathcal{H} and null on the conjugate subspace \mathcal{H}^\perp of vectors orthogonal to vectors in \mathcal{H}' .

DEFINITION 2.4. $[\varphi_i]$ denotes the subspace spanned by the vector φ_i . Similarly $[\varphi_i]_i$ denotes the subspace spanned by the vectors φ_i in the range of the index i .

We then have the following

THEOREM 2.7. 1. Consider a complete orthonormal set $\{\hat{\varphi}_1, \dots, \hat{\varphi}_j, \dots, \hat{\varphi}_m\}$ in \mathcal{H}_S and an initial pointer state $\sum_{i=1}^n w_i P_{[\psi_{0i}]}$. Every operator $U : \mathcal{H}_S \otimes [\psi_{0i}]_i \rightarrow \mathcal{H}_S \otimes [\psi_{ij}]_{ij}$ is unitary and satisfies objectification of the pointer observable $\mathbb{I} \otimes A_M$ if and only if

$$(4.1) \quad U := \sum_{\substack{i=1, \dots, n \\ j=1, \dots, m}} \sum_{k=1}^{l_i} W^S_{ijk} \otimes W^M_{ijk}$$

where $W^S_{ijk}(\cdot) : \mathcal{H}_S \rightarrow \mathcal{H}_S$ is a partial isometry of the form

$$\langle \hat{\varphi}_j | (\cdot) \rangle (\tilde{\varphi}_{ijk})$$

on the subspace $[\hat{\varphi}_j]$ and $W^M_{ijk}(\cdot) : [\psi_{0i}]_i \rightarrow [\psi_{ijk}]_{ijk}$ is a partial isometry of the form

$$\langle \psi_{0i} | (\cdot) \rangle (\psi_{ijk})$$

that maps the (sub)space spanned by the initial states ψ_{0i} (which are fixed for any U through Proposition 2.3) to the (sub)space spanned by final pointer eigenstates ψ_{ijk} , in such a way that ψ_{ijk} , $\psi_{ijk'}$, $\psi_{ij'k}$ are eigenstates of the pointer observable with the same pointer eigenvalue.

2. The following equation also holds:

$$(4.2) \quad \sum_{k=1}^{l_i} W_{ijk}^S \otimes W_{ijk}^M = \sum_{k=1}^{l_i} \langle \hat{\varphi}_j \otimes \psi_{0i} | (\cdot) \rangle (\tilde{\varphi}_{ijk} \otimes \psi_{ijk}).$$

3. For any two different such partial isometries, W and W' , and any vector $\Phi \in \mathcal{H}_S \otimes \mathcal{H}_M$, $\langle W(\Phi) | W'(\Phi) \rangle = 0$.

In particular, if the partial isometries W_{ijk}^M are equal for all j 's and k 's then

$$(4.3) \quad U = \sum_{i=1}^n U_i^S \otimes W_i^M,$$

and if all U_i^S are equal

$$(4.4) \quad U = U^S \otimes U^M$$

PROOF. See Appendix A □

The result establishes the following facts. The main representation of unitary operators, given by equation (4.1), looks quite complicated, and indeed allows for entanglement between the object and apparatus system: the measurement can map an initial factorised object + apparatus state to a final entangled state. The entanglement in question, however, is possible only on the assumption that the pointer observable A_M has multiple eigenvalues. In this case the observable $\mathbb{I} \otimes A_M$ has entangled eigenstates: to see this note first that if ψ_i is an eigenstate of A_M then, for any $\varphi \in \mathcal{H}_S$, $\varphi \otimes \psi_1$ is an eigenstate of $\mathbb{I} \otimes A_M$; then, if ψ_1 and ψ_2 are both eigenstates of A_M with the same eigenvalue, it clearly follows that, for any two φ and φ' , $\alpha_1 \varphi \otimes \psi_1 + \alpha_2 \varphi' \otimes \psi_2$ is also an eigenstate of $\mathbb{I} \otimes A_M$. It is precisely these kinds of eigenstates which might emerge as a result of a unitary interaction satisfying objectification. The indexed l_i which appear in the characterisation of U 'track' the extent to which such an eigenstate is entangled, by determining how many factorised states make up the entangled eigenstate that can appear as a result of the unitary evolution.

It is only to this extent that objectification allows entanglement. When, however, the observable A_M has no multiple eigenstates (which is the case described by equation

(4.3)) then objectification allows for *no* interaction whatsoever between the object system on the one hand, and the apparatus system when this is in the states decomposing the initial apparatus density operator. That is, for such an observable A_M , initial factorised states of the composite system, where the apparatus state is one of the states ψ_{0i} , do not get entangled by the unitary operator satisfying objectification.

The ‘no-information’ result for quantum measurements, stated on page 75 as being Stein’s claim, then follows easily: an operator which satisfies objectification yields probabilities for pointer observable eigenstates which depend solely on the initial apparatus density operator: in particular the weights in the appropriate decomposition of the initial object + apparatus state, fixed by proposition 2.3 and by Corollaries 2.5 and 2.6, determine the final probabilities for eigenstates of the pointer observable $\mathbb{I} \otimes A_M$.

In the context of this result it is worth reviewing Busch and Shimony’s example, mentioned at the end of Section 2.

EXAMPLE 2.3. This example is supposed to be an example of how objectification can be satisfied through the conjugation action of a unitary operator, while illustrating the failure of the probability reproducibility condition in such a case.

Here is how Busch and Shimony present the example:

Let U be of the form $V \otimes V_a$, where V , V_a are unitary operators and V_a commutes with all F_n [F_n is a finite or countable family of mutually orthogonal projections of the apparatus system, characterising pointer observables]. Furthermore let T_a be a mixture of F_n -eigenstates. Then $U(P_\varphi \otimes T_a)U^{-1} = VP_\varphi V^{-1} \otimes V_a T_a V_a^{-1}$ is a mixture of $\mathbb{I} \otimes F_n$ -eigenstates, in fulfilment of [the objectification condition]. But the probabilities $\text{Tr}[\mathbb{I} \otimes F_n U(P_\varphi \otimes T_a)U^{-1}] = \text{Tr}[F_n V_a T_a V_a^{-1}] = \text{Tr}[F_n T_a]$ are independent of φ so that [the probability reproducibility condition] is violated. [17, p. 401]

Theorem 2.7 characterises all operators U that satisfy objectification. Busch and Shimony’s is a special case of such an operator (explained in the third claim of Theorem 2.7) where furthermore the initial state of the apparatus also happens to be a mixture over pointer states. If T_a is a mixture of F_n -eigenstates, let F_i be the projection on any one such eigenstate. If F_i commutes with an operator V_a , then it is a fixed point under the action $V_a F_i V_a^{-1}$, i.e. it is left unchanged by this action. Then, independently of φ , the unitary operator in question is of the form $V \otimes \mathbb{I}_M$ and the final state under

the action given in Busch and Shimony's example will be $VP_\varphi V^{-1} \otimes T_a$. This shows in an evident way what it means to say that unitary operators satisfying objectification cannot transmit any information from the object to the apparatus system: this unitary operator has no effect whatsoever on the apparatus system! In particular, the probabilities of the pointer observable with respect to the final state clearly depend solely on the weights of the initial apparatus density operator. Theorem 2.7 shows how all unitary operators satisfying objectification in a sense do this same thing, and what their general form is.

5. Insolubility proofs in the case of unsharp pointer observables

This section investigates the extent to which the insolubility results discussed in previous sections have counterparts in cases in which the pointer observable is unsharp, and under which conditions. To even begin to address this question, however, we have to reformulate some of the conditions imposed in the previous sections of the chapter. Despite doing this it will be apparent that none of the content of these conditions is substantially changed. Rather the conditions are suitably modified to deal with the conceptual differences which arise once we work with unsharp observables.

Before we address the problem, a brief introduction to unsharp observables in quantum mechanics is required (for more comprehensive treatments, see the classic texts by Ludwig [63] and Davies [25], or more recently Busch, Grabowski and Lahti [13]).

5.1. Unsharp Observables and Unsharp Objectification. Traditionally observables are identified with self-adjoint operators on a Hilbert space \mathcal{H} . A more recent approach identifies observables with mappings from a subset of the power set of the set of possible values of an observable (usually the Borel sets) to projection operators on a Hilbert space (including \mathbb{O} and \mathbb{I} , the null and identity operators on \mathcal{H}). Thus, if the set of values is the real line, one associates a projection operator with every (Borel) subset of the real line. For example, consider the observable 'spin in the z direction'. This has values 1 and -1 . To the set $(0, 2)$ (having a value between 0 and 2) such a mapping will associate the projection operator P_1 onto the eigenstate corresponding to the eigenvalue 1; to the set $[-1, 1]$ the identity \mathbb{I} ; to $[2, \infty]$ the null map \mathbb{O} . Such mappings are known as *projection valued measures*, PV measures for short, and they

characterise so-called sharp observables. These are the standard quantum mechanical observables, and there is a one-to-one correspondence between PV measures and bounded self-adjoint operators on the Hilbert space.

PV measures can be used to define the probability measures which provide quantum mechanical information. However the condition that sets be mapped to projection operators is stricter than is needed to obtain a probability measure. It suffices that the operator that sets are mapped to be a normalised positive operator. We then encounter the generalisation to positive operator valued measures, or POV measures.

Those observables characterised by POV measures which aren't PV measures are called unsharp. Consider the generalisation of the above spin example by means of the observable E . E maps $\{1\} \subset \mathbb{R}$ to the normalised positive self-adjoint operator

$$F_{1\epsilon} = (1 - \epsilon)P_1 + \epsilon P_{-1},$$

$\epsilon \ll \frac{1}{2}$, $\{-1\} \subset \mathbb{R}$ to the operator

$$F_{-1\epsilon} = \epsilon P_1 + (1 - \epsilon)P_{-1},$$

and all subsets of $\mathbb{R} \setminus \{-1, 1\}$ to the null operator \mathbb{O} . It is easy to see how this additively generalises to a mapping for all subsets X of \mathbb{R} . F is not a sharp observable; nevertheless, for any state $T \in \mathcal{T}_1^+(\mathcal{H})$ the quantity $\text{Tr}[TF(X)]$ is a well-defined probability, for $F(X)$ a normalised positive self-adjoint operator associated with the subset X of \mathbb{R} . For instance if $T = P_{[\varphi]}$ with $\varphi = \alpha_1\varphi_1 + \alpha_{-1}\varphi_{-1}$ and φ_1, φ_{-1} eigenstates of the sharp spin observable, then it is easy to see that $\text{Tr}[TF_{1\epsilon}] = (1 - \epsilon)|\alpha_1|^2 + \epsilon|\alpha_{-1}|^2$ which is positive and less than one. This number represents the probability that a certain spin direction is unsharply realisable (in a sense that will be defined shortly) if the object is subjected to a measurement of the unsharp spin observable E . To the generalisation from PV to POV measures there corresponds a generalisation from projections such as P_1 to normalised positive self-adjoint operator such as $F_{1\epsilon}$, which are known as *effects*.

The precise definition is as follows¹⁴. Let Ω be a non empty set and \mathcal{F} a σ -algebra of subsets of Ω so that (Ω, \mathcal{F}) is a measurable space. A POV measure $E : \mathcal{F} \rightarrow \mathcal{L}(\mathcal{H})$ ($\mathcal{L}(\mathcal{H})$ the set of bounded linear operators on \mathcal{H}) on (Ω, \mathcal{F}) is defined through the

¹⁴This particular one is taken from Busch, Lahti and Grabowski [13].

properties:

$$(5.1a) \quad E(X) \geq 0 \quad \text{for all } X \in \mathcal{F}$$

$$(5.1b) \quad E(\Omega) = \mathbb{I}$$

$$(5.1c) \quad E(\cup X_i) = \sum E(X_i) \quad \text{for all disjoint sequences } (X_i) \subset \mathcal{F}$$

where the series in (5.1c) must converge in the weak operator topology of $\mathcal{L}(\mathcal{H})$. For a PV measure the positivity condition (5.1a) becomes the more familiar idempotency condition. Note that the set of effects $E(X)$ associated with a system is in fact nothing but the convex combination of the projection operators *of any dimension* associated with the system (for a proof, see Davies [25, p. 19]). Note also that it is still possible to associate a positive, self-adjoint operator to any POV measure. Intuitively, we can do this simply by adding all the effects which are mapped from the spectral set of the POV measure, multiplied by the spectral values associated with them, together. As effects are bounded and self-adjoint, their sum will also be a bounded, self-adjoint operator. However, the uniqueness of the observables associated with the self-adjoint operator is lost: many different POV measures will be associated with the same operator.

The standard definition of when a property is realised in a quantum state is just a generalisation of the eigenvalue-eigenstate link: if, for a state T and a projection P , $\text{Tr}[TP] = 1$, then the property P is realised in state T . This can be generalised to a notion of unsharp realisability for so called *regular* effects, namely effects whose spectrum extends both above and below $\frac{1}{2}$. For two such effects F and its complement F^\perp (note that the map $F \rightarrow F^\perp$ is an orthocomplementation) there exist states T and T' for which $\text{Tr}[TF] > \frac{1}{2}$ and $\text{Tr}[T'F^\perp] > \frac{1}{2}$. We can say that F is unsharply realised in T and F^\perp is unsharply realised in T' .

Now suppose we allow for an unsharp pointer observable. It is essential to reformulate the objectification condition. Pointer objectification reflects sharp realisation by requiring the final state of a measurement interaction to be a mixture of pointer eigenstates of a self-adjoint operator representing the observable: these are states for which the property “having the pointer point to the value x ” is sharply realised. If the pointer observable is unsharp this condition must be modified. Busch suggests the following condition: we require the components of the final state to be quasi-eigenstates of the effects making up the pointer observable, in the sense of giving probabilities close to 1 for the properties that these effects denote. In other words, given regular effects

F we look for (pure) states T for which $\text{Tr}[TF] \geq 1 - \epsilon$ for some arbitrary selected ϵ close to zero. Such effects will form a smaller set than the one of regular effects defined above.

These states should form a density operator for which it is possible to apply the ignorance interpretation to the reduced state of the apparatus: the quasi-eigenstates of the pointer observable on the composite system should decompose the final density matrix, much as the pointer eigenstates should do for the sharp pointer observable case. The reduced apparatus state should also be a mixture over quasi-eigenstates of the pointer observable.

In order to make clear this condition, some points must be clarified and some assumptions must be made explicit. The first assumption is that, again, the pointer observable E_M is discrete. This now is taken to mean that, for a set of possible values Ω , there exists a countable subset Σ such that $E_M(\Omega \setminus \Sigma)$ is equal to the null projector. The elements of Σ will correspond to measured values. This will induce, much as for the sharp case, a discretised version of the measured observable.

Secondly, what should really count as a quasi-eigenstate? Suppose the pointer observable has the formal properties of the ‘smeared’ spin observable E . If the only condition for a quasi-eigenstate were that its probability with respect to one of the two effects be greater than $1/2$, this would make every state of \mathbb{C}^2 a quasi-eigenstate. For the time being it is assumed that the quasi-eigenstate corresponding to each effect is the state for which the effect yields maximal probability. This will be discussed after the next point.

There are two ways to interpret such an unsharp objectification condition, depending on the distinction between commutative and non-commutative unsharp observables.

Recall that here we consider observables as mappings from a value set to effects. If such effects turn out to form a complete ¹⁵ set of pairwise orthogonal projections, then the observable is sharp. Consider the case of a mapping from a set of values to a set of effects which are not projections. These effects are themselves representable as self-adjoint operators on a Hilbert space and therefore admit of a spectral representation.

In particular, we can associate orthogonal projections with these effects, corresponding to projections on the effects’ eigenstates, their *spectral set*. If effects characterising

¹⁵In the sense that their sum is the identity operator on the Hilbert space under consideration.

an unsharp observable have spectral sets whose union form a complete pairwise orthogonal set of projections, the unsharp observable is said to be commutative¹⁶. Consider the example above of the unsharp generalisation of a spin observable on \mathbb{C}^2 . The spectral set of the effect $E_{1\epsilon}$ is $\{P_1, P_{-1}\}$, equal to the spectral set of $E_{-1\epsilon}$, and obviously to their union. This set is a complete set of orthogonal projections for \mathbb{C}^2 and so the observable is commutative. This is no coincidence: the observable in question is constructed from a sharp one, essentially through a smearing procedure¹⁷.

On the other hand not all unsharp observables are commutative, as the following example shows.

EXAMPLE 2.4. Consider the mapping $F : \{1, 2, 3\} \Rightarrow \mathcal{F}(\mathbb{C}^2)$ (where $\mathcal{F}(\mathbb{C}^2)$ is the set of effects on the space \mathbb{C}^2) defined by

$$F(1) := \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad F(2) := \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad F(3) := \frac{1}{4} \begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix}.$$

F is an unsharp observable, but

$$[F(1), F(2)] = \frac{1}{8} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \frac{i}{8} \sigma_y \neq \mathbb{O},$$

so the observable is not commutative¹⁸.

Returning to the question about quasi-eigenstates, the choice of quasi-eigenstates above makes trivial the consideration of commutative unsharp observables as possible pointer quasi-eigenstates. This is because it is easy to see that the state which will yield the maximal probability for a given effect (that is, a bounded self-adjoint operator which is a convex combination of projections) will be a projection over an eigenstate of the

¹⁶The usual definition of commutativity here is that, for all values x, y in the set of possible values of an observables, the effects associated with these values commute, i.e. $[E_x, E_y] = \mathbb{O}$. We use this other definition because it will be useful to think in terms of the complete pairwise orthogonal set of projections defined above when selecting quasi-eigenstates.

¹⁷It is well known, at least for the finite-dimensional case, both that an unsharp commutative observable can be constructed, by a standard procedure, from a realisation of a sharp observable and from a suitable finite family of functions (see [20]), and that a commutative unsharp observable can be understood as an unsharp realisation of an essentially unique sharp observable (see [21]).

¹⁸This example is taken from Cattaneo, Nisticò and Bacciagaluppi [21].

effect, namely the eigenstate corresponding to the greatest eigenvalue¹⁹. If the quasi-eigenstate is a projection over an eigenstate of effects, clearly the case of commutative observables would lead to a set of pairwise orthogonal pointer quasi-eigenstates, which clearly would pick out a sharp observable as pointer observable, and be subject to insolubility results such as the ones discussed in the previous sections. In order not to make the question trivial it is necessary, therefore, to assume that quasi-eigenstates will not be pairwise orthogonal.

At this point the assumption made is that there is a unique quasi-eigenstate associated with each effect that makes up the object observable to be measured, and such quasi eigenstates are not pairwise orthogonal.

For interpretative reasons it might be appropriate to also assume that such quasi-eigenstates are not linearly dependent, even though this is not a necessary assumption for the proof to be given in the next subsection. Two reasons can be given for this, a technical one and a less strong conceptual one. Such an assumption follows if it is assumed that quasi-eigenstates must yield probability close to one, $1 - \epsilon$ with ϵ very small, for their associated effects. More specifically, suppose, without loss of generality, that there are three quasi-eigenstates,

$$\varphi_1, \quad \psi_1 = \alpha_1 \varphi_1 + \alpha_2 \psi, \text{ and } \chi_1 \frac{\beta_1 \varphi_1 + \beta_2 \psi_1}{\|\beta_1 \varphi_1 + \beta_2 \psi_1\|} = \beta_{\varphi_1} \varphi_1 + \beta_{\psi} \psi$$

with $|\alpha_1|^2 + |\alpha_2|^2 = |\beta_{\varphi_1}|^2 + |\beta_{\psi}|^2 = 1$ and $\varphi_1 \perp \psi$, associated to the effects

$$T_1 = a_1 P_{[\varphi_1]} + a_2 P_{[\varphi_2]}$$

$$T_2 = b_1 P_{[\psi_1]} + b_2 P_{[\psi_2]}$$

$$T = c_1 P_{[\chi_1]} + c_2 P_{[\chi_2]},$$

where, without loss of generality, φ_2 , ψ_2 and χ_2 are taken to be orthogonal to the span of the set $\{\varphi_1, \psi_1\}$ and $a_1 \geq b_1 \geq c_1 \geq 1/2$.

It follows that

$$S = a_1 P_{[\varphi_1]} + b_1 P_{[\psi_1]} + c_1 P_{[\chi_1]} = \begin{pmatrix} a_1 + b_1 |\alpha_1|^2 + c_1 |\beta_{\varphi_1}|^2 & b_1 (\overline{\alpha_1} \alpha_2) + c_1 (\overline{\beta_{\varphi_1}} \beta_{\psi}) \\ b_1 (\alpha_1 \overline{\alpha_2}) + c_1 (\beta_{\varphi_1} \overline{\beta_{\psi}}) & b_1 |\alpha_2|^2 + c_1 |\beta_{\psi}|^2 \end{pmatrix}$$

which has trace $a_1 + b_1 + c_1$. Now the sum of effects making up an observable must be the identity matrix, so the sum of (parts of) effects over a two-dimensional subspace of

¹⁹It is a well-known theorem of functional analysis that the spectrum of a self-adjoint operator has a maximum which is finite, as proved for instance in Riesz and Nagy [72, pp. 231–235].

the space over which the observable is defined must be a projection operator over that subspace, and in particular it must have trace 2. So if the effects T_1 , T_2 and T are to be possible effects of the one observable, it must be that, as a minimum requirement, $a_1 + b_1 + c_1 \leq 2$. But if φ_1 , ψ_1 and χ_1 are to be quasi-eigenstates then they must yield probability close to one with respect to their associated effects, so that, for example,

$$\text{Tr}[T_1 P_{[\varphi_1]}] \approx 1 - \epsilon.$$

Yet even the choice of $a_1 = b_1 = c_1 = 3/4$ results in $a_1 + b_1 + c_1 = 9/4 > 2$. The argument clearly generalises to effects which have associated quasi-eigenstates linearly dependent on a finite number of quasi-eigenstates. However the argument will not work in the case of a quasi-eigenstate which is a sum of infinitely many other quasi-eigenstates.

The assumption does seem to be reasonable, though, simply because it would be strange to think that a measurement would result in pointer states which are superpositions of one another. The quantum theory of measurement emerges to account precisely for the fact that superpositions of properties that are measured aren't observed, so countenancing final pointer states which are linearly dependent on one another might seem an odd way to go about resolving the issues with measurement, for it's unclear how it is that such *pointer states* could be said to be observed.

5.2. A result for unsharp pointer observables. The previous subsection has clarified the following assumptions:

1. The pointer observable is a discrete unsharp observable: that is the observable is of the form $A_M : \mathbb{R} \rightarrow \mathcal{E} := X \rightarrow A_M(X)$ and there is a countable subset C of \mathbb{R} such that $A_M(\mathbb{R} \setminus C) = \mathbb{O}$, the null operator. Elements of C are the readings associated with the pointer observable.
2. Unsharp objectification stands for the condition that the final object + apparatus system be represented by a density operator which has a decomposition in terms of a set of quasi-eigenstates of an observable. Quasi-eigenstates are one-dimensional projection operators whose probabilities with respect to the effect they are associated with (the effect corresponding to the number in C that the quasi-eigenstate indicates is measured) are approximately 1.
3. Objectification implies that the measurement process measures discrete observables, which might be coarse-grained versions of observables with a continuous spectrum in the case of an infinite dimensional system.

4. It is assumed, for the purpose of the present discussion, that, for each effect associated with a single value in the countable set C , there is exactly one quasi-eigenstate corresponding to this effect, the quasi-eigenstate which has maximum probability with respect to this effect. This is equivalent to saying that each such effect is represented as a convex sum $a_1 P_1 + \dots$ of projectors where a_1 is the greatest weight in the convex sum, and its associated projection P_1 is one-dimensional. Then the state P_1 will be the unique quasi-eigenstate associated with the effect according to the above definition. For the moment, all quasi-eigenstates are assumed to be pairwise non-orthogonal.
5. Finally, the condition 5 is assumed; this condition is here interpreted as saying that the final density operator for the apparatus should be decomposable into the quasi-eigenstates which correspond to the effects of the measured object observable having non-zero probability with respect to the initial state.

Consider now an orthonormal basis $\hat{\varphi}_j$ of the object system \mathcal{H}_S which has the property that all the $\hat{\varphi}_j$'s yield non-zero probability with respect to all the effects of the object observable to be measured.

Consider also the set ψ_i of quasi-eigenstates; by (5) the final state of the object + apparatus system must be, for all initial object states $\hat{\varphi}_j$'s, a density operator with a decomposition $\sum_i w_i \tilde{\varphi}_{ij} \otimes \psi_i$.

Now suppose the initial state of the apparatus to be T_M . All initial object + apparatus states of the form $\hat{\varphi}_j \otimes \psi_{ij}$, where ψ_{ij} is in the range of T_M , must be orthogonal for different j 's. Then, assuming the measurement evolution to be given by a unitary operator U , all $\tilde{\varphi}_{ij} \otimes \psi_i$'s must be orthogonal for different j 's, which implies that $\tilde{\varphi}_{ij} = \tilde{\varphi}_j$ and that all such $\tilde{\varphi}_j$'s are pairwise orthogonal.

So the unitary measurement evolution U maps initial states $\hat{\varphi}_j \otimes \psi_{ij}$ to final states $\tilde{\varphi}_j \otimes \psi_i$. As, by unitarity, $\langle \hat{\varphi}_j \otimes \psi_{ij} | \hat{\varphi}_j \otimes \psi_{i'j} \rangle = \langle \tilde{\varphi}_j \otimes \psi_i | \tilde{\varphi}_j \otimes \psi_{i'} \rangle$ for different i and i' , it follows that $\langle \psi_{ij} | \psi_{i'j} \rangle = \langle \psi_i | \psi_{i'} \rangle$. Should all ψ_{ij} 's be equal for different j 's the result would be immediate: it would be easy to show that the only unitary operator that could have such properties would be of the form $U_S \otimes U_M$. It would then follow, just as it does in section 4, that there could be no transmission of information between the object and apparatus system. But the ψ_{ij} 's are not orthogonal, and it is quite possible that there are different sets of them, with inner products fixed, which might all be mapped to the appropriate final states.

We now need to work with the sets of initial pointer states $\{\psi_{ij}\}_i$ decomposing T_M , one for each j denoting a different initial object state $\hat{\phi}_j$. How can they differ for different j 's, while preserving the inner product relations captured by $\langle \psi_{ij} | \psi_{i'j} \rangle = \langle \psi_i | \psi_{i'} \rangle$?

The first thing to notice is that, given a state T_M and the inner product property, the way that the ψ_{ij} can change in the decomposition is via a unitary transformation of the range of T_M which leaves T_M invariant. That is, the states ψ_{ij} must be transformed by a unitary evolution which has as its eigenstates the states projected on by the orthogonal projections decomposing T_M . Consider such a transformation V . It is unitary, and will then act on the state $T_M = \sum_i w_i P_{[\psi_{ij}]}$, in the ψ_{ij} decomposition, by means of the action $V(T_M)V^{-1} = V(\sum_i w_i P_{[\psi_{ij}]})V^{-1} = \sum_i w_i P_{[V(\psi_{ij})]}$.

Crucially, any such transformation leaves invariant the weights w_i ²⁰. Then a unitary measurement operator U will lead, for an initial state $\hat{\phi}_j$ in the set considered above, to the evolution

$$U(\sum_i w_i P_{[\hat{\phi}_j \otimes \psi_{ij}]})U^{-1} = \sum_i w_i P_{[U(\hat{\phi}_j \otimes \psi_{ij})]} = \sum_i w_i P_{[\hat{\phi}_j \otimes \psi_i]},$$

for which it is easy to check that the final reduced state for the apparatus is

$$\sum_i w_i P_{[\psi_i]}$$

independently of the specific form of $\hat{\phi}$. U will fail to convey, therefore, the probabilistic information contained in $\hat{\phi}$, and this establishes an insolubility result.

6. Conclusions

In the first instance this chapter has added to the already considerable literature on quantum measurements and insolubility proofs three new results, and has rehabilitated somewhat Fine and Brown's proof strategy by analysing the idea of real unitary evolution: by giving up an unnecessary part of RUE related to the interpretation of density operators in quantum mechanics and showing that the technical part of RUE is not independent of other assumptions, but rather follows from them, a proof can be given.

²⁰Note how the argument here is just a generalisation of what happens in the case of a density operator with multiplicities in the initial weights given earlier in subsection 2.3, when one asks for different decompositions of the operator which preserve orthogonality relations between the decomposing states, again such different decompositions having no effect on the weights of the density operator.

There are then several different ways of proving the insolubility of the quantum measurement problem. The proofs of Fine, Shimony, Brown and Stein cover the traditional cases as do the proofs in Section 2 and Proposition 2.4. More restricted results can be obtained by explicit characterisation of operators of a different kind. By the latter I mean that, by providing a full characterisation of the unitary operators governing measurement evolutions and satisfying the probability reproducibility condition 4, as is done, for example, in Beltrametti, Cassinelli and Lahti [6], it is easy to see almost by inspection that objectification cannot be satisfied for the cases covered by the characterisation result. We can do this just by checking the form of the projectors on $\mathcal{H}_S \otimes \mathcal{H}_M$ that make up the final state of a measurement evolution. It helps if the characterisation is general, and in particular if it specifies the form of the unitary operator exactly, rather than as an (unspecified) completion of non-unitary operators, as is generally done in the literature; this and related questions are discussed in the next chapter.

More importantly this chapter provides a characterisation of all unitary operators which satisfy objectification, given an initial apparatus state. This enables a clearer understanding of how Stein's result about no possible information transfer in measurements satisfying objectification comes about. By relying on arguments which use the orthogonality of decompositions less explicitly, it also suggests techniques that might be applied to giving a similar result in the case of unsharp pointer observables. A first result along these lines is developed in Section 5 for unsharp noncommutative pointer observables. Generalisation of this result is left for a later stage.

The discussion of this chapter also allows some comments on the relevance of insolubility proofs. Shimony remarks that if RUE were true, the philosophical consequences of the insolubility results would become almost trivial, as such result would follow simply from the dynamics. By this he means that, if given an initial object + apparatus state $\sum_i w_i P_{[\varphi \otimes \psi_{0i}]}$, the pointer eigenstates of the final state must necessarily be of the form $U(\varphi \otimes \psi_{0i})$, then by linearity of the dynamics the final state for the composite system must necessarily be $\sum_i w_i P_{[U(\varphi \otimes \psi_{0i})]}$ and trivially the probabilities with respect to the pointer observable for the combined system will be given by w_i 's or sums of them. In a sense all proofs presented in this chapter aim to establish that insolubility

of the quantum measurement problem *is* simply a consequence of the dynamics. However I don't share Shimony's pessimism about the philosophical, or physical, weight of the conclusion.

There is one sense in which Shimony's remark is important. The more conditions we impose in the search for a certain result, the more likely it is that such result will not be achieved. Adding RUE as a condition for a (yet unknown) measurement evolution makes it less likely that such evolution will exist. If furthermore it can be envisaged that such a condition might not be met in a reasonable measurement (and this seems possible) then Brown's proof, by accepting this condition, does fall short of the required result.

But that such a proof might be essentially a consequence of quantum mechanical unitary dynamics seems neither trivial nor philosophically unimportant, for several reasons. It is technically no less trivial a proof to give; it requires more than just pointing to a mock evolution and showing that it goes wrong, in the manner just outlined. Careful consideration must be given, for example, to the cases when multiplicities arise.

The conceptual importance of such proofs is also unaffected by the methods by which they are achieved. It is helpful to look at what importance these proofs had for their first proponents. Why did Wigner set his mind to proving these results?

Wigner's concern was to definitively establish, in Von Neumann's words, that "the non-causal nature of the process 1. is not produced by any incomplete knowledge of the state of the observer". [82, page 439]

Both Von Neumann and Wigner believed in the fundamental difference between the unitary and non-unitary evolutions seemingly required by quantum mechanics. Von Neumann premises the above quotation with an argument formally very similar to the result of Section 2, as was mentioned before. Both want to put forward a cogent defence of non-unitary evolutions and the proofs, complicated or otherwise, are aimed at establishing this. It is arguable that such an explicit aim would have led to little concern with the techniques used and would have made an easier and more intuitive proof of such fundamental difference quite welcome.

Yet the outcome of establishing an incompatibility between unitarity, objectification and probabilistic conditions is no longer taken to mean that we must give up unitarity of certain quantum evolutions. Superficially at least these proofs assume four conditions for measurements and really work with three: objectification, probability

reproducibility and unitarity of the measurement operator. This clarifies the stakes somewhat: the incompatibility of these conditions must mean that one of them has to be rejected, but is not necessarily an argument against the unitarity of the measurement evolution in the manner that Von Neumann and Wigner had envisaged it.

The probability reproducibility condition is often treated as being pretty much safe. It is not only that it embodies, more than any of the other conditions, what is thought to be important about quantum mechanics, namely its status as an essentially probabilistic theory. It is clear from the above proofs that unitarity and objectification imply a probability distribution over pointer states that is independent of the initial state of the apparatus; in fact a proof can be given which abstracts from the probabilistic content of the probability reproducibility condition, which links probabilities for the object system to probabilities of the apparatus system. So no relaxation of probabilistic conditions alone can achieve much, unless we are prepared to consider the possibility that measurement of an eigenstate of an object observable will not necessarily yield a given value with certainty.

One obvious way to get around the insolubility results is to give up the idea that measurement is a unitary, closed-system interaction, and hence best captured by a bijective, invertible mapping. The next chapter will show a way to construct measurement mappings which satisfy objectification and the probability reproducibility condition, relying on Davies' work on open quantum systems. The idea is simply to construct a mapping on the set of density operators of a composite which is a sum of Davies' pure mappings of type 3. These kinds of mappings are effectively generalisations of Von Neumann type 1. evolutions on composite systems. The problems with this approach are philosophical, more than physical: taking this proposal seriously seems to involve the belief that at some point the evolution is fundamentally non-unitary, rather than a projection on a subspace of a unitary evolution on a larger space. Otherwise the problem just resurfaces.

Another option is to look more closely at objectification. As remarked above, this condition allows the applicability of what Fine calls the eigenvalue-eigenstate link. A number of problems have emerged with this condition specifically, though. One is that it effectively underwrites the possibility of giving an ignorance interpretation of the final apparatus mixture. But if such an interpretation is mistaken, as has been discussed in the chapter, if it indeed amounts to assuming that quantum mechanics

is incomplete, then it is less obvious why we should assume objectification. The final apparatus state certainly can tell us what the probability of a system having a certain property is regardless of how it is generated from a composite state. If the apparatus density operator does not, on the other hand, allow us to claim that the system *is* in one particular pure state, then the motivation for having the right composite state to make this claim seems considerably weakened.

There has indeed been over the past 40 years a long tradition of formulating interpretations of quantum mechanics that renounce this criterion for states in a composite system, starting with Everett's work, and including, for instance, modal interpretations, although these approaches also have many problems, which won't be discussed here.

What insolubility results teach us, from a foundational and philosophical point of view, is the importance of these lines of research, given that the most intuitive approach fails: this for me has been and still is their most important role.

CHAPTER 3

Describing Measurement Interactions in Quantum Measurement Theory

This chapter investigates one of the standard assumptions of the quantum mechanical account of measurement, that *premeasurement* interactions between a system and a measuring apparatus are mathematically described by a unitary operator.

A premeasurement is a transition which takes an initial object + apparatus state to a final state in such a way that the transition satisfies the probability reproducibility condition. It is called a premeasurement since it is not automatically implied that such a transition will satisfy the objectification condition. Indeed, assuming the eigenvalue-eigenstate link and that the object and apparatus systems are both proper quantum systems, the premeasurement cannot be a full measurement in the sense of satisfying objectification, as was shown in the previous chapter. The study of the properties of premeasurement operators goes back at least to Von Neumann [82, pp. 440–445], who gave the first characterisation result for a class of such operators.

It is important to study premeasurements carefully for at least two reasons. Firstly, in those interpretations which drop the eigenvalue-eigenstate link these premeasurements (or perhaps a subset of them, as argued by Van Fraassen [80, pp. 211–2]) yield the complete description of measurement processes; such interpretations include Everett’s relative state interpretation of quantum mechanics [33, 34] and its derivative interpretations, such as many worlds and many minds interpretations (on the latter see for example Albert [1]), as well as modal interpretations of all varieties, (see Kochen [52], Healey [48], Vermaas and Dieks [81] and Bacciagaluppi [3]).

Secondly, it has been argued (for instance by Kronz [54], as seen in Chapter 1) that Von Neumann’s process of state reduction is best applied to a composite object + apparatus system which has undergone the appropriate unitary premeasurement evolution: this can help in giving a description, at the level of theory, of apparently problematic measurement situations, such as destructive photon measurements. For

the rest of this chapter the transitions will be referred to simply as measurements, except in cases where this might create confusion.

The chapter begins by showing that in the Hilbert space representation such measurement interactions are fully described by a unique and ‘minimal’ partially isometric and contractive operator; usual unitary operators are non-unique completions of this operator. It then studies what possible transformations on the space of density operators of a system mirror the properties of the partial isometry characterised for the Hilbert space representation. Several different options are explored and compared, and I prove a new result, showing how the only operations which can satisfy both probability reproducibility and a generalised version of unitarity, which retains its essential content, are those generated by conjugation of the unitary completions just mentioned.

This analysis then enables the detailed investigation of such measurements where the initial apparatus state is a non-pure density operator. The basic informational properties of Von Neumann-Lüders, repeatable and first-kind measurements (to be defined later on) are affected under these circumstances: it will be shown that, when the initial apparatus state is a density operator which is not pure, the measurements under consideration cannot satisfy these informational properties. Furthermore, for a certain class of initial density operators, it is impossible to define any ‘unitary’ measurement at all which satisfies the probability reproducibility condition. It is easy to characterise the basic properties of this class of density operators. This in turn suggests the theoretical possibility of direct tests for the quantum theory of measurement: these cases could well appear in measurements in practice, while the theory rules them out. I will conclude by discussing some of the consequences of this analysis for the quantum measurement problem. In particular the problems raised for the case of an initial apparatus state which is not a pure density operator are linked to the problem of interpreting quantum mixtures, and raise some issues which connect to the discussion in Chapter 1. Both of these merit some comment.

1. Some introductory remarks and definitions

In the standard approach to measurement theory, a measurement interaction between a system S and a measurement apparatus M is thought to be described, in the first instance, by a unitary operator acting on the Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_M$, tensor

product of the Hilbert spaces \mathcal{H}_S and \mathcal{H}_M describing the system and the apparatus respectively (see for instance Beltrametti et al. [6] and Busch et al. [14]).

The setup described is such that information about an observable, fixed for a physical system under observation, must be obtained on the basis of the results given by an observable for a measuring apparatus, prepared in an initial state, after a coupling of system + apparatus and a dynamical evolution of this pair.

The standard quantum mechanical (QM) description of this situation on Hilbert spaces is based on the following assumptions:

1. both system and apparatus observables are described by projection valued (PV) measures (or equivalently, in accordance with von Neumann's spectral theory, by self-adjoint operators);
2. the dynamical evolution of the coupled pair 'system + apparatus' is unitary.

Some attempts to generalise quantum measurements have concentrated on the first assumption, as I have briefly touched upon in the previous chapter:

- 1'. in the mathematical description of observables, generalise from PV-measures (sharp QM) to POV-measures (unsharp QM).

POV-measures, or positive operator measures, are based on the class of so-called effect operators (operators which are linear, positive, and bounded by the identity); they contain as sub-class the family of all projectors, and in this sense the unsharp formulation of QM is an enlargement of the standard sharp one.

It is also possible to discuss a weakening of (2),

- 2'. in the mathematical description of evolution, generalise from unitary dynamics to a non-unitary one.

These generalisations, as was briefly mentioned in the previous chapter, are limited by the following well-known result of Naimark (see [72, Appendix]):

- for any POV-measure on the Hilbert space \mathcal{H} , there exists a larger Hilbert space $\hat{\mathcal{H}}$ (\mathcal{H} turns out to be a subspace of $\hat{\mathcal{H}}$) in which the POV-measure is represented by a PV-measure;
- for any contraction operator (one parameter semi-group of contractions) on the Hilbert space \mathcal{H} , there exists a larger Hilbert space $\hat{\mathcal{H}}$ (\mathcal{H} turns out to be a subspace of $\hat{\mathcal{H}}$) in which the contraction(s) is (are) unitary operator(s).

This theorem raises the question as to whether the generalisation of the unitary evolutions to non-unitary evolutions should be understood in the context of larger systems whose evolutions are unitary; in other words, we might ask whether the apparent non-unitary evolutions are a result of \mathcal{H} becoming entangled with $\hat{\mathcal{H}} \ominus \mathcal{H}$, where $\hat{\mathcal{H}} \ominus \mathcal{H}$ models an environment, or whether the interaction is just irreducibly non-unitary. Similar questions can be asked about POV-measures. The present chapter, however, is not concerned with whether the physics of measurement interactions is best described by unitary operators or by other operators. Assuming that interactions are described by unitary operators, given a specific quantum measurement I show that there are many unitary operators satisfying the standard conditions for an interaction operator for such measurement: they are the ones that share the common mathematical “core” W , a partial isometric contraction which I will shortly define. I also show that, in some more general circumstances, no such unitary operators can be defined. I do not discuss here whether this implies that it is appropriate to abandon the idea of unitary evolution. The physical arguments on this are unclear (for more on this in the context of quantum mechanics see, for instance, Davies [25]; for a discussion of this in the context of the quantum measurement problem, see for example Cartwright [18, Essay 9]); on the other hand the question is surely underdetermined: there are many more options than just rejecting unitarity.

Note that, when the question of objectification is discussed, the previous chapter has shown that POV-measures cannot, in most cases, help to overcome insolubility results. Whether assuming $2'$ can help with the problem of satisfying objectification will be briefly considered later on. The chapter, in any case, does consider weaker versions of (2) in a context sufficiently general so as to not impede unsharp generalisations of (1).

A measurement interaction for an initial *vector* apparatus state will be shown to be minimally described by a unique linear contraction $W : \mathcal{H}_S \otimes \mathcal{H}_M \rightarrow \mathcal{H}_S \otimes \mathcal{H}_M$, with $\|W\| \leq 1$, which is also a partial isometry on a subspace \mathcal{M} of $\mathcal{H}_S \otimes \mathcal{H}_M$. As a consequence we have that for any partial “conjugate” isometry $W^\perp : \mathcal{H}_S \otimes \mathcal{H}_M \rightarrow \mathcal{H}_S \otimes \mathcal{H}_M$ on the subspace \mathcal{M}^\perp , i.e., such that $W(\mathcal{M}) \perp W^\perp(\mathcal{M}^\perp)$ and $W(\mathcal{M}) \oplus W^\perp(\mathcal{M}^\perp) = \mathcal{H}_S \otimes \mathcal{H}_M$, the operator $W + W^\perp : \mathcal{H}_S \otimes \mathcal{H}_M \rightarrow \mathcal{H}_S \otimes \mathcal{H}_M$ is unitary and realises the same measurement interaction. W^\perp , in a sense, ‘completes’ the partial isometry to a unitary operator. In what follows, the question of how minimal operators realising

the measurement process might be so completed will be analysed in several different contexts: this is because, while the minimal requirement of satisfying the probability reproducibility condition is achieved by ‘smaller’ operators, unitarity of the measurement remains an important condition in the debates on the quantum measurement problem.

For the case when $\dim(\mathcal{H}_M) > 2$ there is an infinite class of conjugate, contractive, partial isometries $\hat{W}_1^\perp, \hat{W}_2^\perp, \dots$ such that each of them behaves as the null operator on the vectors describing the prepared initial object + apparatus combination, namely on $\mathcal{H}_S \otimes [\psi_0]$ (note that $[\psi]$ denotes the closed linear subspace spanned by the vector ψ , while $\{\psi_1, \psi_2, \dots, \psi_n\}$ denotes the closed linear subspace spanned by the set of vectors $\{\psi_1, \psi_2, \dots, \psi_n\}$). It is then possible to construct the operators $U_1 = W \oplus \hat{W}_1^\perp, U_2 = W \oplus \hat{W}_2^\perp, \dots$ which are all unitary. Theorem 3.2 will show how to do this.

There are some important physical ideas here, which are not fully investigated in the present chapter, nor in the standard literature. We are used to thinking of the physical content of a quantum mechanical process as being described by the interaction Hamiltonian. This Hamiltonian is not normally calculated explicitly for measurement interactions. The present account of quantum measurement shows that there are going to be many physical processes, all described by different Hamiltonians, which will yield a measurement and suggests that the physically interesting information about measurement may be found in the common core to all these processes, that is our partial isometry W .

Furthermore, it will emerge from the discussion that the partial isometry W depends critically on the initial state of the apparatus. This is a well-known fact, it is for example certainly clear in the analysis that Beltrametti et al. give [6]. It implies that the interaction Hamiltonian itself depends on such a state: for different initial apparatus states, the interactions are different. This is a peculiarity of the measurement process as it is standardly characterised, and one whose significance for interpretive questions of quantum measurement has been little explored. It is nevertheless a consequence of assuming the probability reproducibility condition, as will be clear in the next section. I will not attempt to criticise this feature of measurements here, for it would involve an analysis of the exact structure of Hamiltonians generating the unitary measurements discussed here, and a discussion of the physical relevance of such Hamiltonians. However it is at least important to note that this is unusual with respect to how quantum

mechanics usually works. The Hamiltonian for the particle in a box does not depend on the initial state of the system it describes in any way; this feature of the system is not relevant in determining its energy. In the case of a quantum measurement, on the other hand, a different initial apparatus state, though *not* a different initial object state, implies that the energy of the combined measurement interaction is different: change the initial apparatus states, and the combined system energy eigenstates change. This means that, if we were to observe the process of measuring a given object observable in cases of different initial apparatus states, we would be observing something totally different.

What we can say is that the project of characterising unitary measurement operators satisfying certain further assumptions is not in and of itself a project which aims to characterise how measurements really work, as has been argued in chapter 1. The latter is a different, and further question, one which explicitly addresses the way in which theory connects with what happens in real experiments. I would claim that the physical and philosophical justification for this property of measurement interactions has not been explored, as far as I know, but does not invalidate the first project. On the other hand, any solution of the measurement problem which claims to mirror and rely on what happens in all real measurements would require some account of it.

The next step is to consider possible extensions of the linear contraction W to operations on the space $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ of positive trace class one operators on $\mathcal{H}_S \otimes \mathcal{H}_M$, the density operators.

The possibility of bijective mappings having certain properties will be investigated. The requirement of bijectivity of operations defined on density operators is an obvious one if we wish to generalise the idea of a unitary operator on a Hilbert space. A map $f : A \rightarrow B$, with A and B any two sets, is bijective if and only if it is one-to-one and onto. f is one-to-one if and only if, given any two $x, y \in A$, $x \neq y$, it follows that $f(x) \neq f(y)$. It is onto if for any $y \in B$ there exists an x in A such that $y = f(x)$. All unitary maps on a Hilbert space are bijections, and bijectivity of a mapping $f : A \rightarrow B$ captures the possibility of constructing an inverse mapping $f^{-1} : B \rightarrow A$ which is well defined on all of B . This property is important, in the context of quantum mechanics, for distinguishing the dynamical properties of unitary operators U as opposed to projection operators P : the former describe reversible interactions in the sense that, given U , it is possible to construct an operator U^{-1} which will enable us to tell, in a sense, from

where any state has originated, while this is not possible in the case of a projection operator like P . Bijectivity is a minimal condition for mirroring unitarity, on the other hand, because it doesn't make any allowances for the other important property of unitary operators, preservation of inner product of states. An additional condition for operations might be that they preserve the trace. Nevertheless, any result which we prove for a bijective operation will be valid for a bijective, trace-preserving one, too.

As well as more usual extensions, it is interesting to investigate a mapping W^1 which preserves the trace of the density operators it acts on in a particular subset of the set $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$, namely $\mathcal{T}_1^+(\mathcal{H}_S) \otimes P_{[\psi_0]}$. In general the trace of the operator produced by the action of W^1 will be less than or equal to 1. It is possible to extend W^1 to a trace preserving, but not bijective, mapping on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$, unlike for the operator W . On the other hand, it is easy to see that W^1 satisfies objectification of the pointer observable. Comparison of W^1 with standard extensions of premeasurements to the density operator framework will offer yet another view of insolubility of the quantum measurement problem.

Before moving on to address these questions, it is important to review the reasons for assuming the probability reproducibility condition as a condition for measurements. There are two, at least. First of all, intuitively this condition ensures that the probabilities associated with the individual object states are recovered in the apparatus. In particular, measurement of an eigenstate of an object observable will lead to the pointer observable being in an eigenstate. While this is certainly sufficient to ensure the relevance of this condition, should it be necessary to give it such a central place in the analysis of unitary measurements?

At the very least, we might expect a measurement transferring information from the object to the pointer system to misfire occasionally. Perhaps we shouldn't ask for the probabilities associated with an object system to be reproduced exactly, but only approximately. On the other hand, if the measurement is to count as measurement of an object observable O , it should distinguish initial object states which yield different probabilities with respect to O . This is a possible rationale behind Fine's probabilistic condition, briefly discussed in the previous Chapter. Fine's condition is an interesting one in the context of insolubility proofs, because of its generality: it is implied by all other probabilistic conditions on measurements, so a no-go theorem holding for this condition covers all reasonable cases of measurement. For the discussion in this chapter,

however, it is easily seen to be defective as a condition for the measurement of an object observable. Consider an object observable O_S , a pointer observable A_M , an operation $U : \mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M) \rightarrow \mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$, a pair of initial object states $T, T' \in \mathcal{T}_1^+(\mathcal{H}_S)$ and an initial apparatus state $T_M \in \mathcal{T}_1^+(\mathcal{H}_M)$. Fine [35] proposes two conditions to impose on U :

$$(1.1a) \quad p_T^{O_S} \neq p_{T'}^{O_S} \implies p_{\mathcal{R}_M(U(T \otimes T_M))}^{A_M} \neq p_{\mathcal{R}_M(U(T' \otimes T_M))}^{A_M}$$

or, in addition to 1.1a,

$$(1.1b) \quad p_T^{O_S} = p_{T'}^{O_S} \implies p_{\mathcal{R}_M(U(T \otimes T_M))}^{A_M} = p_{\mathcal{R}_M(U(T' \otimes T_M))}^{A_M},$$

where P_T^O is the probability of observable O in state T and $\mathcal{R}_M : \mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M) \rightarrow \mathcal{T}_1^+(\mathcal{H}_M)$ is the operation of partial trace, ‘taking out’ the object space. A measurement is called an O -discrimination if it satisfies 1.1a, it is called an O -filter if it satisfies 1.1a and 1.1b.

The problem with these conditions, as Busch et al. show [14, p. 30], is that they allow a U which satisfies these conditions *and* is a measurement of many observables, including non-commuting ones. In the extreme case, a discrimination of any informationally complete observable¹ will be a discrimination of all observables. As for filters, clearly an O -filter is a Q -filter for any observable Q which is informationally equivalent to O , i.e. any observable for which, if $p_T^Q = p_{T'}^Q$, then $p_T^O = p_{T'}^O$ for any pair of states T and T' . Fine himself [35, pp. 116–117] makes a number of criticisms: while he is trying to establish what could be the most general probabilistic conditions for a measurement he is aware of several potential difficulties for discriminations and filters. As a consequence of this, it seems reasonable to concentrate the analysis in this chapter on the probability reproducibility condition. Given, however, that I will be highlighting some of the problems that can arise with it, perhaps at some point it would be reasonable to reassess the role of Fine’s more general conditions.

¹An informationally complete observable O is an observable which separates all states, i.e. for any T and T' , $p_T^O \neq p_{T'}^O$. No sharp observable is informationally complete, but there exist unsharp informationally complete observables.

2. The contractive partially isometric evolution and its completion in system-apparatus measurement interactions

Consider a measurement interaction of standard measurement theory involving a system \mathbb{S} and a measurement device \mathbb{M} . The quantum mechanical description is mathematically realised by two complex Hilbert spaces with dimensions equal to n and m , $\mathcal{H}_{\mathbb{S}}$ and $\mathcal{H}_{\mathbb{M}}$ respectively with $\dim(\mathcal{H}_{\mathbb{S}}) = n \leq \dim(\mathcal{H}_{\mathbb{M}}) = m$. Let $\{\varphi_1, \dots, \varphi_n\}$ be an orthonormal basis of vectors in $\mathcal{H}_{\mathbb{S}}$ (which picks out an object observable to be measured) such that φ_i describes a state of the system. The measurement apparatus \mathbb{M} is initialised in a fixed *starting* state described by the vector $\psi_0 \in \mathcal{H}_{\mathbb{M}}$, normalised so that $\|\psi_0\| = 1$. Having chosen the basis $\{\varphi_1, \dots, \varphi_n\}$ and the starting state ψ_0 , select an orthonormal basis $\{\psi_1, \dots, \psi_m\}$ of vectors in $\mathcal{H}_{\mathbb{M}}$ (which picks out an apparatus pointer observable) and a set of vectors $\tilde{\varphi}_1, \dots, \tilde{\varphi}_n \in \mathcal{H}_{\mathbb{S}}$, with $\|\tilde{\varphi}_1\|_{\mathcal{H}_{\mathbb{S}}} = \dots = \|\tilde{\varphi}_n\|_{\mathcal{H}_{\mathbb{S}}} = 1$.

The question is now whether there exists a (not necessarily unitary) linear operator

$$W^{\mathbb{S}, \mathbb{M}} : \mathcal{H}_{\mathbb{S}} \otimes \mathcal{H}_{\mathbb{M}} \rightarrow \mathcal{H}_{\mathbb{S}} \otimes \mathcal{H}_{\mathbb{M}}$$

which mathematically realises the measurement interaction described by the “probability reproducibility condition” [6, Eq. 1]: for any vector $\varphi \in \mathcal{H}_{\mathbb{S}}$, and any eigenstate of the measured observable φ_i , measured by the pointer eigenstate ψ_i ,

$$(2.1) \quad \langle \varphi | P_{[\varphi_i]} \varphi \rangle = \left\langle W^{\mathbb{S}, \mathbb{M}}(\varphi \otimes \psi_0) \left| (I_{\mathbb{S}} \otimes P_{[\psi_i]}) (W^{\mathbb{S}, \mathbb{M}}(\varphi \otimes \psi_0)) \right. \right\rangle.$$

As a consequence, such an operator will satisfy the “interaction conditions” (calibration conditions in the terminology of Busch et al. [14])

$$(2.2) \quad i = 1, \dots, n \quad W^{\mathbb{S}, \mathbb{M}}(\varphi_i \otimes \psi_0) = \tilde{\varphi}_i \otimes \psi_i,$$

$\tilde{\varphi}_i$ any vector in $\mathcal{H}_{\mathbb{S}}$. In fact, suppose $\varphi = \varphi_j$, so that $\langle \varphi | P_{[\varphi_i]} \varphi \rangle = 0$, or 1 when $i = j$; then it must be that $\langle W^{\mathbb{S}, \mathbb{M}}(\varphi \otimes \psi_0) | (I_{\mathbb{S}} \otimes P_{[\psi_i]}) (W^{\mathbb{S}, \mathbb{M}}(\varphi \otimes \psi_0)) \rangle = 0$, or 1 when $\psi_i = \psi_j$. This implies the interaction conditions for $j = 1, \dots, n$. By linearity and (2.2) it follows immediately that $\forall \varphi \in \mathcal{H}_{\mathbb{S}}$, $W^{\mathbb{S}, \mathbb{M}}(\varphi \otimes \psi_0) = \sum_{i=1}^n \langle \varphi_i \otimes \psi_0 | \varphi \otimes \psi_0 \rangle (\tilde{\varphi}_i \otimes \psi_i)$, and more generally

$$(2.3) \quad W^{\mathbb{S}, \mathbb{M}}(\cdot) = \sum_{i=1}^n \langle \varphi_i \otimes \psi_0 | \cdot \rangle (\tilde{\varphi}_i \otimes \psi_i).$$

The mapping $W^{\mathbf{S},\mathbf{M}}$ is clearly dependent, in a crucial way, on the initial apparatus state ψ_0 , and in a way that is entirely a consequence of the probability reproducibility condition.

2.1. Beltrametti, Cassinelli and Lahti's Characterisation Theorem. Beltrametti, Cassinelli and Lahti [6] have proved a well-known characterisation theorem for unitary measurements of discrete quantities. There are several similarities between their paper and this section, so it is worth reviewing their work and explaining in what way the present material differs and is more general.

Beltrametti et al. essentially show that a mapping like (2.3) exists and can be extended to a unitary mapping (possibly in a non-unique way), and then discuss informational properties of these mappings (conditions under which these mappings are of the first kind, repeatable, Von Neumann-Lüders) and problems with the objectification condition and the Wigner-Araki-Yanase result.

The following work begins by restating the result about unitary mappings, with some differences. The first is that it is shown that the mapping (2.3) is a partial isometry, a result not previously established; also much more detail is given of the ways in which completions are constructed, and there is no restriction on representations of the measurement apparatus to spaces of dimension equal to the cardinality of the discrete spectrum of the measured observable.

The latter difference merits some comments. Beltrametti et al. make the assumption that if the discrete spectrum of a self-adjoint operator has cardinality n , then, for $\mathcal{H}_{\mathbf{M}}$ the apparatus space, $\dim(\mathcal{H}_{\mathbf{M}}) = n$. This seems harmless enough, yet the construction of the unitary extension leads to many different problems for the case in which $\dim(\mathcal{H}_{\mathbf{M}}) > \dim(\mathcal{H}_{\mathbf{S}})$. Given that the latter relation is what we would usually expect to be the case when we speak of apparatuses as macroscopic objects (indeed it would often be the case that $\dim(\mathcal{H}_{\mathbf{M}}) \gg \dim(\mathcal{H}_{\mathbf{S}})$), this case merits further attention. It also plays a crucial role in the further generalisation that is considered here, where the initial state of the apparatus is taken to be a nonpure density operator. Several problems emerge here for the quantum theory of measurement, and the impossibility of constructing an operation satisfying the probability reproducibility condition is related to some instances of the case when $\dim(\mathcal{H}_{\mathbf{M}}) > \dim(\mathcal{H}_{\mathbf{S}})$.

The definitions and the notation introduced at the beginning of the section differ from those put forward by Beltrametti et al.: they work with orthogonal projectors $P_{[\varphi_1]}, \dots, P_{[\varphi_n]}$ which are spectral resolutions of the identity, so that $\sum_i P_{[\varphi_i]} = I_S$. Further, they index the values for the observables with two indices, the second tracking the multiplicity of the first index.

Given our framework, it is certainly possible to construct, from the basis $\{\varphi_1, \dots, \varphi_n\}$ of the Hilbert space describing the object system and a corresponding value set $\{a_1, \dots, a_n\}$ of a measured system observable A_S , the self-adjoint object operator A_S as

$$A_S = a_1 P_{[\varphi_1]} + \dots + a_n P_{[\varphi_n]},$$

so that in the notation of spectral representations one has $P_{A_S}(a_i) = P_{[\varphi_i]}$, the event testing if “a measurement of the observable A_S yields the result a_i ”. Furthermore, to the basis $\{\psi_1, \dots, \psi_m\}$ of the Hilbert space describing the apparatus system one can associate an observable A_M , with a set of values $\{m_1, \dots, m_m\}$ for the pointer observable, such that ψ_i mathematically describes the state in which a test of the observable M yields the value m_i with certainty.

The “pointer function” f which correlates the value sets of the system observable and the pointer observable is

$$(2.4) \quad f : \{a_1, \dots, a_n\} \longrightarrow \{m_1, \dots, m_m\}, a_i \longrightarrow f(a_i) = m_i.$$

The pointer function does not usually play a major role in the discussion of quantum measurements: it doesn't play a role in the dynamics of measurements, and is in a sense defined externally of quantum mechanics. In particular, no measurement scheme such as those discussed in this chapter enables recovery of this function. On the other hand it is clear that *only* through such a pointer function can we actually say which specific observable is being measured

2.2. A Characterisation Result for the Case of an Initial Vector State.

The general mathematical requirements needed in order to establish the main result of this section can be spelt out as follows:

1. The dimension of the Hilbert space \mathcal{H}_S describing the system and of the Hilbert space \mathcal{H}_M describing the apparatus are in the relation:

$$\dim(\mathcal{H}_S) \leq \dim(\mathcal{H}_M)$$

In what follows I will assume $I = \{1, 2, \dots, \dim(\mathcal{H}_S) = n\}$ and $J = \{1, 2, \dots, \dim(\mathcal{H}_M) = m\}$, so that clearly $I \subseteq J$;

2. $\{\varphi_i\}_{i \in I}$ is an orthonormal basis of \mathcal{H}_S ;
3. $\{\psi_0^j\}_{j \in J}$ is an orthonormal basis of \mathcal{H}_M ;
4. (a) $\{\hat{\psi}_i\}_{i \in I}$ is an orthonormal family in \mathcal{H}_M
 (b) $\forall j \in I$ and $\forall j' \in J - I$ there exists a $\hat{\psi}_{j'}$ such that $\langle \hat{\psi}_j | \hat{\psi}_{j'} \rangle = 0$ and $\{\hat{\psi}_j\}_{j \in J}$ is an orthonormal basis in \mathcal{H}_M .
5. (a) $\forall j \in I$, $\{\tilde{\varphi}_i^j\}_{i \in I}$ are normalised (not necessarily orthogonal) families from \mathcal{H}_S
 (b) $\forall i \in I, \forall h, k \in I, h \neq k, \langle \tilde{\varphi}_i^h | \tilde{\varphi}_i^k \rangle = 0$.

Under the notation just specified the interaction conditions assume the form, for all $i, j \in I$,

$$(2.5) \quad W^j(\varphi_i \otimes \psi_0^j) = \tilde{\varphi}_i^j \otimes \hat{\psi}_i.$$

Also assume for the moment that, $\forall i \in I$ and $j \in J - I$. Given any orthonormal family $\{v_i\}$ in \mathcal{H}_S , let

$$W^j(\varphi_i \otimes \psi_0^j) = v_i \otimes \hat{\psi}_j.$$

Informally, these requirements fix a number of points concerned mainly with the case where $\dim(\mathcal{H}_S) < \dim(\mathcal{H}_M)$. (2) and (3) fix two observables, one on \mathcal{H}_S which is the observable to be subjected to measurement, and one (of the generally infinitely possible ones) on \mathcal{H}_M , which has as one of its eigenstates the ‘ready-to-measure’ state. In characterising a unitary operator I make use of the fact that it maps an orthonormal basis for the space $\mathcal{H}_S \otimes \mathcal{H}_M$ to another orthonormal basis. (2) and (3) fix an ‘initial’ orthonormal basis for the composite space $\mathcal{H}_S \otimes \mathcal{H}_M$; hence the indexation of the initial ‘ready-to-measure’ apparatus state: we still need a basis for \mathcal{H}_M in order to characterise a basis for $\mathcal{H}_S \otimes \mathcal{H}_M$, in order to characterise a unitary operator $U : \mathcal{H}_S \otimes \mathcal{H}_M \rightarrow \mathcal{H}_S \otimes \mathcal{H}_M$. The fact that, for $\dim(\mathcal{H}_M) > 2$, the possible basis in (3) is fixed only by such an initial apparatus state implies that many unitary operators will realise a measurement of the observable selected through (2), one for each different basis for \mathcal{H}_M that can be chosen in (3). The further fact that W^j , as implied by the probability reproducibility condition, requires only specification of an initial apparatus state should

already be a clue to the fact that this condition is not sufficient to characterise a unitary operator.

Conditions (4a) and (4b) select a set of orthonormal pointer states in the apparatus space \mathcal{H}_M , which in the first instance, barring pointer degeneracies, will have cardinality bounded by the dimension of the space \mathcal{H}_S . In the case in which $\dim(\mathcal{H}_S) < \dim(\mathcal{H}_M)$, this orthogonal set must be completed to an orthonormal basis, as is done in (4b).

(5a) makes it clear that, in principle, we can't rule out, given an initial 'ready-to-measure' apparatus state ψ_0^1 , the possibility of the final object states resulting from the measurement interaction not being mutually orthogonal (as in the by now familiar case of destructive measurements discussed in Chapter 1). However, (5b) states that, given two mutually orthogonal states in the set described in (3), the final object states resulting from the measurement interactions *must* be orthogonal. This is because, if we couple such states ψ_0^1 and ψ_0^2 with an element of the set in (2) (an eigenstate of the measured observable), the probability reproducibility condition forces the final pointer state to be the same for both initial states obtained in this way. Then, in order to preserve the orthogonality of the two initial states that is demanded by the unitarity of the interaction, the final object states must themselves be orthogonal.

In the next theorem (which is work carried out jointly with G. Cattaneo [19]) I will refer to the following

LEMMA 3.1. *Under the above conditions, both*

$$(2.6a) \quad \{\varphi_1 \otimes \psi_0^1, \dots, \varphi_n \otimes \psi_0^1\}$$

$$(2.6b) \quad \{\tilde{\varphi}_1^1 \otimes \hat{\psi}_1, \dots, \tilde{\varphi}_n^1 \otimes \hat{\psi}_n\}$$

and

$$(2.7a) \quad \{\varphi_1 \otimes \psi_0^k, \dots, \varphi_n \otimes \psi_0^k\}_{\substack{k \in J \\ k \neq 1}}$$

$$(2.7b) \quad \{\tilde{\varphi}_1^k \otimes \hat{\psi}_1, \dots, \tilde{\varphi}_n^k \otimes \hat{\psi}_n\}_{\substack{k \in I \\ k \neq 1}}$$

$$(2.7c) \quad \{v_1 \otimes \hat{\psi}_j, \dots, v_n \otimes \hat{\psi}_j\}_{j \in J-I}$$

are orthonormal (not complete) families of vectors in $\mathcal{H}_S \otimes \mathcal{H}_M$. The union of sets (2.6a) and (2.7a), and of (2.6b), (2.7b) and (2.7c) form an orthonormal basis.

The proof is trivial, following from the points under condition 4 above. Completeness for the first union follows from conditions 2 and 3, while for the second they follow from 5, which establishes a one-to-one correspondence between the two union sets.

THEOREM 3.2. *For an apparatus in the starting state ψ_0^1*

(i) *the operator*

$$(2.8) \quad W^1 = \sum_{i \in I} \langle \varphi_i \otimes \psi_0^1 | (\cdot) \rangle (\tilde{\varphi}_i^1 \otimes \hat{\psi}_i)$$

satisfies the probability reproducibility condition (2.1) (and therefore the interaction conditions (2.5)); it is contractive and partially isometric on the subspace $\mathcal{H}_S \otimes [\psi_0^1]$ and null on the subspace $\mathcal{H}_S \otimes [\psi_0^1]^\perp$.

(ii) *The conjugate premeasurement operator for W^1 , fixed by the base selection in condition (3), is*

$$(2.9) \quad W^{1,\perp} := \sum_{\substack{k \in I \\ k \neq 1}} W^k = \sum_{\substack{k \in I \\ k \neq 1}} \left(\sum_{i \in I} \langle \varphi_i \otimes \psi_0^k | (\cdot) \rangle (\tilde{\varphi}_i^k \otimes \hat{\psi}_i) \right)$$

and is itself a partial isometry on the space $\mathcal{H}_S \otimes [\{\psi_0^k\}_{\substack{k \in I \\ k \neq 1}}]$.

(iii) *Given an orthonormal set $\{v_i\}_{i \in I}$ and the operator*

$$W^{1'} := \sum_{i \in I, k \in J-I} \langle \varphi_i \otimes \psi_0^k | (\cdot) \rangle (v_i \otimes \hat{\psi}_k),$$

itself a partial isometry on $\mathcal{H}_S \otimes [\{\psi_0^k\}_{k \in J-I}]$, the operator

$$(2.10) \quad W = W^1 + W^{1,\perp} + W^{1'} = \sum_{\substack{i \in I \\ k \in I}} \langle \varphi_i \otimes \psi_0^k | (\cdot) \rangle (\tilde{\varphi}_i^k \otimes \hat{\psi}_i) \\ + \sum_{\substack{i \in I \\ k \in J-I}} \langle \varphi_i \otimes \psi_0^k | (\cdot) \rangle (v_i \otimes \hat{\psi}_k)$$

is unitary.

PROOF. See Appendix A. □

REMARK 3.1. *As noted on page 103, the probability reproducibility condition implies that operators satisfying it have the form (2.3) when restricted to the subspace of $\mathcal{H}_S \otimes \mathcal{H}_M$ selected by the initial ready-to-measure state. The previous Theorem shows the converse, that all operators of the form (2.3) satisfy the probability reproducibility condition and are in fact contractive partial isometries*

on the subspace. Their unitary extensions characterise measurement operators precisely because they also satisfy the probability reproducibility condition.

REMARK 3.2. It is important to appreciate that, for any (arbitrary) completion

$$\begin{aligned} \{u_{11} = \varphi_1 \otimes \psi_0^1, \dots, u_{n1} = \varphi_n \otimes \psi_0^n, \dots, u_{ij}, \dots, u_{nm}\}_{\substack{i \in I - \{1\} \\ j \in J - \{1\}}} \\ \{\tilde{u}_{11} = \tilde{\varphi}_1 \otimes \psi_1, \dots, \tilde{u}_{21} = \tilde{\varphi}_2 \otimes \psi_2, \dots, \tilde{u}_{ij}, \dots, \tilde{u}_{nm}\}_{\substack{i \in I - \{1\} \\ j \in J - \{1\}}} \end{aligned}$$

the linear operator

$$U : \mathcal{H}_S \otimes \mathcal{H}_M \rightarrow \mathcal{H}_S \otimes \mathcal{H}_M$$

defined by the law

$$(2.11) \quad U(\cdot) := \sum_{\substack{i \in I - \{1\} \\ j \in J}} \langle u_{ij} | \cdot \rangle \tilde{u}_{ij} = W^j(\cdot) + \sum_{\substack{i \in I - \{1\} \\ j \in J - \{1\}}} \langle u_{ij} | \cdot \rangle \tilde{u}_{ij}$$

is unitary and satisfies the interaction rules (2.5) and the probability reproducibility condition (2.1). There are, obviously, as many of these unitary operators realising a unitary measurement for a given initial state ψ_0^1 as there are completions. The particular construction of the unitary operator W discussed in Theorem 3.2 is motivated in section 5, when the possibility of measurements with initial states given by density operators is considered: it is a construction where the components in $W^{1,\perp}$ are themselves partial isometries which are premeasurements of the object observable. Then the probability reproducibility condition implies generalised interaction conditions: more initial ready-to-measure vector states are fixed through the initial density operator, and this fixes more than one of the partial isometries which make up the unitary measurements.

It is worth noting that this construction does not exhaust all the possible extensions. For example, suppose $W^{1,\perp} := \sum_{i \in I, k \in I, k \neq 1} \langle \varphi_i \otimes \psi_0^k | (\cdot) \rangle (v_i \otimes \hat{\psi}_k)$. This is clearly not a premeasurement, as it does not satisfy the probability reproducibility condition. This can be easily seen by its failure to satisfy the interaction conditions 2.5.

REMARK 3.3. The completion discussed in part (iii) of the Theorem details a W^j , part of $W^{1'}$, which is not a premeasurement when $j \in J$. Again this need not be the case in general. The specific situations where there exists a W^j which will necessarily fail to be a premeasurement will be outlined in section 5, together with the problems which arise as a consequence of this.

2.3. Some Examples.

EXAMPLE 3.1. — *The (2,3)-dimensional case.* I give a simple example of the above operator in order to illustrate its construction for the case in which $\dim(\mathcal{H}_S) = 2$ and $\dim(\mathcal{H}_M) = 3$. We have the explicit formula:

$$(2.12) \quad W^1 = \langle \varphi_1 \otimes \psi_0^1 | (\cdot) \rangle (\tilde{\varphi}_1^1 \otimes \hat{\psi}_1) + \langle \varphi_2 \otimes \psi_0^1 | (\cdot) \rangle (\tilde{\varphi}_2^1 \otimes \hat{\psi}_2)$$

describing the perfect measurement interaction:

$$(2.13a) \quad W^1 (\varphi_1 \otimes \psi_0^1) = \tilde{\varphi}_1^1 \otimes \hat{\psi}_1$$

$$(2.13b) \quad W^1 (\varphi_2 \otimes \psi_0^1) = \tilde{\varphi}_2^1 \otimes \hat{\psi}_2$$

The unitary operator $W : \mathcal{H}_S \otimes \mathcal{H}_M \mapsto \mathcal{H}_S \otimes \mathcal{H}_M$ is the sum

$$\begin{aligned} W &= W^1 + \left[\sum_{h=2,3} W^h \right] \\ &= \langle \varphi_1 \otimes \psi_0^1 | (\cdot) \rangle (\tilde{\varphi}_1^1 \otimes \hat{\psi}_1) & + \langle \varphi_2 \otimes \psi_0^1 | (\cdot) \rangle (\tilde{\varphi}_2^1 \otimes \hat{\psi}_2) \\ &+ \langle \varphi_1 \otimes \psi_0^2 | (\cdot) \rangle (\tilde{\varphi}_1^2 \otimes \hat{\psi}_1) & + \langle \varphi_2 \otimes \psi_0^2 | (\cdot) \rangle (\tilde{\varphi}_2^2 \otimes \hat{\psi}_2) \\ &+ \langle \varphi_1 \otimes \psi_0^3 | (\cdot) \rangle (v_1 \otimes \hat{\psi}_3) & + \langle \varphi_2 \otimes \psi_0^3 | (\cdot) \rangle (v_2 \otimes \hat{\psi}_3) \end{aligned}$$

The unitarity follows immediately from the fact that the two families

$$\begin{aligned} &\{\varphi_1 \otimes \psi_0^1, \varphi_2 \otimes \psi_0^1, \varphi_1 \otimes \psi_0^2, \varphi_2 \otimes \psi_0^2, \varphi_1 \otimes \psi_0^3, \varphi_2 \otimes \psi_0^3\} \\ &\{\tilde{\varphi}_1^1 \otimes \hat{\psi}_1, \tilde{\varphi}_2^1 \otimes \hat{\psi}_2, \tilde{\varphi}_1^2 \otimes \hat{\psi}_1, \tilde{\varphi}_2^2 \otimes \hat{\psi}_2, v_1 \otimes \hat{\psi}_3, v_2 \otimes \hat{\psi}_3\} \end{aligned}$$

are orthonormal bases of $\mathcal{H}_S \otimes \mathcal{H}_M$, the latter following in a crucial way from condition (5b) on page 106.

EXAMPLE 3.2. The special case of von Neumann-Lüders measurements is characterised by the fact that $W_L(\varphi_i \otimes \psi_0^1) = \varphi_i \otimes \psi_i$. If we assume that $\dim(\mathcal{H}_S) = n \leq m =$

$\dim(\mathcal{H}_M)$, W_L can be obtained by considering for example

$$\begin{aligned}
 W_L(\varphi_1 \otimes \psi_0^1) &= \varphi_1 \otimes \psi_1 & \dots & & W_L(\varphi_n \otimes \psi_0^1) &= \varphi_n \otimes \psi_n \\
 &\vdots & & & \vdots & \\
 W_L(\varphi_1 \otimes \psi_0^n) &= \varphi_n \otimes \psi_1 & \dots & & W_L(\varphi_n \otimes \psi_0^n) &= \varphi_1 \otimes \psi_n \\
 W_L(\varphi_1 \otimes \psi_0^{n+1}) &= \varphi_1 \otimes \psi_{n+1} & \dots & & W_L(\varphi_n \otimes \psi_0^{n+1}) &= \varphi_n \otimes \psi_{n+1} \\
 &\vdots & & & \vdots & \\
 W_L(\varphi_1 \otimes \psi_0^m) &= \varphi_1 \otimes \psi_m & \dots & & W_L(\varphi_n \otimes \psi_0^m) &= \varphi_n \otimes \psi_m
 \end{aligned}$$

corresponding to a unitary operator like the one in Theorem 3.2 under the assumptions $\tilde{\varphi}_i^j = \varphi_{\pi^{j-1}(i)} \forall i, j \in I$, $\tilde{\varphi}_i^j = \varphi_i \forall j \in J - I$ where $\pi^{j-1}(i)$ is the product of applying to the index i the $(j-1)^{\text{th}}$ power of the permutation $\pi = (1, 2, \dots, n)$. Again it is worth stressing that the completion here is arbitrary and the only relevant part for the measurement interaction is ψ_0^1 in the present case, but again the completions will play crucial roles when discussing the case of an initial apparatus state which is a nonpure density operator.

The following Lemma will be needed later.

LEMMA 3.3. *The adjoint of the operator W is the operator*

$$\begin{aligned}
 (2.14) \quad W^* &= \langle \tilde{\varphi}_1^1 \otimes \hat{\psi}_1 | (\cdot) \rangle (\varphi_1 \otimes \psi_0^1) + \dots + \langle \tilde{\varphi}_n^1 \otimes \hat{\psi}_n | (\cdot) \rangle (\varphi_n \otimes \psi_0^1) \\
 &+ \dots \\
 &+ \langle \tilde{\varphi}_1^m \otimes \hat{\psi}_1 | (\cdot) \rangle (\varphi_1 \otimes \psi_0^m) + \dots + \langle \tilde{\varphi}_n^m \otimes \hat{\psi}_n | (\cdot) \rangle (\varphi_n \otimes \psi_0^m). \\
 &+ \sum_{\substack{i \in I \\ k \in J - I}} \langle v_i \otimes \hat{\psi}_k | (\cdot) \rangle (\varphi_i \otimes \psi_0^k)
 \end{aligned}$$

PROOF. Consider the first component of W , i.e. $\langle \varphi_1 \otimes \psi_0^1 | (\cdot) \rangle (\tilde{\varphi}_1^1 \otimes \hat{\psi}_1)$.

We have that

$$\begin{aligned}
 \langle \Phi | \langle \varphi_1 \otimes \psi_0^1 | \Psi \rangle (\tilde{\varphi}_1^1 \otimes \hat{\psi}_1) \rangle &= \langle \varphi_1 \otimes \psi_0^1 | \Psi \rangle \langle \Phi | \tilde{\varphi}_1^1 \otimes \hat{\psi}_1 \rangle \\
 &= \left\langle \overline{\langle \Phi | \tilde{\varphi}_1^1 \otimes \hat{\psi}_1 \rangle} (\varphi_1 \otimes \psi_0^1) \middle| \Psi \right\rangle \\
 &= \left\langle \left\langle \tilde{\varphi}_1^1 \otimes \hat{\psi}_1 \middle| \Phi \right\rangle (\varphi_1 \otimes \psi_0^1) \middle| \Psi \right\rangle
 \end{aligned}$$

repeating the process for each component of W , by linearity we have the result. \square

In particular we have that, for the partial isometry W^1 which captures the measurement interaction for a given initial state ψ_0^1 , the adjoint is

$$(2.15) \quad W^{1*} = \sum_{i=1}^n \left\langle \hat{\varphi}_i^1 \otimes \hat{\psi}_i \middle| (\cdot) \right\rangle (\varphi_i \otimes \psi_0^1)$$

3. Partially isometric contractive mappings in the density operator formalism

In the next two sections the earlier discussion will be generalised to the density operator formalism, something not previously done in the literature. In particular I examine what sort of *operations* on the set of states, now taken to be density operators, are compatible with the probability reproducibility condition. First some notation is introduced. Let $\mathcal{H}_S \otimes \mathcal{H}_M$ be the space of the coupled object-apparatus system as defined in the previous section. Then define $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ as the set of all positive trace one operators on $\mathcal{H}_S \otimes \mathcal{H}_M$. $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ is the standard choice for the space of physical states; however it is easy to verify that most of the results that follow hold for the general case of action on the set $\mathcal{T}(\mathcal{H}_S \otimes \mathcal{H}_M)$ of trace class operators.

Clearly all the one-dimensional projection operators

$$P_{[\varphi \otimes \psi]} = \langle \varphi \otimes \psi | (\cdot) \rangle (\varphi \otimes \psi) = P_{[\varphi]} \otimes P_{[\psi]}$$

(where $\|\varphi\| = \|\psi\| = 1$) are positive operators of trace one. In general, for any arbitrary Hilbert space \mathcal{H} , one dimensional projectors are the extremal elements of the set $\mathcal{T}_1^+(\mathcal{H})$. This set is σ -convex with respect to its linear structure, so that an element $T \in \mathcal{T}_1^+(\mathcal{H})$ is extremal if the condition $T = wT_1 + \dots + w_nT_n$, with $T_1, \dots, T_n \in \mathcal{T}_1^+(\mathcal{H})$ and $0 < w_i < 1$, $\sum_i w_i = 1$ always implies that $T = T_1 = \dots = T_n$. The set of extremal elements exhausts the whole set $\mathcal{T}_1^+(\mathcal{H})$ in the sense that any $T \in \mathcal{T}_1^+(\mathcal{H})$ can be expressed as a convex combinations of some elements of this set. These elements are the set of states of the system in consideration, the projections being called pure states, and their (non-trivial) convex combinations mixed states.

An induced notion of linearity for $\mathcal{T}_1^+(\mathcal{H})$ is defined by considering a linear mapping $U : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$ on the Banach space of trace class operators $\mathcal{T}(\mathcal{H})$. Its restriction $U : \mathcal{T}_1^+(\mathcal{H}) \rightarrow \mathcal{T}_1^+(\mathcal{H})$, when well-defined, satisfies the corresponding *convex-preserving*

property, i.e. for all $a_1, \dots, a_n \in \mathbb{R}_+$, $a_1 + \dots + a_n = 1$, P_1, \dots, P_n one-dimensional projectors

$$U(a_1 P_1 + \dots + a_n P_n) = a_1 U(P_1) + \dots + a_n U(P_n).$$

Coming back to the space $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$, note the following result for the trace operation:

$$(3.1) \quad \text{Tr}_{\mathcal{H}_S \otimes \mathcal{H}_M}[T_S \otimes T_M] = \text{Tr}_{\mathcal{H}_S}[T_S] \cdot \text{Tr}_{\mathcal{H}_M}[T_M],$$

where $T_S \in \mathcal{T}_1^+(\mathcal{H}_S)$, $T_M \in \mathcal{T}_1^+(\mathcal{H}_M)$ and $T_S \otimes T_M \in \mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$.

The natural way to extend the partially isometric operator W^1 over $\mathcal{H}_S \otimes \mathcal{H}_M$ defined by (2.8) is through the standard conjugation action, formally characterised amongst others by Davies [25] for operations on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$. Consider first, however, the mapping $W^1 : \mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M) \rightarrow \mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$, for all $T \in \mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$, and $P_{[\psi_0]}$ the initial apparatus state:

$$(3.2) \quad W^1(T) = \sum_{i=1}^n \text{Tr}_{\mathcal{H}_S \otimes \mathcal{H}_M} [(P_{[\varphi_i \otimes \psi_0]})(T)] (P_{[\bar{\varphi}_i \otimes \psi_i]})$$

The mapping W^1 is clearly convex-preserving. It satisfies the interaction conditions, i.e., for $i = 1, \dots, n$, $W^1(P_{[\varphi_i \otimes \psi_0]}) = P_{[\bar{\varphi}_i \otimes \psi_i]}$, corresponding to (2.5). Moreover, for $T = P_{[\Phi]}$, $\Phi \in \mathcal{H}_S \otimes \mathcal{H}_M$, $\|\Phi\| = 1$,

$$(3.3) \quad W^1(P_{[\Phi]}) = \sum_{i=1}^n |\langle \varphi_i \otimes \psi_0 | \Phi \rangle|^2 P_{[\bar{\varphi}_i \otimes \psi_i]},$$

which is analogous to (2.8).

Suppose now that the observable selected by the orthonormal basis $\{\varphi_1, \dots, \varphi_n\}$ is being measured by an apparatus which is in the state $P_{[\psi_0]}$ and that $\{\psi_1, \dots, \psi_n\}$ is an orthonormal set in \mathcal{H}_M selecting a pointer observable for the measurement. Suppose the object being measured is in the state $P_{[\varphi]}$, where $\varphi = \alpha_1 \varphi_1 + \dots + \alpha_n \varphi_n$, $|\alpha_1|^2 + \dots + |\alpha_n|^2 = 1$. Hence the combined object system + measuring apparatus in $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ is $T = P_{[\varphi \otimes \psi_0]}$. The result of applying W^1 to T is, by (3.3)

$$(3.4) \quad W^1(P_{[\varphi \otimes \psi_0]}) = \sum_{i=1}^n |\alpha_i|^2 (P_{[\bar{\varphi}_i \otimes \psi_i]})$$

This is again in accordance with the probability reproducibility condition. This example usefully illustrates part (1) of the next theorem. The following lemma will be used in the proof.

LEMMA 3.4. Suppose $\{\varphi_1, \dots, \varphi_n\}$ is a complete orthonormal system in \mathbb{C}^n . Then

$$\text{Tr}[P_{[\varphi_i]} P_{[\alpha_1 \varphi_1 + \dots + \alpha_n \varphi_n]}] = |\alpha_i|^2,$$

$$\alpha_1, \dots, \alpha_n \in \mathbb{C}, |\alpha_1|^2 + \dots + |\alpha_n|^2 = 1.$$

The proof is straightforward and will not be given here.

THEOREM 3.5. The mapping W^1 has the following properties:

1. it satisfies the probability reproducibility condition in the sense that, for a state $T \in \mathcal{T}_1^+(\mathcal{H}_S)$ and for $j = 1, \dots, n$,

$$(3.5) \quad \text{Tr}_{\mathcal{H}_S}[T P_{[\varphi_j]}] = \text{Tr}_{\mathcal{H}_S \otimes \mathcal{H}_M} \left[W^1(T \otimes P_{[\psi_0]}) (I_S \otimes P_{[\psi_j]}) \right],$$

where I_S is the identity map on $\mathcal{T}(\mathcal{H}_S)$ and $P_{[\psi_0]}$ is, as usual, the initial state of the measuring apparatus.

2. it is a partial isometry, in the sense that it preserves the trace of the density operators in the subset of $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ defined as $\mathcal{T}_1^+(\mathcal{H}_S) \otimes P_{[\psi_0]}$, i.e. for all $T_S \otimes P_{[\psi_0]} \in \mathcal{T}_1^+(\mathcal{H}_S) \otimes P_{[\psi_0]}$

$$\text{Tr}_{\mathcal{H}_S \otimes \mathcal{H}_M} [W^1(T_S \otimes P_{[\psi_0]})] = \text{Tr}_{\mathcal{H}_S \otimes \mathcal{H}_M} [T_S \otimes P_{[\psi_0]}] = 1$$

and it is null on the subset of $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ denoted by $\mathcal{T}_1^+(\mathcal{H}_S) \otimes (P_{[\psi_0]})^\perp$, i.e. for all states $T_S \otimes P_{[\psi_0^\perp]} \in \mathcal{T}_1^+(\mathcal{H}_S) \otimes P_{[\psi_0^\perp]}$, where ψ_0^\perp is any vector in \mathcal{H}_M perpendicular to ψ_0

$$\text{Tr}_{\mathcal{H}_S \otimes \mathcal{H}_M} [W^1(T_S \otimes P_{[\psi_0^\perp]})] = 0.$$

3. it is contractive, i.e. for all $T \in \mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$,

$$\text{Tr}_{\mathcal{H}_S \otimes \mathcal{H}_M} (W^1(T)) \leq \text{Tr}_{\mathcal{H}_S \otimes \mathcal{H}_M} (T).$$

PROOF. See Appendix A □

REMARK 3.4. For all $T \in \mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$, setting

$$\lambda_i(T) = \text{Tr}[P_{[\varphi_i \otimes \psi_0]} \circ T];$$

then clearly

$$\lambda_i(T) = \langle \varphi_i \otimes \psi_0 | T(\varphi_i \otimes \psi_0) \rangle \in [0, 1]$$

so we have that

$$W^1(T) = \sum_{i=1}^n \lambda_i(T) P_{[\tilde{\varphi}_i \otimes \psi_i]}$$

It follows that $W^1(T)$ is a quasi-convex combination with weights $\lambda_i(T)$ (in the sense that $0 \leq \sum_i \lambda_i(T) \leq 1$, while for a convex the sum must be equal to one).

Consider again Davies's theorem [25, Theorem 3.1, p.21]:

THEOREM 3.6. *Consider a Hilbert space \mathcal{H} . Every pure positive linear map $T : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$ is of one of the following three forms:*

$$(1) \quad T(T) = BTB^*,$$

where $B : \mathcal{H} \rightarrow \mathcal{H}$ is bounded and linear;

$$(2) \quad T(T) = BT^*B^*,$$

where $B : \mathcal{H} \rightarrow \mathcal{H}$ is bounded and conjugate linear;

$$(3) \quad T(T) = \text{Tr}[TB] (|\psi\rangle\langle\psi|),$$

where $B : \mathcal{H} \rightarrow \mathcal{H}$ is bounded, linear and positive and $\psi \in \mathcal{H}$.

Pure positive linear maps on $\mathcal{T}(\mathcal{H})$ are by definition such that every element in the set of positive trace class operators is mapped into the same set, and furthermore every pure state is mapped to a pure state. Type 3 maps as characterised by the theorem are called degenerate: they map all states to a pure state. Note that the map W^1 is a sum of such degenerate mappings:

$$\begin{aligned} W^1(T) &= \sum_{i=1}^n \langle \varphi_i \otimes \psi_0 | T(\varphi_i \otimes \psi_0) \rangle (|\tilde{\varphi}_i \otimes \psi_i\rangle\langle\tilde{\varphi}_i \otimes \psi_i|) \\ &= \sum_{i=1}^n \text{Tr}[TP_{[\varphi_i \otimes \psi_0]}] (|\tilde{\varphi}_i \otimes \psi_i\rangle\langle\tilde{\varphi}_i \otimes \psi_i|) \end{aligned}$$

At this point a comparison is useful between W^1 and the Davies map of type 1 induced by the partial isometry W^1 , i.e.

$$W_D^1(T) = W^1 T W^{1*}.$$

Both are candidates for constructing a bijective, trace-preserving mapping yielding a measurement on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$. Like the partial isometry W^1 on $\mathcal{H}_S \otimes \mathcal{H}_M$ they satisfy the probability reproducibility condition, and if they can be completed to a map which is one-to-one and onto and preserves the trace of the density operators it is applied to, we will have an equivalent for the space $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ of a unitary mapping on

$\mathcal{H}_S \otimes \mathcal{H}_M$. In what follows I show why this can be done for W_D^1 , but not for W^1 , and connect this with the discussion of the previous chapter. On the other hand, I will also show how W^1 , but not W_D^1 , retains, in the density operator description, important informational properties, normally associated with unitary measurements.

The explicit form of W_D^1 is

PROPOSITION 3.7.

$$W_D^1(T) = \sum_{i,j} \langle \varphi_i \otimes \psi_0 | T(\varphi_j \otimes \psi_0) \rangle (|\tilde{\varphi}_i \otimes \psi_i\rangle \langle \tilde{\varphi}_j \otimes \psi_j|).$$

PROOF. Since $W^j = \sum_i \langle \varphi_i \otimes \psi_0 | (\cdot) \rangle (\tilde{\varphi}_i \otimes \psi_i)$ and $W^{j*} = \sum_j \langle \tilde{\varphi}_j \otimes \psi_j | (\cdot) \rangle (\varphi_j \otimes \psi_0)$, we have, for any $\Phi \in \mathcal{H}_S \otimes \mathcal{H}_M$,

$$\begin{aligned} W^{j*}(\Phi) &= \sum_j \langle \tilde{\varphi}_j \otimes \psi_j | \Phi \rangle (\varphi_j \otimes \psi_0) \\ T(W^{j*}(\Phi)) &= \sum_j \langle \tilde{\varphi}_j \otimes \psi_j | \Phi \rangle T(\varphi_j \otimes \psi_0) \\ W^j T W^{j*}(\Phi) &= \sum_i \langle \varphi_i \otimes \psi_0 | \sum_j \langle \tilde{\varphi}_j \otimes \psi_j | \Phi \rangle T(\varphi_j \otimes \psi_0) \rangle (\tilde{\varphi}_i \otimes \psi_i) \\ &= \sum_{i,j} \langle \tilde{\varphi}_j \otimes \psi_j | \Phi \rangle \langle \varphi_i \otimes \psi_0 | T(\varphi_j \otimes \psi_0) \rangle (\tilde{\varphi}_i \otimes \psi_i), \end{aligned}$$

proving the result. \square

In order to appreciate the difference between W^1 and W_D^1 , consider their action on a pure state

$$T = P_{[\Phi]}, \quad \Phi \in \mathcal{H}_S \otimes \mathcal{H}_M, \quad \|\Phi\| = 1.$$

We have already seen in equation (3.3) that

$$W^1(P_{[\Phi]}) = \sum_i |\langle \varphi_i \otimes \psi_0 | \Phi \rangle|^2 P_{[\tilde{\varphi}_i \otimes \psi_i]},$$

a quasi-convex combination of the pure states $P_{[\tilde{\varphi}_i \otimes \psi_i]}$ with weights given by $\lambda_i(\Phi) = |\langle \varphi_i \otimes \psi_0 | \Phi \rangle|^2$, such that $\sum_i |\langle \varphi_i \otimes \psi_0 | \Phi \rangle|^2 \leq 1$. Now we have the following

PROPOSITION 3.8.

$$\begin{aligned} W_D^1(P_{[\Phi]}) &= P_{[\sum_i \langle \varphi_i \otimes \psi_0 | \Phi \rangle (\tilde{\varphi}_i \otimes \psi_i)]} \\ &= \sum_i |\langle \varphi_i \otimes \psi_0 | \Phi \rangle|^2 P_{[\tilde{\varphi}_i \otimes \psi_i]} + \text{Int}(\dots), \end{aligned}$$

a pure state where $\text{Int}(\dots)$ represents interference terms between the projectors equal to

$$(3.6) \quad \sum_{i \neq j} \langle \Phi | \varphi_j \otimes \psi_0 \rangle \overline{\langle \Phi | \varphi_i \otimes \psi_0 \rangle} |\tilde{\varphi}_i \otimes \psi_i\rangle \langle \tilde{\varphi}_j \otimes \psi_j|.$$

PROOF.

$$\begin{aligned} W_D^1(P_{[\Phi]}) &= \sum_{i,j} \langle \varphi_i \otimes \psi_0 | \langle \Phi | \varphi_j \otimes \psi_0 \rangle \Phi | \tilde{\varphi}_i \otimes \psi_i \rangle \langle \tilde{\varphi}_j \otimes \psi_j | \\ &= \left| \sum_i \langle \varphi_i \otimes \psi_0 | \Phi \rangle (\tilde{\varphi}_i \otimes \psi_i) \right\rangle \left\langle \sum_j \langle \varphi_j \otimes \psi_0 | \Phi \rangle (\tilde{\varphi}_j \otimes \psi_j) \right|. \end{aligned}$$

But

$$\sum_i \langle \varphi_i \otimes \psi_0 | \Phi \rangle (\tilde{\varphi}_i \otimes \psi_i) = \sum_j \langle \varphi_j \otimes \psi_0 | \Phi \rangle (\tilde{\varphi}_j \otimes \psi_j)$$

so

$$W_D^1(P_{[\Phi]}) = P_{[\sum_i \langle \varphi_i \otimes \psi_0 | \Phi \rangle (\tilde{\varphi}_i \otimes \psi_i)]}$$

The proof of the second claim is similar. \square

EXAMPLE 3.3. For the case when $\Phi = (\sum_{i=1}^2 \alpha_i \varphi_i) \otimes \psi_0 \in \mathcal{H}_S \otimes \mathcal{H}_M$ with $\dim(\mathcal{H}_S) = \dim(\mathcal{H}_M) = 2$ and $|\alpha_1|^2 + |\alpha_2|^2 = 1$ we have that

$$\begin{aligned} W_D^1(P_{[(\sum_{i=1}^2 \alpha_i \varphi_i) \otimes \psi_0]}) &= P_{[\sum_{i=1}^2 \alpha_i \tilde{\varphi}_i \otimes \psi_i]} \\ &= \sum_{i=1}^2 |\alpha_i|^2 P_{[\tilde{\varphi}_i \otimes \psi_i]} \\ &\quad + \bar{\alpha}_1 \alpha_2 |\tilde{\varphi}_2 \otimes \psi_2\rangle \langle \tilde{\varphi}_1 \otimes \psi_1| + \alpha_1 \bar{\alpha}_2 |\tilde{\varphi}_1 \otimes \psi_1\rangle \langle \tilde{\varphi}_2 \otimes \psi_2| \end{aligned}$$

The difference between W^1 and W_D^1 is a crucial one: W^1 maps away interference terms like (3.6), while W_D^1 does not. As such it should be intuitively clear that W^1 cannot be extended to a bijective mapping on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ generated by a unitary operator on $\mathcal{H}_S \otimes \mathcal{H}_M$, in the manner of a Davies type 1 conjugation. Nevertheless we can't exclude the possibility that W^1 or some such mapping can be extended to a bijective mapping in general. This question will be treated in detail in the next section.

4. The problem of completion for mappings on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$

In this section I examine the problem of constructing bijective completions of the mappings W^1 and W_D^1 on the whole of $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$. As explained at the beginning of the chapter, such completions are important because quantum mechanical evolutions are assumed and, with respect to quantum measurement, hoped to be invertible. Again, therefore, after introducing mappings which satisfy the probability reproducibility condition, the next step is to see whether these mappings can be ‘embedded’ into invertible mappings. I will begin with some results which hold for any space $\mathcal{T}(\mathcal{H})$, \mathcal{H} some Hilbert space.

- DEFINITION 3.1.**
1. The rank of a density operator is the minimal number of one-dimensional projectors of which it is a convex combination. This is an invariant for the operator and it stratifies the convex set of density operators (for a geometric description of some low-dimensional examples see [7]).
 2. A map is rank-preserving in $\mathcal{T}_1^+(\mathcal{H})$ if it preserves the rank of the density operators on which it acts.

Note that it follows trivially from the above definition that rank-preserving maps must be pure, as all pure states are 1-dimensional projections, and so have rank 1.

Positive, linear rank-preserving mappings on $\mathcal{T}(\mathcal{H})$ are a subset of the pure mappings of Davies type 1 and 2: mappings of type 3 cannot be rank-preserving, as they map all T states, regardless of their rank, to final states of the form $\text{Tr}[TB](|\psi\rangle\langle\psi|)$, which have rank one. The restrictions of such mappings to $\mathcal{T}_1^+(\mathcal{H})$ is clearly well-defined and convex. I will prove a theorem linking bijections to rank-preserving maps in $\mathcal{T}_1^+(\mathcal{H})$. For this I will need the following

LEMMA 3.9. *If U is a pure convex bijection on $\mathcal{T}_1^+(\mathcal{H})$, then so is its inverse U^{-1} .*

PROOF. In order to prove this I must show that U maps no mixed states to a pure state, so that its inverse U^{-1} necessarily maps pure states into pure states. Consider a mixed state $T = \sum_{i=1}^n a_i P_i \in \mathcal{T}_1^+(\mathcal{H})$ such that $U(T) = P$, P a pure state. Without loss of generality suppose $n = 2$. Now let $U(P_i) = T_i$, a pure state. Then we have that, by convexity

$$U(T) = a_1 U(P_1) + a_2 U(P_2) = a_1 T_1 + a_2 T_2.$$

By the definition of a pure state this implies that $T_1 = T_2 = P$, so $U(P_i) = U(T)$ which contradicts bijectivity. \square

THEOREM 3.10. *A convex mapping $U : \mathcal{T}_1^+(\mathcal{H}) \rightarrow \mathcal{T}_1^+(\mathcal{H})$ is bijective if and only if it is rank-preserving for all $T \in \mathcal{T}_1^+(\mathcal{H})$.*

PROOF. The first part of the proof is the claim that bijectivity implies rank-preservation. Note first that every bijection is pure. Suppose the linear mapping U is not pure. Then there exists a pure state $P_{[\varphi]}$ for which

$$U(P_{[\varphi]}) = \sum_i a_i P_{[\varphi_i]}, \quad a_i \in \mathbb{R}_+, \quad \sum_i a_i = 1, \quad P_{[\varphi_i]} \neq P_{[\varphi_j]} \text{ when } i \neq j.$$

Without loss of generality assume that $i = 1, 2$. Then there must exist two distinct states T_1, T_2 such that $U(T_i) = P_{[\varphi_i]}$. Consider the state $a_1 T_1 + a_2 T_2$, clearly distinct from $P_{[\varphi]}$. We have that

$$\begin{aligned} U(a_1 T_1 + a_2 T_2) &= a_1 U(T_1) + a_2 U(T_2) && \text{by linearity} \\ &= a_1 P_{[\varphi_1]} + a_2 P_{[\varphi_2]} \\ &= U(P_{[\varphi]}), \end{aligned}$$

so U cannot be a bijection. Hence every bijection must be pure.

The next step is to show that every pure bijection is rank-preserving. Consider $T = \sum_{i=1}^n c_i P_i \in \mathcal{T}_1^+(\mathcal{H})$, with P_i one-dimensional projectors and hence $\text{Rk}(T) = n$. Then $U(T) = \sum_{i=1}^n c_i U(P_i)$. It follows that $\text{Rk}[U(T)] \leq n$.

Suppose now that there exists a decomposition $U(T) = \sum_{j=1}^m d_j \hat{P}_j$, with $m < n$. It follows that $T = \sum_{j=1}^m d_j U^{-1}(\hat{P}_j)$. By Lemma 3.9 U^{-1} is also a pure bijection, hence we can write $U^{-1}(\hat{P}_j) = P'_j$ and $T = \sum_{j=1}^m d_j P'_j$, which implies $\text{Rk}(T) \leq m < n$, yielding a contradiction. We therefore have that $\text{Rk}[U(T)] = n = \text{Rk}(T)$.

For the converse, assume that U is a rank-preserving map. If it is not a bijection, then there must be at least two distinct states $T_1 = \sum_{i=1}^n a_i P_i$ and $T_2 = \sum_{i=1}^m b_i P'_i$ which map to the same final state T . Rank preservation implies that $n = m$. Convexity implies that

$$U(T_1) = \sum_i a_i U(P_i) = \sum_i b_i U(P'_i) = U(T_2)$$

so the two sets $\{a_i\}$ and $\{b_i\}$ must be equal and each $U(P_i)$ must be equal to some $U(P'_j)$. From this it follows that there must be at least two distinct pure states, call them P_1 and P_2 , which are mapped to the same final pure state P .

Now consider the mixed state $T' = a_1 P_1 + a_2 P_2 \in \mathcal{T}_1^+(\mathcal{H})$. Again by convexity

$$U(T') = a_1 U(P_1) + a_2 U(P_2) = a_1 P + a_2 P = P,$$

which violates rank-preservation and hence yields a contradiction. \square

W^1 , analysed in Theorem 3.5, is clearly not a bijection. It yields a final state for the system + apparatus which is a mixture over final states of the system perfectly correlated with the corresponding pointer states, while achieving the appropriate probabilistic behaviour in order for it to satisfy the probability reproducibility condition. In order to do so it maps all initial states with the same probability distribution with respect to the system observable in question to the same final state, irrespective of the rank of the initial state. For instance, given an observable determined by $\{\varphi_1, \dots, \varphi_n\}$, the two states $P_{[\alpha_1 \varphi_1 + \alpha_2 \varphi_2]}$ and $\sum_{i=1,1} |\alpha_i|^2 P_{[\varphi_i]}$, once tensored with a ready-to-measure pointer state $P_{[\psi_0]}$, will both be mapped to the same final state

$$\sum_{i=1,2} |\alpha_i|^2 P_{[\tilde{\varphi}_i \otimes \psi_i]}.$$

Equally clearly it is not rank-preserving on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$; for example $P_{[\varphi \otimes \psi_0]}$, which has rank 1 since it is a projection onto a one-dimensional subspace, is mapped to a state of rank 2. Any trace preserving extension W on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ of an operator like W^1 would retain such failure of rank preservation. It is possible to construct many such operators, for example

$$(4.1) \quad W(T) = \sum_{i=1}^n \text{Tr}_{\mathcal{H}_S \otimes \mathcal{H}_M} \left[\left(\sum_{j=1}^m P_{[\varphi_i \otimes \psi_0^j]} \right) (T) \right] (P_{[\tilde{\varphi}_i \otimes \psi_i]})$$

But, by Theorem 3.10, no such extension can be a bijection, because of the continued failure of rank preservation.

The positive, linear, trace preserving, pure transformation W_D^1 is equally clearly not a bijection; for instance, it maps all the operators $P_{[\Phi]}$ with $\Phi = \varphi \otimes \psi_0^\perp$, ψ_0^\perp fixed and orthogonal to ψ_0 , to the null operator. However, consider the map

$$(4.2) \quad \begin{aligned} W_D : \mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M) &\longrightarrow \mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M), \\ W_D(T) &= W^* T W; \end{aligned}$$

it is not hard to see that this will be a linear bijection on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$: it is just the unitary operator W “extended” to the set of density operators. Davies’s theorem,

together with Theorem 3.10, makes it clear that mappings like (4.2), where W are the unitary operators characterised in Theorem 3.2, are the *only pure bijective premeasurements* of discrete observables. The Davies result characterises pure mappings; Theorem 3.10 tells us that all bijections must be rank-preserving, therefore pure.

Bijections, therefore, must be of one of the Davies types. They cannot be of type 3, which are essentially projection mappings and therefore not bijective. Given a unitary operator U , type 1 and type 2 characterise a mapping and its inverse as a conjugation action just like the one in (4.2); the pure bijections then will all be of the form W_D for some unitary premeasurement W . Note that no other invertible map other than the ones characterised through Theorem 3.2 can yield a premeasurement through (4.2); straightforwardly, for a pure initial state $P_{[\varphi]}$, the probability reproducibility condition would fail.

As an aside, we know that such mappings do not solve the general measurement problem as defined in the previous chapter. They clearly do not satisfy the objectification condition for the pointer observable, because they fail in all cases to excise the relevant interference terms. This is particularly clear from proposition (3.8) and is obviously independent of whether the initial combined state is a nonpure density operator. Convexity will imply in this case that all individual states will be transformed in accordance with (3.8), thus retaining interference terms between pointer eigenstates in the final state after the premeasurement evolution. So in a sense Theorem 3.2 and Theorem 3.10 also make up an insolubility proof.

Concluding this section, we have seen how the probability reproducibility condition is satisfied by a varied class of operations on density operator spaces, not all of which admit of completions to invertible mappings, which is not the case for operators on a Hilbert space. We have explicitly examined this problem for two operations satisfying the probability reproducibility condition, W^1 and W_D^1 , defined in the previous section: W , an extension of W^1 , can never be an invertible completion, while W_D is indeed an invertible completion of W_D^1 .

5. Some informational results for mixed initial apparatus states

The analysis of this chapter has been considerably more detailed about the implications of assuming the probability reproducibility condition than is the case, for instance, with the work of Beltrametti et al. [6]. I now draw some conclusions from

this further analysis. A consequence of it is that in the density operator formalism, the possibility that the initial state of the apparatus might be of rank *greater than* 1 will make a difference to the kinds of bijective measurement mappings that are possible.

To begin with, note that the only possible rank-preserving maps on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ will be mappings such as W_D , where the generating unitary evolution on $\mathcal{H}_S \otimes \mathcal{H}_M$ must necessarily be given by an equation such as (2.10). Suppose in fact that $\dim(\mathcal{H}_S) = n \leq \dim(\mathcal{H}_M)$ and that we have an apparatus in an initial state given by the density operator $T_0 = \sum_{i=1}^2 w_i P_{[\psi_0^i]}$, $\psi_0^1 \perp \psi_0^2$, measuring an observable determined by the orthonormal basis $\{\varphi_1, \dots, \varphi_n\}$, with the final pointer observable determined by $\{\psi_1, \dots, \psi_n\}$. It is easy to see that, if the natural probability reproducibility condition (3.5) is satisfied for an initial pointer state T_0 , then it must be satisfied even if the initial pointer state were to be either one of the component states $P_{[\psi_0^i]}$. This follows naturally when we consider that the probability reproducibility condition (3.5) in this case takes the form

$$(5.1) \quad \text{Tr}_{\mathcal{H}_S}[T P_{[\varphi_j]}] = \text{Tr}_{\mathcal{H}_S \otimes \mathcal{H}_M} \left[W^1(T \otimes T_0) (I_S \otimes P_{[\psi_j]}) \right],$$

for all $T \in \mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ and for all φ_j 's eigenstates of the measured observable. But by linearity of W^1 and of the trace operation this becomes

$$\text{Tr}_{\mathcal{H}_S}[T P_{[\varphi_j]}] = \sum_{i=1}^2 w_i \text{Tr}_{\mathcal{H}_S \otimes \mathcal{H}_M} \left[W^1(T \otimes P_{[\psi_0^i]}) (I_S \otimes P_{[\psi_j]}) \right].$$

Letting $\mathcal{H}_S \ni T = P_{[\varphi_j]}$ it is easy to see that condition (5.1) can be satisfied only when it is satisfied for each of the pointer states $P_{[\psi_0^i]}$. Therefore the only mappings that can satisfy (5.1) must satisfy probability reproducibility conditions for initial states which are pure states, so must be captured by mappings W_D generated by unitary operators as characterised by Theorem 3.2.

In particular, ψ_0^1 will fix one partial isometry W^1 , but its completion will not be totally arbitrary now. The idea is for W_D to satisfy the natural interaction conditions

$$\begin{aligned} P_{[\varphi_1]} \otimes P_{[\psi_0^1]} &\rightarrow P_{[\tilde{\varphi}_1^1]} \otimes P_{[\psi_1]} \\ P_{[\varphi_1]} \otimes P_{[\psi_0^2]} &\rightarrow P_{[\tilde{\varphi}_2^1]} \otimes P_{[\psi_1]} \end{aligned}$$

which follow from (5.1). These can be put more generally as

$$(5.2) \quad P_{[\varphi_1]} \otimes T_0 \rightarrow \tilde{T} \otimes P_{[\psi_1]}.$$

Now, however, the partial isometry W^2 , corresponding to the initial apparatus state $P_{[\psi_0^2]}$, will be fixed, just as W^1 is, in the way shown by Theorem 3.2. The more pure states $P_{[\psi_0^i]}$ appear in the initial apparatus state, the less freedom we have in specifying unitary completions of partial isometries if we want to satisfy the interaction conditions, until we reach the point where the initial state is $T_0 = \sum_{i=1}^n P_{[\psi_0^i]}$ for $n = \dim(\mathcal{H}_S)$, when the above reasoning shows that *the unitary operator characterising the measurement will be fully specified and unique* given the system observable to be measured, the apparatus pointer observable and the initial state.

This prompts, in the first instance, some observations on the conditions that measurements can satisfy.

- The first is that if the apparatus observable is a density matrix of rank strictly greater than 1, then Von Neumann-Lüders measurements are impossible. Consider the interaction condition (5.2). A Von Neumann-Lüders measurement would require that $\tilde{T} = \varphi_1$. This is clearly impossible. The rank of $P_{[\varphi_1]} \otimes T_0$ is equal to 2 and any one-to-one measurement cannot change this rank, hence the rank of $T \otimes P_{[\psi_1]}$ must also be equal to 2. As the rank of $P_{[\psi_1]}$ is equal to 1, it follows that T must necessarily be of rank 2 (as can be easily calculated using the explicit formulation of a Von Neumann-Lüders measurement in the earlier example).
- The second remark concerns two other important properties that measurements can satisfy, repeatability and the property of being a measurement of the first kind. It is possible to show that, as for Von Neumann-Lüders measurements, no measurement can be of the first kind if the initial apparatus state is a density matrix of rank greater than 1, and hence cannot be repeatable, as all repeatable measurements are of the first kind. The intuition is simple. A measurement of the first kind is a probabilistic non-disturbance measurement: take the final object state after a measurement, obtained through a partial tracing operation, and subject this state to another measurement of the same observable. The end object state after this measurement should exhibit the same probability distribution as the initial object state, for all possible initial states of the measured object. The fixed weights of the non-trivial density operator taken as the apparatus starting state make this impossible.

More precisely, a measurement of an observable A is of the *first kind* if the probability for a result being in a subset X of the set of all possible results is the same before and after a measurement. Therefore, for a state $T_S \in \mathcal{T}_1^+(\mathcal{H}_S)$, for an initial

state $T_M \in \mathcal{T}_1^+(\mathcal{H}_M)$ and for all such subsets X

$$(5.3) \quad p_{T_S}^A(X) = p_{\mathcal{R}_S(U(T_S \otimes T_M))}^A(X),$$

where \mathcal{R}_S is the operation of partial tracing the apparatus system away.

As an example, consider a system in an initial state $\varphi = \sum_i \alpha_i \varphi_i$, where the φ_i 's are eigenstates of a discrete self-adjoint hermitean operator A representing the observable to be measured (and form an orthonormal basis for \mathcal{H}_M). The spectral resolution of the operator is given by $A = \sum_i a_i P_{[\varphi_i]}$. Such a system undergoes a measurement, for which there is a pointer observable Z whose spectral resolution is given by $Z = \sum_i b_i P_{[\psi_i]}$. The starting state of the measured system is given by a density operator $T = \sum_{j=1,2} w_j P_{[\psi_j]}$.

A generic unitary measurement U will map the state $P_{[\varphi]} \otimes T$ to the state $U^{-1}(P_{[\varphi]} \otimes T)U$. This state will be

$$T_f = U^{-1}(P_{[\varphi]} \otimes T)U = \sum_{j=1,2} w_j P_{[\sum_i \alpha_i \tilde{\varphi}_i^j \otimes \psi_i]}$$

with $\tilde{\varphi}_i^j$ orthogonal for different j 's. (5.3) tells us that, for example, if $X = \{a_1\}$ then the probability given by $\text{Tr}_{\mathcal{H}_S}[P_{[\varphi]} P_{[\varphi_1]}]$ is equal to $|\alpha_1|^2$ and this must be equal to $\text{Tr}_{\mathcal{H}_S}[\mathcal{R}_S(U^{-1}(P_{[\varphi]} \otimes T)U)P_{[\varphi_1]}]$ if the measurement is to be a measurement of the first kind.

To calculate this probability first write $\tilde{\varphi}_i^j = \sum_{l=1}^n \beta_{ijl} \varphi_l$, with $\beta_{ijl} = \langle \tilde{\varphi}_i^j | \varphi_l \rangle$. The final state T_f is then equal to

$$\sum_{j=1,2} w_j P_{[\sum_i \alpha_i (\sum_{l=1}^n \beta_{ijl} \varphi_l) \otimes \psi_i]}$$

As the ψ_i form an orthonormal set it is easy to see that

$$\mathcal{R}_S(T_f) = \sum_{j=1,2} w_j \sum_i |\alpha_i|^2 P_{[\sum_{l=1}^n \beta_{ijl} \varphi_l]}$$

Now multiplying this state by $P_{[\varphi_1]}$ yields the operator

$$\sum_{j=1,2} w_j \sum_i |\alpha_i|^2 |\beta_{ij1}|^2 P_{[\varphi_1]}$$

whose trace is

$$(5.4) \quad \sum_{j=1,2} w_j \sum_i |\alpha_i|^2 |\beta_{ij1}|^2.$$

We want to investigate under what conditions this is equal to $|\alpha_1|^2$.

Suppose $\varphi = \varphi_1$; then $|\alpha_1|^2 = 1$ and (5.4) becomes $\sum_{j=1,2} w_j |\beta_{1j}|^2$. From the fact that $\beta_{ij} = \langle \tilde{\varphi}_i^j | \varphi_l \rangle$, this is clearly equal to 1 if and only if, for one of the two possible values of j , $w_j = |\beta_{1j}|^2 = 1$, while for the other value of j we set the constants to 0. This shows that, for a discrete, non-degenerate observable A , the initial state of the apparatus has to be a pure state if we are to be able to define a measurement possessing the first kind property (and hence the repeatability property).

- Finally it is clear that, if the unitary mapping W generating the measurement is uniquely fixed by the initial state of the apparatus being of rank n in the case when $\dim(\mathcal{H}_S) = n$, then it follows that if $\dim(\mathcal{H}_M) = m$, $m > n$, and the rank of the initial apparatus state is greater than n then *no measurement* of any kind is possible if the pointer observable is non-degenerate.

Even assuming degeneracy of the pointer observable, a common physical situation, there are cases where a unitary mapping W cannot be found. An example will be useful to illustrate the point. Suppose $\dim(\mathcal{H}_S) = 2$ and $\dim(\mathcal{H}_M) = 3$: the observable to be measured is a spin direction, represented by an operator whose eigenstates belong to the set $\{\varphi_1, \varphi_2\}$, the pointer observable's eigenstates belong to $\{\psi_1, \psi_2, \psi_3\}$, with ψ_1 indicating a result compatible with φ_1 , ψ_2 indicating φ_2 . Further, let the initial state of the measured system be $\varphi = \sum_{i=1,2} \alpha_i \varphi_i$ and the initial state of the apparatus is a density operator $\sum_{j=1}^3 w_j P_{[\psi_0^j]}$, written in its spectral resolution. A unitary measurement will have the form

$$(5.5) \quad U = \sum_{\substack{i=1,2 \\ j=1,2}} \langle \varphi_i \otimes \psi_0^j | (\cdot) \rangle (\tilde{\varphi}_i^j \otimes \psi_i) + \sum_{i=1,2} \langle \varphi_i \otimes \psi_0^3 | (\cdot) \rangle (v_i \otimes \psi_3),$$

the v_i 's being an arbitrary orthonormal pair in \mathcal{H}_S .

Now $U(\varphi \otimes \psi_0^3) = \sum_i v_i \otimes \psi_3$ so that the apparatus points to 'nothing'. Clearly there is no way to satisfy any sort of interaction condition for the third component of the initial pointer state: for the initial apparatus state ψ_0^3 , the final pointer state will have to be ψ_3 regardless of what the initial object state is, and so can't be taken to be properly pointing to anything.

In particular, since interaction conditions are implied by the probability reproducibility condition, failure of the former implies failure of the latter. Suppose for example that the initial apparatus state for a measurement of φ is the density operator $T_M = \sum_{j=1}^3 w_j P_{[\psi_0^j]}$, then the probability reproducibility condition requires the final

probabilities to be captured by the formula $\text{Tr}[\mathbb{I} \otimes P_{\{\psi_i\}} U(P_{[\varphi]} \otimes T_M) U^*]$ for $i = 1, 2$. But a ‘portion’ w_3 of these probabilities will be lost in the interaction with the initial state $P_{[\psi_3]}$ to the final state ψ_3 . From this it follows that definition of a unitary measurement is impossible in such situations. The reasoning can easily be seen to generalise to cases where $\dim(\mathcal{H}_M) \neq n \dim(\mathcal{H}_S)$, where $n \in \mathbb{N}$.

This may seem a somewhat surprising result. It is in fact implicit in the construction adopted for the operator W in (2.10), where the completion terms in the case when the index for the partial isometries W^j exceeds n have a form which is very different from the initial partial isometry and its ‘natural’ conjugates; the argument just presented shows that this is necessarily the case when $\dim(\mathcal{H}_M) \neq n \dim(\mathcal{H}_S)$, if we want to preserve unitarity. This really highlights the need to examine the question of how the minimal partial isometry is completed. In other results, unitary operators such as W are just assumed to be arbitrary completions of a partial isometry, but the structure of completions is not examined in any detail. In the present analysis it is now evident that this structure is very important.

Note also that this result is independent of both

1. the method for constructing the unitary operator
2. and the way in which the initial nonpure density operator for the apparatus system is resolved.

The first claim highlights the fact that, when writing down a unitary evolution, for example, for the system in the case just discussed, (5.5) is not the only option. There will be different measurements satisfying, for example, the property that one of the partial isometries W^j yields a Von Neumann-Lüders measurement on the appropriate subset of the set of states; one could start with a partial isometry W^1 , or with W^2 , and complete these into a unitary operator. Yet the problem will remain if the rank of the initial density operator is greater than the dimension of the object space.

The second stresses the independence of the result from cases like the following: if the initial state of the apparatus were to be in a density operator some of whose weights appeared with multiplicity greater than 1, one could rewrite the initial state perhaps in some way that might avoid the problem. However suppose that $T_M = \sum_{j=1}^3 w_j P_{[\psi_0^j]}$, with $w_2 = w_3 = w$. Then write the initial state of the apparatus as a density operator with weights w_1 and (twice) w , where the projectors making up the density operators do not include $P_{[\psi_0^3]}$. These projectors will, nevertheless, be projections on a linear

combination of $P_{[\psi_0^2]}$ and $P_{[\psi_0^3]}$, so the evolution will still generate problems because of the effect of the term $P_{[\psi_0^3]}$ when calculating probabilities. It is well known, in any case, that the evolution of density operators is independent of how they are decomposed.

- The non-unique decomposability of density operators, on the other hand, raises another conceptual problem for the definition of bijective, trace preserving measurements: on which decomposition of the initial apparatus density operator should the measurement operator be based?

It might be suggested that, given that unitary measurement operators are naturally defined through an orthonormal set of apparatus vector states, the correct procedure to follow is to rely on an orthonormal decomposition of the initial apparatus density operator. It is true that in the case when multiplicities are present in the initial apparatus state this does not select a unique basis. In a sense this does not make much difference. Suppose that we have a case of T_M having weights with multiplicities as described above, now in a measurement of an object system having three genuinely different eigenstates of the measured observable. Then the initial object plus apparatus state will be

$$T = \sum_{j=1}^3 w_j P_{[\varphi \otimes \psi_0^j]} = \sum_{j=1}^3 w_j P_{[\varphi \otimes \zeta_0^j]},$$

with $P_{[\psi_0^1]} = P_{[\zeta_0^1]}$ and $P_{[\zeta_0^2]}$ and $P_{[\zeta_0^3]}$ projections on orthogonal linear combinations of ψ_0^2 and ψ_0^3 . This would give the following two choices (out of an infinity of possible ones) for a measurement evolution:

$$(5.6a) \quad U = \sum_{\substack{i=1,\dots,3 \\ j=1,\dots,3}} \left\langle \varphi_i \otimes \psi_0^j \middle| (\cdot) \right\rangle (\tilde{\varphi}_i^j \otimes \psi_i)$$

and

$$(5.6b) \quad U = \sum_{\substack{i=1,\dots,3 \\ j=1,\dots,3}} \left\langle \varphi_i \otimes \zeta_0^j \middle| (\cdot) \right\rangle (\tilde{\varphi}_i^j \otimes \psi_i)$$

It is easy to see that the result of applying (5.6a) or (5.6b) to T will be the same.

- On the other hand the non-unique decomposability of mixtures would present genuine problems for arguments which might seek to avoid the problem generated by evolutions such as (5.5). These arguments could claim that, when the initial apparatus state is a density operator, it is correct to treat the measurement process as different measurements for the different initial pure states making up the initial apparatus state,

each having their own unitary measurement associated with them and each evolving independently of one another. After all the problem is generated by the fact that unitary evolutions satisfying the probability reproducibility condition *must* depend on the initial apparatus state, as well as on the object and pointer observables. In the case when the initial state is a pure state there are, as we have seen earlier, plenty of these unitary operators: we need a basis of initial observables for the apparatus space \mathcal{H}_M in order to define a unitary operator, but only one vector in the basis is relevant to the measurement, the rest can be chosen at random. When the initial state is not pure, more and more vectors become relevant to the measurement operator as the rank of the density matrix increases, until we have the case described above, when no unitary premeasurement can be defined. We can, on the other hand, claim that, for every pure state decomposing the density matrix, there is a different unitary operator describing the evolution.

Note that it would then be irrelevant whether the decompositions were orthogonal or not: the idea would be not to define a single unitary operator, thus presenting us with a choice like the one between (5.6a) and (5.6b). Rather a different unitary operator would be responsible for the measurement evolution of each component $P_{[\varphi \otimes \psi_0^j]}$. Furthermore these could be chosen so that they all give rise to same final state: assuming $\varphi = \sum_{i=1}^n \alpha_i \otimes \psi_i$ the final states for the object + apparatus initial state, $P_{[\varphi \otimes \psi_0^j]}$ could be $P_{[\sum_{i=1}^n \alpha_i \hat{\varphi}_i^j \otimes \hat{\psi}_i]}$, again through operators such as (5.6a) and (5.6b).

This strategy has the problem of committing itself to saying *what* the initial apparatus pure states are, if the claim of individual evolutions with pure apparatus states is to make sense. Arguably, in the case where the apparatus system is initially in a proper mixture (i.e. in a case where we *know* what the different pure states making up the mixture are), the problem will not arise: the one measurement *will* in fact be three measurements, each determined by operators fixed solely by the pure state representing, case by case, the initial apparatus state. The other option is of course that the initial density operator might admit of an ignorance interpretation, so that in fact the density operator is the state of a system which is really in a pure state, it's just that we don't know which.

But there are cases for which a density operator does not admit of an ignorance interpretation, and indeed, as we saw in the last chapter, there is a view about how to interpret density operators associated with a quantum system which treats them

purely as statistical states, and downplays the information about which states the system might be in. In such cases we cannot define different measurements for different possible initial apparatus states. If, for example, we accept the interpretation of states that Stein advocates, there is no such thing as *really* pure initial states underlying a nonpure density operator. If mother nature really works like quantum mechanics tells us, and we accept this interpretation, then the problem just discussed cannot be avoided.

The topics of the preceding discussion challenge the quantum theory of measurement, as characterised in subsection 2.2 of this chapter by unitary measurements satisfying the probability reproducibility condition, by suggesting a class of cases in which it would fail to model measurement results.

At first glance this may seem little more than a curiosity. But it is not difficult to think of at least a thought experiment exemplifying it. Suppose we have a photon for which we are trying to measure polarisation (a two-level system) by coupling it to a pointer system, in this case an atom's magnetic moment described by a three level system. Possible unitary premeasurement couplings for this system will be characterised by Theorem 3.2, therefore for an atom suitably prepared in an initial state, perhaps entangled with an environment, the situation just described *can* theoretically arise if the initial apparatus state is a density operator of rank three. This would entail that either unitarity/bijectivity, or the probability reproducibility condition, would have to be given up if a case of this kind is to admit of a quantum mechanical analysis.

On the other hand, a measurement defined as

$$(5.7) \quad \tilde{W}^j : \Phi \rightarrow \frac{1}{\|W^j(\Phi)\|} W^j(\Phi)$$

with W^j a partial isometry on a set $\mathcal{H}_S \otimes [\psi_0^j] \subset \mathcal{H}_S \otimes \mathcal{H}_M$ (characterised as in part (i) of Theorem 3.2), $P_{[\psi_0^j]}$ one of the allowable projections in the possible decompositions of the initial apparatus state, determining a standard conjugation action on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$, and $1/\|W^j(\Phi)\|$ a (non-constant) normalisation factor, would avoid all of the above problems. Suppose we have an initial object + apparatus state $\sum_{j=1}^n P_{[\varphi]} \otimes P_{[\psi_0^j]}$; a measurement \tilde{W}^1 defined as in (5.7) will map $P_{[\varphi]} \otimes P_{[\psi_0^1]}$ to a final state in such a way that the probability reproducibility condition will be satisfied, and will map $\sum_{j=2}^n P_{[\varphi]} \otimes P_{[\psi_0^j]}$ to 0. In other words, it treats the initial state $\sum_{j=1}^n P_{[\varphi]} \otimes P_{[\psi_0^j]}$ as if it were simply $P_{[\varphi]} \otimes P_{[\psi_0^1]}$, therefore avoiding all of the above

problems, associated with the fact that the initial apparatus state was a density operator.

However it is unclear why in such a case we should regard any one W^j as the appropriate operator from which to construct a measurement on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$. These partially isometric operators are normally defined through initial apparatus states; if the initial state is a nonpure density operator, why should the corresponding measurement be defined by one specific pure state as we have just done? And if so, which pure state is physically relevant? Again, if the initial apparatus state consists of a proper mixture we can at least claim that the underlying apparatus state is a specific pure state, though of course the problem here emerges precisely because in so many situations we cannot assume that a density operator does denote a proper mixture. There are, in such cases, many more than one partial isometry to choose from, and no good reason to choose a specific one. The claim that the partial isometry captures what is basic, the ‘core’, of the measurement process, would lose its force.

On the other hand, W^1 and any one of its ‘completions’ such as W defined in equation (4.1) also avoid the problems discussed in this section. In particular W is a completion in the sense that it is trace-preserving on all elements of $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$. It cannot be, as we have seen, an invertible map, but then neither are the mapping defined in equation (5.7), nor W^1 . Its advantage over these is that it can be defined for any initial apparatus state whatsoever.

Note how different a role the problems raised in this section play, with respect to the account of measurements that quantum mechanics gives, when confronted with the criticisms of this account discussed in chapter 1. There the criticisms argue from the richness of the experimental situation to the deficiency of the theoretical account, but they fail to show that what is going on in the real measurements is (negatively) relevant to the theoretical questions that underlie the quantum mechanical accounts of measurements. The problem discussed in this chapter, however, while it may not correspond to any measurement actually performed, puts forward a possible case where it is *impossible* to answer the questions of the theoretical accounts, given the parameters imposed by the probability reproducibility condition and unitarity.

The latter conditions, as I have pointed out in this chapter, are an important part of providing interpretations of quantum mechanics. Their role in the attempts to provide an interpretation of quantum mechanics have been justified in a number of ways. For

the probability reproducibility condition I have pointed out, in particular, that weakening it leads to possible measurements which count as measurements of incompatible observables. This does not mean that, if we relax the probability reproducibility condition, it is possible to construct unitary measurements which are joint measurements of incompatible observables, but rather that there are unitary measurements which transmit information from the object to the apparatus from which we would not be able to distinguish which of a number of incompatible observables are being measured. As many interpretations of quantum theory seek to describe measurements which enable us to read object properties from apparatus properties, they cannot accept such a weakening of the condition without ending up in a position where they can't assign precise properties.

As for unitarity, as I have already remarked, its role in the interpretation of quantum mechanics is related to a demand for deterministic theories. Without getting into the issues of determinism (extensively treated by Earman [30]) it is clear that, if we want quantum mechanics to be something like a deterministic theory, we need some condition like unitarity (Earman in fact discusses how the unitarity of the Schrödinger evolution is *sufficient* for the evolution being deterministic, unlike what happens in classical mechanics and special relativity).

The counterexample highlights a conflict between assuming both unitarity and probability reproducibility for an evolution on the space $\mathcal{H}_S \otimes \mathcal{H}_M$: to account for the example within something like the present framework, we must give up one of the two conditions. As such it offers little in the way of suggestions about what actual measurements in quantum mechanics are like, but is very informative about the limitations of the assumptions we impose when trying to address the interpretive questions raised by quantum measurement.

6. Conclusions

This Chapter has reviewed four types of measurement mappings: the partial isometry W^1 , its unitary completion W , and the two non-invertible operations \mathcal{W}^1 and \mathcal{W} defined on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$.

W^1 is a contractive partial isometry on $\mathcal{H}_S \otimes \mathcal{H}_M$, defined with respect to a measured system observable, a pointer observable and an initial vector state $\psi_0^1 \in \mathcal{H}_M$, and the mapping it generates on the set of states $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ is denoted by \mathcal{W}_D^1 .

This operator satisfies the standard interaction conditions for measurements under all circumstances, it cannot objectify the pointer observable, and is obviously not unitary. Interaction conditions are satisfied under all circumstances, including when the initial apparatus state is a nonpure density operator for which the ignorance interpretation of mixtures is inapplicable. However, it is difficult to see, as I have argued, on what physical grounds this operator, defined with respect to only one pure state, should be the appropriate measurement when the initial apparatus state is composed of a convex combination of pure states. What makes this possible description of a measurement interesting, in the case when the initial state is a pure state, is the fact that it captures all that is relevant in the interaction. Measurements depend crucially on the information represented by the initial apparatus state, and when the latter is a pure state such information is fully used in constructing W^1 : this is the *unique* common part to all unitary operators which are completions of it. When the initial state is nonpure, this feature of W^1 (in its normalised version \tilde{W}^1) is lost: it is no longer unique at all, and the feature which makes it physically interesting in the former case, its property of being the common core of all possible unitary measurements, no longer obtains.

W is a unitary completion of W^1 characterised by Theorem 3.2, and the mapping it generates on the set of states is denoted by W_D . It satisfies the probability reproducibility condition, but fails to objectify the pointer observable and fails to satisfy the interaction conditions (and therefore the probability reproducibility condition) when $\dim(\mathcal{H}_M) \neq n \dim(\mathcal{H}_S)$ and the initial apparatus state T_M has rank such that $\text{Rk}(T_S) \neq n \dim(\mathcal{H}_S)$. As I have discussed in the previous section, *no unitary measurement* can satisfy the probability reproducibility condition in such cases, no matter which W^1 we start from, no matter how we complete it.

W^1 is a positive, contractive partial isometry on the set of states $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$. It is a mapping composed of a sum of pure positive convex operations of Davies type 3, which map all states to a pure projector (as a projection operator would), it satisfies the probability reproducibility condition, it achieves objectification of the pointer observable, but it clearly is not a one-to-one mapping. It is effectively a generalisation of Von Neumann's type I evolutions. No such map can be generalised to a one-to-one mapping on the set of states.

However a trace preserving completion W on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ can be defined which will not violate the probability reproducibility condition in the manner that W does, under any possible initial apparatus state, including ones which emerge when $\dim(\mathcal{H}_M) \neq n \dim(\mathcal{H}_S)$. While this map cannot be a bijection, it avoids the problem of the last section, always satisfying the probability reproducibility condition, and it does so relying on *all* the information encoded in the initial apparatus state, rather than having to argue for one state of the decomposition of the initial apparatus state as being in some sense the relevant one, which we have to do both for W_D^1 and for W^1 . Interestingly, it also satisfies the objectification condition; by giving up unitarity/bijectivity, it avoids both the problems discussed in the previous chapter and in this one.

The difference between mappings like W and those like W_D is a part of the insolubility results and of the work of Busch et al. on objectification of measurements, which was extensively discussed in the previous chapter. The standard requirements made on maps that describe quantum measurements is that, as well as satisfying the probability reproducibility condition, they yield final states satisfying objectification (this is achieved by a mapping like W) and that they be one-to-one (this is realised with a mapping like W_D). It is by now totally clear how these two requirements are contradictory, so long as one is confined to working with a Hilbert space like $\mathcal{H}_S \otimes \mathcal{H}_M$.

On the other hand, results like [14, Thm. 6.3.1], the approach based on superselection rules (on this see, for example, Hepp [49] and Beltrametti and Cassinelli [5]) and Fine's work on "interactions with an aspect" [40] show that *all* the above requirements for measurements can be satisfied by setting one of the two spaces in the tensor product to be the representation of an algebra of commutative observables (either for the system or the apparatus). This approach has its own problems (see amongst others [10, 50, 80]); it can nevertheless be seen, from a mathematical point of view, as an attempt to identify the subsets of $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$ on which two mappings like W and W_D are equivalent, in the sense of yielding the same result when applied to the same state. I will discuss some of the issues raised by these analyses in the next chapter, specifically in the context of Fine's work.

CHAPTER 4

Algebraic Approaches in Fine's Solution to the Measurement Problem

This chapter looks at Fine's solution to the measurement problem, developed in a recent series of papers [40, 42, 43]. It shows how it can be reinterpreted in the framework of Algebraic Quantum Theory, something that Fine has advocated himself as a means of motivating his solution [40]. Such a reinterpretation leads to problems as it shows the solution to be equivalent, from the mathematical point of view, to assuming a superselection rule for one of the subsystems. Fine's account might then be open to the standard criticisms levied against superselection rules. These criticisms are reviewed and their impact and tenability vis-a-vis Fine's work is assessed. The chapter concludes with a further discussion of whether this solution, while being in itself straightforward, is lacking in physical justification regardless of whether it is motivated by assuming a superselection rule; this will be done by considering two further criticisms that can be made of it.

It is natural, on the back of the discussion in the previous chapter, to study Fine's proposed solution to the measurement problem. Recall how the previous chapter identified two kinds of mapping: the first, essentially the one characterised by Beltrametti, Cassinelli and Lahti [6], is *unitary* and satisfies the *probability reproducibility condition*; the second satisfies *objectification* of the pointer observable. It is possible to identify subspaces of the tensored object + apparatus Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_M$ where the two kinds of mappings coincide, and where as a consequence all three properties (unitarity, probability reproducibility and objectification) can be satisfied, if the composite system is thought to be restricted to one of said subspaces. Fine's solution, as will become apparent in the rest of this chapter, selects just one such subspace as the appropriate subspace to analyse quantum measurements.

1. Fine's Solution to the Measurement Problem

The question, once again, is: How does quantum mechanics account for the fact that quantum measuring processes actually produce results? Up to now this thesis has been mainly concerned with providing an exhaustive explanation of what the conceptual problem is, and how it manifests itself in the physical theory. Several results have been proved, and the upshot of the previous discussion is that most of the conditions which are assumed in framing the measurement problem lead to some difficulties. Having discussed how objectification and the probability reproducibility condition lead to problems, this chapter now considers some proposed solutions to the measurement problem. It considers in particular one of the solutions which relax the condition of the coupled systems being composed of two proper quantum systems, where a proper quantum system is a system whose observables are represented by the full set of bounded, self-adjoint operators on a Hilbert space and whose states are represented by the positive, trace class one and self-adjoint density operators on said Hilbert space.

Usually such relaxations are achieved by arguing for the presence of a superselection rule. The focus of the present discussion, on the other hand, is Fine's proposal for a solution of the quantum measurement problem. *Prima facie* Fine's solution does not seem to rely on relaxing the condition that coupled systems be made of proper quantum systems; this is certainly not an explicit part of the argument. However it is soon apparent that formally that is what the solution amounts to, particularly through the investigation of some suggestions that Fine himself makes. This leads to the discussion of problems associated with superselection solutions in the slightly different context of Fine's work, and leads to a number of arguments that can be made against it. The idea is to try to understand what kind of argument is needed to support Fine's proposal, and whether such an argument can be given.

As is customary, and by now familiar, Fine considers the measurement process as an interaction between an object system S upon which a measurement is performed, and a measuring apparatus, M . The measurement interaction is treated in the space $\mathcal{H}_S \otimes \mathcal{H}_M$, (\mathcal{H}_S and \mathcal{H}_M two Hilbert spaces) and is described by a unitary transformation U such that

$$(1.1) \quad \mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M) \ni V \otimes W \xrightarrow{U} U(V \otimes W)U^{-1} = F,$$

where $V \in \mathcal{T}_1^+(\mathcal{H}_S)$ is the initial state of the object (as usual a density operator on \mathcal{H}_S) and $W \in \mathcal{T}_1^+(\mathcal{H}_M)$ is the state of the measurement apparatus before the measurement takes place. U transforms the initial state $V \otimes W$ into $F = U(V \otimes W)U^{-1}$, the final state after measurement. Suppose the observable \hat{E} is being measured and that it is represented by an operator E with projections E_n onto its eigenstates. Also suppose that there is a pointer observable, represented by the operator Q , and projections Q_n for this pointer indicating that E_n has been measured. Then the first requirement for this measurement is taken to be

$$(1.2) \quad \text{Prob}(E_n, V) = \text{Prob}(I \otimes Q_n, F),$$

which is the usual probability reproducibility condition in its simplified form: the probability of finding E_n given the initial state V should equal the probability of finding the pointer position Q_n given the final state F . The second requirement is taken to be that measurements have definite results, i.e. the objectification requirement: F is a density operator admitting a resolution over projectors $P_n \otimes Q_n$, which display no interference between pointer eigenstates and where P_n is any projection operator on \mathcal{H}_S . Assume, without loss of generality for the present discussion, that Q_n has no multiple eigenstates. This means formally that

$$(1.3) \quad F = \sum_n w_{nn} P_{[P_n \otimes Q_n]}, \quad \sum_{nn} w_{nn} = 1.$$

Insolubility theorems, as discussed at length in Chapter 2, show that (1.1), (1.2) and (1.3) in general lead to a contradiction. Hence this approach does not account for quantum measurements.

Fine gives the following informal characterisation of the requirements for his attempted solution:

1. The Rule of Silence [formally (1.3)]: there is in quantum mechanics what Fine calls the eigenvalue-eigenstate link, which says that it is possible to assign as a value the eigenvalue of an operator E to a system that is in the eigenstate of that same operator. The Rule of Silence forbids us to speak of a value for that operator when the system is not in an eigenstate of it, or in a density operator decomposable over such eigenstates.
2. Unitary Dynamics [formally (1.1)]: measurement should be modelled by a unitary evolution given by a Schrödinger-type operator.

3. The requirement that the apparatus after measurement can distinguish probabilistically distinct initial object states (w.r.t. E) [formally (1.2)].
4. Fine's own addition to this list: if the object is initially in the state V and the apparatus is in the state W , the measurement interaction should start with the state $V \otimes W$.

Past solutions to the measurement problem have involved modifications of the first two requirements. For instance modifying the first leads to hidden variable theories or Everett-type interpretations, modifying the second is related to non-unitary evolutions and the theory of open systems. Fine's proposal is that we should modify the fourth requirement: if the object is initially in the state V and the apparatus is initially in state W , the initial state of the interaction should look like

$$V_E \otimes W,$$

where V_E is the so-called " E -aspect" of the V -state. What then is V_E ? Fine introduces an equivalence relation on the set of states in a system in the following way. For any two states V and V' and any observable E , V and V' are said to be E -equivalent, written $V \underset{E}{=} V'$, in the following case:

$$(1.4) \quad V \underset{E}{=} V' \iff \text{Prob}(V, E) = \text{Prob}(V', E).$$

For every operator E this does indeed partition states into equivalence classes, which Fine denotes by $[V]_E$ for the class of all states satisfying $V \underset{E}{=} V'$. "This relation", he says, "lumps together pure states with mixed states, provided they yield the same probability distribution for E " [40, page 493].

To see what this equivalence relation actually does to a set of states, pure and mixed, consider the set of all such states for a spin- $\frac{1}{2}$ system. The states of this system (normally represented by the complex Hilbert space \mathbb{C}^2) can be visualised by a "ball" in \mathbb{R}^3 , for which the surface sphere S^2 is identified with the pure states and the points inside the sphere are identified with the mixed states (for details see, for instance, [7]). The equivalence classes given by Fine's relation are then picked out by intersections of parallel planes with the "ball", such planes being perpendicular to one of the diameters of the sphere. The two points at which the diameter intersects the surface of the sphere represent the two eigenstates of the operator corresponding to the observable E to be measured, say P_{+1} and P_{-1} , which for simplicity we assume to be non-trivial (that is,

not equal to a multiple of the identity operator on the object Hilbert space \mathbb{C}^2). For Fine these equivalence classes group together all those states which respond in the same manner to a measurement of the observable E (see figure 1).

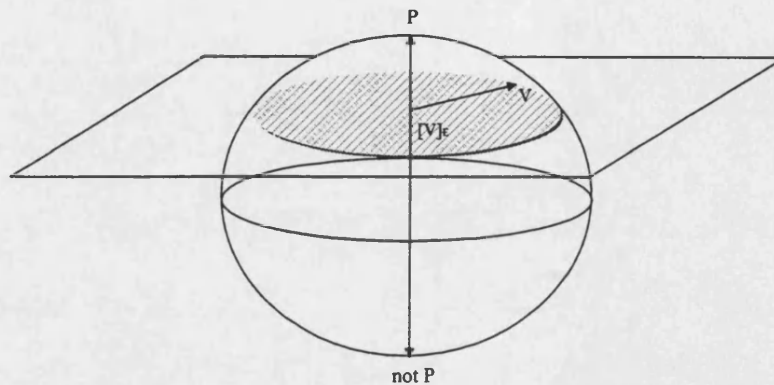


FIGURE 1. The unit ball and a section picked out by Fine's equivalence relation

In interacting with an object to be measured, the measuring apparatus responds only to certain physical aspects of the object, namely those that are relevant to the observable E to be measured. This is crucial in providing the motivation for determining Fine's initial object state V_E . As the Rule of Silence forbids talk of possessed values by the object, the only thing left for it to respond to is the probability distribution associated with the state of the object and the observable E . But then all the states grouped in the above described equivalence classes have the same probability distribution; hence the measurement apparatus interacts with all of them in the same way on Fine's account. Hence we can select from this equivalence class a specific representative V_E for all the states in it. In order to solve the measurement problem, *this representative will be the density operator defined over the pure states which are eigenstates of the observable to be measured* (one such state is always contained in the equivalence class).

It is clear that, if the initial state of an object is either an eigenstate of the measured observable or a mixture of such eigenstates, then it is possible to describe the transition to the appropriate final state by means of a unitary evolution. For instance, in the case of an ideal measurement, unitary operators which satisfy the probability reproducibility condition, discussed in chapter 3, are defined as operators which map an initial state $P_{[\varphi_i]} \otimes P_{[\psi_0]}$, φ_i an eigenstate of the measured observable and ψ_0 the initial apparatus state, to a final state $P_{[\varphi_i]} \otimes P_{[\psi_i]}$, where ψ_i is the pointer eigenstate indicating that

the value associated with the object observable eigenstate φ_i has been measured. So any initial state of the form $\sum_i w_i P_{[\varphi_i]} \otimes P_{[\psi_0]}$ is mapped, by convexity, to a final state $\sum_i w_i P_{[\varphi_i]} \otimes P_{[\psi_i]}$. The problem is solved.

Two ideas are involved in Fine's solution to the measurement problem. The first is that we ought to be fairly pragmatic about how we choose to model physical processes in general. The standard way of treating the interaction of two systems is to start with the system in a state which is the product of the states of the components. But why should we do this if it has failed to give the right (i.e. observed) answer in certain classes of situations? We ought to be prepared to modify this if it doesn't work.

The second is the idea of interacting with an aspect, which underpins the whole strategy. What is it that we know in general about an observable for a given state? If we stick to the interpretative rules that Fine outlines, then we cannot say that we know specific values all the time. What we can know are probability distributions. These are in some way physically recognised by the measuring instruments, which filter out, in a measurement, all the physical features that are not relevant to the particular kind of observation being made.

Fine's idea is original in asking what quantum measurement interactions might actually be doing, something not very common in discussions of the quantum measurement problem. He tries to solve the problem of measurement by providing an answer to this, through the idea of a *selective interaction*. Measurements at the quantum level are interactions where an apparatus records only certain, 'selective', aspects of the object it is measuring. This seems intuitively plausible from our patterns of observation. We observe contextually and if we concentrate on some aspect of what we observe we miss out many other features. It is also true that measuring instruments, like a weighing scale, ignore shape and size. Fine uses *this* idea to derive his conclusions about the states to assign to the system to be measured.

2. Algebraic Quantum Theory and Fine's Solution

In his first paper on this topic [40] Fine mentions the algebraic approach to quantum mechanics as possibly providing a rationale for the particular procedure which models the measurement process. I want to explore this further.

What Fine means by a rationale is that there ought to be a subalgebra of an algebra of operators as defined in Algebraic Quantum Theory, which does the job of "picking

out” the appropriate restriction on the states of a measured system. This section shows how it is possible to define an algebra such that a set of states can be associated with it according to the definition of a quantum system in Algebraic Quantum Theory. It will be clear how this set of states corresponds to Fine’s restricted set. The system composed by the algebra and the states can be used to provide a model for measurement interactions equivalent to Fine’s. The present discussion will also highlight in which way Fine’s modification of the ‘initial state’ condition can be thought of as equivalent to the modification of the ‘quantum systems’ condition.

Fine’s pragmatism about modelling certain interactions is well served by the definition of a system in Algebraic Quantum Theory. In this definition no particular constraints are placed on what the observables ought to be, and no particular representation of states and observables is privileged above others. The only requirement that is given basically ensures that you have enough to calculate expectation values, while not having so much that you might create confusion or redundancy. A system is defined as an algebra satisfying certain postulates, together with a set of states which is “full” with respect to the algebra. States here are intended as maps from the operators in the algebra to the real numbers, which give you expectation values for these operators.

The crucial idea is that the set of states should be full. There will be lots of different functions mapping the operators to the real numbers. Full sets select a subset of these functions satisfying certain criteria. There are different equivalent definitions of the term *full*: in Rieckers [71] we find that a convex subset S_0 of the set of states S is full (with respect to the algebra A on which it is defined) if $a \in A$ and $\langle \varphi; a \rangle \geq 0$, for all $\varphi \in S_0$, implies $a \geq 0$, that is fullness is defined in terms of positivity. Kadison [51] defines a set of states to be full if and only if it is $\sigma(S, A)$ -dense in $S(A)$. However, in order to discuss this it is best to stick to Segal’s definition [76]. Segal defines a convex subset to be full if and only if, for every pair of distinct observables, there exists at least one state for which these observables have different expectation values. Segal shows how, for all systems, it is possible to find a full set of pure states.

EXAMPLE 4.1. Suppose we define our set of observables to be the set \mathcal{O} of 2×2 real matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with $ad - bc \neq 0$, and our states to be the set \mathcal{F} of functions $\varphi_{x,y} := \mathcal{O} \rightarrow \mathbb{R}_+ = ax + dy$, $x, y \geq 0$, $x + y = 1$. Clearly none of these states can distinguish between any of the observables where a and d are fixed, and b and c are variable, as the states calculate the expectation values using only a and d . So this set of states for our observables is not full.

If a set of states is not full with respect to some algebra, this means that we have observables whose expectation values cannot be distinguished (maybe infinitely many, as in the previous example). The set of observables in a quantum system is normally the set of all self-adjoint operators on a Hilbert Space and the set of states is the set of density operators (positive, trace class 1 operators on the same Hilbert Space). Given any pair of observables, there is always a state that will yield a different expectation value for these observables, so the set of density operators is full. But, for a given measurement interaction, Fine restricts the possible states that can be assigned to a system to be measured, to a set essentially like that in example (4.1), and these states will no longer form a full set. Relative to that interaction there will be observables which cannot be distinguished and, given the states we have, they will be for all purposes the same observable. If, as in Fine's case, we want to restrict the set of states that we want to assign to a system, then the set of observables that can be distinguished in such a system will also be restricted and it is not difficult to see what they will be.

EXAMPLE 4.2. Suppose we restrict the set of observables from the previous example to the set

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

Now any two different observables A_1 and A_2 can only differ in a or in b . If they differ in a , then the state $\varphi_{1,0} = a \times 1 + b \times 0$ will lead to different expectation values for A_1 and A_2 ; if they differ in b , then $\varphi_{0,1} = a \times 0 + b \times 1$ again will give different expectation values. The set of states \mathcal{F} is then full with respect to the restriction on the observables that we have outlined.

Practically the same happens to the algebra of observables for the quantum system in Fine's solution to the measurement problem. The restriction in example (4.2) leaves us with a commutative algebra of observables. The observables all share the

same eigenstates. In Fine's solution for quantum measurement we start from a set of observables which we can represent as matrices written in the basis of the observable that we want to measure:

$$\begin{pmatrix} a & b+ic \\ b-ic & d \end{pmatrix}$$

Fine's allowed states can be expressed as in example (4.1) and hence, if we fix a and d , none of the observables so defined can be distinguished by these states. The restriction leads to the subalgebra of observables which commute with the observable which we want to measure, i.e. all operators with $b = c = 0$.

2.1. An example for $\text{spin-}\frac{1}{2}$ systems. The following properties define a C^* -algebra.

DEFINITION 4.1. 1. \mathcal{A} is a distributive and associative algebra over \mathbb{C} if \mathcal{A} is a linear vector space over \mathbb{C} and has an associative product which satisfies for $\lambda_i \in \mathbb{C}$ and $a_i, b \in \mathcal{A}$

$$\begin{aligned} (\lambda_1 a_1 + \lambda_2 a_2)b &= \lambda_1 a_1 b + \lambda_2 a_2 b \\ b(\lambda_1 a_1 + \lambda_2 a_2) &= \lambda_1 b a_1 + \lambda_2 b a_2. \end{aligned}$$

2. \mathcal{A} is a Banach $*$ -algebra if it is an algebra with a norm, i.e. there is a mapping

$$\| \cdot \| : \mathcal{A} \rightarrow \mathbb{R}_+$$

with $(a, b \in \mathcal{A}, \lambda \in \mathbb{C})$

$$(2.1) \quad \|a\| = 0 \iff a = 0$$

$$(2.2) \quad \|a + b\| \leq \|a\| + \|b\|$$

$$(2.3) \quad \|\lambda a\| = |\lambda| \|a\|$$

$$(2.4) \quad \|ab\| \leq \|a\| \|b\|.$$

Furthermore, \mathcal{A} has to be complete with respect to the norm topology. Finally, there exists a mapping

$$* : \mathcal{A} \rightarrow \mathcal{A},$$

called an involution, with

$$(2.5) \quad (a^*)^* = a$$

$$(2.6) \quad (a + b)^* = a^* + b^*$$

$$(2.7) \quad (\lambda a)^* = \lambda^* a^*$$

$$(2.8) \quad (ab)^* = b^* a^*.$$

3. \mathcal{A} is a C^* -algebra if

(a) \mathcal{A} is a Banach * -algebra

(b) $\|a^* a\| = \|a\|^2, \forall a \in \mathcal{A}$.

DEFINITION 4.2. Let \mathcal{A} be a C^* -algebra. A state is defined as a functional

$$\varphi : a \in \mathcal{A} \mapsto \langle \varphi; a \rangle \in \mathbb{C}$$

which satisfies the following properties:

1. $\langle \varphi; \lambda_1 a_1 + \lambda_2 a_2 \rangle = \lambda_1 \langle \varphi; a_1 \rangle + \lambda_2 \langle \varphi; a_2 \rangle$
 $a_i \in \mathcal{A}, \lambda_i \in \mathbb{C}$
2. $\langle \varphi; a^* a \rangle \geq 0, \forall a \in \mathcal{A}$
3. $\langle \varphi; 1 \rangle = 1$.

\mathcal{S} denotes the set of all states. This set is a convex set. The elements of this set that are not expressible as a convex combination of any states are called pure states.

DEFINITION 4.3. A convex subset \mathcal{S}_0 of \mathcal{S} is full (with respect to \mathcal{A}), if for all $a, b \in \mathcal{A}, a \neq b$, there exists a state $\varphi \in \mathcal{S}_0$ such that $\langle \varphi; a \rangle \neq \langle \varphi; b \rangle$.

DEFINITION 4.4. The pair $(\mathcal{A}_{\text{sa}}, \mathcal{S}_0)$, \mathcal{A}_{sa} the self-adjoint part of a (C^*) -algebra, \mathcal{S}_0 a full set of states, is called a description (model) of a (class of) quantum system(s).

Note in particular that under the restriction of \mathcal{A} to \mathcal{A}_{sa} , the states map elements of the restricted set (which is *not* a subalgebra) to \mathbb{R} ; $\langle \varphi; a \rangle$ corresponds to the expectation value of a for the state φ . \mathcal{A}_{sa} does, however, uniquely determine every state in \mathcal{S} : if \mathcal{S} is full with respect to \mathcal{A}_{sa} , it is also full with respect to \mathcal{A} . It should be stressed that Definition 4.2 is not the only way in which states for a quantum system might be defined, and that different ways might lead to different sets of states when the Hilbert space we are considering is infinite dimensional. For the purpose of discussing Fine's

solution to the quantum measurement problem, however, we will only deal with an example involving finite dimensional systems.

2.2. Algebraic approaches to Fine's solution. The set of states that are relevant in the case of a measurement of spin in a spin- $\frac{1}{2}$ system, according to Fine's description, is not full with respect to the set of bounded hermitean operators on the spin space. These states correspond to the functions

$$\varphi : \mathcal{A} \rightarrow \mathbb{R} = \text{Tr}(\rho a),$$

where

$$\rho = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix},$$

$x + y = 1$, and

$$a = \begin{pmatrix} p & H_R - iH_I \\ H_R + iH_I & q \end{pmatrix}$$

Consider two observables A and B , one with $H_R = H_I = 0$, the other with $H_R \neq 0$, say. It is easy to see that no such state will be able to distinguish between the two.

Fine's question can now be phrased in the following way: is there some algebra (hopefully a subalgebra of the C^* -algebra from which the quantum system is defined) with a full set of states isomorphic to the one Fine picks out? The answer is yes.

Consider the set of ordered pairs (a, b) of complex numbers. Define scalar multiplication of the elements of this set by a complex number as $\lambda(a, b) = (\lambda a, \lambda b)$. Define addition by $(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$, and multiplication by $(a_1, b_1)(a_2, b_2) = (a_1 a_2, b_1 b_2)$. This is indeed an algebra according to the above definition¹. It will be denoted by \mathcal{C} .

Now consider the mapping

$$|| \cdot || : (a, b) \rightarrow \mathbb{R}_+ = \max[|a|, |b|],$$

where the operation $|\cdot|$ yields the absolute value of a complex number. This is a norm for the above algebra.

PROOF. Check conditions (2.1, 2.2, 2.3, 2.4).

- (2.1) requires that $||(a, b)|| = 0 \iff (a, b) = 0$, i.e $a = b = 0$. This is clearly the case.

¹As a consequence of this it also satisfies Segal's postulates I , as stated in [76].

- (2.2) requires that $|(a_1 + a_2, b_1 + b_2)| \leq |(a_1, b_1)| + |(a_2, b_2)|$. Suppose without loss of generality that the number on the left hand side is $|a_1 + a_2|$. Two cases arise:

1.

$$\max[|a_i|, |b_i|] = |a_i|, \quad i = 1, 2.$$

Then $|(a_1 + a_2, b_1 + b_2)| \leq |(a_1, b_1)| + |(a_2, b_2)|$ holds if and only if $|a_1 + a_2| \leq |a_1| + |a_2|$, which holds as a property for the operation of taking absolute values of complex numbers.

2. In the second case, one of $|b_1|$ or $|b_2|$ is greater than the respective $|a_i|$; suppose it is $|b_1|$. It clearly follows that $|b_1| + |a_2| > |a_1| + |a_2|$ the latter being of course greater than or equal to $|a_1 + a_2|$, as we have just seen.

This completes the proof.

- (2.3) requires that $|(\lambda a, \lambda b)| = |\lambda| |(a, b)|$. Clearly the case as, for any two complex numbers z and z' , $|z \cdot z'| = |z||z'|$.
- The proof of (2.4) is similar to that of (2.2). □

We define the involution for the algebra \mathcal{C} to be the operation $(a, b)^* = (\bar{a}, \bar{b})$, where \bar{a} is the complex conjugate of a . This obviously satisfies all the postulates described above for such maps, as the operation of complex conjugation is an automorphism of \mathbb{C} . The $*$ -algebra condition then becomes, when this involution map is adopted, $|(a\bar{a}, b\bar{b})| = |(a, b)|^2$. Now suppose $|a| \geq |b|$, then $|(a, b)| = |a|$ and $|(a, b)|^2 = |a|^2$. If $|a| \geq |b|$ then $|a\bar{a}| \geq |b\bar{b}|$ and $|(a\bar{a}, b\bar{b})| = |a\bar{a}|$. But $|a\bar{a}| = |a| \cdot |\bar{a}| = |a| \cdot |a| = |a|^2$, and the condition is satisfied, so the algebra is a C^* -algebra.

The states for this algebra are the linear functions from the algebra to the complex numbers. Consider the functions defined by

$$\mathcal{S}_0 \ni \varphi_{x,y} : (a, b) \rightarrow \mathbb{C} = ax + by,$$

with $\mathbb{R} \ni x, y \geq 0$ and $x + y = 1$. These are indeed states according to the conditions laid out in Def. 2.2: $\langle \varphi; A^*A \rangle = |a|^2x + |b|^2y \geq 0$ clearly, which fulfils the positivity condition; $\langle \varphi; 1 \rangle = x + y = 1$, satisfying the normalisation condition; and

$$\begin{aligned} \langle \varphi_{x,y}; (\lambda_1 a_1 + \lambda_2 a_2, \lambda_1 b_1 + \lambda_2 b_2) \rangle &= (\lambda_1 a_1 + \lambda_2 a_2)x + (\lambda_1 b_1 + \lambda_2 b_2)y \\ \lambda_1 \langle \varphi_{x,y}; (a_1, b_1) \rangle + \lambda_2 \langle \varphi_{x,y}; (a_2, b_2) \rangle &= \lambda_1 a_1 x + \lambda_2 a_2 x + \lambda_1 b_1 y + \lambda_2 b_2 y \end{aligned}$$

which are clearly equal, so linearity is satisfied. Note that there are only two pure states, $\varphi_{1,0}$ and $\varphi_{0,1}$.

This set of states is full with respect to \mathcal{C} . Suppose we have two different elements of \mathcal{R} , A_1 and A_2 , say. If they differ in the first complex number, then $\varphi_{1,0}$ will lead to different expectation values for A_1 and A_2 ; if they differ in the second number, then $\varphi_{0,1}$ again will give different expectation values. So it is always possible to find a state for differing elements of \mathcal{C} that will give different expectation values, and hence the set of states is full. It is worth noting also that these functions exhaust all the possible states for the system.

The self-adjoint part of the algebra \mathcal{C} , namely the subset \mathcal{C}_{sa} of the algebra which is invariant under the involution operation, together with this set of states \mathcal{S}_0 , forms *by definition* a quantum system. What do the representations of the algebra \mathcal{C} look like? I will briefly show that, through a GNS construction, we can reach the two possible representations of this algebra (one is irreducible, the other is not). Through these representations we obtain a Hilbert space realisation of the quantum system $(\mathcal{C}_{sa}, \mathcal{S}_0)$. The GNS method of constructing representations for algebras with sets of states relies on the calculation of left ideals for the algebra. A state is fixed and the representation is constructed with respect to that state. It is a theorem of the theory of quantum algebras that such a representation is irreducible if and only if the corresponding state is pure. In our case there will be two representations that arise from the algebra: one for the pure states and one for the mixed ones. This highlights the structural features of this algebra.

Once a state $\varphi \in \mathcal{S}_0$ is chosen, the first part of the GNS construction consists of calculating the set

$$I_\varphi = \{A \in \mathcal{C} : \langle \varphi; A^* A \rangle = 0\},$$

which is provably a left ideal of the algebra. For present purposes, take the state to be $\varphi_{\frac{1}{2}, \frac{1}{2}}$. Then $I_{\varphi_{\frac{1}{2}, \frac{1}{2}}}$ consists of all the elements (a, b) of \mathcal{C} for which

$$\langle \varphi_{\frac{1}{2}, \frac{1}{2}}; (|a|^2, |b|^2) \rangle = \frac{|a|^2}{2} + \frac{|b|^2}{2} = 0.$$

The only way in which this equality can be satisfied is if $a = b = 0$, hence $I_{\varphi_{\frac{1}{2}, \frac{1}{2}}} = \{(0, 0)\}$. Now write

$$\Phi_A = A + I_{\varphi_{\frac{1}{2}, \frac{1}{2}}}.$$

The set $\mathcal{C}/I_{\varphi_{\frac{1}{2}, \frac{1}{2}}} = \{\Phi_A : A \in \mathcal{C}\}$ with scalar product given by

$$(\Phi_A, \Phi_B) = \langle \varphi_{\frac{1}{2}, \frac{1}{2}}; A^* B \rangle$$

is a Hilbert Space. The mapping

$$\pi_{\varphi_{\frac{1}{2}, \frac{1}{2}}} : \mathcal{C} \rightarrow \mathcal{B}(\mathcal{C}/I_{\varphi_{\frac{1}{2}, \frac{1}{2}}})$$

given by $A \rightarrow \pi_{\varphi_{\frac{1}{2}, \frac{1}{2}}}(A)$,

$$\pi_{\varphi_{\frac{1}{2}, \frac{1}{2}}}(A)\Phi_B = \Phi_{AB}$$

links every A with a bounded operator on $\mathcal{C}/I_{\varphi_{\frac{1}{2}, \frac{1}{2}}}$, and is easily seen to be a *-morphism.

Furthermore let $\Phi_{\mathbb{I}} = \Phi$, \mathbb{I} being the element $(1, 1)$ of the algebra, then

$$(2.9) \quad \langle \varphi_{\frac{1}{2}, \frac{1}{2}}; A \rangle = \langle \varphi_{\frac{1}{2}, \frac{1}{2}}; 1^* A \rangle = (\Phi_{\mathbb{I}}, \Phi_A) = (\Phi, \pi_{\varphi_{\frac{1}{2}, \frac{1}{2}}}(A)\Phi),$$

and Φ can easily be shown to be cyclic².

The Hilbert space is easily seen to be isomorphic to \mathbb{C}^2 , and the representation of the algebra on this space is given by the following set of 2×2 complex matrices on \mathbb{C}^2 :

$$\pi_{\varphi_{\frac{1}{2}, \frac{1}{2}}}(a, b) = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}, a, b \in \mathbb{C}.$$

The involution operation is easily seen to be equivalent to the adjoint operation on these matrices.

This representation is not irreducible, because it is based on a mixed state. To obtain an irreducible representation we can repeat the above procedure for one of the two pure states of the system. It is easily seen³ that this representation will yield the set of linear maps on \mathbb{C} . This is an irreducible representation. What this shows is that the first representation is made of two copies of the irreducible representation which act independently of one another on two subspaces of \mathbb{C}^2 . This means, in quantum mechanical terms, that nothing like an interference term can arise in this system. The first representation has the property that for every unit vector $\Phi \in \mathcal{H}$ the state φ defined by

$$\langle \varphi; (a, b) \rangle = (\Phi, \pi_{\varphi_{\frac{1}{2}, \frac{1}{2}}}(a, b)\Phi)$$

²The condition of being cyclic, as well as condition (2.9), are standardly satisfied by representations constructed through the GNS scheme, though they need not concern us explicitly here.

³From the fact that I_{φ} in this case will have a continuum of elements of the algebra in it.

is a state in \mathcal{S}_0 . Consider the representation of (a, b)

$$\begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}$$

and $\Phi = (x, y)$ with

$$\frac{x^2 + y^2}{2} = 1.$$

Then

$$(\Phi, \pi_{\varphi_{\frac{1}{2}, \frac{1}{2}}}(a, b)\Phi) = (x, y)(ax, by) = \frac{ax^2 + by^2}{2}.$$

Put $x' = \frac{x^2}{2}$, $y' = \frac{y^2}{2}$, then $x' + y' = 1$ and $x', y' \geq 0$, and $\phi_{x', y'}$ clearly defines a state. This means that the representation is a representation for $(\mathcal{C}, \mathcal{S}_0)$ and a fortiori for the quantum system $(\mathcal{A}_{sa}, \mathcal{S}_0)$.

2.3. An answer to Fine's question. It is worth emphasising that the quantum system defined by the self-adjoint part of \mathcal{C} together with the states \mathcal{S}_0 is a system that can model Fine's selective interaction measurements, as its states have the same structure as Fine's equivalence classes.

We have just seen, from an abstract point of view, what the algebra induced by Fine's restriction on states will be. But remember that Fine's original question concerning algebraic approaches was the following: is it possible to find an algebra containing a subalgebra that in some sense induces the restriction on the states? I have just shown that there is an algebra that has a full set of states which is isomorphic to Fine's equivalence classes. What is the algebra that can have this as a subalgebra?

The answer in the case just discussed is simple, and in general relies on the so-called *Gelfand-Naimark theorem* (see [45] and [32, p. 375]). This result states that *all* C^* -algebras are 'concrete', i.e. they can be realised as a C^* -subalgebra of the C^* -algebra of bounded operators on a Hilbert space \mathcal{H} . The obvious algebra for the example is the algebra of bounded operators on \mathbb{C}^2 . This algebra (over the field of complex numbers) is *linearly* generated by the following operators:

$$(2.10) \quad \mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$(2.11) \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \text{ and } \sigma_z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

the identity operator on \mathbb{C}^2 and the Pauli spin matrices. The subalgebra of operators

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

is linearly generated by the two operators

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and is the subalgebra of operators commuting with the spin x observable

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

Note that this subalgebra contains *no* smaller subalgebras within it: a subalgebra must be closed under matrix multiplication as well as addition, but $\sigma_x^2 = \mathbb{I}$ so that $\alpha\sigma_x + \beta\sigma_x^2$ linearly generates the subalgebra commuting with σ_x . In particular, Fine's remark that the algebra of observables which are functions of a given observable, $\mathcal{F}(E)$, is different from the algebra of observables commuting with a given observable is not always correct [40, p. 502–503]: it fails in particular when the observable is maximal, in which case the two are equal. Fine wants to argue that there is a subset of all the bounded, self-adjoint operators of a quantum system (the usual observables), which justifies selecting the states he does for the object system. At the same time he does not want this subset to be the set of all observables commuting with the observable to be measured. This will become important later on, when I argue that the fact that this does not hold in important cases constitutes a problem for the solution presented.

In sum, the answer to Fine's question is the following: given a Hilbert space, and an observable to be measured on the object modelled in such a space, the restriction from the algebra of bounded operators on the Hilbert space to the subalgebra of observables commuting with the observable to be measured induces Fine's restriction to the states he argues are to be assigned to the object in his proposed solution to the measurement problem.

This line of reasoning does more: it also provides a rationale for why the assigned states should be expressed as density operators decomposed over eigenvalues of the measured observable, something which Fine thinks is decided pragmatically: "If E -equivalent states are actually identical relative to the subsystem represented by $\mathcal{F}(E)$, then how can we justify using any one of them, say $W(E)$ [the density operator over

object observable eigenstates], as opposed to any other, say W ? I think the only answer can be that E -equivalence warrants using one or another state, indifferently, provided that the particular use yields satisfactory results" [40, p. 502]. Under such a restriction of the algebra of relevant observables, eigenstates of the measured observable are the *only* pure states remaining⁴: there is *no* state W left in the algebra, there are only such pure states and the convex combinations of them⁵.

3. Arguing for the algebraic rationale

At this stage I will address two issues. The first thing to deal with is whether there is any scope for an algebraic account to back up Fine's proposed solution to the measurement problem, and what problems it might lead to. This, from a formal point of view, involves looking at Fine's solution from the standpoint exposed in the previous section. This means, for instance, that we have to think of the quantum mechanical object system *not* as an ordinary quantum system, but as a system for which considerably fewer observables and states are defined. Otherwise the suggestion that we are looking at this from the algebraic point of view is vacuous. The algebraic definition of a quantum system *is* very weak when compared, for instance, to the idea that quantum mechanics should be modelled by a quantum system of bounded operators on a Hilbert Space: this latter system satisfies the algebraic definition, but so do infinitely many others. What Fine calls "precedents and rules of thumb" come into play in deciding what quantum model is applicable to which physical system. First of all I will explain, in this section, why an *algebraic* justification of Fine's strategy amounts to a justification by means of a superselection rule. I will then examine in the next section whether questions that arise with respect to superselection rules can pose problems for Fine's solution to the measurement problem. Finally in section 5 I will consider whether precedents and rules of thumb can sufficiently back up Fine's strategy, independently of whether its justification comes from algebraic quantum mechanics.

⁴This can be seen in many different ways. For example they are the only states which are eigenstates of some observable of the system; they are the only states which have probability one with respect to some (non-trivial) observables, and so on.

⁵Another way of looking at this is to think that a pure state can be used to represent an ensemble which cannot be split, in accordance with D'Espagnat's definition of splitting. Then it is easy to see that, in the presence of a superselection rule as in Fine's examples, the only pure states are eigenstates of the measured object observable [27, p. 63, Remark 1].

As anticipated on page 140, the discussion of the previous section implies that, if we try to give an algebraic rationale to the solution, then what Fine has done is *formally*, though not necessarily in the details of the argument, equivalent to imposing a superselection rule, therefore modifying the condition that the object and apparatus systems both be proper quantum systems. Simply put, superselection rules forbid superpositions between certain pure states by first restricting the set of observables. This is usually done by claiming that all operators which are to count as observables must commute with a certain operator for the system (the superselection operator), thereby revoking the status of observables for operators which normally would represent observables, namely those that do not commute with the superselection operator. The set of states for the larger algebra of operators containing all original observables now contains states which are equivalent with each other with respect to the observables to which the superselection rule restricts us. This leads to an equivalence class just like Fine's. While the starting point is the opposite of Fine's, in that the restriction begins at the level of observables rather than at the level of states, the end result is formally the same.

I mean this in at least two senses. First of all, if we attempt to justify Fine's strategy by appeal to an algebraic argument, we clearly are in the same position as if we assumed a superselection rule to apply on the object space *when, and only when, a measurement is being carried out*. We might, on the other hand, want to argue that the states we assign when measurements are being carried out represent physical states of the object's aspects, rather than of the object itself. Nevertheless the object's aspects have a state structure equivalent to that imposed by a superselection rule.

It is worth stressing again that it is the same from a formal point of view. The mathematical model of the object system that is relevant in the solution of the measurement problem that Fine proposes is the same, *when accounted for by algebraic reasoning*, as a model induced by a superselection rule on the object system with the measured observable counting as the superselection operator. Because this does not necessarily imply that we have exactly the same physical case as we would with a superselection rule, it is worth considering in some detail whether objections to superselection rules straightforwardly carry over to objections to Fine's account. There are, however, at least some differences between Fine's strategy and superselection rules which clearly cannot rescue Fine.

First of all, it might be argued that Fine thinks measurement evolutions look at aspects, not states, and that states are therefore still distinguishable, while aspects are not. This strategy might seek to block the inference that we are dealing with something formally equivalent to a superselection rule in quantum measurements by claiming that there still is a fact of the matter about which state the system is in while a measurement is performed. This seems to me to fail in at least two ways. Firstly, if we treat aspects in measurement, it is completely irrelevant to measurements themselves whether there are still states attached to object systems or not, for we ignore them completely when accounting for measurements: all the relevant information is encoded in the representatives of aspects and all the final information about the state of the object system *after the measurement* is recovered from the final evolved *aspect*. If we don't deduce the final object state in this way, then we don't solve the measurement problem according to the rules Fine sets out.

Secondly, therefore, if someone were to object that there are two 'facts of the matter', so to speak, about a system in measurement, namely what state it is in, and what aspect it is in, then the present account should be read as an analysis of the object aspect + apparatus interaction. As this is all that is used in Fine's account of measurement, a critique of this analysis will be a critique of Fine's solution regardless. Also we might add that there is simply no way of assigning a state to the object during measurement from any of the physical information available in this process, in fact the whole idea of modelling measurements through aspects is that all information about the original state is lost in the measurement process, apart from the probabilistic information relative to the measured object observable. This seems to make it rather hard to tell what the state of the object system, as distinct from its aspect, might be in the measurement process.

A second objection to the formal identification of the algebraic justification of Fine's state restriction with a superselection rule might be that superselection rules represent *prohibitions* for certain superpositions to exist (Nature just doesn't allow these states to arise), while Fine's interaction with an aspect is the result of knowledge of a specific causal process, and it is this causal process that backs up the restriction of the algebra. While this might be true, it is hard to see what bearing this would have on any problematic aspects of the formal scheme that we are dealing with here. If the formal model is inconsistent or contradictory or unable to deal with the problem at hand in

any sense, it matters little what kind of story we give for its adoption. In the next parts I will be raising problems for the formal aspect that becomes relevant if we try to use algebraic quantum mechanics to justify Fine's solution. Of course any problems that emerge here will not by themselves rule out Fine's solution, but only a reconceptualisation of it in terms of algebras. The distinct arguments that lead to Fine's solution, the causal arguments about interactions with aspects, will be dealt with in section 5.

We have just seen at the end of the previous section, on page 149 that Fine in his original paper tentatively proposes a distinction between two subsets of the algebra of observables of a system: the function algebra associated with an observable E and the algebra of operators commuting with E , the latter being the set analysed in the previous section; the two are supposed to be different, and it is the first such subset that should be of help in defining states for measurement the way that Fine does. But in our \mathcal{H}^2 case, if the function algebra is the subalgebra generated from E by the usual operations of addition and multiplication, it is the same as the commutative subalgebra discussed earlier. The reason for wanting to identify different such subsets is that the set of all operators commuting with a given one contains observables which are qualitatively different to the original one, rather than just, for example, being multiples of it.

This makes things complicated: instead of finding a rationale for assigning states in the way that Fine does, we find that in so doing we have effectively imposed a superselection rule, and invoking superselection rules in order to solve the measurement problem has many difficulties. Note that the problems thrown up by this formal analogy are not solely technical, but also philosophical. There is nothing wrong with Fine's strategy if the aim is to find a way of redefining conditions (1)–(4) in such a way that they are consistent. The objections against Fine's solution have a flavour of “good, but not good enough”, the solution is felt to be cheap, or *ad hoc*. Implicitly, in raising such points, the argument is that the measurement problem is not summarised by the requirement of modifying (1)–(4) to make them consistent; the implication is that there is more to the measurement problem than doing just this.

Recall how the first chapter argues, amongst other things, that measurement is about answering two kinds of questions, about property assignment and unitary evolutions. Chapters 2 and 3 have picked on two important conditions which emerge in discussions of these questions: they are, in this chapter, condition (1.2), the probability reproducibility condition; and condition (1.3), or objectification, which Fine calls the

Rule of Silence. However, the general unease about Fine's solution suggests that it is not sufficient to find ways of dealing with these conditions. The ways must also be acceptable in some other sense, for instance by being physically motivated, rather than just mathematically consistent. Much of the unease with superselection rules as a solution to the measurement problem arises precisely because measurement superselection rules often seem to be not so well motivated, and a lot of work has gone into providing such motivations in terms of well understood physical processes which might generate them. Fine's solution should be appraised in the same way.

The first step is to ask what kinds of answers Fine's solution gives to the two 'technical' questions characterised in Chapter 1: the answers are that properties are assigned according to the eigenvalue-eigenstate rule *and* the interaction evolutions are deterministic. On top of this there is an important 'metaquestion' to tackle: *what is so special about quantum measurement?* This question will in some sense affect the answers to the first ones: a discussion of it requires arguing whether measurement is a special process in quantum mechanics, a process that requires separate treatments in some sense from other kinds of interactions, and showing quite how different the process is. However the connection is not total: for example different modal interpretations give the same kind of technical answers to the two questions in Chapter 1, though they have different view as to how special quantum measurement is.

There is a sense in which measurement is trivially going to be a special interaction: it relies on different Hamiltonians than other interactions do. This would obviously be the case with interpretations of the measurement process such as Everett's. In many interpretations, however, there is more 'specialisation' to measurement than just being described by a different kind of Hamiltonian. In collapse solutions, for example, a standard linear Hamiltonian evolution of the object + apparatus system is followed by a non-linear evolution. When invoking a superselection rule (in the indirect way in which Fine does it as much as, I think, in other superselection approaches) it is also necessary to admit that there is something peculiar about measurement.

For example, in the Coleman-Hepp scheme (see, for instance, [49]), further discussed by Bub [12], suppose that the measurement system, contrary to any other system in non-relativistic quantum mechanics, is properly modelled by an infinite number of Hilbert spaces all direct summed together. In the most recent and perhaps most interesting defence of a superselection solution to these issues, Landsmann [61] argues

that we have in measurement a particular physical process of decoherence which gives rise to superselection rules. Fine effectively says that what is special about measurement is that the measurement interaction only involves aspects of the quantum system. Therefore it is the nature of a measurement interaction that the appropriate way of representing such interactions differ from what it would usually be.

It is important to think of these solutions to the measurement problem as singling out measurement processes and to give a justification as to why this is an acceptable procedure. This is suggested by the general unwillingness of many to accept descriptions of systems with superselection rules at face value, as quantum systems on a par with the standard model of a Hilbert space and its associated set of self-adjoint operators. The point here has been made by Beltrametti and Cassinelli [5] and reiterated by Van Fraassen. Somehow, potential empirical evidence of quantum behaviour with interferences warrants description by a Hilbert space and its associated operators, as is the case with claims of macroscopic superposition; but empirical evidence of absence of interference is not warrant enough for asserting a superselection rule. This is because such liberal use of superselection rules would allow us to call all classical phenomena *de facto* quantum phenomena: “classical behaviour could never disconfirm quantum predictions [80, p. 266]”.

Superselection rules and their like ought, in this picture, to be deducible from some less objectionable condition⁶. It shouldn't be possible to invoke them just because phenomena suggest their presence, as most classical behaviour would then suggest, in a sense, the presence of a superselection rule. If this is reason enough to assume these rules then it is thought that there is no problem with showing that classical mechanics is just a special case of quantum mechanics.

This should make it clear that providing a solution to the measurement problem requires a little more work than just establishing an axiom scheme which is a consistent modification of (1)–(4). The solution also involves, for instance, defending the modification of the axioms from possible criticisms.

With regards to Fine's solution and its formal rationale in terms of algebraic quantum mechanics, a number of objections against superselection solutions to the measurement problem have a bearing on this, and some other objections can arise. The next

⁶A good example of this is Wan and Fountain's derivation of a superselection rule for supercurrents governed by the Josephson equation [84].

section will consider these objections and see whether Fine might be able to answer them. Assuming that a reply to this might be that we can drop the attempt to provide such a rationale, I will then look specifically at Fine's arguments to the effect that measurement is a special physical process in order to raise other problems, independent of the rationale.

4. Problems for Fine and Superselection

One standard problem that people have with adopting superselection rules arises just as much in Fine's proposal.

When we have a superselection rule, our state space is split into supersectors; states in different supersectors cannot exist in superpositions with one another. The unitary evolution generated by a self-adjoint operator O on a Hilbert space has an interesting property: it will not map any superposition of eigenstates of the self-adjoint operator O into an eigenstate of O and it will not map an eigenstate of O into a superposition of said eigenstates. Consider the consequence of applying a superselection rule to the apparatus space of an object + apparatus system $\mathcal{H}_S \otimes \mathcal{H}_M$, by selecting a self-adjoint superselection operator $\mathbb{I} \otimes P$, where P is the apparatus pointer observable. By definition this rule will have as a consequence that the only observables for the pointer system commute with $\mathbb{I} \otimes P$, and hence share the property just described with $\mathbb{I} \otimes P$. Thus a state which is in a superselection sector, which is defined by the eigenstates of $\mathbb{I} \otimes P$ cannot be mapped out of this sector by any unitary operators generated by observables.

Now suppose that we have an initial state $P_{[\varphi \otimes \psi_0]}$, with $\mathcal{H}_S \ni \varphi = \alpha_i \varphi_i$. The initial state of the apparatus ψ_0 might be an eigenstate of the pointer observable recording that no observation has been made. If it is an eigenstate of the pointer observable, and if the pointer observable is a superselection operator, then, by the well known fact outlined in the previous paragraph, no Hamiltonian representing an observable can generate an evolution mapping this state into a state which has anything other than ψ_0 as a pointer "substate" in the subsystem \mathcal{H}_M . The unitary operator which will map the initial state to the appropriate final state which can count as a measurement must then be generated by a self-adjoint operator which is not an observable. Normally one regards the self-adjoint operator generating the unitary evolution of a system as the energy of the system, in this state the composite system $\mathcal{H}_S \otimes \mathcal{H}_M$. The problem

then is that either the energy of the total system is not an observable, or the evolution of the system $\mathcal{H}_S \otimes \mathcal{H}_M$ in measurement is not dictated by an operator representing the energy of the system. This has been thought by many to be an objection to the superselection account; Hughes [50, 9.7], for instance, has argued this.

Suppose we try to finesse this by regarding ψ_0 to be a superposition of pointer reading states, $\mathcal{H}_M \ni \psi = \beta_i \psi_i$, and pointing out that, as only the pure states lie in supersectors $\mathcal{H}_S \otimes \psi_i$, all states must be expressed as convex combinations of these states. Then the initial state before the measurement interaction will be $\sum_i |\beta_i|^2 P_{[\varphi \otimes \psi_i]}$. Now the initial state is already split along the correct supersectors, and we could claim that some sort of solution is at hand. But the probability reproducibility condition will fail. We have to change the probabilities $|\beta_i|^2$ in the measurement evolution to reflect the probabilities given through φ , and it is well known that we still can't do it if the Hamiltonian of the system is an operator which is an observable.

On the other hand some people claim that it is not necessarily clear what observable the self-adjoint operator generating the evolution of the system $\mathcal{H}_S \otimes \mathcal{H}_M$ is supposed to represent. It's not the case that the quantity that we are trying to measure within the system (its energy, say) is not an observable for the joint system. Suppose the observable measured on \mathcal{H}_S is E . Then $E \otimes \mathbb{I}_M$ is an observable for the joint system, whatever superselection rule we impose on the system \mathcal{H}_M . It is the Hamiltonian describing the interaction between the two systems which is not an observable of the system. But the interpretation of such a Hamiltonian is not so straightforward, and there are examples from classical mechanics, for example, where the Hamiltonian of the classical system does not represent its energy. Wan has long made this latter point in several papers, recently in a joint paper with Harrison [83]; Landsmann [61] also thinks this is not such a fatal problem, though his point is less convincing.

The issue, however, is still unresolved; quantum mechanics is very much based on the idea that energy observables dictate evolutions of systems. If they are not so dictated in other theories this does not seem sufficient reason for why they shouldn't be in quantum mechanics. And even in the case that the Hamiltonian oddity is accepted, Hughes argues that it still offers no account of why a superselection rule operates with respect to certain macroscopic systems, but not others. Many arguments for superselection rules explain these in terms of the fact that they are macroscopic. But certain

macroscopic systems (Hughes's example is a pot of liquid helium) do not exhibit behaviour consistent with a superselection rule. Yet measurement invariably does. There still is no answer to the question: "What is so special about Quantum Measurement"?

Van Fraassen [80] further points out that superselection rules might entail different empirical predictions in many ways in quantum mechanics; this need not be a problem, but will require careful assessment of what we do when applying them. Perhaps, in the case of standard pointer superselection rules, the only option really is to give up trying to absolutely define what counts as macroscopic and microscopic and turn this into an empirical fact: measurements exhibit macroscopic behaviour, Josephson junctions (perhaps) and liquid helium pots do not, and that is all there is to it. There still is need for an argument, however, as to why classical mechanics isn't trivially a special case of quantum mechanics, given that, as pointed out in the previous section, most classical behaviour would then suggest, in a sense, the presence of a superselection rule.

This problem affects Fine's restriction of states, as justified through algebraic quantum mechanics, and the corresponding superselection rule it entails, just as much. The solution needs to explain, just as much as a solution based on superselection rules, what basis there is for a state evolution determined by a Hamiltonian for the system which is not an observable, or why this should not be relevant in such solution. Suppose that Fine decides that the problem with Hamiltonians is not so desperate after all, it is still the case that an algebraic reformulation of the solution achieves little in the way of providing a rationale of Fine's state selection, for we cannot show that there is a justification for the state selection unless we give a justification for the algebraic restriction. This we cannot do without addressing more closely the question of interactions with an aspect, what the interactions are, how they really do apply, in the same sense, as I have claimed in the previous paragraph, that we have to carefully assess how superselection rules apply.

A more delicate question is whether, quite apart from an explicit appeal to an algebraic reformulation of Fine's solution, the latter implies a superselection rule on the object system anyway. In a sense, the restriction on states that Fine advocates necessarily restricts observables in the way that is done in a superselection rule, effectively acting as Ockham's razor: it is not possible to tell that there are observables of the system other than the ones that would remain after a superselection rule is imposed, so on what basis can we assume them to be there? Note that the point here

is not that we have a problem in deciding *when* we have a measurement, or whether the object state before a measurement is properly represented by probabilistic aspects; it's about whether, given that we know a measurement is about to take place and we have assigned an aspect state, we can tell whether there are any other observables for this system other than the restricted set of observables commuting with the measured object observable.

One fact is clear: given that we have certain representations of object aspects at the beginning of a measurement, then there is no way that in those states we can distinguish an observable which does not commute with the measured object observable from one that does⁷. A further consequence of this is that the claim that what the measurement evolution on the combined system $\mathcal{H}_S \otimes \mathcal{H}_M$ is about is interacting with an aspect cannot easily be substantiated by appeal to some straightforward relationship between the observable measured on the subsystem \mathcal{H}_S and the measurement Hamiltonian on $\mathcal{H}_S \otimes \mathcal{H}_M$; we can't even see the latter!

In any case, we will now leave aside the question of the extent to which Fine's state selection amounts to the same thing as imposing a superselection rule in order to consider more closely in the next section the question of interaction with an aspect independently of this problem.

5. Two further problems for Fine's account

Fine [43, p. 281], citing Bell, claims to be looking for an approach to solving the measurement problem that satisfies the desideratum that "fundamental theory permit exact mathematical formulation" [4, p. 171]. His solution is certainly mathematically precise: from the perspective explored in the previous sections it amounts to assuming a superselection rule. Von Neumann's projections, though, are just as mathematically precise: there is no problem in understanding their formal properties. The problems emerge when trying to gain a physical understanding of what process projections are supposed to model, and when trying to interpret this process.

⁷Note also that another way of putting this is that, just like we pick an equivalence class of states with respect to the measured object observable, the reduced set of states we work with also establishes an equivalence class on the whole set of self-adjoint operators on a Hilbert space, those that cannot be distinguished by any one of the states in the restricted set, as shown in example 4.1

In the same way it is a mistake to judge Fine's proposals, good or bad as we may think them to be, on the basis of how precise their mathematical formulation is. What is the physics behind these proposals? What interpretive gloss do they put on quantum mechanics?

The gloss is explicitly pragmatic; insofar as the pragmatism holds, it is best to steer away from the polemics it raises: one person's pragmatism is another one's *ad hocness*. I want to take, for the sake of argument, the pragmatism for granted and examine in some detail the first question.

The argument in Fine's solution for why measurement is a physical process deserving of special treatment has two components. A measurement of an observable is looking for probabilistic information about an observable O in an initial object state $P_{[\varphi]} \in \mathcal{T}_1^+(\mathcal{H}_S)$, information that is to be transferred to a final pointer state $F_{\mathcal{H}_M} \in \mathcal{T}_1^+(\mathcal{H}_M)$; this state should emerge by partial tracing of a final state $F \in \mathcal{H}_S \otimes \mathcal{H}_M$ with suitable properties. But there is a whole class of object states which yields, via the Born rule, the same probabilistic information about O . Were the measurement problem to be solved, we would expect, for all such object states, initial object + apparatus states formed from them to be mapped onto a final state $F \in \mathcal{H}_S \otimes \mathcal{H}_M$ for which $F_{\mathcal{H}_M}$ is always the same.

This is one way of looking at the measurement problem: we want a bijective mapping on $\mathcal{T}_1^+(\mathcal{H}_S \otimes \mathcal{H}_M)$, a generalisation of a unitary operator, but we end up asking for all these states which yield the same probabilistic information to be mapped to the same final state for the apparatus, with appropriate probabilities. The requirements force us quite explicitly in the case of ideal measurements to have the same global final state for all such measurements when initial states are E -equivalent. Fine proposes, first of all, that all the states that share the same probabilistic information with respect to an observable to be measured be assigned the same initial state. Measurement is about interacting with an aspect of the system to be measured, namely the aspect concerning the probabilistic information that the measured system yields, and we should encode this aspect in a representative state.

Secondly, which state should be the representative state? The answer is "pragmatic" : we initially assign to the system a state which will guarantee the correct outcome at the end. There is more than a similarity with appeal to superselection rules here. It is the easy way out, some might say. Put more seriously, we might ask

what physical basis there is to such an appeal: simply saying that the apparatus system interacts with an 'aspect' of the object system does not seem, by itself, to provide much grounds for changing the states we assign to the object system; if we have a peculiarly specific interaction we would normally describe it with a peculiarly specific Hamiltonian. Why should a peculiar interaction in the case of measurement be accounted for by revising the initial object states, prepared at the beginning in quite a different way? On the other hand, as we have seen in previous sections, the idea that algebraic considerations might provide a rationale for Fine doesn't solve much, as one is left with a superselection rule and the need to provide a rationale for that.

A response might be that the specific structure of an interaction is not just modelled by a Hamiltonian, but we face the further choice of whether the Hamiltonian is a Hamiltonian for the evolution of standard states or of states representing aspects. In any case, we need to look a bit more closely at how selective interactions work with revising initial object states. I want to raise two problems for Fine's solution which do not strictly depend on assuming that a superselection rule is in operation on the object system when a measurement is in progress.

The first problem is related with the selection of representative states, and concerns, for a change, the way that the solution deals with non-ideal measurements. The issue has been raised before, by Stairs, and Fine has responded to this. The way the issue is discussed in the previous debate concerns approximate measurements, but as chapter 1 makes clear, non-ideal measurements need not be interpreted as approximate measurements. In this case there is still a problem to discuss.

The second problem raises the possibility that the intuition that lies behind the solution, that of interaction with an aspect, is not adequately captured by the formal scheme that effects the solution, and so focuses on the first component of Fine's argument. This intuition is what really distinguishes Fine's approach from a superselection based scheme, and it wants to be a *physical* intuition. It forms the basis of the argument for modifying the initial state. The claim is essentially that it is hard to see how the mathematics invoked by Fine is connected to this intuition, and it is hard because we don't really understand what the Hamiltonian of the joint system is doing, and because Fine's state restriction is restriction to the probabilistic aspect of not one observable, but of an infinite number of them, enough in some cases to describe a whole classical system.

5.1. Fine's solution and non-ideal measurements. If the measurement process is to be distinguished in Fine's reading of it by the fact that it formally introduces a superselection rule on the observed system *at the beginning of measurement*, it is to be expected that such a rule will stand throughout the measurement process; indeed we would not normally expect to be able to switch superselection rules on and off. Still in general, even if we don't tie Fine's solution to the measurement problem with the idea of superselection, it seems reasonable to assume that *at least* during the whole measurement interaction the only features that bear on the evolution of the object subsystem are those represented by the representative states, from beginning to end of the interaction, as one is indeed interacting with an aspect throughout the measurement. It is difficult to see why, if such process of interacting with an aspect justifies assigning Fine's representative states at the beginning of the interaction, it should not do so throughout the interaction.

This will work well for ideal measurements, but will leave us unable to predict the final outcome of some non-ideal measurements. The argument is again somewhat similar to the well known and often discussed argument of Albert and Loewer. Stairs has raised some questions about possible Albert and Loewer measurements, questions that Fine has answered on the basis of a disagreement about the correct way to model *imperfect* measurements. The example I have in mind has more to do with genuinely non-ideal measurements, and nothing to do with inaccurate or imperfect measurements: rather than considering inaccuracies with measurement operations resulting from detectors 'misfiring', I will discuss examples which, if they are to be captured by standard quantum measurement techniques at all, must be captured by non-ideal measurements, along the lines of the examples in Chapter 1.

Fine claims that we have "interaction with an aspect" or, put in another way, "*E* selective interaction". What entitles us to talk of $[V]_E$, as defined above, as the appropriate state at the start is at least the fact that this state is going to be transformed in a selective way.

I think that we can't make the *E* aspect state selection go away with any kind of argument until *after* the final state has been reached; little seems to change throughout the measurement that would make the selective interaction argument go away. This would, however, entail that the final object state when a measurement ends will still be a mixture of eigenstates of the object observable, as it was at the beginning. A

non-ideal measurement then will result in a failure to predict the correct final object states for certain particular kinds of measurements, like for example destructive photon measurements.

To recapitulate what we mean by a non-ideal measurement, consider for example the mixed state

$$\sum_{i=1,2} |\alpha_i|^2 P_{[\varphi_i \otimes \psi_0]} \in \mathbb{C}^2 \otimes \mathbb{C}^2,$$

representing the initial state of a measurement interaction with the initial object state $\sum_{i=1,2} |\alpha_i|^2 P_{[\varphi_i]}$ being the representative state of $\varphi = \alpha_1 \varphi_1 + \alpha_2 \varphi_2 \in \mathbb{C}^2$ for an observable picked by the set $\{\varphi_1, \varphi_2\}$. The system is subjected to a non-ideal measurement with the pointer in the initial state ψ_0 , yielding

$$\sum_{i,j=1,2} |\alpha_{ij}|^2 P_{[\varphi_j \otimes \psi_i]},$$

where $\sum_j \alpha_{ij} \varphi_j = \tilde{\varphi}_i$.

If this is a non-ideal measurement, the states $\tilde{\varphi}_i$ may well not be orthogonal; in the case of destructive measurements, as considered initially by Margenau and further discussed by Kronz, all such states will be equal to a 'vacuum' state φ_0 , as we have seen in chapter 1. Other examples of situations where measurements are necessarily non-ideal are discussed by Ruetsche [74] who, in the context of modal interpretations, analyses measurements of object system observables represented by operators commuting with operators representing conserved quantities of the system, as a consequence of the Wigner-Araki-Yanase result [2, 87]. This result in fact establishes that ideal measurements of such object observables are impossible.

Supposing we have an initial photon polarisation state $\varphi = \alpha_1 \varphi_1 + \alpha_2 \varphi_2 \in \mathbb{C}^2$ and we measure an object observable selected by the object system basis $\{\varphi_1, \varphi_2\} \subset \mathbb{C}^2$, by means of a pointer observable picked out by the orthonormal basis of the apparatus space $\{\psi_1, \psi_2\} \subset \mathbb{C}^2$. At least two things can happen in Fine's account under a unitary, non-ideal measurement evolution U applied to the initial state $\sum_{i=1,2} |\alpha_i|^2 P_{[\varphi_i \otimes \psi_0]}$.

1. The final state has the form

$$(5.1) \quad \sum_{i,j=1,2} |\alpha_i \langle \tilde{\varphi}_i | \varphi_j \rangle|^2 P_{[\varphi_j \otimes \psi_i]}$$

2. the final state has the form

$$(5.2) \quad \sum_{i=1,2} |\alpha_i|^2 P_{[\tilde{\varphi}_i \otimes \psi_i]}.$$

Note that, for the two states, the final object states are respectively

$$\sum_{i=1,2} \left(\sum_{k=1,2} |\alpha_k \langle \tilde{\varphi}_k | \varphi_i \rangle|^2 \right) P_{[\varphi_i]}$$

and

$$\sum_{i=1,2} |\alpha_i|^2 P_{[\tilde{\varphi}_i]}$$

which are clearly different states, as can be easily checked by calculating, for instance, the transition probability of the two states with respect to $P_{[\tilde{\varphi}_1]}$.

In the first case clearly the final state fails to predict in any way what the final object state will be: state (5.1) suggests that the final state of the object will be a density operator decomposed by eigenstates of the object observable that is being measured, not necessarily what we expect from an irreducibly non-ideal measurement.

On the other hand, the final state might be thought to be (5.2), the correct one for predicting the final object state. Then a curious situation arises. Right at the beginning of the selective measurement interaction Fine tells us that, as measurements interact selectively, we should write object states as mixtures over the measured observables. During the selective interaction itself, however, this would have to no longer apply in order to have a final state such as the one in (5.2). This in itself would be a strange feature of selective interactions, but we could claim that, for pragmatic reasons, state assignment is done before the actual interaction arises. It is possible, however, to 'freeze-frame' the interaction, which happens over a period of time from 0 to t , at a time t' such that $0 < t' < t$ and ask what is the state at time t' , before the remaining part of the selective interaction takes place. At this point there seems little argument for not applying the same state assignment rule that we apply at the beginning of the interaction. This sort of argument rules out, however, the possibility that the final object state be a state other than an eigenstate of the measured object observable, or a density operator over such eigenstates.

The problem can be stated as 'how long should the E aspect be relevant?' and calls for some reflection. The interesting idea in Fine's solution, as remarked earlier, is that he wants to consider what the measurement interaction is actually doing. Given

that the postulated nature of the interaction (a 'selective' one) is what accounts for the solution, the answer to 'how long should the E aspect be relevant?' should be 'at least as long as the measurement is going on and continues to be the kind of interaction that involves aspects', regardless of what happens before or after (and regardless of whether we explicitly claim that there is a superselection rule in operation).

This is difficult to assess as the question turns on whether the fact that the interaction relies on aspects at the start truly *implies* that it relies on aspects all the way. In a sense it seems that there need not be any implication of this sort, and it might be a purely physical fact that interactions are interactions with an aspect for the first x 'instants', and thereafter cease to be so. Perhaps the interaction only needs x instants to pick out the aspect and then goes its own way, so to speak. While this being *physically* the case would undoubtedly avoid all the problems just described which are associated with non-ideal measurements, I think it will need a much more clear analysis of quite how measurement operations are interactions with aspects than we have seen so far: how would the distinction between what is going on in the first x instants and what is going on afterwards be captured in the formal model? And to what different physically well understood, perhaps causal process do these two types of interactions correspond to? I actually think that the distinction will be hard to maintain, and will give some reasons for this in the next subsection.

As we have just seen, if Fine is forced to concede that interactions are interactions with aspects throughout the measurement, this will lead to problems: if measurement specifically interacts with E aspects, what is the argument for saying that it might yield something which is informationally useful insofar as it isn't encoded in the E aspect story (like knowledge of the difference between states which share the same probability distribution with respect to E), as happens with the final states after a non-ideal interaction? Yet if we selected state (5.2) we would be expecting the final state to encode information which is not obtainable in the E aspect approach: it will make a claim for the object state to be in a specific state which is not an E aspect state normally selected by Fine's argument.

It might be argued that because in the case of non-ideal measurements there is a clear case for needing to produce information, at the end of a measurement, incompatible with the E aspect of the object subsystem, then we simply *shouldn't* in general expect the measurement itself to concentrate on and produce an outcome expressed

only in terms of the E aspect. But this will leave the E aspect as relevant only at the point of initial state assignment, independent of what happens in the measurement. I find this difficult to accept in any case; empirically, and in Fine's own argument, this is not an initial state assigned for any sort of measurement or indeed any interaction, but only for a concrete measurement of a very specific observable, and it is the measurement interaction itself which selects the aspect. If we then assume that in measurement the object state is not confined to the E aspect, but is free to produce, for example, whatever final state of the object we want compatible with a probabilistic condition being satisfied for the final apparatus state, this seems to take away a good deal of the force of the argument.

5.2. What are measurement evolutions selectively interacting with? The second problem for Fine comes from the nature of the measurement operator. If I am placed in front of a final mixture resulting from a measurement and shown the measurement interaction operator, I have no idea what observable was being measured here. This is because in constructing the measurement interaction operator all I rely on is the orthonormal system resulting from an observable, not its eigenvalues. But this implies that all observables which share the same set of eigenstates are "measured" by the same operator. Normally this is not a very important consideration, perhaps wrongly; not much is said about the measurement operator other than the fact that it "kicks" pointers into place. Also, as stated by equation (2.4) in Chapter 3, the standard account of what makes up a measurement requires that we have, as well as an operator characterising the measurement evolution, a function linking the pointer observable's eigenstates, accessible to us by "direct" observation, with the object observable ones, thus specifying in full the measured observable, rather than just its eigenstates. I will return to this soon.

In Fine's story we are told what measurements are supposed to do: a measurement effects the recording of a particular aspect of an observable, its probabilistic one. But this aspect is common to all observables found in the algebra picked out by a superselection rule in the following way. Recall from Chapter 3, page 106 that one of the conditions necessary to define a premeasurement operator (which for Fine will be the measurement operator) is that there is a fixed orthonormal basis of the object system

\mathcal{H}_S , made up of eigenstates of the measured observable (where the basis is not necessarily unique if the measured observable is not maximal). This premeasurement operator measures all observables which share the orthonormal basis as a set of eigenstates. This set of observables forms an algebra picked out by choosing any maximal operator in this set to be a superselection operator. This is clearly a commutative algebra, for the observables are diagonalised by the same orthonormal basis. The system is therefore 'classical', as can be seen in a number of different ways. For instance the following theorem, due to Segal, applies:

THEOREM 4.1 (Segal [76]). *A commutative system is isomorphic (algebraically and metrically) with a system of all real-valued continuous functions on a compact Hausdorff space.*

More intuitively what is happening is that every algebra of commutative observables has a number of properties that indicate its classicality. The sense in which the system is classical is given by the fact that all pure states for such a system are dispersion-free, or that the set of observables separates the set of probability measures, so that the phenomenon of non-unique decomposability of mixtures does not arise for such a system, while it is typical of quantum systems.

So the measurement operator can't be interacting with an aspect of a particular observable, as there is nothing whatsoever in it which distinguishes between an infinite set of observables which is, moreover, non-trivial in that it is equivalent to a classical system.

There is a sense in which this is a problem for Fine's account of measurement, and not for other accounts. The unitary operators effecting measurements are standardly derived in the manner discussed in the previous chapter: assume some conditions which you think a measurement operator should satisfy, and look for which operators *do* satisfy this condition. The 'essentialist' question, about what the measurements actually are, is not even raised. No discussion whatsoever is given of the nature of the interaction. Think of Everett's solution, or of the Kochen-Dieks modal interpretation: what is proposed is a rule for property assignment, given that the measurement is characterised by unitary evolutions equivalent to the ones discussed in the previous chapter. These measurements happen to be such that the interaction measures a whole big set

of observables. Whether this is the case or not becomes a matter of empirical investigation, perhaps. And perhaps the function which, in the standard definition of what constitutes a measurement, links eigenstates of the pointer observable to those of the object observable, as defined in equation (2.4), has a physical interpretation such that a single observable can be said to be measured. But to point out that the unitary operator does not select a specific observable is not to make a criticism of the approach to solving the measurement problem embodied in Everett's or Kochen's work, for *no assumption* is made there about the nature of the interactions.

The very fact that no assumption is made about the physical nature of the measurement interaction might in itself be the basis of a criticism that covers very many approaches to quantum measurement; this came up briefly in chapter 1. But that is not the argument here. The point here is that Fine builds his solution to the measurement problem on a specific intuition. This intuition is that we measure contextually, so that if two objects have the same weight (in the everyday sense of the term), this tells us very little about the objects as a whole: *when* we perform a measurement, then, we interact only with a specific aspect of the object, the one relevant to its weight. But the measurement interactions that are implicitly used in Fine's solution (those that don't violate (1.2), (1.3) and (1.1)) cannot be assumed to be doing anything of the sort, because of the problem just presented: they do not measure energy alone, for instance, but a whole set of observables commuting with energy.

Note that this is not meant to be a slide into realist speak. When we assume that the Hamiltonian for the harmonic oscillator is a good Hamiltonian to *use* in certain situations, we do so because of its properties, and this can be read in any which way we want, as Fine himself has argued in other contexts [39, 41]: such a fact, that the Hamiltonian is good for modelling certain systems, is independent of whether we argue for this on the basis of its reliability or its closeness to the truth. On similar grounds we can argue against measurement operators being good operators to model selective interactions. So the point is not that selective interactions don't exist, or are an unreasonable way of thinking of measurement interactions, but that measurement operators picked out by (1.1), (1.2) and (1.3) have the wrong sort of properties to do justice to the intuitions that Fine invites us to think about.

It is, I think, quite hard to see what interacting with an aspect might mean here. Fine himself, in his original paper [40, p. 503], wants to distinguish between interaction

with an aspect of a single observable from interaction with an aspect characterised as the joint measurement of all observables commuting with a given one; these would be two different processes, leading to a theory based on individual measurements on the one hand, and to a generalised theory of joint measurements. There is no distinction between these two kinds of processes, though, on the analysis given here, for the important cases in which the observable is maximal, and this includes observables such as spin, angular momentum and many others.

Again this creates problems for the selective interaction idea. We can make this idea plausible by pointing to common experiences of sense perception. When we try to ascertain the colour of an object, we are not much concerned with its size; similarly when we try to measure the hardness of some material we are not worried by its colour. But quantum measurement evolutions don't focus on a specific observable, they measure effectively a whole class of them, those that share the same eigenstates. This is possibly more confusing as an intuitive physical idea of measurement than interaction with an aspect is, and the latter isn't served well by the former; this is not, of course, an argument against the viability of selective interactions as an approach to quantum measurements, but an argument about the difficulty of reconciling the intuitive motivation and the general argument of the approach with the mathematical properties of the operators actually describing the interaction.

6. Conclusions

Fine's idea is original in asking what quantum measurement might actually be doing, something not very common in discussions of quantum measurement. He tries to solve the problem of measurement by providing an answer to this, through the idea of selective interaction. This seems intuitively plausible from our patterns of observation: we observe contextually, if we concentrate on some aspect of what we observe we miss out many other features. It is also true that measurements, like weighing, ignore shape and size. Fine uses this idea to derive some conclusions about the states to assign to the system to be measured, which can be interpreted as a justification for a superselection rule.

On the whole, though, I think this does not work. It doesn't work not because the intuitive idea is wrong, but because the formal scheme of quantum measurement as set up by Von Neumann bears little relation to the intuitive ideas that Fine suggests. A

measurement operator is not measuring a specific observable; it is not at all clear what physical observable the Hamiltonian generating it stands for; it can assume all sorts of different forms, all essentially characterising the same measurement process, as shown in chapter 3, Theorem 3.2. All these militate in some way or another against Fine's philosophical argument for selective interaction, but I think the questions remain, and remain interesting: what do measurement operators physically stand for? what exactly are they supposed to measure? are they a reasonable abstraction of real measurements?

CHAPTER 5

Conclusions

It is useful at this stage to briefly review what this thesis has accomplished.

In the first instance, Chapters 2, 3 and 4 have proved a series of results, which I briefly review.

Chapter 2 expands both the kinds of results known as insolubility proofs of the quantum measurement problem, and their scope. I have discussed two new proofs of the original result. Neither add mathematically to the scope of the well known results of the past forty years. Both, however, highlight important facts about the way in which the objectification condition for quantum measurement constrains the possibility of transferring probabilistic information about the measured observable from the object to the apparatus system through a unitary operator.

The first, detailed in sections 2 and 2.3, shows in a fairly intuitive way how objectification implies that the probabilities collected from the apparatus system for a pointer observable after a measurement has been performed is just the information that we could have collected from the apparatus *before* any measurement had been performed on the combined object + apparatus system. Objectification therefore kills off the possibility of any meaningful sense of *unitary* measurement being possible. The second result, detailed in section 4, is a trivial consequence of Theorem 2.7, which characterises all unitary operators satisfying objectification. This result shows how the previous claim, that a unitary operator satisfying objectification makes no difference to probabilities for the pointer observable, is actually possible: in the simplest case, for example, the theorem shows the unitary operator to be of the form $U_S \otimes U_M$, with no interaction term present.

Another consequence of this result is that, in the course of proving it, I establish that the insolubility result due to Fine and Brown is in fact just as valid as other ones presented, even without the use of the condition that Brown has called Real Unitary Evolution. The discussion of the issues revolving around this particular version of the proof is reviewed in section 3.

Finally the chapter provides, in section 5, a discussion of how the objectification condition might be interpreted in the context of unsharp pointer observables, and gives a first insolubility proof for this case. This genuinely extends the scope of the original proofs.

Chapter 3, on the other hand, addresses the status of the probabilistic condition usually assumed to hold for quantum measurement independently of objectification holding. It begins with a more specific result, Theorem 3.2 (developed with Gianpiero Cattaneo) along the lines of Beltrametti, Cassinelli and Lahti. This characterises, in subsection 2.2, unitary operators satisfying the *probability reproducibility condition* for the Hilbert space formalisation of quantum mechanics. Sections 3 and 4 look at how this kind of result might be extended to the density operator formalism. A number of results are proved, and the general conclusion is that the only one-to-one mappings that satisfy the probability reproducibility condition in this context are still the ones characterised through Theorem 3.2.

Making use of the analysis of sections 3 and 4, section 5 raises some problems for the possible unitary operators satisfying the probability reproducibility condition. In particular I show that there are cases for which *no such operators* can exist, and discuss the consequences of this.

Chapter 4 is an analysis and critique of Fine's solution to the measurement problem. The starting point of the chapter is a discussion of a proposal by Fine, to ground his solution in an algebraic framework. The formal implications of this are analysed in section 2. If we presuppose an algebraic viewpoint, Fine's solution is equivalent to the imposition of a superselection rule on the *object* space. The problems that this raises are discussed in section 4. Section 5 looks at some other difficulties with Fine's solution which are independent of the appeal to an algebraic grounding.

I want to conclude this work with some remarks on how the material just outlined fits in with the view of quantum measurement I have defended in Chapter 1, and with some open problems that are raised from this view. Recall that in chapter 1 I have defended the view that the discussion of the quantum measurement problem is a fundamentally theoretical enterprise. Chapter 2 discusses a classic aspect of such work, the consistency proof. The conditions that I analyse there are put forward with an attempt to answer the questions I have marked in Chapter 1, about properties and

dynamics. From a theoretical point of view it is crucial to establish whether such conditions are consistent. That they are not is, as I have said in conclusion to chapter 2, crucial in motivating further research on how we might answer the two questions.

Chapter 3 on the other hand provides us with a particular instance of a problem, rather than a global no-go theorem. It flags for attention a special class of examples for which unitarity and probability reproducibility cannot both hold. The problem here is not global inconsistency, as is the case in chapter 2. What we have is a specific counterexample, but one which has, in my opinion, greater force against the project of the quantum theory of measurement in virtue of originating from a theoretical analysis. This does not require us, then, to enter into arguments about the extent to which theoretical accounts of measurement mirror real measurements, for the counterexample is not drawn from a real measurement situation, but generated internally.

Chapter 4 highlights the difficulty in reconciling the claims of the theoretical analysis of measurement with what happens in real measurements. Fine's support for his solution depends crucially at a conceptual level on the claim that his solution accounts for what happens in real measurements, namely interactions with aspects. The problem is that he uses the technical apparatus provided by the standard theoretical account of measurement, and it is quite difficult to see in what way this carries out Fine's interactions. This is because, in general, the quantum theory of measurement pays relatively little attention to the nature of the interactions that generate measurements. I think this is of no major consequence if the game is to answer questions about properties and dynamics. However, if a solution to the measurement problem is built on the claim that measurement interactions are of a specific kind, and moreover that this is the kind of interaction that one always sees happening in measurement, then it ought to show that this is the case.

This is an open problem in general. There is little interest normally in the status of quantum measurements as real interactions. I think this is of crucial importance to the possible applicability of the formalism of quantum measurement, but also to the typically philosophical questions about the relationship between theory and experiment.

Another problem that emerges from the view I put forward in chapter one is of a more historical nature. There is no properly developed account of how we have moved from discussing interpretations of experiments at the beginning of the century to the more abstract topic that is quantum measurement today. If the analogy with Kuhn's

description of a thought experiment, briefly alluded to in chapter 1, is to hold, we will need to understand this much better.

While this thesis has not addressed these two issues, I think much of the work in it leads to them. I hope it will be of use in this way, as well as of value for all the things contained in it.

APPENDIX A

Proofs of Theorems

PROOF OF COROLLARY 2.5. Two mutually exclusive and exhaustive cases must be distinguished:

1. φ, φ' are nonorthogonal initial states in \mathcal{H}_S ;
2. φ and φ' are orthogonal.

The first case is the one discussed in Proposition 2.4. This Proposition says that assuming that pointer states with different eigenvalues are present in a possible decomposition of the final object + apparatus state will lead to a violation of unitarity, on the assumption that $\langle \varphi | \varphi' \rangle \neq 0$. So the pointer eigenstates decomposing the final state must be associated with the same set of eigenvalues if unitarity and objectification are to be satisfied.

The second case can be treated by considering a third state, for instance

$$\chi = \frac{\varphi + \varphi'}{\|\varphi + \varphi'\|},$$

which is not orthogonal to the two orthogonal states. Then applying the previous argument to φ and χ , and then to φ' and χ , yields that the pointer eigenstates associated with initial states φ and χ are the same, and pointer eigenstates associated with φ' and χ are the same. The result then follows for φ and φ' . \square

PROOF OF COROLLARY 2.6. Two cases must be distinguished, according to whether the pointer observable has multiple eigenvalues or not.

For the first case, consider an orthonormal basis of possible initial states in \mathcal{H}_S , for instance $\{\hat{\varphi}_1, \dots, \hat{\varphi}_n\}$, such that $\langle \hat{\varphi}_i | \hat{\varphi}_j \rangle = 0$ for $i \neq j$. The set of initial pointer states $\{\psi_{0i}\}$ is also an orthonormal set. Therefore the set of states in $\mathcal{H}_S \otimes \mathcal{H}_M$ consisting of vectors of the form $\hat{\varphi}_j \otimes \psi_{0i}$ is also an orthonormal set. Furthermore the subspaces spanned by sets $\hat{\varphi}_k \otimes \psi_{0i}$ and $\hat{\varphi}_l \otimes \psi_{0i}$ with k, l fixed and distinct, and i variable, are orthogonal subspaces. If the evolution applied to initial states $\hat{\varphi}_k \otimes \psi_{0i}$ and $\hat{\varphi}_l \otimes \psi_{0i}$ is unitary, the final density operators for initial object states $\hat{\varphi}_j$ with different j 's must also span orthogonal spaces.

Assume that the final states admit of a representation over eigenstates of the pointer observable $I \otimes A_M$ where the different ψ_i 's, eigenstates of the pointer observable A_M , number only $m - 1$, and suppose in the first instance that the eigenvalues of A_M have multiplicity 1. Given Corollary 2.5, it follows that, for initial states $\hat{\varphi}_j$, the final object + apparatus states after the measurement evolution will have two different projections of the form $P_{[\tilde{\varphi}_{jk}]} \otimes P_{[\psi_i]}$, with i fixed for each j and $k = 1$ or 2 . Consider one such final state corresponding to the case in which $j = 1$, $i = 1$. Then the states $\tilde{\varphi}_{1k}$, for different k 's, must be distinct in order to have distinct projections $P_{[\tilde{\varphi}_{1k}]} \otimes P_{[\psi_1]}$; failure to have distinct projections would lead to a violation of conservation of rank.

All other final states must have at least one pointer eigenstate of the form $P_{[\tilde{\varphi}_{jk}]} \otimes P_{[\psi_1]}$, k varying for the different initial states, again because of Corollary 2.5. Consider the projections $P_{[\tilde{\varphi}_{j1}]} \otimes P_{[\psi_1]}$, j varying. These projections must be mutually orthogonal as they belong to subspaces which are unitarily evolved from initial orthogonal subspaces. This implies that the $\tilde{\varphi}_{j1}$'s must form an orthonormal basis for \mathcal{H}_S . But then the projection $P_{[\tilde{\varphi}_{1,2}]} \otimes P_{[\psi_1]}$ associated with the initial object state $j = 1$ cannot be orthogonal to projections of the form $P_{[\tilde{\varphi}_{j1}]} \otimes P_{[\psi_1]}$, $j \neq 1$, as it is distinct from $P_{[\tilde{\varphi}_{1,1}]} \otimes P_{[\psi_1]}$. This gives a contradiction. Therefore the $P_{[\psi_i]}$'s must form a set of m distinct, orthogonal projections in the final evolved states.

If the pointer observable A_M has multiple eigenvalues, the claims still hold. Note, first of all, that the argument of the previous paragraph is not substantially changed if we have ψ_1 appearing twice as an eigenstate of A_M in the final state for an initial object state $\hat{\varphi}_1$, and ψ_2 appearing in the final state for some $\hat{\varphi}_2$, where ψ_1 and ψ_2 share the same eigenvalue and are not orthogonal. What happens if on the other hand ψ_1 and ψ_2 appear as pointer eigenstates indicating the same eigenvalue, and are orthogonal? A straightforward application of the argument in Proposition 2.4 and Corollary 2.5 shows that, if ψ_1 and ψ_2 appear in any one final apparatus state decomposition, then they must appear in all of them. Separate arguments can then be made for them, similar to the ones just given in the previous paragraph.

As for the second part of the claim, suppose that we do have two pointer eigenstates ψ_1 and $\psi_2 \in \mathcal{H}_M$ appearing in the decomposition of the final state, which are eigenstates of A_M associated with the same eigenvalue and are not orthogonal. Then we clearly have two projections in the decomposition over distinct vectors in \mathcal{H}_M , which however are not orthogonal. An orthogonal resolution of the final state, spanned by the projections

containing ψ_1 and ψ_2 , must exist, however, by the same reasoning given on page 70. This finally establishes the result. \square

PROOF OF THEOREM 2.7. $[\Rightarrow]$: The previous results, culminating in Corollary 2.6, establish that if a unitary operator $U : \mathcal{H}_S \otimes [\psi_{0i}]_i \rightarrow \mathcal{H}_S \otimes [\psi_{ijk}]_{ijk}$ satisfies objectification, then it must map an orthonormal basis $\hat{\varphi}_j \otimes \psi_{0i}$ to the orthonormal basis $\sum_{k=1}^{l_i} \tilde{\varphi}_{ij} \otimes \psi_{ijk}$, where i, j and l are variable and $\langle \psi_{ijk} | \psi_{i'jk} \rangle = 0$ for all $i \neq i'$ and any j, k .

For a fixed, the ψ_{ajk} 's must be eigenstates of the same eigenvalue of the apparatus pointer observable A_M , for if they were not the initial object state $\alpha_1 \hat{\varphi}_1 + \alpha_2 \hat{\varphi}_2$ would, when coupled to ψ_{0a} , be mapped to the final state $(\sum_{k=1}^{l_i} \alpha_1 \tilde{\varphi}_{i1k} \otimes \psi_{i1k}) + (\sum_{k=1}^{l'_i} \alpha_2 \tilde{\varphi}_{i2k} \otimes \psi_{i2k})$, which can only be an eigenstate of $\mathbb{I} \otimes A_M$ if ψ_{i1} and ψ_{i2} are eigenstates of A_M associated to the same eigenvalue.

Then, for a, b fixed,

$$\begin{aligned} U(\hat{\varphi}_b \otimes \psi_{0a}) &= \sum_{k=1}^{l_i} \tilde{\varphi}_{abk} \otimes \psi_{abk} \\ &= \sum_k (\langle \hat{\varphi}_b | \hat{\varphi}_b \rangle (\tilde{\varphi}_{abk}) \otimes \langle \psi_{0a} | \psi_{0a} \rangle (\psi_{abk})) \\ &= \sum_k \left(W_{ab}^S \otimes W_{ab}^M \right) (\hat{\varphi}_b \otimes \psi_{0a}). \end{aligned}$$

Repeating this for each a, b and using linearity shows that (4.1) holds.

The operators composing U in (4.1) are clearly partial isometries; the proof is given for one of them. Consider $W_{abc}^S = \langle \hat{\varphi}_b | (\cdot) \rangle (\tilde{\varphi}_{abc})$. Its action on any operator $\alpha \hat{\varphi}_b$ is given by $\langle \hat{\varphi}_b | \alpha \hat{\varphi}_b \rangle (\tilde{\varphi}_{abc}) = \alpha \langle \hat{\varphi}_b | \hat{\varphi}_b \rangle (\tilde{\varphi}_{abc}) = \alpha \tilde{\varphi}_{abc}$, so the operator is unitary as a mapping from $[\hat{\varphi}_j]$ to $[\tilde{\varphi}_{abc}]$.

Also for any vector $\sum_{i \in I} \alpha_i \hat{\varphi}_i$, where $I = \{1, \dots, b-1, b+1, \dots, m\}$,

$$\left\langle \hat{\varphi}_b \left| \sum_I \alpha_i \hat{\varphi}_i \right. \right\rangle (\tilde{\varphi}_{abc}) = \sum_I \alpha_i \langle \hat{\varphi}_b | \hat{\varphi}_i \rangle (\tilde{\varphi}_{abc}) = 0;$$

the operator is clearly linear. It follows that W^S_{abc} is indeed a partial isometry. Next, note that

$$\begin{aligned} \sum_{k=1}^{l_i} \langle \hat{\varphi}_j \otimes \psi_{0i} | (\hat{\varphi}_j \otimes \psi_{0i}) \rangle (\tilde{\varphi}_{ijk} \otimes \psi_{ijk}) &= \sum_k \tilde{\varphi}_{ijk} \otimes \psi_{ijk} \\ &= \sum_k W^S_{ijk}(\hat{\varphi}_j) \otimes W^M_{ijk}(\psi_{0i}) \\ &= \sum_k W^S_{ijk} \otimes W^M_{ijk}(\hat{\varphi}_j \otimes \psi_{0i}), \end{aligned}$$

establishing (4.2). It is easy to show that any two such operators W, W' are partial isometries, and that $\langle W(\Phi) | W'(\Phi) \rangle = 0$.

[\Leftarrow]: Consider U given by (4.1). U is then obviously linear. In order to show that it is unitary, consider an arbitrary vector $\Phi \in \mathcal{H}_S \otimes [\psi_{0i}]_i$. Then

$$\begin{aligned} \|U(\Phi)\|^2 &= \left\langle \sum_{\substack{f=1,\dots,n \\ g=1,\dots,m}} \sum_{h=1}^{l_f} W^S_{fgh} \otimes W^M_{fgh}(\Phi) \middle| \sum_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \sum_{k=1}^{l_i} W^S_{ijk} \otimes W^M_{ijk}(\Phi) \right\rangle \\ &= \sum_{\substack{f,i=1,\dots,n \\ g,j=1,\dots,m}} \sum_{h=1}^{l_f} \sum_{k=1}^{l_i} \langle \langle \hat{\varphi}_g \otimes \psi_{0f} | (\Phi) \rangle (\tilde{\varphi}_{fgh} \otimes \psi_{fgh}) | \langle \hat{\varphi}_j \otimes \psi_{0i} | (\Phi) \rangle (\tilde{\varphi}_{ijk} \otimes \psi_{ijk}) \rangle \\ &= \sum_{\substack{f,i=1,\dots,n \\ g,j=1,\dots,m}} \overline{\langle \hat{\varphi}_g \otimes \psi_{0f} | (\Phi) \rangle} \langle \hat{\varphi}_j \otimes \psi_{0i} | (\Phi) \rangle \left\langle \sum_h \tilde{\varphi}_{fgh} \otimes \psi_{fgh} \middle| \sum_k \tilde{\varphi}_{ijk} \otimes \psi_{ijk} \right\rangle \\ &= \sum_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \overline{\langle \hat{\varphi}_j \otimes \psi_{0i} | (\Phi) \rangle} \langle \hat{\varphi}_j \otimes \psi_{0i} | (\Phi) \rangle \\ &= \|\Phi\|^2. \end{aligned}$$

Therefore U is an isometry.

Furthermore U is one-to-one and onto: given any $\Psi \in \mathcal{H}_S \otimes [\psi_{ijk}]_{ijk}$ of the form

$$\Psi = \sum_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \sum_{k=1}^{l_i} \alpha_{ijk} \tilde{\varphi}_{ijk} \otimes \psi_{ijk}$$

there is a unique $\Phi \in \mathcal{H}_S \otimes [\psi_{0i}]_i$ such that $U(\Phi) = \Psi$, namely

$$\Phi = \sum_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \alpha_{ij} \hat{\varphi}_j \otimes \psi_{0i}.$$

That U satisfies objectification is straightforward by its definition.

Now suppose the partial isometries $W_{ijk}^{\mathbf{M}}$ are the same for all j 's and k 's. Then

$$U = \sum_{\substack{i=1, \dots, n \\ j=1, \dots, m}} \left(\sum_{k=1}^{l_i} W_{ijk}^{\mathbf{S}} \right) \otimes W_{\mathbf{M}_i} = \sum_{i=1}^n \left(\sum_{j=1}^m \sum_{k=1}^{l_i} W_{ijk}^{\mathbf{S}} \right) \otimes W_{\mathbf{M}_i}.$$

It is necessary to show that $\sum_{j=1}^m \sum_{k=1}^{l_i} W_{ijk}^{\mathbf{S}}$ is a unitary operator on $\mathcal{H}_{\mathbf{S}}$. The condition that

$$\left\langle \left(\sum_{k=1}^{l_i} W_{ijk}^{\mathbf{S}} \right) \otimes W_{\mathbf{M}_i}(\Phi) \middle| \left(\sum_{k=1}^{l_i} W_{ij'k}^{\mathbf{S}} \right) \otimes W_{\mathbf{M}_i}(\Phi) \right\rangle = 0$$

for $\Phi \in \mathcal{H}_{\mathbf{S}} \otimes [\psi_{0i}]_i$ and for each $j \neq j'$ clearly implies that

$$\left\langle \sum_{k=1}^{l_i} W_{ijk}^{\mathbf{S}}(\varphi) \middle| \sum_{k=1}^{l_i} W_{ij'k}^{\mathbf{S}}(\varphi) \right\rangle = 0$$

for each $\varphi \in \mathcal{H}_{\mathbf{S}}$, from which it is easy to show that $\sum_{j=1}^m \sum_{k=1}^{l_i} W_{ijk}^{\mathbf{S}} = U_i^{\mathbf{S}}$ is a unitary operator on $\mathcal{H}_{\mathbf{S}}$.

A similar argument applies for the case in which all $U_i^{\mathbf{S}}$ are equal to a unitary operator $U^{\mathbf{S}}$ and establishes the final claim. \square

PROOF OF THEOREM 3.2. The probability reproducibility condition (2.3) is easily established for W^1 by noting that

$$\begin{aligned} & \langle W^1(\varphi \otimes \psi_0^1) | (I_{\mathbf{S}} \otimes P_{[\hat{\psi}_i]}) W^1(\varphi \otimes \psi_0^1) \rangle \\ &= \left\langle \sum_h \langle \varphi_h | \varphi \rangle (\bar{\varphi}_h^1 \otimes \hat{\psi}_h) \middle| (I_{\mathbf{S}} \otimes P_{[\hat{\psi}_i]}) \left(\sum_k \langle \varphi_k | \varphi \rangle (\bar{\varphi}_k^1 \otimes \hat{\psi}_k) \right) \right\rangle \\ &= \langle \varphi | P_{[\varphi_i]} \varphi \rangle. \end{aligned}$$

Also trivially $\mathcal{H}_{\mathbf{S}} \otimes [\psi_0^1] := \{\varphi \otimes \alpha \psi_0^1 : \varphi \in \mathcal{H}_{\mathbf{S}}, \alpha \in \mathbb{C}\}$ is a subspace of $\mathcal{H}_{\mathbf{S}} \otimes \mathcal{H}_{\mathbf{M}}$; indeed, $\forall \varphi_1 \otimes (\alpha_1 \psi_0^1), \dots, \varphi_n \otimes (\alpha_n \psi_0^1) \in \mathcal{H}_{\mathbf{S}} \otimes [\psi_0^1]$, one gets $\varphi_1 \otimes (\alpha_1 \psi_0^1) + \dots + \varphi_n \otimes (\alpha_n \psi_0^1) = \alpha_1(\varphi_1 \otimes \psi_0^1) + \dots + \alpha_n(\varphi_n \otimes \psi_0^1) = (\alpha_1 \varphi_1) \otimes \psi_0^1 + \dots + (\alpha_n \varphi_n) \otimes \psi_0^1 = (\alpha_1 \varphi_1 + \dots + \alpha_n \varphi_n) \otimes \psi_0^1 \in \mathcal{H}_{\mathbf{S}} \otimes [\psi_0^1]$; furthermore $\forall \beta \in \mathbb{C}$, $\beta[\varphi_1 \otimes (\alpha_1 \psi_0^1)] = \varphi_1 \otimes [(\beta \alpha_1) \psi_0^1] \in \mathcal{H}_{\mathbf{S}} \otimes [\psi_0^1]$. In similar fashion it can be proved that $\mathcal{H}_{\mathbf{S}} \otimes [\psi_0^1]^\perp := \{\varphi \otimes \alpha \psi_0^{1,\perp} : \varphi \in \mathcal{H}_{\mathbf{S}}, \alpha \in \mathbb{C}, \text{ with } \psi_0^{1,\perp} \text{ any unit vector orthogonal to } \psi_0^1 (\langle \psi_0^1 | \psi_0^{1,\perp} \rangle = 0)\}$, is also a subspace of $\mathcal{H}_{\mathbf{S}} \otimes \mathcal{H}_{\mathbf{M}}$.

Moreover, $\forall \tilde{\varphi}, \hat{\varphi} \in \mathcal{H}_{\mathbf{S}}, \forall (\alpha \psi_0^1) \in [\psi_0^1], \forall (\beta \psi_0^{1,\perp}) \in [\psi_0^1]^\perp$

$$\langle \tilde{\varphi} \otimes (\alpha \psi_0^1) | \hat{\varphi} \otimes (\beta \psi_0^{1,\perp}) \rangle = \bar{\alpha} \beta \langle \tilde{\varphi} | \hat{\varphi} \rangle \langle \psi_0^1 | \psi_0^{1,\perp} \rangle = 0$$

concluding that $(\mathcal{H}_{\mathbf{S}} \otimes [\psi_0^1]) \perp (\mathcal{H}_{\mathbf{S}} \otimes [\psi_0^1]^\perp)$.

Note that $\dim(\mathcal{H}_S \otimes \mathcal{H}_M) = \dim(\mathcal{H}_S) \cdot \dim(\mathcal{H}_M)$, $\dim(\mathcal{H}_S \otimes [\psi_0^1]) = \dim(\mathcal{H}_S)$, and that it is also true that $\dim(\mathcal{H}_S \otimes [\psi_0^1]^\perp) = \dim(\mathcal{H}_S) \cdot [\dim(\mathcal{H}_M) - 1]$, so that $\dim(\mathcal{H}_S \otimes \mathcal{H}_M) = \dim(\mathcal{H}_S \otimes [\psi_0^1]) + \dim(\mathcal{H}_S \otimes [\psi_0^1]^\perp)$, finally concluding that

$$\mathcal{H}_S \otimes \mathcal{H}_M = (\mathcal{H}_S \otimes [\psi_0^1]) \oplus (\mathcal{H}_S \otimes [\psi_0^1]^\perp).$$

Let $\varphi = \alpha_1 \varphi_1 + \dots + \alpha_n \varphi_n \in \mathcal{H}_S$, and $\alpha \psi_0^1 \in [\psi_0^1]$; then

$$\begin{aligned} W^1(\varphi \otimes (\alpha \psi_0^1)) &= W^1(\alpha \alpha_1(\varphi_1 \otimes \psi_0^1) + \dots + \alpha \alpha_n(\varphi_n \otimes \psi_0^1)) \\ (0.1) \quad &= \alpha \alpha_1(\tilde{\varphi}_1^1 \otimes \hat{\psi}_1) + \dots + \alpha \alpha_n(\tilde{\varphi}_n^1 \otimes \hat{\psi}_n) \end{aligned}$$

where the last equality follows from conditions (2.5) and the probability reproducibility condition (2.1). From (0.1) and the above Lemma it follows that

$$\begin{aligned} \|W^1(\varphi \otimes (\alpha \psi_0^1))\|^2 &= |\alpha \alpha_1|^2 \cdot \|\tilde{\varphi}_1^1 \otimes \hat{\psi}_1\|^2 + \dots + |\alpha \alpha_n|^2 \cdot \|\tilde{\varphi}_n^1 \otimes \hat{\psi}_n\|^2 \\ (0.2) \quad &= |\alpha|^2 \cdot (|\alpha_1|^2 + \dots + |\alpha_n|^2) \end{aligned}$$

But, making use of the fact that $\|\psi_0^1\| = 1$ it is also the case that

$$(0.3) \quad \|\varphi \otimes (\alpha \psi_0^1)\|^2 = |\alpha|^2 \cdot \|\varphi\|^2 \cdot \|\psi_0^1\|^2 = |\alpha|^2 \cdot (|\alpha_1|^2 + \dots + |\alpha_n|^2)$$

A comparison of (0.2) and (0.3) yields the conclusion that W^1 is an isometry on the subspace $\mathcal{H}_S \otimes [\psi_0^1]$.

On the other hand, $W^1(\varphi_i \otimes \psi_0^{1,\perp})$ is easily seen to be the null vector $\mathbf{0} \in \mathcal{H}_S \otimes \mathcal{H}_M$, so that $\forall \varphi = \alpha_1 \varphi_1 + \dots + \alpha_n \varphi_n \in \mathcal{H}_S$ and $\forall \beta \psi_0^{1,\perp} \in [\psi_0^1]^\perp$

$$W^1(\varphi \otimes (\beta \psi_0^{1,\perp})) = \beta \alpha_1 W^1(\varphi_1 \otimes \psi_0^{1,\perp}) + \dots + \beta \alpha_n W^1(\varphi_n \otimes \psi_0^{1,\perp}) = 0$$

establishing the partial isometric and null properties.

Let us denote by $\{u_1 = \varphi_1 \otimes \psi_0^1, \dots, u_n = \varphi_n \otimes \psi_0^1\}$ and by $\{\tilde{u}_1 = \tilde{\varphi}_1^1 \otimes \hat{\psi}_1, \dots, \tilde{u}_n = \tilde{\varphi}_n^1 \otimes \hat{\psi}_n\}$ the two (non complete) orthonormal systems of $\mathcal{H}_S \otimes \mathcal{H}_M$ of the previous Lemma. Then, for all $\Psi \in \mathcal{H}_S \otimes \mathcal{H}_M$ it is clear that $W^1(\Psi) = \sum_{i=1}^n \langle u_i | \Psi \rangle \tilde{u}_i$ from which it follows that

$$\begin{aligned} \|W^1(\Psi)\|^2 &= \left\langle \sum_{k=1}^n \langle u_k | \Psi \rangle \tilde{u}_k \left| \sum_{i=1}^n \langle u_i | \Psi \rangle \tilde{u}_i \right. \right\rangle \\ (0.4) \quad &= \sum_{i,k=1}^n \overline{\langle u_k | \Psi \rangle} \langle u_i | \Psi \rangle \langle \tilde{u}_k | \tilde{u}_i \rangle \\ (0.5) \quad &\leq \|\Psi\|^2 \end{aligned}$$

where the last (Bessel) inequality is a consequence of the fact that $\{u_i\}_I$ is not complete as an orthonormal system of vectors in $\mathcal{H}_S \otimes \mathcal{H}_M$. This shows the contractive property and completes the proof of part (i).

The proof that $W^{1,\perp}$ is a conjugate contractive partial isometry (that is, an isometry on $\mathcal{H}_S \otimes [\psi_0^1]^\perp$ and the null operator on $\mathcal{H}_S \otimes [\psi_0^1]$) satisfying (2.5) and (2.1) is similar, having noticed that $W^{1,\perp}$ is defined on the basis of the orthonormal, not complete sets (2.7a) and (2.7b). A similar proof can establish that $W^{1'}$ is a partial isometry.

It remains to show that the operator $W = W^1 + W^{1,\perp} + W^{1'}$ is unitary. This is easily done in two stages. First of all the previous lemma shows that the union of the orthonormal, non-complete systems defining the operators W^1 , $W^{1,\perp}$ and $W^{1'}$ is a complete orthonormal system. By working as in (0.4) above from this completeness condition it follows that $\forall \Phi \in \mathcal{H}_S \otimes \mathcal{H}_M$, $\|W(\Phi)\| = \|\Phi\|$, i.e. W is an isometry. Secondly, for any $\Psi = \sum_{i=1}^{n,m} \alpha_i \tilde{u}_i \in \mathcal{H}_S \otimes \mathcal{H}_M$ there exists $\Phi = \sum_{i=1}^{n,m} \alpha_i u_i \in \mathcal{H}_S \otimes \mathcal{H}_M$ such that $W(\Phi) = \Psi$: let the \tilde{u}_i 's belong to one of the sets (2.6b), (2.7b), (2.7c) and let the u_i belong to one of the sets (2.6a), (2.7a), the claim then follows. Therefore W is onto. From these two results it follows that W is unitary. \square

PROOF OF THEOREM 3.5. To establish the first claim, consider a state

$$T = \sum_{i=1}^n w_i P_{[\varphi'_i]} \quad w_1 + \dots + w_n$$

on $\mathcal{T}(\mathcal{H}_S)$, where $\varphi'_i = \alpha_{1i}\varphi_1 + \dots + \alpha_{ni}\varphi_n$ and $\sum_{j=1}^n |\alpha_{ji}|^2 = 1$. Then

$$\begin{aligned} \text{Tr}_{\mathcal{H}_S}(T P_{[\varphi_j]}) &= \text{Tr}_{\mathcal{H}_S} \left(\sum_{i=1}^n w_i P_{[\varphi'_i]} P_{[\varphi_j]} \right) \\ &= \sum_{i=1}^n w_i |\alpha_{ji}|^2. \end{aligned}$$

Now consider the state $W^1(T \otimes P_{[\psi_0]})$. This is equal to

$$\sum_{k=1}^n \left(\sum_{i=1}^n w_i |\alpha_{ki}|^2 \right) (P_{[\tilde{\varphi}_k]} \otimes P_{[\psi_k]}),$$

so the right hand side of (3.5) is equal to

$$\sum_{k=1}^n \left(\sum_{i=1}^n w_i |\alpha_{ki}|^2 \right) (P_{[\tilde{\varphi}_k]} \otimes P_{[\psi_k]})(I_S \otimes P_{[\psi_j]}).$$

By convexity, repeated application of (3.1) to the products $P_{[\tilde{\varphi}_k]} I$ leaves the calculation

$$\mathrm{Tr}_{\mathcal{H}_{\mathbf{M}}} \left[\left(\sum_{k=1}^n \sum_{i=1}^n w_i |\alpha_{ki}|^2 P_{[\psi_k]} \right) P_{[\psi_j]} \right] = \sum_{i=1}^n w_i |\alpha_{ji}|^2,$$

hence (3.5) holds. Next let

$$T_{\mathbb{S}} = \sum_{i=1}^n w_i P_{[\varphi'_i]}, \quad w_1 + \dots + w_n = 1$$

be a trace one operator on $\mathcal{H}_{\mathbb{S}}$, and consider the density matrix $T_{\mathbb{S}} \otimes P_{[\psi_0]} \in \mathcal{T}_1^+(\mathcal{H})$.

Then $W^1(T_{\mathbb{S}} \otimes P_{[\psi_0]})$ yields the density matrix

$$\sum_{j=1}^n \mathrm{Tr}_{\mathcal{H}_{\mathbb{S}} \otimes \mathcal{H}_{\mathbf{M}}} \left[(P_{[\varphi_j]} \otimes P_{[\psi_0]}) \left(\sum_{i=1}^n w_i P_{[\varphi'_i]} \otimes P_{[\psi_0]} \right) \right] (P_{[\tilde{\varphi}_j]} \otimes P_{[\psi_j]})$$

Calculating the inner products the result is

$$\begin{aligned} & \mathrm{Tr}_{\mathcal{H}_{\mathbb{S}} \otimes \mathcal{H}_{\mathbf{M}}} \left[(P_{[\varphi_j]} \otimes P_{[\psi_0]}) \left(\sum_{i=1}^n w_i P_{[\varphi'_i]} \otimes P_{[\psi_0]} \right) \right] \\ (*) \quad &= \mathrm{Tr}_{\mathcal{H}_{\mathbb{S}}} \left(P_{[\varphi_j]} \sum_{i=1}^n w_i P_{[\varphi'_i]} \right) \mathrm{Tr}_{\mathcal{H}_{\mathbf{M}}} (P_{[\psi_0]} P_{[\psi_0]}) \quad \text{by (3.1)} \\ &= \sum_{i=1}^n w_i |\alpha_{ji}|^2 \quad \text{by Lemma 3.4} \end{aligned}$$

Hence the total result of applying the mapping W^1 is

$$\sum_{j=1}^n \left(\sum_{i=1}^n w_i |\alpha_{ji}|^2 \right) (P_{[\tilde{\varphi}_j]} \otimes P_{[\psi_j]})$$

This is a density matrix in $\mathcal{T}_1^+(\mathcal{H}_{\mathbb{S}} \otimes \mathcal{H}_{\mathbf{M}})$ with the tensored projectors corresponding to orthogonal projectors on $\mathcal{H}_{\mathbb{S}} \otimes \mathcal{H}_{\mathbf{M}}$. Hence the total trace can be calculated by simply adding together the coefficients, giving

$$\sum_{j=1}^n \left(\sum_{i=1}^n w_i |\alpha_{ji}|^2 \right) = \sum_{i=1}^n w_i \left(\sum_{j=1}^n |\alpha_{ji}|^2 \right) = \sum_{i=1}^n w_i = 1$$

hence W^1 is trace preserving on the subset $\mathcal{T}(\mathcal{H}_{\mathbb{S}}) \otimes P_{[\psi_0]}$.

The second part of the claim is established in a similar manner, by noticing that in (*), for all $T \in \mathcal{T}(\mathcal{H}_{\mathbb{S}}) \otimes (P_{[\psi_0]})^\perp$ one has to calculate terms such as $\mathrm{Tr}_{\mathcal{H}_{\mathbf{M}}} [P_{[\psi_0]} P_{[\psi_0^\perp]}]$, clearly equal to 0.

As for the last claim, consider a trace one operator T given by

$$T = \sum_{i,j=1}^n w_{ij} P_{[\varphi'_i]} \otimes P_{[\psi'_j]}, \quad \sum_{i,j=1}^n w_{ij} = 1,$$

where φ'_i is defined as before and $\psi'_j = \beta_j \psi_0 + \beta_j^\perp \psi_0^\perp$, ψ_0^\perp denoting without loss of generality the vectors obtained by projecting ψ'_j onto the subspace of \mathcal{H}_M orthogonal to ψ_0 and $|\beta_j|^2 + |\beta_j^\perp|^2 = 1$.

In order to obtain $W^1(T)$, calculate the coefficients

$$\text{Tr}_{\mathcal{H}_S \otimes \mathcal{H}_M} \left[(P_{[\varphi_k]} \otimes P_{[\psi_0]}) \left(\sum_{i,j=1}^n w_{ij} P_{[\varphi'_i]} \otimes P_{[\psi'_j]} \right) \right].$$

Rewriting this as

$$\text{Tr}_{\mathcal{H}_S \otimes \mathcal{H}_M} \left[(P_{[\varphi_k]} \otimes P_{[\psi_0]}) \left(\sum_{i,j=1}^n w_{ij} P_{[\alpha_{1i}\varphi_1 + \dots + \alpha_{ni}\varphi_n]} \otimes P_{[\beta_j \psi_0 + \beta_j^\perp \psi_0^\perp]} \right) \right]$$

it is easy to show that the coefficients will be equal to

$$\sum_{i,j=1}^n w_{ij} |\alpha_{ki}|^2 |\beta_j|^2.$$

We then have that

$$W^1(T) = \sum_{k=1}^n \left(\sum_{i,j=1}^n w_{ij} |\alpha_{ki}|^2 |\beta_j|^2 \right) (P_{[\varphi_k]} \otimes P_{[\psi_k]}),$$

the trace of which clearly is

$$\sum_{k=1}^n \left(\sum_{i,j=1}^n w_{ij} |\alpha_{ki}|^2 |\beta_j|^2 \right).$$

From this it follows that

$$\begin{aligned} \text{Tr}_{\mathcal{H}_S \otimes \mathcal{H}_M} (W^1(T)) &= \sum_{i,j=1}^n w_{ij} |\beta_j|^2 \left(\sum_{k=1}^n |\alpha_{ki}|^2 \right) \\ &= \sum_{i,j=1}^n w_{ij} |\beta_j|^2 \leq 1 \end{aligned}$$

where clearly the equality holds if and only if $|\beta_1|^2 = \dots = |\beta_n|^2 = 1$, which completes the proof. \square

Bibliography

- [1] D. ALBERT, *Quantum Mechanics and Experience*, Harvard University Press, Cambridge, Mass., 1992.
- [2] H. ARAKI AND M. YANASE, *Measurement of quantum mechanical operators*, Physical Review, 120 (1960), pp. 622–26, reprinted in Wheeler and Zurek [85].
- [3] G. BACCIAGALUPPI, *Topics in the Modal Interpretation of Quantum Mechanics*, PhD thesis, University of Cambridge, Cambridge, UK, 1996.
- [4] J. S. BELL, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, Cambridge, U.K., 1987.
- [5] E. BELTRAMETTI AND G. CASSINELLI, eds., *The Logic of Quantum Mechanics*, Addison-Wesley, Reading, Mass., 1981.
- [6] E. BELTRAMETTI, G. CASSINELLI, AND P. LAHTI, *Unitary measurements of discrete quantities in quantum mechanics*, Journal of Mathematical Physics, 31 (1990), pp. 91–98.
- [7] F. J. BLOORE, *Geometrical description of the convex sets of states for systems with spin- $\frac{1}{2}$ and spin-1*, Journal of Physics, 9 (1976), pp. 2059–2067.
- [8] M. BORN, *On the quantum mechanics of collisions*, Zeitschrift für Physik, 37 (1963), pp. 863–867, reprinted in Wheeler and Zurek [85].
- [9] V. B. BRAGINSKI, F. Y. KHALILI, AND K. S. E. THORNE, *Quantum Measurement*, Cambridge University Press, Cambridge, UK, 1992.
- [10] T. BREUER, A. AMMAN, AND N. LANDSMANN, *Robustness in quantum measurements*, Journal of Mathematical Physics, 34 (1993).
- [11] H. BROWN, *The insolubility proof of the quantum measurement problem*, Foundations of Physics, 16 (1986), pp. 857–870.
- [12] J. BUB, *How to solve the measurement problem of quantum mechanics*, Foundations of Physics, 18 (1988), pp. 701–22.
- [13] P. BUSCH, M. GRABOWSKI, AND P. J. LAHTI, *Operational Quantum Physics*, Springer-Verlag, Berlin, 1995.
- [14] P. BUSCH, P. LAHTI, AND P. MITTELSTAEDT, *The Quantum Theory of Measurement*, Springer-Verlag, Berlin, 1996. 2nd revised edition.
- [15] P. BUSCH AND P. J. LAHTI, *Correlation properties of quantum measurements*, Journal of Mathematical Physics, 37 (1996), pp. 2585–2601.

- [16] P. BUSCH AND P. MITTELSTAEDT, *The problem of objectification in quantum mechanics*, Foundations of Physics, 21 (1991), pp. 889–904.
- [17] P. BUSCH AND A. SHIMONY, *Insolubility of the quantum measurement problem for unsharp observables*, Studies in the History and Philosophy of Modern Physics, 27 (1996), pp. 397–404.
- [18] N. CARTWRIGHT, *How the Laws of Physics Lie*, Oxford University Press, Oxford, 1983.
- [19] G. CATTANEO AND M. DEL SETA, *Description of perfect premeasurement interactions in quantum measurement theory*. Preprint, June 1994.
- [20] G. CATTANEO, T. MARSICO, AND G. NISTICÒ, *Sharp and unsharp descriptions of physical magnitudes in finite-dimensional quantum mechanics*. Preprint, 1996.
- [21] G. CATTANEO, G. NISTICÒ, AND G. BACCIAGALUPPI, *A concrete procedure for obtaining sharp reconstructions of unsharp observables in finite-dimensional quantum mechanics*, Foundations of Physics, 27 (1997), pp. 1323–1343.
- [22] H. CHANG, *Measurement and the Disunity of Quantum Physics*, PhD thesis, Stanford University, Stanford, Ca., 1993.
- [23] ———, *On the applicability of the quantum measurement formalism*, Erkenntnis, 46 (1997), pp. 143–163.
- [24] R. CLIFTON, *Independently motivating the Kochen-Dieks interpretation of quantum mechanics*, The British Journal for the Philosophy of Science, 46 (1995), pp. 33–58.
- [25] E. DAVIES, *Quantum Theory of Open Systems*, Academic Press, London, 1976.
- [26] E. B. DAVIES AND J. T. LEWIS, *An operational approach to quantum probability*, Communications in Mathematical Physics, 17 (1970), pp. 239–260.
- [27] B. D'ESPAGNAT, *Conceptual Foundations of Quantum Mechanics*, Addison-Wesley Publishing Company, New York, 1976.
- [28] B. S. DEWITT AND N. GRAHAM, eds., *The Many-Worlds Interpretation of Quantum Mechanics*, Princeton University Press, Princeton, New Jersey, 1973.
- [29] P. DIRAC, *The Principles of Quantum Mechanics*, Clarendon Press, Oxford, 1968.
- [30] J. EARMAN, *A Primer on Determinism*, Dordrecht-Reidel, Dordrecht, Neth., 1986.
- [31] J. EARMAN AND A. SHIMONY, *A note on measurement*, Nuovo Cimento, 54B (1968), pp. 332–4.
- [32] G. EMCH, *Mathematical and Conceptual Foundations of 20th-Century Physics*, North-Holland, Amsterdam, 1984.
- [33] H. EVERETT III, *"Relative state" formulation of quantum mechanics*, Reviews of Modern Physics, 29 (1957), pp. 454–62, reprinted in Wheeler and Zurek [85].
- [34] ———, *The theory of the universal wave function*, in DeWitt and Graham [28], pp. 3–140.
- [35] A. FINE, *On the general quantum theory of measurement*, Proceedings of the Cambridge Philosophical Society, 65 (1969), pp. 111–122.
- [36] ———, *Insolubility of the quantum measurement problem*, Physical Review, 2 (1970), pp. 2783–2787.
- [37] ———, *Probability and the interpretation of quantum mechanics*, British Journal for the Philosophy of Science, 24 (1973), pp. 1–37.

- [38] ———, *The two problems of quantum measurement*, in *Proceedings of the Fourth International Congress for Logic, Methodology and Philosophy of Science*, P. Suppes, ed., North-Holland, Amsterdam, 1973, pp. 567–581.
- [39] ———, *The Shaky Game: Einstein, Realism and the Quantum Theory*, The University of Chicago Press, Chicago, 1986.
- [40] ———, *With complacency or concern: Solving the quantum measurement problem*, in *Kelvin's Baltimore Lectures and Modern Theoretical Physics: Historical and Philosophical Perspectives*, R. Kargon and P. Achinstein, eds., MIT Press, Cambridge, MA, 1987.
- [41] ———, *Piecemeal realism*, *Philosophical Studies*, 61 (1991), pp. 79–96.
- [42] ———, *Resolving the measurement problem: a reply to Allen Stairs*, *Foundations of Physics Letters*, 5 (1992), pp. 125–138.
- [43] ———, *Measurement and quantum silence*, in *Correspondence, Invariance and Heuristics*, S. French and H. Hamminga, eds., Dordrecht, 1993, Kluwer Academic Publishers.
- [44] P. GALISON, *Image and Logic: A Material Culture of Microphysics*, University of Chicago Press, Chicago, IL, 1997.
- [45] I. GELFAND AND M. NAIMARK, *On the embedding of normed rings into the rings of operators in Hilbert space*, *Matematicheskii Sbornik n.s.*, 12 (1943), pp. 197–213.
- [46] G. C. GHIRARDI, *I fondamenti concettuali e le implicazioni epistemologiche della meccanica quantistica*, in *Filosofia della Fisica*, G. Boniolo, ed., Bruno Mondadori, Milan, Italy, 1997, pp. 337–608.
- [47] I. HACKING, *Representing and Intervening*, Cambridge University Press, Cambridge, UK, 1983.
- [48] R. HEALEY, *The Philosophy of Quantum Mechanics: An Interactive Interpretation*, Cambridge University Press, Cambridge, UK, 1989.
- [49] K. HEPP, *Quantum theory of measurement and macroscopic observables*, *Helvetica Physica Acta*, 45 (1972), pp. 237–48.
- [50] R. HUGHES, *The Structure and Interpretation of Quantum Mechanics*, Harvard University Press, Cambridge, Mass., 1989.
- [51] R. V. KADISON, *Transformations of states in operator theory and dynamics*, *Topology*, 3 (1965), pp. 177–198. Supplement 2.
- [52] S. KOCHEN, *A new interpretation of quantum mechanics*, in Lahti and Mittelstaedt [59], pp. 151–69.
- [53] H. KRIPS, *The Metaphysics of Quantum Mechanics*, Cambridge University Press, Cambridge, UK, 1989.
- [54] F. KRONZ, *Quantum entanglement and nonideal measurements: a critique of Margenau's objections to the projection postulate*, *Synthese*, 89 (1991), pp. 229–51.
- [55] T. KUHN, *The Essential Tension: Selected Studies in Scientific Tradition and Change*, University of Chicago Press, Chicago, IL, 1977.
- [56] ———, *A function for thought experiments*, in *The Essential Tension: Selected Studies in Scientific Tradition and Change* [55], pp. 240–265.

- [57] ———, *The function of measurement in modern physical science*, in *The Essential Tension: Selected Studies in Scientific Tradition and Change* [55], pp. 174–224.
- [58] ———, *Mathematical versus experimental traditions in the development of physical science*, in *The Essential Tension: Selected Studies in Scientific Tradition and Change* [55], pp. 31–65.
- [59] P. J. LAHTI AND P. MITTELSTAEDT, eds., *Symposium on the Foundations of Modern Physics*, World Scientific, Singapore, 1987.
- [60] I. LAKATOS, *The methodology of scientific research programmes*, in *Criticism and the Growth of Knowledge*, I. Lakatos and A. Musgrave, eds., Cambridge University Press, Cambridge, UK, 1974.
- [61] N. LANDSMANN, *Observation and superselection in quantum mechanics*, *Studies in the History and Philosophy of Modern Physics*, 26B (1995), pp. 45–73.
- [62] F. LONDON AND E. BAUER, *The theory of observation in quantum mechanics*, *Actualités scientifiques et industrielles: Exposés de physique générale*, 775 (1939), reprinted in Wheeler and Zurek [85].
- [63] G. LUDWIG, *Foundations of Quantum Mechanics*, Springer-Verlag, Berlin, 1983.
- [64] H. MARGENAU, *Critical points in modern physical theory*, *Philosophy of Science*, 4 (1937), pp. 337–370.
- [65] ———, *Philosophical problems concerning the meaning of measurement in physics*, *Philosophy of Science*, 25 (1958), pp. 23–34.
- [66] ———, *The Nature of Physical Reality*, Ox Bow Press, Woodbridge, Conn., 1977.
- [67] R. A. MILLIKAN, *A direct photoelectric determination of planck's "h"*, *Physical Review*, 7 (1916), pp. 355–388.
- [68] E. NAGEL, *The Structure of Science: Problems in the Logic of Scientific Explanation*, Hackett, Indianapolis, IN, 1960.
- [69] J. L. PARK, *Quantum theoretical concepts of measurement, parts 1 and 2*, *Philosophy of Science*, 19 (1968), pp. 205–231.
- [70] O. PESSOA JR., *Simple realisations of generalised measurements in quantum mechanics*, *Foundations of Physics Letters*, 7 (1994), pp. 447–457.
- [71] A. RIECKERS, *Fundamentals of algebraic quantum theory*, in *Groups, Systems and Many-Body Physics*, P. Kramer and M. Dal Cin, eds., Vieweg, Braunschweig, Wiesbaden, 1980.
- [72] F. RIESZ AND B. SZ.-NAGY, *Functional Analysis*, Frederick Ungar Publishing, New Your, 1955.
- [73] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
- [74] L. RUETSCHKE, *Measurement error and the Albert-Loewer problem*, *Foundations of Physics Letters*, 8 (1995), pp. 327–44.
- [75] E. SCHEIBE, *The Logical Analysis of Quantum Mechanics*, Pergamon Press, Oxford, 1973.
- [76] I. E. SEGAL, *Postulates for general quantum mechanics*, *Annals of Mathematics*, 48 (1947), pp. 930–948.
- [77] A. SHIMONY, *Approximate measurement in quantum mechanics, part ii*, *Physical Review D*, 9 (1974), pp. 2321–23, reprinted in *Search for a Naturalistic World View* [78].
- [78] ———, *Search for a Naturalistic World View*, Cambridge University Press, Cambridge, 1993.

- [79] H. STEIN, *Maximal extension of an impossibility theorem concerning quantum measurement*, in Potentiality, Entanglement and Passion at a Distance, R. Cohen and J. Stachel, eds., Kluwer Academic Publishers, Dordrecht, 1997.
- [80] B. VAN FRAASSEN, *Quantum Mechanics: An Empiricist's View*, Oxford University Press, Oxford, 1991.
- [81] P. VERMAAS AND D. DIEKS, *The modal interpretation of quantum mechanics and its generalisation to density operators*, Foundations of Physics, 25 (1995), pp. 145–58.
- [82] J. VON NEUMANN, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton, 1955.
- [83] K. WAN AND F. HARRISON, *Superconducting rings, superselection rules and quantum measurement problems*, Physics Letters A, 174 (1993), pp. 1–8.
- [84] K. K. WAN AND R. H. FOUNTAIN, *Quantization by parts, selfadjoint extensions and a novel derivation of the josephson equation in superconductivity*, Foundations of Physics, (1997).
- [85] J. WHEELER AND W. ZUREK, eds., *Quantum Theory and Measurement*, Princeton University Press, Princeton, 1983.
- [86] E. P. WIGNER, *Remarks on the mind-body question*, in I. J. Good, The Scientist Speculates, Heinemann, London (1961), pp. 284–302, reprinted in Wheeler and Zurek [85].
- [87] ———, *The problem of measurement*, American Journal of Physics, 31 (1963), pp. 6–15, reprinted in Wheeler and Zurek [85].