

Classical Observables, Measurement, and Quantum Mechanics

Thomas Breuer
Peterhouse

Submitted to the University of Cambridge
in fulfilment of the requirements for the Doctor of
Philosophy
in the Department of History and Philosophy of Science

September 1994

Ágimagimnak

Contents

Introduction	7
1 Self-Reference and the Universal Validity of Quantum Mechanics	13
1.1 Previous arguments against deterministic universally valid theories	15
1.2 The impossibility of accurate self- measurements	21
1.3 Universal Validity Revisited	35
1.4 Application to quantum measurements	39
1.4.1 EPR-correlations and the universal validity of quantum mechanics	39
1.4.2 Subjective Decoherence	43
1.4.3 Wigner's friend	51
2 Phase Space Quantisation	55
2.1 Classical Mechanics	58
2.1.1 Phase space and configuration space as homogeneous spaces	58
2.1.2 Lattice theoretic and algebraic formulation of classical mechanics	63
2.2 Quantisation on Configuration Space	72

2.2.1	Systems of Imprimitivity	78
2.2.2	The commutation relations	80
2.2.3	Classification of configuration space quantisations.	82
2.2.4	Algebraic description of systems of imprimitivity	85
2.3	Phase space quantisation of the Weyl system	87
2.3.1	Phase space quantisation by systems of covariance	88
2.3.2	Phase space representation of the Weyl system	91
2.3.3	Systems of covariance in geometric quantisation.	105
2.4	Generalised systems of covariance	117
2.4.1	Phase space quantisation by generalised systems of covariance	127
2.4.2	Existence of transitive generalised systems of co- variance	129
2.5	An extension and an imprimitivity theorem	134
2.5.1	An extension theorem	136
2.5.2	An imprimitivity theorem	140
3	The Classical Limit	145
3.1	Classical observables in quantum mechanics	147
3.1.1	Hilbert space formulation	147
3.1.2	Algebraic formulation	151
3.1.3	The infinite tensor product	158
3.2	Simultaneous Sharp Measurability of Position and Mo- mentum in Infinite Quantum Systems	163
3.2.1	Joint position-momentum observables	165
3.2.2	Elementary Systems	167
3.2.3	Finite Systems	170
3.2.4	Infinite systems	176

4	The Classical Pointer in Quantum Measurements	183
4.1	Previous arguments for and against classical pointer observables	184
4.1.1	The projection postulate	185
4.1.2	The ignorance interpretation	190
4.2	Inaccurate measurements of non-normal states?	195
4.2.1	Non-normal states in traditional quantum mechanics	197
4.2.2	Measurability of non-normal states in algebraic quantum mechanics	203
4.3	Inaccurate experiments can only be described by a classical pointer	208
4.3.1	The requirement of finite measurement accuracy	210
4.3.2	W^* -version of the proof	214
4.3.3	C^* -version of the Proof	217
4.3.4	The continuous classical pointer	221
4.4	The time-evolution of the classical pointer	225
4.4.1	Infinite time limit for automorphic measurements with classical pointer	225
4.4.2	Non-automorphic time evolution	229
	References	237
	Acknowledgements	254

Introduction

Are universally valid theories possible?

A universally valid theory of material reality seems to be the ultimate goal of scientific enquiry. This was for example expressed in the dream of Laplace's demon: if a universally valid theory were deterministic, the demon could use it to calculate any future state of the universe from its present state.

Chapter 1 addresses an aspect of measurement which is common to all universally valid theories—the problem of self-reference. It will turn out that it is impossible for an observer to distinguish all states of a system in which she or he is contained. With this kind of non-self-measurability it is impossible for us to know the present state of the universe exactly. Therefore, even if the ultimate universally valid theory were deterministic, we could not use it to calculate the exact future state of the universe.

The fact that it is impossible for us to distinguish all states of the universe seriously impairs the empirical relevance of universally valid theories. But a physical realist would rather not dismiss a theory just for this reason. In his opinion, there are entities which in some sense are independent of human knowledge. Statements about these entities should not be conflated with statements about knowledge of the entities. From this point of view, the fact that no experiment can, not

even in principle, distinguish all states is not in itself objectionable. Accordingly, a physical realist would not take this argument as sufficient reason to exclude the possibility of universally valid theories. We have to accept, however, that a universally valid theory is of ontological relevance rather than empirical.

Is quantum theory universally valid?

Experimental evidence does not suggest that there is a clear border beyond which quantum theory does not apply. Also, there is nothing in the formalism which makes it inapplicable to 10^{23} particles. On the other hand, we perceive the macroscopic world around us as being correctly described by classical physics. If quantum theory is universally valid, it should also be able to account for this fact.

Is quantum theory consistent with our experience of the macroscopic world? Can we regard macroscopic bodies as consisting of a large number of quantum systems?

There are famous historical arguments why this should be impossible. Such an argument is for example Schrödinger's cat paradox. These arguments intend to show that

- (1) at least some properties of macroscopic bodies (such as a cat being alive or dead) are definite in all states; (I will call such properties classical; and properties for which there are states in which they are not definite, quantum)
- (2) quantum theory does not allow for classical properties.

Assuming that quantum theory is universally valid, (1) and (2) contradict each other. If one wants to avoid this contradiction, and at

the same time stick to the universal validity of quantum theory, then clearly one has to drop (1) or (2).

(1) should not be taken as a stronger claim than it actually is. It is not claimed that *all* properties of macroscopic bodies are definite in all states. The claim is just that macroscopic systems have *some* classical properties. It is consistent with (1) that macroscopic bodies display some quantum properties. Superfluidity and superconductivity probably are such macroscopic quantum effects.

From a strictly empirical point of view, even this weak claim is vulnerable. For it is probably consistent with our experience of the macroscopic world that states in which macroscopic properties are not definite are rare. If one can, by this or a similar argument, deny (1), classical mechanics does not have any place in a correct physical theory. It is simply a wrong description of the real quantum world, but happens to give on the macroscopic level a good numerical approximation.

(1) can be denied in another way, albeit at the cost of modifying the quantum mechanical time evolution. States in which macroscopic properties are not definite need not be rare; but perhaps we do not perceive such states because our observation changes the state of the system into one where the observed property is definite. This observation-induced state change is usually called “reduction of the wave packet”.

There is another way in which quantum mechanics can be consistent with macroscopic perception: one can deny (2). The argument for this is that macroscopic quantum systems can have classical properties. This approach will be taken in Chapter 3. There I discuss the emergence of classical properties in quantum systems with infinitely many degrees of freedom.

One could, of course, take the point of view that such an emergence

of classical properties violates the universal validity of quantum theory. I would like to take a different point of view. The emergence of classical properties in infinite systems is natural in the sense that at no point one assumes quantum theory not to be valid anymore. The derivation of classical properties, however, relies on the assumption that the *quasilo-cal* observables form the algebra of observables of the infinite quantum system. This will be discussed in more detail in section 3.1.3.

(Chapter 2 addresses the reverse problem: the transition from classical theories to quantum theories by various quantisation procedures. If quantum theory is universally valid, quantisation is not of immediate ontological relevance. Rather it highlights similarities and differences in the mathematical structure of classical and quantum theories.)

The contradiction arising from (1), (2), and the universal validity of quantum theory is part of the measurement problem. In which sense the denial of (2) solves the measurement problem will be discussed in Chapter 4. The approach to the measurement problem presented there violates (2) by using a classical pointer observable belonging to an infinite apparatus or an infinite environment.

There is a third alternative: the contradiction arising from (1), (2), and the universal validity of quantum theory can be avoided if we renounce the universal validity of quantum theory. This is done for example by insisting—in the spirit, perhaps, of Bohr—that the realm of our immediate experience is by described classical physics and not by quantum physics.

One final remark about style. In order to present the problems in a form to which definite answers can be given, I have been at pains to give rigorous mathematical formulations. At the same time I have

attempted to keep as clear a distinction as possible between physical motivation and purely mathematical considerations. This is an exercise, not just in the mathematical foundations of quantum mechanics, but more importantly in the philosophy of physics, a discipline that is not a compromise between philosophy and physics.

Chapter 1

Self-Reference and the Universal Validity of Quantum Mechanics

Chapter Abstract

In the central argument of this chapter, it is shown that it is impossible for an observer to distinguish all states of a system in which she or he is contained, irrespective of whether this system is a classical or a quantum mechanical one. This implies that absolutely universally valid theories are not fully justifiable from an operational point of view. Still, such theories can have ontological relevance. For an operationally fully justifiable theory claims of universal validity can only be maintained in an observer dependent form: for each part of the world some observer can apply the theory to it. In an application to quantum measurements, the central result implies subjective decoherence: For an observer it is not possible to determine whether the joint system consisting of himself and the observed system is in a statistical state with or without interference terms.

Chapter Overview

In this chapter I shall analyse the consequences of postulating universal validity for a physical theory. As far as quantum mechanics is concerned, von Neumann assumed the theory to be universally valid and thus was led to the measurement problem. Bohr denied the universal validity of quantum mechanics for “purely logical reasons”, and thereby avoided confrontation with the measurement problem. It has often¹ been suggested that self-reference problems for universally valid theories may pose serious difficulties for a quantum mechanical description of the measurement apparatus. The aim of this chapter is to investigate these suggestions.

I will say that a theory is universally valid in the absolute sense if it is true of the whole world, without any reference to observers. In section 1.1 some arguments claiming that absolutely universally valid theories cannot be deterministic will be reviewed and criticised.

In section 1.2 I give the central argument why no apparatus can distinguish all states of a system in which it is properly contained. This section follows Breuer (1994a). Properties of self-reference play a crucial rôle in the argument. Whether the system is a quantum mechanical one or a classical one, and whether the time evolution is deterministic or stochastic, is irrelevant.

In section 1.3 the central result is applied to the question of how we have to conceive of universally valid theories. It leads to the conclusion that absolutely universally valid theories at least partially lack operational justification. Still they might be ontologically meaningful.

¹See for example Dalla Chiara (1977), Peres and Zurek (1982), Roessler (1987), Finkelstein (1988), Penrose (1989), Mittelstaedt (1993), Primas (1990), Svozil (1993).

An operationally fully justified theory can be universally valid only in a relative sense: for any subsystem of the world some observer can apply the theory to it. Such a position is sometimes adopted with respect to quantum mechanics, where the final observer in a measurement chain is either not described at all by quantum mechanics (this is Bohr's point of view) or, because of the projection postulate she or he does not follow the usual unitary quantum mechanical time evolution. But the argument presented here shows additionally two things: Firstly, in quantum mechanics these conclusions do not rely on the deterministic character of the Schrödinger time evolution as opposed to the stochastic character of the projection postulate. Secondly, we have the same conclusions in classical theories.

In section 1.4 I turn to self-measurements in quantum mechanics. It is shown that no apparatus can measure the Einstein-Podolsky-Rosen-correlations between itself and an outside system. This then leads to the concept of subjective decoherence: For an apparatus it is not possible to determine whether the joint system consisting of itself and the observed system is in a statistical state with or without interference terms. Subjective decoherence sheds new light on the problem of Wigner's friend, and on the question whether the observer's mind can be described by quantum mechanics.

1.1 Previous arguments against deterministic universally valid theories

Absolute universal validity. As I take it, the thesis of *absolute universal validity* of a physical theory says that such a theory is true of the whole "world", or of the whole "universe", without any reference to

observers. Such a theory is universally valid in the sense that no part of the “universe” is excluded from its domain of validity. Still, I would not like to call it a theory of everything because it need not describe phenomena at all levels of complexity, from high energy physics to sociology. I call it universally valid in an *absolute* way because it does not make any reference to observers.

Non-self-predictability and non-self-measurability. An absolutely universally valid theory of material reality seems at first sight to be the ultimate goal of scientific inquiry. This was expressed, for example, in the dream (or nightmare) of Laplace’s demon. If a universally valid theory were deterministic, then the demon could use it to calculate any future state of the universe from its present state.

Popper (1950) argued that, however complete the information provided to the demon about its own past or present state, there will always be questions about its own future state which the demon cannot answer. This is the thesis of non-self-predictability. The demon can make accurate predictions only about the outside world. Therefore, if one wants to maintain—in Laplace’s spirit—that in presence of a deterministic time evolution the demon should be able to predict the future of the whole universe, then one has to exclude the demon from the universe. (This is perhaps why Laplace’s demon is a demon. If it can make predictions about the whole universe, but still not himself, then it must be a truly supernatural being outside our material world. It is not just demonic because of its great calculational abilities.) Non-self-predictability implies that, even in a deterministic theory, accurate predictions about subsystems of the universe may at most be feasible for *some* observer, namely for one outside the system whose behaviour

is to be predicted.² But there will be no observer able to predict *everything* accurately. Every deterministic theory must admit the existence of unpredictable events when a predictor applies it to himself. Even in classical mechanics with a deterministic time evolution we have this kind of unpredictability.

Rothstein (1952, 1964) showed that if, in the spirit of Szilard's (1929) solution of the problem of Maxwell's demon³, the demon is considered as part of the thermodynamic system, the second law of thermodynamics imposes restrictions on the accuracy of measurements which the demon can perform on the system.⁴ In classical thermodynamics exact state self-measurements are impossible. In section 1.2 I will arrive at the same conclusion in a more general framework.

²I think it is doubtful whether an outside predictor, even if he obtains precise information about the state of the system whose behaviour he has to predict, could be able to make accurate predictions. After all, it may be that there are some interactions between the predictor and the system. In this case complete knowledge of the system's environment is necessary to exploit for predictions the deterministic time evolution. But, as Popper's (1950, section III) argument about the "Oedipus effect" intends to show, it is impossible for the predictor to obtain complete information about his present state and therefore about the system's environment (in which he is contained). Therefore the only way to ensure that the predictor can make accurate predictions about the system is to exclude interactions. This is difficult if the predictor is subject to the laws of physics.

³Similarly to Laplace's demon, Maxwell's demon must not be subject to the laws of physical reality if his presence is not to destroy the phenomenon he should illustrate. This seems to be a general property of demons.

⁴Rothstein (1964) draws from his result the following conclusion. The impossibility of exact state self-measurability implies that we never can fully verify the allegedly deterministic character of an absolutely universally valid theory. For this reason he dismisses the problem of free will and determinism as "pseudo-problem".

Although I agree with the result, I think that the conclusion drawn by Rothstein is controversial. Not being able to assess whether or not the time evolution of the world is deterministic, the problem whether free will and determinism are compatible surely loses some of its physical relevance. But the fact that the assumption of a deterministic time evolution of the universe cannot be checked does *not* mean that it is not decidable whether free will and determinism are compatible. That might be decided by conceptual analysis alone.

I believe the conclusions of non-self-predictability and non-self-measurability are correct. In this chapter, they will follow in a more general framework from the fact that no observer can obtain or store information sufficient to distinguish all states of a system in which he or she is contained.⁵ But of course non-predictability, or non-self-predictability, does not disprove determinism.⁶

⁵Similarly, Popper (1950, section III, pp. 189-191) argued that impossibility of having complete up-to-date self-information is the most conclusive argument for non-self-predictability. He gives two arguments why it is impossible to have complete up-to-date self-information. The first one is what he calls the “Oedipus effect”: the reception of information changes the state of the predictor. (This is a truism if the information is encoded in the physical state of the predictor. If the information is encoded in the mental state of the predictor one needs additionally an assumption of psycho-physical supervenience: different mental states correspond to different physical states (but not necessarily vice versa).) Popper does not yet consider this argument as conclusive because a very superior outside being might foresee the state change caused by the reception of information, and give the predictor not accurate information about its state but some information designed ingeniously in such a way as to make the predictor’s predictions about himself accurate. Therefore Popper gives a second argument: no finite piece of information can be exact self-information. The reason for this is that the self-information must contain as one of its parts a description of itself, which is impossible because there is no bijective map from a finite set to a proper subset. Popper thinks that an infinite piece of information could be complete up-to-date self-information, but such an infinite piece cannot be used for prediction if the predictor works at finite speed.

My central argument will also be to the effect that it is impossible to have complete up-to-date self-information. But, unlike Popper, I will not restrict myself to finite pieces of informations. Neither do I need to invoke assumptions about determinism, or finite processing speed of the predictor.

⁶Remember Earman’s (1986) warning: Predictability is not a full-fledged substitute for determinism. The latter term is about how things really are, the former about our relation to them.

If we were to follow Popper (1950) and take determinism as predictability with a certain degree of precision (depending on the precision of knowledge about the initial state), then determinism could fail for classical dynamical systems. In classical dynamical systems, the time evolution is described by a one-parameter family of automorphisms of the state space. Thus the initial state uniquely determines the state at any later time. In this sense classical dynamical systems are deterministic. Still, such systems may have unstable manifolds along which the distance of orbits grows exponentially. In this case, however small the inaccuracy of knowledge δ

Relative universal validity. Peres and Zurek (1982) present an argument against absolute universal validity which is particularly simple and seemingly convincing. They argue that no physical theory can at the same time fulfil the three requirements of absolute universal validity, experimental verifiability and determinism.

Their argument runs as follows: For the experimental verifiability of a physical theory they regard it as a necessary condition that the experimenter can freely choose which experiment he is going to make. A universally valid theory which is deterministic precludes the free choice of the observer and thus experimental verifiability.

To illustrate this they look at interpretations of quantum mechanics which drop one or the other of the three requirements. Firstly, one can have universal validity and determinism, but drop experimental verifiability. Everett's (1957) relative state interpretation is a theory of this kind in that the universe is completely described by quantum mechanics and follows a deterministic unitary time evolution. (Determinism

about the initial state x , and however large the error ϵ we accept for our predictions, there will be a state x' closer to x than δ evolving into a state $x'(t)$ farther than ϵ from $x(t)$ for sufficiently large t . Thus for classical dynamical systems with unstable manifolds predictability with a certain degree of precision fails, but the present state still determines all future states.

So predictability with a certain degree of precision is not a good criterion for deterministic behaviour of a system. But is predictability from hypothetically precise initial data a possible way to formulate our intuition about "ontological" determinism? By hypothetically precise initial data I mean initial data which could be gathered by an ideal observer without technological limitations on the accuracy of measurements he can perform.

One of the lessons of this chapter will be that hypothetically precise initial data gathered by an observer about a system in which he is contained cannot specify the state uniquely, because there are limitations of principle (rather than technological ones) on the accuracy of self-measurements. Therefore, if the time evolution of our universe is deterministic, but the phase space has unstable manifolds, even an ideal internal observers cannot predict the future state with arbitrarily small inaccuracy. Thus predictability, even for ideal observers as long as they are *in* the universe, is not a good criterion for determinism.

does not hold for the single branches of the universe.)

Secondly, one can consider theories which are universally valid and grant the observer free choice. Such theories cannot be deterministic. Quantum mechanics, with the observer considered as a quantum mechanical system, and using the projection postulate to describe the measurement process, would be such a theory. This theory applies to the apparatus as well, but the stochastic behaviour described by the projection postulate defies determinism. (Perhaps von Neumann's (1932) quantum theory can be regarded as such a theory. This depends on whether one wants to consider as universally valid a theory which can perhaps describe the whole material reality, but not the observer's mind.)

Peres and Zurek conclude that if quantum mechanics is universally valid⁷ at all, then it is so only in the *relative* sense that every observer can perhaps apply it to any selected part of the world, except himself. It supposedly applies to Schrödinger's cat, Wigner's friend and Wigner himself under the condition that they lose their status of observer and are observed by something or somebody else. (It is not the point here whether quantum mechanics can really be applied to *every* phenomenon of the external world. Quantum theory just serves as an example of a theory which might be universally valid.)

I agree with the conclusion of Peres and Zurek, but in my view their argument must be challenged. Firstly, I think that they are unnecessarily restrictive by assuming that a universally valid theory describes phenomena at all levels of complexity, including mental phenomena like free will. This concept of absolute universal validity is stronger than

⁷Peres and Zurek (1982) actually do not use the terms absolute and relative universal validity. Rather they refer to relatively universally valid theories as 'universal', and to absolutely universally valid theories as 'closed'.

the one I use, but of course Peres and Zurek are free to use such a concept. Secondly, and this is more important, they take it as granted, that an observer who is described by a deterministic theory does not have the freedom to choose his experimental set-up. Determinism and free will are assumed as mutually exclusive. This may be so but it is controversial.⁸ Peres and Zurek need the assumption of determinism only to preclude the observer in a universally valid theory from having free will. Thereby they exclude experimental verifiability of a universally valid deterministic theory. If determinism does not necessarily exclude free will of the observer, then the argument fails.

But, as indicated, I think the conclusion that an absolutely universally valid theory (even in the weaker sense I use this concept) is not fully justifiable from an operational point of view, can be obtained with a different argument, one where determinism has no rôle to play.

1.2 The impossibility of accurate self-measurements

In this section I am going to present an argument why it is impossible for an observer to distinguish all states of a system in which he or she is contained. The argument exploits self-reference properties, but it does not make any assumptions about the character of the time evolution. It is valid for quantum mechanical as well as for classical theories.

Self-reference in physical theories. It has often⁹ been suggested that

⁸Arguments for the compatibility of free will and determinism can be found for example in Hume (1758) and Schlick (1939, chapter VII).

⁹See Popper (1950), Rothstein (1964), Dalla Chiara (1977), Mittelstaedt (1993),

there is a possible connection between self-reference properties of formal systems and restrictions on measurability in universally valid theories. Since my argument will exploit self-reference properties, let me first make some remarks about similarities and differences between Gödel's (1931) theorem and my argument.

Propositions about physical systems can be reformulated by saying "The state of the system has this and this property". So instead of speaking about propositions we can equally well speak about sets of states: to each proposition there corresponds the set of states for which the proposition is true. The way to test propositions about physical systems is to make experiments on the system. Experiments give information about the state of the system, and it can then be checked whether or not this information is compatible with the proposition under consideration. So experiments serve to at least partially constitute the semantics of physical theories. In this sense observation is a semantical concept.

Call a language semantically closed in the sense of Tarski (1956, 1969) if it contains (1) semantic concepts and (2) expressions referring to its own propositions. The language of a physical theory describing experiments can be closed semantically: If apparatus and object system, as well as their interaction, can be described by the theory, then the semantic concept of observation can be introduced into the language of the theory. Also, replacing propositions by states provides the language of the theory with expressions for its propositions. Additionally, there are propositions referring to other propositions: the apparatus states after the experiment are not only states in their own

Peres and Zurek (1982), Roessler (1987), Finkelstein (1988), Primas (1990a), and Svozil (1993).

right, they also refer to states of the observed system. Thus one arrives at a theory whose language is semantically closed.

The language of the formal system used by Gödel is not semantically closed: its language does not contain any expressions referring explicitly to metatheoretical concepts. But after assigning numbers to the propositions, these numbers can be *interpreted* as expressions of the language referring to its own propositions. Also, one observes that some propositions in the formal system hold if and only if the system has certain properties. By *interpreting* these propositions as propositions *about* the system, one transfers metatheoretical concepts down to the level of the theory. For example provability, restricted to meaning the existence of a formal proof, is such a metaconcept transferred down to the level of the formal system. These two interpretational steps make it possible to intuitively regard the system as semantically closed, although strictly speaking it is not.

In a semantically closed language it is possible to formulate self-referential propositions. The self-reference may be paradoxical or consistent. In the formal system used by Gödel a self-referential proposition with the following intuitive meaning is formulated: “This proposition is not provable.” The formal expression of this is the Gödel formula. It is neither refutable nor provable within the formal system. (This is Gödel’s theorem.) But the Gödel formula is true by the standards of informal number theory.

Similarly, in the language of a physical theory describing observations, there will be paradoxically self-referential propositions, or rather paradoxically self-referential states. Since the reference is from apparatus states to states of the observed system, self-reference can and will occur if the apparatus is contained in the observed system. The

result that it is impossible for an apparatus in the observed system to discriminate all states of the observed system somewhat resembles Gödel's theorem. Also the fact that an apparatus outside the observed systems in general can distinguish all states seems to be analogous to the fact that the Gödel formula is true in informal number theory.

In spite of these similarities, there are important differences between Gödel's proof and my result. In fact, the most important parts of Gödel's proof do not have a parallel in my argument. Firstly, the formal system used by Gödel is not semantically closed in the strict sense. In my argument there is nothing similar to Gödel's ingenious idea to do without semantic closure by introducing the Gödel numbers. Secondly, Gödel proved his result without assuming that provable statements are true.¹⁰ It is probably no exaggeration to say that my argument does not have much more in common with Gödel's proof than the use of self-reference.

Description of measurements. Let us assume that we have a physical theory whose formalism specifies for the systems it describes sets of possible states. These states may refer to individual systems or to statistical ensembles. In an individual formulation of classical mechanics, for example, the states would correspond to the points on phase space, whereas in a statistical formulation they would be probability distributions, i.e. normalised real-valued L^1 -functions on phase space. In quantum mechanics the individual states would be pure, i.e. extremal,

¹⁰Guido Bacciagaluppi pointed out to me that this considerably strengthens his result. It does not rely on the controversial concept of 'truth in informal number theory'. Assuming provable statements to be true in informal number theory would have discredited Gödel's result in the eyes of the intuitionists. But both formalists and intuitionists accepted the informal concept of truth in finite (or constructive) number theory. This and ω -consistency was all that Gödel needed. (In modern proofs the latter can in fact be replaced by the weaker requirement of consistency.)

normalised, positive, linear functionals on the observables, whereas the statistical states would be σ -weakly continuous and therefore correspond to the normal, positive, normalised, linear functionals on the observables.

A measurement performed by an apparatus A on some observed system O is an interaction establishing certain relations between the states of A and of O . After a measurement, we infer information about the state of the observed system from information we have about the state of the apparatus. I will take it that the states of A and of O refer to the *same* time after the experiment. To describe this inference, let us use a map θ from the power set $\mathcal{P}(\mathcal{S}_A)$ of the set \mathcal{S}_A of apparatus states into the power set $\mathcal{P}(\mathcal{S}_O)$ of the set \mathcal{S}_O of system states. (Note that \mathcal{S}_A denotes the set of all apparatus states, whereas I use S_A for arbitrary sets of apparatus states.) θ assigns to every set S_A of apparatus states (except the empty set) the set $\theta(S_A)$ of object states compatible with the information that the apparatus after the experiment is in one of the states in S_A . This defines the inference map θ which depends on the kind of measurement we are making. θ is different in different measurement situations. But when the observer chooses the experimental set-up, he also chooses a map θ describing how he is going to interpret the pointer reading after the experiment. This map is fixed throughout the measurement. The states in $\theta(\mathcal{S}_A)$ are the possible states of O after the experiment; usually not every state of O is a possible state after the experiment. We have $\theta(\mathcal{S}_A) \subset \mathcal{S}_O$.

Knowing that if the apparatus after the experiment is in a state s_A the observed system must be in a state in $\theta(\{s_A\})$, one infers from the information that the apparatus after the experiment is in one of the states in S_A that the state of the observed system must be in

$\bigcup_{s_A \in S_A} \theta(\{s_A\})$. So $\theta(S_A) = \bigcup_{s_A \in S_A} \theta(\{s_A\})$.

Definition 1 *I will say that in an experiment with inference map θ a state $s_o \in \mathcal{S}_O$ is exactly measurable if after the measurement there exists a set $S_A \in \mathcal{P}(\mathcal{S}_A)$ of apparatus states referring uniquely to the state s_o , i.e. $\theta(S_A) = \{s_o\}$.*

An experiment with inference map θ is said to be able to distinguish between the states s_1, s_2 if there is one set S_A^1 of final apparatus states referring to s_1 , but not to s_2 , and another set S_A^2 referring to s_2 but not to s_1 : $\theta(S_A^1) \ni s_1 \notin \theta(S_A^2)$ and $\theta(S_A^1) \not\ni s_2 \in \theta(S_A^2)$.

If a state s_o is exactly measurable, we can say that if the apparatus is in one of the states in S_A , then the measured system is with certainty in the state s_o . (In general S_A will consist of several apparatus states, because we usually do not make the inference from the exact state of the apparatus, but rather just from the pointer value.)

If a state is exactly measurable, it can be distinguished from any other possible final state, i.e. from any other state in $\theta(\mathcal{S}_A)$. But for two states to be distinguishable it is not necessary that either of them is exactly measurable. If all possible final states are distinguishable from each other, then they are all exactly measurable.

Distinguishability in measurements from outside. Let me say something about the relevance of what I am going to prove. The concept of exact measurability is strong. The results to follow—namely the impossibility of distinguishing from inside all states—do not deny the possibility of observers knowing *something* about their own states.

In classical mechanics all individual states (i.e. points in phase space) at least in principle can be distinguished by a joint measurement of position and momentum. There is no lower bound to the

accuracy of such a measurement. Also, statistical states (i.e. probability distributions on phase space) can be distinguished in statistical experiments. Note, however, that individual states even in principle cannot be measured exactly in a statistical experiment.¹¹ But I do not think that this is a problem. After all, it is only due to the fact that in a statistical description of experiments one tries to use a concept of state which describes individual systems.

In quantum mechanics the situation is different. No single experiment can distinguish all states of an individual system: the only pure states which by a first kind measurement are exactly measurable in the sense above are the eigenstates of the measured observable. On the statistical level everything is alright again: there are statistical experiments which can distinguish all statistical states. For spinless particles this can be done for example in unsharp joint measurements of position and momentum.¹² These are informationally complete¹³ and therefore

¹¹The reason for this (see Primas (1980)) is the following. In a statistical experiments we measure probability distributions which are σ -additive. Defining—in a somewhat operationalist spirit—statistical states to be what you can measure in statistical experiments, one takes the statistical states to be those which induce probability distributions. For states on von Neumann algebras this is equivalent to both, σ -weak-continuity or normality. Taking $L^\infty(\Gamma)$ as the algebra of observables of the classical system with phase space Γ , the statistical states of a classical system correspond to the normalised elements of the predual $L^1(\Gamma)$. They are probability measures on the phase space. Since there is no normalised L^1 -function on Γ whose support is just one point, no individual state is a statistical state and vice versa. Therefore statistical experiments cannot distinguish individual states.

¹²See Busch 1982, Mittelstaedt *et al.* 1987.

¹³A positive-operator-valued measure a on the value space \mathbb{R}^n is called *informationally complete* if $\text{tr}(a(\Delta)\rho_1) = \text{tr}(a(\Delta)\rho_2)$ for all Borel subsets Δ of \mathbb{R}^n is only possible if the density matrices ρ_1, ρ_2 are equal. The POV-measures on \mathbb{R}^n describe generalised observables with values in \mathbb{R}^n . Observables in the traditional sense are self-adjoint operators on the Hilbert space and induce via their spectral resolution a PV-measure on \mathbb{R} . Such an observable can never be informationally complete. A set of traditional observables $\{A_1, A_2, \dots\}$ can be informationally complete: this means that $\text{tr}(A_i\rho_1) = \text{tr}(A_i\rho_2)$ for all i implies $\rho_1 = \rho_2$.

can distinguish all statistical states.

Traditionally it has been considered a peculiarity of quantum mechanics that no single experiment can distinguish all pure states. But the argument I am going to present shows that the same occurs for any measurement where the observer is properly included in the observed system. This is true for classical theories as well as for quantum theories, and irrespective of the character of the time evolution. So many quantum mechanical lessons about the rôle of the observer are perhaps not so specific to quantum mechanics. Rather they seem to reflect a more general problem.

Measurements from inside. Now let us return to the argument. To bring self-reference into the game consider the case where the apparatus is measuring a system in which it is contained. So the observed system O is composed of the apparatus A and of a residue R . We assume that the observed system has strictly more degrees of freedom than the apparatus and contains it. This can be formulated in an *assumption of proper inclusion*:

$$(\exists s, s' \in \mathcal{S}_O) : s|_A = s'|_A, s \neq s'.$$

Here $s|_A$ denotes the state of A which is determined by restricting the state s of O to the subsystem A . So $|_A$ describes a map from the states of O to the states of A . (Later on I will, by slight abuse of notation, also denote by $|_A$ the map from $\mathcal{P}(\mathcal{S}_O)$ into $\mathcal{P}(\mathcal{S}_A)$ defined by $S_o|_A := \{s|_A : s \in S_o\}$.) In classical mechanics, for example, a map $|_A$ is defined by discarding coordinates which refer to degrees of freedom of O which are not in A . In quantum mechanics, one can take $|_A$ to be for example the partial trace over R . For our purposes it is enough to take an arbitrary but fixed map.

Whether the assumption of proper inclusion is satisfied or not depends not only on the sets $\mathcal{S}_A, \mathcal{S}_O$ but also on the restriction map $|_A$. One can give examples of sets $\mathcal{S}_A, \mathcal{S}_O$ and two restriction maps such that the assumption of proper inclusion is satisfied with respect to one but not the other.¹⁴ This may seem odd but it is not. After all, the elements of \mathcal{S}_A and \mathcal{S}_O are states of different systems. Therefore, even if \mathcal{S}_A is some subset of \mathcal{S}_O , we cannot infer that A is a subsystem of O ; an arbitrary subset of \mathcal{S}_O can in general not be interpreted to be the set of states of a subsystem of O . The restriction map $|_A$ gives physical information which is not reflected in the structure of the sets \mathcal{S}_A or \mathcal{S}_O , namely the fact that A is a subsystem of O . That A is a subsystem of O does not only depend on the abstract structure of A (and of O), but on *which* system A is. If A and A' are isomorphic and A is a subsystem of O , it does not follow that A' is a subsystem of O .

The assumption of proper inclusion seems trivial in the sense that the bigger system O needs more parameters to fix its state. But it excludes situations in which each physically possible state of the whole system is uniquely determined by a state of a subsystem together with some constraint. (I take constraint to mean that states violating the constraint are physically impossible in the sense that the system can never be in such a state.)

Albert's results on self-measurement. The quantum mechanical automata described by Albert (1983, 1987) measure also something about themselves, but they do not attempt to determine their own state ex-

¹⁴Take for example as \mathcal{S}_O the natural numbers and as \mathcal{S}_A the even natural numbers. If one takes as restriction map $\mathcal{S}_O \rightarrow \mathcal{S}_A : n \mapsto 2n$, then the assumption of proper inclusion is not satisfied. If the restriction map associates to every $n \in \mathcal{S}_O$ two times the biggest natural number less or equal to $n/2$, then the assumption of proper inclusion is satisfied.

actly. In Albert's description the apparatus A is composed of several subsystem A_1, A_2, A_3, \dots . Measurement results of non-commuting observables A, B of a system S are displayed by pointer observables P_A of A_1 and P_B of A_2 . Since $[P_A \otimes \mathbf{1}_{A_2}, \mathbf{1}_{A_1} \otimes P_B] = 0$ but $[A, B] \neq 0$, measurement results for A and B can be displayed simultaneously but they cannot both be accurate. In Albert's kind of self-measurement an observable $B^{(1)}$ of the system $A_1 \cup S$ is measured by the subsystem A_3 of the apparatus with the pointer observable $P_{B^{(1)}}$. The crucial point is now that even if $[\mathbf{1}_{A_1} \otimes A, B^{(1)}] \neq 0$, the apparatus $A_1 \cup A_3$ can measure them both simultaneously with full accuracy if $[B^{(1)}, \mathbf{1}_{A_1} \otimes A] = [B^{(1)}, P_A \otimes \mathbf{1}_S]$. This peculiarity is due to the fact that the measurement of the observable $B^{(1)}$ involves a measurement on the system A_1 . Since the apparatus $A_1 \cup A_3$ makes this measurement, it is surely a self-measurement in the sense that the apparatus $A_1 \cup A_3$ is partially contained in the observed system $A_1 + S$. But it is not fully contained in the observed system. Therefore Albert's measurements are not self-measurements in the stronger sense that the assumption of proper inclusion is fulfilled. Albert's conclusions are therefore not related to the results of this chapter.

A first attempt. For exact measurability of *all* states it is necessary (but not sufficient) that there is a surjective map from the states of A onto the states of O . But if additionally A is a properly included in O , there are strictly more states of O than states of A . If O has only finitely many possible states, this already excludes the possibility of exact measurement of all states from inside the observed system.

If we deal with systems with infinitely many possible states, it would be natural to require *continuity* of the mapping: if two states of A are

close then the corresponding states of O should also be close to each other. This additional requirement in classical mechanics implies that the phase spaces of A and O must have the same dimension, because there could not be a continuous bijection between the phase spaces if they were of different, finite dimension. But the phase spaces of A and O cannot have the same dimension if A is properly included in O . So under the assumption of continuity exact measurability of all states from inside is, in classical mechanics, at most feasible if the phase space \mathcal{S}_O is infinite dimensional. Since this case is difficult to handle I will drop the assumption of continuity altogether. Instead I take an entirely different approach.

A consistency condition. The states of the apparatus after the measurement are self-referential: they are states in their own right, but they also refer to states of the observed system in which they are contained. This leads to a *meshing condition* for the inference map θ which must be satisfied lest the inference map be contradictory:

For every apparatus state $s_A \in \mathcal{S}_A$, the restriction of the system states $\theta(\{s_A\})$ to which it refers should again be the same apparatus state s_A .

So meshing can be written:

$$\forall s_A \in \mathcal{S}_A : \{s|_A : s \in \theta(\{s_A\})\} = \{s_A\}.$$

(By a slight abuse of notation I will write $\theta(\{s_A\})|_A$ instead of $\{s|_A : s \in \theta(\{s_A\})\}$.)

From the physical point of view the meshing condition is not a restrictive requirement. Rather it is motivated by logic: it ensures that we cannot arrive at contradictory conclusions about the apparatus state. Assume that the meshing condition is violated and that therefore

there is a state $s' \in \theta(\{s_A\})$ such that $s'|_A \neq s_A$. Then knowing that after the experiment the apparatus is in the state s_A , we would conclude that O is in one of the states in $\theta(\{s_A\})$, possibly in s' . From this in turn we conclude that A can be in the state $s'|_A$, which contradicts the assumption that A is in the state s_A . Note that the meshing condition has to be imposed because both s_A and $\theta(\{s_A\})|_A$ describe the state of A at one given time. This reflects the fact that self-reference problems only occur if an observer wants to know his *present* state.

The main result. With the meshing condition at hand we can now establish that not all states of a system can be measured exactly by an internal observer. The intuitive reason for this is that the meshing condition and the assumption of proper inclusion prevent the existence of a bijection from \mathcal{S}_A to \mathcal{S}_O .¹⁵

Theorem 1 *The assumption of proper inclusion and the meshing condition imply that not all states of a system can be measured exactly by an internal observer:*

$$\exists s_o \in \mathcal{S}_O, \forall S_A \in \mathcal{P}(\mathcal{S}_A) : \theta(S_A) \neq \{s_o\}.$$

PROOF: To prove this by reductio, suppose that the observer can measure all states of O exactly: $(\forall s \in \mathcal{S}_O)(\exists S_A \in \mathcal{P}(\mathcal{S}_A)) : \theta(S_A) = \{s\}$.

¹⁵There could, of course, be bijections *not* fulfilling the meshing condition. For example, if \mathcal{S}_A is infinite but countable, and if for every state $s_A \in \mathcal{S}_A$ there are only finitely many $s \in \mathcal{S}_O$ such that $s|_A = s_A$, then there is a bijection ϕ between \mathcal{S}_A and \mathcal{S}_O . But since $\exists s_A \in \mathcal{S}_A : \phi(\{s_A\})|_A \neq \{s_A\}$ the meshing condition is not fulfilled. Therefore ϕ describes a contradictory inference.

There could also be bijections between \mathcal{S}_A and \mathcal{S}_O fulfilling the meshing condition but violating the assumption of proper inclusion. Take for example as \mathcal{S}_O the natural numbers and as \mathcal{S}_A the even natural numbers. If one takes as restriction map $\mathcal{S}_O \rightarrow \mathcal{S}_A : n \mapsto 2n$, and as inference map $\theta : \mathcal{P}(\mathcal{S}_A) \rightarrow \mathcal{P}(\mathcal{S}_O), \{2n\} \mapsto \{n\}$, then the meshing condition is satisfied because $\theta(\{2n\})|_A = \{2n\}$. But the assumption of proper inclusion is not satisfied. This is natural because $\mathcal{S}_A \subset \mathcal{S}_O$ does *not* imply that A is a subsystem of O . See the discussion p. 28.

This assumption together with the meshing condition will lead to a contradiction.

From the assumption of proper inclusion it follows that $(\exists s, s' \in \mathcal{S}_O) : s|_A = s'|_A, s \neq s'$. By assumption there are $S_A, S'_A \in \mathcal{P}(\mathcal{S}_A)$ such that $\theta(S_A) = \{s\}, \theta(S'_A) = \{s'\}$. Since $\bigcup_{s_A \in S_A} \theta(\{s_A\}) = \theta(S_A) = \{s\}$ there is a $s_A \in S_A$ such that $\theta(\{s_A\}) = \{s\}$. Similarly, there is a $s'_A \in S'_A$ such that $\theta(\{s'_A\}) = \{s'\}$. Repeated application of the meshing condition yields

$$\begin{aligned} \{s\} &= \theta(\{s_A\}) = \theta(\{\theta(\{s_A\})|_A\}) = \theta(\{s|_A\}) = \\ &= \theta(\{s'|_A\}) = \theta(\{\theta(\{s'_A\})|_A\}) = \theta(\{s'_A\}) = \{s'\}, \end{aligned}$$

contradicting $s \neq s'$. □

Lemma 1 *The meshing condition implies that*

$$(\forall s_A) : \theta(\{s_A\}) = \{s \in \mathcal{S}_O : s \in \theta(\mathcal{S}_A), s|_A = s_A\}.$$

PROOF: Let $s \in \theta(\{s_A\})$, then the meshing condition implies $s|_A \in \theta(\{s_A\})|_A = \{s_A\}$. So $s|_A = s_A$ and $\theta(\{s_A\}) \subset \{s \in \mathcal{S}_O : s \in \theta(\mathcal{S}_A), s|_A = s_A\}$. Conversely, let $s \in \mathcal{S}_O$ be such that $s|_A = s_A$ for some s_A and $s \in \theta(\mathcal{S}_A)$. Then there is a $s'_A \in \mathcal{S}_A$ such that $s \in \theta(\{s'_A\})$. Then again from the meshing condition we conclude that $s|_A = s'_A$. So $s_A = s'_A$ and $s \in \theta(\{s_A\})$. □

Theorem 2 *Let s_1, s_2 be two states of O fulfilling $s_1|_A = s_2|_A$. Then there is no inference map θ , and thus no measurement using as apparatus A , which can distinguish s_1 and s_2 :*

$$(\forall \theta) : \left((\exists S_A^1, S_A^2 \in \mathcal{P}(\mathcal{S}_A)) : \theta(S_A^1) \ni s_1 \notin \theta(S_A^2), \theta(S_A^1) \not\ni s_2 \in \theta(S_A^2) \right).$$

PROOF: Assume that there exists an inference map θ , and sets S_A^1, S_A^2 of apparatus states such that $\theta(S_A^1) \ni s_1 \notin \theta(S_A^2), \theta(S_A^1) \not\supseteq s_2 \in \theta(S_A^2)$. This will lead to a contradiction.

From $s_1 \in \theta(S_A^1) = \bigcup_{s \in S_A^1} \theta(\{s\}) \not\supseteq s_2$ we conclude that there is a $s_A^1 \in S_A^1$ such that $s_1 \in \theta(\{s_A^1\})$ and that $s_2 \notin \theta(\{s_A\})$ for all $s_A \in S_A^1$. $s_2 \in \theta(S_A^2)$ implies that $s_2 \in \theta(\mathcal{S}_A)$. Using the Lemma we conclude from $s_1 \in \theta(\{s_A^1\})$ that $s_1|_A = s_A^1$. Since $s_2 \in \theta(\mathcal{S}_A)$ and $s_2|_A = s_1|_A = s_A^1$ we conclude from the Lemma also that $s_2 \in \theta(\{s_A^1\})$. This is in contradiction with the fact that $s_2 \notin \theta(\{s_A\})$ for all $s_A \in S_A^1$. \square

The result of Theorem 1 can be reformulated in a way reflecting the analogy with Gödel's result.

Corollary 1 *Under the assumption of proper inclusion, if all states are exactly measurable from inside the system then the inference map θ is contradictory, i.e. the meshing condition is violated.*

The corollary can be written as

$$\begin{aligned} & ((\forall s \in \mathcal{S}_O) (\exists s_A \in \mathcal{S}_A) : \theta(\{s_A\}) = \{s\}) \Rightarrow \\ & \Rightarrow ((\exists s_o \in \mathcal{S}_O) : \theta(\{s_o|_A\})|_A \neq \{s_o|_A\}). \end{aligned}$$

PROOF: From the assumption of proper inclusion we know that there are $s, s' \in \mathcal{S}_O$ such that $s \neq s', s|_A = s'|_A$. By the antecedent, there are states $s_A, s'_A \in \mathcal{S}_A$ such that $\theta(\{s_A\}) = \{s\}$ and $\theta(\{s'_A\}) = \{s'\}$. Assume that θ satisfies the meshing condition for s'_A : $\theta(\{s'_A\})|_A = \{s'_A\}$. Then $\{s'|_A\} = \theta(\{s'_A\})|_A = \{s'_A\}$. So $s'_A = s'|_A$ and $\theta(\{s'|_A\}) = \{s'\}$. This leads to $\{s\} \neq \{s'\} = \theta(\{s'|_A\}) = \theta(\{s|_A\})$. Therefore, if the meshing condition is satisfied for $s'|_A$, then it is *not* satisfied for $s|_A$. \square

The state $s_o|_A$ plays a rôle analogous to the Gödel-formula. Since $\theta(\{s_o|_A\})|_A \neq \{s_o|_A\}$, this state is self-referential in a paradoxical way. A second analogy becomes apparent when we reformulate the main result in still another way. Recall that an observable is called *informationally complete*¹⁶ if by measuring it one can distinguish all the states. Now the main result can be formulated in a way reminiscent of Gödel's incompleteness theorem: *No measurement from inside the observed system can be informationally complete.*

These conclusions are valid for classical systems as well as for quantum systems, and irrespective of the character of the time evolution. In spite of the similarities with Gödel's proof we should not forget the fundamental differences between the two situations.

1.3 Universal Validity Revisited

Now let us return to the question of how we should conceive of universally valid theories. If a theory is universally valid in the absolute sense, it does not allow for an observer not described by the theory. Take O_u to be the biggest system described by an absolutely universally valid theory. (O_u might be called the "world", or the "universe".) As all potential observers are described by the theory, O_u does not have any outside observer.¹⁷ If, and this is a slightly stronger assumption, the union of all observers fulfils the assumption of proper inclusion, then by Theorem 1 there are some states of O_u which cannot be measured exactly by any observer, not even by all of them together. (It does not help to try and share out the work of measuring the state of O_u

¹⁶See footnote 13.

¹⁷In the terminology of Roessler (1987), Finkelstein (1988), and Primas (1990a) a system without external observer is called *endophysical*.

between several observers. For if the union of observers still obeys the assumptions of proper inclusion and the meshing condition, then Theorem 1 holds.) *So no experiment can distinguish all states of O_u .* Is it acceptable that an absolutely universally valid theory describes systems for which there are no experiments, which at least in principle can distinguish all states? How you answer this question depends on your philosophical proclivities.

A physical realist would rather not dismiss a theory just because it does not make sufficient reference to test procedures. In his opinion, there are entities which in some sense are independent of human knowledge. Statements about these entities should not be conflated with statements about knowledge of the entities. From this point of view, the fact that no experiment can, even in principle, distinguish all states is not in itself objectionable. Accordingly, a physical realist would not take the above argument as sufficient reason to exclude the possibility of absolutely universally valid theories.

An extreme operationalist would say that a physical theory is meaningless unless it is linked to procedures for obtaining knowledge. So he might insist on using the term “state” in a way which guarantees that there is some experiment which at least in principle can distinguish all states, even if technical problems make this difficult in practice. Consequently, the extreme operationalist would think that a theory should be operational in the sense that there is some experiment conceivable which can distinguish all states. But, as we have seen above, no experiment can distinguish all states of the largest system O_u described by an absolutely universally valid theory.¹⁸ From this extremely op-

¹⁸ Still it may be possible that for any two states of O_u , there is some experiment able to distinguish these two states. In classical mechanics this possibility cannot be denied. In quantum mechanics, however, there are states whose restrictions to all

erationalist point of view the possibility of absolutely universally valid theories, deterministic or not, would have to be rejected.

The operationalist, being forced to deny the possibility of *absolutely* universally valid theories, has to find a different, weaker concept of universal validity. The first thing to realise is that for an external observer the assumption of proper inclusion is violated, and therefore the conclusion of Theorem 1 does not apply. So an external observer, or an at least partially external one, may be able to distinguish all states of the observed system. Let us take O to be the biggest system described by the operationalist's theory. Since the operationalist requires that some experiment must be at least conceivable which can distinguish all states of O , he has to admit observers partially outside O . In what sense can a theory that admits observers outside the biggest system it can describe, be universally valid?

Interpreting “describe” in the ontic sense of “is true of”, a theory with an observer outside the biggest system it can describe is not universally valid at all. But an operationalist would prefer to interpret “describe” in the epistemic sense of “can be applied by an observer so as to lead to asserted sentences”. What a theory can describe therefore depends on the observer applying the theory. The above results imply that no observer can apply the theory to the whole world: if he applies it to a system he is properly contained in, then with no experiment can he distinguish all states of the system. For each observer, the biggest system to which he can apply the theory does not contain himself. Still, the theory might be universally valid in the sense that for every part

subsystems coincide but which are still different. By Theorem 2, such states are not distinguishable by any internal apparatus. For O_u there is no possible apparatus which at least partially is outside O_u . Thus, if O_u is a quantum system, it has states which cannot be distinguished by any experiment. See also section 1.4.

of the world some observer can apply the theory to it. Since—for an operationalist—the range of applicability of a theory depends on the observer, I call such a theory *universally valid in this relative sense*. Operationally fully justified theories can be universally valid at most in this relative sense.

Note that this position is often taken for quantum mechanics. According to many interpretations, as for example the one of Bohr, or the one of London and Bauer (1939) and Wigner (1961, 1963), or even perhaps¹⁹ the one of von Neumann (1932), the “true” observer (or his mind) cannot be described by quantum mechanics. These authors say that if quantum mechanics is universally valid at all, then it is so only in the *relative* sense that every observer can perhaps apply it to any selected part of the world, except himself. It supposedly applies to Schrödinger’s cat, Wigner’s friend and Wigner himself under the condition that they lose their status of observer and are observed by something or somebody else.

The difference here is that the result that operationally fully justified theories can be universally valid at most in the relative sense was derived without any assumptions on the character of the time evolution. At least if one adopts an operationalist point of view, one can explain by self-reference problems alone why quantum mechanics can be universally valid at most in the relative sense. Also, the result does not rely on having a quantum mechanical theory.²⁰ The whole argument holds true for classical mechanics as well. Even in classical mechanics, experiments from inside are subject to the restrictions on the measur-

¹⁹I mean von Neumann’s remark that the observer’s mind should not be described by quantum mechanics.

²⁰But in the case of quantum mechanics one has the even stronger results. See footnote 18 and section 1.4.

ability of states as described by Theorems 1 and th-distinguishability. It seems that some quantum mechanical lessons about the rôle of the observer reflect, at least in an operationalist framework, a more general problem: self-reference.

1.4 Application to quantum measurements

The main results presented until now are true for classical and for quantum mechanics, and irrespective of the character of the time evolution. Stronger results hold when we take into account particular features of the quantum mechanical situation. This is what I will deal with now.

1.4.1 EPR-correlations and the universal validity of quantum mechanics

EPR-correlations. Consider again an observed system O containing the apparatus A and some environment or residue R , $O = A \cup R$. If the systems A and R have Hilbert spaces \mathcal{H}_A and \mathcal{H}_R as state spaces, then the EPR-correlations in the vector state $\psi \in \mathcal{H}_A \otimes \mathcal{H}_R$ can be obtained for example from the coefficients of ψ in Schmidt's (1908) biorthonormal decomposition. ψ can be expanded as $\psi = \sum \lambda_n \alpha_n \otimes \beta_n$, where $\{\alpha_n\}$ and $\{\beta_n\}$ are sets of orthonormal vectors in \mathcal{H}_A and \mathcal{H}_R respectively. The EPR-correlations vanish if and only if ψ is a product state, i.e. if all coefficients λ_n except one vanish. The phases ϕ_n of the coefficients $\lambda_n =: |\lambda_n| e^{i\phi_n}$ describe the EPR-correlations: if the coefficients λ_n^1, λ_n^2 of two different states ψ^1, ψ^2 of the joint system differ only in their phases ϕ_n^1, ϕ_n^2 , then their restrictions to any one of the two subsystems obtained by partial tracing are the same (mixed) states:

$$\psi^1|_A = \sum_n |\lambda_n^1|^2 |\alpha_n\rangle \langle \alpha_n| = \sum_n |\lambda_n^2|^2 |\alpha_n\rangle \langle \alpha_n| = \psi^2|_A,$$

and similarly for $\psi^1|_R, \psi^2|_R$. The EPR-correlations encode information not contained in the subsystems taken separately: The information obtainable from experiments just on A (resp. on R) is contained in the density matrix $\sum_n |\lambda_n|^2 |\alpha_n\rangle\langle\alpha_n|$ (resp. in $\sum_n |\lambda_n|^2 |\beta_n\rangle\langle\beta_n|$). Information obtainable from both subsystems separately is therefore described by the density matrix

$$\left(\sum_n |\lambda_n|^2 |\alpha_n\rangle\langle\alpha_n| \right) \otimes \left(\sum_n |\lambda_n|^2 |\beta_n\rangle\langle\beta_n| \right),$$

whereas the information encoded in the system as a whole is described by the density matrix

$$|\psi\rangle\langle\psi| = \sum_{n,m} \lambda_n \lambda_m^* |\alpha_n \otimes \beta_n\rangle\langle\alpha_m \otimes \beta_m|.$$

Generalising this property of EPR-correlations in pure states of a composite system, we will say that two arbitrary states differ only in the EPR-correlations between the subsystems if and only if the states are different but the restrictions by partial trace of both states to any of the subsystems coincide. The correlations have been named after Einstein, Podolsky, and Rosen, because in the version of their (1935) argument presented by Bohm (1951, sections 15-19, chapter 22), the antisymmetric spin state of two electrons with total spin zero has this property.

The assumption of proper inclusion. I said that the assumption of proper inclusion excludes situations in which each physically possible state of the whole system is uniquely determined by a state of a subsystem together with some constraint. (As already noted, I take constraint to mean that states violating the constraint are physically impossible in the sense that the system can never be in such a state.)

But after an ideal quantum measurement of an apparatus A on an external observed system R , the assumption of proper inclusion is fulfilled: A is properly included in the composite system $A \cup R$. Agreed, such a measurement establishes strict correlations between a certain quantity of A (the pointer value) and the measured quantity, so that after the experiment some states of A may determine uniquely some states of R and also of $A \cup R$. Still, this is not the case for all states of $A \cup R$: states of the composite system in which these strict correlations do not obtain are physically possible. (Usually before the measurement the joint system is in such a state.) These states guarantee that the assumption of proper inclusion is satisfied.

But what is more, due to the existence of EPR-correlations, the assumption of proper inclusion is satisfied in a more radical way than in classical mechanics. In classical mechanics, the restrictions to A and R of a pure states of O determine this state uniquely. In quantum mechanics, there are uncountably many pure states of O whose restrictions to A and R coincide. I will now present some implications of this.

An observer cannot measure his EPR-correlations with an outside system. Consider two arbitrary states s_1, s_2 of the joint system $A \cup R$ which differ only in the EPR-correlations between A and R . Can the apparatus A distinguish these states? I will argue that this is impossible.²¹

EPR-correlations cannot be measured in experiments just on the external system R . Such experiments can at most determine the reduced density matrix of R . This density matrix does not encode any information about the EPR-correlations between A and R . Therefore correlation experiments have to be measurements on the joint system

²¹In the context of quantum field theories a similar result was shown by Komar (1964).

$A \cup R$. Since we require that A should make these measurements, the measuring apparatus is properly contained in the observed system. We are thus in a position to apply Theorem 2. It implies that for the apparatus A there is no inference map θ , and thus no measurement, such that there is one set S_A^1 of final apparatus states referring, possibly not uniquely, to s_1 , but not to s_2 , and another set S_A^2 referring to s_2 but not to s_1 . According to Definition 1, this is a necessary condition for A to be able to distinguish s_1 and s_2 .

We therefore conclude that A cannot distinguish between states of O which differ only in the EPR-correlations between A and R . But of course an observer only partially or not at all contained in $A \cup R$ could measure the EPR-correlations between A and R .

The universal validity of quantum theory. In section 1.3 I discussed which implications self-reference problems have for the universal validity of physical theories. The starting point of the argument was that, if the union of all observers is properly included in the universe O_u , no experiment can distinguish all states of O_u . This was a consequence of Theorem 1.

In quantum mechanics we have the particular situation that there exist many states of O_u which differ only by the EPR-correlations between the subsystems of O_u . In particular there are many different states which differ only by the EPR-correlations between all potential observers. The restriction of all these states to the observers coincide. Then it follows from Theorem 2 that *given two such states, there is no experiment able to distinguish between them.* This conclusion is stronger than the one of section 1.3.

An operationalist might try maintain absolute universal validity of

a classical theory by renouncing the requirement that there be an experiment able to distinguish all states. Instead he could just require that for any two different states of his theory there is some experiment able to distinguish between them. This option is not open in quantum mechanics: there are states of O_u which cannot be distinguished by any experiment. Therefore even this more modest operationalist would have to admit that quantum mechanics can be universally valid at most in the relative sense.

1.4.2 Subjective Decoherence

Decoherence in quantum measurements. Let us turn to the measurement problem. Here I will sketch the problem only very roughly. A more detailed discussion can be found in section 4.1.1. To keep things simple assume that a discrete non-degenerate observable of R is measured by the apparatus A . Assume, again for the sake of simplicity, that the pointer observable is also discrete and non-degenerate. Let $\{\phi_i\}$ be the eigenstates of the measured observable and $\{\Phi_i\}$ the eigenstates of the pointer observable. Assume also that the measurement interaction establishes strict correlations between the states ϕ_i and the corresponding apparatus states Φ_i .

If the initial state of R is $\sum_i \lambda_i |\phi_i\rangle$, then the final state of the joint system, as produced by the unitary time evolution, will be

$$s_1 := \left| \sum_i \lambda_i \phi_i \otimes \Phi_i \right\rangle \left\langle \sum_j \lambda_j \phi_j \otimes \Phi_j \right|.$$

In this state the interference terms $|\phi_i \otimes \Phi_i\rangle \langle \phi_j \otimes \Phi_j|$, $i \neq j$ do not vanish. If, however, in each single measurement the outcome were one of the unambiguous final states $\phi_i \otimes \Phi_i$, then the density matrix of the

ensemble would be

$$s_2 := \sum_i |\lambda_i|^2 |\phi_i \otimes \Phi_i\rangle \langle \phi_i \otimes \Phi_i|.$$

In this mixed state the interference terms $|\phi_i \otimes \Phi_i\rangle \langle \phi_j \otimes \Phi_j|$, $i \neq j$ do vanish. The transition from s_1 to s_2 is described by the non-selective version of the projection postulate, also called Lüder's rule. As a first step to a solution of the measurement problem it is often required that the time evolution leads to s_2 as final state.²² In such a case we would say that the time evolution brought about complete *decoherence* between the components $\phi_i \otimes \Phi_i$.

Subjective Decoherence. Observe that the density matrices s_1, s_2 both lead by partial tracing to the same apparatus state

$$s_1|_A = s_2|_A = \sum_i |\lambda_i|^2 |\Phi_i\rangle \langle \Phi_i|.$$

Thus we can conclude from Theorem 2 that there is no inference map θ and no two sets S_A^1, S_A^2 of apparatus states fulfilling $\theta(S_A^1) \ni s_1 \notin \theta(S_A^2)$ and $\theta(S_A^1) \not\ni s_2 \in \theta(S_A^2)$. There is no way in which the apparatus A can distinguish the decohered final states s_2 from the superposed final state s_1 . It is impossible for the apparatus A to tell whether decoherence between itself and the observed system has taken place or not.

If the time evolution is always unitary, the joint system after experiment is in the pure state s_1 . But there is no inference map θ and no set S_A^1 from which the observer could infer that the state of the joint system is not s_2 . So the observer cannot exclude that the final state is decohered. Therefore it is possible for the observer to systematically

²²This is only a first step which is sufficient but not necessary. I will come back to this shortly.

mistake the final state to be decohered. This is what I call *subjective decoherence*. Note that subjective decoherence is only possible if s_1 and s_2 are indistinguishable for A but the converse is not true. For with the latter it would also be possible that, while the joint system is actually in the mixed state s_2 , the apparatus has the mistaken impression that the system is in the pure state s_1 . This might be called ‘subjective coherence’.

Theorem 2 only implies indistinguishability by A of s_1, s_2 . It does not favour subjective decoherence over subjective coherence, nor vice versa. What speaks for subjective decoherence is just experimental evidence: Deviations from the unitary quantum mechanical equations of motion have not been observed; so we have evidence for assuming that the state of the apparatus in fact is s_1 . Also, we as observers have the impression of registering unambiguous measurement results in individual experiments; this is evidence that on the statistical level the observer’s impression is described by the mixed state s_2 . Because s_1 and s_2 are not distinguishable by A , *there is no contradiction between the unitary Schrödinger time evolution for statistical states and the fact that the statistics of A ’s impression about the final state of the joint system is described by the decohered state s_2* . This should not be taken as a solution of the measurement problem, not even on the statistical level. The argument merely relates to the vanishing of interference terms in s_2 , but s_2 can in general not be given an ignorance interpretation. This will be discussed on p. 49.

The phenomenon of subjective decoherence does not appear at the level of single experiments. The reason for this is the following. The final state $\phi_i \otimes \Phi_i$ in a single run of the experiment has as reduced density matrix $|\Phi_i\rangle\langle\Phi_i|$ which does not coincide with the reduced density ma-

trix of the individual state s_1 produced by the unitary time evolution. Since $|\Phi_i\rangle\langle\Phi_i| \neq s_1|_A$, Theorem 2 does not apply. A can distinguish the individual state s_1 produced by the unitary time evolution from the individual state $\phi_i \otimes \Phi_i$. There *is* a contradiction between the unitary Schrödinger time evolution for individual systems and A 's impression that the final state of the joint system after a single run is some $\phi_i \otimes \Phi_i$.

Why subjective decoherence is subjective. Even if the real statistical state of the joint system is s_1 , the apparatus A can have the impression that the statistical state of the joint system is s_2 . But it is only the apparatus A which—being properly contained in the observed system $A \cup R$ —cannot distinguish s_1, s_2 . An apparatus outside $A \cup R$ can measure the interference terms $|\phi_i \otimes \Phi_i\rangle\langle\phi_j \otimes \Phi_j|$. From outside it *is* possible to distinguish the states s_1, s_2 . Therefore subjective decoherence takes only place relative to an internal observer.

Implicit in this discussion is the assumption that it is possible to assign an objective meaning to the quantum mechanical state. Although I personally think that this assumption can be justified, there are other interpretations of quantum mechanics which hold that the quantum mechanical state is essentially an epistemic notion and merely describes our knowledge of the system.²³ On such an epistemic reading of the quantum mechanical state, subjective decoherence is very spectacular: the indistinguishability by A of s_1 and s_2 implies that in both cases A assigns subjectively the same epistemic state to the joint system. For A , s_1 and s_2 are *one* epistemic state. This is not quite the same as

²³See von Weizsäcker (1941, 1961), Schlieder (1968), Süßmann (1957). Such interpretations have proved particularly useful in relativistic descriptions of the measurement process (Schlieder (1968)). On a purely epistemic interpretation of the quantum mechanical state the non-local correlations in EPR-type experiments are much less troublesome.

the mind's ability to collapse the quantum state, as claimed by London and Bauer (1939) or Wigner (1963), because these authors do not necessarily share the purely epistemic interpretation of the quantum state.

Comparison with Everett's relative state interpretation How is subjective decoherence related to Everett's (1957) relative state interpretation? I will briefly point to some similarities and to some differences.

There are some obvious similarities between Everett's interpretation and subjective decoherence. In both theories it is essential that the observer makes his observations from within the observed system, which in Everett's theory is the universe. Also, the trick in both theories is that the observer cannot see the state of the universe correctly. Thereby it is explained why the observer still can have his traditional perception of a world which in reality is the quantum world. The collapse postulate, which is usually employed to explain our traditional perception in a quantum world governed by the Schrödinger equation, therefore is redundant in both approaches.

But there are some important differences. The first seems rather technical, but it implies a second which is of conceptual importance. Let us compare the ways in which a state of a composite system determines the state of a subsystem. For subjective decoherence this restriction map is the partial trace. It is independent of the states of the subsystems and leads to *mixed* states of the subsystems. In Everett's theory a state of the composite system determines the state of one subsystem only *relative* to the state of the other subsystem. Everett thereby arrives at *pure* states of the subsystems.

This technical difference leads to an important conceptual differ-

ence of the two approaches. Subjective decoherence leads to a mixed state of the composite system, and mixed states are assigned to the subsystems as well. Therefore this approach is necessarily *statistical* in its language and intention.²⁴ Everett's approach retains the pure state of the composite system and assigns also pure states to the subsystems. Therefore it can be (and is) interpreted individually.

Why objective decoherence is still necessary. Decoherence is an essential part of several approaches to the measurement problem.²⁵ If subjective decoherence alone can explain why the final statistical state seems to be s_2 , what do we need objective decoherence for? Also, at the first glance, the relevance of objective decoherence seems further impaired by the fact that in finite time it is usually only approximate, whereas subjective decoherence is instantaneous and strict.

However, objective decoherence still is an essential feature of the quantum theory of large systems. Subjective decoherence alone is not enough to explain for example, why it is so difficult to measure the (EPR-)interference terms of macroscopic objects with their environment. Measurements of such interference terms are made from outside: the interference terms are between parts of an external system and not between the apparatus used in the measurement and some other system R . Only these latter interferences are unobservable according to subjective decoherence. The interferences within an external system are observable and are not subject to subjective decoherence.

There is another problem which subjective decoherence cannot solve

²⁴As argued above, although it is statistical in spirit, subjective decoherence alone, due to the inadmissibility of the ignorance interpretation of the statistical states, cannot give a satisfactory statistical description.

²⁵See for example Zurek (1982), Dieks (1989), Hannabus (1984), Gell-Mann and Hartle (1988).

alone, and where objective decoherence is essential: A measurement changes the probabilities of subsequent experiments. This real change cannot be explained by subjective decoherence alone. Objective decoherence is necessary to explain why after the interaction with a *microscopic* system the interference terms, *as measured from outside*, are considerable, whereas these terms, again as measured from outside, are very small after the interaction with a *macroscopic* system.

Decoherence, subjective or objective, is not enough. Decoherence is the transition from a density matrix s_1 to s_2 . If in each single measurement on an ensemble we get a definite outcome $\phi_i \otimes \Phi_i$, and we get the outcome i with probability $|\lambda_i|^2$, then the density matrix of the statistical state is s_2 . But it is not possible to go the other way: the density matrix s_2 does not only describe ensembles where each single constituent is in one of the states $|\phi_i \otimes \Phi_i\rangle\langle\phi_i \otimes \Phi_i|$. This problem of the ignorance interpretability of the mixed state s_2 is often referred to as the “problem of the preferred basis” and will be discussed in more detail in section 4.1.2. Here I will just make a few remarks on this problem.

There are many other ensembles with the same density matrix; a mixed state in general can be decomposed in many ways into a convex sum of pure states.²⁶ The problem is that the space of statistical quantum mechanical states is not a simplex. Therefore mixed states can be decomposed in many different ways into pure states. This means that many different ensembles (decompositions) can lead to the same statistics (mixed state).

²⁶The corresponding theorem (see e.g. Bratteli and Robinson (1979), theorem 4.2.3.) says that the decomposition of a state ω into pure states is unique if and only if the commutant of the GNS-representation associated with the state is abelian.

Even if one additionally imposes the condition that the pure states in the decomposition are orthogonal, there are cases where the decomposition is not unique.²⁷ Additionally, the requirement of orthogonality is hard to justify: it is no problem to experimentally prepare ensembles whose constituents are not in mutually orthogonal states. (This can be done for example by mixing two beams of electrons leaving Stern-Gerlach devices oriented at non-orthogonal angles.)

Thus the density matrix s_2 describes also ensembles whose constituents are not eigenstates of the pointer and the measured observable. Therefore it is not possible to justify an ignorance interpretation of the statistical state s_2 : We cannot say that if the ensemble is described by s_2 , then each single system in the ensemble is in an unknown but definite eigenstate of the pointer and the measured observable. Decoherence alone is not enough because the density matrix does not determine the ensemble, although the ensemble determines the density matrix.

(Note that the above argument against the ignorance interpretability of s_2 is different from d’Espagnat’s argument against the ignorance interpretability of the mixed states of subsystems which arise from taking the partial trace over the rest. The above argument also applies to mixed states of closed systems.)

For this reason decoherence alone, even on the statistical level, does

²⁷This is the case when at least two coefficients in the Schmidt decomposition are equal. The argument for this is well known. A simple example is the following one. Take the case that $i \in \{1, 2\}$, $|\lambda_1|^2 = |\lambda_2|^2 = \frac{1}{2}$. Then we have

$$\begin{aligned} s_2 &= \frac{1}{2}|\phi_1 \otimes \Phi_1\rangle\langle\phi_1 \otimes \Phi_1| + \frac{1}{2}|\phi_2 \otimes \Phi_2\rangle\langle\phi_2 \otimes \Phi_2| \\ &= \frac{1}{2}|\phi_+ \otimes \Phi_+\rangle\langle\phi_+ \otimes \Phi_+| + \frac{1}{2}|\phi_- \otimes \Phi_-\rangle\langle\phi_- \otimes \Phi_-|, \end{aligned}$$

where $\phi_{\pm} = \frac{1}{\sqrt{2}}(\phi_1 \pm \phi_2)$ and similarly for Φ_{\pm} .

not give a satisfying description. To overcome this difficulty it is necessary to either have a classical pointer observable or to change the interpretation of the quantum mechanical formalism.²⁸

1.4.3 Wigner's friend

Subjective decoherence suggests a new way of looking at the so-called paradox of Wigner's friend (1961). Wigner's paper points out that von Neumann's (1932) argument for the arbitrariness of the cut between observed system and observer is problematic.

Wigner considered the situation where an intermediate system is put between himself and the observed system. If this intermediate system is a microscopic system then the joint system consisting of the observed plus the intermediate systems is, at a time before Wigner registered the measurement result, in a state with non-vanishing interference terms, similar²⁹ to s_1 . But when the intermediate system is a human observer like Wigner's friend, then, according to the friend's impression about his own state and to Lüders's rule, the statistical state of the joint system is the mixed state s_2 with vanishing interference terms. Von Neumann's argument about the arbitrariness of the cut is applicable, and therefore, according to von Neumann, the result registered in the end by Wigner himself does not depend on whether the interference

²⁸I am thinking here of modal interpretations of quantum mechanics which in one way or another assign values to some observables of which the state is not an eigenstate. If the pointer observable is classical, then every pure state is an eigenstate of it. Thus all the constituents of all the ensembles describing an arbitrary statistical states are eigenstates of the pointer. Under these circumstances every density matrix can be given an ignorance interpretation. This does not require the assignment of values to observables of which the state is not an eigenstate.

²⁹The important point is that the state of the joint system would still display interference terms. In this respect it is similar to s_1 . But being a subsystem of the bigger system including Wigner, the joint system would probably not be in a pure state. In this respect its state is not similar to s_1 .

terms between the observed and the intermediate systems vanish. Still, the situation seems paradoxical because the state of the joint (observed plus intermediate) system depends on whether the intermediate system is microscopic or a human observer.³⁰

As a way out of this strange situation Wigner (1961) suggested that quantum mechanics only describes inanimate objects. Quantum mechanics in his opinion is not universally valid. The non-linearity displayed by the projection postulate reflects, according to Wigner, the conscious character of human observer.

Objective decoherence suggests a less radical solution to the problem of Wigner's friend. It can explain why the magnitude of the interference terms decreases sharply as the size of the intermediate system increases. Therefore, objective decoherence explains why, after waiting sufficiently long, the actual state s_1 can only with great difficulty be distinguished from the state s_2 . But in the models usually considered³¹ the interference terms do not strictly vanish in finite time. Therefore objective decoherence cannot explain why the statistical state describing the observer's impression in finite time actually *is* s_2 . The problem of Wigner's friend being a matter of principle rather than a practical problem, the solution offered by objective decoherence is not satisfying.

I claim that subjective decoherence offers a solution to the problem of Wigner's friend which, at least in the form the problem was presented by Wigner,³² is satisfying. According to Theorem 2, Wigner's friend

³⁰Actually, the situation becomes paradoxical only if one takes into account that no unitary time evolution can lead to the final state s_2 . This was shown in Komar (1962), Wigner (1963), d'Espagnat (1966), Earman and Shimony (1968), Fine (1970), Busch et al. (1992) and in many other places. Therefore *any* state s_1 produced by a unitary time evolution has to be different from s_2 .

³¹See e.g. Zurek (1982).

³²Wigner asked how it is possible that the interference terms between the observed system and the intermediate system seem to vanish if the intermediate system is

cannot distinguish between situations where the joint system (observed system plus friend) is in a state with non-vanishing interference terms from a situation where the joint system is in a state s_2 with vanishing interference terms. *There is no contradiction between the real statistical state of Wigner's friend being one with non-vanishing interference terms, and the statistical state of his impression about the joint system being one with vanishing interference terms.* (Note, however, that for Wigner as outside observer it is decidable, at least in principle whether the joint system is in a mixed state or not. To achieve this he has to measure the interference terms between observed system and friend.)

This solution does not require that the mind behaves in a non-linear way, or that it is not described by quantum mechanics. Rather it suggests that objective decoherence is not necessary to understand the problem of Wigner's friend (as posed by Wigner). Subjective decoherence is sufficient for that. The friend's impression that his statistical state is s_2 can be explained merely by the fact that he is the last member of the measurement chain leading to *his* perception. His perceptions are measurements from inside, and therefore the restrictions on state-measurability from inside apply. These restrictions lead to subjective decoherence.

the friend, but do not vanish if the intermediate system is an atom. This question is answered by subjective decoherence. Wigner did not ask whether the final statistical state of the observer is ignorance interpretable. This question, which is very important, is not answered by subjective decoherence.

Chapter 2

Phase Space Quantisation

Chapter Abstract

This chapter discusses phase space quantisation by means of systems of covariance. For the Weyl system, the resulting quantisation procedure is compared to configuration space quantisation by systems of imprimitivity, and to geometric quantisation. Then systems of covariance are introduced more generally as POV-measures transforming covariantly with respect to automorphic group representations on von Neumann algebras. A necessary and a sufficient condition for the existence of such systems of covariance is given. Then I prove an extension theorem and an imprimitivity theorem for such systems of covariance.

Chapter Overview

I take quantisation to be a mathematical procedure by which one derives from the classical description of a system a quantum description of the same system. By “same” I mean that all the essential features of the classical system, except its classicality, should be present in the quantised description. There have been disputes about what constitutes the “classicality” of a description. For the purpose I have in

mind, namely phase space quantisation, I take it that a system is classical if its position and momentum observables commute, and quantum mechanical if they do not.

If one wants to give a realist interpretation to physical theories, then quantisation preferably is not regarded as a process which real physical systems undergo. A given physical system in the real world is suitably described either classically or quantum mechanically, or in some other way. But this given system cannot be described classically *and* quantum mechanically, since I took these descriptions to be mutually exclusive. Quantisation, the transition from a classical to a quantum description, is therefore the transition from a correct to the wrong description, or from the wrong to the correct description, or from one wrong description to another wrong one.

Instead I prefer to think of quantisation as an enterprise, perhaps mathematical in nature, of finding parallels between classical and quantum descriptions, and finding out where the parallels break down. This may provide hints under which circumstances a given system may be described classically, and when it may be described quantum mechanically. The transition from the classical the quantum mechanical description is not a transition undergone by the system, but a change in the perspective from which we look at the system.

In this chapter I will deal with phase space quantisation by means of systems of covariance. The first three sections of this chapter review and comment on well-known material on quantisation. Their purpose is to provide a deeper understanding of the new mathematical results of sections 2.4 and 2.5.

In section 2.1 I will introduce some basic notions of classical mechanics needed in this chapter. In section 2.2 I will discuss configuration

space quantisation by Mackey's systems of imprimitivity. This should emphasise the contrast with phase space quantisation by systems of covariance. In section 2.3 I describe, in the example of the Weyl system, phase space quantisation by systems of covariance. The resulting quantisation procedure is compared to geometric quantisation.

In section 2.4, I first introduce generalised systems of covariance. The essence of this generalisation is that automorphic group representations on arbitrary von Neumann algebras are considered, and not just the actions on $\mathcal{B}(\mathcal{H})$ derived from unitary ray representations of a group. Then I describe phase space quantisation by generalised systems of covariance. Finally I give a necessary and a sufficient condition for the existence of generalised systems of covariance. I am not able to give a single condition which is necessary and sufficient. Integrability of automorphic group representations will be the key concept.

In section 2.5, I prove extension and imprimitivity theorems for generalised systems of covariance. This leads to a new outlook on phase space quantisation by generalised systems of covariance. The Naimark extension appears as inverse of quantisation: it leads from the quantum to the classical system.

The quantisation scheme presented can be applied to cases where the quantised system is finite and possesses some classical properties (superselection rules). But the emergence of superselection is more natural for infinite systems (see Chapter 3). Unfortunately infinite systems cannot be treated in the framework presented here because we have to make the assumption of local compactness for the phase space.

2.1 Classical Mechanics

In this section I introduce some notions needed in the following sections of the chapter. First I will discuss how the phase space and the configuration space of classical mechanics can be considered as homogeneous spaces. Then I will give a lattice theoretic and an operator algebraic description of classical mechanics. Finally I will introduce some general problems of quantisation, namely the ordering problem and the Dirac problem.

2.1.1 Phase space and configuration space as homogeneous spaces

Homogeneous spaces. I will first introduce some basic mathematical notions about homogeneous spaces.

Definition 2 *Let G be a locally compact second countable group. A Borel space $(X, \Sigma(X))$ is a G -space if each $g \in G$ acts as a group of Borel automorphism $x \mapsto gx$ of X such that*

$$(i) \quad ex = x \text{ for any } x \in X,$$

$$(ii) \quad g_1(g_2x) = (g_1g_2)x \text{ for any } x \in X \text{ and } g_1, g_2 \in G.$$

The G -space X is transitive or homogeneous if

$$(\forall x_1, x_2 \in X)(\exists g \in G) : gx_1 = x_2.$$

Transitivity is equivalent with the fact that the empty set and X are the only subsets of X which are globally invariant under the automorphism group G .

For a closed subgroup H of G , G/H is the space of left cosets of H in G . It is endowed with the quotient topology. Denote the elements

of G/H by gH . G acts on G/H from the left by $k : gH \mapsto kgH$ of G on G/H .

If X is a transitive standard Borel G -space, then X is homeomorphic to G/H , where H is the stabiliser subgroup $H := \{g \in G : gx_0 = x_0\}$ for an arbitrary point $x_0 \in X$.¹ Choosing a different x_0 leads to a different stabiliser subgroup H . But since all stabiliser subgroups are conjugate to each other, the corresponding left coset spaces are homeomorphic. Under the homeomorphism between X and G/H , the actions $x \mapsto gx$ of G on X and $g : sH \mapsto gsH$ on G/H correspond to each other.

A measure μ on G/H is said to be quasi-invariant under the action $g : x \mapsto gx$ of G on G/H if, for all $g \in G$, $\mu(g\Delta) = 0$ is equivalent to $\mu(\Delta) = 0$. If μ is quasi-invariant then any measure equivalent² to μ is also quasi-invariant. Thus all the measures in the measure class of μ are quasi-invariant. But in general there will be many quasi-invariant measure classes.

Let $L^\infty(X, \mu)$ be the von Neumann algebra of (equivalence classes of) μ -essentially bounded functions on X with values in the complex numbers. G acts as σ -weakly continuous automorphism group $\text{Ad}\lambda$ on $L^\infty(X, \mu)$ by

$$(\text{Ad}\lambda(g)f)(x) := f(g^{-1}x), \quad g \in G, x \in X.$$

If G acts transitively on X , then the requirement of quasi-invariance characterises the measure class uniquely: all the quasi-invariant measures on X are in the same measure class. In that case the definition of $L^\infty(X, \mu)$ does not depend on μ . The von Neumann algebras $L^\infty(X, \mu)$ are all isomorphic and I will denote them simply by $L^\infty(G/H)$.

¹See e.g. Varadarajan 1970, Theorem 8.11.

²Two measures are called equivalent if they have the same null sets.

The phase space as a homogeneous space. The concept of phase space comes from classical point mechanics. The equations of motion of a system of n point particles in three dimensional Euclidean space have $6n$ variables. In the Hamiltonian formulation these are the position coordinates q and the canonically conjugate momenta p . The $6n$ -tuple $\omega(t) = (p(t), q(t))$ is interpreted as the individual state of the system at time t . The phase space Γ is the space of all the possible individual states of the system.

Assume that a locally compact group G acts transitively on the phase space Γ . If G leaves the symplectic structure of the phase space invariant, it might appropriately be called a symmetry group. This is a natural requirement, but for the considerations to follow it is not necessary. All we need is that Γ is homeomorphic to G/H , where H is the subgroup of G leaving fixed some arbitrary but fixed point of the phase space.

In the simple case of a single non-relativistic particle, the phase space Γ is \mathbb{R}^6 , the symmetry group G would be the Galilei group \mathcal{G} , and the stabiliser subgroup H would be the semidirect product $\mathcal{T} \triangleleft \mathcal{R}$ of the time translations \mathcal{T} and the rotation group \mathcal{R} . (I use \triangleleft for the semidirect product.) An element g of the Galilei group is of the form $g = (b, a, v, R)$, where b is a time and a is a spatial translation, v a velocity boost, and R a spatial rotation. According to Lévy-Leblond (1971), the phase space $\Gamma = \mathbb{R}^6$ may be obtained as the coset space

$$\Gamma \cong \mathcal{G}/(\mathcal{T} \triangleleft \mathcal{R}).$$

This is a consequence of the fact that each $g \in \mathcal{G}$ admits a decomposition

$$g = (b, a, v, R) = (0, a - vb, v, 1)(b, 0, 0, R).$$

Thus, if the particle in question has mass m , we may use the parametrisation

$$q := a - vb, \quad p := mv$$

to denote the coset to which the group element g belongs. Hence points $(q, p) \in \mathbb{R}^6 = \Gamma$ may be used to parametrise the coset space $G/H = \mathcal{G}/(\mathcal{T} \triangleleft \mathcal{R})$.

In the case of a free relativistic particle, G would be the Poincaré group \mathcal{P}_+^\uparrow and H again would be the semidirect product of the time translations \mathcal{T} and the rotation group \mathcal{R} . According to Ali (1979), \mathcal{P}_+^\uparrow admits the decomposition

$$\Gamma \cong \mathcal{P}_+^\uparrow / (\mathcal{T} \triangleleft \mathcal{R}).$$

Note that the representation of Γ as G/H is not unique. One can always extend G and H by the same group K to arrive at groups G', H' satisfying $\Gamma = G'/H'$. But the choice of G could be restricted by the possible requirement that it preserves the symplectic structure. The efficiency of the phase space quantisation mechanism to be presented in section 2.3 partly rests on the fact that the phase space Γ need not be homeomorphic to some Euclidean space \mathbb{R}^n . It therefore allows for the description of topologically non-trivial phase spaces.

Often a homogeneous phase space G/H additionally has a *complex* homogeneous structure. Then the action of G on G/H is additionally required to be analytic. For example, the phase space $\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$ of a free particle in Euclidean space is isomorphic to the three dimensional complex space \mathbf{C}^3 . The homogeneous space G/H is called a Kähler manifold if it has a complex manifold structure such that the imaginary part of the complex inner product defines a symplectic form. This symplectic form is essential for the Hamiltonian formalism.

Configuration space as a homogeneous space. Configuration spaces Q sometimes³ are homogeneous spaces as well. For example, Q might be a line, a circle, a sphere, three dimensional Euclidean space, etc.

If G preserves the metric of the configuration space, it might appropriately be called a symmetry group. But for our purposes we do not need to make this assumption. All we need is that the configuration space is a homogeneous space with respect to G .

In the case of a single particle moving in Euclidean space \mathbb{R}^3 , the symmetry group acting on Q is the Euclidean group \mathcal{E} , which is the semidirect product of the group of space translations and the group of space rotations \mathcal{R} . Configuration space $Q = \mathbb{R}^3$ can be written as

$$Q \cong \mathcal{E}/\mathcal{R}.$$

The representation of a given configuration space Q as G/H for some group G and some subgroup H is also non-unique. (Requiring that G preserves the metric would restrict the choice of possible G .) For instance, one may form arbitrary non-trivial extensions E of G by a group K (so $G = E/K$) and let E act on Q via the canonical epimorphism $p : E \rightarrow G$. Such extensions can be classified by cohomological methods. Inequivalent extensions lead to different quantisations.⁴ This is of great physical importance since topological quantum effects can be described by a non-minimal choice of G .

The configuration space Q need not be a Euclidean space. Physical systems with non-Euclidean configuration spaces are of interest because they exhibit topological quantum effects. Well known examples of such

³Configuration spaces can, of course, *always* be regarded as homogeneous groups with respect to the diffeomorphism group of Q . But we cannot exploit this because the diffeomorphism group is not locally compact.

⁴See Landsman (1990a).

effects are the Aharanov-Bohm effect and the Dirac magnetic monopole charge quantisation effect.⁵ In these situations the configuration space arises if one deletes a line, or a point respectively, from \mathbb{R}^3 .

2.1.2 Lattice theoretic and algebraic formulation of classical mechanics

States and observables in classical mechanics. The pure states ω of the system are represented by points in the phase space Γ . They refer to states of individual systems. The observables of an individual system are represented by the real valued functions on Γ . The value of an observable f in a pure state ω is $f(\omega)$.

When one does not know precisely the state ω of the system, but knows that it is more likely to be in some parts of Γ than in others, then one describes the system by a statistical state ρ . Statistical states are probability measures on Γ . $\rho(\Delta)$ is the probability that the actual state ω is in the Borel set $\Delta \subset \Gamma$. A pure state ω may be identified with a statistical state ρ_ω by

$$\rho_\omega(\Delta) := \begin{cases} 1, & \omega \in \Delta; \\ 0 & \text{otherwise.} \end{cases}$$

The statistical state ρ_ω corresponding to a pure state ω is singular in the sense that it is not countably additive⁶ but just finitely additive.

As observables one can take the real valued Borel functions on Γ with the right continuity properties. In the context of differential geometry one often takes the self-adjoint part of $C^\infty(\Gamma)$; in a W^* -algebraic framework one takes $L^\infty(\Gamma)$ with respect to the Liouville measure; in a C^* -framework one usually takes $C_0(\Gamma)$, the continuous functions on Γ

⁵See Landsman (1990b).

⁶A state ρ is countably additive (or σ -additive) if for all sequences of mutually disjoint Borel sets Δ_i we have $\sum_{i=1}^{\infty} \rho(\Delta_i) = \rho(\bigcup_{i=1}^{\infty} \Delta_i)$.

vanishing at infinity. A given statistical state ρ assigns to each observable f a probability measure on the real line by

$$\Delta \mapsto \rho(f^{-1}(\Delta)), \quad \Delta \in \Sigma(\mathbb{R}).$$

If our knowledge of the system is described by the state ρ , then $\rho(f^{-1}(\Delta))$ is the probability that the value of the observable f is in the Borel set $\Delta \subset \mathbb{R}$.

Formulating classical mechanics in terms of lattices. The notions of statistical state and observable depend in no way upon the fact that the phase space Γ is a point set. One can define the statistical states and observables without referring to the points of Γ , but just to the Borel sets of Γ . More generally, it is possible to define states and observables with reference just to orthomodular partially ordered sets. The states and observables of classical mechanics then emerge from the fact that the Borel subsets of the phase space Γ form a Boolean lattice \mathcal{L}_Γ , and thus an orthomodular partially ordered set.

Definition 3 *On an orthomodular partially ordered set \mathcal{L} one can define statistical states and observables in the following way. A statistical state is a probability measure on \mathcal{L} , i.e. a function $\rho : \mathcal{L} \rightarrow [0, 1]$ such that*

- (i) $\sum_{i=1}^{\infty} \rho(a_i) = \rho(\bigcup_{i=1}^{\infty} a_i)$ for any sequence of orthogonal a_i ,
- (ii) $\rho(\mathbf{1}) = 1$.

An observable is an \mathcal{L} -valued function $a : \Delta \mapsto a_\Delta$ on the Borel sets of the real line such that

- (i) $a_{\Delta_1} \perp a_{\Delta_2}$ if $\Delta_1 \cap \Delta_2 = \emptyset$,
- (ii) $\sum_{i=1}^{\infty} a_{\Delta_i} = a_{\bigcup_{i=1}^{\infty} \Delta_i}$ for any sequence of disjoint Δ_i ,
- (iii) $a_\emptyset = \mathbf{0}$.

Note that this lattice theoretic⁷ definition of state only covers truly statistical states: in the case of a classical system with lattice \mathcal{L}_Γ , the statistical states ρ_ω corresponding to a pure state $\omega \in \Gamma$ are not σ -additive and therefore violate condition (i) above. A statistical state ρ and an observable a determine a probability measure on the real line by

$$\Delta \mapsto \rho(a_\Delta).$$

Formulating classical mechanics in terms of operator algebras. In the spirit of Koopman's (1957) Hilbert space formulation of classical mechanics, one can formulate classical mechanics in terms of operator algebras.⁸ As Hilbert space one takes $L^2(\Gamma)$, the complex functions on Γ which are square integrable with respect to the Liouville measure. In the operator algebraic tradition one tries to find an algebra of bounded operators on a Hilbert space such that the bounded observables are the self-adjoint elements of the algebra, and that the states are positive, continuous, normalised linear functionals on the algebra of observables.

The algebra of observables of an individual classical system is taken to be the C^* -algebra of continuous complex functions on phase space vanishing at infinity. The elements of $C_0(\Gamma)$ act on $L^2(\Gamma)$ as multiplication operators. The bounded observables are then the self-adjoint elements of $C_0(\Gamma)$, i.e. the real continuous functions on Γ vanishing at infinity. Note that the so-called algebra of observables $C_0(\Gamma)$ contains operators which are not observables, and that it does not contain all the

⁷The term "lattice theoretic" is a little bit unfortunate because Definition 3 refers to orthomodular partially ordered sets, which are not necessarily lattices. An alternative term would be "quantum logical" but this is equally misleading because \mathcal{L} may describe classical systems.

⁸For a detailed description of classical mechanics in terms of operator algebras see e.g. Reed and Simon (1980, II.5 and 1975, X.14).

operators corresponding to observables: Only the self-adjoint elements of $C_0(\Gamma)$ are interpreted as observables; unbounded real functions on Γ describe observables, although they are not in $C_0(\Gamma)$.

It can be shown⁹ that the extremal continuous linear functionals on $C_0(\Gamma)$ are in one-to-one relation with the points of phase space.¹⁰ Thus to each pure state ρ_ω there corresponds exactly one point $\omega \in \Gamma$ such that the value of the state ρ_ω on the observable f is

$$\rho_\omega(f) = f(\omega).$$

The strictly C^* -algebraic framework is not suitable for a statistical description. In the lattice theoretic context of Definition 3, a statistical state was introduced as a probability measure on an orthomodular partially ordered set \mathcal{L} . To make sense of this definition in an operator algebraic framework, it is necessary that the projectors in the algebra of observables form an orthomodular partially ordered set. This is not in general the case for C^* -algebras. The projections in a von Neumann algebra, however, fulfil this condition and even form a lattice. Therefore in a statistical description, we should take the algebra of observables to be a von Neumann algebra.

Call a state (in the operator algebraic sense) on a von Neumann algebra \mathcal{M} statistical iff it is a statistical state (in the lattice theoretic sense) on the lattice $\mathcal{L}_\mathcal{M}$ of projections in \mathcal{M} . A state ρ on a von Neumann algebra is said to be completely additive if $\rho(\sum_{k \in I} P_k) = \sum_{k \in I} \rho(P_k)$ for any family I of mutually orthogonal projections P_k . Therefore, according to Definition 3, the lattice theoretic states on $\mathcal{L}_\mathcal{M}$

⁹See Kadison and Ringrose 1983, 4.4.

¹⁰This one-to-one correspondence between pure states and phase space points depends sensitively on the choice of $C_0(\Gamma)$ as algebra of observables. If one chooses as algebra of observables the bounded functions on Γ , then the pure states on this algebra by far outnumber the phase space points.

are exactly the completely additive states on \mathcal{M} . A state on a von Neumann algebra is completely additive if and only if it is σ -weakly continuous.¹¹ Thus we have to take as statistical states on \mathcal{M} exactly the states which are not only in the dual \mathcal{M}^* , but also in the predual \mathcal{M}_* of \mathcal{M} . They induce probability measures on the lattice of projections in \mathcal{M} . Some consequences of requiring complete (or countable, which is the same in the case of separable Hilbert spaces) additivity for states will be discussed in section 4.2.

In the classical case, the standard choice for a von Neumann algebra of functions on Γ is $L^\infty(\Gamma, \mu)$, where μ is the Liouville measure. Since $L^\infty(\Gamma, \mu)_* \cong L^1(\Gamma, \mu)$, the above definition of statistical state on $L^\infty(\Gamma, \mu)$ leads exactly to the normalised real-valued elements of $L^1(\Gamma, \mu)$. This coincides with the idea that statistical states of a classical system are represented by the countably additive probability measures on Γ . An observable $f \in L^\infty(\Gamma, \mu)$ associates to every statistical state $\rho \in L^1(\Gamma, \mu)$ a real number

$$f(\rho) = \int_{\Gamma} f(q, p) \rho(q, p) d\mu(q, p).$$

This number is the expectation value of the observable f in the statistical state ρ . (ρ on the left hand side is regarded as an element of $L^\infty(\Gamma, \mu)_*$, whereas on the right hand side it is regarded as an element of $L^1(\Gamma, \mu)$.)

Now assume that a symmetry group G acts on the phase space Γ . So Γ should be a (locally compact) Borel G -space. Let μ be a quasi-invariant measure on Γ . G acts as σ -weakly continuous automorphism group $\text{Ad}\lambda$ on $L^\infty(\Gamma, \mu)$ by

$$(\text{Ad}\lambda(g)f)(x) := f(g^{-1}x), \quad g \in G, x \in \Gamma. \quad (2.1)$$

¹¹See Kadison and Ringrose (1986, theorem 7.1.12).

If the phase space Γ is homogeneous, and it was argued in section 2.1.1 that this is often the case, then there is a unique quasi-invariant measure class on Γ and we may simply write $L^\infty(\Gamma)$ instead of $L^\infty(\Gamma, \mu)$.

This is an example of a classical W^* -systems in the sense of the following definition.

Definition 4 A classical W^* -system (\mathcal{M}, G, α) consists of a commutative W^* -algebra \mathcal{M} , a locally compact, separable group G , and a representation $\alpha : G \rightarrow \text{Aut}(\mathcal{M})$ of G as a group of automorphisms of \mathcal{M} such that

- (i) $\alpha_{g_1 g_2} = \alpha_{g_1} \alpha_{g_2}$ and
- (ii) for all operators $x \in \mathcal{M}$ the function $g \mapsto \alpha_g(x)$ is σ -weakly continuous.

A classical C^* -system is defined in the same way, except that \mathcal{M} is taken to be a commutative C^* -algebra, and that the map $g \mapsto \alpha_g(x)$ is required to be norm continuous.

The Dirac Problem. A quantum description of the system replaces the classical observables by operators on a Hilbert space. Let O be a quantisation map associating Hilbert space operators to functions on phase space. It might seem sensible to require that O has the following properties.

- (1) O is linear.
- (2) $O(1) = \mathbf{1}$.
- (3) $[O(f), O(g)] = -i\hbar O(\{f, g\})$.
- (4) The quantum mechanical position and momentum operators act irreducibly on \mathcal{H} , i.e. $\{O(p), O(q)\}'' = \mathcal{B}(\mathcal{H})$.

(Here I denote by $\{O(p), O(q)\}''$ the von Neumann algebra generated by (the projections in the spectral measures of) the operators $O(p), O(q)$. By $\mathcal{B}(\mathcal{H})$ I denote the algebra of all bounded operators on the Hilbert space \mathcal{H} .)

Condition 4, as it is formulated here, makes sense only if classical position and momentum coordinates q, p are defined globally.

Condition 3 is the requirement that O is a Lie algebra morphism. It formulates the following idea of Dirac: Since the classical observables generate via Hamilton's equations the evolution of the classical system, and since the quantum observables generate via Stone's theorem the evolution of the quantum system, the quantisation map should associate the generators of the classical evolution to the generators of the quantum evolution in such a way that the Lie algebra structure of the generators is preserved.

The fact that the observables are operators on a Hilbert space does not guarantee that the observables describe a quantum system. For example, in the above operator algebraic formulation of classical mechanics, the classical observables are multiplication operators on the Hilbert space of L^2 -functions on the phase space. Classical position and momentum are represented by the multiplication operators q, p . Since they commute and their Poisson bracket does not vanish, in this formulation of classical mechanics, condition 3 is violated. Condition 3 ensures that the position and momentum operators $O(p), O(q)$ fulfil the commutation relations $[O(q), O(p)] = -i\hbar\mathbf{1}$.¹² It is therefore justified to say that the observables $O(f)$ describe a *quantum* system.

¹²Again, speaking of quantum mechanical position and momentum operators $O(p), O(q)$ contains the tacit assumption that classical position and momentum coordinates can be defined globally. In the sequel I will not mention this proviso any more.

Since the Koopman position and momentum operators do not act irreducibly on $L^2(\Gamma)$, the operator algebraic formulation of classical mechanics also violates condition 4. If condition 3 is satisfied, then $O(q), O(p)$ satisfy the commutation relations, but they need not generate an irreducible von Neumann algebra.¹³ If the configuration space Q is Euclidean, then condition 4, together with requiring normality of the representation, implies via the Stone-von Neumann uniqueness theorem that O is unitarily equivalent to the Schrödinger representation.¹⁴

Condition 4 can be given a second motivation. The n -fold direct product of the additive group of spatial translations and boosts acts separately on each of the n particles of the system. This group action is transitive on the phase space: no proper submanifold of the phase space is invariant under the full group. Condition 4 requires something similar for the corresponding quantum system: there should be no proper subspace of \mathcal{H} which is left invariant by all operators representing quantum mechanical observables. One often calls a classical system elementary if the kinematical group acts transitively on the phase space; also, quantum systems associated to irreducible representations of the kinematical group are called elementary. Condition 4 thus can be viewed as the requirement that the quantum system corresponding to an elementary classical system should also be elementary.

However, conditions 1-4, reasonable as they may seem, are not compatible, even in the simple case where the configuration space is Euclidean: van Hove (1951) has shown that the Schrödinger representation can not be extended to non-quadratic polynomials in such a way that condition 3 still holds. By the Stone-von Neumann uniqueness theo-

¹³This is the case for example with the von Neumann algebra generated by the prequantised position and momentum operators of the Weyl system. See page 109.

¹⁴See Mackey 1949, and 1978, ch. 18.

rem, condition 4 requires that any quantisation map O should yield position and momentum operators unitarily equivalent to the ones of Schrödinger. Thus, no quantisation map O defined on all polynomials can simultaneously satisfy conditions 3 and 4.

The quantisation schemes described in the sequel of this chapter can be classified according to which of the conditions 3 and 4 they relax in order to satisfy the other.

The ordering problem. There is considerable arbitrariness in the choice of a quantisation map O . Suppose one has a real-valued smooth function $f(q, p)$ on phase space. Assume that f has a Taylor expansion as

$$f(q, p) = \sum_{m, n=0}^{\infty} c_{mn} q^m p^n,$$

for some suitable coefficients c_{mn} . After quantisation, the variables p, q are replaced by some self-adjoint unbounded operators $O(p), O(q)$ on a Hilbert space fulfilling the commutation relations. This does not determine completely a quantisation map O on $C^\infty(\Gamma)$. One might try to associate to a classical observable f the operator

$$O(f) := \sum_{m, n=0}^{\infty} c_{mn} O(q)^m O(p)^n.$$

But apart from some mathematical problems (e.g. one has to give a meaning to this infinite sum of unbounded operators, and $O(f)$ has to be a self-adjoint operator on a dense domain), one has to decide on a certain ordering of the non-commuting operators $O(q), O(p)$ in $O(q)^m O(p)^n$. It is, for instance, quite arbitrary whether the classical observable qp^2 should correspond to $\frac{1}{2}(O(q)O(p)^2 + O(p)^2O(q))$, or to $\frac{1}{4}(O(q)O(p)^2 + 2O(p)O(q)^2O(p) + O(p)^2O(q))$, or to some other self-adjoint combination.

This arbitrariness of quantisation is called the ordering problem. Different orderings have been discussed in the literature for example by Agarwal and Wolf (1970a and 1970b), and by Srinivas (1975). Ali and Doebner (1990) have argued that although most orderings are mathematically arbitrary and in some sense equivalent, different orderings are suitable for the description of different experimental situations. In the next sections I will often come back to the arbitrariness in the choice of an ordering, and to the associated non-uniqueness in the choice of quantisation maps.

2.2 Quantisation on Configuration Space

Quantisation from a lattice theoretic point of view. One obtains the quantum description of a system when we replace the orthomodular lattice \mathcal{L}_Γ of Borel sets of the phase space by the orthomodular lattice $\mathcal{P}(\mathcal{H})$ of projectors on a Hilbert space \mathcal{H} .¹⁵

Replacing \mathcal{L}_Γ by $\mathcal{P}(\mathcal{H})$, one can still apply the lattice theoretical Definition 3 of statistical states. Gleason's theorem says that the statistical states on $\mathcal{P}(\mathcal{H})$ (for $\dim \mathcal{H} > 2$) are all of the form $\text{tr}(\rho \cdot)$ for some trace class operator ρ of trace norm one. Thus the statistical

¹⁵The replacement of \mathcal{L}_Γ by some orthocomplemented poset is quite natural. This was pointed out by Mackey (1963). He showed the following. Let \mathcal{O} be the observables of the system, and let \mathcal{S} be the set of statistical states. For each Borel set Δ on the real line, denote by $p(a, \rho, \Delta)$ the probability that a measurement of the observable a in the state ρ yields a result in Δ . If the function $p(a, \rho, \Delta)$ fulfils some plausible conditions, there is a unique poset \mathcal{L} , a one-to-one mapping from \mathcal{O} onto the \mathcal{L} -valued functions on the Borel sets of the real line, and a one-to-one mapping from \mathcal{S} into a convex subset of the probability measures on \mathcal{L} such that $p(a, \rho, \Delta) = \rho(a(\Delta))$. The choice of $\mathcal{P}(\mathcal{H})$ rather than some other orthocomplemented poset is less compelling (see Mackey (1963), p.72-74.) Indeed, the choice of some general orthocomplemented poset rather than $\mathcal{P}(\mathcal{H})$ is the starting point of quantum logical generalisations of traditional quantum mechanics. (For an overview see Primas (1980), for a monograph see Beltrametti and Casinelli (1981).)

states on $\mathcal{P}(\mathcal{H})$ are exactly the density matrices.

Also, one can define observables as $\mathcal{P}(\mathcal{H})$ -valued measures on the real line. By the spectral theorem, there is a one-to-one correspondence between these projection-valued measures on \mathbb{R} and the self-adjoint operators on \mathcal{H} . Thus on the lattice $\mathcal{P}(\mathcal{H})$ the observables in the sense of Definition 3 are exactly the self-adjoint operators.

So far we have seen that the replacement of the lattice \mathcal{L}_Γ by $\mathcal{P}(\mathcal{H})$ yields the usual Hilbert space formalism of quantum mechanics. But the replacement of \mathcal{L}_Γ by *some* lattice of projections on a Hilbert space does *not* guarantee that the system described is really a quantum mechanical one. For example, consider the map $E : \mathcal{L}_\Gamma \rightarrow \mathcal{P}(L^2(\Gamma))$ defined by

$$E(\Delta)\psi(q, p) := \chi_\Delta(q, p)\psi(q, p).$$

E is a lattice *morphism* between \mathcal{L}_Γ and a Boolean sublattice of $\mathcal{P}(L^2(\Gamma))$. This Hilbert space representation is a lattice *morphism* and therefore leads to a classical system. This is a lattice theoretic description of Koopman's Hilbert space formulation of classical mechanics.

In order to guarantee that a system described in the Hilbert space formalism is quantum mechanical, one has to impose further conditions on the observables. For example, one can require that the observables generated by quantum mechanical position and momentum act irreducibly. This is condition 4 of the Dirac problem. A different condition would be that the operators corresponding to position and momentum should fulfil the commutation relations $[O(p), O(q)] = -i\hbar\mathbf{1}$, or that the set of all observables acts irreducibly on the Hilbert space. (These last two conditions do not imply condition 4 of the Dirac problem.) All these conditions exclude Koopman's (1957) Hilbert space formulation of classical mechanics.

Speaking lattice theoretically, there are two possible ways to avoid the system described by a Hilbert space formulation being classical: A map from the lattice \mathcal{L}_Γ of all classical properties into $\mathcal{P}(\mathcal{H})$ should not be a morphism: such a map must not preserve the Boolean character of \mathcal{L}_Γ . Alternatively, a Boolean sublattice of \mathcal{L}_Γ can be mapped into a Boolean sublattice of $\mathcal{P}(\mathcal{H})$ by a morphism.

Configuration space quantisation as described in this section takes the second route: spatial localisation properties (but not momentum space localisation properties) of the classical system are mapped by a morphism to spatial localisation properties of the quantum system. Phase space quantisation by systems of covariance as described in sections 2.3 and 2.4 takes the first route.

Configuration space quantisation. Consider a classical particle moving in a homogeneous configuration space $Q = G/H$. The action of G on the configuration space can be represented equivalently in two ways. From a lattice theoretical point of view, it is described by a representation of G as the automorphism group $\Delta \mapsto g\Delta$ of the Boolean σ -algebra \mathcal{L}_Q . From an operator algebraic point of view, the symmetry group G was represented by the representation $\text{Ad}\lambda$ of G as automorphism group of $L^\infty(G/H)$ (see equation 2.1).

A first step in configuration space quantisation is to take a representation of the localisation properties as projectors on a Hilbert space. From a lattice theoretical point of view the Boolean σ -algebra $\Sigma(Q) = \mathcal{L}_Q$ is represented by a Boolean sublattice of the lattice $\mathcal{P}(\mathcal{H})$ of projectors on \mathcal{H} . From an algebraic point of view, the commutative W^* -algebra $L^\infty(Q)$ is represented by a commutative W^* -algebra $\pi(L^\infty(Q))$ of bounded operators on \mathcal{H} .

The second step is to represent on \mathcal{H} the action of the symmetry group G . In the lattice theoretical language, the action $\Delta \mapsto g\Delta$ is replaced by a unitary representation U of G on \mathcal{H} . In the operator algebraic language, the action $\text{Ad}\lambda$ is replaced by the automorphic representation

$$\alpha_g(\pi(f)) := U_g \pi(f) U_g^* \quad f \in L^\infty(Q)$$

of G on $\pi(L^\infty(Q))$.

Localisability and the Hilbert space representation of $\Sigma(Q)$. Let me make some brief remarks about the first step, in which the Boolean σ -algebra $\Sigma(Q)$ is represented as a Boolean sublattice of $\mathcal{P}(\mathcal{H})$. From a mathematical point of view, such a representation might seem very natural. From a physical point of view it is much more problematic. Since Q is the configuration space of the particle, such a representation in fact is a translation of the classical localisation properties to the quantum domain.

Localisability of a particle means that the proposition system must contain a class of propositions which answer the question whether the particle is in this volume element or in that. Assumptions about localisation properties enter in a crucial way into the choice of propositions representing localisation. In classical mechanics, such propositions are represented by the Boolean algebra of Borel sets in the three dimensional physical space. This choice reflects the assumptions that position measurements in different regions do not disturb each other and that such measurements are exactly repeatable. In classical mechanics these assumptions are natural. For the time being, one can make the same assumptions for the localisation of quantum particles:

- (1) Propositions about localisation of non-relativistic quantum systems in different space time regions are compatible.
- (2) Localisation experiments are repeatable.

These assumptions are highly questionable. For example, Hegerfeldt's theorem¹⁶ shows that localisation experiments on relativistic particles are *not repeatable*. In the non-relativistic theory, we can tentatively make these assumptions.

Assumptions (1), (2) imply that propositions about localisation are described by a projection valued (PV) measure E on configuration space.¹⁷

Definition 5 *A PV-measure on a Borel space $(X, \Sigma(X))$ is a mapping E from $\Sigma(X)$ into $\mathcal{P}(\mathcal{H})$ such that*

- (i) $E(\Delta_1 \cap \Delta_2) = E(\Delta_1)E(\Delta_2)$,
- (ii) $E(\Delta_1 \cup \Delta_2) = E(\Delta_1) + E(\Delta_2) - E(\Delta_1)E(\Delta_2)$,
- (iii) $\sum_{i=1}^{\infty} E(\Delta_i) = E(\bigcup_{i=1}^{\infty} \Delta_i)$ for any sequence of disjoint Δ_i , where the sum on the left hand side converges weakly,
- (iv) $E(X) = \mathbf{1}$.

The PV-measure E describing localisation is a lattice morphism from the Boolean lattice \mathcal{L}_Q of Borel subsets of the configuration space into a Boolean sublattice of $\mathcal{P}(\mathcal{H})$. It can be extended to a representation π of $L^\infty(Q)$ as a commutative von Neumann subalgebra of $\mathcal{B}(\mathcal{H})$ such that $\pi(\chi_\Delta) = E(\Delta)$.

The covariance condition. Let us now come to the second step of configuration space quantisation: The action of G on the configuration

¹⁶See Hegerfeldt (1974), Hegerfeldt and Ruijsenaars (1980).

¹⁷The precise argument for this can be found in Jauch (1968).

space $Q = G/H$ is replaced by a unitary representation U of G on the Hilbert space. Accordingly, the action $g : sH \mapsto gsH$ of $g \in G$ on $Q = G/H$ is represented on the lattice of localisation operators by

$$g : E(\Delta) \mapsto U_g E(\Delta) U_g^*, \quad \Delta \in \Sigma(Q).$$

Since G is a symmetry of the configuration space, the simultaneous action of G on the configuration space and of $U(G) \cdot U(G)^*$ on the localisation propositions should not have observable consequences. Thus we have to require commutativity of the diagram

$$\begin{array}{ccc} \Delta & \longrightarrow & g\Delta \\ \downarrow E & & \downarrow E \\ E(\Delta) & \xrightarrow{U \cdot U^*} & U_g E(\Delta) U_g^* \end{array} .$$

The commutativity requirement implies the covariance condition

$$U_g E(\Delta) U_g^* = E(g\Delta), \quad \Delta \in \Sigma(Q), g \in G. \quad (2.2)$$

If we take as configuration space Q three dimensional Euclidean space, and as symmetry group G the Euclidean group \mathcal{E} , then the covariance condition implies isotropy and homogeneity in the physical sense.¹⁸

This covariance condition can be interpreted in the following way. Transforming the system actively changes its state from ψ to $U_g \psi$. Also, the map $E(\Delta) \mapsto E(g\Delta)$ describes an active transformation of the apparatus measuring the observable E . Using the covariance condition one calculates

$$\langle U_g \psi | E(g\Delta) | U_g \psi \rangle = \langle U_g \psi | U_g E(\Delta) U_g^* | U_g \psi \rangle = \langle \psi | E(\Delta) | \psi \rangle.$$

¹⁸See Jauch 1968.

Thus the covariance condition guarantees that an active transformation of the system and a simultaneous active transformation of the apparatus does not change the expectation values of experiments. This is a statement of symmetry of the configuration space under G .

2.2.1 Systems of Imprimitivity

The covariance condition implies that the PV-measure E together with the unitary representation U forms a system of imprimitivity in the following sense.

Definition 6 *A PV-measure E on a Borel space X on which G acts transitively, together with a measurable unitary ray representation U of G is called an transitive system of imprimitivity¹⁹ (E, U) based on X if E acts covariantly with respect to U ,*

$$U_g E(\Delta) U_g^* = E(g\Delta), \quad \Delta \in \Sigma(X), g \in G.$$

A system of imprimitivity (E, U) is called irreducible if the von Neumann algebra $\mathcal{M} := \{E(\Delta)U_g : \Delta \in \Sigma(X), g \in G\}$ generated by E and by U is irreducible.

A transitive system of imprimitivity (E, U) is irreducible if and only if U is irreducible. A non-transitive system of imprimitivity may be irreducible even if U is reducible.

A physical system is called elementary if the corresponding system of imprimitivity is irreducible. This is the case if only the identity operator commutes with both the configuration observables E and with

¹⁹The term “imprimitivity” has been introduced by S. Lie and F. Engel (1888, p. 220) in the context of transformation group theory.

(the self-adjoint generators of) U . This is in harmony with the traditional picture that elementary systems are described by irreducible representations of the kinematical group G .

Systems of imprimitivity induced from representation of the little group H . Given a unitary ray representation $L : h \mapsto L_h$ of the closed subgroup H on some Hilbert space \mathcal{H}_0 one can construct a transitive system of imprimitivity based on G/H in the following way.

Let μ be a measure on G/H which is quasi-invariant with respect to translations from the left. Let \mathcal{H}_L be the Hilbert space of functions f on G with values in \mathcal{H}_0 satisfying

1. $g \mapsto \langle f(g), \psi \rangle$ is a measurable function for all $\psi \in \mathcal{H}_0$,
2. $f(gh) = L_h^{-1} f(g)$,
3. $\int_{G/H} \|f(g)\|^2 d\mu(gH) < \infty$.

Because of requirement (2) and the unitarity of L_h we have $\|f(gh)\| = \|L_h^{-1} f(g)\| = \|f(g)\|$. Therefore the quantity $\|f(g)\|$ is constant on all left cosets. In this way requirement (3) makes sense. The scalar product on \mathcal{H}_L is defined by $\langle f, f' \rangle := \int_{G/H} \langle f(g), f'(g) \rangle d\mu(gH)$.

Now define a representation U^L of G on \mathcal{H}_L . Denote by $\rho_g(s) := d\mu(g^{-1}sH)/d\mu(sH)$ the Radon-Nikodym derivative of the translated measure $\mu_{g^{-1}}$ with respect to μ . Then define a unitary representation U^L of G on \mathcal{H}_L by

$$U_g^L f(s) := \sqrt{\rho_g(s)} f(g^{-1}s), \quad g, s \in G, f \in \mathcal{H}_L. \quad (2.3)$$

U^L is called the representation of G on \mathcal{H}_L *induced* by the representation L of H . The Radon-Nikodym factor $\rho_g(s)$ compensates for the non-invariance of the measure μ in such a way as to make U_g^L a unitary

operator. Note that changing μ to another measure in the same class does not change the equivalence class of the representation U^L .

Define a projection-valued measure E^L on G/H by

$$(E^L(\Delta)f)(s) := \chi_\Delta(sH)f(s), \quad g, s \in G, f \in \mathcal{H}_L, \Delta \in \Sigma(G/H). \quad (2.4)$$

(E^L, U^L) satisfies the covariance condition

$$(U_g^L E^L(\Delta) U_{g^{-1}}^L f)(s) = E^L(g^{-1}\Delta)f(s), \quad s \in G$$

and therefore is a transitive system of imprimitivity based on G/H . It is called the *system of imprimitivity induced from the representation L* .

2.2.2 The commutation relations

To see why configuration space systems of imprimitivity describe quantum systems in the usual sense, I shall briefly explain how Mackey (1949, and 1978, ch. 18) derives the commutation relations from the covariance condition of Euclidean systems of imprimitivity.

To take the simplest example, consider a particle moving freely in three dimensional Euclidean space $\mathbb{R}_{\text{pos}}^3$. As symmetry group of Euclidean space take the Euclidean group \mathcal{E} , and write $\mathbb{R}_{\text{pos}}^3 = \mathcal{E}/\mathcal{R}$. Making this assumption of Euclidean symmetry of configuration space means that one assumes physical space to be homogeneous and isotropic. Let us choose a rectangular coordinate system $\{x, y, z\}$ in $\mathbb{R}_{\text{pos}}^3$. According to Mackey's (1949) imprimitivity theorem (see below, p. 83), every irreducible system of imprimitivity is unitarily equivalent to a system of imprimitivity (E^L, U^L) induced from some representation L of the rotation group \mathcal{R} . So when considering a system of imprimitivity we may take it to be (E^L, U^L) and drop the superscripts L . Applying equation

(2.3) one defines on \mathcal{H}_L (now called \mathcal{H}) a unitary representation U of the additive group \mathbb{R} by

$$(U_{a,x}f)(x, y, z) := f(x - a, y, z).$$

The Radon-Nikodym factor $\rho_g(s)^{1/2}$ equals one because the Lebesgue measure is translation invariant. Mackey (1958) showed that because of *Euclidean* symmetry of E (not just translation symmetry), every representation of the Euclidean group, when restricted to the translation group, is unitarily equivalent to an ordinary (i.e. not projective) unitary representation of the translation group. Therefore the phase factors of $U_{a,x}$ above can be chosen in such a way that $\{U_{a,x} : a \in \mathbb{R}\}$ forms a continuous ordinary representation of the additive group \mathbb{R} . According to Stone's theorem such a group uniquely determines a self-adjoint operator P_x such that in the weak sense we have

$$U_{a,x} = \exp(-iaP_x).$$

By equation (2.4) one defines a projection valued measure E_x on \mathbb{R} , which in turn defines a self-adjoint operator Q_x on \mathcal{H} by

$$Q_x := \int_{\mathbb{R}} \lambda E_x(d\lambda).$$

Q_x generates a strongly continuous unitary one parameter group $V_{b,x}$ by

$$V_{b,x} := \exp(ibQ_x) = \int_{\mathbb{R}} \exp(ib\lambda) E_x(d\lambda).$$

This equation again holds in the weak sense. Omitting the index x , from the covariance condition we obtain

$$\begin{aligned} U_a V_b U_a^{-1} &= \int_{\mathbb{R}} \exp(ib\lambda) U_a E(d\lambda) U_a^{-1} \\ &= \int_{\mathbb{R}} \exp(ib\lambda) E(d\lambda - a) \end{aligned}$$

$$\begin{aligned}
&= \int_{\mathbb{R}} \exp(ib(\lambda + a))E(d\lambda) \\
&= \exp(iab)V_b.
\end{aligned}$$

These are the Weyl commutation relations.²⁰

Thus Euclidean invariance of (E, U) implies that the Heisenberg commutation relations are fulfilled by $Q := \int_{\mathbb{R}} \lambda E(d\lambda)$ and the self-adjoint generator P of U . Furthermore, the assumption of Euclidean symmetry implies the commutation relations for angular momentum. Different irreducible Euclidean systems of imprimitivity are induced by different representations of the rotation group and correspond to different spins.²¹

2.2.3 Classification of configuration space quantisations.

A system of imprimitivity on configuration space Q consists of a Hilbert space representation E of the Boolean lattice \mathcal{L}_Q and a representation U of the action of G on $Q = G/H$. Inequivalent representations (E, U) describe inequivalent quantisations of the classical systems with the same configuration space.²² Therefore one can classify the configuration

²⁰The self-adjoint generator P_x is defined on all wave functions $f \in \mathcal{H}$ for which the limit $\lim_{a \rightarrow 0} \frac{1}{ia}(U_{a,x} - \mathbf{1})f =: P_x f$ exists. The domain of the self-adjoint generator Q_x is defined in a similar way. From these definitions and the Weyl commutation relations one can obtain the Heisenberg form of the commutation relations

$$[Q_x, P_x]f = if.$$

This equation holds for all f in the intersection of the domains of P_x and Q_x . The intersection is dense and the restrictions of P_x and Q_x to the intersection are essentially self-adjoint.

²¹For a detailed description of spin by Euclidean systems of covariance, see Mackey 1978.

²²As pointed out by Robson (1994a,b), these systems do *not* have the same phase space. Rather they correspond to systems with different constraints. I will report on this later.

space quantisations by a classification of systems of imprimitivity.

This is achieved by Mackey's (1949) imprimitivity theorem: Let G be locally compact second countable²³ group and H a closed subgroup of G . Let (E, U) be a transitive system of imprimitivity based on G/H . Then there exists a unitary representation L of H on a Hilbert space \mathcal{H}_0 such that U is unitarily equivalent to the representation U^L of G on \mathcal{H}_L induced by L , and E is unitarily equivalent to the PV-measure E^L . The equivalence class of the representation L is uniquely determined by (E, U) . So there is a one-to-one correspondence between the set of all equivalence classes of systems of imprimitivity based on G/H and equivalence classes of unitary representations of the little group H . (E, U) is irreducible or a direct sum of irreducible systems of imprimitivity, if and only if L is.

Consider a system of imprimitivity based on a given configuration space $Q = G/H$. It follows from Mackey's imprimitivity theorem that these different configuration space systems of imprimitivity are in a one-to-one relation to the different representations of the little group H . Traditionally these have been regarded as inequivalent representations of the same classical system with phase space T^*Q . Recent work by Landsman (1993) and Robson (1994a,b), however, stresses that these are quantisations of classical systems with the same configuration space but with different phase spaces.

Superselection rules associated with inequivalent quantisations. Consider a Galilean elementary particle moving without constraints in \mathbb{R}^3 . The unitary ray representations of the Galilei group are in one-to-one relation with the unitary vector representation of the central exten-

²³The assumption that G is second countable could actually be dropped. Without this assumption the imprimitivity theorem was proved by Blattner (1961, 1962).

sion of the covering group of the Galilei group by some abelian group (Bargmann 1954). Take as symmetry group G this extended Galilei group. As algebra of kinematical observables $\mathcal{M} := \{U(g) : g \in G\}$ take the von Neumann algebra generated by the infinitesimal generators of the representation U of G . Bargmann (1954) pointed out that the generator M of the extension commutes with all other generators of the extended Galilei group and can be regarded as mass operator. Apart from the mass operator M , there is only one additional central element in \mathcal{M} .²⁴ This is $S^2 := (J + Q \times P)^2$, where J is the vector of the three generators of the rotation group, Q the generator of the boosts and P the generator of the translations. If U is irreducible, the central elements of \mathcal{M} are constant multiples of the identity operator. Denote the value of the mass operator by m , and the value of S^2 by $s(s + 1)$ where $s = 0, 1/2, 1, 3/2, \dots$. The numbers m, s label the irreducible representations of G . In all pure states of a given representation M and S^2 have a dispersion free value.

This is just one example of a more general phenomenon. Inequivalent systems of imprimitivity, or in general inequivalent quantisations, lead to superselection rules in the following sense: there are quantities which have a dispersion free value in all pure states.

Whether inequivalent quantisations are described by a superselection rule or by a parameter is, to a certain extent, a matter of taste. Consider, for example, an elementary Galilei system. If mass is regarded as a fixed parameter, one uses the corresponding representation of the Galilei group and works with an irreducible Hilbert space. Alternatively, one can regard mass as an observable which can take different values. Then one works on a direct sum Hilbert space, i. e. simultane-

²⁴See Bargman 1954, Lévy-Leblond (1963, 1971).

ously with all the possible representations of the group.

The description of inequivalent quantisations by superselection rules will soon be related to the usual picture of superselection rules as arising from different inequivalent representations of an abstract algebra of observables.

2.2.4 Algebraic description of systems of imprimitivity

It is also possible to give an operator algebraic description of systems of imprimitivity. I will outline how this is done, following Landsman (1990a), who elaborated results of Glimm (1962).

As a first step, the commutative lattice $\mathcal{L}_Q = \Sigma(G/H)$ and the action $\Delta \mapsto g\Delta$ is replaced by a C^* -dynamical system. Denote by $C_0(G/H)$ the C^* -algebra of continuous complex-valued functions on G/H vanishing at infinity. The action $\Delta \mapsto g\Delta$ naturally carries over to an action $\text{Ad}\lambda$ of G on $C_0(G/H)$ by

$$(\text{Ad}\lambda(g)f)(sH) := f(g^{-1}sH).$$

$(C_0(G/H), G, \text{Ad}\lambda)$ forms a classical C^* -system in the sense of Definition 4. It describes the action of G on the configuration space.

Then, in a second step, one can define²⁵ the C^* -algebra $C^*(G, G/H)$ as the C^* -crossed product of $C_0(G/H)$ and G with respect to the action $\text{Ad}\lambda$. $C^*(G, G/H)$ is called the transformation group algebra of the classical system. Note that there is by no means a one-to-one correspondence between $C^*(G, G/H)$ and the pair $G, G/H$: Whenever the duals of H_1 and H_2 are homeomorphic, $C^*(G_1, G_1/H_1)$ is isomorphic to $C^*(G_2, G_2/H_2)$.

²⁵See e.g. Bratteli and Robinson (1979), chapter 1.3.6). In the literature $C^*(G, G/H)$ is sometimes denoted by $C_0(G/H) \otimes_{\text{Ad}\lambda} G$.

$C^*(G, G/H)$ is already the abstract algebra of quantum mechanical observables. The operators in $C^*(G, G/H) = C_0(G/H) \otimes_{\text{Ad}\lambda} G$ representing $C_0(G/H)$ describe the quantum mechanical localisation properties, which are the same as the classical ones. The representation of $C_0(G/H)$ in $C^*(G, G/H)$ plays the same rôle as the PV-measure E . The operators in $C^*(G, G/H)$ representing the action $\text{Ad}\lambda$ describe the quantum mechanical generators of translations. They play the same rôle as the representation $U(G)$ in the traditional configuration space system of imprimitivity (E, U) . $C^*(G, G/H)$ can be regarded as abstract C^* -algebra of kinematically determined quantum mechanical observables.

To recover the familiar Hilbert space description of systems of imprimitivity, we take, in a third step, an irreducible Hilbert space representation of the C^* -algebra $C^*(G, G/H)$. The important point now, proved by Glimm (1962), is that the non-degenerate (but possibly non-faithful) covariant representations²⁶ of the classical C^* -system $(C_0(G/H), G, \text{Ad}\lambda)$ are in one-to-one correspondence with the non-degenerate representations of the C^* -algebra $C^*(G, G/H)$. Thus this third step amounts to choosing a covariant Hilbert space representation of the classical C^* -system $(C_0(G/H), G, \text{Ad}\lambda)$.

Non-degenerate covariant representations (π, U) of $(C_0(G/H), G, \text{Ad}\lambda)$ are naturally related to systems of imprimitivity. The von Neumann algebras $\pi(C_0(G/H))''$ and $\{E(\Delta) : \Delta \in \Sigma(G/H)\}''$ are both isomorphic to $L^\infty(G/H)$. Therefore the W^* -systems $(\{E(\Delta) : \Delta \in \Sigma(G/H)\}'', G, U \cdot U^*)$ and $(\pi(C_0(G/H))'', G, U \cdot U^*)$ obtained from the covariant representation (π, U) and from the system of imprimitivity (E, U) are isomorphic.

²⁶A covariant representation (π, U) of the classical C^* -system $(C_0(G/H), G, \text{Ad}\lambda)$ is given by a representation π of the C^* -algebra $C_0(G/H)$ on a Hilbert space \mathcal{H} and a unitary representation U of G on \mathcal{H} such that $U_g \pi(f) U_g^* = \pi(\text{Ad}\lambda(g)f)$.

The non-degenerate covariant representations of the classical system $(C_0(G/H), G, \text{Ad}\lambda)$ are in one-to-one correspondence with the non-degenerate representations of the C^* -algebra $C^*(G, G/H)$. Therefore the inequivalent configuration space quantisations, which are in one-to-one relation with the inequivalent representations of the little group H , are also in one-to-one relation to the inequivalent non-degenerate representations of the abstract C^* -algebra $C^*(G, G/H)$. The superselection rules induced by inequivalent quantisations can then be regarded as arising from inequivalent representation of this abstract C^* -algebra of kinematic observables. This is in harmony with the usual picture of superselection rules as arising from inequivalent representation of abstract C^* -algebras of observables.

2.3 Phase space quantisation of the Weyl system

In the previous section we considered systems of imprimitivity (E, U) based on configuration space. The PV-measure E described the quantum mechanical position observable, whereas the quantum mechanical momentum observable emerged as self-adjoint generator of the unitary representation U . So position and momentum were quantised in strikingly different ways.

Phase space quantisation tries to treat position and momentum on equal footing. In order to achieve this, the configuration space quantisation scheme can be extended in two different directions. Firstly, the quantisation of position *and* momentum can be described by a unitary representation of some appropriate group. This approach is taken by the geometric quantisation procedure or by Isham's (1983) quantisation

scheme. The aim of Isham's quantisation scheme is to find a Lie group and to associate to classical observables as quantum observables the self-adjoint generators of an irreducible unitary representation of the Lie group. Secondly, the quantisation of position *and* momentum can be described by a POV-measure on phase space. This is the approach I will take in this chapter.

2.3.1 Phase space quantisation by systems of covariance

Recall how systems of imprimitivity are used in configuration space quantisation: the configuration space PV-measure represents the lattice \mathcal{L}_Q of Borel sets in the configuration space by a Boolean sublattice of the projection lattice $\mathcal{P}(\mathcal{H})$ of some Hilbert space. This lattice morphism was described by a PV-measure E on configuration space fulfilling

$$E(\Delta_1)E(\Delta_2) = E(\Delta_1 \cap \Delta_2) = E(\Delta_2)E(\Delta_1). \quad (2.5)$$

In view of the fact that it is only the phase space Γ which carries a full description of the system, it would be desirable to regard quantisation as a map from the Boolean lattice \mathcal{L}_Γ of Borel subsets of the phase space into the lattice $\mathcal{P}(\mathcal{H})$. But quantum mechanical position and momentum observables do not commute. Therefore a measure representing localisation on the phase space cannot fulfil equation (2.5). Thus the quantisation map cannot be a projection-valued measure on phase space. The impossibility of sharp localisation on phase space is consequence of the non-commutativity of quantum mechanical position and momentum operators, or of the uncertainty relations.

Having to drop requirement (2.5), a measure representing phase space localisation can at most be a POV-measure as used by Ludwig

(1967, 1968, 1972) and Davies and Lewis (1970).

Definition 7 A POV-measure on a Borel space $(X, \Sigma(X))$ is a mapping a from $\Sigma(X)$ into $\mathcal{B}(\mathcal{H})$ such that

(i) $a(X) = \mathbf{1}$,

(ii) $\mathbf{0} \leq a(\Delta) \leq \mathbf{1}$ for all $\Delta \in \Sigma(X)$, and

(iii) for each family $\{\Delta_i\}$ of mutually disjoint Borel sets of X we have $a(\bigcup_{i=1}^{\infty} \Delta_i) = \sum_{i=1}^{\infty} a(\Delta_i)$, where the right hand side converges weakly.

A POV-measure on the phase space describes a mapping from the Boolean lattice \mathcal{L}_Γ into the partially ordered set of effect operators²⁷ on \mathcal{H} . This mapping is of course not a lattice morphism. How a phase space POV-measure can be characterised lattice-theoretically is still an open question.

Systems of covariance. It was argued in section 2.1.1 that it is sometimes justified to assume that the phase space Γ is a homogeneous space, which I will denote it again by G/H . (Of course G and H are not the groups denoted by G and H when I referred to a homogeneous configuration space G/H .) Assume also that a unitary ray representation U of G on the Hilbert space is given.

Since G is a symmetry of the phase space, the action of G on the phase space and of $U(G) \cdot U(G)^*$ on the phase space observables should not have any observable consequences. Thus we have to require commutativity of the diagram

²⁷Effect operators are self-adjoint operators whose spectrum is contained in the interval $[0, 1]$.

$$\begin{array}{ccc}
\Delta & \xrightarrow{g} & g\Delta \\
\downarrow a & & \downarrow a \\
a(\Delta) & \xrightarrow{U_g \cdot U_g^*} & U_g a(\Delta) U_g^*
\end{array}
.$$

The commutativity requirement implies the covariance condition

$$U_g a(\Delta) U_g^* = a(g\Delta), \quad \Delta \in \Sigma(\Gamma), g \in G. \quad (2.6)$$

Definition 8 *A POV-measure a on a homogeneous G -space X , together with a weakly measurable unitary ray representation U is called a transitive system of covariance (a, U) based on X if the covariance condition (2.6) is satisfied.*

This definition of systems of covariance is the traditional one.²⁸ It allows for both a description of covariant phase space localisation and of unsharp configuration space localisation of quantum systems. In the latter case the POV-measure a is commutative. In this chapter I will use system of covariance for phase space localisation.

Let us see how a covariant POV-measure a on phase space Γ can be used for quantisation. The von Neumann algebra generated by the characteristic functions $\{\chi_\Delta : \Delta \in \Sigma(\Gamma)\}$ is $L^\infty(\Gamma)$. Then by the functional calculus, the map $\Delta \mapsto a(\Delta)$ can be extended σ -weak continuously to a map O from $L^\infty(\Gamma, \mu)$ into $\{a(\Delta) : \Delta \in \Sigma(\Gamma)\}''$ satisfying

$$O(\chi_\Delta) = a(\Delta).$$

If (a, U) is a system of covariance then O is a covariant embedding in the following sense.

²⁸See e.g. the review of Ali (1985).

Definition 9 Let X be a Borel space on which G acts transitively, μ a quasiinvariant measure on X , and U be a weakly measurable unitary ray representation of G on some separable Hilbert space \mathcal{H} . A covariant embedding of $L^\infty(X, \mu)$ into $\mathcal{B}(\mathcal{H})$ is a map $O : L^\infty(X, \mu) \rightarrow \mathcal{B}(\mathcal{H})$ satisfying

- (i) *linearity*: $O(cf + g) = cO(f) + O(g)$ for $f, g \in L^\infty(X, \mu), c \in \mathbf{C}$,
- (ii) *covariance*: $O(\text{Ad}\lambda(g)f) = \alpha_g(O(f))$ for all $f \in L^\infty(X, \mu)$,
- (iii) *positivity*: $O(f) \geq 0$ for all $f \in L^\infty(X, \mu)$ with $f(x) \geq 0$ a.e.,
- (iv) $O(f)$ is self-adjoint if f is real
- (v) *normalisation*: $O(1) = \mathbf{1}$.

Note that positivity of O only requires that $O(f)$ is a positive operator if f is a positive function. It does not mean that every positive operator in $O(L^\infty(X, \mu))$ is the image of a positive function. In general there are positive operators in $O(L^\infty(X, \mu))$ which are not the image of positive functions. The term ‘embedding’ might be somewhat misleading, since O preserves only the linear, but not the multiplicative structure of $L^\infty(X, \mu)$.

Now we are going to see how covariant embeddings arising from phase space systems of covariance are quantisation maps.

2.3.2 Phase space representation of the Weyl system

Because in this section I merely want to explain the main ideas of quantisation by phase space systems of covariance, I will restrict attention to the case of a non-relativistic spin zero particle. The theory has been developed more generally by Ali (1985) for traditional systems of covariance on homogeneous phase spaces. In a different direction, namely for the systems of covariance defined with respect to von Neumann al-

gebras, the theory was generalised by Amann (1986a). Phase space quantisation by systems of covariance, as presented here, is also known under the name of coherent state quantisation.²⁹

Since we are dealing with a spinless particle, it is sufficient to take as symmetry group G not the full Galilei group but just the space translations and the velocity boosts. So we specify as G the additive group $\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$ corresponding to the phase space. The Haar measure on $\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$ is the Lebesgue measure μ . Let $\{U_a, V_b : a, b \in \mathbb{R}^3\}$ be a weakly measurable irreducible representation of the Weyl relations

$$U_a V_b = \exp(iab) V_b U_a$$

on the Hilbert space $L^2(\mathbb{R}^3)$ (ab denotes the scalar product of the vectors a and $b \in \mathbb{R}^3$). Consider the action α of $\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$ on $\mathcal{B}(L^2(\mathbb{R}^3))$ defined by

$$\alpha_{(a,b)}(x) := U_a V_b x V_b^* U_a^*, \quad x \in \mathcal{B}(L^2(\mathbb{R}^3)), (a, b) \in \mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3. \quad (2.7)$$

This action is pointwise σ -continuous, since the representation of the Weyl relations is weakly measurable. Due to the irreducibility of the representation of the Weyl relations, only multiples of the identity fulfil $\alpha_{(a,b)}(x) = x$ for all a, b .

Now observe that for each character χ on $\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$, $\chi(a, b) = \exp(iat + isb)$, there is a unitary operator $u_\chi \in \mathcal{B}(\mathcal{H}_1)$, namely $u_\chi = U_t V_s$, such that u_χ transforms under α according to χ :

$$\alpha_{(a,b)}(u_\chi) = \chi(a, b) u_\chi.$$

²⁹For more information on the relation to coherent states, see Ali 1985 and Amann 1986b.

Since only multiples of the identity are fixed under all $\alpha_{(a,b)}$, it follows³⁰ that the action α on $\mathcal{B}(L^2(\mathbb{R}^3))$ is integrable.³¹

By definition of integrability (see Definition 16 on p. 130) there is some atomic projection

$$P_e = |e\rangle\langle e|, \quad e \in L^2(\mathbb{R}^3)$$

such that the integral $\int_{\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3} \alpha_{(a,b)}(P_e) d\mu(a,b)$ exists in the weak sense and is in $\mathcal{B}(L^2(\mathbb{R}^3))$. Since the Lebesgue measure μ is translation invariant, the integral is also invariant under all phase space translations:

$$\alpha_{a',b'} \left(\int_{\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3} \alpha_{(a,b)}(P_e) d\mu(a,b) \right) = \int_{\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3} \alpha_{(a,b)}(P_e) d\mu(a,b),$$

for all $(a', b') \in \mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$. This implies that the integral is a scalar multiple of the identity operator. By multiplying P or e with an appropriate complex number we can achieve

$$\int_{\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3} \alpha_{(a,b)}(P_e) d\mu(a,b) = \mathbf{1}. \quad (2.8)$$

Now we can define a covariant POV-measure a_e on $\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$ by

$$a_e(\Delta) := \int_{\Delta} \alpha_{(a,b)}(P_e) d\mu(a,b) = \int_{\Delta} |U_a V_b e\rangle\langle U_a V_b e| d\mu(a,b), \quad (2.9)$$

where $\Delta \in \Sigma(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)$. The corresponding quantisation map O_e is given by

$$O_e(f) := \int_{\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3} \alpha_{(a,b)}(P_e) f(a,b) d\mu(a,b), \quad (2.10)$$

where $f \in L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)$. In the sense of Definition 9, O_e is a covariant embedding of $L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)$ into $\mathcal{B}(L^2(\mathbb{R}^3))$.

³⁰See de Schreye (1983).

³¹Integrability of a W^* -system will be introduced in Definition 16.

Phase space representation. The quantisation map O_e induces a dual map $O_e^* : \mathcal{B}(L^2(\mathbb{R}^3))^* \rightarrow L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)^*$ defined by

$$O_e^* : \rho \mapsto \rho \circ O_e.$$

Thus we have $(O_e^*(\rho))(f) = \rho(O_e(f))$ for $f \in L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)$. Assume that O_e maps $\mathcal{B}(L^2(\mathbb{R}^3))_* \subset \mathcal{B}(L^2(\mathbb{R}^3))^*$ into $L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)_* \subset$

$L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)^*$. Since $L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)_* \cong L^1(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)$, for every normal state $\rho \in \mathcal{B}(L^2(\mathbb{R}^3))_*$ there is a μ -integrable function $O_*^e(\rho)$ such that

$$\rho(O_e(f)) = O_e^*(\rho)(f) = \int_{\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3} O_*^e(\rho)(a, b) f(a, b) d\mu(a, b) \quad (2.11)$$

for any $f \in L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)$. $O_*^e(\rho)$ is given by

$$O_*^e(\rho)(q, p) = \text{tr} \left(\rho \alpha_{(q,p)}(P_e) \right), \quad (q, p) \in \mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3 \quad (2.12)$$

where, by a slight abuse of notation, the same symbol ρ refers on the right hand side to a density matrix and on the left hand side to the linear functional $\text{tr}(\rho \cdot)$.

O_*^e associates to every density matrix ρ an L^1 -function $O_*^e(\rho)$ on the phase space such that (2.11) is fulfilled. Since O_e maps positive functions on $\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$ into positive operators, O_*^e maps positive functionals in $\mathcal{B}(L^2(\mathbb{R}^3))_*$ into positive L^1 -functions. So the phase space probability distribution $O_*^e(\rho)$ is positive almost everywhere. Since states are normalised (i.e. $\rho(\mathbf{1}) = 1$), we have

$$\int_{\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3} O_*^e(\rho)(a, b) d\mu(a, b) = 1.$$

Therefore $O_*^e(\rho)$ is a probability density on the phase space.³² As we will see later (eqs. (2.13), (2.14)), it can be regarded as smeared Wigner function.

The plurality of quantisation maps. Different choices of e lead to different quantisation maps O_e .

Many choices are possible: Since the action α on $\mathcal{B}(L^2(\mathbb{R}^3))$ is integrable, the set of $x \in \mathcal{B}(L^2(\mathbb{R}^3))$ for which the integral $\int_{\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3} \alpha_{(a,b)}(x) d\mu(a,b)$ exists and is finite, is very large indeed.³³ Since the Lebesgue measure μ is translation-invariant, for every x for which the integral exists and is finite, the integral is fixed under all $\alpha_{(a,b)}$. This implies that the integral is a multiple of the identity. Multiplying x with an appropriate complex number one arrives at operators x' which fulfil $\int \alpha_{(a,b)}(x') d\mu(a,b) = \mathbf{1}$. All these operators define covariant embeddings $O_{x'}$ by $O_{x'}(f) := \int \alpha_{(a,b)}(x') f(a,b) d\mu(a,b)$.

One has to impose an additional condition which guarantees that the range $O(L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3))$ of the quantisation map exploits $\mathcal{B}(L^2(\mathbb{R}^3))$ as far as possible. Following Amann (1986a), we require quantisation maps to be maximal in the following sense.

³²In the terminology of Araki (1960), $O_e^*(\rho)$ is the *state generating functional* of the state ρ . Araki (1960) showed that every complex valued function f on \mathbb{R}^6 satisfying

- (i) $f(0) = 1$,
- (ii) for all $(p, q) \in \mathbb{R}^6$, the map $\mathbb{R} \rightarrow \mathbf{C}$ defined by $r \mapsto f(rp, rq)$ is continuous,
- (iii) f is positive semidefinite in the sense that for all natural numbers N , for all $z_k \in \mathbf{C}$, for all $(p_k, q_k) \in \mathbb{R}^6$ ($k=1, 2, \dots, N$), we have

$$\sum_{j,k=1}^N z_j^* z_k^* f(p_k - p_j, q_k - q_j) e^{-\frac{2\pi i}{\hbar}(p_j q_k - p_k q_j)} \geq 0,$$

is the state generating function of some state ρ , $f = O^*(\rho)$.

³³There exists a set n_α which is σ -weakly dense in \mathcal{M} such that the integral exists and is finite for arbitrary x of the form $x = y_1^* y_2$ with $y_1, y_2 \in n_\alpha$.

Definition 10 A covariant embedding O of $L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)$ into $\mathcal{B}(L^2(\mathbb{R}^3))$ is called maximal if for every covariant embedding O' ,
 $O(L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)_+) \subseteq O'(L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)_+)$ implies
 $O(L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)_+) = O'(L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)_+)$.

If x' is a multiple of an atomic projection, then the quantisation map $O_{x'}$ is maximal in the sense of Definition 10.³⁴ (It is not clear whether the converse holds, namely that every maximal normal embedding O is of the form $O_{x'}$ for some x' which is a multiple of an atomic projection.) An atomic projection x' can be written as $x' = |e\rangle\langle e|$ for some vector $e \in L^2(\mathbb{R}^3)$. In this case we write O_e instead of $O_{x'}$. So by taking quantisation maps of the form O_e we are sure that the range $O_e(L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)_+)$ is maximal.

If O' is any embedding with $O'(L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)_+) = O_e(L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)_+)$, then O' is of the form

$$O'(f) = O_{U_{a_0}V_{b_0}e}(f) = O_{\alpha_{(a_0, b_0)}(P_e)}(f), \quad f \in L^\infty(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)$$

for some suitable $(a_0, b_0) \in \mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$.³⁵ $O_{U_{a_0}V_{b_0}e}$ is the covariant embedding associated with the translated function $U_{a_0}V_{b_0}e$.

Interpretation of e in O_e . The map O_e defines operators $O_e(q), O_e(p)$ corresponding to classical position and momentum. Since O_e was defined (in equations (2.9) and (2.7)) via a representation of the Weyl commutation relations, the operators $O_e(q), O_e(p)$ fulfil the Heisenberg commutation relations. The choice of e fixes how the quantisation map extends from position and momentum to all functions on phase space. Therefore, choosing e amounts to a choice of operator ordering.³⁶

³⁴See Amann 1986b, Theorem 1.

³⁵See Amann (1986b), Theorem 1.

³⁶See e.g. Ali and Doebner (1990).

In the context of coherent state quantisation, the choice of e is the choice of a fiducial vector determining a family $\{U_a V_b e : (a, b) \in \mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3\}$ of coherent states.³⁷ Equation (2.8) is then the completeness property for a family of coherent states. Coherent states in this generalised sense were introduced by Perelomov (1972). Often the choice of e is related to the Hamiltonian. (For example, the usual e in the coherent state approaches is the ground state of the harmonic oscillator. A family of coherent states is then given by $\{\exp(\lambda a^* - \lambda^* a)e : \lambda \in \mathbf{C}\}$, where a is the destruction operator and a^* is the creation operator. These states minimize the product of the dispersions in position and momentum.) In the context of phase space quantisation the Hamiltonian does not play any rôle in the choice of e . I will discuss the relevance of e on p. 101.

The marginals of the phase space probability density satisfy the uncertainty relations. Ali (1985, his eq. 4.52) showed that the phase space probability density $O_*^e(q, p)$ satisfies the marginality conditions

$$\begin{aligned} \int_{\mathbb{R}_{\text{mom}}^3} O_*^e(q, p) dp &= \int_{\mathbb{R}_{\text{pos}}^3} |e(q' - q)|^2 \langle q' | \rho | q' \rangle dq', \\ \int_{\mathbb{R}_{\text{pos}}^3} O_*^e(q, p) dq &= \int_{\mathbb{R}_{\text{mom}}^3} |\hat{e}(p' - p)|^2 \langle p' | \rho | p' \rangle dp', \end{aligned} \quad (2.13)$$

where $q, q' \in \mathbb{R}_{\text{pos}}^3$, $p, p' \in \mathbb{R}_{\text{mom}}^3$. \hat{e} denotes the Fourier transform of the function e , $\langle q' | \rho | q' \rangle$ is the position representation of the density matrix, and $\langle p' | \rho | p' \rangle$ is its momentum representation.

Define the standard deviations of the marginals (2.13) by

$$\sigma(q_k) := \sqrt{\int_{\mathbb{R}_{\text{pos}}^3} q_k^2 |e(q)|^2 dq - \left(\int_{\mathbb{R}_{\text{pos}}^3} q_k |e(q)|^2 dq \right)^2}, \quad k = 1, 2, 3$$

³⁷See Klauder and Skagerstam 1984, p. 21.

$$\sigma(p_k) := \sqrt{\int_{\mathbb{R}_{\text{mom}}^3} p_k^2 |\hat{e}(p)|^2 dp - \left(\int_{\mathbb{R}_{\text{mom}}^3} p_k |\hat{e}(p)|^2 dp \right)^2}. \quad k = 1, 2, 3$$

They satisfy the uncertainty relations

$$\sigma(q_k)\sigma(p_i) \geq \frac{\hbar}{2}\delta_{ik}.$$

It might seem surprising that the phase space probability distribution O_e^* is positive. After all, Wigner's (1932) phase space distribution function

$$w_\rho(q, p) := \int_{\mathbb{R}_{\text{pos}}^3} \exp(-ip \cdot x) \langle q + \frac{x}{2} | \rho | q - \frac{x}{2} \rangle dx, \quad p \in \mathbb{R}_{\text{mom}}^3, q, x \in \mathbb{R}_{\text{pos}}^3$$

is not positive. The marginals of the Wigner distribution satisfy

$$\begin{aligned} \int_{\mathbb{R}_{\text{mom}}^3} w_\rho(q, p) dp &= \langle q | \rho | q \rangle \\ \int_{\mathbb{R}_{\text{pos}}^3} w_\rho(q, p) dq &= \langle p | \rho | p \rangle. \end{aligned} \quad (2.14)$$

$\langle q | \rho | q \rangle$ is the probability density for finding the particle, if it in the state ρ , sharply localised at the point q . Similarly, $\langle p | \rho | p \rangle$ is the probability of the particle having sharp momentum p . Therefore equation (2.14) suggest that $w_\rho(q, p)$ should be interpreted as a probability density for finding the particle at the phase space point (q, p) . However, $w_\rho(q, p)$ can assume also negative values. So it is not possible to interpret it as a probability distribution. This situation reflects the fact that it is impossible for a quantum mechanical particle to be sharply localised at q and at the same time to have sharp momentum p . It was proved by Wigner (1971) that it is impossible to find a linear map $\rho \mapsto w_\rho$ such that all w_ρ satisfies the marginality conditions (2.14) and are positive.

The marginals (2.13) of the probability density $O_*^e(\rho)(p, q)$ are obtained from the marginals (2.14) of the Wigner distribution by smearing with the functions $|e|^2, |\hat{e}|^2$. Since the smearing in position and

momentum space is by (squares of the moduli of) functions which are Fourier transforms of each other, the smeared marginals satisfy the uncertainty relations. Because the marginals of $O_*^e(\rho)(p, q)$ fail to fulfil the sharp marginal conditions (2.14), Wigners no-go theorem cannot prevent $O_*^e(\rho)$ from being positive.

Interpretation of the marginals. Since $O_*^e(q, p)$ is positive and normalised, it is a probability density. What is the interpretation of $O_*^e(\rho)$ and its marginals?

Recall why it would have been desirable to interpret the Wigner function w_ρ as a probability distribution: $\langle q|\rho|q\rangle$ is the probability for finding the particle, if it is in the state ρ , sharply localised at the point q . Similarly, $\langle p|\rho|p\rangle$ is the probability of the particle having sharp momentum p . Therefore equations (2.14) suggest that $w_\rho(q, p)$ should be interpreted as a probability density for finding the particle at the phase space point (q, p) . So the probability interpretation of w_ρ would have been motivated because the marginals of w_ρ can be interpreted as probabilities of localisation in position or in momentum space.

In the same spirit, one might be tempted to base a probability interpretation of $O_*^e(\rho)$ on a probability interpretation of its marginals (2.13). But it is difficult to give a probability interpretation to the smeared marginals (2.13) since the functions $|e|^2, |\hat{e}|^2$ by which the sharp marginals (2.14) are smeared are *independent of the state* ρ . The obvious way out is to base a probability interpretation of $O_*^e(\rho)$ not on a probability interpretation of its marginals, but to proceed in the opposite direction: A probability interpretation of the marginals will be based on a probability interpretation of the joint distribution $O_*^e(\rho)$.

$O_*^e(\rho)$ can be given a straightforward probability interpretation in

terms of phase space localisation: (2.11) can be rewritten as

$$\mathrm{tr}(\rho a_e(\Delta)) = \int_{\Delta} (O_*^e(\rho))(q, p) d\mu(q, p), \quad \Delta \in \Sigma(\mathbb{R}_{\mathrm{pos}}^3 \times \mathbb{R}_{\mathrm{mom}}^3). \quad (2.15)$$

a_e is the covariant POV-measure on $\mathbb{R}_{\mathrm{pos}}^3 \times \mathbb{R}_{\mathrm{mom}}^3$ corresponding to O_e by $a_e(\Delta) = O_e(\chi_{\Delta})$. It is a covariant POV-measure describing phase space localisation: By (2.15), $\mathrm{tr}(\rho a_e(\Delta))$ is the probability to find the expectation values of position and momentum of the system, if it is in the state ρ , in the phase space region Δ .

The probability interpretation of $O_*^e(\rho)$ can be carried over to its marginals (2.13) in the following way: The phase space probability measure a_e gives rise to “marginal” POV-measures a_e^Q, a_e^P on position and on momentum space by

$$\begin{aligned} a_e^Q(\Delta) &:= a_e(\Delta \times \mathbb{R}_{\mathrm{mom}}^3), & \Delta \in \Sigma(\mathbb{R}_{\mathrm{pos}}^3), \\ a_e^P(\Delta') &:= a_e(\mathbb{R}_{\mathrm{pos}}^3 \times \Delta'), & \Delta' \in \Sigma(\mathbb{R}_{\mathrm{mom}}^3). \end{aligned} \quad (2.16)$$

a_e^Q, a_e^P are POV-measures on position resp. momentum space. By integration they lead to operators

$$\begin{aligned} Q_e^k &:= \int_{\mathbb{R}_{\mathrm{pos}}^3} q_k a_e^Q(dq) & q \in \mathbb{R}_{\mathrm{pos}}^3, k = 1, 2, 3 \\ P_e^k &:= \int_{\mathbb{R}_{\mathrm{mom}}^3} p_k a_e^P(dp) & p \in \mathbb{R}_{\mathrm{mom}}^3, k = 1, 2, 3 \end{aligned} \quad (2.17)$$

where the integral is taken in the weak sense. As already remarked above, the operators Q_e^k, P_e^k fulfil the commutation relations

$$[Q_e^k, P_e^j] = i\hbar \delta_{jk} \mathbf{1}.$$

Furthermore, Q_e^k transforms covariantly under translations and is invariant under boosts, whereas P_e^k is invariant under translations and

transforms covariantly under boosts. Therefore Q_e^k, P_e^j can be called position and momentum operators. Since they were derived from POV-measures, and not from PV-measures, they are regarded as *unsharp* position and momentum operators.

The marginals (2.13) of $O_*^e(\rho)$ are related to the marginal POV-measures a_e^Q, a_e^P by

$$\begin{aligned} \text{tr}(\rho a_e^Q(\Delta)) &= \int_{\Delta} dq \left(\int_{\mathbb{R}^3} dp O_*^e(q, p) \right) \\ &= \int_{\Delta} dq \left(\int_{\mathbb{R}_{\text{pos}}^3} |e(q' - q)|^2 \langle q' | \rho | q' \rangle dq' \right) \end{aligned} \quad (2.18)$$

$$\begin{aligned} \text{tr}(\rho a_e^P(\Delta')) &= \int_{\Delta'} dp \left(\int_{\mathbb{R}_{\text{mom}}^3} |\hat{e}(p' - p)|^2 \langle p' | \rho | p' \rangle dp' \right) \\ &= \int_{\Delta'} dp \left(\int_{\mathbb{R}_{\text{pos}}^3} dq O_*^e(q, p) \right) \end{aligned}$$

It is easier to give a probability interpretation to the marginal POV-measures a_e^Q, a_e^P than to the marginals of the phase space probability density $O_*^e(\rho)$ of the state ρ . The difficulty with a probability interpretation of the marginals of the $O_*^e(\rho)$ was that the smearing function e does not have anything to do with the state ρ . Each state ρ is smeared in the same way (2.13) by e . e is not related to the different states having different spreading in position and momentum.

This suggests two possible interpretations of the function e : Following Prugovečki (1984), Ali (1985), or Ali and Doebner (1990), one can explain the ρ -independence of the smearing function e by saying that e characterises the measurement apparatus performing the simultaneous unsharp position-momentum measurements described by the POV-measure a_e . Since e refers to the measuring apparatus and not to the system described by ρ , it is natural that the smearing function e does not depend on ρ .

Carrying operationalism even further, one might define space to be what we measure in position measurements. Then the function e does not only characterise position measurement instruments, but is an intrinsic property of space itself. Spelling this out, Prugovečki (1984) is thus led to the concept of *stochastic spacetime*.

There is a second way to interpret the function e : As above, the e -dependence of $O_*^e(\rho)$ can be shifted from the states to the observables. The result of this are the smeared marginal POV-measures a_e^Q, a_e^P . The arbitrariness in the choice of e then simply reflects arbitrariness of quantisation maps, as expressed also by the ordering problem. In this way e refers to the quantisation and not to the system described by ρ . Then it is natural that the smearing function e in the marginal $O_e^*(\rho)$ does not depend on ρ .

There are reasons to consider the first interpretation preferable, as do for example Ali and Doebner (1990). Associating the choice of e , and therefore the choice of an ordering, with the characteristics of a measurement apparatus certainly has the advantage that it provides physical reasons for the choice of one of the mathematically equivalent orderings. An operationalist should like this.

A realist might find it difficult to accept that the arbitrary function e specifies a property of space. Being realist about space, the one real space distinguishes one function e . It is not satisfactory that there is one function e distinguished by nature, and that the formalism is unable to reflect this distinction. The realist can choose the second interpretation, which shifts the arbitrariness in the choice of e to the quantisation map.

(a_e^Q, U_a) and (a_e^P, V_b) are systems of covariance on position and mo-

mentum space. The POV-measures a_e^Q on position space (defined in equation (2.16)) has the covariance properties

$$U_a a_e^Q(\Delta) U_a^* = a_e^Q(\Delta + a), \quad \Delta \in \Sigma(\mathbb{R}_{\text{pos}}^3). \quad (2.19)$$

(a_e^Q, U_a) is therefore a system of covariance on position space. Note that since a_e^Q arose from integration over momentum space, it is invariant under boosts

$$V_b a_e^Q(\Delta) V_b^* = a_e^Q(\Delta), \quad \Delta \in \Sigma(\mathbb{R}_{\text{pos}}^3).$$

(a_e^Q, U_a) is not a system of *imprimitivity* on position space since a_e^Q is not projection valued. However, a_e^Q is still a special kind of POV-measure since it is *commutative*,

$$a_e^Q(\Delta_1) a_e^Q(\Delta_2) = a_e^Q(\Delta_2) a_e^Q(\Delta_1), \quad \Delta_1, \Delta_2 \in \Sigma(\mathbb{R}_{\text{pos}}^3).$$

To see how the position space POV-measure a_e^Q is related to the configuration space PV-measure E defined by

$$(E(\Delta)\psi)(q) = \chi_\Delta(q)\psi(q), \quad \psi \in L^2(\mathbb{R}_{\text{pos}}^3),$$

compare equation (2.18),

$$\text{tr}(\rho a_e^Q(\Delta)) = \int_\Delta dq \left(\int_{\mathbb{R}_{\text{pos}}^3} |e(q' - q)|^2 \langle q' | \rho | q' \rangle dq' \right)$$

with

$$\text{tr}(\rho E(\Delta)) = \int_\Delta dq \langle q | \rho | q \rangle, \quad \Delta \in \mathbb{R}_{\text{pos}}^3.$$

Obviously $a_e^*(\Delta)$ is obtained from smearing the sharp position operator $E(\Delta)$ with the probability measure $|e|^2$,

$$(a_e^Q(\Delta)\phi)(x) = (\chi_\Delta * |e(x)|^2)\phi(x), \quad \phi \in L^2(\mathbb{R}_{\text{pos}}^3). \quad (2.20)$$

(Note that all commutative POV-measures can be written as smeared PV-measures.³⁸) Because a_e^Q arises from smearing the sharp localisation PV-measure E , it can be considered as an *unsharp* localisation in position space.

Comparison of phase space systems of covariance and configuration space systems of imprimitivity. As marginals of the phase space POV-measure a_e we arrived at POV-measures a_e^Q, a_e^P which are *not* projection valued. We were thus led to *unsharp* localisation in position and in momentum space. This contradicts the assumption (2) made on page 76 when we introduced systems of imprimitivity by sharp localisation in configuration space.

Another important difference between phase space systems of covariance and configuration space systems of imprimitivity is the following one: the phase space POV-measure a_e is covariant with respect to boosts and translations, whereas the configuration space PV-measures are covariant only with respect to the translations but invariant with respect to boosts.

One has to pay a price for the more general covariance properties of phase space POV-measures. As mentioned above, configuration space systems of imprimitivity can be very nicely described from a lattice theoretic point of view: The configuration space PV-measure E is simply a lattice isomorphism from the Boolean σ -algebra of the Borel sets of configuration space into a Boolean sublattice of $\mathcal{P}(\mathcal{H})$. In contradistinction to this, a satisfying lattice theoretic description of phase space POV measures is still wanting.

³⁸See Ali (1984).

2.3.3 Systems of covariance in geometric quantisation.

In this section I will point out some connections between geometric quantisation and phase space quantisation by systems of covariance. I will sketch the different steps of the geometric quantisation procedure at different levels of generality: I start with some remarks on the choice of the prequantum bundle for general symplectic manifolds. Then I will describe prequantisation for cotangent bundles. Thirdly, geometric quantisation will be made explicit for the Weyl system. Since I am mainly interested in the connection arising between quantum and classical mechanics at the intermediate level of prequantisation, I will not describe the geometrical technicalities connected to the choice of polarisation. These aspects are covered, for example, in the textbook of Woodhouse (1976).

The main aim of this section is to establish, for the simple case of the Weyl system, a connection between geometric quantisation and phase space quantisation by systems of covariance.

Prequantisation in general. The geometric quantisation programme proceeds in two steps: In the prequantisation step, one finds a Lie algebra morphism from the functions of the symplectic manifold to the Hamiltonian vector fields on some other manifold, called the prequantum bundle. This morphism maps the position and momentum coordinate functions into operators fulfilling the commutation relations. Since we have a Lie algebra morphism, requirement 3 of the Dirac problem is fulfilled. But requirement 4 is violated: the vector fields corresponding via the morphism to position and momentum coordinate functions are first order differential operators, and the von Neumann algebra they

generate is not irreducible. Therefore, in a second step, one chooses via a polarisation an irreducible subrepresentation of the commutation relations. In doing so, one gives up requirement 3 in order to fulfil requirement 4.

For a general symplectic manifold (Γ, σ) , prequantisation is achieved in the following way. First one tries to find a principal $U(1)$ -bundle Y over Γ with a connection form α and curvature σ . Y is called the prequantum bundle. For an arbitrary symplectic manifold it is not necessarily possible to find a prequantum bundle. But one can formulate an integrality condition which is necessary and sufficient for the existence of a prequantum bundle.³⁹ Given the prequantum bundle (Y, α) , for every function $f \in C^\infty(\Gamma)$ there is a unique vector field V_f on Y such that $\alpha(V_f) = f$ ⁴⁰ and that the image of V_f under the projection $Y \rightarrow \Gamma$ is the hamiltonian vector field X_f .

Note that even if a prequantum bundle (Y, α) exists, it need not be unique. A necessary and sufficient condition for the uniqueness of the prequantum bundle is that Γ is simply connected. Non-uniqueness of the prequantum bundle can be used to explain topological quantum effects, as for example the Aharanov-Bohm effect.⁴¹

Prequantisation of a cotangent bundle. If the symplectic manifold is a cotangent bundle, $\Gamma = T^*Q$, for some configuration space Q , one possible choice for the prequantum bundle is $T^*Q \times U(1)$. This is the only choice if Q is simply connected. A smooth function f on $T^*Q = \Gamma$ defines a Hamiltonian vector field X_f . In a local trivialisation

³⁹See Woodhouse (1992, section 8.3).

⁴⁰Here α is considered a 1-form; I use α for the connection form and the connection.

⁴¹See Woodhouse (1992, p. 188).

the symplectic form σ is $\sigma = \sum_i dp_i \wedge dq_i$, and the vector field X_f is given by

$$X_f(q, p) := \left(\frac{\partial f}{\partial p}(q, p), -\frac{\partial f}{\partial q}(q, p) \right).$$

The mapping $f \mapsto X_f$ is a mapping of Lie algebras and has as kernel the constant functions on Γ .

In traditional quantum mechanics, states are represented by points in the projective Hilbert space, or equivalently by circles on the unit sphere of \mathcal{H} . The prequantum bundle is a reconstruction of this situation in classical mechanics. Define a space Y and a projection $\pi : Y \rightarrow T^*Q$ in local coordinates by

$$Y := T^*Q \times U(1), \quad \pi(q, p, \exp(i\delta)) := (q, p).$$

We use the angle δ as a coordinate on $U(1)$, and denote a vector field X on Γ as a mapping $\gamma \rightarrow \mathbb{R}^{6n}$, and a vector field V on Y as a mapping $V : Y \rightarrow \mathbb{R}^{6n+1}$. To every smooth function on Γ one associates a Hamiltonian vector field V_f on Y , defined locally by

$$V_f(q, p, \delta) := \left(\frac{\partial f}{\partial p}(q, p), -\frac{\partial f}{\partial q}(q, p), f(q, p) - p \cdot \frac{\partial f}{\partial p}(q, p) \right). \quad (2.21)$$

This definition ensures that the map $f \mapsto V_f$ is an *injective* Lie algebra morphism from the Poisson algebra to vector fields on Y , and that the image of V_f under the projection π is X_f . The 1-form α on Y defined by $\alpha = p \cdot dq + d\delta$ satisfies $d\alpha = \sigma$. Thus Y is a principal $U(1)$ -bundle over Γ with connection form α and curvature σ .

So far we have just incorporated phase factors into classical mechanics and obtained the prequantum bundle. Now we will use the prequantum bundle for prequantisation. We want to associate to the smooth functions on Γ operators on a Hilbert space. Associating to

functions f the differential operator $i\hbar X_f$ on $L^2(\Gamma)$ is no good, because requirement 2 of the Dirac problem is not satisfied: the constant functions are mapped into the zero operator. In a second attempt, one can associate to smooth functions f on Y the differential operator $i\hbar V_f$ on the Hilbert space $L^2(Y)$ of functions square integrable with respect to the canonical measure $dpdq d\delta$. Again, requirement 2 of the Dirac problem is not satisfied: although $i\hbar V_1$ is not the zero operator on $L^2(Y)$, it is not the identity operator either; it is $-i\hbar(\partial/\partial\delta)$. However, one can define the Hilbert space \mathcal{H}_{PQ} (PQ stands for prequantisation) to be the subspace of $L^2(Y)$ for which $-i\hbar V_1$ is the identity operator. \mathcal{H}_{PQ} is the Hilbert space

$$\mathcal{H}_{PQ} = \left\{ g \in L^2(Y) : g(q, p, \delta) = f(q, p) e^{i\delta/\hbar}, f \in L^2(\Gamma) \right\} \cong L^2(\Gamma). \quad (2.22)$$

We thus arrive at a prequantisation map O_{PQ} which associates to every smooth function f on Y an operator on \mathcal{H}_{PQ} by

$$f \mapsto O_{PQ}(f) := -i\hbar V_f. \quad (2.23)$$

Prequantisation of the Weyl system. Now we will start to give an explicit account of geometric quantisation of the Weyl system. This is a special case of what was dealt with above. The reason why I give the explicit formulas is that I will need them in order to display the relation between geometric quantisation and phase space quantisation by systems of covariance.

First I will briefly describe prequantisation of the Weyl systems, following Ali and Emch (1986). Instead of working on the Hilbert space \mathcal{H}_{PQ} , I will use the Hilbert space $L^2(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3)$, which will be denoted by \mathcal{H}_Γ . On \mathcal{H}_Γ , the prequantised position and momentum operators

can be constructed in the following way. On \mathcal{H}_Γ , a representation W_Γ of the Weyl relations is given by

$$(W_\Gamma(a, b)\psi)(p, q) := \exp((i/\hbar)a(q - b))\psi(p - a, q - b), \quad \psi \in \mathcal{H}_\Gamma. \quad (2.24)$$

Define $O_{PQ}(p), O_{PQ}(q)$ as self-adjoint generators of W_Γ ,

$$W_\Gamma(a, b) =: \exp(-(i/\hbar)aO_{PQ}(p)) \exp((i/\hbar)bO_{PQ}(q)). \quad (2.25)$$

By differentiating this, one sees that the prequantised position and momentum operators $O_{PQ}(p), O_{PQ}(q)$ are given on \mathcal{H}_Γ by

$$O_{PQ}(q^k) = q^k + i\hbar \frac{\partial}{\partial p^k}, \quad O_{PQ}(p^k) = -i\hbar \frac{\partial}{\partial q^k}, \quad (2.26)$$

where $k = 1, 2, 3$. (Using equations (2.23), (2.21), and (2.22) one can check that, on \mathcal{H}_{PQ} , the prequantised position and momentum operators have the same form as in equation (2.26).) They fulfil the commutation relations

$$i\hbar O_{PQ}(\{q^i, p^k\}) = [O_{PQ}(q^i), O_{PQ}(p^k)] = i\hbar \delta_{ik} \mathbf{1},$$

where $i, k \in \{1, 2, 3\}$. $O_{PQ}(p^k), O_{PQ}(q^i)$ are the position and momentum operators produced by prequantisation. Since they fulfil the commutation relations, they can be regarded as quantum mechanical position and momentum operators. They are, however, not the position and momentum operators of the Schrödinger representation. The von Neumann algebra $\{O_{PQ}(q^i)O_{PQ}(p^k)\}''$ they generate is not irreducible. Therefore requirement 4 of the Dirac problem is not fulfilled.

To functions $f \in C^\infty(\Gamma)$ on the phase space $\Gamma = \mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$, the geometric quantisation map associates operators on \mathcal{H}_Γ , which are, according to equation (2.21), given by

$$O_{PQ}(f) = -i\hbar \frac{\partial f}{\partial p} \frac{\partial}{\partial q} + i\hbar \frac{\partial f}{\partial q} \frac{\partial}{\partial p} + f - p \cdot \frac{\partial f}{\partial p}.$$

Observe that O_{PQ} satisfies requirement 3 of the Dirac problem,

$$[O_{PQ}(f), O_{PQ}(g)] = -i\hbar O_{PQ}(\{f, g\}).$$

Note that the von Neumann algebra $\{O_{PQ}(f) : f \in C^\infty(\Gamma)\}''$ generated by the prequantisation operators of *all* classical observables is irreducible.

The second step in the geometric quantisation of the Weyl system. The von Neumann algebra $\{O_{PQ}(q^i)O_{PQ}(p^k)\}'' \subset \mathcal{B}(\mathcal{H}_\Gamma)$ generated by the prequantised position and momentum operators is not irreducible. In order to fulfil requirement 3 of the Dirac problem, the second step in the geometric quantisation procedure is the choice of a polarisation. Algebraically, this is described by the choice of a projector P in the commutant of $\{O_{PQ}(q^i)O_{PQ}(p^k)\}''$ which projects onto an irreducible subspace \mathcal{H}_P of \mathcal{H}_Γ . (Such projections will later be explicitly constructed.) Define a quantisation map

$$f \mapsto O_{GQ}^P(f) := PO_{PQ}(f)P, \quad (2.27)$$

where O_{GQ}^P is an operator on \mathcal{H}_P . O_{GQ}^P is the geometric quantisation map. By construction of O_{GQ}^P , the von Neumann algebra $\{O_{GQ}^P(q^i)O_{GQ}^P(p^k)\}''$ is irreducible. By the Stone-von Neumann uniqueness theorem, O_{GQ}^P is therefore unitarily equivalent to the Schrödinger representation, which I will denote by O_S . Note that P commutes with the prequantised position and momentum operators. Therefore, on \mathcal{H}_P , we have $O_{GQ}^P(p^k) = O_{PQ}(p^k)$ and $O_{GQ}^P(q^k) = O_{PQ}(q^k)$,

$$O_{GQ}^P(q^k) = q^k + i\hbar \frac{\partial}{\partial p^k}, \quad O_{GQ}^P(p^k) = -i\hbar \frac{\partial}{\partial q^k}. \quad (2.28)$$

These are the position and momentum operators produced by geometric quantisation. Thus, as far as position and momentum are concerned,

O_{GQ}^P is still a Lie algebra morphism, $i\hbar O_S(\{q^i, p^k\}) = [O_{GQ}(q^i), O_{GQ}(p^k)] = i\hbar\delta_{ik}\mathbf{1}$. But for a classical observable f with $O_{PQ}(f) \notin \{O_{PQ}(q^i)O_{PQ}(p^k)\}''$ we do not have in general $[O_{PQ}(f), P] = 0$. Thus on \mathcal{H}_P , $O_{GQ}^P(f)$ does not coincide with $O_{PQ}(f)$. So O_{GQ}^P is not a Lie algebra morphism on the whole of $C^\infty(\Gamma)$: the Poisson bracket is not represented by the commutator. This can be considered as a consequence of van Hove's (1951) result that the Schrödinger representation cannot be extended to a Lie algebra morphism of $C^\infty(\Gamma)$.

Explicit construction of projectors in the commutant of $\{O_{PQ}(q^i)O_{PQ}(p^k)\}''$. Following Ali and Emch (1986), we will now construct explicitly projectors P in the commutant of $\{O_{PQ}(q^i)O_{PQ}(p^k)\}''$. The Schrödinger system is defined on $L^2(\mathbb{R}^3)$ by means of the irreducible representation W_S of the Weyl commutation relations,

$$W_S(a, b)\psi(x) := \exp((i/\hbar)xb)\psi(x - a), \quad (2.29)$$

The generators of W_S are the Schrödinger position and momentum operators

$$O_S(q^k) = x^k, \quad O_S(p^k) = -i\hbar \frac{\partial}{\partial x^k}. \quad (2.30)$$

The algebra of observables of the Schrödinger quantised Weyl system is $\{W_S(a, b)\}'' = \mathcal{B}(L^2(\mathbb{R}^3))$ on which $\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$ acts by $W_s \cdot W_s^*$.

Every $e \in L^2(\mathbb{R}^3)$ fulfilling (2.8) defines by equation (2.10) a covariant embedding of $L^\infty(\Gamma)$ into $\mathcal{B}(L^2(\mathbb{R}^3))$. The choice of P in the geometric quantisation procedure can be related to the choice of e used in (2.10) of a phase space system of covariance. This is done in the following way.

For an arbitrary $e \in L^2(\mathbb{R}^3)$ fulfilling (2.8) define a partial isometry

$V_e : L^2(\mathbb{R}^3) \rightarrow \mathcal{H}_\Gamma$ by

$$(V_e\psi)(p, q) := \langle W_S(p, q)e, \psi \rangle, \quad \psi \in L^2(\mathbb{R}^3). \quad (2.31)$$

Using (2.24), (2.29), (2.31) one can show by explicit calculation that V_e intertwines W_Γ, W_S ,

$$W_\Gamma(p, q)V_e = V_eW_S(p, q). \quad (2.32)$$

This and (2.25), (2.30) imply

$$V_e^*O_{PQ}(p^k)V_e = O_S(p^k) \quad \text{and} \quad V_e^*O_{PQ}(q^k)V_e = O_S(q^k). \quad (2.33)$$

Compare this with

$$P_eO_{PQ}(p^k)P_e = O_{GQ}^{P_e}(p^k) \quad \text{and} \quad P_eO_{PQ}(q^k)P_e = O_{GQ}^{P_e}(q^k), \quad (2.34)$$

which follows from the definition of $O_{GQ}^{P_e}$ in (2.27). $V_eV_e^* =: P_e$ is a projector onto a subspace of \mathcal{H}_Γ , which will be denoted by $\mathcal{H}_e := P_e\mathcal{H}_\Gamma$. The geometric quantisation map $O_{GQ}^{P_e}$ (2.27) and the Schrödinger representation O_S are related by the unitary operator P_eV_e .

Since $V_e^*V_e$ is the identity operator on $L^2(\mathbb{R}^3)$, one obtains from (2.32)

$$P_eW_\Gamma(p, q) = W_\Gamma(p, q)P_e \quad (2.35)$$

Thus P_e is a projector in the commutant of $\{O_{PQ}(q^i)O_{PQ}(p^k)\}''$.

Connection of prequantisation to the Koopman representation of classical mechanics. Now I will point out, for the Weyl system, the connection between geometric quantisation and the Koopman formalism. In doing so I will again follow Ali and Emch (1986).

In the Koopman formalism, the classical observables are represented as multiplication operator on \mathcal{H}_Γ by

$$(M(f)\psi)(p, q) := f(p, q)\psi(p, q), \quad f \in C^\infty(\Gamma).$$

We assume that the expectation value of $O_S(p^k)$ and $O_S(q^k)$ in the state e vanishes for all $k = 1, 2, 3$. (This can be achieved by choosing the origin in $\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$ appropriately.) Then it is possible⁴² to show that

$$V_e^* M(p^k) V_e = O_S(p^k), \quad \text{and} \quad V_e^* M(q^k) V_e = O_S(q^k), \quad (2.36)$$

$$P_e M(p^k) P_e = O_{GQ}^{P_e}(p^k) \quad \text{and} \quad P_e M(q^k) P_e = O_{GQ}^{P_e}(q^k), \quad (2.37)$$

Combining equations (2.36), (2.33), (2.37), and (2.34) we obtain

$$V_e^* O_{PQ}(p^k) V_e = V_e^* M(p^k) V_e, \quad \text{and} \quad V_e^* O_{PQ}(q^k) V_e = V_e^* M(q^k) V_e.$$

$$P_e O_{PQ}(p^k) P_e = P_e M(p^k) P_e, \quad \text{and} \quad P_e O_{PQ}(q^k) P_e = P_e M(q^k) P_e.$$

(Of course, since $[P_e, O_{PQ}(p^k)] = 0$, we have $P_e O_{PQ}(q^k) P_e = P_e O_{PQ}(q^k)$.) We therefore arrive at the following: $P_e \cdot P_e$ projects the classical Koopman observables *and* the prequantised observables to the geometrically quantised observables. $V_e^* \cdot V_e$ maps the classical Koopman observables *and* the prequantised observables to the Schrödinger quantised observables. $P_e V_e$ is the unitary operator establishing the equivalence of the geometrically quantised and the Schrödinger quantised observables.

Resulting picture. The map O_{PQ} preserves the Poisson structure, but not the multiplicative structure: since it maps the classical position and momentum operators q^k, p^k onto operators fulfilling the commutation relations. In contradistinction to this, M preserves the multiplicative structure, but not the Poisson structure: the commutators of the functions q^k, p^k vanish, but their Lie brackets are not zero. The situation can be pictured as follows.

⁴²See Ali and Emch 1986, derivation of their equation 2.27.

$$\begin{array}{ccc}
p, q & \xrightarrow{M} & M(p), M(q) \\
\downarrow O_{PQ} & & \downarrow P_e \cdot P_e \\
O_{PQ}(p), O_{PQ}(q) & \xrightarrow{P_e \cdot P_e} & O_{GQ}(p), O_{GQ}(q)
\end{array} .$$

In the upper line of this diagram we have the classical situation, in the lower line the quantum situation. The transition from the upper to the lower line is achieved by the (pre)quantisation maps; this transition destroys the multiplicative structure. The transition from the left column to the right column destroys the Lie algebra structure. The choice of an irreducible subrepresentation, pictured by the map $P_e \cdot P_e$, in the geometric quantisation procedure has the function of restoring irreducibility, whereas in the case of phase space quantisation, it is the quantisation map. In the end, both routes lead to the geometrically quantised position and momentum operators, which are unitarily equivalent to the Schrödinger operators.

Extension and imprimitivity theorems in geometric quantisation. Now we are in position to point out the rôle played by extension and imprimitivity theorems in geometric quantisation. In the generality needed for the Weyl systems such theorems have been proved, in different versions, by Neumann (1972), Scutaru (1977), Cattaneo (1979), Castrigiano and Henrichs (1980), and Ali (1984). Further generalisations of these theorems will be presented in sections 2.5.1 and 2.5.2.

Define a POV-measure a_S on $\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$ with values in $\mathcal{B}(L^2(\mathbb{R}^3))$ by

$$a_S(\Delta) := \int_{\Delta} dpdq W_S(p, q) |e\rangle \langle e| W_S(p, q)^* . \quad (2.38)$$

By construction a_S is covariant with respect to the representation W_S . (a_S, W_S) forms a system of covariance.

Define a PV-measure E_{cl} on Γ with values in $\mathcal{B}(\mathcal{H}_\Gamma)$ by

$$(E_{cl}(\Delta)\psi)(p, q) := \chi_\Delta(p, q)\psi(p, q), \quad \Delta \in \Sigma(\Gamma), \psi \in \mathcal{H}_\Gamma. \quad (2.39)$$

The Koopman representation $f \mapsto M(f)$ of classical mechanics can be regarded as the extension of E_{cl} because

$$M(\chi_\Delta) = E_{cl}(\Delta). \quad (2.40)$$

Using equations (2.39) and (2.24) one verifies that

$$\begin{aligned} W_\Gamma(p, q)E_{cl}(\Delta)W_\Gamma(p, q)^*\psi(p', q') &= \chi_\Delta(p - p', q - q')\psi(p - p', q - q') \\ &= E_{cl}(\Delta + (p, q))\psi(p', q'). \end{aligned}$$

Therefore (E_{cl}, W_Γ) is a system of imprimitivity. This phase space system of imprimitivity encodes: in E_{cl} the classical multiplicative structure; and in W_Γ the classical Poisson structure.

The rôle played by the classical phase space system of imprimitivity (E_{cl}, W_Γ) in geometric quantisation can be understood in terms of extension and imprimitivity theorems. Let me elaborate this.

Naimark's (1943) extension theorem⁴³ and Mackey's imprimitivity theorem⁴⁴ can be combined in the following way to arrive at a generalised imprimitivity theorem. Let (a, U) be a system of covariance on some G -space $X = G/H$ in a Hilbert space \mathcal{H} (see Definition 8). Then there is a unitary representation L of H and an isometric mapping V from \mathcal{H} into \mathcal{H}_L (defined on p. 79) such that

$$VU_g = U_g^L V, \quad g \in G, \quad (2.41)$$

⁴³ This theorem says the following. For any POV-measure a on a Borel space X in a Hilbert space \mathcal{H} there is a unique minimal Hilbert space \mathcal{H}' containing $\mathcal{H} = P\mathcal{H}'$ as a subspace and a PV-measure E on X in \mathcal{H}' such that $a(\Delta) = PE(\Delta)P$.

⁴⁴See p. 83 for a simplified version of Mackey's theorem.

$$Va_S(\Delta) = E^L(\Delta)V, \quad \Delta \in \Sigma(G/H), \quad (2.42)$$

where U^L is the induced representation of (2.3) and E^L is the PV-measure of (2.4). The Hilbert space \mathcal{H}_L is minimal in the sense of Naimark's extension theorem. The system of imprimitivity (E^L, U^L) is determined uniquely up to unitary equivalence by the system of covariance (a, U) . (Defining the projector $P := VV^*$ on \mathcal{H}_L one relates (2.42) to Naimark's extension theorem (footnote 43).)

Let us apply this generalised imprimitivity theorem to the system of covariance (a_S, W_S) defined in equation (2.38). The imprimitivity theorem now states that there is a Hilbert space \mathcal{H}_L , an isometry $V : L^2(\mathbb{R}^3) \rightarrow \mathcal{H}_L$, and a system of imprimitivity (E^L, U^L) on $\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$ such that $VU_g = U_g^L V$ and $Va(\Delta) = E^L(\Delta)V$. But since the subgroup H is trivial in this case, \mathcal{H}_L is $L^2(\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3) = \mathcal{H}_\Gamma$ and equation (2.4) takes the form

$$E^L(\Delta)\psi(p, q) = \chi_\Delta(p, q)\psi(p, q). \quad (2.43)$$

Comparing this with equation (2.39) one sees that $E_{cl} = E^L$. Since U_L is determined up to unitary equivalence, U^L coincides with W_Γ . This can easily be verified by comparing equations (2.32) and (2.41).

The resulting situation can be pictured as follows.

$$\begin{array}{ccc} M & \xleftrightarrow{(2.40)} & (E_{cl}, W_\Gamma) & \xleftrightarrow{(2.25)} & O_{PQ} \\ & & \uparrow (2.42) & & \uparrow (2.32) \\ & & (a_S, W_S) & & \end{array}$$

In the bottom line of this diagram we have the system of covariance (a_S, W_S) describing the Schrödinger quantised Weyl system. This

system of covariance is extended to the system of imprimitivity of the top line. The extension of the subrepresentation W_S to the full W_Γ leads from the Schrödinger quantised system to the prequantised system (2.33). This extension restores the classical Poisson structure of prequantisation. The extension of the POV-measure a_S to the PV-measure E_{cl} restores the classical multiplicative structure of the Koopman representation. The extension of (a_S, W_S) to (E_{cl}, W_Γ) is an inverse of quantisation, leading from the quantum system to the classical.

2.4 Generalised systems of covariance

W-systems.* In the example of the elementary Weyl system, the action $x \mapsto U_a V_b x V_b^* U_a^*$ of $\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$ on $\mathcal{B}(L^2(\mathbb{R}^3))$ is pointwise σ -weakly continuous. $(\mathcal{B}(L^2(\mathbb{R}^3)), \mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3, UV \cdot V^* U^*)$ is thus a W^* -system in the sense of the following definition.

Definition 11 *A W^* -system (\mathcal{M}, G, α) consists of a W^* -algebra \mathcal{M} with separable predual, a locally compact, separable group G , and a representation $\alpha : G \rightarrow \text{Aut}(\mathcal{M})$ of G as a group of automorphisms of \mathcal{M} such that*

- (i) $\alpha_{g_1 g_2} = \alpha_{g_1} \alpha_{g_2}$ and
- (ii) for all operators $x \in \mathcal{M}$ the function $g \mapsto \alpha_g(x)$ is σ -weakly continuous.

The assumption that \mathcal{M} has separable predual is equivalent⁴⁵ to the fact that \mathcal{M} is isomorphic to a von Neumann algebra on a separable Hilbert space. Although for some of the mathematical results this assumption is not necessary, it is usually made in physical applications. Similarly, the assumption that G is separable is not necessary for most of the

⁴⁵See Pedersen (1979, 3.9.9).

results. It can be dropped if one is willing to enter into more intricate mathematical arguments.⁴⁶

EXAMPLE 1: The systems of traditional quantum mechanics can be regarded as W^* -systems in the following way. A quantum mechanical system is specified by a unitary ray representation U of a kinematical group G , which is for example the Poincaré group or the Galilei group. Associated to U is a representation α of G as a group of automorphisms of $\mathcal{B}(\mathcal{H})$ defined by

$$\alpha_g(x) := U_g x U_g^*, \quad g \in G, x \in \mathcal{B}(\mathcal{H}).$$

Then $(\mathcal{B}(\mathcal{H}), G, \alpha)$ is a W^* -system. The choice of $\mathcal{B}(\mathcal{H})$ for \mathcal{M} is a consequence of von Neumann's irreducibility postulate.

Conversely, every W^* -system, where \mathcal{M} is a type I factor, can be brought into the form $(\mathcal{B}(\mathcal{H}), G, U \cdot U^*)$. This is due to the fact that all automorphisms α of a type I factor are inner: they are produced by a unitary operator $U \in \mathcal{M}$ via $\alpha(x) = UxU^*$.

EXAMPLE 2: Every classical W^* -system in the sense of Definition 4 is a W^* -system in the sense of Definition 11.

As these examples show, W^* -systems generalise quantum mechanical and classical systems. They can also describe systems which are partially classical and partially quantum mechanical.

Elementary systems. A quantum mechanical system is said to be elementary if it is described by an irreducible representation of the kinematical group. Following Amann (1978), it is possible to generalise this notion of elementarity to arbitrary W^* -systems.

⁴⁶See Blattner (1961,1962).

Definition 12 A W^* -system (\mathcal{M}, G, α) describes an elementary system if the action of α on \mathcal{M} is ergodic, i.e. if the condition

$$\alpha_g(x) = x, \quad \forall g \in G$$

is satisfied only by multiples $x = \lambda \mathbf{1}$ of the identity operator.

In traditional quantum mechanics an elementary system is described by an irreducible ray representation of the symmetry group G . In this case the two notions of elementary system coincide: A type I factor system (\mathcal{M}, G, α) is ergodic if and only if the corresponding representation $U(G)$ acts irreducibly on \mathcal{H} .⁴⁷

Traditionally, a classical system with symmetry group G is called elementary if G acts transitively on the phase space Γ . Mackey (1962, theorem 3) showed that if the Borel G -space Γ is standard, then the W^* -system $(L^\infty(\Gamma), G, \alpha)$ is ergodic (in the sense defined above) if and only if the action of G on Γ is *metrically transitive* (or ergodic in the sense of traditional dynamical systems), i.e. if for any $\Delta \in \Sigma(X)$, $s_g(\Delta) = \Delta$ for all $g \in G$ implies that $\mu(\Delta) = 0$ or $\mu(\Gamma - \Delta) = 0$. Thus the notions of ergodicity above and in the sense of traditional dynamical systems coincide. But a metrically transitive system is not always transitive.⁴⁸ Thus in the case of classical systems the above definition of elementarity as ergodicity is much broader than the traditional definition of elementarity as transitivity. This is perhaps disturbing.

But if one restricts oneself to a statistical description, then even for classical systems the above definition of elementarity is in full agreement with the traditional one. The reason for this is as follows. In a

⁴⁷See Berberian (1974), theorem 67.2.

⁴⁸This is of paramount importance in classical ergodic theory, which would be inconceivable without measure zero orbits which are dense in the state space. Such systems are ergodic (metrically transitive).

statistical description of classical mechanics, take as algebra of observables $L^\infty(\Gamma, \mu)$. Associate to a Borel subset Δ of Γ as observable the multiplication operator χ_Δ . Then to the sets with $\mu(\Delta) = 0$ the null operator is associated. Thus both transitive and metrically transitive actions on Γ lead to ergodic actions on $L^\infty(\Gamma, \mu)$. The algebra of statistical observables cannot distinguish between transitive and metrically transitive actions. Therefore, in a statistical description of classical mechanics, it does not matter whether one chooses the traditional notion of elementarity as transitivity or the above notion of elementarity as metric transitivity.

Note that elementarity is defined with respect to a group G . A system which is elementary with respect to some group G need not be elementary with respect to another group G' .

Consider for example an n particle Weyl system with algebra of observables $\otimes^n \mathcal{B}(L^2(\mathbb{R}^3))$. The additive group $\mathbb{R}^3 \times \mathbb{R}^3$ of boosts and translations does *not* act ergodically on $\otimes^n \mathcal{B}(L^2(\mathbb{R}^3))$. With respect to this group the n particle system is therefore not elementary. In contradistinction to this, the n -fold direct product $\oplus^n (\mathbb{R}^3 \times \mathbb{R}^3)$ does act ergodically on $\otimes^n \mathcal{B}(L^2(\mathbb{R}^3))$ by extension to linear combinations of

$$\alpha_{\oplus_{i=1}^n (a_i, b_i)} \left(\bigotimes_{i=1}^n x_i \right) := \bigotimes_{i=1}^n U_{a_i} V_{b_i} x_i V_{b_i}^* U_{a_i}^*, \quad x_i \in \mathcal{B}(L^2(\mathbb{R}^3)).$$

The n particle Weyl system is therefore elementary with respect to the group $\oplus^n (\mathbb{R}^3 \times \mathbb{R}^3)$. (The direct product $\oplus^n (\mathbb{R}^3 \times \mathbb{R}^3)$ corresponds to the classical phase space of the n particle system. Its action describes space translations and boosts of each particle *separately*.)

In order not to have to call the n particle Weyl system elementary, one might wish to say that a system is elementary only if one partic-

ular group acts ergodically on the observables. This would eliminate the group-dependence of the notion of elementarity. Such a stipulation seems possible, but it deprives us of the flexibility to call systems elementary which behave as basic building blocks only *in specific situations*. Protons for example can be regarded as basic building blocks in chemical experiments, but they behave as complex composite systems in high energy collisions. In order to retain the flexibility of calling a proton elementary in one situation but not in another, I want to keep the group-dependent definition of elementarity.

Generalised systems of covariance. Using W^* -systems one arrives at a generalised notion of systems of covariance, which will be used in the sequel.

Definition 13 *Let (\mathcal{M}, G, α) be a W^* -system and let X be a standard Borel G -space. A quasi-invariant POV-measure a on X with values in \mathcal{M} , together with the automorphic representation α of G on \mathcal{M} is called a generalised system of covariance (a, α) based on X if a acts covariantly with respect to α ,*

$$\alpha_g(a(\Delta)) = a(g\Delta), \quad \Delta \in \Sigma(X).$$

If G acts transitively on X , then (a, α) is called a transitive generalised system of covariance.

Why is this definition more general than the traditional Definition 8? Definition 8 used a unitary ray representation U of G instead of an automorphic representation. U leads to an automorphism group of $\mathcal{B}(\mathcal{H})$. Thus every system of covariance in the sense of Definition 8 can be regarded as a generalised system of covariance with respect to the

W^* -system $(\mathcal{B}(\mathcal{H}), G, U \cdot U^*)$. But the converse is not true: an automorphic representation of G on \mathcal{M} cannot be replaced by a unitary ray representation of G on \mathcal{H} . The present definition is more general than Definition 8 because the automorphism group α need not be of the form $\alpha_g = U_g \cdot U_g^*$ for any unitary ray representation U . For this there are several reasons. I will briefly make some remarks which automorphic group representations can be replaced with a unitary ray representations and which not.

Unitary implementability of automorphic group representations. An automorphic group representation α cannot be replaced with a unitary ray representation if there is some α_g which is not spatial. (An automorphism α of a von Neumann algebra $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ is called spatial if there is a unitary operator U on \mathcal{H} , U not necessarily in \mathcal{M} , such that $\alpha = U \cdot U^*$.) Non-spatial automorphisms are essential in the description of spontaneous symmetry breaking: A symmetry, which is described by an automorphism α of a C^* -algebra \mathcal{A} of observables commuting with the dynamics, is said to be broken in a representation π of \mathcal{A} if $\pi \circ \alpha$ is not a spatial automorphism (Streater 1965) or if it is not an automorphism at all (Morchio and Strocchi 1987).⁴⁹

⁴⁹In the traditional (non-algebraic) formulation of quantum field theories, spontaneous symmetry breaking is said to occur if the “formal” charge $Q(x_0) = \int_{\mathbb{R}^3} d\mathbf{x} j_0(x)$ associated with a conserved current j_μ does not exist as an operator and therefore cannot be the generator of a symmetry, i.e. of a unitary group of transformation generated by a self-adjoint operator commuting with the momentum operator and the S -matrix. The algebraic and the traditional definition are not in a straightforward relation, but they are connected by the idea that in the representation specified by a vacuum of the traditional theory there is no automorphism group although the existence of a conserved current suggests a certain symmetry of the Hamiltonian. For a review of spontaneous and intrinsic symmetry breaking in the traditional formalism see Orzalesi (1970). For a discussion in the algebraic framework see Streater (1965) or Swieca (1970), or the more recent paper of Buchholz *et al.* (1992).

Let me quote some partial results on when automorphisms are spatial. If \mathcal{M} is a factor of type I or III, any automorphism is spatial. But if \mathcal{M} is a factor of type II_∞ with commutant of type II_1 , then there are examples⁵⁰ of automorphisms which are not spatial, even if \mathcal{H} is separable. It is possible to give a necessary and sufficient condition⁵¹ when an automorphism of a type II_∞ factor with commutant of type II_1 is spatial. If \mathcal{M} is not a factor, then one can guarantee that an automorphism is spatial if the central decomposition of \mathcal{M} does not contain any factor of type II and if the automorphism leaves every central element fixed. This last condition is a severe restriction in physical applications. It can be dropped if for example the commutant of \mathcal{M} is abelian⁵², or if the commutant is of the form $C \otimes F$, where C is an abelian von Neumann algebra and F is a factor,⁵³ or if \mathcal{M} is of type III and operates on a separable Hilbert space.⁵⁴

Up to now we only dealt with the unitary implementability of automorphisms. But even if all α_g are spatial it is in general not possible to replace α by a unitary ray representation. This is the practically most relevant aspect of generalising from unitary ray representations of G to automorphic representations. Assume that each automorphism α_g can be implemented by a unitary operator U_g . By glueing together the U_g one obtains a unitary representation $U(G)$ *with multiplier in \mathcal{M}'* .⁵⁵ Thus, only if \mathcal{M} is irreducible the U_g form necessarily a unitary ray representation. (If \mathcal{M} is of type I then the multipliers can be cho-

⁵⁰See Kadison and Ringrose (1986, 13.4.3).

⁵¹See Kadison and Ringrose (1986, 9.6.33).

⁵²See Dixmier (1957), III.3.2, Corollary to Proposition 2.

⁵³See Dixmier (1957), III.3.2, Proposition 2.

⁵⁴See Dixmier (1957), III.8.6, Corollary 8.

⁵⁵The resulting non-abelian cohomology theory has been described in Streater (1990).

sen to be in the centre of \mathcal{M} .⁵⁶) Furthermore the multipliers need not commute with the U_g .⁵⁷

Covariant embeddings. Let (a, α) be a system of covariance based on X . Since \mathcal{M} was assumed to have separable predual, it is σ -finite, and therefore there exists a vector ξ which is cyclic and separating for \mathcal{M} .⁵⁸ Define a scalar-valued measure μ on X by

$$\mu(\Delta) := \langle \xi | a(\Delta) | \xi \rangle, \quad \Delta \in \Sigma(X).$$

We have $\mu(\Delta) = 0$ iff $a(\Delta) = 0$. Therefore, since a was assumed to be quasi-invariant, μ is quasi-invariant as well.

The map $\Delta \mapsto a(\Delta)$ can be extended σ -weak continuously to a map O from $L^\infty(X, \mu)$ into \mathcal{M} satisfying

$$O(\chi_\Delta) = a(\Delta).$$

If (a, α) is a system of covariance, then O is a covariant embedding in the following sense.

Definition 14 *A covariant embedding of the classical W^* -system $(L^\infty(X, \mu), G, Ad\lambda)$ into a W^* -system (\mathcal{M}, G, α) is a map $O : L^\infty(X, \mu) \rightarrow \mathcal{M}$ satisfying*

- (i) *linearity:* $O(cf + g) = cO(f) + O(g)$ for $f, g \in L^\infty(X, \mu), c \in \mathbf{C}$,
- (ii) *covariance:* $O(Ad\lambda(g)f) = \alpha_g(O(f))$ for all $f \in L^\infty(X, \mu)$,
- (iii) *positivity:* $O(f) \geq 0$ for all $f \in L^\infty(X, \mu)$ with $f(x) \geq 0$ a.e.,
- (iv) $O(f)$ is self-adjoint if f is real,
- (v) *normalisation:* $O(1) = \mathbf{1}$.

⁵⁶See Theorem 1 of Streater 1990.

⁵⁷When such non-trivial multipliers occur between a one-parameter symmetry group and the time evolution group, then the generator of the symmetry does not commute with the Hamiltonian, and is thus not a conserved quantity. Models with such anomalies can be found in Chen *et al.* (1989) and Streater (1990).

⁵⁸See e.g. Bratteli and Robinson (1987, Proposition 2.5.6).

This definition is a straightforward generalisation of Amann's (1986a) definition of covariant embeddings of $L^\infty(G)$ into a W^* -system. Note that positivity of O only implies that $O(f)$ is a positive operator if f is a positive function. It does not mean that every positive operator in $O(L^\infty(X, \mu))$ is the image of a positive function. In general there are positive operators in $O(L^\infty(X, \mu))$ which are not the image of positive functions.

The term 'embedding' might be somewhat misleading, since O preserves only the linear, but not the multiplicative structure of $L^\infty(X, \mu)$. If the multiplicative structure is preserved by O then O is a covariant representation in the following sense.

Definition 15 *A covariant embedding O is called a covariant representation of $L^\infty(X, \mu)$ in \mathcal{M} if it additionally satisfies condition*

(v) $O(fg) = O(f)O(g)$ for all $f, g \in L^\infty(X, \mu)$.

Proposition 1 *To every covariant embedding O of $L^\infty(X, \mu)$ into a W^* -system (\mathcal{M}, G, α) there corresponds a generalised system of covariance (a, α) based on X , and vice versa. To every covariant representation of $L^\infty(X, \mu)$ in a W^* -system (\mathcal{M}, G, α) there corresponds a system of imprimitivity based on X , and vice versa.*

PROOF: Let (a, α) be a generalised system of covariance based on X . Since \mathcal{M} was assumed to be σ -finite, there exists a vector ξ which is cyclic and separating for \mathcal{M} .⁵⁹ Define a scalar-valued measure μ on X by

$$\mu(\Delta) := \langle \xi | a(\Delta) | \xi \rangle, \quad \Delta \in \Sigma(X).$$

We have $\mu(\Delta) = 0$ iff $a(\Delta) = 0$.⁶⁰ Therefore, since a was assumed to be quasi-invariant, μ is quasi-invariant as well. a can be extended

⁵⁹See e.g. Bratteli and Robinson (1987, Proposition 2.5.6).

⁶⁰See e.g. Conway (1985, IX. 8. 3).

σ -weak continuously to a covariant embedding of $L^\infty(X, \mu)$ into the W^* -system (\mathcal{M}, G, α) . Covariance of a implies covariance of O . O is a representation if a is projection-valued.

A covariant embedding O defines a covariant POV-measure a on X by

$$a(\Delta) := O(\chi_\Delta),$$

where $\Delta \in \Sigma(X)$ and where χ_Δ is the characteristic function of Δ . The POV-measure a is projection-valued if O is a representation. \square

By this Proposition it is admissible to identify generalised systems of covariance (a, α) with covariant embeddings into a W^* -system (\mathcal{M}, G, α) . In the sequel we will use these terms as synonyms.

The non-uniqueness of covariant embeddings. Not every covariant embedding O into a given W^* -system (\mathcal{M}, G, α) is appropriate for describing a quantisation procedure. For example, if G is compact so that the Haar-measure can be normalised, then one can define a covariant embedding O of $L^\infty(G)$ into an arbitrary W^* -system (\mathcal{M}, G, α) by

$$O(f) := \left(\int_G f(g) dg \right) \mathbf{1}, \quad f \in L^\infty(G).$$

This covariant embedding is trivial in the sense that its range is $\{\lambda \mathbf{1} : \lambda \in \mathbf{C}\}$.

One has to impose an additional condition which guarantees that the range $O(L^\infty(X, \mu))$ exploits \mathcal{M} as far as possible. Following Amann (1986a), we require quantisation maps to be covariant embeddings O which are maximal in the sense of Definition 10 (with $\mathbb{R}_{\text{pos}}^3 \times \mathbb{R}_{\text{mom}}^3$ replaced by G).

In general there exist many maximal covariant embeddings of $(L^\infty(X, \mu), G, \text{Ad}\lambda)$ into a given W^* -system (\mathcal{M}, G, α) . The most important reason for this is that $O(L^\infty(X, \mu)_+)$ is usually *properly* contained in \mathcal{M}_+ .

Another reason why in general there are many covariant embeddings is that for a given covariant embedding O there are many embeddings O' with $O(L^\infty(X, \mu)_+) = O'(L^\infty(X, \mu)_+)$. Consider for instance the case where (\mathcal{M}, G, α) is ergodic and integrable.⁶¹ Assume that there is a maximal covariant embedding O of $(L^\infty(G), G, \text{Ad}\lambda)$ into (\mathcal{M}, G, α) . According to Amann (1986a, Theorem II.6), to every $g \in G$ there corresponds a covariant embedding O_g defined by

$$O_g(f) := O(\text{Ad}\rho(g^{-1})f)$$

which fulfils $O(L^\infty(X, \mu)_+) = O'(L^\infty(X, \mu)_+)$. Conversely, for every covariant embedding O' with $O(L^\infty(X, \mu)_+) = O'(L^\infty(X, \mu)_+)$ there is a $g \in G$ such that $O'(f) = O_g(f)$ for all $f \in L^\infty(G)$.

2.4.1 Phase space quantisation by generalised systems of covariance

Suppose now that the phase space Γ of a classical system is homogeneous and therefore of the form G/H for some locally compact group G . As earlier, denote by $L^\infty(\Gamma)$ the complex-valued functions on phase space which are essentially bounded with respect to some measure μ in the unique quasi-invariant measure class on Γ . Let (a, α) be a system of covariance based on Γ . By Proposition 1, a can be extended to a covariant embedding O of $L^\infty(\Gamma)$ into (\mathcal{M}, G, α) . $L^\infty(\Gamma)$ is the algebra of classical statistical observables. For each classical observable $f \in L^\infty(\Gamma)$, $O(f)$ is a bounded operator on a Hilbert space.

⁶¹Integrability of a W^* -system will be introduced in Definition 16.

In order to ensure that the observables $O(f)$ really describe a quantum mechanical system, one has to impose further conditions on (\mathcal{M}, G, α) . Which conditions are regarded as sufficient depends on the situation. Sometimes ergodicity is appropriate, sometimes one would require some commutation relations. If suitable conditions are fulfilled, O can be regarded as a quantisation map.

O is a more general version of the quantisation map O_e introduced in equation (2.10). It is more general because the algebra $O(L^\infty(\Gamma))$ of observables of the quantum mechanical system need not be $\mathcal{B}(\mathcal{H})$. Instead it can be an arbitrary von Neumann algebra \mathcal{M} . This generalisation allows for quantised systems with superselection rules.

In a similar way as in section 2.3.2 for the Weyl system, but now in a more general framework, it is possible to introduce phase representations of the quantised system. A generalised covariant embedding O induces a dual map $O^* : \mathcal{M}^* \rightarrow L^\infty(\Gamma)^*$ defined by

$$O^* : \rho \mapsto \rho \circ O.$$

Thus we have $(O^*(\rho))(f) = \rho(O(f))$ for $f \in L^\infty(\Gamma)$. The preduals \mathcal{M}_* , $L^\infty(\Gamma)_*$ satisfy

$$\mathcal{M}_* \subseteq \mathcal{M}^*,$$

$$L_1(\Gamma) \cong L^\infty(\Gamma)_* \subseteq L^\infty(\Gamma)^*.$$

A covariant embedding is called *normal* if it satisfies $O^*(\mathcal{M}_*) \subseteq L_1(\Gamma)$. For every $O^*(\rho) \in L^\infty(\Gamma)^*$ there is a μ -integrable function $O_*(\rho)$ such that $O^*(\rho)(f) = \int_\Gamma O_*(\rho)(x)f(x)d\mu(x)$ for any $f \in L^\infty(\Gamma)$. Thus for every statistical state ρ of the quantum system there is a μ -integrable function $O_*(\rho)$ on Γ such that

$$\rho(O(f)) = \int_\Gamma (O_*(\rho)(x))f(x)d\mu(x) \quad f \in L^\infty(\Gamma, \mu). \quad (2.44)$$

Thus the normal covariant embedding O induces a *phase space representation* O_* . O_* associates to every statistical state $\rho \in \mathcal{M}_*$ an L^1 -function $O_*(\rho)$ on the phase space such that (2.44) is fulfilled. Since O maps positive functions on Γ into positive operators, O_* maps positive functionals in \mathcal{M}_* into positive L^1 -functions. Since states are normalised (i.e. $\rho(\mathbf{1}) = 1$), we have

$$\int_{\Gamma} O_*(\rho) d\mu(x) = 1.$$

Therefore $O_*(\rho)$ is a probability density on the phase space.

The covariant POV-measure a on the phase space Γ corresponding to O by $a(\Delta) := O(\chi_{\Delta})$ fulfils

$$\rho(a(\Delta)) = \int_{\Delta} (O_*(\rho))(x) d\mu(x), \quad \Delta \in \Sigma(\Gamma). \quad (2.45)$$

a is a covariant POV-measure describing phase space localisation. By (2.45) $\rho(a(\Delta))$ is the probability to find the system, if it is in the state ρ , in the phase space region Δ .

2.4.2 Existence of transitive generalised systems of covariance

Now I will turn to some results on the existence of transitive generalised systems of covariance with values in a given W^* -algebra \mathcal{M} . The results presented in this section are straightforward generalisations of the existence theorems of Amann (1986a) and Homma (1988) for generalised systems of covariance on a group G . This section follows Breuer (1994b).

Integrability of automorphic group representations on von Neumann algebras First I will introduce the notion of integrability for an auto-

morphic group representation on a von Neumann algebra. This definition is due to Connes and Takesaki (1977) and generalises the notion of square-integrability for unitary ray representations of a group on a Hilbert space.

Let (\mathcal{M}, G, α) be a W^* -system. The restriction of the left Haar measure dg to a compact subset K of G is finite. For each fixed $x \in \mathcal{M}$ there is exactly one $y \in \mathcal{M}$ such that $\rho(y) = \int_K \rho(\alpha_g(x))dg$ for all $\rho \in \mathcal{M}_*$. The integral $\int_K \alpha_g(x)dg$ is said to exist in the weak sense and one writes $y = \int_K \alpha_g(x)dg$. The Haar measure on G is not finite if G is not compact. One says that the integral $\int_G \alpha_g(x)dg$ exists if the set

$$\left\{ \int_K \alpha_g(x)dg : K \subset G, K \text{ compact} \right\}$$

is bounded in \mathcal{M} . In this case one writes $\int_G \alpha_g(x)dg < \infty$ and defines $\int_G \alpha_g(x)dg := \sup_K \int_K \alpha_g(x)dg$, where the supremum is taken over all compact subsets of G .

Definition 16 A W^* -system (\mathcal{M}, G, α) is called integrable if the set

$$n_\alpha := \left\{ y \in \mathcal{M} : \int_G \alpha_g(y^*y)dg < \infty \right\}$$

is σ -weakly dense in \mathcal{M} .

Let m_α be the linear hull of the set $m_\alpha^+ := \{x \in \mathcal{M}_+ : \int_G \alpha_g(x)dg < \infty\}$. The integral $\int_G \alpha_g(\cdot)dg$ can be extended from m_α^+ to m_α . We have $m_\alpha = n_\alpha^* n_\alpha := \{x^*y : x, y \in n_\alpha\}$. If $x \in m_\alpha$, then the translation invariance of the Haar measure implies that the integral $\int_G \alpha_g(x)dg$ is in the fixed point algebra

$$\mathcal{M}^\alpha := \{x : x \in \mathcal{M}, \alpha_g(x) = x \text{ for all } g \in G\}.$$

If (\mathcal{M}, G, α) is ergodic, then $\int_G \alpha_g(x)dg$ is a multiple of the identity.

If \mathcal{M} is the von Neumann algebra $\mathcal{B}(\mathcal{H})$ of all bounded operators on a Hilbert space, then the automorphic representation α can be written as

$$\alpha_g(x) := U_g x U_g^*, \quad x \in \mathcal{M},$$

where U is a Borel measurable unitary ray representation of G on \mathcal{H} .

Definition 17 *A Borel measurable unitary ray representation U of G on a Hilbert space \mathcal{H} is square-integrable if there exist at least two non-zero vectors $\xi, \eta \in \mathcal{H}$ such that*

$$\int_G |\langle U_g \eta | \xi \rangle|^2 dg$$

exists.

Note that if G is not unimodular, there are always some pairs $\xi', \eta' \in \mathcal{H}$ such that $\langle U_g \eta' | \xi' \rangle$ is not a square integrable function on G .⁶² A unitary ray representation U is square integrable if and only if the automorphic representation $U \cdot U^*$ is integrable on $\mathcal{B}(\mathcal{H})$.⁶³

Assume that we have two W^* -systems (\mathcal{M}, G, α) and (\mathcal{N}, G, β) satisfying $\mathcal{M} \subset \mathcal{N}$ and $\alpha_g(x) = \beta_g(x)$ for all $x \in \mathcal{M}$. The following then is immediately obvious from the definition of integrability: The integrability of (\mathcal{N}, G, β) implies the integrability of (\mathcal{M}, G, α) , but not vice versa.

We then arrive at the following. Assume that \mathcal{M} is a von Neumann subalgebra of $\mathcal{B}(\mathcal{H})$, and that there exists a Borel measurable unitary ray representation U such that $\alpha_g(x) := U_g x U_g^*$, where the U_g are not necessarily in \mathcal{M} . Then the integrability of (\mathcal{M}, G, α) does *not* imply the square-integrability of U .

⁶²See Duflo and Moore (1976).

⁶³See Amann (1986b, theorem 2).

A sufficient condition. I will now show that integrability of (\mathcal{M}, G, α) is a sufficient condition for the existence of a system of covariance based on G/H with values in \mathcal{M} . This is an obvious application of theorems of Amann (1986a) for ergodic W^* -systems and of Homma (1988) for general W^* -systems showing that the integrability of \mathcal{M} is equivalent to the existence of a covariant embedding of $(L^\infty(G), G, \text{Ad}\lambda)$ into (\mathcal{M}, G, α) .

Proposition 2 *Let (\mathcal{M}, G, α) be a W^* -system, and let H be a closed subgroup of G . If (\mathcal{M}, G, α) is integrable, then there exists a covariant POV-measure on G/H with values in \mathcal{M} .*

PROOF: To each $f \in L^\infty(G/H)$ associate a function $i(f) \in L^\infty(G)$ by

$$i(f)(s) := f(sH), \quad s \in G.$$

According to Homma (1988), the integrability of (\mathcal{M}, G, α) implies the existence of a covariant embedding O of $(L^\infty(G), G, \text{Ad}\lambda)$ into (\mathcal{M}, G, α) . Define a map $O' : L^\infty(G/H) \rightarrow \mathcal{M}$ by

$$O'(f) := O(i(f)), \quad f \in L^\infty(G/H).$$

Normalisation, linearity, and positivity of O' are obviously implied by O having these properties. The covariance of O' follows from

$$\begin{aligned} O'(\text{Ad}\lambda(g)f)(s) &= O(i(\text{Ad}\lambda(g)f))(s) = O(f(g^{-1}sH)) = \\ &= O(\text{Ad}\lambda(g)i(f))(s) = \alpha_g(O(i(f)))(s) = \alpha_g(O'(f))(s). \end{aligned}$$

A covariant POV-measure a on G/H is defined by taking $a(\Delta) := O'(\chi_\Delta)$ for any Borel subset of G/H . □

A necessary condition. In general the existence of a covariant POV-measure on G/H does not imply integrability of (\mathcal{M}, G, α) . But if H is a *compact* closed subgroup of G , then this implication holds.

Proposition 3 *Let (\mathcal{M}, G, α) be a W^* -system, and let H be a closed subgroup of G . If there exists a covariant POV-measure on G/H with values in \mathcal{M} and if H is compact, then (\mathcal{M}, G, α) is integrable.*

PROOF: Since H is compact, the right Haar measure dh on H can be normalised, $\int_H dh = 1$. To every function $f \in L^\infty(G)$ associate a function \bar{f} in $L^\infty(G/H)$ by

$$\bar{f}(gH) := \int_H f(gh)dh.$$

By assumption there exists a covariant POV-measure a on G/H with values in \mathcal{M} . a can be extended to a covariant embedding O' of $(L^\infty(G/H), G, \text{Ad}\lambda)$ into (\mathcal{M}, G, α) . Define a map $O : L^\infty(G) \rightarrow \mathcal{M}$ by

$$O(f) := O'(\bar{f}), \quad f \in L^\infty(G).$$

O is a covariant embedding of $(L^\infty(G), G, \text{Ad}\lambda)$ into (\mathcal{M}, G, α) : from $\overline{\text{Ad}\lambda(g)f} = \text{Ad}\lambda(g)\bar{f}$ and from covariance of O' we have

$$O(\text{Ad}\lambda(g)f) = O'(\overline{\text{Ad}\lambda(g)f}) = O'(\text{Ad}\lambda(g)\bar{f}) = \alpha_g(O'(\bar{f})) = \alpha_g(O(f)).$$

Normalisation of the Haar measure dh and of O' imply $O(1) = O'(\int_H 1dh) = O'(1) = \mathbf{1}$. Positivity and linearity of O also follow from the fact that O' has these properties.

By Homma (1988) the existence of a covariant embedding $O : L^\infty(G) \rightarrow \mathcal{M}$ implies that (\mathcal{M}, G, α) is integrable. \square

Ali, Antoine, and Gazeau (1991a,b) introduced a weaker form of square integrability for unitary group representations. They defined square integrability with respect to a given subgroup H and a cross-section $\sigma : G/H \rightarrow G$. Square integrability in the traditional sense implies square integrability with respect to all subgroups and cross-sections. But square integrability with respect to some subgroup and some cross-section does not imply square integrability in the traditional sense. They showed that square integrability with respect to a subgroup H and a cross-section σ is equivalent to the existence of a quasi-covariant⁶⁴ POV-measure on G/H .

In this spirit, it should be possible to introduce the notion of integrability of an automorphic representation with respect to a subgroup and a cross-section. One might conjecture that integrability of α with respect to a subgroup H and a cross-section σ is equivalent to the existence of a (quasi)covariant embedding of $L^\infty(G/H)$ into (\mathcal{M}, G, α) .

2.5 An extension and an imprimitivity theorem

Various imprimitivity and extension theorems. Mackey's (1949) imprimitivity theorem says that a transitive system of imprimitivity is unitarily equivalent to a system of imprimitivity induced from a representation L of the subgroup H . Furthermore, there is a one-to-one correspondence of the equivalence classes of systems of imprimitivity based on G/H and the equivalence classes of unitary representations of the little group H .

⁶⁴Quasi-covariant here means invariant up to some multiplier. See Ali, Antoine, and Gazeau (1991a,b).

Blattner (1961, 1962) generalised the imprimitivity theorem to locally compact groups which are not necessarily separable. Takesaki (1968, 1973) gave generalisations of Mackey's theorem to C^* - and W^* -systems. He constructed W^* -systems which are induced from W^* -systems of a subgroup H and gave sufficient conditions for W^* -systems to be isomorphic to induced W^* -systems.

The extension theorem of Naimark (1943) says that every POV-measure can be extended to a projection valued measure on an enlarged Hilbert space (see footnote 43). Stinespring (1955) gave a generalisation of Naimark's theorem saying that one need not necessarily consider embeddings of commutative C^* -algebras like $C_0(X)$ into $\mathcal{B}(\mathcal{H})$. A necessary and sufficient condition for an embedding of an arbitrary C^* -algebra \mathcal{A} into $\mathcal{B}(\mathcal{H})$ to be of the form $p\pi(\mathcal{A})p$ for some projector p and some *representation* π is complete positivity.⁶⁵

Amann (1986a, theorem IV.2) proved a theorem which says that every covariant embedding of $(L^\infty(G), G, \text{Ad}\lambda)$ into a W^* -system (\mathcal{M}, G, α) can be extended to a covariant representation of $(L^\infty(G), G, \text{Ad}\lambda)$ into a W^* -system (\mathcal{N}, G, β) containing (\mathcal{M}, G, α) . Amann's theorem is a generalisation of Naimark's theorem in the sense that the embedding is into an arbitrary von Neumann algebra \mathcal{M} , and not necessarily into $\mathcal{B}(\mathcal{H})$. Amann also proves covariance properties of the relevant measures. Amann's theorem is a special case of Naimark's in the sense that only POV-measures on G are considered, and not POV-measures on arbitrary Borel spaces X .

Neumann (1972), Scutaru (1977), Cattaneo (1979), and Castrigiano

⁶⁵A map $d : \mathcal{A} \rightarrow \mathcal{B}$ between C^* -algebras is said to be completely positive if the maps $d \otimes \mathbf{1}_n : \mathcal{A} \otimes M_n \rightarrow \mathcal{B} \otimes M_n$ are positive for all natural numbers n . Here M_n denotes the algebra of complex $n \times n$ matrices and $\mathbf{1}_n : M_n \rightarrow M_n$ denotes the identity map.

and Henrichs (1980), and Ali (1984) combined the results of Mackey and Naimark to arrive at various versions of the imprimitivity and extension theorems formulated on p. 115.

In this section I will first give a straightforward generalisation of Naimark's and Amann's extension theorems. Then I will prove an imprimitivity theorem for *generalised* systems of covariance. In how far this is a generalisation was discussed on pages 122-124.

2.5.1 An extension theorem

Here I will give an extension theorem for generalised systems of covariance. This is a partial generalisation of Naimark's (1943) theorem because embeddings into von Neumann algebras are considered. Naimark's theorem is more general in the sense that it applies to POV-measures on arbitrary Borel spaces, not just on Borel G -spaces. But the fact that we consider POV-measures on G -spaces allows for the construction of a group action with respect to which the extended PV-measure transforms covariantly.

The following extension theorem also generalises Amann's (1986, theorem IV.2) extension theorem, since it applies to covariant POV-measures on arbitrary Borel G -spaces X , and not just to those on G . This generalisation, however, is straightforward, because Amann's proof can be easily adapted.

Theorem 3 *Let (\mathcal{M}, G, α) be a W^* -system. Let O be a covariant embedding of $(L^\infty(X, \mu), G, Ad\lambda)$ into (\mathcal{M}, G, α) . Then there is a W^* -system (\mathcal{N}, G, β) ;
a representation π of $L^\infty(X, \mu)$ in \mathcal{N} fulfilling the covariance condition*

$$\beta_g(\pi(f)) = \pi(Ad\lambda(g)f) \quad f \in L^\infty(X, \mu), g \in G; \quad (2.46)$$

a projector p in the fixed point algebra $\mathcal{N}^\beta := \{x \in \mathcal{N} : \beta_g(x) = x \text{ for all } g \in G\}$;

an isomorphism $i : \mathcal{M} \rightarrow p\mathcal{N}p$ of the W^* -algebras \mathcal{M} and $p\mathcal{N}p$,

$$i(\mathcal{M}) = p\mathcal{N}p, \quad (2.47)$$

fulfilling

$$i(O(f)) = p\pi(f)p, \quad f \in L^\infty(X, \mu). \quad (2.48)$$

The projector p is an atom of \mathcal{N}^β if α acts ergodically on \mathcal{M} . If α is ergodic and if the covariant embedding O is extremal in the set of all covariant embeddings of $L^\infty(X, \mu)$ into (\mathcal{M}, G, α) , then $\pi(L^\infty(X, \mu))$ is maximally abelian in \mathcal{N} : $\pi(L^\infty(X, \mu))' \cap \mathcal{N} = \pi(L^\infty(X, \mu))$.

REMARK: The assumption that \mathcal{M} has separable predual, which enters Theorem 3 through Definition 11, could be replaced by the weaker assumption that \mathcal{M} is σ -finite. The assumption that G is separable could be dropped at the expense of more intricate mathematical arguments, as in Blattner (1961, 1962). But the assumption of local compactness for G is essential.

PROOF: *Step 1: Representation of (\mathcal{M}, G, α) on a Hilbert space.* The construction of the W^* -crossed product (see e.g. Bratteli and Robinson (1987, p. 139)) shows that there exist a representation π_1 of \mathcal{M} and a strongly continuous unitary representation u of G on a Hilbert space \mathcal{H} such that

$$\pi_1(\alpha_g(x)) = u_g\pi_1(x)u_g^*. \quad x \in \mathcal{M}, g \in G \quad (2.49)$$

Step 2: The enlarged Hilbert space. $\pi_1 \circ O : L^\infty(X, \mu) \rightarrow \mathcal{B}(\mathcal{H})$ is a completely positive map from the commutative W^* -algebra $L^\infty(X, \mu)$ into $\mathcal{B}(\mathcal{H})$. Therefore a theorem of Stinespring (1955) or the one of

Naimark (1943) implies that there is a Hilbert space \mathcal{H}' , a bounded linear transformation $V : \mathcal{H} \rightarrow \mathcal{H}'$, and a representation $\pi : L^\infty(X, \mu) \rightarrow \mathcal{B}(\mathcal{H}')$ such that

$$(\pi_1 \circ O)(f) = V^* \pi(f) V, \quad f \in L^\infty(X, \mu) \quad (2.50)$$

$$[\pi(L^\infty(X, \mu) V \mathcal{H})] = \mathcal{H}' \quad (2.51)$$

i.e. \mathcal{H}' is the Hilbert generated by $\pi(L^\infty(X, \mu) V \mathcal{H})$.

Step 3: V is an isometry between \mathcal{H} and a subspace of \mathcal{H}' . Since $(\pi_1 \circ O)(1) = \mathbf{1}_{\mathcal{H}}$, $\pi(1) = \mathbf{1}_{\mathcal{H}'}$, and $(\pi_1 \circ O)(1) = V^* \pi(1) V$, we have $V^* V = \mathbf{1}_{\mathcal{H}}$.

Step 4: Definition of the von Neumann algebra \mathcal{N} . Define $p := V V^*$ to be the projector onto $V \mathcal{H} \subset \mathcal{H}'$. Then $i : \mathcal{M} \rightarrow \mathcal{B}(\mathcal{H}')$ defined by

$$i(x) := V \pi_1(x) V^*, \quad x \in \mathcal{M}$$

is a representation of \mathcal{M} in the W^* -algebra $p \mathcal{B}(\mathcal{H}') p$. We define \mathcal{N} as the von Neumann subalgebra of $\mathcal{B}(\mathcal{H}')$ generated by $\pi(L^\infty(X, \mu))$ and $i(\mathcal{M})$. So from equation (2.50) we conclude that $V^* \mathcal{N} V = \pi_1(\mathcal{M})$. This implies

$$p \mathcal{N} p = V(V^* \mathcal{N} V) V^* = V \pi_1(\mathcal{M}) V^* = i(\mathcal{M}),$$

which proves equation (2.47). From equation (2.50) we also have

$$p \pi(f) p = V(V^* \pi(f) V) V^* = V \pi_1(O(f)) V^* = i(O(f)),$$

which proves equation (2.48).

Step 5: Definition of the automorphic representation β of G on \mathcal{N} . The maps $U_g, g \in G$ are defined on the dense subspace $\pi(L^\infty(X, \mu) V \mathcal{H})$ of the Hilbert space \mathcal{H}' by

$$U_g : \pi(f) V \psi \mapsto \pi(\text{Ad} \lambda(g) f) V u_g \psi, \quad f \in L^\infty(X, \mu), \psi \in \mathcal{H}.$$

Using equations (2.50, 2.51) one proves that the operators U_g are well defined, that they leave the scalar product on \mathcal{H}' invariant, that they can be extended to the whole of \mathcal{H}' , and that they fulfil

- (a) $U_g^* = U_{g^{-1}}$
- (b) $U_{g_1 g_2} = U_{g_1} U_{g_2}$
- (c) $U_g \pi(f) U_g^* = \pi(\text{Ad}\lambda(g)f)$, for $f \in L^\infty(X, \mu)$
- (d) $U_g V = V u_g$
- (e) $U_g V \pi_1(x) V^* U_g^* = V u_g \pi_1(x) u_g^* V^*$, for $x \in \mathcal{M}$
- (f) the map $g \mapsto U_g$ is strongly continuous.

Properties (c) and (e) imply that the $U_g \cdot U_g^*$ leave \mathcal{N} globally invariant. Define the automorphism β_g of \mathcal{N} to be the restriction of $U_g \cdot U_g^*$ to \mathcal{N} . Property (f) implies that β is σ -weakly continuous. Property (c) is the covariance property (2.46). Because of property (e) the projector $p := VV^*$ is an element of the fixed point algebra \mathcal{N}^β .

Step 6: p is an atom of \mathcal{N}^β if α acts ergodically on \mathcal{N} . If this were not the case, then there would be a projector $r \in \mathcal{N}^\beta$ such that $r < p$, $r \neq 0$. The projector $prp \in p\mathcal{N}p$ is invariant under the restriction of β to $p\mathcal{N}p$. Therefore $i^{-1}(r)$ is a projector in \mathcal{M} , which is smaller than the identity. Because of property (e) and equation (2.49), $i^{-1}(r)$ is invariant under all α_g . This contradicts ergodicity of α .

Step 7: If α is ergodic and the covariant embedding O is extremal in the set of all covariant embeddings of $L^\infty(X, \mu)$ into (\mathcal{M}, G, α) , then $\pi(L^\infty(X, \mu))$ is maximally abelian in \mathcal{N} : $\pi(L^\infty(X, \mu))' \cap \mathcal{N} = \pi(L^\infty(X, \mu))$.

According to Nakagami and Takesaki (1979, theorem II.2.2), $\pi(L^\infty(X, \mu))^C := \pi(L^\infty(X, \mu))' \cap \mathcal{N}$ is the von Neumann algebra generated by $\pi(L^\infty(X, \mu))$ and the fixed point algebra $(\pi(L^\infty(X, \mu))^C)^\beta = \mathcal{N}^\beta \cap \pi(L^\infty(X, \mu))^C$. Therefore it is sufficient to show that $(\pi(L^\infty(X, \mu))^C)^\beta =$

C1. Let r be a projector in \mathcal{N}^β which commutes with $\pi(L^\infty(X, \mu))$. Then the map $O_r : L^\infty(X, \mu) \rightarrow \mathcal{M}$ defined by

$$O_r(f) := i^{-1}(pr\pi(f)p)$$

is a positive, normal, linear, covariant map from $L^\infty(X, \mu)$ into \mathcal{M} . Ergodicity of α and covariance of O_r imply that $O_r(1) = c\mathbf{1}$, where c is a real number between 0 and 1. From Arveson (1969, theorem 1.4.2) and the extremality of O we deduce that r is either 0 or 1. Since $(\pi(L^\infty(X, \mu))^C)^\beta$ is generated by its projectors, this implies $(\pi(L^\infty(X, \mu))^C)^\beta = \mathbf{C1}$. \square

2.5.2 An imprimitivity theorem

Now I will prove an imprimitivity theorem for transitive generalised systems of covariance. This generalises the imprimitivity theorem for systems of covariance (see p. 115) by Neumann (1972), Scutaru (1977), Cattaneo (1979), Castrigiano and Henrichs (1980), and Ali (1984), because it applies to automorphic group representations and not only to projective unitary ones.

Construction of induced W^ -systems.* In the construction of induced W^* -systems I will follow Takesaki (1973). Let $(\mathcal{N}_0, H, \gamma)$ be a W^* -system of a closed subgroup H of G . \mathcal{N}_0 is isomorphic to a weakly closed subalgebra of bounded operators on a not necessarily separable Hilbert space \mathcal{H}_0 . We consider the tensor-product $L^\infty(G) \otimes \mathcal{N}_0$ of \mathcal{N}_0 and the abelian von Neumann algebra $L^\infty(G)$. The elements of $L^\infty(G) \otimes \mathcal{N}_0$ are regarded as bounded \mathcal{N}_0 -valued functions x on G with the following properties.

1. For all pairs $\xi, \eta \in \mathcal{H}_0$ the function $g \mapsto \langle \xi, x(g)\eta \rangle$ is Haar measurable.

2. dg -ess $\sup_{g \in G} \|x(g)\| < \infty$.

On $L^\infty(G) \otimes \mathcal{N}_0$ define actions γ' of H and β of G by

$$(\gamma'_h(x))(s) := \gamma_h(x(sh)), \quad s \in G, h \in H, x \in L^\infty(G) \otimes \mathcal{N}_0,$$

$$(\beta_g(x))(s) := x(g^{-1}s), \quad g, s \in G, x \in L^\infty(G) \otimes \mathcal{N}_0.$$

Let \mathcal{N} denote the fixed point algebra of $L^\infty(G) \otimes \mathcal{N}_0$ under $\{\gamma'_h : h \in H\}$. Since γ'_h and β_g commute for all $h \in H, g \in G$, \mathcal{N} is invariant under β_g . The restriction of β_g to \mathcal{N} is also denoted by β_g . We say that the W^* -system (\mathcal{N}, G, β) is *induced from the W^* -system $(\mathcal{N}_0, H, \gamma)$* and write $(\mathcal{N}, G, \beta) = \text{Ind}_H^G(\mathcal{N}_0, H, \gamma)$.

Theorem 4 *Let G be a locally compact group, and let H be a closed subgroup of G . Let a be a POV-measure based on G/H , and denote by \mathcal{M} the W^* -algebra $\{a(\Delta) : \Delta \in \Sigma(G/H)\}''$. Let α be a pointwise σ -weakly continuous representation of G as automorphism group of \mathcal{M} . Assume that a is covariant with respect to α ,*

$$\alpha_g(a(\Delta)) = a(g\Delta), \quad g \in G, \Delta \in \sigma(G/H).$$

Then there is

a classical W^ -system (\mathcal{N}, G, β) (see Definition 4);*

a PV-measure E on G/H generating $\mathcal{N} = \{E(\Delta) : \Delta \in \Sigma(G/H)\}''$ and fulfilling the covariance condition

$$\beta_g(E(\Delta)) = E(g\Delta), \quad f \in L^\infty(G/H), g \in G;$$

a projector p in the fixed point algebra \mathcal{N}^β ;

and an isomorphism $i : \mathcal{M} \rightarrow p\mathcal{N}p$ of the W^ -algebras \mathcal{M} and $p\mathcal{N}p$ such that*

$$i(a(\Delta)) = pE(\Delta)p, \quad \Delta \in \Sigma(G/H).$$

Furthermore, there is a W^* -system $(\mathcal{N}_0, H, \gamma)$ such that (\mathcal{N}, G, β) is isomorphic to $\text{Ind}_H^G(\mathcal{N}_0, H, \gamma)$.

PROOF: The POV-measure a gives rise to a covariant embedding O of $L^\infty(G/H)$ into $\mathcal{M} = \{a(\Delta) : \Delta \in \Sigma(G/H)\}''$ by σ -weakly continuous extension of $O(\chi_\Delta) := a(\Delta)$. The previous proposition then yields everything claimed by the theorem, except that (\mathcal{N}, G, β) is an induced W^* -system.

Because of $\mathcal{M} = \{a(\Delta) : \Delta \in \Sigma(G/H)\}''$ we have

$$O(L^\infty(G/H)) = \mathcal{M}.$$

By equation (2.48) this implies that $i(\mathcal{M}) = i(O(L^\infty(G/H))) = p\pi(L^\infty(G/H))p$. \mathcal{N} was defined as the von Neumann algebra generated by $\pi(L^\infty(G/H))$ and by $i(\mathcal{M})$. $i(\mathcal{M}) = p\pi(L^\infty(G/H))p$ implies that

$$\mathcal{N} = \pi(L^\infty(G/H)).$$

Since $\pi(L^\infty(G/H))$ is commutative, it coincides with its centre. Therefore π is an covariant isomorphism from $L^\infty(G/H)$ into the centre of \mathcal{M} . Thus it follows from Takesaki (1973), Theorem 10.5, that there exists a W^* -system $(\mathcal{N}_0, H, \gamma)$ such that (\mathcal{N}, G, β) is isomorphic to $\text{Ind}_H^G(\mathcal{N}_0, H, \gamma)$. \square

An alternative look at phase space quantisation. From this imprimitivity theorem one gains an alternative view on phase space quantisation.

Assume that the phase space Γ is a homogeneous space $\Gamma = G/H$. The algebra of classical observables in a statistical description is $L^\infty(G/H)$. The quantisation map O associates to these classical observables the W^* -algebra \mathcal{M} of quantum observables.

According to Theorem 3, there exists a representation π of $L^\infty(G/H)$ and a projector p such that

$$O(f) = p\pi(f)p.$$

Since $\pi(L^\infty(G/H))$ is isomorphic to $L^\infty(G/H)$, one can equally well regard $\pi(L^\infty(G/H))$ as algebra of classical statistical observables. From this point of view, the transition from the quantum to the classical system is achieved by the map

$$p \cdot p : \pi(f) \mapsto p\pi(f)p.$$

This generalises to more general systems equations (2.36) and (2.37) which described for the Weyl system how the Koopman observables are related to the geometrically quantised and to the Schrödinger observables. The Naimark extension of the covariant embedding (i.e. of the quantisation map) describes the classical system. In this sense the Naimark extension is the inverse of the quantisation map: it leads from the quantum system to the classical system.

Chapter 3

The Classical Limit

Chapter abstract

The emergence of classical observables in infinite quantum systems is examined. The framework and assumptions for this emergence are discussed in general terms. Then we deal with the simultaneous measurability of position and momentum in infinite quantum systems. Positive-operator-valued measures transforming covariantly under velocity boosts and space translations are regarded as joint position-momentum observables. It is shown that in the finite Weyl system there exist only unsharp joint position-momentum observables, whereas in the infinite Weyl system there are sharp joint position-momentum observables.

Chapter overview

Sometimes it is said that “in the classical limit” $\hbar \rightarrow 0$ quantum mechanics leads to classical mechanics. One way in which this statement can be regarded as true is that the commutators of quantum mechanical observables are usually proportional to \hbar . Therefore, if $\hbar = 0$ all observables commute, and we are in the classical situation.

In another sense, it is very misleading to say that classical mechanics

emerges from quantum mechanics in the limit $\hbar \rightarrow 0$. Many of the conceptual difficulties in the interpretation of quantum mechanics are fully present for very small $\hbar > 0$, but suddenly disappear for $\hbar = 0$. This “discontinuous” behaviour makes the limit very problematic from a conceptual point of view. Furthermore, \hbar is a constant in our world and not a variable parameter. Investigations of how the classical situation is gradually recovered as $\hbar \rightarrow 0$ are therefore mathematical in nature rather than physical.

I prefer to think of the classical limit as the description of situations in which quantum mechanical systems display classical behaviour. This is possibly, but not necessarily, the case if the masses or the number of degrees of freedom involved get very large. But there might also be other circumstances in which the quantum mechanical systems behave classically. I will leave these situations out of consideration. Instead I will deal just with the behaviour of macroscopic quantum systems.

In section 3.1 I introduce classical observables as those which commute with all other observables. In order to motivate this definition, it is discussed in Propositions 4-6 how such observables display properties which are usually associated to classical systems. The introductory section 3.1 closes with a discussion of von Neumann’s (1938) work on the infinite tensor product. Von Neumann’s results can be interpreted as implying the emergence of classical observables in infinite quantum systems. But such an interpretation relies on choosing the quasilocal observables to be the algebra of observables.

In section 3.2 I discuss the simultaneous measurability of position and momentum in elementary, finite, and infinite quantum systems. Joint position-momentum observables are defined to be covariant POVM-measures on the group $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ of space translations and velocity

boosts. For elementary or finite quantum systems, it is shown that there is no covariant PV-measure on $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ into the von Neumann algebra generated by average position and momentum. So these cannot be measured sharply at the same time. For an infinite quantum system, we construct a covariant PV-measure on $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ into the von Neumann algebra generated by average position and momentum. Therefore average position and momentum of an infinite quantum system can be measured sharply at the same time. Section 3.2 follows Breuer (1994c).

3.1 Classical observables in quantum mechanics

We now introduce classical observables as those which commute with all other observables. This is done first in the Hilbert space formalism, and then in the algebraic formalism.

3.1.1 Hilbert space formulation

In the Hilbert space formalism of quantum mechanics superselection rules are present if the Hilbert space decomposes into subspaces which are not connected by observables. This is done in the following way.

Assume that the Hilbert space describing a quantum system is a

discrete¹ direct sum of subspaces

$$\mathcal{H} = \bigoplus_{\substack{i \in I \\ j \in J_i}} \mathcal{H}_{ij},$$

where I and $J_i, i \in I$ are index sets, and the subspaces $\mathcal{H}_{ij}, \mathcal{H}_{i'j'}$ are isomorphic for $i = i'$. (They also may, but need not, be isomorphic for $i \neq i'$.) Assume furthermore that the observables of the quantum system are of the form

$$A = \bigoplus_{i \in I} \left(\bigoplus_{j \in J_i} x_{ij} \right), \quad x_{ij} = x_{i'j'} \text{ if } i = i',$$

where $x_{ij} \in \mathcal{B}(\mathcal{H}_{ij})$. Denote the set of such operators A by \mathcal{M} . \mathcal{M} leaves the subspaces \mathcal{H}_{ij} invariant.

Obviously we have

$$\langle \psi_{ij} | A | \psi_{i'j'} \rangle = 0 \tag{3.1}$$

for all observables $A \in \mathcal{M}$, whenever $\psi_{ij} \in \mathcal{H}_{ij}, \psi_{i'j'} \in \mathcal{H}_{i'j'}$ and $i \neq i'$ or $j' \neq j'$. This is usually interpreted by saying that a transition between the states $\psi_{ij}, \psi_{i'j'}$ is not observable, and that no observable operation can transform ψ_{ij} into a state from which there would be a non-vanishing transition probability to $\psi_{i'j'}$. The Hilbert spaces \mathcal{H}_{ij} describe the superselection sectors of the theory.

Let Ψ be any unit vector in \mathcal{H} . Ψ can be written as

$$\Psi = \bigoplus_{\substack{i \in I \\ j \in J_i}} c_{ij} \psi_{ij}, \quad \psi_{ij} \in \mathcal{H}_{ij},$$

¹The assumption that I is discrete is necessary for the following reason. If I is discrete, then the projection lattice of \mathcal{M} is atomic. (A lattice is called atomic if every element is the join of minimal elements.) Then these minimal elements are unique and \mathcal{M} as a whole is uniquely a direct sum of irreducible factors. Thus the subspaces \mathcal{H}_{ij} are uniquely determined. This would not be the case if I were continuous. If the centre of I were not discrete (i.e. if we had a direct integral instead of a direct sum), we could not even be sure that \mathcal{M} has a non-trivial centre. See e.g. Dixmier (1964b) and Mackey (1978, chaps 7,8).

3.1. CLASSICAL OBSERVABLES IN QUANTUM MECHANICS 149

where $\|\psi_{ij}\| = 1$ and the c_{ij} are properly normalised coefficients. Then, because of (3.1), for any observable A we have

$$\begin{aligned} \langle \Psi | A \Psi \rangle &= \left\langle \bigoplus_{\substack{i \in I \\ j \in J_i}} c_{ij} \psi_{ij} \left| \bigoplus_{\substack{i \in I \\ j \in J_i}} x_i \right| \bigoplus_{\substack{i \in I \\ j \in J_i}} c_{ij} \psi_{ij} \right\rangle \\ &= \sum_{\substack{i \neq i' \\ j \neq j'}} \langle c_{ij} \psi_{ij} | x_{i'j'} | c_{i'j'} \psi_{i'j'} \rangle \\ &\quad + \sum_{\substack{i \in I \\ j \in J_i}} \langle c_{ij} \psi_{ij} | x_{ij} | c_{ij} \psi_{ij} \rangle \\ &= \sum_{\substack{i \in I \\ j \in J_i}} |c_{ij}|^2 \langle \psi_{ij} | x_{ij} | \psi_{ij} \rangle. \end{aligned}$$

Therefore all observables A have the same expectation values in the states Ψ and $\sum_{ij} |c_{ij}|^2 |\psi_{ij}\rangle \langle \psi_{ij}|$. If one defines states to be positive normalised linear functionals on the observables, then Ψ and $\sum_{ij} |c_{ij}|^2 |\psi_{ij}\rangle \langle \psi_{ij}|$ are the same states. Thus the vector Ψ describes a convex mixture of states ψ_{ij} . A vector $\Psi \in \mathcal{H}$ describes a *pure* state on \mathcal{M} only if all coefficients c_{ij} vanish except one.

Operators of the form

$$S = \bigoplus_{\substack{i \in I \\ j \in J_i}} s_{ij} \mathbf{1}_{ij}, \quad s_{ij} \in \mathbf{C}$$

commute with all observables $A \in \mathcal{M}$. The operators S are not in \mathcal{M} unless $s_{ij} = s_{i'j'}$ for $i = i'$. Such operators $S \notin \mathcal{M}$ do not describe observables. But operators of the form

$$Z = \bigoplus_{i \in I} \left(\bigoplus_{j \in J_i} z_i \mathbf{1}_{ij} \right), \quad z_i \in \mathbf{C}$$

commute with all observables and are themselves observables.

Every pure state is an eigenstate of such operators Z, S : $\langle \Psi | S | \Psi \rangle = s_{i_0 j_0} \Psi$ if all c_{ij} except $c_{i_0 j_0}$ vanish. Thus the values of the observables

Z in any pure state is dispersion-free. Also, the operators S associate to each vector in \mathcal{H}_{ij} the same number s_{ij} . The number s_{ij} associated by S to a subspace \mathcal{H}_{ij} plays the rôle of a quantum number, but is in general not an observable.

Let us now compare this to classical mechanics. The continuous functions on the phase space Γ vanishing at infinity form a C^* -algebra of classical observables. Define the states to be the positive normalised linear functionals on these functions. The convex combination of states is again a state. Call a state pure if it can not be written as a non-trivial convex combination of other states. It can be shown that there is a one-to-one relation between the pure states and the points $\omega \in \Gamma$. The value of a pure state ω on a classical observable f is given by $\omega(f) = f(\omega)$. Since $\omega(f^2) = \omega(f)^2$ the value of all observables f in all pure states ω is dispersion-free.

This is exactly the situation encountered for the observables Z of the quantum mechanical system. The points $i \in I$ are the quantum analogue of the points $\omega \in \Gamma$. So the following definition seems justified.

Definition 18 *Assume that the algebra of observables is a C^* - or a W^* -algebra of operators on a Hilbert space. The minimal non-trivial projections in the commutant describe superselection sectors. A superselection operator which is itself an observable is called a classical observable.*

In this terminology, every classical observable is described by a superselection operator, but superselection operators need not be observable. If the algebra of observables acts irreducibly on the Hilbert space, then there is only one superselection sector, and the only classical observables are the constant multiples of the identity operator. If the centre of the algebra of observables contains only multiples of the

identity, then the algebra is called a factor. In this case there are only trivial classical observables. But since a factor may be reducible, there may be non-trivial superselection operators.

For example, permutations in a system of n identical particles should not be observable. If the observables are represented as operators on the full (i.e. not symmetrised) n -fold tensor product Hilbert space, then they are required to commute with the unitary operators implementing permutations. These unitaries must not themselves be observables, because permutations should not be observable. There is another reason why the permutations cannot be observable: If they commute with all observables and were themselves observables they would have to commute with each other. But for more than two elements the permutation group is not abelian. Thus the unitaries implementing permutation on the full tensor product lead, in the above terminology, to superselection operators but not to classical observables. (What one does in this case is, of course, to take the symmetric or antisymmetric tensor product Hilbert space instead of the full tensor product.)

3.1.2 Algebraic formulation

In the algebraic formulation of quantum mechanics, the distinction between superselection operators and classical observables becomes unnecessary. In this formulation one assumes all physically relevant information to be encoded in an abstract C^* - or W^* -algebra \mathcal{A} . ‘Abstract’ means that the algebra need not be an algebra of operators on a Hilbert space. Of course, the abstract algebra is always isomorphic to an algebra of operators on *some* Hilbert space, but this Hilbert space is usually non-separable. To recover the Hilbert space formalism one takes a representation of the abstract algebra on a separable Hilbert space. Under

certain circumstances *inequivalent* representations are possible. The occurrence of inequivalent representations is closely linked to the existence of classical observables. I will come back to this later.

Now, let us see why the distinction between superselection operators and classical observables becomes unnecessary in algebraic quantum mechanics.

Definition 19 *Let us say that two representations π_1, π_2 of the abstract algebra of observables \mathcal{A} are physically equivalent if the statistical states in the two representations can be identified. By this I mean that for every given density matrix D_1 on \mathcal{H}_1 there exists a density matrix D_2 on \mathcal{H}_2 such that*

$$\text{tr}(D_1\pi_1(A)) = \text{tr}(D_2\pi_2(A))$$

for all $A \in \mathcal{A}$.

But if this is the case, then there exists a W^* -isomorphism between the von Neumann algebras $\pi_1(\mathcal{A})''$ and $\pi_2(\mathcal{A})''$.² (Here I denote again by $\pi_1(\mathcal{A})''$ the double commutant of $\pi_1(\mathcal{A})$. According to von Neumann's double commutant theorem, $\pi_1(\mathcal{A})''$ coincides with the closure of $\pi_1(\mathcal{A})$ in the weak operator topology.) This, however, is the case if and only if π_1 and π_2 are quasiequivalent, i.e. if they are both direct sums of the same irreducible representation. Thus representations are *physically equivalent* if and only if they are unitarily equivalent *up to multiplicities*.

How can we see that in the algebraic formulation, the distinction between superselection operators and classical observables becomes unnecessary? Let us consider the situation where we do not have any classical observables but still have several superselection sectors. (This

²See Bratteli and Robinson 1987, 2.4.26.

3.1. CLASSICAL OBSERVABLES IN QUANTUM MECHANICS 153

is the case if, in the notation used above, I consists of only one element i , but J_i has more than one element.) In this case the observables form a reducible factor \mathcal{M} . \mathcal{M} decomposes into irreducible factors. Since two factors are either quasi-equivalent or disjoint, each of the irreducible factors is quasiequivalent to \mathcal{M} . Thus there is a W^* -isomorphism between each of the irreducible factors and \mathcal{M} . Therefore, according to Definition 19, the irreducible factors contain exactly the same physical information as \mathcal{M} . (In the example of p. 147, the operators $A_i = \bigoplus_{j \in J_i} x_{ij}$ act reducibly on $\bigoplus_{j \in J_i} \mathcal{H}_{ij}$ but they form a factor \mathcal{M}_i . The W^* -isomorphism between the irreducible factors $\mathcal{B}(\mathcal{H}_{ij})$ and the reducible factor \mathcal{M}_i is given by $x \mapsto \bigoplus_{j \in J_i} x$.)

From this point of view, irreducibility is a representation dependent property without physical significance. Superselection rules which do not correspond to classical observables are described by *reducible* factors. But since the irreducibility of a factor is without physical significance, the same is true for superselection rules which do not correspond to classical observables.

We are thus led to the following definition.

Definition 20 *Let \mathcal{A} be the abstract C^* -algebra of observables of a physical system, and let $\mathcal{A}^{**} \supset \mathcal{A}$ be the space of continuous linear functional on the states. The non-trivial elements of the centre of \mathcal{A}^{**} describe the classical observables of the system. States on which a classical observable takes different values are separated by a superselection rule.*

In this definition we did not make the assumption that the classical observables have to be discrete. This will always be fulfilled because classical observables are taken as elements of \mathcal{A}^{**} , which can only be represented faithfully on a non-separable Hilbert space. On this Hilbert

space: all pure states, normal and non-normal, are represented by vectors; and all observables, even those with an uncountable range of possible values, have a discrete spectrum.

Note that it may happen that \mathcal{A} has trivial centre, whereas \mathcal{A}^{**} does not. (This is for example the case for the quasilocal algebra of an infinite quantum system, see sections 3.1.3 and 3.2.4.) In this case non-trivial central elements can only be found in the weak closure of particular representations. Is therefore the concept of classical observable representation dependent and without physical significance, similar to the notion of irreducibility?

I will argue that this is not the case. Note first that the above definition of classical observables does not make reference to any particular representation. (But certainly whether and which classical observables emerge in a particular representation depends on the representation.) Since representations π_1, π_2 corresponding to states in which some $Z \in \mathcal{A}^{**}$ has a different value are not quasi-equivalent, there is no W^* -isomorphism between $\pi_1(\mathcal{A})''$ and $\pi_2(\mathcal{A})''$. Thus the representations π_1 and π_2 are *not* physically equivalent in the sense of Definition 19. Therefore the choice of representation *is* of physical significance if classical observables are involved. Only if no classical observables are involved, the choice of representation is of no physical significance.

Properties of classical observables. I will now mention some elementary properties of classical observables. Neither of the following propositions is new or technically difficult. I display them as Propositions only to stress their conceptual importance.

Proposition 4 *If ρ is a pure state on a C^* -algebra \mathcal{A} , and Z is in*

3.1. CLASSICAL OBSERVABLES IN QUANTUM MECHANICS 155

the centre of \mathcal{A}^{**} , then $Z^2(\rho) = (Z(\rho))^2$. Thus the classical observables have dispersion-free expectation values in all pure states. Also, classical observables have dispersion free expectation values in all states which are linear superpositions of pure states with the same expectation values on all classical observables.

PROOF: Let π_ρ be the GNS-representation of \mathcal{A} with respect to the state ρ . There is a normalised vector $\Omega \in \mathcal{H}_\rho$ such that for all $A \in \mathcal{A}$ we have $\rho(A) = \langle \Omega, \pi_\rho(A)\Omega \rangle$. (See e.g. Bratteli and Robinson (1987, Theorem 2.3.16)). Since \mathcal{A}^{**} is canonically isomorphic to the weak closure of the universal representation (Dixmier (1977, 12.1.3), and π_ρ is a subrepresentation of the universal representation, for every $A \in \mathcal{A}^{**}$ there is a $x \in \pi_\rho(\mathcal{A})''$ such that $A(\rho) = \langle \Omega, x\Omega \rangle$. Denote this x by $\pi_\rho(A)$.

Since ρ is pure, $\mathcal{M} := \pi_\rho(\mathcal{A})''$ is irreducible (see Bratteli and Robinson (1987, theorem 2.3.19)) and therefore a factor. \mathcal{M}' and the centre of \mathcal{M} consist only of multiples of the identity. Since Z is in the centre of \mathcal{A}^{**} , $\pi_\rho(Z)$ is in the centre of \mathcal{M} and therefore $\pi_\rho(Z) = z\mathbf{1}$.

We thus have $Z(\rho) = \langle \Omega, z\mathbf{1}\Omega \rangle = z\langle \Omega, \Omega \rangle$ and $Z^2(\rho) = \langle \Omega, z^2\mathbf{1}\Omega \rangle = z^2\langle \Omega, \Omega \rangle$. Thus $Z^2(\rho) = Z(\rho)^2$. \square

Proposition 5 *If ρ is an arbitrary state on a C^* -algebra \mathcal{A} with identity, and Z is in the centre of \mathcal{A}^{**} , then for all observables $A \in \mathcal{A} \subset \mathcal{A}^{**}$ we have $AZ(\rho) = A(\rho)Z(\rho)$. Thus in all states the joint expectation values of classical observables and arbitrary observables factorise.*

PROOF: Let π_ρ be again the GNS-representation of \mathcal{A} with respect to the state ρ . There is a normalised vector $\Omega \in \mathcal{H}_\rho$ such that for all $A \in \mathcal{A}$ we have $\rho(A) = \langle \Omega, \pi_\rho(A)\Omega \rangle$. For $A \in \mathcal{A}^{**}$, denote again by $\pi(A)$ the $x \in \pi_\rho(\mathcal{A})''$ such that $A(\rho) = \langle \Omega, x\Omega \rangle$.

Since Ω is normalised, we have $\langle \Omega, \Omega \rangle = 1$. Because Z is in the centre of \mathcal{A}^{**} and the projector $|\Omega\rangle\langle\Omega|$ is in $\pi_\rho(A)$, we have $[\pi_\rho(Z), |\Omega\rangle\langle\Omega|] = 0$. Thus

$$\begin{aligned} A(\rho)Z(\rho) &= \langle \Omega | \pi_\rho(A) | \Omega \rangle \langle \Omega | \pi_\rho(Z) | \Omega \rangle \\ &= \langle \Omega | \pi_\rho(A) \pi_\rho(Z) | \Omega \rangle \langle \Omega | \Omega \rangle \\ &= \langle \Omega | \pi_\rho(AZ) | \Omega \rangle \\ &= AZ(\rho). \end{aligned}$$

□

Let ρ and ϕ be states on a C^* -algebra \mathcal{A} . Following Roberts and Roepstorff (1969), define the transition probabilities between pure³ states ρ and ϕ by

$$p(\rho, \phi) := 1 - \frac{1}{4} \|\phi - \rho\|^2,$$

where $\|\phi - \rho\| := \sup_{A \in \mathcal{A}} |\rho(A) - \phi(A)| / \|A\|^2$ is the norm distance between the states ρ and ϕ . (Note that for two unit vectors in the Hilbert space, $\|\phi - \rho\|$ is not the length of the vector $\rho - \phi$.)

If the pure states ρ, ϕ assign the same expectation values to all classical observables, then this definition reduces to the traditional definition of transition probabilities: In this case there is an irreducible

³Roberts and Roepstorff (1969) explicitly resist the temptation to use the same formula to define transition probabilities between mixed states. They give the following argument: The transition probability from a mixed state with density matrix ρ into a pure state represented by a vector ψ could be defined as the proportionality constant $(P_\psi \rho P_\psi) / P_\psi$, where $P_\psi := |\psi\rangle\langle\psi|$. $P_\psi \rho P_\psi$ and P_ψ are always proportional, irrespective of the choice of ρ . But if the transition is into a mixed state ϕ , then there is no observable $P(\phi)$ such that $P(\phi)\rho P(\phi)$ is proportional to $P(\phi)$, independent of the choice of ρ . Thus although one can define transition probabilities from mixed states into pure states, it is not possible to do so for transitions into mixed states. In order to preserve the symmetry between initial and final state, Roberts and Roepstorff refrain also from defining transition probabilities from mixed into pure states.

3.1. CLASSICAL OBSERVABLES IN QUANTUM MECHANICS 157

representation π of \mathcal{A} and two vectors $\Omega_\rho, \Omega_\phi \in \mathcal{H}_\pi$ such that $\rho(A) = \langle \Omega_\rho | \pi(A) | \Omega_\rho \rangle$ and $\phi(A) = \langle \Omega_\phi | \pi(A) | \Omega_\phi \rangle$.⁴ Then⁵ one finds that

$$p(\rho, \phi) = |\langle \Omega_\phi | \Omega_\rho \rangle|^2.$$

Following Haag and Kastler (1964), one can define an observable operation to be a state transformation induced by an observable $B \in \mathcal{A}$ via

$$\rho \mapsto \rho_B(A) := \begin{cases} \frac{\rho(B^*AB)}{\rho(B^*B)} & \text{for all } A \in \mathcal{A} \text{ if } \rho(B^*B) > 0, \\ \rho(A) & \text{for all } A \in \mathcal{A} \text{ if } \rho(B^*B) = 0. \end{cases}$$

(In case \mathcal{A} has a quasilocal structure, ρ_B is called a local perturbation of the state ρ .) Note that the states ρ, ρ_B are quasi-equivalent. In the Hilbert space formalism, if we choose $\mathcal{A} = \mathcal{B}(\mathcal{H})$, and represent a pure state ρ by a unit vector ψ , then ρ_B is represented by the unit vector $B\psi / \|B\psi\|^2$ for any B with $B\psi \neq 0$.

Referring to the quantity $p(\rho, \phi)$ as “transition probability” contains the implicit assumption that the time evolution can only be induced by observables. This assumption is questionable. But since the term is firmly established, I will continue to use it, keeping in mind that under certain circumstances the time evolution can bring about transitions not describable as $\rho \mapsto \rho_A$ for any $A \in \mathcal{A}$.

Proposition 6 *If ρ, ϕ are pure states on \mathcal{A} with $Z(\rho) \neq Z(\phi)$ for some classical observable $Z \in \mathcal{A}^{**}$, then the transition probability between ρ and ϕ vanishes. Also, no observable operation can transform ρ into a state with non-vanishing transition probability into ϕ : $p(\rho_A, \phi) = 0$ for all $A \in \mathcal{A}$.*

⁴See e.g. Roberts and Roepstorff 1969, Proposition 4.2.

⁵See e.g. Roberts and Roepstorff 1969, Proposition 4.6.

PROOF: Assume that $p(\rho_A, \phi) > 0$ for some $A \in \mathcal{A}$. Thus $\|\rho_A - \phi\| < 2$. Hence (see Glimm and Kadison 1960, Corollary 9) ρ_A and ϕ give rise to unitarily equivalent irreducible representations π_{ρ_A} and π_ϕ of \mathcal{A} . Since the π_{ρ_A} and π_ρ are quasi-equivalent, so are π_ρ and π_ϕ . This is in contradiction to the disjointness of ρ, ϕ which is a consequence of $Z(\rho) \neq Z(\phi)$. \square

These properties of classical observables are nice enough to justify the predicate “classical” for them. But their nice behaviour only reflects the strong condition we imposed on them. The condition that the centre of \mathcal{A}^{**} be non-trivial is so strong that it is violated in traditional Hilbert space quantum mechanics: the algebra $\mathcal{B}(\mathcal{H})$ has trivial centre.

So far I only presented the mathematical formalism which could describe a system with quantum and classical properties. But the real world could be one without superselection rules or classical observables. Classical observables exist if finite systems are not correctly described by algebras $\mathcal{B}(\mathcal{H})$, but rather by non-simple algebras of observables. But even if this is not the case and finite systems are fully quantum mechanical, infinite quantum systems may have classical properties. In the rest of this chapter I will focus on the superselection rules arising for infinite systems. In doing so I neglect the superselection rules like charge⁶ or mass⁷ which can already be present in finite systems.

3.1.3 The infinite tensor product

The main reasons for the emergence of superselection rules in infinite quantum systems lie in the natural structure of the infinite tensor product Hilbert space. Following the classical paper of von Neumann (1938),

⁶See e.g. Wick, Wigner, Wightman (1952).

⁷See Bargman (1954).

3.1. CLASSICAL OBSERVABLES IN QUANTUM MECHANICS 159

I will quickly describe this structure. For simplicity I will consider the case where all factor Hilbert spaces are identical.

Let Π be an infinite index set, not necessarily countable. Denote by $\{\psi_n\}_{n \in \Pi}$, or simply by $\{\psi\}$, an infinite sequence (or more generally a net) of vectors $\psi_n \in \mathcal{H}$. The scalar product of the two sequences $\{\psi\}, \{\rho\}$ is defined by

$$\langle \{\psi\}, \{\rho\} \rangle := \prod_{n \in \Pi} \langle \psi_n, \rho_n \rangle = \prod_{n \in \Pi} \exp(i\phi_n) |\langle \psi_n, \rho_n \rangle|, \quad (3.2)$$

where the $\phi_n := \arg \langle \psi_n, \rho_n \rangle$ and it is assumed that $\sum_{n \in \Pi} |\phi_n| < \infty$. If $\sum_{n \in \Pi} |\phi_n| = \infty$, i.e. if the phases do not converge, we define the scalar product to be zero. The scalar product $\langle \{\psi\}, \{\rho\} \rangle$ is also zero if $\langle \psi_n, \rho_n \rangle = 0$ for some n or if $\prod_{n \in \Pi} |\langle \psi_n, \rho_n \rangle| = 0$.

One then forms the quotient space with respect to zero sequences. As zero sequences one regards those for which $\prod_{n \in \Pi} \langle \psi_n, \psi_n \rangle$ converges to zero, in particular those for which one factor is zero. On the quotient space define a separating norm by

$$\langle \{\psi\}, \{\psi\} \rangle := \prod_{n \in \Pi} \langle \psi_n, \psi_n \rangle. \quad (3.3)$$

The quotient space can then be completed to a Hilbert space, which I will denote by $\otimes^\infty \mathcal{H}$. The Hilbert space $\otimes^\infty \mathcal{H}$ is not separable.

On $\otimes^\infty \mathcal{H}$ we define weak and strong equivalence classes

$$C\{\psi\} := \{ \{\psi'\} : \sum_n |\langle \psi_n, \psi'_n \rangle - 1| < \infty \}$$

$$C_w\{\psi\} := \{ \{\psi'\} : \sum_n ||\langle \psi_n, \psi'_n \rangle| - 1| < \infty \}.$$

Completing these equivalence classes with respect to the norm (3.3) we obtain a separable Hilbert space $\mathcal{H}_{C\{\psi\}}$ and a nonseparable Hilbert

space $\mathcal{H}_{C_w\{\psi\}}$. The strong $\mathcal{H}_{C\{\rho\}}$ are sometimes called incomplete infinite tensor product Hilbert space. They are only defined with respect to a sequence $\{\psi\}$. Any weak $\mathcal{H}_{C_w\{\psi\}}$ is the uncountable direct sum of all strong $\mathcal{H}_{C\{\psi'\}}$ for which $\sum_n ||\langle\psi_n, \psi'_n\rangle| - 1| < \infty$ but $\sum_{n \in \Pi} |\phi_n| = \infty$ (and the $\{\psi'\}$ are mutually strongly inequivalent). If Π is finite, there is only one weak and strong equivalence class. In this case the whole construction collapses to the unique finite tensor product of Hilbert spaces.

We can pass from one strong $\mathcal{H}_{C\{\psi\}}$ to another $\mathcal{H}_{C\{\psi'\}}$ in the same weak $\mathcal{H}_{C_w\{\psi\}}$ by a unitary operator $U(\{z\})$ defined by

$$U(\{z\})\{\psi\} := \bigotimes_{n \in \Pi} z_n \{\psi_n\},$$

where $|z_n| = 1, \sum_n |\arg z_n| = \infty$. But no $U(\{z\})$ can take a vector from a weak equivalence class into another one. The vectors $U(\{z\})\{\psi\}$ and $\{\psi\}$ belong to the same strong equivalence class if and only if $\sum_n |\arg z_n| < \infty$. The complete tensor product Hilbert space $\bigotimes^\infty \mathcal{H}$ is the direct sum of all weak $\mathcal{H}_{C_w\{\rho\}}$.

The emergence of strong $\mathcal{H}_{C\{\psi\}}$ and weak $\mathcal{H}_{C_w\{\psi\}}$ is a consequence of the convention to define $\{\psi, \rho\} = \prod_{n \in \Pi} \langle\psi_n, \rho_n\rangle$ to be zero if $\prod_{n \in \Pi} |\langle\psi_n, \rho_n\rangle|$ converges but the phases do not converge, $\sum_{n \in \Pi} |\arg \langle\psi_n, \rho_n\rangle| = \infty$. It would be possible, to be more restrictive and insist on the convergence of both moduli and phases. Then we would arrive at an infinite tensor product Hilbert space $\mathcal{H}_{C\{\psi\}}$ depending on a reference sequence $\{\psi\}$. This would have the advantage that one works only on separable Hilbert spaces. On the other hand, the definition of the infinite tensor product would not be unique any more, but rather depend on a reference sequence. In being less restrictive and admitting not one but all $\mathcal{H}_{C\{\psi\}}$ we obtain the complete infinite ten-

3.1. CLASSICAL OBSERVABLES IN QUANTUM MECHANICS 161

tensor product $\otimes^\infty \mathcal{H}$. This provides for a definition of the infinite tensor product which is independent of reference sequences.

The quasilocal algebra. Consider a system of n particles, each with algebra of observables $\mathcal{B}(\mathcal{H}_i)$. An observable x_i of the i -th particle can be represented on the Hilbert space $\otimes_{i=1}^n \mathcal{H}_i$ by

$$\iota x_i := \mathbf{1}_1 \otimes \mathbf{1}_2 \otimes \dots \otimes \mathbf{1}_{i-1} \otimes x_i \otimes \mathbf{1}_{i+1} \otimes \dots \otimes \mathbf{1}_n.$$

The C^* -algebra and the W^* -algebra generated by the one particle observables ιx_i is $\mathcal{B}(\otimes_{i=1}^n \mathcal{H}_i)$. Thus *the observables of the n particle system is the C^* - or W^* -algebra generated by the one particle observables.*

Let us make the same assumption for the infinite system: take as algebra of observables the C^* - or W^* -algebra generated by the one particle observables. This coincides with the algebra generated by the observables relating to finitely many particles. Call the C^* -algebra generated by the one particle observables \mathcal{A} . It is the norm closed algebra of operators on the infinite tensor product Hilbert space. Let \mathcal{A}'' be its weak (= strong) closure. \mathcal{A} is called the algebra of quasilocal observables.

An element of \mathcal{A} sends no vector $\{\psi\}$ of $\otimes^\infty \mathcal{H}$ out of its strong equivalence class: It can be approximated in norm by operators x with only finitely many entries different from $\mathbf{1}$, and $x\{\psi\}$ is in the same strong equivalence class as $\{\psi\}$. Therefore every strong $\mathcal{H}_{C\{\psi\}}$ (and also every weak $\mathcal{H}_{C_w\{\psi\}}$) is a subspace of $\otimes^\infty \mathcal{H}$ invariant under \mathcal{A} , and also under \mathcal{A}'' .

What is more, von Neumann (1938, theorems IX, X) showed that \mathcal{A}'' consist exactly of those bounded operators on $\otimes^\infty \mathcal{H}$ which leave all strong $\mathcal{H}_{C\{\psi\}}$ invariant and commute with all $U(\{z\})$, where $\{z_n\}_{n \in \mathbb{N}}$

is an arbitrary sequence of complex numbers of modulus one. Thus a bounded operator on $\otimes^\infty \mathcal{H}$ is in \mathcal{A}'' if and only if

1. it leaves all weak $\mathcal{H}_{C_w\{\psi\}}$ invariant, and
2. within the weak $\mathcal{H}_{C_w\{\psi\}}$ it leaves all strong $\mathcal{H}_{C\{\psi\}}$ invariant and *its behaviour on all $\mathcal{H}_{C\{\psi\}}$ in a weak equivalence class is the same.*

One can define representations of the C^* -algebra \mathcal{A} by restriction to weak and strong subspaces. For $x \in \mathcal{A}$ one takes

$$\pi_{C\{\psi\}}(x) := x|_{\mathcal{H}_{C\{\psi\}}},$$

$$\pi_{C_w\{\psi\}}(x) := x|_{\mathcal{H}_{C_w\{\psi\}}}.$$

From the above conditions on x it follows that all representations $\pi_{C\{\psi\}}(\mathcal{A})''$ are irreducible and that all representations $\pi_{C_w\{\psi\}}(\mathcal{A})''$ are reducible factor representations. The projectors onto the weak subspaces are in the centre of \mathcal{A}'' . In the Hilbert space sense of Definition 18, the strong and weak subspaces are the superselection sectors. In the algebraic sense of Definition 20 only the weak subspaces are superselection sectors.

We have seen that choosing for the observables an algebra which is generated by the one particle observables leads to the emergence of classical observables in an infinite system but not in any finite one. Alternatively it would be conceivable, although this is not usually done in the theory of infinite quantum systems, to choose for the observables the algebra of all bounded operators on the tensor product Hilbert space $\otimes^\infty \mathcal{H}$. Then even in the infinite system there are no classical observables. In this case the fact that $\mathcal{A} \subset \mathcal{B}(\otimes^\infty \mathcal{H})$ would be interpreted by saying that for infinite systems there are observables which cannot

be approximated by observables of finite subsystems, or, equivalently, by sums and products of one particle observables.

Thus in the description of infinite quantum systems one faces the choice of whether to interpret $\mathcal{A} \subset \mathcal{B}(\otimes^\infty \mathcal{H})$ either as emergence of classical observables or as the existence of observables which cannot be approximated by observables of finite systems. In the first case one would choose \mathcal{A} or \mathcal{A}'' as algebra of observables, in the second case $\mathcal{B}(\otimes^\infty \mathcal{H})$. I do not know of anybody opting for the second alternative. It seems that the phenomenon of having observables which cannot be approximated by observables of elementary systems is too much at odds with some reductionist preconception to be acceptable. Rather the emergence of classical observables in infinite quantum systems is accepted.

3.2 Simultaneous Sharp Measurability of Position and Momentum in Infinite Quantum Systems

Position and momentum of an elementary quantum system obey (on a dense domain of common self-adjointness) the commutation relations

$$[P, Q] = i\hbar \mathbf{1}. \quad (3.4)$$

Therefore the accuracy of joint position and momentum measurements is bound from below by $\Delta p \Delta q \geq \hbar/2$. Average position and momentum Q^n, P^n of an n -particle quantum system obey the commutation relations

$$[P^n, Q^n] = \frac{i\hbar}{n} \mathbf{1}. \quad (3.5)$$

In the limit of an infinite quantum system it is clear that the right hand side of (3.5) tends to zero. But it is not immediately clear how to interpret the left hand side because the norm limits of the operators P^n, Q^n do not exist.

A solution to this problem is provided in algebraic quantum mechanics. The limits of the averages P^n, Q^n exist in the strong operator topology and are global observables. As such they commute with all other observables and are classical.

In this section we treat the problem of joint measurability of position and momentum in the framework of covariant POV-measures sketched in Chapter 2.⁸ It is necessary to use this more flexible framework because in the traditional (i.e. von Neumann's) formalism non-commuting observables do not have joint probability distributions. So in von Neumann's formalism it is not even possible to talk about joint measurability of position and momentum. In the framework of covariant POV-measures, non-commuting observables can have *unsharp* joint probability measures.⁹ In this framework it is possible to describe joint measurements of non-commuting observables, as long as the measurement is sufficiently unsharp.

In section 3.2.1 a joint position-momentum observable is defined to be a covariant POV-measure on the additive group $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$. Then it is shown that there is no covariant PV-measure on $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ into the von Neumann algebra generated by average position and momentum of elementary (section 3.2.2) and finite (section 3.2.3) quantum systems. So these cannot be measured sharply at the same time. In section 3.2.4

⁸POV-measures were introduced into the mathematical literature by Davies and Lewis (1970), and Ludwig (1967, 1968, 1972). For a review of the theory of covariant POV-measures see Ali 1985.

⁹See Busch and Lahti 1986, Busch 1987.

we construct the strong operator limits of the average position and momentum observables. Also, we construct a covariant PV-measure on $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ into the von Neumann algebra generated by the average position and momentum of infinite quantum systems. This means that the average position and momentum of an infinite quantum system can be measured sharply at the same time and are classical observables. The main argument of this section follows Breuer (1994c).

3.2.1 Joint position-momentum observables

Let us start with an

EXAMPLE 1: Let $\{U_a, V_b : a, b \in \mathbb{R}^3\}$ be an irreducible σ -weakly measurable unitary representation of the Weyl relations $U_a V_b = \exp(iab)V_b U_a$ on the Hilbert space $L^2(\mathbb{R}^3)$ (ab denotes the scalar product of the vectors $a, b \in \mathbb{R}^3$). U_a represents the space translations and V_b the velocity boosts.

On $\mathcal{B}(L^2(\mathbb{R}^3))$ the additive group $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ is represented by an automorphic action α ,

$$\alpha_{(a, \hat{b})}(x) := U_a V_b x V_b^* U_a^* =: \alpha_{(a, b)}(x), \quad x \in \mathcal{B}(L^2(\mathbb{R}^3)). \quad (3.6)$$

(The characters \hat{b} on \mathbb{R}^3 and the elements $b \in \mathbb{R}^3$ are in a one-to-one relation by $\hat{b}(s) = \exp(ibs)$ for $s \in \mathbb{R}^3$. So one can say that $\alpha_{(a, b)}$ is an automorphic representation of $\mathbb{R}^3 \times \mathbb{R}^3$ and that $\alpha_{(a, \hat{b})}$ is an automorphic representation of $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$.)

Since the representation $\{U_a, V_b\}$ is σ -weakly measurable, the action α is pointwise σ -weakly continuous. Therefore $(\mathcal{B}(L^2(\mathbb{R}^3)), \mathbb{R}^3 \times \widehat{\mathbb{R}^3}, \alpha)$ is a W^* -system in the sense of Definition 11. The action α is ergodic (Definition 12) on $\mathcal{B}(L^2(\mathbb{R}^3))$ since, due to the irreducibility of the

representation $\{U_a, V_b\}$, the only $x \in \mathcal{B}(L^2(\mathbb{R}^3))$ which fulfil $\alpha_{(a,\hat{b})}(x) = x$ for all a, \hat{b} are multiples of the identity.

Denote by P the self-adjoint generator of the unitary group U_a and by Q the self-adjoint generator of the unitary group V_b . Each character χ of $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ can be written as $\chi(a, \hat{b}) = \exp(iat + isb)$ for some $s, t \in \mathbb{R}^3$. Now observe that for each character χ there is a unitary operator $u_\chi \in \mathcal{B}(L^2(\mathbb{R}^3))$, namely $u_\chi = U_t V_s$, such that u_χ transforms under α according to χ :

$$\alpha_{(a,\hat{b})}(u_\chi) = \chi(a, \hat{b})u_\chi. \quad (3.7)$$

According to de Schreye (1986), since $(\mathcal{B}(L^2(\mathbb{R}^3)), \mathbb{R}^3 \times \widehat{\mathbb{R}^3}, \alpha)$ is ergodic, it follows from (3.7) that the automorphic representation α is integrable in the sense of Connes and Takesaki (see Definition 16). This in turn implies¹⁰ that there is a POV-measure a on $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ with values in $\mathcal{B}(L^2(\mathbb{R}^3))$ which fulfils the covariance condition

$$\alpha_{(a,\hat{b})}(a(\Delta)) = a([\Delta] + (a, \hat{b})), \quad (3.8)$$

for any Borel subset Δ of $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ and for any $(a, \hat{b}) \in \mathbb{R}^3 \times \widehat{\mathbb{R}^3}$. (Such a POV-measure was explicitly defined in equation (2.9).) The POV-measure a was derived from the unitary groups U_a, V_b whose generators are the position and momentum operators Q, P .

a transforms covariantly under space translations and velocity boosts. Also, its marginals a_e^Q (resp. a_e^P), defined in equation (2.16), transform covariantly under space translations (resp. velocity boosts), see eq. (2.19). The position space POV-measure a_e^Q and the momentum space POV-measure a_e^P lead to unsharp position and momentum observables Q_e, P_e (defined in eq. (2.17)). Q_e is a smeared version of the sharp

¹⁰See Amann 1986, Theorem II.7.

Schrödinger position operator, see eq. (2.20). For these reasons it is justified to regard a as joint position-momentum observable.

This motivates the following central

Definition 21 *Let $(\mathcal{M}, \mathbb{R}^3 \times \widehat{\mathbb{R}}^3, \alpha)$ be a W^* -system of the additive group $\mathbb{R}^3 \times \widehat{\mathbb{R}}^3$. A POV-measure a on $\mathbb{R}^3 \times \widehat{\mathbb{R}}^3$ with values in \mathcal{M} fulfilling the covariance condition (3.8) is called a joint position-momentum observable. A covariant PV-measure on $\mathbb{R}^3 \times \widehat{\mathbb{R}}^3$ is called a sharp joint position-momentum observable.*

REMARKS: (1) We use the additive group $\mathbb{R}^3 \times \widehat{\mathbb{R}}^3$ instead of the additive group $\mathbb{R}^6 = \mathbb{R}^3 \times \mathbb{R}^3$ because we do not want to consider a 6-dimensional position observable as a joint position-momentum observable. $\mathbb{R}^3 \times \widehat{\mathbb{R}}^3$ describes the group of space translations and boosts.

(2) As illustrated by Example 1, the justification of calling a covariant POV-measure on $\mathbb{R}^3 \times \widehat{\mathbb{R}}^3$ a joint position-momentum observable stems from the fact that the operators in its image have the transformation properties of position and momentum. This idea of defining observables through their transformation properties is due to Weyl.

3.2.2 Elementary Systems

We say that a physical system with algebra of observables \mathcal{M} is *elementary* if the kinematical group acts ergodically on \mathcal{M} (see Definition 12).

This definition (see Amann 1986) generalises the quantum mechanical notion of elementarity: In traditional quantum mechanics an elementary system is described by an irreducible ray representation U of the kinematical group G . Its algebra of observables $\mathcal{M} = \{U(g) : g \in G\}''$ is a type I factor. In this case the two notions of elementarity coincide: Every automorphism α_g is inner; it can be written as $\alpha_g = U_g \cdot U_g^*$

for some $U_g \in \mathcal{M}$. According to (Berberian 1974, theorem 67.2) a type I factor system (\mathcal{M}, G, α) is ergodic if and only if the corresponding ray representation $U(G)$ is irreducible.

Theorem 5 *Let (\mathcal{M}, G, α) be a W^* -system with respect to a locally compact separable group G . If (\mathcal{M}, G, α) is ergodic and \mathcal{M} is not abelian, then there exists no covariant PV-measure on G with values in \mathcal{M} .*

PROOF (see Amann 1986, IV.1): A covariant PV-measure on G with values in \mathcal{M} can be extended to a covariant $*$ -morphism $\pi : L^\infty(G) \rightarrow \mathcal{M}$. According to Amann (1986, II.2.2.), the existence of such a covariant $*$ -morphism is equivalent to the existence of a W^* -algebra \mathcal{N} and of a coaction δ on \mathcal{N} such that $\{\mathcal{M}, \alpha\} \cong \{\mathcal{N} \otimes_\delta G, \hat{\delta}\}$. Therefore $\mathcal{M}^\alpha \cong \mathcal{N}$. But since (\mathcal{M}, G, α) is ergodic, $\mathcal{M}^\alpha = \{\lambda \mathbf{1}\}$. Therefore $\mathcal{M} \cong \{\lambda \mathbf{1}\} \otimes_\delta G \cong L^\infty(G)$, which contradicts the assumption that \mathcal{M} is not abelian. \square

Corollary 2 *Let $\{U_a, V_b : a, b \in \mathbb{R}^3\}$ be an irreducible σ -weakly measurable unitary ray representation of the Weyl relations on the Hilbert space $L^2(\mathbb{R}^3)$. There exists no covariant PV-measure on $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ with values in $\mathcal{B}(L^2(\mathbb{R}^3))$. There exists a covariant POV-measure on $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ with values in $\mathcal{B}(L^2(\mathbb{R}^3))$.*

PROOF: The non-existence of a covariant PV-measure follows from Theorem 5 and the fact that $(\mathcal{B}(L^2(\mathbb{R}^3)), \mathbb{R}^3 \times \widehat{\mathbb{R}^3}, U_a V_b \cdot V_b^* U_a^*)$ is an ergodic non-abelian W^* -system. The existence of a covariant POV-measure was proved in Example 1.

REMARKS: (1) For elementary systems there exists a sharp joint position-momentum observable iff \mathcal{M} is abelian, i.e. iff the system is classical.

(2) One should be cautious not to infer prematurely from the existence of an unsharp joint position-momentum observable that position and momentum can be measured at the same time. Caution is necessary because for two observables the notion of being measurable *at the same time* carries the connotation of being transitive, whereas for two observables the existence of a joint unsharp observable is *not* transitive.

(3) Taking as G the phase space, Theorem 5 generalises to arbitrary von Neumann algebras results of Ali and Prugovečki (1977) that for (elementary) quantum systems there exist no covariant *projection* valued measures on the *phase* space. There exist, however, covariant PV-measures on the *configuration* space of an elementary particle. These are the transitive systems of imprimitivity as introduced by Mackey (1949). Also, there exist covariant POV-measures on the phase space if and only if α is integrable.¹¹ These are unsharp joint position-momentum observables.

(4) Note that the problem is not to find covariant PV-measure on G with values in *some* W^* -system of the group G . It is always possible take the W^* -system $(\mathcal{B}(L^2(G)), G, U \cdot U^*)$ where U is the representation of G on $L^2(G)$ defined by $U_{g'}\psi(g) := \psi(g - g')$, and to take as PV-measure E on G simply $E(\Delta)\psi(g) := \chi_\Delta(g)\psi(g)$, where χ_Δ is the characteristic function of the subset Δ of G . (Note that this W^* -system is not ergodic.) The problem is to find a covariant POV-measure with values in a *given* W^* -system.

Theorem 5 suggests that there are two scenarios in which it is possible to have sharp joint position-momentum observables: Either the von Neumann algebra \mathcal{M} should be abelian, or the action α should not be

¹¹This proved in Homma 1988 or Amann (1986, II.7) for the case that the phase space is a group. Propositions 2 and 3 generalise this to homogeneous phase spaces.

ergodic.

3.2.3 Finite Systems

The assumption of ergodicity is essential in Theorem 5. This is illustrated by the following

EXAMPLE 2: As in Example 1, let $\{U_a, V_b : a, b \in \mathbb{R}^3\}$ be an irreducible representation of the Weyl relations $U_a V_b = \exp(iab)V_b U_a$ on $L^2(\mathbb{R}^3)$. It describes the one particle Weyl system. Consider the two particle Weyl system whose algebra of observables is $\mathcal{B}(L^2(\mathbb{R}^6)) = \mathcal{B}(L^2(\mathbb{R}^3)) \otimes \mathcal{B}(L^2(\mathbb{R}^3))$. Define a representation $\bar{\alpha}$ of $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ by extension to linear combinations of

$$\bar{\alpha}_{(a,\hat{b})}(x_1 \otimes x_2) := U_a V_b x_1 V_b^* U_a^* \otimes U_a V_b x_2 V_b^* U_a^*.$$

This W^* -system is not ergodic since $1/2(\mathbf{1} + W)$, where $W(\psi_1 \otimes \psi_2) := \psi_2 \otimes \psi_1$ is the unitary operator implementing a permutation of the two particles, is a projection invariant under all $\bar{\alpha}_{(a,b)}$.

The condition of ergodicity being violated, we can construct a PV-measure on $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ which is covariant with respect to $\bar{\alpha}$. Let Q be the self-adjoint generator of the unitary group V_b and let E^Q be its spectral measure: $Q =: \int_{\mathbb{R}^3} \lambda E^Q(d\lambda)$. Since $a \mapsto U_a$ is a strongly continuous unitary representation of the additive group \mathbb{R}^3 , a corollary of Stone's theorem¹² implies that there exists a PV-measure \hat{E}^P on $\widehat{\mathbb{R}^3}$ with values in $\mathcal{B}(L^2(\mathbb{R}^3))$ such that $\langle \phi, U_a \psi \rangle = \int_{\widehat{\mathbb{R}^3}} \hat{\lambda}(a) \langle \phi, \hat{E}^P(d\hat{\lambda}) \psi \rangle$ for all $\phi, \psi \in L^2(\mathbb{R}^3)$. Then

$$\Delta_1 \times \Delta_2 \mapsto E^Q(\Delta_1) \otimes \hat{E}^P(\Delta_2)$$

¹²See e.g. Reed and Simon 1972, Theorem VIII.12.

3.2. SHARP MEASURABILITY OF POSITION AND MOMENTUM 171

can be extended to a PV-measure on $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ with values in $\mathcal{B}(L^2(\mathbb{R}^3)) \otimes \mathcal{B}(L^2(\mathbb{R}^3))$. It is covariant since

$$\begin{aligned} U_a V_b E^Q(\Delta_1) V_b^* U_a^* &\otimes U_a V_b \hat{E}^P(\Delta_2) V_b^* U_a^* = \\ &= E^Q([\Delta_1] + a) \otimes \hat{E}^P([\Delta_2] + \hat{b}). \end{aligned}$$

This sharp joint position-momentum observable describes the position of one particle and the momentum of the other.

We should exclude such trivial joint sharp measurements of position and momentum for non-elementary systems. We do this by requiring that the operators in the range of the POV-measure is permutation invariant. Since we will only consider average position and momentum observables, the requirement of permutation invariance will automatically be satisfied.

We will now set out to show that average position and momentum of a finite Weyl system cannot be measured sharply at the same time. But first we will define the position and momentum operators of the one particle system. Usually these are taken to be the selfadjoint generators P of U_a and Q of V_b . Since P, Q are unbounded, they do not belong to $\mathcal{B}(L^2(\mathbb{R}^3))$. Rather they are affiliated to it. To make this mathematically precise, write $P = P_+ - P_-$ and $Q = Q_+ - Q_-$ as differences of positive unbounded selfadjoint operators P_\pm, Q_\pm , also affiliated with $\mathcal{B}(L^2(\mathbb{R}^3))$. So (see e.g. Haagerup 1979) there exist mappings (which I will also denote by P_\pm, Q_\pm) from $\mathcal{B}(L^2(\mathbb{R}^3))_1^*$ into the positive reals which are linear and semicontinuous from below. For every $\epsilon > 0$, $(1 + \epsilon P_\pm)^{-1} P_\pm$, $(1 + \epsilon Q_\pm)^{-1} Q_\pm$ are in $\mathcal{B}(L^2(\mathbb{R}^3))$ (see e.g. Pedersen 1979, 5.3.10). For every normal state ρ in $\mathcal{B}(L^2(\mathbb{R}^3))_1^*$ we have

$$P_\pm(\rho) = \lim_{\epsilon \rightarrow 0} \rho((1 + \epsilon P_\pm)^{-1} P_\pm), \quad Q_\pm(\rho) = \lim_{\epsilon \rightarrow 0} \rho((1 + \epsilon Q_\pm)^{-1} Q_\pm).$$

P and Q transform under α covariantly according to

$$\alpha_{(a,\hat{b})}(P) = P - b\mathbf{1}, \quad \alpha_{(a,\hat{b})}(Q) = Q - a\mathbf{1}. \quad (3.9)$$

They describe position and momentum of the one-particle system.

Let us now consider a system consisting of finitely many particles. Let Λ be the set of finite subsets of an infinite index set Π , and for $\Lambda \in \Lambda$ let $|\Lambda|$ be the cardinality of Λ . Let $\{U_a, V_b : a, b \in \mathbb{R}^3\}$ be an irreducible representation of the Weyl relations $U_a V_b = \exp(iab)V_b U_a$ on $L^2(\mathbb{R}^3)$. Denote by $u_n : L^2(\mathbb{R}^3) \rightarrow \mathcal{H}_n$ unitary mappings from $L^2(\mathbb{R}^3)$ onto copies \mathcal{H}_n of $L^2(\mathbb{R}^3)$. Define $\pi_n(x) := u_n x u_n^{-1}$ for $x \in \mathcal{B}(L^2(\mathbb{R}^3))$. Now let

$$\mathcal{A}^\Lambda := \bigotimes_{n \in \Lambda} \pi_n(\mathcal{B}(L^2(\mathbb{R}^3))),$$

be the W^* -tensorproduct of $|\Lambda|$ copies of $\mathcal{B}(L^2(\mathbb{R}^3))$. \mathcal{A}^Λ is the algebra of observables of a system consisting of $|\Lambda|$ non-relativistic particles.

As before, a pointwise σ -weakly continuous representation of the additive group $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ on $\mathcal{B}(L^2(\mathbb{R}^3))$ is given by $\alpha_{(a,\hat{b})}(x) := U_a V_b x V_b^* U_a^*$. On \mathcal{A}^Λ , define an action $\bar{\alpha}$ of $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ by extension to linear combinations of

$$\bar{\alpha}_{(a,\hat{b})} \left(\bigotimes_{n \in \Lambda} \pi_n(x_n) \right) := \bigotimes_{n \in \Lambda} \pi_n \left(\alpha_{(a,\hat{b})}(x_n) \right). \quad (3.10)$$

$\bar{\alpha}$ describes translations and boosts of the system as a whole.

Next construct the average position and momentum operators for the system of $|\Lambda|$ particles. For $x \in \mathcal{B}(L^2(\mathbb{R}^3))$ define the *averaged* element $x^\Lambda \in \mathcal{A}^\Lambda$ by

$$x^\Lambda := \frac{1}{|\Lambda|} \sum_{n \in \Lambda} \mathbf{1} \otimes \dots \otimes \mathbf{1} \otimes \pi_n(x) \otimes \mathbf{1} \otimes \dots \otimes \mathbf{1}. \quad (3.11)$$

3.2. SHARP MEASURABILITY OF POSITION AND MOMENTUM 173

$((1 + \epsilon P_{\pm})^{-1} P_{\pm})^{\Lambda}$ and $((1 + \epsilon Q_{\pm})^{-1} Q_{\pm})^{\Lambda}$ are in \mathcal{A}^{Λ} for all $\epsilon > 0$. Therefore we can define positive unbounded self-adjoint operators $P_{\pm}^{\Lambda}, Q_{\pm}^{\Lambda}$ affiliated with \mathcal{A}^{Λ} by taking

$$P_{\pm}^{\Lambda}(\rho) := \lim_{\epsilon \rightarrow 0} \rho \left(\left((1 + \epsilon P_{\pm})^{-1} P_{\pm} \right)^{\Lambda} \right),$$

$$Q_{\pm}^{\Lambda}(\rho) := \lim_{\epsilon \rightarrow 0} \rho \left(\left((1 + \epsilon Q_{\pm})^{-1} Q_{\pm} \right)^{\Lambda} \right),$$

for any normal state $\rho \in (\mathcal{A}^{\Lambda})_{*}^1$. Define $P^{\Lambda} := P_{+}^{\Lambda} - P_{-}^{\Lambda}$ and $Q^{\Lambda} := Q_{+}^{\Lambda} - Q_{-}^{\Lambda}$. Formally (i.e. ignoring the fact that π_n is only defined on bounded operators) one can write

$$Q^{\Lambda} = \frac{1}{|\Lambda|} \sum_{n \in \Lambda} \mathbf{1} \otimes \dots \otimes \mathbf{1} \otimes \pi_n(Q) \otimes \mathbf{1} \otimes \dots \otimes \mathbf{1}, \quad (3.12)$$

$$P^{\Lambda} = \frac{1}{|\Lambda|} \sum_{n \in \Lambda} \mathbf{1} \otimes \dots \otimes \mathbf{1} \otimes \pi_n(P) \otimes \mathbf{1} \otimes \dots \otimes \mathbf{1}. \quad (3.13)$$

Readers less worried about technical details can take these last equations to define Q^{Λ}, P^{Λ} . Q^{Λ}, P^{Λ} describe the average position and momentum of the Weyl system with $|\Lambda|$ particles.

Lemma 2 *The P^{Λ}, Q^{Λ} defined above transform under the action $\bar{\alpha}$ defined in (3.10) covariantly according to*

$$\bar{\alpha}_{(a,\hat{b})}(P^{\Lambda}) = P^{\Lambda} - b\mathbf{1}^{\Lambda}, \quad \bar{\alpha}_{(a,\hat{b})}(Q^{\Lambda}) = Q^{\Lambda} - a\mathbf{1}^{\Lambda}. \quad (3.14)$$

PROOF: Using the proper definition of P^{Λ} one verifies the following formal calculation

$$\begin{aligned} \bar{\alpha}_{(a,\hat{b})}(P^{\Lambda}) &\stackrel{(3.10,3.13)}{=} \frac{1}{|\Lambda|} \sum_{n \in \Lambda} \mathbf{1} \otimes \dots \otimes \mathbf{1} \otimes \pi_n(\alpha_{(a,\hat{b})}(P)) \otimes \mathbf{1} \otimes \dots \otimes \mathbf{1} \\ &\stackrel{(3.9)}{=} \frac{1}{|\Lambda|} \sum_{n \in \Lambda} \mathbf{1} \otimes \dots \otimes \mathbf{1} \otimes \pi_n(P - b\mathbf{1}) \otimes \mathbf{1} \otimes \dots \otimes \mathbf{1} \\ &\stackrel{(3.11,3.13)}{=} P^{\Lambda} - b\mathbf{1}^{\Lambda}. \end{aligned}$$

Similarly for Q^{Λ} .

Lemma 3 $\exp(iaP^\Lambda), \exp(ibQ^\Lambda)$ fulfil the commutation relations

$$\exp(iaP^\Lambda) \exp(ibQ^\Lambda) = e^{iab/|\Lambda|} \exp(ibQ^\Lambda) \exp(iaP^\Lambda).$$

PROOF: Follows by straight calculation from the Weyl relations $\exp(ia|\Lambda|^{-1}P) \exp(ib|\Lambda|^{-1}Q) = e^{iab|\Lambda|^{-2}} \exp(ib|\Lambda|^{-1}Q) \exp(ia|\Lambda|^{-1}P)$.

Define \mathcal{M}^Λ to be the von Neumann subalgebra of \mathcal{A}^Λ generated by $\exp(iaP^\Lambda)$ and $\exp(ibQ^\Lambda)$:

$$\mathcal{M}^\Lambda := \{\exp(iaP^\Lambda) \exp(ibQ^\Lambda) : a, b \in \mathbb{R}^3\}'' \subset \mathcal{A}^\Lambda. \quad (3.15)$$

Loosely speaking (i.e. neglecting the fact that P^Λ, Q^Λ are unbounded and therefore not contained in \mathcal{M}^Λ) one can say that \mathcal{M}^Λ is the von Neumann algebra generated by average position and momentum Q^Λ, P^Λ . \mathcal{M}^Λ contains exactly the averaged observables: $\mathcal{M}^\Lambda = \{x^\Lambda : x \in \mathcal{B}(L^2(\mathbb{R}^3))\}$. For $|\Lambda|$ finite, \mathcal{M}^Λ is a factor of type I_∞ .

Lemma 4 For every finite index set Λ , the action $\bar{\alpha}$ defined in (3.10) acts ergodically on the von Neumann algebra \mathcal{M}^Λ defined by (3.15).

PROOF: Assume that

$$\bar{\alpha}_{(a,\hat{b})} \left(\sum_{n \in \Lambda} \mathbf{1} \otimes \dots \otimes \pi_n(x) \otimes \dots \otimes \mathbf{1} \right) = \sum_{n \in \Lambda} \mathbf{1} \otimes \dots \otimes \pi_n(x) \otimes \dots \otimes \mathbf{1} \quad (3.16)$$

for all $(a, \hat{b}) \in \mathbb{R}^3 \times \widehat{\mathbb{R}^3}$. We will prove that $\sum_{n \in \Lambda} \mathbf{1} \otimes \dots \otimes \pi_n(x) \otimes \dots \otimes \mathbf{1}$ is a multiple of the identity.

First of all note that if $\mathcal{M}_1, \mathcal{M}_2$ are von Neumann algebras, then for any $A \in \mathcal{M}_2, B \in \mathcal{M}_1, \mathbf{1}_1 \otimes A = B \otimes \mathbf{1}_2$ implies that $A = c\mathbf{1}_2, B = c\mathbf{1}_1$ for some complex number c .

3.2. SHARP MEASURABILITY OF POSITION AND MOMENTUM 175

Denoting by n_0 the first element of the index set Λ , (3.16) implies

$$\begin{aligned} & \left(\pi_{n_0}(x) - \pi_{n_0}(\alpha_{(a,\hat{b})}(x)) \right) \otimes \mathbf{1} \dots \otimes \mathbf{1} = \\ & = \mathbf{1} \otimes \left(\sum_{\substack{n \in \Lambda \\ n \neq n_0}} \mathbf{1} \otimes \dots \otimes (\pi_n(\alpha_{(a,\hat{b})}(x)) - \pi_n(x)) \otimes \dots \otimes \mathbf{1} \right). \end{aligned}$$

This in turn implies that

$$\pi_{n_0}(x) - \pi_{n_0}(\alpha_{(a,\hat{b})}(x)) = c\mathbf{1} \quad (3.17)$$

$$\sum_{\substack{n \in \Lambda \\ n \neq n_0}} \mathbf{1} \otimes \dots \otimes (\pi_n(\alpha_{(a,\hat{b})}(x)) - \pi_n(x)) \otimes \dots \otimes \mathbf{1} = c\mathbf{1} \otimes \dots \otimes \mathbf{1}. \quad (3.18)$$

From (3.17) we infer that

$$x - \alpha_{(a,\hat{b})}(x) = c\mathbf{1} \quad (3.19)$$

which implies that $\pi_n(x) - \pi_n(\alpha_{(a,\hat{b})}(x)) = c\mathbf{1}$ for all $n \in \Lambda$. So (3.18) can be written as

$$- \left(\sum_{n \in \Lambda, n \neq n_0} c\mathbf{1} \otimes \dots \otimes \mathbf{1} \right) = c\mathbf{1} \otimes \dots \otimes \mathbf{1}.$$

This implies that $-|\Lambda|c\mathbf{1} \otimes \dots \otimes \mathbf{1} = 0$ which can only be the case if $c = 0$. Thus (3.19) reads $x = \alpha_{(a,\hat{b})}(x)$ for all $(a, \hat{b}) \in \mathbb{R}^3 \times \widehat{\mathbb{R}^3}$, which by ergodicity of α on $\mathcal{B}(L^2(\mathbb{R}^3))$ implies that x is a multiple of the identity. The same is true for $\sum_{n \in \Lambda} \mathbf{1} \otimes \dots \otimes \pi_n(x) \otimes \dots \otimes \mathbf{1}$. \square

From Lemmata 2, 4, Theorem 5, and an obvious generalisation of Example 1 we obtain the following

Theorem 6 *For $|\Lambda|$ finite, there is no PV-measure on $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ with values in \mathcal{M}^Λ which is covariant with respect to the action $\bar{\alpha}$ defined by (3.10). There is a covariant POV-measure on $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ with values in \mathcal{M}^Λ .*

This shows that average position and momentum of a finite quantum system cannot be measured sharply at the same time.

3.2.4 Infinite systems

Let \mathcal{A} be the C^* -inductive limit of the net $\{\mathcal{A}^\Lambda\}_{\Lambda \in \mathbf{\Lambda}}$. \mathcal{A} is simple and has a quasilocal structure. One can embed each \mathcal{A}^Λ into \mathcal{A} by $\iota : x \mapsto x \otimes \mathbf{1} \otimes \mathbf{1} \otimes \dots$. The action $\bar{\alpha}$ on \mathcal{A}^Λ defined by equation (3.10) can be extended continuously to an action on \mathcal{A} , and further to one on \mathcal{A}^{**} . Let us denote this action also by $\bar{\alpha}$.

Using the framework of Amann and Müller-Herold (1986) and of Amann (1987), one can construct a classical momentum operator on a separable Hilbert space. This classical momentum operator arises through spontaneous breaking of the boost symmetry. Due to its classicality it commutes with all possible position operators. This implies the simultaneous sharp measurability of position and this momentum operator.

Here we will take a different route and consider *average* position and momentum operators. One idea for how to treat the case of infinitely many particles would be to take the norm limit of the average position and momentum operators Q^Λ, P^Λ . From Lemma 3 one would expect that their commutator tends to zero, and that therefore one can measure them sharply at the same time. But we will see that there is a problem: the norm limits of the average position and momentum operators do not exist.

Lemma 5 *For any $y \in \mathcal{A}$, $\lim_{\Lambda} \|[y, \iota x^\Lambda]\|$, if it exists, is zero.*

PROOF: Let $\epsilon > 0$, $y \in \mathcal{A}$, and $x \in \mathcal{B}(L^2(\mathbb{R}^3))$ be arbitrary but fixed. I will show that there is a $\Lambda_0 \in \mathbf{\Lambda}$ such that $\|[y, \iota x^\Lambda]\| \leq \epsilon$ for all $\Lambda \supset \Lambda_0$.

3.2. SHARP MEASURABILITY OF POSITION AND MOMENTUM 177

Every element y in the quasilocal algebra \mathcal{A} can be approximated in norm by local elements. So there is a $\Lambda_1 \in \mathbf{\Lambda}$ and a $y_1 \in \mathcal{A}^{\Lambda_1}$ such that $\|y - \iota y_1\| \leq \epsilon/(8\|x\|)$. Every element y_1 of the tensor product \mathcal{A}^{Λ_1} , can be approximated in norm by elements which are finite sums. So there is an element $z_1 \in \mathcal{A}^{\Lambda_1}$ of the form $z_1 = \sum_{i=1}^n (\otimes_{k \in \Lambda_1} z_k^i)$ fulfilling $\|y_1 - z_1\| \leq \epsilon/(8\|x\|)$. ιz_1 can be written as $\sum_{i=1}^n (\otimes_{k \in \Pi} z_k^i)$ if we take $z_k^i := \pi_k(\mathbf{1})$ for all i whenever $k \notin \Lambda_1$. Then $\|\iota y_1 - \iota z_1\| \leq \epsilon/(8\|x\|)$. Thus $\|y - \iota z_1\| \leq \epsilon/(4\|x\|)$. We thus have

$$\begin{aligned} \|[\iota x^\Lambda, y - \iota z_1]\| &\leq 2\|\iota x^\Lambda\| \|y - \iota z_1\| \\ &= 2\|x\| \|y - \iota z_1\| \\ &\leq 2 \frac{\epsilon}{4} = \frac{\epsilon}{2}. \end{aligned}$$

For an arbitrary $\Lambda_0 \in \mathbf{\Lambda}$ with $\Lambda_1 \subset \Lambda_0$ one calculates

$$\begin{aligned} \|[\iota z_1, \iota x^{\Lambda_0}]\| &= \left\| \left[\sum_{i=1}^n (\otimes_{k \in \Pi} z_k^i), \frac{1}{|\Lambda_0|} \sum_{j \in \Lambda_0} \mathbf{1} \otimes \dots \otimes \pi_j(x) \otimes \mathbf{1} \otimes \dots \right] \right\| \\ &= \frac{1}{|\Lambda_0|} \left\| \sum_{i=1}^n \sum_{k \in \Lambda_0} [z_k^i, x] \right\| = \frac{1}{|\Lambda_0|} \left\| \sum_{i=1}^n \sum_{k \in \Lambda_1} [z_k^i, x] \right\| \\ &=: \frac{1}{|\Lambda_0|} c, \end{aligned}$$

where the number $c := \|\sum_{i=1}^n \sum_{k \in \Lambda_1} [z_k^i, x]\|$ is finite and does not depend on Λ_0 . Choosing a Λ_0 with $\Lambda_1 \subset \Lambda_0$ and $|\Lambda_0| \geq 2c/\epsilon$, we get $\|[\iota z_1, \iota x^\Lambda]\| \leq \epsilon/2$ for all $\Lambda \supset \Lambda_0$.

We therefore have

$$\begin{aligned} \|[\iota x^\Lambda, y]\| &= \|[\iota x^\Lambda, y - \iota z_1] + [\iota x^\Lambda, \iota z_1]\| \\ &\leq \|[\iota x^\Lambda, y - \iota z_1]\| + \|[\iota x^\Lambda, \iota z_1]\| \\ &= \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

for all $\Lambda \supset \Lambda_0$. □

Since \mathcal{A} is simple, Lemma 5 implies that any norm limit of x^Λ would be a multiple of the identity. This can only be the case if x is a multiple of the identity.¹³ Therefore the norm limits of the average position and momentum observables P^Λ, Q^Λ do not exist.

But it can be shown¹⁴ that the strong operator limits

$$s - \lim_{\Lambda \in \mathbf{\Lambda}} \pi_\omega(x^\Lambda) \in \pi_\omega(\mathcal{A})'',$$

where π_ω denotes the GNS-representation of \mathcal{A} associated to the state ω , exist for many states ω on \mathcal{A} , in particular for the permutation invariant states. The existence of this limit is equivalent to the existence of the limit with ω replaced by a state quasi-equivalent to ω .

Lemma 6 *For any normal state ω on \mathcal{A} and for any $x \in \mathcal{B}(L^2(\mathbb{R}^3))$, the strong operator limits $s - \lim_{\Lambda \in \mathbf{\Lambda}} \pi_\omega(x^\Lambda)$ of the average elements x^Λ , if they exist, are in the centre of $\pi_\omega(\mathcal{A})''$.*

PROOF: Follows from Lemma 5 because for all $y \in \mathcal{A}$ we have

$$[\pi_\omega(y), s - \lim_{\Lambda \in \mathbf{\Lambda}} \pi_\omega(x^\Lambda)] = s - \lim_{\Lambda \in \mathbf{\Lambda}} \pi_\omega([y, \iota x^\Lambda]) = 0.$$

Now we will define the average position and momentum observables Q^Π, P^Π of the infinite quantum system. Denote by π_u the universal representation of \mathcal{A} and by \mathcal{H}_u its Hilbert space. For any state ω on \mathcal{A} we identify $\pi_\omega(x)$ with $\pi_u(x)c(\pi_\omega)$, where $c(\pi_\omega)$ is the central support

¹³See Bóna 1988, sect. 2.

¹⁴See van Hemmen 1978.

3.2. SHARP MEASURABILITY OF POSITION AND MOMENTUM 179

of the representation π_ω . Define P_G as the largest central projection C in \mathcal{A}^{**} for which all the strong limits

$$P_{\pm, \epsilon}^{\Pi} := s - \lim_{\Lambda} \pi_u((1 + \epsilon P_{\pm}^{\Lambda})^{-1} P_{\pm}^{\Lambda})C,$$

$$Q_{\pm, \epsilon}^{\Pi} := s - \lim_{\Lambda} \pi_u((1 + \epsilon Q_{\pm}^{\Lambda})^{-1} Q_{\pm}^{\Lambda})C$$

exist. $P_G \mathcal{A}^{**}$ is the direct sum of all GNS-representations associated to states ω for which the strong operator limits $s\text{-}\lim_{\Lambda \in \Lambda} \pi_\omega(x^\Lambda)$ of the averaged observables x^Λ exist. From Lemma 6 we know that $P_{\pm, \epsilon}^{\Pi}, Q_{\pm, \epsilon}^{\Pi}$ are in the centre of $P_G \mathcal{A}^{**}$. Then we can define positive unbounded self-adjoint operators $P_{\pm}^{\Pi}, Q_{\pm}^{\Pi}$ affiliated with the centre of $P_G \mathcal{A}^{**}$ by setting

$$\rho(P_{\pm}^{\Pi}) := \lim_{\epsilon \rightarrow 0} \rho(P_{\pm, \epsilon}^{\Pi}), \quad \rho(Q_{\pm}^{\Pi}) := \lim_{\epsilon \rightarrow 0} \rho(Q_{\pm, \epsilon}^{\Pi})$$

for every $\rho \in (P_G \mathcal{A}^{**})_*^1$. Now take $P^{\Pi} := P_+^{\Pi} - P_-^{\Pi}$ and $Q^{\Pi} := Q_+^{\Pi} - Q_-^{\Pi}$. Q^{Π} and P^{Π} can formally be written as

$$Q^{\Pi} = s\text{-}\lim_{\Lambda \in \Lambda} \pi_u(Q^{\Lambda})P_G, \quad (3.20)$$

$$P^{\Pi} = s\text{-}\lim_{\Lambda \in \Lambda} \pi_u(P^{\Lambda})P_G. \quad (3.21)$$

Again, the reader less worried about technical details can think of P^{Π}, Q^{Π} as being defined by (3.20), (3.21). They are the average position and momentum observables of the infinite system. By construction they are affiliated to the centre of $P_G \mathcal{A}^{**}$ and therefore commute with each other.

Define \mathcal{M}^{Π} to be the von Neumann subalgebra of $P_G \mathcal{A}^{**}$ generated by $\exp(iaP^{\Pi})$ and $\exp(ibQ^{\Pi})$:

$$\mathcal{M}^{\Pi} := \{\exp(iaP^{\Pi}) \exp(ibQ^{\Pi}) : a, b \in \mathbb{R}^3\}'' \subset P_G \mathcal{A}^{**}. \quad (3.22)$$

Loosely speaking (i.e. neglecting the fact that P^{Π}, Q^{Π} are unbounded and therefore not contained in \mathcal{M}^{Λ}) one can say that \mathcal{M}^{Π} is the von

Neumann algebra generated by average position and momentum Q^Π, P^Π . \mathcal{M}^Π contains exactly the strong limits of the averaged observables (3.11), and is contained in the centre of $P_G\mathcal{A}^{**}$.

So we arrive at

Theorem 7 *There exists a covariant PV-measure E_G on $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ with values in the von Neumann algebra \mathcal{M}^Π generated by the average position and momentum. E_G can be extended to a (covariant) isomorphism between*

$L^\infty(\mathbb{R}^3 \times \widehat{\mathbb{R}^3})$ and \mathcal{M}^Π .

PROOF: The only thing which we did not prove yet is the existence of a covariant PV-measure on $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ with values in \mathcal{M}^Π .

Due to the linear dependence of aP^Π on a and of bQ^Π on b ,

$$(\hat{b}, a) \mapsto \exp(ibQ^\Pi) \exp(iaP^\Pi)$$

is a strongly continuous unitary representation of the additive group $\widehat{\mathbb{R}^3} \times \mathbb{R}^3$ (the dual of $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$) in the Hilbert space $P_G\mathcal{H}_u$. According to a corollary of Stone's theorem¹⁵ there is a unique projection valued measure E_G on the dual group of $\widehat{\mathbb{R}^3} \times \mathbb{R}^3$ with values in the central projectors of $P_G\mathcal{A}^{**}$ such that

$$\langle \phi, \exp(ibQ^\Pi) \exp(iaP^\Pi) \psi \rangle = \int_{\mathbb{R}^3 \times \widehat{\mathbb{R}^3}} (s, \hat{t})(\hat{b}, a) \langle \phi, E_G(ds, d\hat{t}) \psi \rangle$$

for all $\phi, \psi \in P_G\mathcal{H}_u$. (Here we identified $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$ with the dual of $\widehat{\mathbb{R}^3} \times \mathbb{R}^3$ and wrote $(s, \hat{t})(\hat{b}, a)$ for the value of the character (s, \hat{t}) on the group element $(\hat{b}, a) \in \widehat{\mathbb{R}^3} \times \mathbb{R}^3$.) The support of E_G is equal to the spectra of its selfadjoint generators¹⁶ P^Π, Q^Π , and so is equal to the whole of $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$.

¹⁵See e.g. Reed and Simon 1972, Theorem VIII.12.

¹⁶See e.g. Bratteli and Robinson (1987), Proposition 3.2.40, (6).

3.2. SHARP MEASURABILITY OF POSITION AND MOMENTUM 181

From the definition of P_G it is obvious that it is invariant under $\bar{\alpha}$. Also, using the proper definition of P^Π, Q^Π one verifies the following formal calculation

$$\begin{aligned}
 \bar{\alpha}_{(a,\hat{b})}(P^\Pi) &\stackrel{(3.21)}{=} \bar{\alpha}_{(a,\hat{b})}(\text{s-lim}_{\Lambda \in \mathbf{\Lambda}} \pi_u(P^\Lambda)P_G) \\
 &\stackrel{(3.14,3.10)}{=} \text{s-lim}_{\Lambda \in \mathbf{\Lambda}} \pi_u(P^\Lambda - b\mathbf{1}^\Lambda)P_G \\
 &\stackrel{(3.21)}{=} P^\Pi - b \text{s-lim}_{\Lambda \in \mathbf{\Lambda}} \pi_u(\mathbf{1})P_G \\
 &= P^\Pi - b\mathbf{1}^\Pi,
 \end{aligned}$$

and similarly for Q^Π :

$$\bar{\alpha}_{(a,\hat{b})}(Q^\Pi) = Q^\Pi - b\mathbf{1}^\Pi.$$

(Here $\mathbf{1}^\Pi$ denotes the identity in the von Neumann algebra $P_G\mathcal{A}^{**}$.)

This can be written as covariance of the measure E_G :

$$\bar{\alpha}_{(a,\hat{b})}(E_G(\Delta)) = E_G([\Delta] - (a, \hat{b}))$$

for all Borel sets Δ of $\mathbb{R}^3 \times \widehat{\mathbb{R}^3}$. □

REMARKS: (1) The action $\bar{\alpha}$ on the von Neumann algebra \mathcal{M}^Π is ergodic. (This can be seen by taking the strong limit of Lemma 4. The proof of Lemma 4 does not depend on $|\Lambda|$.) Let us compare this with Theorem 5: The existence of a covariant representation in spite of the ergodicity of the action is due to the fact that \mathcal{M}^Π is abelian, whereas for $|\Lambda|$ finite, \mathcal{M}^Λ is a factor.

(2) E_G is a joint position-momentum observable derived from the limit of the average position and momentum observables. It describes the average position and momentum of the infinite quantum system. Since E_G is a projection valued measure, it describes a *sharp* joint position-momentum observable. Furthermore, since E_G takes values in the *centre* of $P_G\mathcal{A}^{**}$, average position and momentum are classical observables.

Chapter 4

The Classical Pointer in Quantum Measurements

Chapter Abstract

The use of classical pointer observables in quantum measurements is discussed. It is shown that inaccurate experiments can be described if and only if the pointer observable is classical. The main open problem of the approach is the existence of time evolutions which lead in finite time to different pointer readings. It is suggested that long range interactions can give rise to non-automorphic time evolutions achieving this.

Chapter Overview

In this last chapter I shall discuss the rôle classical pointer observables plays in the description of quantum measurements. Section 4.1 shortly reviews recent literature concerned with this question. In particular I shall explain which aspects of the measurement problem are solved by assuming the pointer observable to be classical, which problems remain, and which new problems arise.

The idea of inaccurate state-measurements will be further developed

in sections 4.2 and 4.3. In section 4.2 I will discuss how the existence of non-normal states and of von Neumann algebras with atom free projection lattice is related to measurement inaccuracy. In section 4.3 a new argument will be presented in favour of the classical pointer observable. I will formulate one notion of inaccuracy and show that quantum measurements can be inaccurate in this sense if and only if the pointer observable is classical. Section 4.3 follows Breuer (1992, 1993), and Breuer, Amann, and Landsmann (1993a, 1993b).

Finally, section 4.4 I will deal with the time evolution of the classical pointer. The main problem is that under an automorphic time evolution states of the same superselection sector (corresponding to pointer state “ready”) cannot evolve in finite time into states in which the classical pointer observable has different values. I briefly discuss arguments for and against the infinite time limit as a way to circumvent this problem. Another way around the problem is to consider non-automorphic time evolutions, which arise for example in the presence of long range interactions. Generalising von Neumann’s (1932, p. 236) measurement interaction, I give an example of a time evolution which arrives at disjoint final states in finite time.

4.1 Previous arguments for and against classical pointer observables

I will set the scene by a brief description of the measurement problem. Following the pioneering work of Hepp (1972) the use of classical pointer observables in quantum measurements has been extensively¹

¹See e.g. Bóna (1973), Frigerio (1974), Whitten-Wolf and Emch (1976), Machida and Namiki (1980), Araki (1980, 1986), Kudaka *et al.* (1989), Namiki and Pascazio (1991), Landsman (1991), Nakazato and Pascazio (1991, 1992). For philosophical

studied. The main purpose of this section is to describe which problems are solved by the introduction of a classical pointer observable, which are not solved, and which arise newly.

4.1.1 The projection postulate

Description of the problem. In the original formulation of the measurement problem, the algebra of observables for the observed system plus the apparatus is taken to be $\mathcal{B}(\mathcal{H}_S) \otimes \mathcal{B}(\mathcal{H}_M)$. (\mathcal{H}_S and \mathcal{H}_M are the Hilbert spaces of the observed system and of the apparatus, respectively.) For the time being assume that the measured observable and the pointer observable are discrete and non-degenerate. Assume also that the measurement established perfect correlations between the measured quantity and the pointer observable. Although these assumptions are an idealisation, the measurement problem does not disappear if drop them. I make these assumption in order to present the problem in its simplest form.

Denote: by o_n the eigenstates of the measured observable; by I the initial state of the apparatus; and by p_n the eigenstates of the pointer observable. It is possible² to find a Hamiltonian such that under the Schrödinger equation the initial states $|o_n\rangle \otimes |I\rangle$ evolves into $|o_n\rangle \otimes |p_n\rangle$. So if after the measurement the apparatus is in the state $|p_n\rangle$, one knows that before the measurement the observed system was in the state $|o_n\rangle$. But if the initial state of the observed system is $\sum c_n |o_n\rangle$, then, by

discussions of the classical pointer observable in quantum measurements see Bub (1988, 1989), Landsman (1995), and Robinson (1990, 1994).

²This is for example achieved by the Hamiltonian proposed by von Neumann (1932, p. 236).

linearity of the Schrödinger equation, the final state will be

$$s_1 := \sum c_n |o_n\rangle \otimes |p_n\rangle.$$

In the basis $\{o_i \otimes p_j\}$ the density matrix of the state s_1 has non-vanishing elements off the diagonal.

We have, however, the impression that after an experiment the pointer observable has an unambiguous value. If p_k is apparatus state after a single experiment, the final state of the joint system is

$$s_3 := |o_k \otimes p_k\rangle$$

with the probability $|c_k|^2$. In the basis $\{o_i \otimes p_j\}$ the density matrix of the state s_3 has only one entry which does not vanish, and this entry is on the diagonal.

Assuming unambiguous measurement results the density operator describing the statistics of many experiments should be the mixture

$$s_2 := \sum |c_n|^2 |o_n \otimes p_n\rangle \langle o_n \otimes p_n|.$$

In the basis $\{o_i \otimes p_j\}$ the state s_2 has vanishing elements off the diagonal.

In a description of individual experiments, the stochastic transition from the initial state $(\sum c_n |o_n\rangle) \otimes |I\rangle$ to the final state s_3 has to be explained. One can think of the transition $(\sum c_n |o_n\rangle) \otimes |I\rangle \rightarrow s_3$ as formally³ consisting of two parts: the Schrödinger time evolution $(\sum c_n |o_n\rangle) \otimes |I\rangle \rightarrow s_1$ and a transition $s_1 \rightarrow s_3$. I will call the transition $s_1 \rightarrow s_3$, the *reduction of the wave packet*.

In the description of a statistical experiment, one has to explain the transition from the initial state $(\sum c_n |o_n\rangle) \otimes |I\rangle$ to the final state s_2 ,

³This should not be taken to mean that the physical dynamics necessarily transforms the initial state into s_1 and subsequently s_1 into s_3 .

and show that in each single experiments the pointer observable has an unambiguous value. One can think of the transition $(\sum c_n |o_n\rangle) \otimes |I\rangle \rightarrow s_2$ as formally consisting of two parts: the Schrödinger time evolution $(\sum c_n |o_n\rangle) \otimes |I\rangle \rightarrow s_1$ and a transition $s_1 \rightarrow s_2$. I will call the transition $s_1 \rightarrow s_2$ *projection postulate*. In this terminology, reduction of the wave packet in single experiments implies the projection postulate for statistical experiments. The converse need not be true. Note that many authors use the terms projection postulate and reduction of the wave packet as synonyms.

Von Neumann introduced the projection postulate as a second kind of time evolution, independent of the usual Schrödinger time evolution. The problem now is the following: In a description of individual experiments, is it possible to derive the transition $(\sum c_n |o_n\rangle) \otimes |I\rangle \rightarrow s_3$ from the Schrödinger time evolution alone? If one is more modest, and is satisfied with a description of statistical experiments, one has to answer the following question: Is it possible

- (1) to derive the transition $(\sum c_n |o_n\rangle) \otimes |I\rangle \rightarrow s_2$ from one kind of time evolution, and
- (2) to show that in each single experiments the pointer observable has an unambiguous value?

In the traditional framework of quantum mechanics as proposed by von Neumann (1932) even part (1) of the more modest question is not soluble: The linearity of the Schrödinger time-evolution and the unrestricted validity of the superposition principle prevent a transition from $(\sum c_n |o_n\rangle) \otimes |I\rangle$ into s_2 .⁴ The Schrödinger time evolution is incompat-

⁴This was proved by Komar (1962), Wigner (1963), d'Espagnat (1966), Earman and Shimony (1968), Fine (1970), Busch, Lahti, and Mittelstaedt (1992), an probably many others.

ible with the final state being s_2 . Note that even if, for example by considering the joint system as open system evolving non-unitarily, one arrives by natural dynamics at a final state s_2 , it still would have to be shown that (2) is satisfied, i.e. that in each single experiment the pointer observable has an unambiguous value. This is the problem of whether or not the mixed state s_2 admits an ignorance interpretation.

Remarks on the above formulation of the measurement problem. The measurement problem only arises if after the experiment we are still interested in the state of the system. If we were only interested in the state of the apparatus and then infer from it the states of the system before the experiment, then the whole problem would not even arise. Both s_1 and s_2 yield the same state $\sum |c_n|^2 |p_n\rangle\langle p_n|$ when restricted to the apparatus by partial tracing. The decisive difference between the superposed and the mixed final states is the existence of correlations between apparatus and system in the superposed state. These correlations cannot be observed if one is only interested in the final state of the apparatus.

Let me remark here also that the notion of a measurement of the first kind actually changes the traditional idea of measurement. Traditional measurement are retrospective, whereas a measurement of the first kind is state-preparing, at least if the measured observable is non-degenerate. This distinction between state preparing and retrospective measurements is not important in classical physics, where the disturbances introduced by the act of measurement can be controlled and where one can safely assert that a quantity already had a certain value before the measurement. In quantum physics the situation is different. The introduction of state-preparing measurements instead of ret-

rospective measurements has serious ontological consequences. It is not always possible to say that *before* a state preparing measurement the measured quantity had an unambiguous value.

The introduction of state-preparing measurements is equivalent to the repeatability postulate. Nowadays it is known that measurement of observables which do not commute with all conserved quantities are never exactly repeatable.⁵ Furthermore, measurements of observables with continuous spectrum cannot be exactly repeatable.⁶ Also, realistic measurements are not instantaneous and not exactly repeatable. Rather, they are nowadays described by instruments⁷. In this context the assumption of repeatability is relaxed to allow for ϵ -repeatable, ϵ -preparatory, and ϵ -ideal measurements.⁸ But it would be false to blame the idealising character of first kind measurements for the measurement problem. The more realistic formulation of the measurement problem in terms of POV-measures does not solve the problem.⁹ Historically, it is the merit of the unrealistic idealisation of measurements as measurements of the first kind, that these novel features of quantum mechanics have been focused upon immediately.

The projection postulate and the classical pointer observable. If the pointer observable is classical, all observables have the same expectation values in the states s_1 and s_2 . The argument for this goes as follows. Let

⁵See Wigner (1952), Araki and Yanase (1960), and Yanase (1961).

⁶See Davies and Lewis (1970), Ozawa (1984), Holevo (1985).

⁷An *instrument* is defined in the following way. A statistical experiment with an input state ρ leads to a probability distribution $\Delta \mapsto p_\rho(\Delta)$ on the real line. For every Δ there is a linear map $I(\Delta)$ —an operation—transforming the ρ into the (non-normalised) final states such that $\text{tr}(I(\Delta)\rho) = p_\rho(\Delta)$. The operation-valued measure $\Delta \mapsto I(\Delta)$ is called an instrument.

⁸For $\epsilon \neq 0$, these notions are not equivalent. See for example Busch (1986), Busch and Lahti (1990).

⁹See Busch, Lahti, and Mittelstaedt (1991).

\mathcal{A} be the universal representation of the algebra of observables of the joint system. So \mathcal{A} is a C^* -algebra of operators on some non-separable Hilbert space \mathcal{H} . All states on \mathcal{A} are represented by vectors in \mathcal{H} . Denoting by $\mathbf{1} \otimes P$ the pointer observable we have $(\mathbf{1} \otimes P)|o_n \otimes p_n\rangle = \alpha_n|o_n \otimes p_n\rangle$, where α_n are the eigenvalues of the pointer observable. Assume that the pointer observable is classical, i.e. that it commutes with all other observables. From

$$\begin{aligned} \alpha_i \langle o_i \otimes p_i | A | o_k \otimes p_k \rangle &= \langle o_i \otimes p_i | (\mathbf{1} \otimes P) A | o_k \otimes p_k \rangle \\ &= \langle o_i \otimes p_i | A (\mathbf{1} \otimes P) | o_k \otimes p_k \rangle \\ &= \alpha_k \langle o_i \otimes p_i | A | o_k \otimes p_k \rangle \end{aligned}$$

and $\alpha_i \neq \alpha_k$ we conclude that

$$\langle o_i \otimes p_i | A | o_k \otimes p_k \rangle = 0 \quad \forall A \in \mathcal{A}$$

whenever $i \neq k$. Therefore

$$\begin{aligned} \text{tr}(A s_1) &= \sum_{ik} c_i \bar{c}_k \langle o_k \otimes p_k | A | o_i \otimes p_i \rangle \\ &= \sum_i |c_i|^2 \langle o_i \otimes p_i | A | o_i \otimes p_i \rangle \\ &= \text{tr}(A s_2). \end{aligned}$$

Thus, if the pointer observable is classical, then all observables have the same expectation values in the states s_1 and s_2 . Defining states as linear functionals which associate expectation values to observables, s_1 and s_2 are the same state. Therefore the projection postulate is not necessary if the pointer observable is classical.

4.1.2 The ignorance interpretation

For a satisfying description of statistical experiments we required that (1) the final state is s_2 , and that (2) in every single experiment the

pointer observable has an unambiguous value.

I will say that some statistical state admits an ignorance interpretation *with respect to the pointer observable* if the pointer observable has an unambiguous value in every constituent of all the ensembles described by the state. This is an ignorance interpretation with respect to quantities. Following d’Espagnat the term ignorance interpretation is often taken to refer something different: a statistical state admits an ignorance interpretation in the sense of d’Espagnat if every constituent of the ensembles described by the state is in a pure state. D’Espagnat showed that the mixed states of subsystems which result from partially tracing a pure entangled state of the composite system do not admit an ignorance interpretation in his sense.

In the sequel I will assume that observed system and apparatus form a closed system. This requires that the environment is regarded as part of the apparatus. I do this because in the case of open systems additional complications arise, due to d’Espagnat’s argument: it may happen that some state does not admit an ignorance interpretation in d’Espagnat’s sense. In order to exclude these complications I will only speak about closed systems. Statistical states of closed systems always admit an ignorance interpretation in d’Espagnat’s sense. Still, we have to deal with the problem that statistical states of closed systems do not in general admit an ignorance interpretation with respect to most observables.

In traditional quantum mechanics the decomposition of mixed states into pure states is in general not unique. This follows from the following theorem¹⁰: The decomposition of a state into pure states is unique if and only if the commutant of the GNS-representation associated with

¹⁰See Bratteli and Robinson (1979), theorem 4.2.3.

the state is abelian. This theorem implies that in general the decomposition of a state on a non-abelian algebra into pure states is not unique.

Only very special states can be decomposed in a unique way into pure states: If a state is pure, then its GNS-representation is irreducible, so its commutant equals the scalar multiples of the identity, which is abelian. Thus the decomposition of a pure state into pure states is unique. Furthermore, any state on a maximally abelian algebra has a unique decomposition into pure states. Thirdly, a mixed state ρ which is a discrete convex combination of mutually disjoint pure states can *only* be decomposed into these pure states.¹¹

Non-uniqueness of the decomposition into pure states is a major obstacle for an ignorance interpretation of the mixed state: in general many *different* ensembles are described by the same density matrix. The ensemble determines uniquely the density matrix, but the density matrix does not in general determine a unique ensemble. The different ensembles are equivalent for the purposes of prediction, but they are different in the sense that they are different collections of system states. In particular, there are ensembles described by the density matrix s_2 whose constituents are not eigenstates of the pointer observable. In many traditional interpretations of quantum mechanics, it is assumed that an observable has an unambiguous value only in eigenstates. In such interpretations it is not possible to say that if an ensemble is described by the density matrix s_2 , the pointer observable has in every single experiment an unambiguous value.

The ignorance interpretation and the classical pointer observable. According to Proposition 4 of Chapter 3, every pure state is an eigenstate

¹¹See Landsman 1991.

of all classical observables. Thus every decomposition into pure states is a decomposition into states in which all classical observables have dispersion free values. Since the individual constituents of every ensemble are in pure states, we arrive at the conclusion that a classical pointer observable has an unambiguous value in all constituents of all ensembles. Therefore it is justified to say that the pointer observable in each single experiment has a well-defined but perhaps unknown value.

Note that this result does not rely on any new interpretative rule (as for example those introduced by modal interpretations). It only relies on the traditional rule that an observable has an unambiguous value if its expectation value is dispersion free.

Let me relate this to the uniqueness of decomposing a mixed state. In a fully classical system, every state has a unique decomposition into pure states. Thus the ignorance interpretation does not make any problems in classical mechanics. If the system has classical and quantum properties, the situation is more intricate. In this case the decomposition into pure states is not in general unique. Even in the presence of classical observables a general density matrix does not determine a unique ensemble. States, pure or mixed, in which all classical observables have a dispersion free value are called factor states. The central decomposition theorem¹² says that the decomposition of every state into factor states is unique. Therefore any density matrix, although it does not specify a unique ensemble, specifies uniquely the *classical* properties of all ensembles described by it.

Summary and open problems. To sum up the results of section 4.1: If the pointer observable is classical, then (1) the projection postulate

¹²See e.g. Dixmier (1964a), théorème 8.4.2.

is not necessary, and (2) the pointer observable has a dispersion free value in every single experiment. Therefore both requirements for a satisfactory description of experiments are satisfied. But there is no reason to be euphoric. There are still many problems.

Firstly, these strong results only reflect how strong the assumption is that the pointer observable is classical. The real problem is whether and if yes when this assumption is justified. How can we have classical observables in a quantum world? In this chapter I will not deal with this problem. The emergence of classical observables in infinite¹³ quantum systems was discussed in Chapter 3.

Let me also mention that the results of chapter 1 suggest another line of argument for the existence of classical observables. In chapter 1 I argued that for an internal observer it is impossible to distinguish all states of a system in which she or he is contained: for each observer, states of the big system are indistinguishable if they have the same restriction to the observer. Which properties of the universe cannot be observed thus depends on the observer and on what he chooses to measure, but no observer can, not even in principle, measure all the properties of the universe. Even in a finite universe which is correctly described by traditional quantum mechanics, every observer can, as a matter of principle, discriminate only a restricted set of properties of the universe. Thus for each observer the algebra of operators which she or he can in principle observe, is smaller than the algebra of all properties of the universe. This, perhaps, can be turned into an argument for the existence of observer-dependent superselection rules.

But remember that the argument of chapter 1 relied on several

¹³For a debate on whether or not it is justified to assume that the apparatus has infinitely many degrees of freedom, see Bub (1988, 1989) and Robinson (1990, 1994).

assumptions. Firstly, the assumption of proper inclusion was made. If the observed system is at least partially outside the observer, then the conclusion does not apply. (But in this case, again the observer does not observe some part of the universe.) Secondly, the argument only applies if the observer insists on measuring present (and not past) properties of the universe. The algebra of accessible observables is thus time dependent and observer dependent, but it cannot be the algebra of all observables of the universe. I will not develop this line of argument further.

The second problem is related to the time evolution: How can the pure initial state $(\sum c_n |o_n\rangle) \otimes |I\rangle$ be transformed in finite time into the mixed state s_2 ? How can the components $|o_n\rangle \otimes |I\rangle$, which are in the same superselection sector evolve in finite time into different sectors? I will address this problem in section 4.4.

The problems mentioned so far just refer to a description of statistical experiments. In a description of individual experiments many new problems arise, in particular with respect to the stochastic time evolution. I will say nothing about these problems.

4.2 Inaccurate measurements of non-normal states?

According to von Neumann (1932, chapter III.3) an observable can be measured exactly if it possesses a purely discrete spectrum. Von Neumann admits that measurements of observables with at least partially continuous spectrum have to be inaccurate.

In this section I will investigate von Neumann's claim that measurements of quantities with uncountable spectrum are necessarily in-

accurate. Then I will discuss von Neumann's idea that only normal states are experimentally accessible, and discuss its relevance in the more general framework of algebraic quantum mechanics. This more general framework allows for a description of classical systems, and the analogue of von Neumann's claim for classical systems can be discussed. Finally I will examine assumptions implicit in the claim that only normal states are experimentally accessible. The basic notions used in this section are collected in the following

Definition 22 *Let \mathcal{M} be a von Neumann algebra. A positive normalised linear functional on \mathcal{M} is called a state. The set of all states is convex, its extremal points are called pure states. Every von Neumann algebra \mathcal{M} is the dual of a Banach space \mathcal{M}_* . A state ρ is called normal if one of the following three equivalent¹⁴ conditions is fulfilled.*

1. ρ is in \mathcal{M}_* .
2. ρ is completely additive¹⁵, i.e. $\rho(\sum_{k \in I} P_k) = \sum_{k \in I} \rho(P_k)$ for any family of mutually orthogonal projections $P_k \in \mathcal{M}$.
3. ρ is σ -weakly continuous.

Denote the set of states by \mathcal{S}^* and the set of normal states by \mathcal{S}_* . In general $\mathcal{S}_* \subseteq \mathcal{S}^*$.

In a lattice theoretic framework, countable additivity was introduced in Definition 3 in Chapter 2.

¹⁴See Kadison and Ringrose (1986, theorem 7.1.12).

¹⁵If \mathcal{M} operates on a separable Hilbert space, complete additivity reduces to countable additivity. Instead of countably additive I will sometimes write σ -additive.

4.2.1 Non-normal states in traditional quantum mechanics

In traditional quantum mechanics the set of observables is identified with the set of all self-adjoint operators on a separable Hilbert space. The normal states on $\mathcal{B}(\mathcal{H})$ are exactly those ρ for which there is a positive trace-class operator T on \mathcal{H} with trace-norm one such that

$$\rho(A) = \text{tr}(TA), \quad A \in \mathcal{B}(\mathcal{H}).$$

So the normal states are in one-to-one correspondence with the density matrices. It is important that $\mathcal{S}_* \neq \mathcal{S}^*$ if and only if the Hilbert space \mathcal{H} is infinite dimensional.¹⁶

The notion of state as a positive linear functional on the observables is a mathematical one. It is not clear whether all states in this mathematical sense should be regarded as states of a physical system. Since, on an infinite dimensional Hilbert space, $\mathcal{S}_* \subset \mathcal{S}^*$, one faces the choice of whether or not to require states of a physical system to be normal. Correspondingly one chooses \mathcal{S}_* or \mathcal{S}^* as set of physical states. Both choices are supported by good arguments.

One argument in favour of \mathcal{S}^* goes as follows. Neighbourhoods in the weak*-topology are defined with respect to finite sets of observables by

$$\mathcal{U}_{A_1, \dots, A_n; \epsilon}(\rho) := \{\rho' \in \mathcal{B}(\mathcal{H})^* : |\rho(A_i) - \rho'(A_i)| < \epsilon, i = 1, 2, \dots, n\}.$$

Such a neighbourhood contains all states ρ' whose expectation values for the observables A_1, \dots, A_n differ by less than ϵ from those of ρ . The weak*-topology on the state space is claimed to be the experimentally relevant topology because it is determined by the observables

¹⁶See e.g. Emch 1984, p. 374.

A_1, \dots, A_n which are actually measured. In contradistinction to this, the neighbourhoods in the norm-topology on $\mathcal{B}(\mathcal{H})^*$ are determined by all observables, measured or not. In the weak*-topology, \mathcal{S}^* is closed, but \mathcal{S}_* is not. In fact \mathcal{S}^* is the weak*-closure of \mathcal{S}_* . If we want to require that a Cauchy sequence of physical states should converge to a physical state, then we should take \mathcal{S}^* as space of physical states.

I think that this argument is vulnerable. There are two reasons for this. Firstly, a realist might prefer to judge similarity of states with reference to all properties of the physical system, not just with respect to those which happen to be measured. From this point of view, the norm topology on $\mathcal{B}(\mathcal{H})^*$ is more appropriate. If one requires closure of the set of physical states with respect to the norm topology, then \mathcal{S}_* does not fare worse than \mathcal{S}^* . They are both closed with respect to the norm topology. Secondly, it seems that even an operationalist need not support the argument. Surely, the operationalist would choose the weak*-topology to judge the similarity of states. But for him there is no reason to require that the set of physical states should be closed. Every state in \mathcal{S}^* can be approximated with arbitrary precision by states in \mathcal{S}_* , and this may be enough for the operationalist. After all, a full-fledged operationalist would like to rely only on realistic experiments, which are, as a rule, inaccurate.

Perhaps a stronger argument for the choice of \mathcal{S}^* rather than \mathcal{S}_* is that non-normal states are widely used when we deal with observables with continuous spectrum. Let A be a self-adjoint operator with purely continuous spectrum on a separable Hilbert space \mathcal{H} . It can be shown¹⁷ that for every a in the spectrum of A there is a pure state ρ_a such that $\rho_a(A) = a$ and $\rho_a((A - a\mathbf{1})^2) = 0$. Therefore $\rho_a(A^2) = \rho_a(A)^2$. Thus the

¹⁷See Segal (1947b) or Emch (1984, p. 374).

observable A has a dispersion free expectation value a in the pure state ρ_a . But ρ_a is not of the form $|\psi\rangle\langle\psi|$ for any vector $\psi \in \mathcal{H}$.¹⁸ Therefore observables with uncountable spectrum do not have eigenvectors in the strict sense, if they are represented by self-adjoint operators on a separable Hilbert space.

The notion of an eigenvector is very important in most traditional interpretations of quantum mechanics: There is not much disagreement about the interpretative rule that if a state is an eigenvector of a certain self-adjoint operator, then the corresponding observable has a well defined value in this state.¹⁹ In this formulation the rule cannot be applied to observables with continuous spectrum, because these do not have eigenvectors.

It is obvious how to generalise the above interpretational rule so that it can be applied to observables with continuous spectrum. One observes that if a vector ψ_a is an eigenstate for the eigenvalue a of an observable A with discrete spectrum, then the expectation value of A in ψ_a is dispersionfree, $\langle\psi_a|A^2|\psi_a\rangle = a^2 = \langle\psi_a|A|\psi_a\rangle^2$. This is the property which can be generalised to continuous observables: One says that if an observable has a dispersion free expectation value in a certain state, it *has* this value. (Such states are called generalised, or improper, eigenvectors.)

States in which observables with continuous spectrum have a dispersionfree expectation value are not normal. Such observables have a

¹⁸An intuitive argument for this is very simple. Assume that all the states ρ_a , $a \in \sigma(A)$, are of the form $|\psi_a\rangle\langle\psi_a|$, where ψ_a is a normalised eigenvector of A . Then $\langle\psi_a|\psi_b\rangle = 0$ for $a \neq b$. But in a separable Hilbert space there cannot be uncountably many orthogonal unit vectors.

¹⁹Even modal interpretations, as for example those of Kochen (1985), Dieks (1989), Healey (1989), van Fraassen (1991), and Bub (1992, 1993), accept this rule. They deny however the converse.

dispersionfree value only in non-normal states. According to the above rule, such observables *have* a certain value only if they are in non-normal states.

So we arrive at a stronger argument against the choice \mathcal{S}_* as the set of physical states: If one excludes non-normal states from the set of physical states, one is led to the strange situation that observables with continuous spectrum do not have a well defined value in any state.

In spite of these arguments, von Neumann (1932) identified the states of a quantum mechanical system with the density matrices and thus excluded non-normal states. What are the reasons to choose \mathcal{S}_* rather than \mathcal{S}^* as set of physical states in quantum mechanics? A first point is that non-normal states do not necessarily satisfy the Jauch-Piron condition. A state ρ on $\mathcal{B}(\mathcal{H})$ (or more generally on a von Neumann algebra) is said to satisfy the Jauch-Piron condition if $\rho(P) = \rho(Q) = 1$ implies $\rho(P \wedge Q) = 1$ for all projections $P, Q \in \mathcal{B}(\mathcal{H})$. For states which do not fulfil the Jauch-Piron condition there are projections P, Q both with expectation value one, which are not both true. Amann (1987b) proved that all pure states fulfilling the Jauch-Piron condition are normal; furthermore all normal states satisfy the Jauch-Piron condition.

Another argument for the choice of \mathcal{S}_* is that non-normal states are not completely additive. (On separable Hilbert spaces this is equivalent to the breakdown of countable additivity.) The requirement of countable additivity comes from probability theory. Assume that the statistics of experiments are described by a classical probability space (Ω, Σ, μ) . Here Ω denotes the sample space and consists of the possible results in single experiments. The space of events Σ is taken to be a

Boolean σ -algebra. In classical probability theory Σ is a σ -complete²⁰ Boolean lattice. μ is a σ -additive probability measure on the event space Σ . It assigns to every event Δ in Σ the probability $\mu(\Delta) \in [0, 1]$ of this event. Since μ is σ -additive, it fulfils $\mu(\bigcup_{k=1}^{\infty} \Delta_k) = \sum_{k=1}^{\infty} \mu(\Delta_k)$ for every finite or countably infinite set of mutually disjoint Δ_k .

In Hilbert space quantum mechanics the set $\mathcal{P}(\mathcal{H})$ of projectors on subspaces of \mathcal{H} is an orthocomplemented lattice. This is taken to be the event space Σ . A statistical state ρ now can be regarded²¹ as a probability measure assigning to every event P in the lattice the probability $\rho(P)$ that the state ρ of the system is in the subspace $P\mathcal{H} \subset \mathcal{H}$. If we want ρ to induce a probability measure, then this measure should be, by analogy to classical probability theory, σ -additive: $\rho(\bigcup_{k=1}^{\infty} P_k) = \sum_{k=1}^{\infty} \rho(P_k)$ for any finite or countably infinite family of mutually orthogonal events. This condition is fulfilled if and only if ρ is normal.

If one describes results of experiments by classical probability theory, one requires states to be normal. A non-normal state does not induce a σ -additive measure on $\mathcal{P}(\mathcal{H})$. Therefore, as long as one is exclusively concerned with the description of statistical experiments, it is justified to choose \mathcal{S}_* as the set of physical states.

Now we are in a position to understand von Neumann's (1932) claim that observables with continuous spectrum cannot be measured exactly. Von Neumann makes the assumptions that

- (1) The Hilbert spaces in question are separable and that
- (2) The state space of a physical system can be identified with the

²⁰A lattice is said to be σ -complete if every *countable* set of elements has an infimum and a supremum.

²¹See Definition 3 in Chapter 2.

density matrices.

Assumption (2) amounts to the choice of \mathcal{S}_* as set of physical states. Under these two assumptions, his claim that observables with continuous spectrum cannot be measured exactly emerges as a necessary consequence of the fact that states in which continuous quantities have a dispersion free value can not be represented as density matrices on separable Hilbert spaces. Thus there are no physical states in which a quantity with continuous spectrum has dispersion free values.

Even with a weaker assumption than (2) it is possible to give an argument why observables with continuous spectrum cannot be measured exactly. Instead of assumption (2) one can make the weaker assumption

(2') There exist normal states which reproduce the experimental probability distributions.

One can argue for (2') by saying that non-normal states do not induce a probability measure on the observables, and that the outcome of a statistical experiment is always represented by a probability measure. Once (2') is accepted, then the argument against the exact measurability of observables with continuous spectrum goes through smoothly: (1) implies that states in which such observables have dispersion free values are not normal, and (2') says that these states cannot be a possible outcome of a statistical experiment.

Assumption (2') obviously is weaker than assumption (2): it allows for non-normal states as possible states of physical systems, although they cannot be specified by experiments. (But they can be arbitrarily well approximated by states which can be specified experimentally.) Replacing (2) with (2') makes it possible that there are physical states in which an observable with continuous spectrum has dispersion free values. For this reason assumption (2') is perhaps preferable.

Still, assumption (2') is not weak. There are two scenarios in which it could be violated. Firstly, it may be the case that not every experiment is statistical. (But probably statistics is essential in order to make sense of the requirement that experiments should be in some sense repeatable.) Secondly, even if we restrict ourselves to statistical experiments, assumption (2') reflects the assumption that experimental probability distributions should be σ -additive. At the end of section 4.2.2, I will briefly discuss the consequences of giving up σ -additivity in uncountable statistical models. Also, it should be mentioned, and I will come back to this later, that the whole argument breaks down if we drop assumption (1) and do not restrict ourselves to separable Hilbert spaces.

4.2.2 Measurability of non-normal states in algebraic quantum mechanics

When we pass from $\mathcal{B}(\mathcal{H})$ to arbitrary von Neumann algebras \mathcal{M} , large parts of the above discussion still apply. There will, however, be two further complications. Firstly, if the algebra of observables is not $\mathcal{B}(\mathcal{H})$ but some other von Neumann algebra, one has to face the possibility that *no* pure state is normal. This is the case if the von Neumann algebra is a factor of type II or III. Secondly, the projection lattice of a general von Neumann algebra may not be atomic (so there may be propositions P such that there is no atom Q with $0 < Q < P$). In particular this is the case if the von Neumann algebra is a factor of types II or III, because the projection lattice of such von Neumann algebras does not contain any atoms.

Pattern of the argument. Central to the argument is now the following

Proposition 7 *Let \mathcal{M} be a von Neumann algebra, and ρ a state on \mathcal{M} . Let A be an observable such that the projection lattice of $\{A\}''$ is atom free. If $\rho(A)$ is dispersion free, $\rho(A)^2 = \rho(A^2)$, then ρ is not normal.*

PROOF: Denote by \mathcal{A} the commutative von Neumann algebra $\{A\}''$. Since $\rho(A)$ is dispersion free, $\rho(B)$ is dispersionfree for any $B \in \mathcal{A}$. So ρ is dispersion free on \mathcal{A} . Segal (1947a) showed that a state which is dispersion free on a von Neumann algebra is pure. Thus ρ is pure on \mathcal{A} . Plymen (1968) showed that atom free von Neumann algebras do not have any pure normal states. Therefore, since ρ is pure on \mathcal{A} , it cannot be normal on \mathcal{A} . A state which is normal on \mathcal{M} is also normal on any von Neumann subalgebra of \mathcal{M} . As ρ is not normal on $\mathcal{A} \subset \mathcal{M}$ it cannot be normal on \mathcal{M} . \square

The basic pattern of arguments is as follows. Assume (2') above. Together with Proposition 7 it implies that observables A for which the projection lattice of $\{A\}''$ is atom free cannot be measured exactly. (In the case of $\mathcal{M} = \mathcal{B}(\mathcal{H})$ this is precisely the argument against exact measurability of observables with continuous spectrum.)

Examples. There are two prominent examples which follow the pattern of this argument. The first concerns observables with purely continuous spectrum. The spectral resolution of such an observable is an abelian von Neumann algebra with atom free projection lattice. Then the above argument just reduces to the argument of section 4.2.1 that observables with continuous spectrum cannot be measured exactly.

The second example is provided by classical mechanics. As algebra of observables take $L^\infty(\Gamma)$, the functions on phase space essentially bounded with respect to the Liouville measure μ . There

is no point $\omega \in \Gamma$ with $\mu(\{\omega\}) \neq 0$. $L^\infty(\Gamma)$ is represented on the Hilbert space $L^2(\Gamma, \mu)$ of square integrable functions on the phase space by $(f\psi)(x) := f(x)\psi(x)$. Since $L^\infty(\Gamma)_* \cong L^1(\Gamma)$, the normal states on $L^\infty(\Gamma)$ correspond to the real-valued integrable functions ρ on Γ with $\int_\Gamma \rho(x)d\mu(x) = 1$. They are the probability measures on phase space. The projection lattice of $L^\infty(\Gamma)$ is given by the Borel subsets of Γ . To every Borel set Δ associate the projector P_Δ defined by $(P_\Delta\psi)(x) := \chi_\Delta(x)\psi(x)$, where χ_Δ is the characteristic function of the set Δ . The expectation value of a projector P_Δ in a state ρ is given by $\int_\Gamma \chi_\Delta(x)\rho(x)d\mu(x)$. P_Δ is the zero operator if $\mu(\Delta) = 0$.

In this example it is clear that (i) the projection lattice is atom-free and (ii) the pure states on $L^\infty(\Gamma)$ are not normal. The argument for (i) goes as follows: Take a projection $E(\Delta) \neq 0$ in the projection lattice of $L^\infty(\Gamma)$. $E(\Delta) \neq 0$ is equivalent with $\mu(\Delta) \neq 0$. (So Δ cannot consist of just one point.) Therefore it is possible to find a Borel set Δ' with $0 < \mu(\Delta') < \mu(\Delta)$ and $\Delta' \subset \Delta$. We have $0 < E(\Delta') < E(\Delta)$. To see (ii) one argues as follows: Plymen (1968, Lemmata 3.2, 3.3) showed that there is a one-to-one correspondence between the normal states on a von Neumann algebra in which all observables have a dispersion free value and the minimal projectors in the centre. Since $L^\infty(\Gamma)$ is commutative, it equals its centre. By (i) it does not have any atoms, and therefore there are no normal states on $L^\infty(\Gamma)$ in which all observables take dispersion free values. Every pure state on $L^\infty(\Gamma)$ takes a dispersion free value on all observables, and therefore cannot be normal. Thus there are no pure normal states on $L^\infty(\Gamma)$.

Assumption (2') together with (ii) above implies that in classical mechanics normal states cannot specify exactly phase space points. But for two reasons I think that, in this example, the conclusion is put

in by hand.

Firstly, it is not necessary to use an atom free lattice in classical mechanics. It is possible to use for classical mechanics an event space with atoms. Usually the event space Ω in a statistical description is a σ -complete Boolean algebra. According to a representation theorem of Loomis (1947) and Sikorski (1960, p. 117), any σ -complete Boolean algebra is isomorphic to a σ -complete Boolean algebra Σ/Δ of points sets Σ modulo a σ -ideal in that algebra. Given a phase space Γ , in order to arrive at an event space, one has to choose a σ -algebra Σ of subsets of Γ and an ideal Δ of negligible subsets such that Σ/Δ is the event space. It would be possible to choose as Σ all subsets of Γ and as Δ the trivial null ideal. (This amounts to choosing a discrete topology on Γ , and thus giving up separability of Σ/Δ .) Then the event space would consist of all Borel subsets of Γ , and this Boolean lattice has atoms .

Secondly, the result that in classical mechanics normal states cannot specify exactly phase space points is simply a consequence of taking $L^\infty(\Gamma)$ with respect to the Liouville measure μ . Birkhoff and von Neumann (1936) argue that it is experimentally unrealistic to identify each subset of the phase space with a proposition. (Furthermore this leads to non-separable event spaces.) Instead they propose to regard sets of Lebesgue measure zero as experimentally irrelevant and choose Δ to be the ideal of Lebesgue (or Liouville) null sets. $E(\{\omega\}) = 0$ is just a consequence of the choice of the Liouville measure, which in turn reflects the *assumption* that the points in phase space are not experimentally accessible. If $E(\{\omega\}) = 0$ were violated, it would not be possible to show that the projection lattice of $L^\infty(\Gamma)$ is atomfree and that there are no pure normal states on $L^\infty(\Gamma)$.

How conclusive is the argument? Let me now briefly discuss the assumptions entering into the argument.

One main assumption was (2'): Experiments only specify normal states. As justification of this assumption it is often taken that the probability distributions resulting from experiments should be σ -additive. Why should σ -additivity be a necessary condition for a probability measure on $\mathcal{P}(\mathcal{M})$ to be specifiable by a statistical experiment? It might seem that for practical purposes finite additivity could be sufficient.²² If this is the case then the non-normal states could not be excluded from describing the results of statistical experiments: *every* state ρ on a von Neumann algebra is finitely additive on its projection lattice. If we only require that statistical states are described by finitely additive measures, then any state, normal or not, is a statistical state and can describe the results of statistical experiments.

But probability theory of continuous models is difficult without σ -additivity: the Lebesgue monotone convergence theorem breaks down, and expectation values cannot be properly defined. Furthermore, even if, for sufficiently regular probability distributions, the expectation value can be defined, the strong law of large numbers does not hold. One cannot be almost certain that the empirical average of a long series of independent trials converges to the expectation value. This makes the interpretation of probabilities and expectation values much more difficult. Therefore, in a continuous model, I would rather stick to the requirement of σ -additivity.

Another way to circumvent the problem of non-normal states is to restrict oneself to discrete, or even finite models. Primas (1990a) asserts that all experiments can be executed in a digital manner and

²²This point of view was taken by de Finetti (1972).

have to be described by a *finite* Boolean algebra. If the sample space of an analogue output signal is uncountable, one is forced to use some classification method together with some statistical decision procedure. By introducing a finite partition $\{B_i\}$ of the sample space Ω (i.e. $B_i \cap B_k = \emptyset$ for $i \neq k$, $\bigcup_{i=1}^n B_i = \Omega$), the observational data are classified into disjoint groups. Obviously on a finite Boolean algebra finite additivity and σ -additivity coincide. Thus all states can be regarded as statistical states.

Let me finally note that the whole problem of non-normality would not arise if we admitted non-separable Hilbert spaces. For then one can replace the von Neumann algebra \mathcal{M} by its universal representation to which it is isomorphic. The universal representation equals the bidual \mathcal{M}^{**} . Since $\mathcal{M}^* = (\mathcal{M}^{**})_*$, every state on \mathcal{M} corresponds to a normal state on the universal representation. The price one has to pay for this nice feature is that, as a rule, the universal representation is on a non-separable Hilbert space.

4.3 Inaccurate experiments can only be described by a classical pointer

In section 4.1 I discussed some known arguments for and against the classical pointer observable. In this chapter I will give another reason why the pointer observable should be classical: inaccurate measurements can only be described by a classical pointer.

Apart from the fact that realistic experiments are always inaccurate, there are some arguments why experiments, as a matter of principle, cannot be accurate. Whether measurements of observables with continuous spectrum can be accurate was discussed in the previous section.

As far as observables with discrete spectrum are concerned, von Neumann (1932) claimed that observables with discrete spectrum can be measured exactly. This claim about the exact measurability of discrete observables is relativised by results of Wigner (1952) and Araki and Yanase (1960), who proved that even a discrete observable cannot be measured exactly if it does not commute with all conserved quantities. The well known theorem of Wigner (1952), Araki and Yanase (1960) says that it is impossible to measure exactly the value of an observable with discrete spectrum, if this observable does not commute with all conserved quantities. An approximate measurement of such an observable can be made with the help of an apparatus which is large enough in the sense that the mean square value of the conserved quantity is large.²³ The bigger this value, the more precise the measurement can be at least in principle. Ozawa (1984), and in a different context Holevo (1985), showed that on a separable Hilbert space there is no exactly repeatable measurement of an observable with continuous spectrum.

One might think of other reasons why experiments are inaccurate. Remember the result of Chapter 1: for an observer it is not possible to distinguish all states of a system in which she or he is properly contained. But this argument cannot in general be taken to imply the impossibility of accurate experiments. Firstly, the argument merely shows that it is impossible, under certain circumstances, to measure *states* exactly. But it leaves open the possibility of measuring quantities exactly. (The only case when an exact measurement of a quantity is also an exact state measurement is an ideal first kind measurement of a quantum observable with non-degenerate spectrum. Such a mea-

²³See Yanase 1961.

surement determines some states exactly, but not all.) Secondly, the argument only applies if the observed system properly contains the measurement apparatus. It does not apply if at least some part of the apparatus is not part of the observed system. Thirdly, the argument does not apply if the apparatus does not determine the present state of the observed system, but rather some earlier state.

In section 4.3.1 I will precisely formulate what I mean by an inaccurate measurement. Then, in the central sections 4.3.2 and 4.3.3, I will show that measurements which are inaccurate in this sense can only be described by a classical pointer observable.

4.3.1 The requirement of finite measurement accuracy

In this section I give two examples of inaccurate experiments and then abstract from them a notion of inaccuracy which will be used in this section.

First example: consider the case where a digital pointer is used to measure a quantity A with continuous spectrum. So the spectrum of A is partitioned into finitely many intervals. Let K be the index set for the intervals, and let each interval be labeled by a number k in it. The pointer reading k is the same for all states with a dispersion free expectation value in a given interval k . After a measurement, when the pointer reads k , we register as final state of the observed system the one with dispersion free value k , even if the actual final state is one in which A has some other value in the interval.

Second example: we measure the spin of a particle in a given direction z . A Stern-Gerlach experiment is inaccurate in the following way. If the receptor of the counter has finite area, then the counter registers

particles even if the direction of the magnetic field deviates by a small angle from z and the particle therefore takes a slightly different trajectory. Therefore these particles give rise to the same pointer reading as particles with spin exactly in z direction.

These examples suggest the following intuitive notion of inaccuracy: If an experiment is inaccurate, then to a given pointer reading we associate not only one state but a neighbourhood in the state space.

Let us try to formulate this notion of inaccuracy in a more general framework. Assume that the observables of the measured system and of the apparatus generate von Neumann algebras \mathcal{A}_S , \mathcal{A}_M respectively. Let ρ_k be some state of the joint system after an experiment leading to the result $k \in K$. Let K be the set of all possible results. (In the example of a Stern-Gerlach experiment on a spin 1/2 particle, K would be a set with two elements, or four if firing of no counter or of both counters is regarded a possible result.) Assume that the states ρ_k are normal and factorial.²⁴ Furthermore, assume that there is one observable $P \in \mathcal{A}_S \overline{\otimes} \mathcal{A}_M$ specified which associates the measurement results $\rho_k(P)$ to the typical final states ρ_k ; this is the pointer observable. It is convenient but not necessary to take $\rho_k(P) = k$.

Definition 23 *Let $\mathcal{A}_S \overline{\otimes} \mathcal{A}_M$ be a von Neumann algebra describing the observables of the joint system. An experiment is said to be of finite accuracy if there exists a $\delta > 0$ and a non-empty family $\{\rho_k\}_{k \in K}$ of normal factor states of the joint system, such that the pointer observable P has the same expectation value in all pure states ρ of the joint system*

²⁴A state is called factorial or factor state, if all classical observables have a dispersion free value in this state. If all macroscopic observables are classical, one can say that factor states are “macroscopically pure”. All pure states are factor states. In traditional quantum mechanics, i.e. if there are no classical observables, all states, pure or mixed, are factor states.

satisfying $\|\rho - \rho_k\| < \delta$ for some k .

If the joint system is described by a C^* -algebra it is very artificial to assume the states ρ_k to be normal. Instead they are taken to be pure.

Definition 24 *Let $\mathcal{A}_S \otimes \mathcal{A}_M$ be a C^* -algebra describing the observables of the joint system. An experiment is said to be of finite accuracy if there exists a $\delta > 0$ and a non-empty family $\{\rho_k\}_{k \in K}$ of pure states of the joint system, such that the pointer observable P has the same expectation value in all pure states ρ of the joint system satisfying $\|\rho - \rho_k\| < \delta$ for some k .*

For brevity I will refer to the conditions in Definition 23 and 24 as FMA, for finite measurement accuracy. I will conclude this subsection with some remarks on these definitions.

Only for pure states in the neighbourhood is it required that they have the same expectation value as ρ_k . This is necessary for the following reason: If we admitted arbitrary mixed states, then one could easily violate FMA by taking a mixed state $c\rho_k + (1-c)\rho$ for some state ρ with $\rho_k(P) \neq \rho(P)$ and c close enough to one.

Secondly, the states ρ_k are assumed to be factor states (Definition 23) or pure states (Definition 24) because otherwise it might not be possible to find factor states or pure states in small neighbourhoods of ρ_k .

Thirdly, the assumption in Definition 23 that ρ_k be normal is not a restriction: in both the norm and weak*-topology, the normal states are dense in the state space of a von Neumann algebra. So for every neighbourhood one can choose a normal state in it.

Fourthly, it is worthwhile to note that the distance of states is taken here in the norm topology. It is defined by $\|\rho - \rho'\| = \sup_{A \in \mathcal{A}_S \otimes \mathcal{A}_M} |\rho(A) -$

$\rho'(A)/\|A\|$. For states in such a δ -norm neighbourhood, *all* observables have expectation values which differ by less than δ . Such states are indeed very similar. If the inaccuracy of our measurements is δ then states in such a norm-neighbourhoods cannot be distinguished by experimental means.

If the distance of states were taken in the weak*-topology, then FMA could never be satisfied. Neighbourhoods in the weak*-topology are defined with respect to sets of observables A_1, A_2, \dots, A_n by $|\rho(A_i) - \rho'(A_i)| < \delta$ for $i = 1, \dots, n$. If states are close by standards of the observables A_1, \dots, A_n , they may be very far apart by the standards of other observables. In fact, since the norm-topology is strictly finer than the weak*-topology, in every weak*-neighbourhood $\mathcal{U}_{A_1, \dots, A_n; \delta}(\rho)$ there is a state ρ' and an observable B such that $|\rho(B) - \rho'(B)| > \delta$. So a measurement of B with inaccuracy δ can still distinguish between ρ and ρ' . It would be too strong a requirement that measurements with inaccuracy δ cannot distinguish between states in a weak*-neighbourhood $\mathcal{U}_{A_1, \dots, A_n; \delta}$.

Fifthly, I do not claim that experiments which violate FMA are perfectly accurate. Take for example the case where there are several possible pointer observables, and we register a definite result if one of them has a definite value. This kind of inaccuracy is sometimes invoked in connection with the description of imperfect measurements in the modal interpretation of quantum mechanics. Another example of inaccuracy which does not meet the requirement FMA is a degenerate, quantum mechanical pointer observable. Because of its degeneracy, it cannot specify states uniquely: all states in an eigenspace correspond to one pointer reading. But of course there are pure states arbitrarily close to such a eigenspace which are still not in it. In such a state the

expectation value of the pointer is different and thus FMA is violated. Finally, one could give a different, weaker definition of inaccuracy by requiring that states in a certain neighbourhood lead to very similar but perhaps not equal expectation values. (The theorems to follow do not hold for this weaker notion of inaccuracy.)

Last but not least note that in FMA it is *not* required that the pointer observable has a dispersionfree value. This will follow as a corollary.

4.3.2 W^* -version of the proof

In this subsection we will assume that there is a W^* -algebra $\mathcal{A}_S \overline{\otimes} \mathcal{A}_M$ given and an observable P in $\mathcal{A}_S \overline{\otimes} \mathcal{A}_M$ which is the pointer observable. The following theorem shows that the condition FMA in Definition 23 can be fulfilled if and only if $\mathcal{A}_S \overline{\otimes} \mathcal{A}_M$ has non-trivial centre and a classical observable $\overline{P} \in \mathcal{Z}(\mathcal{A}_S \overline{\otimes} \mathcal{A}_M)$ yields the same pointer reading as P for the typical final states. So the typical final states cannot distinguish the pointer observable from the classical observable \overline{P} . Furthermore it shows that then the typical final states are disjoint. Thus the pointer observable takes *dispersion free* values in the typical final states and in all pure states in their neighbourhood. The proofs of the following Lemma and Theorem follow Breuer (1993a).

Lemma 7 *Let $\mathcal{A}_S, \mathcal{A}_M$ be W^* -algebras and $\{\rho_k\}_{k \in K}$ be a set of normal factor states on $\mathcal{A}_S \overline{\otimes} \mathcal{A}_M$. Let $P \in \mathcal{A}_S \overline{\otimes} \mathcal{A}_M$. Assuming that for all $Z \in \mathcal{Z}(\mathcal{A}_S \overline{\otimes} \mathcal{A}_M)$ there is a ρ_k with $\rho_k(Z) \neq 0$, the condition FMA in Definition 23 can be fulfilled if and only if $P \in \mathcal{Z}(\mathcal{A}_S \overline{\otimes} \mathcal{A}_M)$.*

PROOF: Denote by $\pi_{\rho_k}(\mathcal{A}_S \overline{\otimes} \mathcal{A}_M)$ the GNS-representation with respect to the state ρ_k of $\mathcal{A}_S \overline{\otimes} \mathcal{A}_M$ on the Hilbert space \mathcal{H}_{ρ_k} , where there is a

cyclic vector Ω_{ρ_k} such that $\rho_k(A) = \langle \Omega_{\rho_k} | \pi_{\rho_k}(A) | \Omega_{\rho_k} \rangle$ for all $A \in \mathcal{A}_S \bar{\otimes} \mathcal{A}_M$.

Step 1: We show that $\bigcap_{k \in K} \ker \pi_{\rho_k} = \{0\}$.

Take $A \in \bigcap_{k \in K} \ker \pi_{\rho_k}$. First of all we observe that A cannot be in $\mathcal{Z}(\mathcal{A}_S \bar{\otimes} \mathcal{A}_M)$: If A were in $\mathcal{Z}(\mathcal{A}_S \bar{\otimes} \mathcal{A}_M)$, then there would be a ρ_k with $\rho_k(A) \neq 0$ and therefore $A \notin \ker \pi_{\rho_k}$.

Since the states ρ_k are normal, $\bigcap_{k \in K} \ker \pi_{\rho_k}$ is a weakly closed two-sided ideal of $\mathcal{A}_S \bar{\otimes} \mathcal{A}_M$. It follows that there is a projection $Q \in \mathcal{Z}(\mathcal{A}_S \bar{\otimes} \mathcal{A}_M)$ such that $\bigcap_{k \in K} \ker \pi_{\rho_k} = Q(\mathcal{A}_S \bar{\otimes} \mathcal{A}_M)Q$ (see Bratelli and Robinson (1981), Proposition 2.4.22). So $\bigcap_{k \in K} \ker \pi_{\rho_k} \neq \{0\}$ would imply that $\bigcap_{k \in K} \ker \pi_{\rho_k} \cap \mathcal{Z}(\mathcal{A}_S \bar{\otimes} \mathcal{A}_M) \neq \{0\}$. But since there is no central element in $\bigcap_{k \in K} \ker \pi_{\rho_k}$, $\bigcap_{k \in K} \ker \pi_{\rho_k} = \{0\}$.

Step 2: $P \notin \mathcal{Z}(\mathcal{A}_S \bar{\otimes} \mathcal{A}_M)$ implies $(\forall \delta)(\exists k \in K)(\exists \rho \text{ pure state}) : \|\rho - \rho_k\| < \delta, \rho(P) \neq \rho_k(P)$.

$P \notin \mathcal{Z}(\mathcal{A}_S \bar{\otimes} \mathcal{A}_M)$ implies that there exists a $B \in \mathcal{A}_S \bar{\otimes} \mathcal{A}_M$ such that $[P, B] \neq 0$. From the first step it follows that there is a k with $\pi_{\rho_k}([P, B]) \neq 0$. So $\pi_{\rho_k}(P) \notin \mathcal{Z}(\pi_{\rho_k}(\mathcal{A}_S \bar{\otimes} \mathcal{A}_M))$. Since $\pi_{\rho_k}(\mathcal{A}_S \bar{\otimes} \mathcal{A}_M)$ is a factor, $\pi_{\rho_k}(P)$ cannot be constant. Take any vector state $\psi \in \mathcal{H}_{\rho_k}$ with $\langle \psi | \pi_{\rho_k}(P) | \psi \rangle \neq \langle \Omega_{\rho_k} | \pi_{\rho_k}(P) | \Omega_{\rho_k} \rangle = \rho_k(P)$. Define $\psi_\alpha := \alpha \Omega_{\rho_k} + (1 - \alpha)\psi$ for $0 < \alpha < 1$. Take $\rho_\alpha(A) := \langle \psi_\alpha | \pi_{\rho_k}(A) | \psi_\alpha \rangle$ for all $A \in \mathcal{A}_S \bar{\otimes} \mathcal{A}_M$. Now it can be checked that all ρ_α are pure states satisfying $\rho_\alpha(P) \neq \rho_k(P)$. If we choose α close enough to one, then $\|\rho_\alpha - \rho_k\| < \delta$.

Step 3: $P \in \mathcal{Z}(\mathcal{A}_S \bar{\otimes} \mathcal{A}_M)$ implies $(\exists \delta)(\forall k \in K)(\forall \rho \text{ pure}) : \|\rho - \rho_k\| < \delta \Rightarrow \rho(P) = \rho_k(P)$.

Take $\delta < \inf_{i \neq k} \|\rho_i - \rho_k\|/2 < 2$ and let ρ be a pure state with $\|\rho - \rho_k\| < \delta < 2$ for some k . From Glimm and Kadison (1960) it follows that the representations π_ρ and π_{ρ_k} are quasi-equivalent. All

classical observables have the same value in the states ρ and ρ_k , so $\rho(P) = \rho_k(P)$. \square

Theorem 8 *Let $\mathcal{A}_S, \mathcal{A}_M$ be W^* -algebras and $\{\rho_k\}_{k \in K}$ be a set of normal factor states on $\mathcal{A}_S \overline{\otimes} \mathcal{A}_M$. Let $P \in \mathcal{A}_S \overline{\otimes} \mathcal{A}_M$. The condition FMA in Definition 23 can be fulfilled if and only if there is an observable \overline{P} in $\mathcal{Z}(\mathcal{A}_S \overline{\otimes} \mathcal{A}_M)$ with $\rho_k(P) = \rho_k(\overline{P})$, $\forall k \in K$. In this case the states $\{\rho_k\}_{k \in K}$ are mutually disjoint because for $i \neq k$ and $\overline{P} \in \mathcal{Z}(\mathcal{A}_S \overline{\otimes} \mathcal{A}_M)$, $\rho_k(\overline{P}) \neq \rho_i(\overline{P})$.*

PROOF: The family $\{\rho_k\}_{k \in K}$ defines a two-sided ideal $\mathcal{I} \subset \mathcal{A}_S \overline{\otimes} \mathcal{A}_M$ by

$$\mathcal{I} := \{A \mid \rho_k(B^*AC) = 0, \forall B, C \in \mathcal{A}_S \overline{\otimes} \mathcal{A}_M, \forall k \in K\}.$$

Note that $\mathcal{I} = \bigcap_{k \in K} \ker \pi_{\rho_k}$. There exists central projection $Q \in \mathcal{A}_S \overline{\otimes} \mathcal{A}_M$ such that $\mathcal{I} = (\mathcal{A}_S \overline{\otimes} \mathcal{A}_M)Q = Q(\mathcal{A}_S \overline{\otimes} \mathcal{A}_M)Q$. $\mathcal{A}_S \overline{\otimes} \mathcal{A}_M$ can be written as $\mathcal{A}_S \overline{\otimes} \mathcal{A}_M = (\mathcal{A}_S \overline{\otimes} \mathcal{A}_M)(\mathbf{1} - Q) \oplus (\mathcal{A}_S \overline{\otimes} \mathcal{A}_M)Q =: \mathcal{A}_1 \oplus \mathcal{A}_2$. Since for every $Z' \in \mathcal{Z}(\mathcal{A}_1)$ there is a ρ_k with $\rho_k(Z') \neq 0$, we have $\bigcap_{k \in K} \ker \pi_{\rho_k}(\mathcal{A}_1 \oplus 0) = \{0\}$. Every observable $A \in \mathcal{A}_S \overline{\otimes} \mathcal{A}_M$ can be written as $A = A' \oplus A''$ with $A' \in \mathcal{A}_1, A'' \in \mathcal{A}_2$; every state ρ on $\mathcal{A}_S \overline{\otimes} \mathcal{A}_M$ as $\rho = \rho' + \rho''$.

As $\bigcap_{k \in K} \ker \pi_{\rho_k}(\mathcal{A}_1 \oplus 0) = \{0\}$, everything now follows from the proof of the Lemma: If there is a $\overline{P} \in \mathcal{Z}(\mathcal{A}_S \overline{\otimes} \mathcal{A}_M)$ with $(\forall k \in K) : \rho_k(P) = \rho_k(\overline{P})$, then all factor states ρ with $\|\rho - \rho_k\| < \delta$ satisfy $\rho(\overline{P}' \oplus 0) = \rho'(\overline{P}') = \rho'_k(\overline{P}') = \rho_k(\overline{P}' \oplus 0)$. So FMA is satisfied for the pointer observable $\overline{P}' \oplus 0 =: \overline{P}$. If there is no $\overline{P} \in \mathcal{Z}(\mathcal{A}_S \overline{\otimes} \mathcal{A}_M)$ with $(\forall k \in K) : \rho_k(P) = \rho_k(\overline{P})$, then $P' \notin \mathcal{Z}(\mathcal{A}_1)$. From the second step of the proof of the Lemma it follows that $(\forall \delta)(\exists k \in K)(\exists \rho'$ factor state on $\mathcal{A}_1) : \|\rho' - \rho'_k\| < \delta$ but $\rho'_k(P') \neq \rho'(P')$. Taking $\rho := \rho' + \rho''_k$ it follows that $\|\rho - \rho_k\| < \delta$ but $\rho(P) \neq \rho_k(P)$. \square

Corollary 3 *If the condition FMA of Definition 23 is fulfilled, then the pointer observable P has a dispersion-free value in the typical final states ρ_k .*

The drawback of Theorem 8 is that as long as the typical final states ρ_k are assumed to be normal and the W^* -algebra $\mathcal{A}_S \overline{\otimes} \mathcal{A}_M$ can be represented faithfully on a separable Hilbert space, there cannot be a continuous observable on which they take different dispersionfree values. Therefore, this pure W^* -framework is not suitable for the description of a continuous pointer observable, except perhaps in Piron's (1976) approach which always uses a discrete topology but can represent observables with uncountable spectrum only on non-separable Hilbert spaces.

4.3.3 C^* -version of the Proof

Theorem 8 uses a pure W^* -framework. The assumption that the typical final states are normal is sensible in a statistical description because otherwise these states would not be σ -additive and therefore would not induce probability measures.

A C^* -algebraic version of Theorem 8 meets two difficulties: Firstly, the assumption that the typical final states ρ_k are normal cannot be justified. In fact normality is not a very natural concept for states on C^* -algebras. The C^* -framework is not suitable for a statistical description but rather for an individual one. In an individual description it is justified to assume the typical final states to be pure. Secondly, the basic C^* -algebra of quasilocal observables usually has trivial centre so that we do not have any classical observables. In this section I try to weld together these two problems and give a C^* -algebraic version of Theorem 8. In doing so I will follow Breuer, Landsman, and Amann (1993a).

As an abstract, representation independent C^* -analogue of Theorem 8 one could show the following.

Let $\{\rho_k\}_{k \in K}$ be a family of pure states on a C^ -algebra $\mathcal{A}_S \otimes \mathcal{A}_M$. This family defines a two-sided ideal \mathcal{I} by*

$$\mathcal{I} := \{A \in \mathcal{A}_S \otimes \mathcal{A}_M : \rho_k(B^*AC) = 0, \forall k \in K, \forall B, C \in \mathcal{A}_S \otimes \mathcal{A}_M\}.$$

Then the condition FMA in Definition 24 can be satisfied if and only if there exists an observable $\bar{P} \in \mathcal{Z}(\mathcal{A}_S \otimes \mathcal{A}_M/\mathcal{I})$ such that $\rho_k(P) = \rho_k(\bar{P})$ for all $k \in K$.

I will not prove this proposition because it is not very useful. The algebra of quasilocal observables, obtained as a C^* -inductive limit, usually has trivial centre and is simple. Since it has trivial centre we cannot find a classical pointer observable in it. Since it is simple the ideal \mathcal{I} defined by the $\{\rho_k\}_{k \in K}$ is the null-ideal, so $\mathcal{Z}(\mathcal{A}_S \otimes \mathcal{A}_M/\mathcal{I})$ is again trivial. Therefore in most quantum theories the above proposition is empty.

This brings our attention to a fundamental problem rooted in the use of a classical pointer observable: How does a classical observable arise in a quantum system? Usually the algebra of quasilocal observables is simple. Classical observables are often only in the weak closure of a particular representation appropriate to describe the particular situation under consideration. In the same way, the pointer observable will usually not be an element of the basic C^* -algebra $\mathcal{A}_S \otimes \mathcal{A}_M$. Rather it will be in the weak closure of a representation appropriate to describe measurement situations. The main goal is to find such a representation. So it is natural to change strategy and abandon the search for an abstract, representation independent version of Theorem 8. It is

more appropriate to look for a representation of the basic C^* -algebra $\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}}$ such that the pointer observable is contained at least in the weak closure and that we recover the W^* -algebraic situation of Theorem 8.

To see the situation more clearly, let us summarise first the typical features of realistic measurement situations. The basic C^* -algebra $\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}}$ generated by the local observables of the joint system is given. Furthermore, we have a family $\{\rho_k\}_{k \in K}$ of typical final states. In an individual description we will require them to be pure. There is a pointer value functional p on the states (so p is an element of the bidual $(\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}})^{**}$) which has different values on the typical final states. Furthermore, we know that realistic measurements are inaccurate in the way described in section 4.3.1: there is a $\delta > 0$ such that for all pure states ρ on $\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}}$ we have $\|\rho - \rho_k\| < \delta$ for some $k \Rightarrow \langle p, \rho_k \rangle = \langle p, \rho \rangle$.

The goal now is to find a representation of the basic C^* -algebra $\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}}$ which can incorporate these features. The existence of a representation describing this situation is guaranteed by the following

Theorem 9 *Let $\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}}$ be the simple C^* -algebra of observables of the joint system, \otimes denoting the minimal $*$ -tensor product. Assume that there is a set $\{\rho_k\}_{k \in K}$ of pure typical final states on $\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}}$ and a pointer functional $p \in (\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}})^{**}$ on the states such that the requirement FMA of Definition 24 is fulfilled:*

$$((\exists \delta > 0)(\forall \rho \text{ pure states})) : \|\rho - \rho_k\| < \delta \text{ for some } k \Rightarrow \langle p, \rho_k \rangle = \langle p, \rho \rangle .$$

Then there is a representation π_0 of $\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}}$ such that the states ρ_k can be extended to normal states $\overline{\rho}_k$ on $\pi_0(\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}})''$, and that there is an observable $P \in \pi(\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}})''$ with $\overline{\rho}_k(P) = \langle p, \rho_k \rangle$ for

all k . Furthermore, P is in the centre of $\pi_0(\mathcal{A}_S \otimes \mathcal{A}_M)''$ and therefore classical. The states $\overline{\rho}_k$ (as well as the ρ_k) are mutually disjoint, because $\overline{\rho}_k(P) = \langle \rho_k, p \rangle \neq \overline{\rho}_i(P) = \langle \rho_i, p \rangle$ for $i \neq k$.

PROOF: Embed $\mathcal{A}_S \otimes \mathcal{A}_M$ canonically into its enveloping von Neumann algebra \mathcal{B} , which is the weak closure of its universal representation. As a Banach space \mathcal{B} is isomorphic to the bidual $(\mathcal{A}_S \otimes \mathcal{A}_M)^{**}$. To this imbedding there corresponds an isometric isomorphism of $(\mathcal{A}_S \otimes \mathcal{A}_M)^*$ onto the predual of \mathcal{B} , associating to each state ρ on $\mathcal{A}_S \otimes \mathcal{A}_M$ a normal state $\tilde{\rho}$ on \mathcal{B} . Since the states ρ_k on $\mathcal{A}_S \otimes \mathcal{A}_M$ are pure, the W^* -algebraic version of FMA is satisfied for the pointer functional p and the typical final states $\tilde{\rho}_k$. From Theorem 8 it follows that there is a central projection $Q \in \mathcal{Z}(\mathcal{B})$ such that $\mathcal{Z}(Q\mathcal{B})$ contains an observable P on which the restrictions of the $\tilde{\rho}_k$ take the same values as the pointer functional P does on the ρ_k . Now take as π_0 the subrepresentation of the universal representation specified by $\pi_0(\mathcal{A}_S \otimes \mathcal{A}_M)'' = Q\mathcal{B}$. Since $\mathcal{A}_S \otimes \mathcal{A}_M$ is simple it is isomorphic to $\pi_0(\mathcal{A}_S \otimes \mathcal{A}_M)$. π_0 satisfies the conditions we mentioned in the proposition: the states ρ_k extend to states $\overline{\rho}_k$ which are defined by $\overline{\rho}_k(QA) = \tilde{\rho}_k(A)$, $A \in \mathcal{B}$, the pointer observable P is in $\pi_0(\mathcal{A}_S \otimes \mathcal{A}_M)''$ and fulfils $\overline{\rho}_k(P) = \langle p, \rho_k \rangle$, and FMA is satisfied. □

Let me make some remarks on this result. Firstly, even if $\mathcal{A}_S \otimes \mathcal{A}_M$ has a trivial centre, $(\mathcal{A}_S \otimes \mathcal{A}_M)^{**}$ does not if $\mathcal{A}_S \otimes \mathcal{A}_M$ has disjoint representations. This is for example the case if $\mathcal{A}_S \otimes \mathcal{A}_M$ is the C^* -algebra arising as C^* -inductive limit of finite systems. Secondly, knowing from experimental experience that measurements are inaccurate, Theorem 9 guarantees the existence of a representation which can describe this sit-

uation. It turns out that in such a representation the pointer observable is classical.

Theorem 9 is not satisfactory because from a physical point of view it would have been justified to require the representation π to be separable. Theorem 9 does not prove the existence of a separable representation fulfilling all the other requirements. If the pointer functional has a continuous range of possible values, then the representation π_0 constructed in the proof of Theorem 9 is surely not separable. Another drawback of Theorem 9 is that the states $\overline{\rho_k}$ are normal. This is not what is needed in the case of a continuous pointer observable, because normal states cannot take a dispersionfree value on an observable with continuous spectrum. The consequence of Theorem 9 primarily used in the sequel is that the typical final states ρ_k are disjoint.

4.3.4 The continuous classical pointer

Neither Theorem 8 nor Theorem 9 can offer a satisfactory description of a continuous pointer. Theorem 8 had to assume the normality of the typical final states, which is a condition that cannot be satisfied for a continuous pointer observable on a separable Hilbert space. Theorem 9 led to non-separable representations. Furthermore, the extended states $\overline{\rho_k}$ constructed there were normal. This is fine in the case of a discrete pointer, but not for a continuous pointer. (The points in the continuous phase space M of a classical system do not correspond to normal states on $L^\infty(M)$.) So a representation can only be appropriate for the description of an experiment with continuous pointer if the states ρ_k can be extended to states which are not normal on the centre of the representation.

Now we will try to construct, under some additional assumptions,

a measurement representation which is appropriate for the description of classical pointer observables with continuous spectrum. The starting point is the abstract C^* -algebra $\mathcal{A}_S \otimes \mathcal{A}_M$ and the set $\{\rho_k\}_{k \in K}$ of pure typical final states. From Theorem 9 we know that the ρ_k are mutually disjoint if the described experiment is inaccurate. Assume also that the Hilbert spaces \mathcal{H}_{ρ_k} of the GNS-representations π_{ρ_k} are all isomorphic to a separable Hilbert space \mathcal{H}_ρ .

In the discrete case the direct sum representation

$$\pi := \bigoplus_{k \in K} \pi_{\rho_k}$$

can be taken as a measurement representation. It has the functions on the discrete set K as the centre. $\pi(\mathcal{A}_S \otimes \mathcal{A}_M)''$ is isomorphic to the tensor-product of this centre with a factor π_{ρ_k} .

In the case of a continuous pointer we will expect something similar, except that the commutative part of the representation will correspond to functions on a continuous configuration space K . In this case the states ρ_k will not be extendable to normal states on the representation. The singular character of the extensions is, however, of a harmless kind. Only their restrictions to the commutative part are singular, on the factor they are normal. (It would be much worse to have a singular state on a type II or III factor, since these states do not fulfil the Jauch-Piron condition.²⁵)

For a continuous pointer the index set K of typical final states contains continuously many elements; assume it to be the real line \mathbb{R} . Taking μ to be a probability measure on the real line, for example $d\mu(k) = \pi^{-1/2} \exp(-k^2) dk$, define $\omega := \int_{\mathbb{R}} d\mu(k) \rho_k$. Then

$$\pi_\omega = \int_{\mathbb{R}}^\oplus d\mu(k) \pi_{\rho_k}$$

²⁵See Amann (1987b).

is the direct integral of the mutually disjoint factor representations π_{ρ_k} .

The main problem now is that, in contradistinction to the discrete case, the diagonal operators (that is, those of the form $\int_{\mathbb{R}}^{\oplus} d\mu(k)f(k)\mathbf{1}_k$) are not necessarily in the weak closure of the direct integral representation π_{ω} . The direct integral of disjoint representations π_{ω} may have trivial centre.²⁶ But under some additional assumptions (which are sufficient, not necessary) the centre of the direct integral representation π_{ω} can be shown to be non-trivial.

Firstly, this is the case if $\pi_{\omega}(\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}})''$ is of type I or II_1 : *If the direct integral over irreducible (or type II_1) and mutually disjoint representations of a C^* -algebra is again of type I (II_1), then the centre of the direct integral representation is non-trivial.*²⁷ (Conversely, the components in a direct integral decomposition of a von Neumann algebra of type A which contains the diagonal operators, are again of type A ($A = \text{I}, \text{II}_1$; in fact this converse holds for $A = \text{II}_{\infty}, \text{III}$) as well.)

For the representation π_{ω} constructed above one can give a simple argument to show that

$\mathcal{Z}(\pi_{\omega}(\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}})'') \cong L^{\infty}(\mathbb{R})$ if $\pi_{\omega}(\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}})''$ is of type I. If $\pi_{\omega}(\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}})''$ is of type I then it is quasi-equivalent to a multiplicity-free representation.²⁸

As the π_{ρ_k} in the integral π_{ω} are mutually disjoint and irreducible, π_{ω} is itself multiplicity-free. Therefore²⁹ the commutant $\pi_{\omega}(\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}})'$ is abelian and the composition of π_{ω} into irreducible representations is unique.³⁰ Since $\pi_{\omega}(\mathcal{A}_{\mathcal{S}} \otimes \mathcal{A}_{\mathcal{M}})'$ is abelian, each factor in the central decomposition of π_{ω} is irreducible. Making this decomposition in our special case of an integral representation over the real line, one sees that

²⁶See e.g. Dixmier (1964b) and Mackey (1978, chaps 7,8).

²⁷See Guichardet (1964).

²⁸See Dixmier (1964a), 5.4.6.

²⁹See Dixmier (1964a), 5.4.4.

³⁰See Bratteli and Robinson (1979), Theorem 4.2.3.

$\pi_\omega(\mathcal{A}_S \otimes \mathcal{A}_M)' \cong L^\infty(\mathbb{R})$. Since the commutant is abelian it equals the centre. Thus $\mathcal{Z}(\pi_\omega(\mathcal{A}_S \otimes \mathcal{A}_M)'') \cong L^\infty(\mathbb{R})$.

If the direct integral representation $\pi_\omega(\mathcal{A}_S \otimes \mathcal{A}_M)''$ is not of type I, then it may be badly behaved. In case $\pi_\omega(\mathcal{A}_S \otimes \mathcal{A}_M)''$ is of types II or III the above argument that the centre is not trivial breaks down. The commutant of such representations cannot be abelian since the commutant is always of the same type as the representation and abelian von Neumann algebras are of type I.

Under the assumption that $\pi_\omega(\mathcal{A}_S \otimes \mathcal{A}_M)''$ is of type I, this representation is indeed appropriate for the description of measurements with continuous pointer. The extensions of the states ρ_k are singular on the centre $L^\infty(\mathbb{R})$. Since $L^2(\mathbb{R}, \mu)$ and all $\mathcal{H}_{\rho_k} = \mathcal{H}_\rho$ are separable, so is $\mathcal{H}_\omega = \int_{\mathbb{R}}^{\oplus} d\mu(k) \mathcal{H}_{\rho_k}$. Furthermore, if the representations π_{ρ_k} are separable, then the direct integral representation π_ω is separable. But the assumption that the direct integral representation is of type I is very strong. In many applications to thermodynamics or field theories one deals with factors of type III.

Secondly, by integration over disjoint factors one can obtain a non-trivial centre if the system is asymptotically abelian. More precisely, Amann (1987a) showed that if the system is *asymptotically abelian* in the norm topology (i.e.,

$\lim_{|\mathbf{q}| \rightarrow \infty} \|[\alpha_{|\mathbf{q}|}(A), B]\| = 0$, where $\alpha : \mathbb{R}^3 \rightarrow \text{Aut}(\mathcal{A}_S \otimes \mathcal{A}_M)$ is the representation of the group \mathbb{R}^3 of space translations), and if there is a translation invariant state ϕ , and if the action of the symmetry group G commutes with the space translations, then the centre of the direct integral representation obtained from averaging the factor π_ϕ over G/H is isomorphic to the abelian von Neumann algebra $L^\infty(G/H)$. (Here H is the subgroup of G leaving ϕ invariant.) We have $\pi_\omega(\mathcal{A}_S \otimes \mathcal{A}_M)'' \cong$

$$L^\infty(G/H) \otimes \pi_\phi(\mathcal{A}_S \otimes \mathcal{A}_M)''.$$

4.4 The time-evolution of the classical pointer

4.4.1 Infinite time limit for automorphic measurements with classical pointer

Hamiltonian and automorphic time evolutions. In traditional Hilbert space quantum mechanics, the time-evolution is generated by the self-adjoint Hamiltonian operator. In the Heisenberg-picture the time evolution of an observable A is given by

$$U_t A U_t^* =: \alpha_t(A),$$

where $U_t = \exp(itH)$.

The maps $\alpha_t : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ are *-automorphisms of the von Neumann algebra $\mathcal{B}(\mathcal{H})$ because they have the following properties:

- (i) $\alpha_t(\lambda A + \mu B) = \lambda \alpha_t(A) + \mu \alpha_t(B)$,
- (ii) $\alpha_t(AB) = \alpha_t(A)\alpha_t(B)$,
- (iii) $\alpha_t(A^*) = (\alpha_t(A))^*$
- (iv) $\alpha_{t+s} = \alpha_t \circ \alpha_s$, and
- (v) $t \mapsto \alpha_t(A)$ is, for all $A \in \mathcal{B}(\mathcal{H})$, continuous in the σ -weak topology on $\mathcal{B}(\mathcal{H})$.

Restricted to the C^* -algebra $\mathcal{K}(\mathcal{H})$ of compact operators on \mathcal{H} , the maps $\alpha_t : \mathcal{K}(\mathcal{H}) \rightarrow \mathcal{K}(\mathcal{H})$ form a even a C^* -automorphism group because they fulfil requirements (i)-(iv) above and additionally

- (v') $t \mapsto \alpha_t(A)$ is, for all $A \in \mathcal{K}(\mathcal{H})$, continuous in the norm topology on $\mathcal{K}(\mathcal{H})$.

On $\mathcal{B}(\mathcal{H})$ norm continuity holds only if the Hamiltonian is bounded.

Generalising the concept of Hamiltonian time-evolution, *automorphic time-evolution* may be defined to be pointwise norm continuous one

parameter $*$ -automorphism-group of a C^* -algebra. This definition provides for a real generalisation of Hamiltonian time-evolution: Firstly, it may happen that a given one parameter group of automorphisms does not allow a strongly continuous unitary implementation. Secondly, even if the automorphisms are unitarily implementable in some representations, the Hamiltonian may depend on the representation.

The classical pointer and automorphic time evolution. In spite of the fact that automorphic time evolutions of a C^* -algebra are more general than Hamiltonian ones, automorphic time-evolutions have difficulties in describing the evolution of a classical pointer observable. This was pointed out by Hepp (1972, Lemma 2). He showed that if two states ρ_1, ρ_2 on a C^* -algebra are in the same superselection sector, and if α_t is an automorphism, then $\rho_1 \circ \alpha_t$ and $\rho_2 \circ \alpha_t$ will again be in the same sector. Applied to measurements, this has the following consequence: The initial states of the joint system are all in the superselection sector corresponding to pointer “ready”; if the time evolution is described by a one parameter automorphism group of the C^* -algebra, then for every finite time the time evolved states will still be in the same sector. Therefore every classical observable, in particular the pointer observable, will have the same value in the time evolved states. The value of a classical pointer observable does not change in finite time.

There are two ways to circumvent Hepp’s no-go theorem: Firstly one can take a time evolution which is not an automorphism of the C^* -algebra. I will deal with this approach in section 4.4.2. Secondly, and this way of escape was emphasised by Hepp, one can take the limit $t \rightarrow \infty$: Even if ρ_1, ρ_2 are in the same sector and the time evolution is described by a norm-continuous automorphism group α_t , $\rho_1 \circ \alpha_t$ and

$\rho_2 \circ \alpha_t$ can converge for $t \rightarrow \infty$ to disjoint states.

But for two reasons there are problems with this second way of escape. Firstly, if one insists that the experimentally relevant pointer observable really is classical, then it can change its value just in the infinite time limit. Real measurements, however, only take finite time. Secondly, convergence to the disjoint final states $\rho_1^\infty, \rho_2^\infty$ only takes place in the weak*- topology on the state space. This means that for any *fixed* set of observables A_1, \dots, A_n we have

$$\lim_{t \rightarrow \infty} (\rho_1(\alpha_t(A_i)) - \rho_1^\infty(A_i)) = 0,$$

and similarly for ρ_2 . But for any given time and any given ϵ it is possible to find an observable $B(t)$ for which the interference terms between $\rho_1 \circ \alpha_t$ and $\rho_2 \circ \alpha_t$ are bigger than ϵ . (The operator $B(t)$ can simply be chosen in such a way that the Heisenberg time evolution is undone.) In the norm topology no limit at all is approached and the measurement will never even approximately be completed. It was for this reason that Bell (1975) doubted the physical significance of Hepp's result.

Peres (1980, 1986) responded to Bell's criticism by pointing out that the choice of $B(t)$ requires an inversion of the equations of motion. For large systems this is practically impossible. Landsman (1995) argued that it is not really necessary to calculate the inverse time evolution because the initial state, which is the result of the inverse time evolution, is already known. Instead, Landsman argues, the observed time evolution is irreversible due to the fact that we monitor only some selected degrees of freedom. This irreversibility makes it impossible to apply Bell's argument.

Other models which use the infinite time limit. Let me now discuss the

first argument against taking the infinite time limit: If one insists that the experimentally relevant pointer observable really is classical, then an automorphic time evolution can change its value just in the infinite time limit. Real measurements, however, only take finite time.

This is of course an argument which can be raised against all measurement models using the infinite time limit. For example, Whitten-Wolfe and Emch (1976), Hannabus (1984), Lockhart and Misra (1986), Frigerio (1974) presented models in which the correlations between apparatus and system vanish in the infinite time limit. Joos and Zeh (1985), as well as Zurek (1982) include the environment into their considerations and arrive in the infinite time limit at a state whose partial trace over the environment is the wanted final state s_2 . At any finite time, partial tracing leads to a state with non-vanishing off-diagonal terms in the pointer basis. The rôle of the environment is to accelerate the convergence of the off-diagonal terms to zero. Furthermore, if the environment is infinite, Poincaré recurrence can be excluded.

But unlike Hepp's model, these above models have trivial superselection structure, although they often take as environment an infinite system. The arguments against the infinite time limit are less harmful to these models because the pointer observable takes different values already at finite time. It is just the interference terms which take an infinite time to vanish. But such models fare worse than Hepp's model on a different account: they have difficulties in justifying the ignorance interpretation.

Irreversible time evolutions. A non-linear irreversible time evolution of the apparatus and object system is often obtained as reduced dynamics: the feedback effects of the environment on the system are described

by a *non-linear* equation. In an important paper Hannabus (1984) showed the following. Take any contractive semi-group describing the irreversible time-evolution of apparatus and object system. (The transition $(\sum c_n |o_n\rangle) \otimes |I\rangle \rightarrow s_2$ can be achieved by such a time evolution.) Then it is possible to find an environment and an interaction such that the dynamical semigroup is approximated arbitrarily well by reducing the unitary dynamics of the bigger system.

The measurement models making reference to an unobserved environment can, even if they fully succeed, explain only the emergence of observer-dependent facts in a quantum world. They are however not able to satisfy the less modest realist demanding the explanation of observer-independent facts.

4.4.2 Non-automorphic time evolution

Up to now non-automorphic time evolutions were discussed for reduced dynamics. In this final section I would like to discuss the possibility of non-automorphic time evolutions for *closed* systems. I will argue that for systems with long-range interactions non-automorphic time evolutions are very natural. Finally I will give an example of a time evolution which has some of the features displayed by long range dynamics and leads in finite time to disjoint final states.

Hepp's Big Bang time evolution. Hepp (1972) proposed the following “big bang model” for measurements. Other time evolutions achieving the same have been proposed by Bub (1988, 1989) and Wan (1980). Consider an infinite spin chain. Assume that the first spin (at site $n = 0$) serves as the observed system, and all the other spins (at sites $n = 1, 2, 3, \dots$) as measurement apparatus. So $\mathcal{H}_S = \mathbf{C}^2$ and $\mathcal{H}_M =$

$\bigotimes_{n=1}^{\infty} \mathbf{C}_n^2$. Let the unitary operator $U(t) : \mathcal{H}_M \rightarrow \mathcal{H}_M$ be defined by linear extension of

$$U(t) \bigotimes_{n=1}^{\infty} \psi_n := \bigotimes_{n=1}^{\infty} (\exp(i\sigma^1 t) \psi_n).$$

Assume that the measured quantity is the spin σ^3 in 3-direction, and denote by P_+ and P_- the projectors onto the eigenstates ψ^+, ψ^- of σ^3 . Then Hepp defines a unitary operator $W(t) : \mathcal{H}_S \otimes \mathcal{H}_M \rightarrow \mathcal{H}_S \otimes \mathcal{H}_M$ by

$$W(t) := P_+ \otimes \mathbf{1} + P_- \otimes U(t).$$

Take as initial states $\psi_0^+ \otimes \phi^+$ (resp. $\psi_0^- \otimes \phi^+$), where $\phi^+ = \bigotimes_{n=1}^{\infty} \psi_n^+$. Then

$$W(t) \psi_0^+ \otimes \phi^+ = \psi_0^+ \otimes \phi^+$$

and

$$W(t) \psi_0^- \otimes \phi^+ = \psi_0^- \otimes \bigotimes_{n=1}^{\infty} (\exp(i\sigma^1 t) \psi_n)$$

are disjoint for arbitrarily small $t > 0$. Thus $W(t)$ transforms the states $\psi_0^+ \otimes \phi^+$ and $\psi_0^- \otimes \phi^+$, which are in the same superselection sector, in arbitrarily short time into states which are not in the same superselection sector.

The time evolution $W(t)$ used by Hepp is the infinite system analogue of the first measurement time evolution defined by von Neumann (1932, p. 235). Von Neumann chose the pointer basis $\{p_n\}$, $n = 0, \pm 1, \pm 2, \dots$, for \mathcal{H}_M and a basis of eigenvectors for the observed quantity $\{o_m\}$, $m = 0, \pm 1, \pm 2, \dots$, for \mathcal{H}_S . He then defined a unitary map $\mathcal{H}_S \otimes \mathcal{H}_M \rightarrow \mathcal{H}_S \otimes \mathcal{H}_M$ by linear extension of

$$o_m \otimes p_n \mapsto o_m \otimes p_{m+n}. \quad (4.1)$$

This unitary map achieves exactly the coupling

$$\sum c_m o_m \otimes p_0 \rightarrow \sum c_m o_m \otimes p_m.$$

If there are just two possible values of the measured quantity ($m = 1, 2$) then one can replace (4.1) with

$$o_m \otimes p_n \mapsto o_m \otimes p_{(n+m+1) \bmod 2}.$$

This extends precisely to Hepp's $W(t)$ for $t = \pi$.

*Hepp's Big Bang time evolution does not lead to an automorphism of \mathcal{A}^{**} .* $W(t)$ maps equivalent states in finite time into disjoint states, and therefore cannot, by Hepp's (1972) Lemma 2, be an automorphism of the C^* -algebra \mathcal{A} of quasilocal operators on $\mathcal{H}_S \otimes \mathcal{H}_M$. But $W(t)$ is not even an automorphism of the W^* -algebra \mathcal{A}^{**} (which is isomorphic to the weak closure of the universal representation of \mathcal{A}). To see this one argues as follows.

Choose a product state $\phi_1 \in \mathcal{H}_S \otimes \mathcal{H}_M$ and a $t > 0$ such that $(P_- \otimes U(t))\phi_1 =: \phi_2$ and $(P_- \otimes U(t)^*)\phi_1 =: \phi_0$ are disjoint from ϕ_1 and from each other. Since ϕ_1 is a product state it is pure. The same is true for ϕ_0 and ϕ_2 . The operators $P_\pm \otimes \mathbf{1}$, $P_\pm \otimes U(t)$, and $P_\pm \otimes U(t)^*$ map ϕ_1 and ϕ_2 into pure states. Pure states can only be in the same superselection sector (then we write \sim) or disjoint (then we write $\not\sim$). We then have the following relations $(P_\pm \otimes \mathbf{1})\phi_i \sim \phi_i$, $(P_\pm \otimes U(t))\phi_1 \sim \phi_2$, $(P_\pm \otimes U(t)^*)\phi_2 \sim \phi_1$. If $\phi \sim \phi'$ then there exists an $A \in \mathcal{A}$ such that $\langle \phi | A | \phi' \rangle \neq 0$. If, however, $\phi \not\sim \phi'$ then $\langle \phi | A | \phi' \rangle = 0$ for all $A \in \mathcal{A}$. Putting all this together we obtain

$$\begin{aligned} \langle \phi_1 | W(t) A W(t)^* | \phi_2 \rangle &= \langle \phi_1 | (P_+ \otimes \mathbf{1} + P_- \otimes U(t)) A (P_+ \otimes \mathbf{1} + P_- \otimes U(t))^* | \phi_2 \rangle \\ &= \langle \phi_1 | (P_+ \otimes \mathbf{1}) A (P_+ \otimes \mathbf{1})^* | \phi_2 \rangle \end{aligned}$$

$$\begin{aligned}
& + \langle \phi_1 | (P_- \otimes U(t)) A (P_- \otimes U(t)^*) | \phi_2 \rangle \\
& + \langle \phi_1 | (P_- \otimes U(t)) A (P_+ \otimes \mathbf{1})^* | \phi_2 \rangle \\
& + \langle \phi_1 | (P_+ \otimes \mathbf{1}) A (P_- \otimes U(t)^*) | \phi_2 \rangle.
\end{aligned}$$

Since all four summands on the right hand side, except the last, vanish, there is an $A \in \mathcal{A}$ such that

$$\langle \phi_1 | W(t) A W(t)^* | \phi_2 \rangle \neq 0.$$

But for all operators $B \in \mathcal{A}^{**}$ we have $\langle \phi_1 | B | \phi_2 \rangle = 0$. Therefore $W(t) A W(t)^*$ cannot be in \mathcal{A}^{**} . Thus the unitary time evolution $W(t)$ does not lead to an automorphism of \mathcal{A}^{**} .

Time evolution in the presence of long range interactions. Non-automorphic time evolutions arise in some lattice models with long-range interactions.³¹ These models do not admit a time evolution described by a one-parameter group of automorphisms of the quasilocal algebra. Landsman (1991) suggested that the kind of dynamics arising in the presence of long range interactions might be used to describe measurement evolutions.

The typical picture of the dynamics in the presence of long range interactions, as described in Morchio and Strocchi (1987), looks as follows. One starts with a series of Hamiltonians H^Λ . The parameter Λ refers usually to a finite volume cut-off or to an infrared cut-off. Finally one has to take the limit $\Lambda \rightarrow \infty$. Then one defines a Λ -dependent time evolution by automorphisms α_t^Λ of \mathcal{A}

$$\alpha_t^\Lambda(x) := \exp(itH^\Lambda)x \exp(-itH^\Lambda), \quad x \in \mathcal{A}.$$

³¹Examples are the mean field models of Hepp and Lieb (1972), Morchio and Strocchi (1985), Bóna (1988), Unnerstall (1990).

If the interactions is of sufficient long range³² the limit $\lim_{\Lambda \rightarrow \infty} \alpha_t^\Lambda(x)$, $x \in \mathcal{A}$ does not exist in norm. Rather, there is a class F of physically relevant states such that

$$\lim_{\Lambda \rightarrow \infty} \rho(\alpha_t^\Lambda(x)) =: \rho_t(x)$$

exists for all $x \in \mathcal{A}$ only if $\rho \in F$. Passing to the Schrödinger picture this defines a time evolution which is not an automorphism group of \mathcal{A} . Since it is not possible to write $\rho_t = \rho \circ \alpha_t$ for some automorphism α_t , Hepp's (1972) Lemma 2 does not apply. Thus it is not excluded that two states ρ, ω on \mathcal{A} which are in the same superselection sector evolve in finite time t into disjoint states ρ_t, ω_t .

But the evolutions studied by Morchio and Strocchi (1987) are not general enough for our purposes. The reason for this is that their interactions lead to an automorphic time evolution of some von Neumann algebra \mathcal{M} , $\mathcal{A} \subset \mathcal{M} \subset \mathcal{A}^{**}$. But then Hepp's (1972) Lemma 2 can be applied to states on \mathcal{M} .

A proposal for a non-automorphic time evolution. Finally I would like to give an example of a measurement model which displays some of the structure typical for long range interactions, and which leads to disjoint pointer states in finite time. This time evolution is a limit of automorphic evolutions, but is itself not an automorphism of a \mathcal{A} , \mathcal{A}^{**} , or \mathcal{M} . The basic idea is that a long range interaction (in the sense of Morchio and Strocchi (1987) or Bóna (1988)) of the apparatus is coupled to the measured observable.

Consider the infinite Weyl system described in chapter 3. Take as observed system the 0-th particle and as apparatus the other particles

³²Sufficiently long range means in the example of spin systems with two-body interaction that the potential must decay slower than r^{-3} . See Haag, Hugenholtz, and Winnink (1967), or Robinson (1968).

1, 2, \dots. Let $\mathcal{H} := L^2(\mathbb{R}^3)$ and let \mathcal{H}_n , $n = 0, 1, 2, \dots$, be copies of \mathcal{H} . So $\mathcal{H}_S = \mathcal{B}(\mathcal{H}_0)$ and $\mathcal{H}_M = \bigotimes_{n=1}^{\infty} \mathcal{B}(\mathcal{H}_n)$. As measured observable take the position of particle 0, and as pointer observable take the position $\mathbf{1} \otimes Q^{\text{II}}$ of the centre of mass of the particles 1, 2, \dots. In chapter 3 it was shown that $\mathbf{1} \otimes Q^{\text{II}}$ is classical. Assume that masses are big enough for the kinetic energy terms to be negligible.

As finite volume Hamiltonian define

$$H_N := q \sum_{n=1}^N \frac{\partial}{\partial r_n}.$$

(So N plays the rôle of the cut-off parameter Λ above.) A solution Ψ of the Schrödinger equation

$$\frac{\partial}{\partial t} \Psi_t(q, r_1, r_2, \dots) = -H_N \Psi_t(q, r_1, r_2, \dots)$$

must be of the form

$$\Psi_t(q, r_1, r_2, \dots) = f(q, r_1 + tq, \dots, r_N + tq, r_{N+1}, \dots)$$

where $f(q, r_1, r_2, \dots) = \Psi_0(q, r_1, r_2, \dots)$. If we choose as initial state the product state

$$\Psi_0(q, r_1, r_2, \dots) := \phi_0(q) \psi_1(r_1) \psi_2(r_2) \dots$$

then the state at time t will be the product state

$$\phi(q) \psi_1(r_1 + tq) \dots \psi_N(r_N + tq) \psi_{N+1}(r_{N+1}) \dots = \exp(itH_N) \Psi_0.$$

The map

$$x \mapsto \exp(itH_N)x \exp(-itH_N) =: \alpha_t^N(x), \quad x \in \mathcal{A}$$

defines an automorphism of \mathcal{A} which can be extended to an automorphism of \mathcal{A}^{**} .

Let us now take the limit $N \rightarrow \infty$ of the finite volume dynamics. For all states Ψ_0 for which the limit $\lim_{N \rightarrow \infty} \langle \Psi_0 | \alpha_t^N(x) | \Psi_0 \rangle$ exists for all $x \in \mathcal{A}$ one has

$$\lim_{N \rightarrow \infty} \exp(itH_N) \Psi_0(q, r_1, r_2, \dots) = \Psi_0(q, r_1 + tq, r_2 + tq, \dots) =: \Psi_t(q, r_1, r_2, \dots).$$

The pointer observable $\mathbf{1} \otimes Q^\Pi$ has different dispersion free expectation values in the states Ψ_0 and Ψ_t for any $t > 0$. These states are disjoint after an arbitrary short time.

This model is very crude for several reasons. Firstly, it is not clear whether the above intuitive picture persists if the kinetic energy terms are included in the Hamiltonian. Secondly, it would have to be investigated for which states Ψ and for which x the limit $\lim_{N \rightarrow \infty} \langle \Psi | \alpha_t^N(x) | \Psi \rangle$ exists. Thirdly, the model is essentially non-relativistic.³³ Finally, it does not seem probable that any realistic measurement interaction is of the simple form described in the model. (This is perhaps the least disturbing feature since the Hamiltonian H_N is exactly the N -particle version of von Neumann's (1932, p. 236) measurement Hamiltonian.)

Observe that the proposed interaction cannot be an automorphism group of \mathcal{A} , nor of \mathcal{A}^{**} , nor of any von Neumann algebra $\mathcal{A} \subset \mathcal{M} \subset \mathcal{A}^{**}$. In this respect the proposed interaction differs from the long range interactions studied by Morchio and Strocchi (1987), or the mean field interactions of Bóna (1988). The decisive point of the interaction given here is that it couples a microscopic observable and a macroscopic one.

³³It would perhaps be possible to have long range interactions in a relativistic field theory if one chooses a non-relativistic gauge. Because of gauge invariance the theory can still be relativistic.

References

Accardi L. (1988), “Foundations of Quantum Mechanics: A quantum probabilistic approach”, in G. Tarozzi And A. van der Merwe (eds.): *The nature of quantum paradoxes*, Dordrecht: Kluwer

Accardi L. and Fedulo A. (1982), *Lett. Nuovo Cim. A* **34**, 161

Accardi L., Frigiero A. and Gorini V.(eds.), *Quantum Probability and Applications to Quantum Theoary of Irreversible Processes*, Lecture Notes in Mathematics 1055, Berlin: Springer

Agarwal G. S., E. Wolf (1970a), *Phys. Rev. D* **2**, 2161

Agarwal G. S., E. Wolf (1970b), *Phys. Rev. D* **2**, 2182

Albert D. Z. (1983), *Phys. Lett. A***98**, 246

Albert D. Z. (1987), *Philosophy of Science* **57**, 577

Ali S.T. (1979), *J. Math. Phys.* **20**, 1385

Ali S. T. (1984), *Can. Math. Bull.* **27**, 390

Ali S. T. (1985), *Riv. Nuovo Cim.* **8** (11), 1-128

Ali S. T. , J.-P. Antoine, J.-P. Gazeau (1991a), Ann. Inst. H. Poincaré **55**, 829

Ali S. T. , J.-P. Antoine, J.-P. Gazeau (1991b), Ann. Inst. H. Poincaré **55**, 857

Ali S. T., Doebner H.-D. (1990), Phys. Rev. A **41**, 1199

Ali S. T., Emch G. G. (1986), J. Math. Phys. **27**, 2936

Ali S.T., Prugovečki E. (1977), J. Math. Phys. **18** (2), 219

Amann A. (1978), *Über eine Verallgemeinerung des quantenmechanischen Elementaritätsbegriffs*, Diplomarbeit ETH Zürich,

Amann A. (1986a), Fortschr. Phys. **34**, 167-215

Amann A. (1986b), J. Math. Phys. **27**, 2282

Amann A. (1987a), Helv. Phys. Acta **60**, 384

Amann A. (1987b), J.Math.Phys. **28**, 2384

Amann A., Müller-Herold U. (1986), Helv. Phys. Acta **59**, 1311

Araki H. (1960), J. Math. Phys. **1**, 492

Araki H. (1980), *Prog. Theor. Phys.* **64**, 719

Araki H. (1986), p. 348 in M. Namiki *et al.* (eds.): *Proceedings of the 2nd International Symposium on the Foundations of Quantum Mechanics*, Tokyo: Physical Society of Japan,

Araki H. and Yanase M. (1960), Phys.Rev. **120**, 622

- Arveson W. B. (1969), *Acta Math* **123**, 142
- Ballentine L. (1970), *Rev. Mod. Phys.* **42**, 358
- Bargmann V. (1954), *Ann. Math.* **59**, 1
- Barut A. O, Rączka (1986), *Theory of Group Representations and Applications*, Singapore: World Scientific
- Beltrametti E., Casinelli G. (1981), *The Logic of Quantum Mechanics*, Reading, Mass.: Addison-Wesley
- Berberian S. K. (1974), *Lectures in Functional Analysis and Operator Theory*, New York: Springer
- Birkhoff G., J. von Neumann (1936), *Ann. Math.* **37**, 823-843
- Blattner R. J. (1961), *Amer. J. Math.* **83**, 79
- Blattner R. J. (1962), *Bull. Amer. Math. Soc.* **68**, 585
- Bohm D. (1951), *Quantum Theory*, Englewood Cliffs: Prentice Hall
- Bóna P. (1973), *Acta Phys. Slov.* **27**, 101-112
- Bóna P. (1988), *J. Math. Phys.* **29**, 2223
- Bratteli O., Robinson D. W. (1979, 1981), *Operator Algebras and Quantum Statistical Mechanics*, Vols. 1 and 2, New York: Springer, Second Edition of Vol. 2 (1987)
- Breuer T. (1992), *Phys. Lett. A* **168**, 233
- Breuer T. (1993), *Int. J. Theor. Phys.* **32**, 2253

- Breuer T. (1994a), Philosophy of Science, forthcoming
- Breuer T. (1994b), in V. B. Belavkin, O. Hirota, R. L. Hudson (eds.): *Quantum Communication and Measurement*, New York: Plenum
- Breuer T. (1994c), Phys. Rev. A **50**, 4565
- Breuer T., Amann A., Landsman N. P. (1993a), J. Math. Phys. **34**, 5441
- Breuer T., Amann A., Landsman N. P. (1993b), p. 118-126 in P. Busch, P. Lahti, P. Mittelstaedt (eds.): *Symposium on the Foundations of Modern Physics 1993*, Singapore: World Scientific
- Bub J. (1988), pp. 134-144 in *PSA 1988*, Volume 2, East Lansing: Philosophy of Science Association
- Bub J. (1989), in M. Kafatos (ed.): *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, Boston: Kluwer
- Bub J. (1992), *Found. Phys.* **22**, 737
- Bub J. (1993), p. 125-136 in P. Busch, P. Lahti, P. Mittelstaedt (eds.): *Symposium on the Foundations of Modern Physics 1993*, Singapore: World Scientific
- Buchholz D., Doplicher S., Longo R., and Roberts J.E. (1992), Rev. Math. Phys. (Special Issue December 1992, Dedicated to R. Haag)
- Busch P, Lahti P., Annalen der Physik **47**, 369
- Busch P. (1982), Int. J. Theor. Phys. **24**, 63

Busch P., Phys. Rev. D **33**, 2253

Busch P. (1987), Found. Phys. **17**, 905

Busch P., Lahti P. J. (1986), Phys. Lett. **A115**, 259

Busch P., Lahti P., Mittelstaedt P. (1992), *Quantum Theory of Measurement*, Lecture Notes in Physics **m2**, Heidelberg: Springer

Castrigiano D. P. L., R. W. Henrichs (1980), Lett. Math. Phys. **4**, 169

Cattaneo U. (1979), Comment. Math. Helv. **54**, 629

Chen Y.H., Wilczek F., Witten E., Halperin P. I. (1989), Int. J. Mod. Phys. B **3**, 1001

Connes A., Takesaki M. (1977), Tôhoku Math. J. **29**, 473

Conway J. B. (1985), *A Course in Functional Analysis*, New York: Springer

Dalla Chiara M.L. (1977), J. Phil. Logic **6**, 331

Davies E. B., Lewis J. T. (1970), Comm Math. Phys. **17**, 239

Dieks D. (1989), Phys. Lett. **A142**, 439

Dixmier J. (1950), Ann. of Math. **51**, 387-408

Dixmier J. (1957), *Les algèbres d'opérateurs dans l'espace Hilbertien (Algèbres de von Neumann)*, Paris: Gauthier-Villars

Dixmier J. (1964a), *Les C^* -algèbres et leurs représentations*, Cahiers Scientifiques, **29**, Paris: Gauthier-Villars

- Dixmier J. (1964b), C. R. Acad. Sc. Paris **258**, 4184
- Duffo M., Moore C. C. (1976), J. Funct. Anal. **21**, 209-243
- Earman J. (1986), *A Primer on Determinism*, Dordrecht: D. Reidel
- Earman J., Shimony A. (1968), Nuovo Cimento **B 54**, 332
- Einstein A., Podolsky B., Rosen N. (1935), Phys. Rev. **47**, 777
- Emch G. G. (1990), pp. 356-364 in V. V. Dodonov, V. I. Man'ko (eds.): *Group Theoretical Methods in Physics (Moscow 1990)*, Berlin: Springer, 1991
- Espagnat B. de (1966), Nuovo Cimento Suppl. **4**, 828
- Everett H. III (1957), " 'Relative State' Formulation of Quantum Mechanics", *Reviews in Modern Physics* **29**, 454-462
- Fine A. (1972), Phys.Rev. D **2**, 2783
- Finetti, de B.(1972), *Theory of Probability*, New York: Wiley
- Finkelstein D. (1988), "Finite Physics", in R.Herken (ed.): *The Universal Turing Machine. A half-Century Survey*, Oxford: University Press
- Frigiero A. (1974), Ann.Inst.H.Poincaré A **3**, 259
- G. W. Mackey (1978), *Unitary Group Representations in Physics, Probability, and Number Theory*, Providence: Addison-Wesley
- Gödel K. (1931), "Über formal unentscheidbare Sätze der Principia

Mathematica und verwandter Systeme”, Monatshefte für Mathematik und Physik **38**, 173

Gel'fand I.M. and Naimark M.A. (1943), Mat.Sb. **12**, 197

Gell-Mann M., Hartle J. (1990), “Quantum Mechanics in the Light of Quantum Cosmology”, in W. H. Zurek (ed.): *Complexity, Entropy, and the Physics of Information*, SFI Studies in the Science of Complexity, vol. VIII, Addison-Wesley

Glimm J. (1962), *Pacific J. Math.* **12**, 855

Glimm J., Kadison R.V. (1960), *Pacific J. Math.* **10**, 547

Guichardet A. (1964), C. R. Acad. Sc. Paris, **258**, 768

Haag R., D. Kastler (1964), *J. Math. Phys* **5**, 848

Haag R., Hugenholtz N. M., Winnink M. (1967), *Comm. Math, Phys.* **5**, 215

Haag R., V. Kadison, D. Kastler (1970), *Comm. Math. Phys.* **16**, 81

Haagerup U. (1979), *J. Functional Analysis* **32**, 175-206

Hannabus K.C. (1984), *Helv. Phys. Acta* **57**, 610-620

Healey R. (1989), *The Philosophy of Quantum Mechanics: An Interactive Interpretation*, Cambridge: University Press

Hegerfeldt G. C. (1974), *Phys. Rev.* D**10**, 3320

Hegerfeldt G. C., Ruijsenaars S. N. M. (1980), *Phys. Rev.* D**22**, 377

- Hepp K. (1972), *Helv.Phys.Acta* **45**, 237
- Hepp K., E. H. Lieb (1973), *Helv. Phys. Acta* **46**, 573
- Holevo A. S. (1982), *Probabilistic and Statistical Aspects of Quantum Theory*, Amsterdam: North Holland
- Holevo A. S. (1985), *Theor. Math. Phys.* **65**, 1250
- Homma K. (1988), *Yokohama Math. J.* **36**, 105
- Hume D. (1758), *An Enquiry Concerning Human Understanding*
- Isham C. J. (1983), “Topological and global aspects of quantum theory”, in B. S. Dewitt, R. Stora (eds.): *Relativity, Groups, and Topology 2*. Les Houches Summer School 1983, Amsterdam: North Holland, 1984
- Jauch J. M. (1968), *Foundations of Quantum Mechanics*, Reading, Mass.: Addison-Wesley
- Joos E. and Zeh H.D. (1985), *Z.Phys* **B59**, 223
- Kümmerer B. (1990), *Construction and Structure of Markov Dilations on W^* -algebras*, Habilitationsschrift, Universität Tübingen
- Kadison R.V. (1965), *Topology* **3**, 177
- Kadison R.V., Ringrose (1986), *Fundamentals of Operator Algebras*, Vols. I, II, New York: Academic Press
- Kirillov A. A. (1974), *Eléments de la théorie des représentation*, Moscow: Mir

Klauder J. R., B. S. Skagerstam (1984), *Coherent States. Applications in Physics and Mathematical Physics*, Singapore: World Scientific

Kochen S. (1985), p. 151 in P. Lahti, P. Mittelstaedt (eds.): *Symposium on the Foundations of Modern Physics 1985*, Singapore: World Scientific

Komar A. (1962), *Phys.Rev.* **126**, 365

Komar A. (1964), *Phys. Rev.* **133B**, 542-544

Koopman B.O. (1957), pp. 97-102 in *Proceedings of Symposia in Applied Mathematics*, Am. Math. Soc., Providence, Rhode Island

Kudaka S., Matsumoto S., Kakazu K. (1989), *Prog. Theor. Phys.* **82**, 665

Lévy-Leblond J.-M. (1963), *J. Math. Phys.* **4**, 776

Lévy-Leblond J.-M. (1971), in E. M. Loeb (ed.): *Group Theory and its Applications*, Vol. 2, New York: Academic Press

Landsman N. P. (1990a), *Rev. Math. Phys.* **2**, 45

Landsman N. P. (1990b), *Rev. Math. Phys.* **2**, 73

Landsman N.P. (1991), *Int. J. Mod. Phys.* **A6, 30**, 5349

Landsman N. P. (1993), *J. Geom. Phys.* **12**, 93

Landsman N.P. (1995), "Observation and superselection in quantum mechanics", *Stud. Hist. Phil. Mod. Phys.*, *forthcoming*

- Leggett A.J. (1987), in B.J.Hiley and F.D. Peat (eds.): *Quantum Implications*, London: Routledge and Kegan Paul
- Lie S., Engel F. (1888), *Theorie der Transformationsgruppen*, Vol. 1, Leipzig: Teubner
- Lockhart C.N., Misra B. (1986), *Physica* **136A**, 47
- London F. and Bauer E. (1939), *La théorie de l'observation en mécanique quantique*, Paris: Hermann
- Loomis L. H. (1947), *Bull. Amer. Math. Soc.* **53**, 757-760
- Ludwig G. (1967), *Comm. Math. Phys.* **4**, 331
- Ludwig G. (1968), *Comm. Math. Phys.* **9**, 1
- Ludwig G. (1972), *Comm. Math. Phys.* **26**, 78
- Machida S. and Namiki M. (1980), *Prog.Theor.Phys.* **63**, 1457–1473, 1833–1847
- Mackey G. W. (1949), *Duke Math. J.* **16**, 313
- Mackey G.W. (1949), *Proc. Natl. Acad. Sci. USA* **35**, 537
- Mackey G. W. (1952), *Ann. of Math.* **55**, 101
- Mackey G. W. (1955), *Acta Math.* **99**, 265
- Mackey G. W. (1962), *Ill. J. Math.* **6**, 327
- Mackey G. W. (1963), *Foundations of Quantum Mechanics*, Reading, Mass.: Benjamin-Cummings

Mackey G.W. (1978), *Unitary Group Representations in Physics, Probability, and Number Theory*, Reading, Mass.: Benjamin-Cummings

Mittelstaedt P. (1993), "Measurement induced interrelations between quantum theory and its interpretation", in P. Busch, P. Mittelstaedt, P. Lahti: *Proceedings of the Symposium on the Foundations of Modern Physics 1993*, Singapore: World Scientific

Mittelstaedt P., Prieur A., Schieder R. (1987), Unsharp Particle Wave Duality in a Photon Split-Beam Experiment, *Foundations of Physics* **17** (9), 891-903

Morchio G., F. Strocchi (1985), *Comm. Math. Phys.* **99**, 153

Morchio G., F. Strocchi (1987), *J. Math. Phys.* **28**, 622

Naimark M. A. (1943), *Dokl. Akad. Nauk. SSSR* **41**, 359

Nakagami Y., Takesaki M. (1979), *Duality for Crossed Products of von Neumann Algebras*, *Lecture Notes in Mathematics* **731**, Berlin: Springer

Nakazato H., Pascazio S. (1991), *Phys. Lett.* **156A**, 386

Nakazato H., Pascazio S. (1992), *Phys. Rev. A* **45**, 4355

Namiki M., Pascazio S. (1991), *Found. Phys. Lett.* **4**, 203

Neumann H. (1972), *Helv. Phys. Acta* **45**, 811

Neumann J. von (1932), *Mathematische Grundlagen der Quantenmechanik*, Berlin: Springer

- Neumann J. von (1938), *Composit. Math.* **6**, 1
- Orzalesi C. A. (1970), *Rev. Mod. Phys.* **42**, 381
- Ozawa M. (1984), *J.Math.Phys.* **25**, 79
- Pedersen G. K. (1979), *C*-algebras and their Automorphism Groups*, London: Academic Press
- Penrose R. (1989), *The Emperor's New Mind*, Oxford: University Press
- Perelomov A. M. (1972), *Comm. Math. Phys.* **26**, 222
- Peres A. (1980), *Phys. Rev.* **D22**, 879
- Peres A. (1986), *Amer. J. Phys.* **54**, 688
- Peres A. and Zurek W.H. (1982), *Am. J. Phys.* **50** (9), 807
- Piron C. (1976), *Foundations of Quantum Physics*, Reading: Benjamin
- Plymen R. J. (1968), *Nuovo Cimento A* **54**, 862
- Popper K.R. (1950), *Brit. J. Phil. Sci.* **1**, 173
- Primas H. (1980), "Foundations in Theoretical Chemistry", in R.G. Woolley (ed.): *Quantum Dynamics of Molecules—The new experimental challenge for the theorist*, NATO Advanced Studies Series, Vol. 57, New York: Plenum
- Primas H. (1981), *Chemistry, Quantum Mechanics, and Reductionism*, Berlin: Springer, Second Edition 1983
- Primas H. (1990a), "The measurement process in the individual inter-

pretation of quantum mechanics”, in M. Cini and J. M. Lévy-Leblond (eds.): *The measurement problem of quantum theory*, Bristol: IOP

Primas H. (1990b), “Induced non-linear time evolution of open quantum objects”, in A.I.Miller: *Sixty-two years of uncertainty*, New York: Plenum

Reed M., B. Simon (1972), *Methods of Modern Mathematical Physics* Vol. I, New York: Academic Press, (2nd edition 1980)

Reed. M, B. Simon (1975), *Methods of Modern Mathemanical Physics*, Vol. II, New York: Academic Press

Riesz, B. Sz.-Nagz (1960), *Functional Analysis*, New York: Frederick Ungar Publishing Company

Roberts J. E., G. Roepstorff (1969), *Comm. Math. Phys.* **11**, 321

Robinson D. (1990), pp. 251-261 in *PSA 1990*, Volume 1, East Lansing: Philosophy of Science Association

Robinson D. (1994), *Brit. J. Phil. Sci.* **45**, 79

Robinson D. W. (1968), *Comm. Math. Phys.* **7**, 337

Robson M. A. (1994a), “Geometric quantisation of the phase space of a particle in a Yang-Mills field”, Preprint DAMTP 94-36

Robson M. A. (1994b), “Geometric quantisation on homogeneous spaces and the meaning of ‘inequivalent’ quantisations”, Preprint DAMTP 94-45

- Roessler O.E. (1987), “Endophysics”, in J.L.Casti and A.Karlqvist (eds.): *Real Brains—Artificial Minds*, New York: North-Holland
- Rothstein J. (1952), *Phys. Rev.* **86**, 640
- Rothstein J. (1964), “Thermodynamics and some undecidable physical questions”, *Philosophy of Science* **31**, 40
- Süssmann G. (1957), “An analysis of measurement”, in S. Körner (ed.): *Observation and Interpretation*, London: Butterworth, (reprinted Dover: New York (1962)), 131-136
- Sakai S. (1971), *C*-Algebras and W*-Algebras*, Berlin: Springer
- Schlick M. (1939), *Problems of Ethics*, New York
- Schlieder S. (1968), *Commun. Math. Phys.* **7**, 305
- Schreye D. de (1983), *Math. Scand.* **53**, 265-280
- Schmidt E. (1908), *Mathematische Annalen* **63**, 433-476, and **64**, 161-174
- Scutaru H. (1977), *Lett. Math. Phys.* **2**, 101
- Segal I. E. (1947a), *Bull. Amer. Math. Soc.* **53**, 73
- Segal I.E. (1947b), *Annals of Mathematics* **48**, 930
- Sikorski R. (1960), *Boolean algebras*, Berlin: Springer, 3rd edition 1968
- Stinespring W. F. (1955), *Proc. Amer. Math. Soc.* **6**, 211
- Stratila S., Zsidó L. (1975), *Lectures on von Neumann Algebras*, trans-

lated from Romanian (1975), Tunbridge Wells: Abacus Press, (1979)

Streater R. F. (1965). Proc. R. Soc. A **287**, 510

Streater R. F. (1990), Comm. Math. Phys. **132**, 201

Svozil K. (1993), *Randomness and Undecidability in Physics*, Singapore: World Scientific

Swieca J.A. (1970), in *Cargèse Lectures in Physics Vol. 4* (ed. Kastler, D.), p. 215, New York: Gordon and Breach (1970)

Szilard L. (1929), Z. Phys. **53**, 840-856

Takesaki M. (1968), Acta Math. **119**, 273

Takesaki M. (1973), Acta Math. **131**, 249

Takesaki M. (1979), *Theory of Operator Algebras I*, Springer: New York

Tarski A. (1956), *Logics, Semantics, Metamathematics*, Oxford: Clarendon

Tarski A. (1969), "Truth and Proof", *Scientific American* **220**, 63

Unnerstall T. (1990), Comm. Math. Phys. **130**, 237

van Fraassen B. (1991), *Quantum Mechanics: An Empiricist View*, Oxford: Clarendon

van Hemmen L. (1978), Fortschr. Phys. **26**, 397

van Hove L. (1951), Mémoires de l'Académie Royale de Belgique **26** (6)

Varadarajan V. S. (1968, 1970), *The Geometry of Quantum Theory*, Vols. 1 and 2, North Holland

Wan K.K. (1980), *Canad. J. Physics* **58**, 976

Weizsäcker, C. F. von (1941), “Zur Deutung der Quantenmechanik”, *Zeitschrift für Physik*, **118**, 489-509

Weizsäcker, C. F. von (1961), “Die Einheit der Physik”, in F. Bopp (ed.): *Werner Heisenberg und die Physik unserer Zeit*, Braunschweig: Vieweg, 23-46

Weyl H. (1939), *The classical groups*, Princeton: University Press Whitten-

Wolfe, B. and Emch, G.G. (1976), *Helv. Phys. Acta* **49**, 45-55

Wick G. C., A. S. Wightman, E. P. Wigner (1952), *Phys. Rev.* **88**, 101

Wigner E. P. (1932), *Phys. Rev.* **40**, 749

Wigner E.P. (1952), *Z. Phys.* **131**, 101

Wigner E. P. (1961), “Remarks on the Mind-Body Question”, in I. J. Good (ed.): *The Scientist Speculates*, London: Heinemann, 284-302

Wigner E. P. (1963), *Amer. J. Phys.* **31**, 6-15

Wigner E. P. (1971), in W. Yourgrau , A. van der Merve (eds.): *Perspectives in Quantum Theory*, Cambridge, Mass.: MIT Press

Woodhouse N. M. J. (1976), *Geometric Quantisation*, Oxford: Clarendon, 2nd edition, (2nd edition 1992)

Yanase M. (1961), Phys.Rev. **123**, 666

Zeh H.D. (1970), Found. Phys. **1**, 69

Zeh H.D. (1971), “On the irreversibility of time and observation in quantum theory”, p. 263 in B. d’Espagnat: *Proceedings of the International School of Physics “Enrico Fermi”, Course 49: Foundations of Quantum Mechanics*, New York: Academic Press

Zurek W. K. (1982), Phys. Rev. D **26**, 1862

Acknowledgements

Köszönöm, Ágnes, a szeretetedet, a türelmedet és az erődöt.

I am greatly indebted to my outstanding supervisors, Professor Michael Redhead and Dr Jeremy Butterfield for guiding me through this project. Their friendly support was most valuable both intellectually and personally. They also gathered around them the stimulating philosophy of physics group at Cambridge in which this project grew.

This work also owes much to its “inofficial supervisors”, Anton Amann and Klaas Landsman. They commented competently on many of the more technical parts. I am also grateful to various people who have helped in conversation and correspondence to sharpen certain points of my analysis: David Albert, Twareque Ali, J.-P. Antoine, Guido Bacciagaluppi, Joseph Berkovitz, Pavel Bóna, Ole Bratteli, Jeffrey Bub, Paul Busch, Gianni Casinelli, Gianpiero Cattaneo, Dennis Dieks, Lajos Diósi, Gérard Emch, Brigitte Falkenburg, David Finkelstein, Bas van Fraassen, Nicolas Gisin, Jonathan Halliwell, Keith Hannabus, Meir Hemmo, Jean-Marc Lévy-Leblond, Joseph Melia, Peter Mittelstaedt, Heide Narnhofer, Dénes Petz, Hans Primas, Helmut Rauch, Miklós Rédei, Ray Streater, Abner Shimony, Walter Thirring, Jos Uffink, Lev Vaidman, Paul Weingartner, Arthur Wightman, Carl Friedrich von Weizsäcker, and Wojtek Zurek.

This project was supported financially by the Austrian Ministry of Science and Research, the Governing Body of Peterhouse, the Austrian Chamber of Commerce, the Overseas Research Students Awards Scheme, the Arnold Gerstenberg Trust, the Government of Vorarlberg.

The Institute of Theoretical Physics at the University of Cologne provided the facilities to bring this dissertation into its final form. I am also grateful to Professor Hans Primas and Professor Peter Mittelstaedt for their hospitality during my stays at Zürich and Cologne.