The Impossibility of Accurate State Self-Measurements
Author(s): Thomas Breuer
Source: *Philosophy of Science*, Vol. 62, No. 2 (Jun., 1995), pp. 197-214
Published by: The University of Chicago Press on behalf of the Philosophy of Science Association
Stable URL: http://www.jstor.org/stable/188430
Accessed: 04-02-2017 12:53 UTC

THE IMPOSSIBILITY OF ACCURATE STATE SELF-MEASUREMENTS*

THOMAS BREUER†‡

Department of History and Philosophy of Science
University of Cambridge, UK

It is shown that it is impossible for an observer to distinguish all present states of a system in which he or she is contained, irrespective of whether this system is a classical or a quantum mechanical one and irrespective of whether the time evolution is deterministic or stochastic. As a corollary, this implies that it is impossible for an observer to measure the EPR-correlations between himself or herself and an outside system. Implications of the main result are discussed for how we have to conceive of universally valid theories.

1. Introduction. In this paper I shall analyze 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 often (see for example Dalla Chiara 1977, Peres and Zurek 1982, Roessler 1987, Finkelstein 1988, Penrose 1989, Primas 1990) 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 paper 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 2 some arguments claiming that absolutely universally valid theories cannot be deterministic will be reviewed and criticised.

In section 3 I give the central argument why no apparatus can distinguish all states of a system in which it is properly contained. Self-reference properties 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.

Then in section 4 I turn to quantum mechanics and arrive at the additional conclusion that no quantum mechanical apparatus can measure

*Received July 1993; revised July 1994.

†I am grateful to Jeremy Butterfield, Dennis Dieks, Joseph Melia, Peter Mittelstaedt, Michael Redhead, and an anonymous referee for comments on drafts of this paper.

‡Send reprint requests to the author, Peterhouse, University of Cambridge, Cambridge, CB2 1RD, England.

Philosophy of Science, 62 (1995) pp. 197–214

the Einstein-Podolsky-Rosen correlations between itself an external system.

In section 5 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 in the sense that there is no experiment able to distinguish all states. Still an absolutely universally theory might be ontologically meaningful. From an operational point of view, theories can at most be universally valid in a relative, observer dependent sense. This corroborates conjectures made in the context of quantum mechanics by Peres and Zurek 1982, Roessler 1987, Finkelstein 1988, and Primas 1990, saying that self-reference problems might be the reason why quantum mechanics is not applicable to the observer. But it shows two additional things: first, in quantum mechanics these conclusions do not rely on the deterministic or linear character of the time evolution; and second, we have similar conclusions in classical theories.

2. Received Arguments Against Deterministic Universally Valid Theories. *2.1. Absolute Universal Validity.* In a rather vague and strong formulation, 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 nuclear physics to sociology. I call it universally valid in an *absolute* way because it does not make any reference to observers.

At first sight, an absolutely universally valid theory of material reality seems 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, the demon could use it to calculate any future state of the universe from its present state.

Popper (1950) argues that however complete the information provided to the demon about its own past or present state, there will always be some 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 the 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 itself, then it must be a truly supernatural being outside our material world. It is not just demonic because of its great calculational abilities.) Rothstein (1964) shows that the second law of thermodynamics imposes restrictions on the accuracy of measurements which Maxwell's demon, considered as part of the thermodynamic system, can perform. Dalla Chiara (1977) and Peres and Zurek (1982) present different arguments for why a *deterministic* theory cannot be universally valid in the absolute sense. They also arrive at similar conclusions: non-self-predictability is inevitable. Every deterministic theory must admit the existence of unpredictable events when a predictor applies it to himself. This unpredictability is present even in classical mechanics with a deterministic time evolution.

I believe the conclusion of non-self-predictability is correct. It will follow from the fact that no observer can obtain or store information sufficient to distinguish all states of a system in which he is contained. (Non-self-predictability implies that we can never fully verify the allegedly deterministic character of an absolutely universally valid theory. Being unable to assess whether or not the time evolution of the world is deterministic, the problem of whether free will and determinism are compatible loses some of its relevance. But the fact that the assumption of a deterministic time evolution of the universe cannot be checked does *not* preclude deciding whether free will and determinism are compatible. This may be decidable by conceptual analysis alone.) But of course non-predictability, or non-self-predictability, does not disprove determinism.

2.2. 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.

The reason they give is the following. 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. They argue that a universally valid theory that 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. First, 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 does not hold for the single branches of the universe.)

Second, 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 behavior 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 (1982) conclude 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 Schrödingers 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 point here is not that quantum mechanics can really be applied to *every* phenomenon of the external world. Quantum theory merely serves as an example of a theory which might be universally valid.) They conclude that experimentally fully-justifiable theories can be universally valid at most in the relative sense.

I agree with Peres and Zurek's conclusion, but in my view their argument must be challenged. First, 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 the one I use, but of course Peres and Zurek are free to do that. Second, 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. Peres and Zurek need the assumption of determinism only to preclude the observer in a universally valid theory from having free will, thereby excluding experimental verifiability of a universally valid deterministic theory. If determinism does not necessarily exclude the free will of the observer, then the argument fails.

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

3. The Central Argument. In this section I present an argument as to 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.

3.1. Self-reference in Physical Theories. Popper 1950, Rothstein 1964, Dalla Chiara 1977, Peres and Zurek 1982, Roessler 1987, Finkelstein 1988, Primas 1990, and Mittelstaedt 1993 allude to 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 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. Good experiments give information about the state of the system, and whether or not this information is compatible with the proposition under consideration can then be checked. So good experiments serve to at least partially constitute the semantics of physical theories. In this sense, observation is a semantic concept.

Tarski (1956, 1969) calls a language semantically closed 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 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. First, 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. Second, Gödel proved his result without assuming that provable statements are true. This considerably strengthens his result, because it does not rely on the controversial concept of "truth in informal number theory." (Guido Bacciagaluppi pointed out to me that 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. Gödel relied only on this concept of truth.) ω-consistency was all that Gödel needed, and in modern proofs it can in fact be replaced by the weaker requirement of consistency. It is probably not exaggerated to say that my argument does not have much more in common with Gödel's proof than the use of self-reference.

3.2. Description of Measurements. Let us assume that we have a physical theory whose formalism specifies sets of possible states for the systems it describes. 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., normalized L^1-functions on phase space. In quantum mechanics, the individual states would be pure, i.e., extremal, normalized, positive, linear functionals on the observables, whereas the statistical states would be σ-weakly continuous and therefore correspond to the normal, positive, normalized, 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 $\mathcal{P}(\mathcal{S}_{\mathcal{A}})$ of the set $\mathcal{S}_{\mathcal{A}}$ of apparatus states into the power set $\mathcal{P}(\mathcal{S}_{\mathcal{O}})$ of the set $\mathcal{S}_{\mathcal{O}}$ of system states.[1] 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 $\theta(\mathcal{S}_{\mathcal{A}})$ are the possible states of O after the experiment; usually not every state of O is a possible state after the experiment. We have $\theta(\mathcal{S}_{\mathcal{A}}) \subset \mathcal{S}_{\mathcal{O}}$.

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

I will say that in an experiment with inference map θ a state $s_o \in \mathcal{S}_{\mathcal{O}}$ is *exactly measurable* if after the measurement there exists a set $S_A \in \mathcal{P}(\mathcal{S}_{\mathcal{A}})$ 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 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 $\theta(S_A^1) \not\ni s_2 \not\ni \theta(S_A^2)$.

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

[1]Note that the curly symbol $\mathcal{S}_{\mathcal{A}}$ refers to the set of *all* apparatus states, whereas S_A refers to some set of apparatus states. Similarly for $\mathcal{S}_{\mathcal{O}}$ and S_o.

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 $\theta(\mathcal{S}_{\mathcal{A}})$. But for two states to be distinguishable it is not necessary that either of them be exactly measurable. If all possible final states are distinguishable from one another, then they are all exactly measurable.

The concepts of exact measurability and distinguishability are strong. The results to follow—namely the impossibility of distinguishing from inside all states—do not deny the possibility of internal observers knowing *something* about their own states.

Let me say something about distinguishability of states for external observers in classical and quantum theory. 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.[2] But I do not think that this is a problem since it is only due to the fact that in a statistical description of experiments one uses a concept of state which describes individual systems.

In quantum mechanics the situation is different. No measurement of the first kind can distinguish all states of an individual system: the only pure states which are a possible measurement outcome are the eigenstates of the measured observable. If a pure state is either not an eigenstate of the measured observable or is an eigenstate belonging to a degenerate eigenspace, it is not exactly measurable. On the statistical level everything is all right 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 (see Busch 1982, Mittelstaedt et al. 1987). These measurements are informationally complete[3] and therefore can distinguish all statistical states.

[2]The reason for this (see Primas 1979) is the following: In 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 $L^{\infty}(\Omega)$ as the algebra of observables of the classical system with phase space Ω, the statistical states of a classical system correspond to the normalized elements of the predual $L^1(\Omega)$. They are probability measures on the phase space. Since there is no normalised L^1-function on Ω whose support is just one point, no individual state is a statistical state and vice versa. Therefore no statistical experiment can specify a unique individual state.

[3]A positive-operator-valued measure a on the value space $\mathbf{R}^n$ is called *informationally complete* if $\mathrm{tr}(a(\Delta)\rho_1) = \mathrm{tr}(a(\Delta)\rho_2)$ for all Borel subsets Δ of $\mathbf{R}^n$ is only possible if the

Traditionally it has been considered a peculiarity of quantum mechanics that no single experiment can distinguish all pure states. But the argument I will 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 role of the observer are perhaps not so specific for quantum mechanics, rather they seem to reflect a more general problem.

3.3. Measurements from Inside. Let us now 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 rest R. We assume that the observed system has more degrees of freedom than the apparatus and contains it. This can be formulated in an *assumption of proper inclusion*:

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

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 surjective map from the states of O to the states of A. (Later I will, by a slight abuse of notation, also denote by $|_A$ the map from $\mathcal{P}(\mathcal{S}_{\mathcal{O}})$ into $\mathcal{P}(\mathcal{S}_{\mathcal{A}})$ 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 $\mathcal{S}_{\mathcal{A}}$, $\mathcal{S}_{\mathcal{O}}$ but also on the restriction map $|_A$. One can give examples of sets $\mathcal{S}_{\mathcal{A}}$, $\mathcal{S}_{\mathcal{O}}$ and two restriction maps such that the assumption of proper inclusion is satisfied with respect to one but not the other. (Take for example as $\mathcal{S}_{\mathcal{O}}$ the natural numbers and as $\mathcal{S}_{\mathcal{A}}$ the even natural numbers. If one takes as restriction map $\mathcal{S}_{\mathcal{O}} \rightarrow \mathcal{S}_{\mathcal{A}} : n \mapsto 2n$, then the assumption of proper inclusion is not satisfied. If the restriction map associates to every $n \in \mathcal{S}_{\mathcal{O}}$ two times the biggest natural number less or equal to $n/2$, then the assumption of proper inclusion is satisfied.) This may seem odd but it is not. After all, the elements of $\mathcal{S}_{\mathcal{A}}$ and $\mathcal{S}_{\mathcal{O}}$ are states of different systems. Therefore, even if $\mathcal{S}_{\mathcal{A}}$ is some subset of $\mathcal{S}_{\mathcal{O}}$,

density matrices ρ_1, ρ_2 are equal. The positive-operator-valued measures on $\mathbf{R}^n$ describe generalized observables with values in $\mathbf{R}^n$. Observables in the traditional sense are self-adjoint operators on the Hilbert space and induce via their spectral resolution a projection-valued measure on $\mathbf{R}$. Such an observable can never be informationally complete.

we cannot infer automatically that A is a subsystem of O; an arbitrary subset of $\mathcal{S}_{\mathcal{O}}$ cannot in general 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 $\mathcal{S}_{\mathcal{A}}$ or $\mathcal{S}_{\mathcal{O}}$, namely the fact that A is a subsystem of O. That A is a subsystem of O not only depends 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.)

Let me briefly say something about the connection of my results to the self-measurements of Albert (1983, 1987). The quantum mechanical automata described by Albert measure also something about themselves, but they do not attempt to determine their own state exactly. In Albert's description the apparatus A is composed of several subsystem A_1, A_2, A_3, ... 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.

3.4. 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 exludes 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 in classical mechanics is at most feasible if the phase space $\mathcal{S}_{\mathcal{O}}$ 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.

3.5. A Consistency Condition and the Main Results. 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 $s_A \in \mathcal{S}_{\mathcal{A}}$, 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: $\forall s_A \in \mathcal{S}_{\mathcal{A}} : \{s|_A : s \in \theta(\{s_A\})\} = \{s_A\}$.

(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 logics: it just guarantees 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.

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 $\mathcal{S}_{\mathcal{A}}$ to $\mathcal{S}_{\mathcal{O}}$.[4]

PROPOSITION 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.* $\exists s_o \in \mathcal{S}_{\mathcal{O}}, \forall S_A \in \mathcal{P}(\mathcal{S}_{\mathcal{A}}) : \theta(S_A) \neq \{s_o\}$.

Proof: To prove this indirectly, suppose now that the observer can measure all states of O exactly: $(\forall s \in \mathcal{S}_{\mathcal{O}})(\exists S_A \in \mathcal{P}(\mathcal{S}_{\mathcal{A}})) : \theta(S_A) = \{s\}$. This assumption together with the meshing condition will lead to a contradiction.

From the assumption of proper inclusion it follows that $(\exists s, s' \in \mathcal{S}_{\mathcal{O}}) : s|_A = s'|_A, s \neq s'$. By assumption there are $S_A, S'_A \in \mathcal{P}(\mathcal{S}_{\mathcal{A}})$ such that $\theta(S_A) = \{s\}$, $\theta(S'_A) = \{s'\}$. Since $\cup_{s_A \in S_A} \theta(\{s_A\}) = \theta(S_A) = \{s\}$ 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 $\{s\} = \theta(\{s_A\}) = \theta(\{\theta(\{s_A\})|_A\}) = \theta(\{s|_A\}) = \theta(\{s'|_A\}) = \theta(\{\theta(\{s'_A\})|_A\}) = \theta(\{s'_A\}) = \{s'\}$, contradicting $s \neq s'$. QED.

LEMMA: The meshing condition implies that

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

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 $\theta(\{s_A\}) \subset \{s \in \mathcal{S}_{\mathcal{O}} : s \in \theta(\mathcal{S}_{\mathcal{A}}), s|_A = s_A\}$. Conversely, let $s \in \mathcal{S}_{\mathcal{O}}$ be such that $s|_A = s_A$ for some s_A and $s \in \theta(\mathcal{S}_{\mathcal{A}})$. Then there is a $s'_A \in \mathcal{S}_{\mathcal{A}}$ 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 $s \in \theta(\{s_A\})$. QED.

PROPOSITION 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:*

$$(\forall \theta) : ((\nexists S^1_A, S^2_A \in \mathcal{P}(\mathcal{S}_{\mathcal{A}})) : \theta(S^1_A) \ni s_1 \notin \theta(S^2_A), \theta(S^1_A) \ni s_2 \not\ni \theta(S^2_A)).$$

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

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

Proof: Assume that there exists an inference map θ, and sets S_A^1, S_A^2 of apparatus states such that $\theta(S_A^1) \ni s_1 \notin \theta(S_A^2)$, $\theta(S_A^1) \ni s_2 \in \theta(S_A^2)$. This will lead to a contradiction.

From $s_1 \in \theta(S_A^1) = \cup_{s \in S_A^1} \theta(\{s\}) \not\ni s_2$ 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 $s_2 \in \theta(\mathcal{S}_{\mathcal{A}})$. Using the Lemma we conclude from $s_1 \in \theta(\{s_A^1\})$ that $s_1|_A = s_A^1$. Since $s_2 \in \theta(\mathcal{S}_{\mathcal{A}})$ 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 $s_A \in S_A^1$. QED.

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

Corollary: 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). $(\forall s \in \mathcal{S}_{\mathcal{O}})(\exists s_A \in \mathcal{S}_{\mathcal{A}}) : \theta(\{s_A\}) = \{s\}$ *implies* $(\exists s_o \in \mathcal{S}_{\mathcal{O}}) : \theta(\{s_o|_A\})|_A \neq \{s_o|_A\}$.

Proof: From the assumption of proper inclusion we know that there are $s, s' \in \mathcal{S}_{\mathcal{O}}$ such that $s \neq s'$, $s|_A = s'|_A$. By the antecedent, there are states s_A, $s_A' \in \mathcal{S}_{\mathcal{A}}$ 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 $s|_A$. QED.

The state $s_o|_A$ plays a role 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* if by measuring it one can distinguish all the states (for a precise definition see footnote 2.) 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.* In spite of the intuitive similarities with Gödel's theorem we should not forget the fundamental differences between the two situations.

4. Measurement of EPR-correlations. 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.

EPR-correlations and Their Measurability from Inside. Consider again an observed system O containing the apparatus A and some environment or residue R, $O = A \cup R$. Assume that all these systems are correctly described by quantum mechanics. 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 $\psi = \Sigma\ \lambda_n \alpha_n \otimes \beta_n$, 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 $\psi^1|_R$, $\psi^2|_R$.

Generalizing this property of EPR-correlations in pure states of a composite system, we will say that two (possibly mixed) 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.

After a first kind 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, the states of A do not determine uniquely all possible states of R: states of the composite system in which the strict correlations do not obtain are physically possible. (Usually the joint system is in such a state before the measurement.) 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.

Now 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. I will argue that for the apparatus A it is impossible to distinguish the states s_1, s_2. (In the context of quantum field theories a similar result was shown by Komar (1964).)

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 $A \cup R$. 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 Proposition 2. It implies that for the apparatus A there is no inference map θ, und 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. We therefore conclude that A is unable to distinguish s_1 and s_2.

Hence A cannot distinguish 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.

5. Universal Validity Revisited. *5.1.* 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. (In the terminology of Roessler (1987), Finkelstein (1988), and Primas (1990) a system without external observer is called endophysical.) If, and this is a slightly stronger assumption, the union of all observers fulfills the assumption of proper inclusion, then, according to Proposition 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 share out the work of measuring the state of O_u between several observers, for if the union of observers still obeys the assumptions of proper inclusion and the meshing condition, then the Proposition 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 one answers this question depends on one's 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 the knowledge of the entities. From this point of view, the fact that no experiment, not even in principle, can 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 that 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 able to distinguish all states. But, as we have seen above, absolutely universally valid theories do not have this property. From this 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 realize is that for an external observer, or one who is at least partially external, the assumption of proper inclusion is violated, and thus Proposition 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 having an observer outside the biggest system it can describe be universally valid?

If we interpret "describe" in the ontic sense of "is true of," a theory having 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, every system to which he can apply the theory must not contain himself. Still, the theory might be universally valid in the sense that for every part 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 the relative sense*.

Nowhere in these considerations does the assumption enter that we are dealing with a quantum mechanical system; the whole argument holds true for classical mechanics as well. As long as one adopts a strictly operationalist position, one can conclude by the above argument that even classical mechanics can be universally valid at most in the relative sense.

5.2. The Universal Validity of Quantum Mechanics. It has often been claimed that traditional quantum mechanics cannot describe the observer and that therefore it is universally valid at most in the relative sense. This conclusion was briefly discussed in section 2 in connection with the argument of Peres and Zurek (1982).

Applied to quantum mechanics, my arguments lead to two novel aspects. First, the conclusion of relative universal validity does not depend on the deterministic or on the linear character of the Schrödinger 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.

Second, for quantum mechanics one arrives at stronger conclusions. Up to now, I have discussed implications of self-reference problems for the universal validity of general physical theories. The starting point of the argument was that, if the union of all potential observers is properly included in the universe O_u, no experiment can distinguish all states of O_u. This was a consequence of Proposition 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. Therefore, 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 Proposition 2 that *given two such states, there is no experiment able to distinguish between them*. This is stronger than the conclusion that no observer can distinguish all states.

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.

REFERENCES

Albert, D. Z. (1983), "On Quantum-mechanical Automata", *Physics Letters A 98*: 249–252

———. (1987), "A Quantum Mechanical Automation", *Philosophy of Science 57*: 577–585

Bohm, D. (1951), *Quantum Theory*. Englewood Cliffs: Prentice Hall

Busch, P. (1982), "Indeterminacy Relations and Simultaneous Measurements in Quantum Theory", *International Journal of Theoretical Physics 24*: 63–92

Dalla Chiara, M. L. (1977), "Logical Self-Reference, Set Theoretical Paradoxes and the Measurement Problem", *Journal of Philosophical Logic 6*: 331–347

Einstein, A.; Podolsky, B.; and Rosen, N. (1935), "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Physical Review 47*: 777–780

Finkelstein, D. (1988), "Finite Physics", in R. Herken (ed.): *The Universal Turing Machine. A Half-Century Survey*, Oxford: Oxford University Press, pp. 349–375

Komar, A. (1964), "Undecidability of Macroscopically Distinguishable States in Quantum Field Theory", *Physical Review 133B*: 542–544

Mittelstaedt, P. (1993), "Measurement induced interrelations between quantum theory and its interpretation", in P. Busch, P. Mittelstaedt, and P. Lahti: *Proceedings of the Symposium on the Foundations of Modern Physics 1993*, Singapore: World Scientific, pp. 269–280

Mittelstaedt, P., Prieur, A., and Schieder, R. (1987), "Unsharp Particle Wave Duality in a Photon Split-Beam Experiment", *Foundations of Physics 17*: 891–903

Penrose, R. (1989), *The Emperor's New Mind*. Oxford: Oxford University Press

Peres, A. and Zurek, W. H. (1982), "Is Quantum Mechanics Universally Valid?", *American Journal of Physics 50*: 807–810

Popper K. R. (1950), "Indeterminism in Classical Physics and Quantum Physics", *British Journal for the Philosophy of Science 1*: 173–195

Primas, H. (1979), "Foundations of Theoretical Chemistry", in *Quantum Dynamics of Molecules: The new Experimental Challenge to the Theorist*, NATO Advanced Studies Series Vol. 57, New York: Plenum Press, pp. 41–113

———. (1990), "Mathematical and Philosophical Questions in the Theory of Open Quantum Systems", in A. I. Miller (ed.): *Sixty-two Years of Uncertainty: Historical, Philosophical and Physics Inquiries into the Foundations of Quantum Mechanics*, New York: Plenum

Roessler, O. E. (1987), "Endophysics", in J. L. Casti and A. Karlqvist (eds.): *Real Brains—Artificial Minds*, New York: North-Holland, pp. 25–46

Rothstein, J. (1964), "Thermodynamics and Some Undecidable Physical Questions", *Philosophy of Science 31*: 40–48

Schmidt, E. (1908), "Zur Theorie der linearen und nichtlinearen Integralgleichungen", *Mathematische Annalen 63*: 433–476, and *64*: 161–174

Tarski, A. (1956), *Logics, Semantics, Metamathematics*. Oxford: Clarendon

———. (1969), "Truth and Proof", *Scientific American 220*: 63–77