2

Quantum Measurement, Probability, and Logic

In the centuries preceding the development of quantum mechanics, the conception of mechanical systems as objects existing independently of conscious agents and possessing physical magnitudes that can, in principle, be arbitrarily well specified was rarely questioned by physicists. In classical mechanics, that is, that of the tradition of Newton, Lagrange, and Hamilton, the full set of physical magnitudes describing each physical system is precisely determined at all times by a collection of six parameters, the *dynamical variables* of vector position $\mathbf{q}$ and vector momentum $\mathbf{p}$, together constituting the state $(\mathbf{q}, \mathbf{p})$, in accordance with Hamilton's partial differential equations for the Hamiltonian function $H(\mathbf{q}, \mathbf{p})$, and all imprecision of state specification is entirely due to the ignorance of agents as to this objective state.

In classical mechanics, the domain of the dynamical variables constitutes the *state space*, in which states evolve in time along precisely specifiable trajectories in such a way that each magnitude is at every time uniquely determined by the state at any earlier time, that is, there is strict *causality*. Bohm described the introduction of causality as a principle of physics as follows.

> "[O]ne begins to consider the possibility that in processes by which one thing comes out of others, the constancy of certain relationships is no coincidence. Rather, we interpret this constancy as signifying that such relationships are *necessary*, in the sense that they could not be otherwise, because they are inherent and essential aspects of what things are. The necessary relationships between objects, events, conditions, or other things at a given time and those at later times are then termed *causal laws*." ([53], pp. 1-2)

The state of a compound system is also taken to be fully described through those of its subsystems as specified by their dynamical variables, that is, their (generalized) positions and momenta subject to the appropriate form of the causal law of motion. This is the view of the world as a clockwork. In the event that the state of a classical system is incompletely known to a conscious

agent, the agent can describe the system statistically through the behavior of a corresponding ensemble of possible copies of the system in states compatible with his knowledge, each copy possessing precise values for its physical magnitudes some or all of which are not precisely known by that agent. Such ensembles can be used by another agent with more complete knowledge of the states of the individual copies for communication with that agent by sending these more precisely specified members of the ensemble as signals the receipt of which provide information to the more ignorant agent when received.

Hans Reichenbach argued that inferences to determinism, such as captured by Bohm's simple reconstruction, are suspect, especially when they involve inferences from behavior at one length scale to determinism at another.

> "The idea of determinism, i.e., of strict causal laws governing the elementary phenomena of nature, was recognized as an extrapolation inferred from the causal regularities of the macrocosm. The validity of this extrapolation was questioned as soon as it turned out that macrocosmic regularity is equally compatible with irregularity in the microcosmic domain, since the law of great numbers will transform the probability character of the elementary phenomena into the practical certainty of statistical laws. Observations in the macrocosmic domain will never furnish any evidence for causality of atomic occurrences so long as only effects of great numbers of atomic particles are considered." ([373], p. 1)

Already early in the twentieth century, the deterministic conception of mechanics was brought into question not only in the realm of quantum theory but also in relation to the mechanics of classical systems having evolutions with extreme sensitivity to initial conditions for arbitrary initial conditions [384]. It is noteworthy that radioactive decay, observed as early as 1896, has often retrospectively been viewed as pivotal for the early acceptance of *indeterminism*, even though that appears not to have actually been the case. In particular, "van Brakel (1985) has surveyed the literature and come to the conclusion that 'before 1925 there is no publication in which the 'indeterministic' nature of radioactive decay is considered to be a remarkable aspect of the phenomenon'... On the other hand, he finds many publications claiming such a role, all written *after* 1928" ([479] p. 140, [459]). That is, there is a distinct possibility that the introduction of the New quantum mechanics of microscopic systems motivated a revisionist account of the initial significance of radioactivity for the increased acceptance of indeterminism.

The traditional pre-quantum mechanical picture of matter included causal relations between systems located in space and time and involved in processes under which they move between initial and final states in a *continuous* manner and may interact with one another. A classic statement of determinism,

following from causality, was made by Pierre Simon, Marquis de Laplace in reference to the entire universe within space and time as a physical system:[1]

> "We ought then to regard the present state of the universe as the effect of its anterior state and as the cause of the one which is to follow. Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain, and the future, as well as the past, would be present to its eyes." ([123], pp. 3-4)

Laplace's characterization has two aspects, that of the first sentence, which is metaphysical, and that of the second, which is epistemic in that it has to do with what is in principle *calculable*. What is *in practice* calculable may be significant as well, because, for example, the universe may have a finite number of components that could be used by agents, such as ourselves, for the performance of calculations, a point which is addressed later in Chapter 4.

For the evolution of a complex system, even a classical one, to be deterministic, a number of assumptions must be valid, including the assumption that the system of interest is truly *closed*. Although a finite complete universe is a closed system by definition, any part that human beings might actually comprehend may never be closed, as reflected in the view of Émile Borel that the classical description of gas "composed of molecules with positions and velocities which are rigorously determined at a given instant is. . . a *pure abstract fiction*" because one is driven, for the purposes of practical physics, to consider the *external* forces acting on them as indeterminate [66]. Furthermore, von Smoluchowski's 1918 model of radioactive decay based on sensitivity to initial conditions suggested early on that causality and random phenomena are not inherently incompatible.

> "As an explanation of the origins of the random variables observed in classical systems, he suggested what is known as sensitive dependence on initial conditions. In order to show that there is no contradiction between 'lawlike' causes and 'random' effects, von Smoluchowski constructs a mechanical model reproducing precisely the exponential law of radioactive decay. To complicate matters further, he says that he 'of course does not believe radium atoms really possess such a structure' [as that of a tiny planetary system]. Instead, radioactivity can be taken 'as the most complete type of 'randomness'." ([479], p. 140)

---

[1] Also see [244], Section 2.6. It is interesting to compare this with the perspective of the Collapse-Free interpretation discussed in the following chapter, which assumes the validity of a universal quantum state deterministically evolving.

The later popular idea that macroscopic behavior is essentially deterministic whereas microscopic behavior is essentially indeterministic can be viewed as the result of the long history of successes of classical physics in the macroscopic realm and the necessity of explaining later 'anomalous' phenomena in terms of the quantum mechanical behavior of microscopic systems, against a background of deeply embedded philosophical assumptions. As Dudley Shapere has pointed out,

> "Determinism was a guiding principle...because it was based on more general abstract or philosophical ideas, largely inherited from the Greeks (or before) about what an explanation ought to be. One could not have an explanation unless it explained every detail of experience and allowed the specific prediction or retrodiction of every detail of experience. That ideal of what an explanation ought to be and must be if the theory is to be explanatory is rejected in quantum mechanics." ([99], p. 148)

Although irreducible randomness, as opposed to limitations on practical descriptions free of stochastic elements, was increasingly accepted after the arrival of the New quantum theory, Schrödinger began his 1935 cat paper by reviewing the basic elements of the traditional approach, which he described as "an ideal of the exact description of nature," as a basis for criticizing the new state of affairs in physics vis-à-vis its conceptual foundations after the formalization of quantum mechanics had been essentially completed, although he strongly criticized "naive realist" interpretations of the theory [394].

Einstein was also gravely concerned about the status of quantum mechanics in relation to these long-standing ideals of natural philosophy. "That business about causality causes me a great deal of trouble...I would be very unhappy to renounce complete causality" ([72], p. 23). However, as Arthur Fine quotes Einstein saying in one of his letters to E. Zeisler, "For us causal connections only exist as features of the theoretical constructs" ([174], p. 87). Indeed, by 1950, Einstein had certainly come to see realism as a more crucial element of a proper physics than causality. "In the center of the problematic situation I see not so much the question of causality but the question of reality (in a physical sense)" (Quote from a 1950 letter to Jerome Rothstein, [174] p. 87). It is not that Einstein had a prejudice against probabilistic laws, but rather only against irreducibly probabilistic laws in a *fundamental* theory (*cf.*, [479], Chapter 4). Furthermore, as John Stachel has shown, "it was radical non-locality that most bothered Einstein, even more than its probabilistic element" ([429], p. 246, *cf.* [428]). Einstein sought *both* realism *and* locality in physical theory.

Few, if any, of the founders of quantum mechanics were entirely inflexible in the face of the new phenomena and the successes of quantum theory. Rather, they sought the best balance of fundamental principles that still allowed for the preservation of scientific naturalism. Thus, for example, Bohr's approach to quantum mechanics was an attempt to describe quantum phe-

nomena within the broad outlines of traditional scientific methodology that emphasized exactly that the scope of some principles may be limited by that of others. This is most clear in his position that there is complementarity between the continuous space-time description of microscopic systems and their causal description. Related to but distinct from indeterminism, which is metaphysical and regards the relation between system magnitudes of different times, are *indeterminacy* (or *indefiniteness*), which is metaphysical and regards magnitudes at just one time, and *uncertainty*, which is epistemic and may relate to one time or several times.

> "Bohr uses the concept of 'complementarity' at several places in the interpretation of quantum theory. The knowledge of the position of the particle is complementary to the knowledge of its velocity or momentum... ; still we must know both for determining the behavior of the system. The space-time description of the atomic events is complementary to their deterministic description... [The change in the course of time of the probability function] is completely determined by the quantum mechanical equation, but it does not allow a description in space and time. The observation, on the other hand, enforces the description in space and time but breaks the determined continuity of the probability function by changing our knowledge of the system." ([219], pp. 49-50)

His disciple Heisenberg, just quoted, was quite willing to abandon causality.

> "Since all experiments obey the quantum laws and, consequently, the indeterminacy relations, the incorrectness of the law of causality is a definitively established consequence of quantum mechanics itself." ([216], p. 197)

The differences of applicability of different sorts of 'Heisenberg relation' can lead to confusion in regard to the distinction between indeterminism and uncertainty. This relates to the fact that probabilities associated with quantum states ρ may have both an objective and a subjective aspect, with the boundary between them depending on the interpretation of the formalism assumed.

Eugene Wigner noted that in standard quantum mechanics there is "the possibility of an observation giving various possible results even on a system with a well defined and completely known state," where "the acausality of the theory manifests itself only at the observations undertaken" [502]. However, the concept of *objective indefiniteness*, namely, that physical magnitudes inhere in an object without being simultaneously definite allows both indeterminism and indefiniteness to obtain without essentially involving the mind in the description of measurement, as Wigner ultimately suggested. This idea has been explicated more recently by Shimony, who has articulated it in the following simple and striking context. "If... we concede that [a Bell state] Ψ is a complete description of the polarization state of the pair of photons [involved in demonstrating the violation of Bell's inequality], then we must accept

the *indefiniteness* of the [relevant projections of] polarization of each... as an objective fact, not as a feature of the knowledge of one scientist or of all human beings collectively." Furthermore, "[w]e must also acknowledge *objective chance* and *objective probability*, since the outcome of the polarization analysis of each photon is a matter of probability" ([412], pp. 177-178). Shimony made the following suggestion for interpreting the quantum state.

> "It is convenient to use a term of Heisenberg to epitomize objective indefiniteness together with the objective determination of probabilities of the various possible outcomes; the polarizations of the photons are *potentialities*. The work initiated by Bell has the consequence of making virtually inescapable a philosophically radical interpretation of quantum mechanics: that there is a modality of existence of physical systems which is somehow intermediate between bare logical possibility and full actuality, namely, the mode of potentiality." ([412], pp. 177-178; also *cf.* [419], p. 108)

The interpretation of quantum theory, that is, the drawing out of the epistemic, metaphysical, and operational significance of the elements of the theory, especially of the state function $|\psi\rangle$, is necessary because the formalism, like that of any other physical theory is not, as is often suggested, self-interpreting.

In this chapter, various ways that physical magnitudes and measurements have been quantum mechanically characterized are considered. This will equip us with much of what we will need to appreciate more fully the import of interpreting quantum mechanics. After setting out pertinent elements of logic and probability theory, we consider fundamental results pointing out the above restrictions and explore some apparently paradoxical situations that arise in the application of the theory. For now, let us keep in mind the view of quantum mechanics that prevailed in 1935, as summarized by Schrödinger.

> "Continuing to expound the official teaching, let us turn to the already mentioned ψ-function. It is now the means for predicting probability of measurement results. In it is embodied the momentarily-attained sum of theoretically based future expectation, somewhat as laid down in a *catalog*. It is the relation- and determinacy-bridge between measurements and measurements, as in the classical theory the model and its state were. With this latter the ψ-function moreover has much in common. It is, in principle, determined by a finite number of suitably chosen measurements on the object, half as many as were required in the classical theory. Thus the catalog of expectations is initially compiled. From then on it changes with time, just as the state of the model of classical theory, in constrained and unique fashion ("causally")—the evolution of the ψ-function is governed by a partial differential equation (of first order in time and solved for $\partial\psi/\partial t$). This corresponds to the undisturbed motion of the model in classical theory. But this goes on only until one again carries out any measurement." ([394])

2.1 Logic and Mechanics

In characterizing a physical system at a given moment, one can both formally and operationally consider the set of *propositions*, that is, true/false questions about its magnitudes and their definiteness; one traditionally speaks of a system having or not having any specific value of any magnitude by virtue of that value lying or not lying in the corresponding Borel subset of the real numbers.

In the case of a classical mechanical system, there is a specific real spatial position vector of the particle and a specific real momentum vector, that is, a location in phase space, as long as this particle exists. For classical systems, complex propositions can be formed from the simplest ones via Boolean logical connectives; for any proposition regarding system properties, there is a subset of state space for which the proposition is true, any two propositions being physically equivalent if the same subset of state space attributes the same value to them, modulo sets of Lebesgue measure zero. One can always provide a characteristic function for all Borel subsets of the state space that provides the desired logical mapping to propositions. These functions are idempotent, that is, they are equal to their squares. However, when one attempts to extend this approach to the relation between states and properties to quantum mechanical systems, it fails, and does so in specific ways. In particular, the quantum state does not fully determine specific values of all their physical magnitudes: from the logical point of view, the distributive law fails.

The failure of the distributive law can be related to the term *observable* taking the place of the term *property* for the characterization of the physical magnitudes of quantum-mechanical systems. Recall that the quantum state of a system at a given time is capable of precisely specifying only a corresponding subset of the physical magnitudes, that is, the currently "observed" quantities and the others that are functions only of them, in the intervals between measurement events, this being so only as long as quantities other than these remain unmeasured, that is, "unobserved." As Dirac put it,

> "The expression that an observable 'has a particular value' for a particular state is permissible in quantum mechanics in the special case when a measurement of the observable is certain to lead to the particular value, so that the state is an eigenvalue of the observable. . . In the general case we cannot speak of an observable having a value for a particular state, but we can speak of its having an average value for the state. We can go further and speak of the probability of its having any specified value for the state, meaning the probability of this specified value being obtained when one makes a measurement of the observable." ([139], pp. 46-47)

These probabilities are the expectation values provided by the Born rule, although there is some debate as to propriety of this procedure in the context of traditional probability theory.

Bell argued that the use of the term *observable* can be misleading, much as Bethe argued that the term *uncertainty* can be.

> "There is indeed much talk of 'observables' in quantum theory books. And from some popular presentations the general public could get the impression that the very existence of the cosmos depends on our being here to observe the observables. I do not know that this is wrong. I am inclined to hope that we are indeed that important. But I see no evidence that this is so in the success of contemporary quantum theory." ([24], p. 170)

Bell, Einstein, and Schrödinger sought a formulation of quantum theory lying unequivocally in the traditional realist approach to physics in which physical properties are independent of observing minds, and did so not simply for the sake of tradition. Heisenberg described Einstein's similar concern as follows.

> "He pointed out to me that the very concept of observation was itself problematic. Every observation, so he argued, presupposes that there is an unambiguous connection known to us, between the phenomenon to be observed and the sensation which eventually penetrates into our consciousness. But we can only be sure of this connection, if we know the natural laws by which it is determined. If, however, as is obviously the case in modern atomic physics, these laws have to be called in question, then even the concept of 'observation' loses its clear meaning." ([224], p. 114)

Recall that, in standard quantum mechanics, the observables are Hermitian linear operators on separable Hilbert spaces, which is a rather indirect representation by comparison to the direct representation of physical magnitudes as functions in classical mechanics. A difference between classical and quantum mechanics also arises in that, in the quantum case, the dynamical variables, that is, the observables are constrained by commutation relations, and form a non-Abelian (non-commutative) algebra; in the case of simple position Q and momentum P, for example, one has $[Q, P] = i\hbar$.

In order to enable a logical interpretation of the quantum state $|\psi\rangle$ along the lines of that of classical mechanics, physical magnitudes could be understood via a *value state*, which is a Boolean value (0 or 1) attributed to the system for each pairing of observable and value of that observable. However, according to the Born rule, the Hilbert-space ray for pure quantum states $|\psi\rangle$ instead provides only *probabilities* for all observables at any one time, in particular, to the idempotent observables (projectors) which correspond to propositions [67]; even in the case of pure states, only the proposition corresponding to one *Hilbert-space ray* which takes the value 1, and the propositions corresponding to rays orthogonal to it, which take the value 0, have truth values. The probabilities provided by the Born rule are not all definable as measures over properties on a classical Kolmogorovian probability space, but are well defined as such only in relation to quantities that have been measured

and not subsequently disturbed by further measurements of non-commuting quantities [244, 439]. As a result, the probabilities provided by quantum mechanics are best not defined on simple sets but instead on quantum events or, operationally, experimental questions, that is, on pairings of individual observables (defined on Hilbert space) and individual Borel subsets of the real line (associated with values of measurement outcomes). Thus, the sort of probability that appears in quantum mechanics is a *generalized* probability.

As with probability, which is taken up in detail in the following section, it has been suggested that a generalization of standard *logic* is required for the proper description the quantum world [166, 373, 437, 480]. In particular, it has been suggested by Hilary Putnam and others that such a revision might lend greater coherence to physical theory through the adoption of a conception of logical truth under an empiricist approach to logic, within which logic is capable of revision when our knowledge of the world increases as it has with the emergence of quantum theory [366]; Redhead has suggested this view of logic might be called *instrumentalist* [371]. By contrast with the case of probability, however, the position that logic is empirical has come under considerable criticism. One objection arises from the fact that it eliminates the universality of logic in that, for example, human reasoning is already well described by classical propositional logic but is apparently *not* by quantum logic; there then would be, at the very least, different logics in different realms [427].[2] Other problems, as pointed out by Allen Stairs, are that

> " 'Quantum Logic' means so many things to so many people that it has almost ceased to be a useful term. . . There is disagreement as to whether quantum logic has anything to do with *logic*, and even among those who think it does, there is disagreement concerning the nature of logic in general and what quantum logic in particular can do for us in our quest to understand quantum theory. . . In what I take to be the most interesting version of quantum logic, the word 'logic' is taken very seriously. The aim is to present quantum mechanics as a theory that posits novel relations among events or states of affairs. . . which are reflected in logical structures of a strongly non-classical character. On this view, logic functions in *explanations:* it is part of an account of the strange behavior which quantum systems exhibit." ([432])

Quantum logic in the broadest sense has a history stretching back to the early 1930's and the work of Garrett Birkhoff and von Neumann, who demonstrated the possibility of connecting the mathematics of lattice theory and Hilbert space [48]. Their manner of associating logical states with quantum systems was to straightforwardly assign binary values to closed linear sub-

---

[2] An extensive but accessible introduction to more contemporary quantum logic can be found in Chapter 7 of [244]. A useful bibliography is [340].

spaces of the Hilbert space of the quantum system as mentioned above.[3] In the context of classical physics, Boolean logic and the traditional understanding of the logical connectives and of negation interpretable in set-theoretic terms can be unproblematically used, with the atomic propositions being those corresponding to a system being in one of the subsets of classical phase space. Again, however, the projection operators representing propositions regarding a quantum system form a non-Boolean algebra. Thus, to construct a quantum logic, one must use different structures, for example, those related to *partial Boolean algebras*.[4]

The quantum logic of von Neumann and Birkhoff, often called the *logic of subspaces*, arises from the set of Hilbert subspaces of the complex Hilbert space $\mathcal{H}$ describing the quantum system of interest, as follows. Each subspace $\bar{h}$ is identified with the operator $P_{\bar{h}}$ that projects onto the subspace. The lattice $\bar{L}(\mathcal{H})$ of closed linear subspaces of a Hilbert space $\mathcal{H}$ is seen to be equivalent to the lattice of projection operators on $\mathcal{H}$. One can define the two operations $\wedge$ (*meet*) and $\vee$ (*join*) acting pairwise on any two projectors P_1 and P_2 by $P_1 \wedge P_2 \equiv P_1 P_2$, $P_1 \vee P_2 \equiv P_1 + P_2 - P_1 P_2$, and identify the zero as the projector $\mathbb{O}$ onto the zero vector $\mathbf{0}$ and the identity as the projector $\mathbb{I}$ onto all of $\mathcal{H}$; $\vee$ corresponds to the linear span, $\wedge$ to intersection. The rays of $\mathcal{H}$ are considered to be the atomic propositions of $\bar{L}(\mathcal{H})$.[5] Compound propositions formed from them correspond to higher-dimensional closed linear subspaces. The conjunction, $\wedge$, is essentially the same as the conjunction of classical logic. The central conclusion of Birkhoff and von Neumann was that

> "one can reasonably expect to find a calculus of propositions which is formally indistinguishable from the calculus of linear subspaces with respect to *set products, linear sums*, and *orthogonal complements*— and resembles the usual calculus of propositions with respect to *and, or*, and *not*." ([48])

However, the disjunction $\vee$ (join) and negation $'$ behave much differently from to set-theoretic disjunction and negation even if, as Putnam claims, these are taken to be synonymous with those of traditional propositional logic. In particular, Putnam has argued that one can reasonably claim that "adopting quantum logic is *not* changing the meanings of the connectives, but merely changing our minds about [the distributive law]" [366].

---

[3] Given this, it is especially important to distinguish quantum logic from the implementation of Boolean logic in quantum information science, which is based on the manipulation of computational basis states by 'quantum logic gates'; *cf.* Section 4.5.

[4] A detailed summary of the elements of Boolean and quantum logic are given in section 5 of the Appendix.

[5] An excellent critical résumé of the quantum-logical approach to quantum theory can be found in [427]. For up-to-date reviews of quantum logic, see [117, 440].

The propositions of quantum logic refer to the state of the physical system at a given time. Their semantic interpretation involves no reference to the preparation or measurement of the system. Following Birkhoff and von Neumann, a quantum logic of events involving a system can be "read off" the Hilbert space that describes it. The inherent algebraic structure of the projection operators on the Hilbert space, that of the partial Boolean algebra, is formed by a family of Boolean algebras $B^{(i)}$ if the following conditions are satisfied.

(1) The set-theoretical intersection, $B^{(i)} \cap B^{(j)} = B^{(k)}$, of two members $B^{(i)}, B^{(j)}$ is a member of the family.

(2) If three elements of the partial Boolean algebra are such that two of them belong to a given member of the family, then there exists a Boolean algebra of which all three are members.

The complement of any element of the partial Boolean algebra is its complement with respect to any of the members of a family to which it belongs; the complement of the projector P_i is the operator $\tilde{P}_i = \mathbb{I} - P_i$ such that $\tilde{P}_i \wedge P_i = \mathbb{O}$ and $\tilde{P}_i \vee P_i = \mathbb{I}$. This complement is then unique and belongs to any family to which the element belongs. Importantly, the matching of truth values with closed linear subspaces of Hilbert space does *not* require that the corresponding propositions are necessarily either true or false; this mathematical correspondence is compatible with the indefiniteness of truth values of statements regarding physical properties inherent in quantum mechanics discussed in Chapter 1.[6] Furthermore, the truth of an elementary quantum proposition is *insufficient* to determine the value of all other propositions.

Again, the central move of the quantum logic approach is to consider the lattice of propositions defined on Hilbert space, as described above, rather than Boolean lattice of propositions defined on classical phase space or, equivalently, to consider the related partial Boolean algebra. One can then attempt to provide a truth-valuation, mapping subspaces onto the truth values of B_2, to arrive at a quantum propositional calculus in a manner resembling that in classical propositional calculus. Such a valuation is subject to an *admissibility criterion*: A truth-valuation is admissible if and only if there is a ray R such that for every subspace S the value is 1 if and only if the former is a subspace of the latter. Another peculiarity, in comparison with traditional logic, is that the logic depends *on the particular quantum system in question*. Furthermore, for Hilbert spaces of dimension greater than two, the valuations are not homomorphisms from the lattice corresponding to the logic of subspaces $\bar{L}(\mathcal{H})$ to B_2; the valuations are *not* truth-functional, in that the values of compound propositions are not determined by their components [371, 427].

---

[6] It is also noteworthy in this regard that Reichenbach introduced a third logical value as a basis on which to approach quantum mechanics ([373], Section 30).

Recall that, most distinctively, the lattice structure of quantum logic is *non-distributive.* As a relatively concrete illustration of this, consider the propositional structure of a quantum two-level system, the case of spin-1/2. The propositions corresponding to spin along a given direction z can be written $L_z = \{0, p_{up}, p_{down}, 1\}$. For another spatial orientation, $\bar{z}$, one has a similar system of propositions $L_{\bar{z}} = \{\bar{0}, \bar{p}_{up}, \bar{p}_{down}, \bar{1}\}$. By identifying the least upper bounds (lub's) and greatest lower bounds (glb's) of these two sets, one obtains the "horizontal sum" $L_z \otimes L_{\bar{z}}$, which can be endowed with a modular orthocomplemented structure. Even in this simple case involving the 'pasting together' of Boolean subalgebras, one obtains a structure differing from a Boolean algebra because the distributive law fails to hold. For example, beginning with the complex proposition $p_{down} \vee (\bar{p}_{down} \wedge \bar{p}'_{down})$ and applying the distributive law, one finds

$$p_{down} \vee (\bar{p}_{down} \wedge \bar{p}'_{down}) = (p_{down} \vee \bar{p}_{down}) \wedge (p_{down} \vee \bar{p}'_{down})$$
$$p_{down} \vee 0 = 1 \wedge 1$$
$$p_{down} = 1,$$

which is erroneous. Moreover, beginning with the complex proposition of the same form as that above, but with conjunction and disjunction interchanged, one can also similarly obtain $p_{down} = 0$ [440].

In order understand the evolution of quantum systems in time, one considers the quantum state space, which was described by Birkhoff and von Neumann as follows.

"[Points in this space] correspond to the so-called 'wave functions,' and hence [phase space] is again a function space—usually assumed to be Hilbert space... the law of propagation is contained... in quantum mechanics, in equations due to Schrödinger. In any case, the law of propagation may be imagined as inducing a steady fluid motion in the phase-space. It may be noted that in quantum mechanics the flow conserves distances (*i.e.* the equations are unitary)." ([48])

Notably, they viewed this as entirely compatible with causality.

"The [phase space] point p_0 associated with [a system] S at time t_0, together with a prescribed mathematical 'law of propagation,' fix the point p_t associated with S at any later time t; this assumption evidently embodies the principle of *mathematical causation.*" ([48])

This reflects a key element of von Neumann's interpretation of quantum mechanics, which is discussed in the following chapter. Birkhoff and von Neumann then made the following comments with regard to measurement that touches on the distinction between quantum and classical behavior.

"Now before a phase-space can become imbued with reality, its elements and subsets must be correlated in some way with 'experimental propositions' (which are subsets of different observation-spaces)... There is an obvious way to do this in dynamical systems of the classical type. [*their footnote*: "Like systems idealizing the solar system or projectile motion"] One can measure position and its first time-derivative velocity—and hence momentum—explicitly, and so establish a one-one correspondence which preserves inclusion between subsets of phase-space and subsets of a suitable observation-space." ([48])

They then explicitly distinguished this from classical statistical theory.

"In the kinetic theory of gases and of electromagnetic waves no such simple procedure is possible, but it was imagined for a long time that 'demons' of small enough size could by tracing the motion of each particle... measure quantities corresponding to every coördinate of the phase-space involved. In quantum theory not even this is imagined, and the possibility of predicting in general the readings from measurements on a physical system S from a knowledge of its 'state' is denied; only statistical predictions are always possible. This has been interpreted as a renunciation of the doctrine of pre-determination; a thoughtful analysis shows that another and more subtle idea is involved. The central idea is that physical quantities are *related*, but are not all computable from a number of *independent basic* quantities (such as position and velocity)." ([48])

It is noteworthy that Birkhoff and von Neumann considered the requirements of imbuing the space of states with "reality," echoing the vocabulary used by EPR the year before, and that they reflect on causality; one motive for pursuing quantum logic in recent times, as is typical of interpretations of the quantum formalism, has been to remove traditional paradoxes associated with the theory's description of the world while being straightforwardly compatible with 'realism' in the EPR sense.[7] Again, however, the atomic propositions must be attributed definite truth values for this to work. Putnam proposed mapping a new set of propositions $\{\Delta_Q\}$, each associated with the value of each observable Q lying in an Borel subset Δ of its set of possible values, onto corresponding projection operators; these projection operators $P_Q(\Delta)$ are those associated with the subspaces the ranges of which are the subspaces spanned by all eigenvectors $|q_i\rangle$ corresponding to the eigenvalues $q_i \in \Delta$, the set of corresponding measurement outcomes associated with measuring the observable Q. One thereby associates with the operator the proposition 'The state of the system is in the range of $P_Q(\Delta)$' [366]. This is problematical if the state of the system, which Redhead calls the *Putnam state* [371], is not

---

[7] The usage of the term *realism* is taken up in detail in Chapter 3.

in this range and one wishes to assert the value of the elementary proposition in order to retain 'realism.'

The Kochen–Specker theorem, considered in a later section, presents a significant obstacle to straightforwardly providing a truth valuation under these circumstance. In an attempt to enable an admissible truth valuation, Putnam asserted that in quantum logic "every observable Q has a value, but there is no value which it has." Although consistent, this move undermines the intuitions motivating 'realism' in the first place, to such an extent that its appeal is almost entirely lost. The underlying components of the 'realist' understanding of quantum mechanics, are (i) the *value-definiteness thesis*, the idea that every physical magnitude has a definite value at all times [433], and (ii) that measurements reveal those values. These run contrary to the idea objective indefiniteness, which allows a less conflicted description of the behavior of quantum systems. As Richard Healey has argued,

> "The *content* of the claim that every dynamical variable has a precise value is quite obscure once one has adopted quantum logic, given that the standard (classical) inferences can no longer be drawn from it. And, in particular, it is by no means clear that the truth of this claim suffices to permit a naive realist reading of the Born rules, according to which the object system possessed the specific value revealed by a measurement prior to, or independent of, the occurrence of that measurement." ([211], p. 22)

Thus, as Peter Mittelstaedt has put it in relation to the historical development of the interpretation of the quantum probability rule itself,

> "[The] original Born interpretation, which was formulated for scattering processes, was... not tenable in the general case... The probabilities must not be related to the system S in state ϕ, since in the preparation [of the quantum state] ϕ the value a_i of an observable A is in general not subjectively unknown but objectively undecided. Instead, one has to interpret the formal expressions $p(\phi, a_i)$ as the probabilities of finding the value a_i after measurement of the observable A of the system S with preparation ϕ. In this improved version, the statistical or Born interpretation is used in the present-day literature.... On the other hand,... the meaning... for an individual system is highly problematic." ([319], p. 41)

2.2 Probability and Quantum Mechanics

From the perspective of logic, one sees that complications arise in quantum mechanics because not all propositions are compatible, that is, the full set of events in quantum mechanics is non-Boolean in a specific way. The probabilities arising in quantum mechanics are probabilities generalizing those of

the traditional kind. Nonetheless, conceptions deriving from those of classical probability can still be brought to bear on the generalized probabilities given by the Born rule as a postulate of standard quantum mechanics. In this section, we review the basic conceptions of probability. Two fundamental results that help contextualize and characterize probability within quantum theory, namely, Gleason's theorem and the Kochen-Specker theorem, are then considered in the following section. It is valuable to have all these in mind when later considering various surprising or arguably problematic quantum mechanical situations, that is, the "quantum paradoxes" that have been contemplated.

There are several different conceptions of probability, most significantly for our purposes the classical, relative frequency, propensity, and subjective conceptions. Although there are difficulties associated with each, they typically do not bear directly on the situation in quantum mechanics *per se*.[8] Let us survey these conceptions in the mathematical context of the Kolmogorov axiomatization of the probability calculus in modern form; in particular, the third axiom as presented below is a more general one than that introduced by Kolmogorov but reduces to it when the set of events is finite rather than merely countable.[9] In the Kolmogorov axiomatization, one is given the events $A, B, C, \ldots$ and thus the sample space S of events (the unit event being identified in quantum mechanics with the projector $\mathbb{I}$) defined as their union. The triple (S, F, p), where F is a field of subsets of S, is referred to as a *Kolmogorovian probability space* when the following conditions are satisfied by p, in particular, taking $p(E_i) \in \mathbb{R}$ as the *probability* of the event E_i.

(1) For any set of events $\{E_i\}$: $0 \leq p(E_i) \leq 1$.

(2) $p(S) = 1$.

(3) For any countable sequence of mutually disjoint events $E_1, E_2, \ldots$,
$p(E_1 \cup E_2 \cup \cdots) = \sum_i p(E_i)$ (σ-additivity).

The probability of one event, B, conditional on another, A, is written $p(B|A) = p(B \cap A)/p(A)$. If two events A and B are such that $p(B|A) = p(B)$ and $p(A|B) = p(A)$, then they are *probabilistically independent*. In the context of probability theory, one is primarily interested in *experiments* consisting of a sequence of *trials,* each having an elementary event as an outcome. If the trials are all independent, such a sequence is referred to as a *Bernoulli sequence*. If the trials are such that the probability of each event may *not* be independent of its predecessor but *is* independent of all others of the sequence, it is referred to as a *Markov sequence.*

The *classical conception* of probability, which predates the Kolmogorov axiomatization, arose by abstraction from practical situations in which all outcomes are in some sense equally possible, the probability of any one event

---

[8] For an illuminating summary of difficulties in interpreting probability, from which this survey has greatly benefited, see [203].

[9] In some versions of the subjective interpretation of probability, this axiom is brought into question.

being the fraction of the total number of events that it represents, as in the appearance of a quantity as a sum obtained in the rolling of a pair of fair dice by adding together the value of the two upward faces. The introduction of the principle of indifference, namely, that whenever there exists no evidence that favors one possibility over another possibility they have equal probabilities, helps avoid a circularity in the classical conception of probability. However, it may be problematic to explicate the idea of "equally good evidence" for two uncertain events without making use of probability. There is also an issue of the impact of the *number* of events. In some applications, most importantly here being quantum mechanical ones, probabilities may take irrational values; this conception of probability requires the consideration of an infinite number of events for these values to be defined. In such cases, one may appeal to a generalization of the principle of indifference, in the form of the maximum entropy principle. On this principle, one takes from the set of all distributions consistent with background knowledge the one maximizing the information-theoretic entropy. Although this is viable in the countably infinite case, there is a certain arbitrariness to the procedure that is particularly problematic in the case of uncountably infinite sets of events, because the principle of indifference can be applied in ways that are incompatible with each other.

The *frequency conception* of probability is based on the direct identification of the probability of events with their relative frequency of occurrence in the total set (reference class) of actual events. Advocates of the Bayesian subjective interpretation of probability consider this identification a category mistake; that subjective view has been adopted by advocates of the Radical Bayesian interpretation of quantum mechanics, discussed in the following chapter.[10] A distinguishing element here is the consideration of *actual* outcomes as opposed to possible outcomes. Thus, on this conception, probability is defined *operationally*. This poses an immediate problem in cases where irrational values of probability might be considered necessary because such values clearly cannot exist for finite sets of events, such as *physical* measurements, which clearly cannot ever constitute an infinite class if measurement is to be carried out by agents. This problem is typically avoided by considering this probability as an ideal limit as the number of events becomes infinite, which is counterfactual in character, at some cost to its operational character.

The *propensity conception* of probability, by contrast, takes probability to be a physical disposition or tendency of a situation in the world to provide each kind of outcome, or a limiting relative frequency for each such outcome. This conception is particularly practical when one desires to attribute probabilities to events that by definition can only occur *once*, something quite unnatural for the conceptions considered above. For example, Karl Popper considered the probability of an outcome of a given type to be a propensity of a repeatable experiment to produce the given outcome with just that limiting relative

---

[10] A category mistake is committed when one discusses a matter in terms appropriate only to matters of a significantly different character (*cf.* [386], p. 16).

frequency [360]. Propensity may also derive its meaning from the role it plays in theories of interest, including quantum mechanics. Bernard d'Espagnat has argued that quantum "probabilities are both intrinsic and 'of appearing,' which mean [*sic*], they are 'probabilities to appear to observers.' " ([126], p. 326). Typically, however, those advocating the propensity theory of quantum probabilities typically wish *not* to view them in this way.

The *subjective conception* of probability identifies probability not with any objective property of the world but directly with the degrees of belief of relevant agents about events, subject to chosen constraints such as rationality, which involves at least consistency. Traditionally, the subjective conception of probability assumes that events are definite and that probabilities arise due to the ignorance of subjects. A set of alternatives are considered that are in a certain sense symmetric relative to this ignorance, with the result that probability is uniformly divided over the elements of this typically finite set [479].

The specific subjectivist approach to probability due to de Finetti has been taken on board by 'Radical Bayesian' interpreters of the theory. The approach of de Finetti takes an agent's degree of belief in an event to be the probability p if and only if p *units of utility* is the price (the so-called 'fair price') the agent would buy or sell a wager that pays one unit of utility if E occurs and 0 otherwise, assuming that there is precisely one such price, an assumption that is often challenged. One considers a 'Dutch book,' which is a series of bets against an agent that the agent considers acceptable; such a series of bets can be avoided by the agent if his subjective probabilities obey the Kolmogorov axioms, that is, are *coherent* [268]. This provides an operational definition of probability. Probabilities are considered to differ categorically from propositions, so that probability assignments are not considered propositions within the theory. One considers an *arbitrary sum* as being the reward of betting on E. Similarly to the situation in other conceptions of probability, one finds that this sum must then be infinitely divisible in principle in order to guarantee full precision of probability measurement; utility must also depend linearly on the sums to avoid dependency of betting on probability assignments. Upon learning new facts, agents probabilities are updated in accordance with Bayes' rule and are dependent on their prior probability assignments.

With these various conceptions of probability in mind, let us now consider a number of central technical results in the foundations of quantum mechanics that have been obtained in relation to quantum mechanical propositions and related probabilities.

2.3 The Completeness of Quantum Mechanics

The traditional approach to physics before the arrival of quantum mechanics was, as seen above, one in which each state of the system attributes a

definite value to *all* physical magnitudes, that is, one that fulfilled the value-definiteness condition: each proposition regarding the system, which is of the form "$O \in \Delta$" where O is a quantity describing the physical magnitude and Δ is a Borel subset of the real numbers, is assigned a definite truth value. Statistical states are then given as probability measures $\mu_O : \Delta \to p(O, \mu, \Delta)$ on a phase space specifying the probability that a measurement of the magnitude O will lie in Δ when the system is in the state μ. The Kochen-Specker theorem shows this approach to be impossible within the structure of standard quantum mechanics. The related theorem of Gleason specifies the precise form of the admissible quantum mechanical probability measures as a functional of the quantum state.

Gleason's theorem justifies, *contra* EPR, the claim that the state description of standard quantum mechanics is complete, by identifying the form of all admissible measures with that of standard quantum mechanics [193], as assumed, for example, under the Basic interpretation of the theory discussed in Section 3.2 below. In the process, it eliminates an entire class of conceivable local hidden-variables theories that had been considered to complete the description of physical systems and would have relegated quantum mechanics to the status of a trivially statistical theory and so would have demoted it from the status of a fundamental theory of physics.

To understand the completeness of quantum state descriptions, one can begin by considering a map $p(P_i)$ from sets of quantum projectors $\{P_i\}$ to the real numbers between 0 and 1, $p : P_i \mapsto p(P_i)$, such that $p(\mathbb{O}) = 0$ and $p(\mathbb{I}) = 1$, such that $P_1 P_2 = 0$ implies $p(P_1 + P_2) = p(P_1) + p(P_2)$, taking p to be a *countably additive* probability measure, where $\mathbb{O}$ is the projector onto the zero vector **0** and $\mathbb{I}$ projects onto the entirety of the Hilbert space pertaining to the quantum system in question. In the course of his proof, Gleason provided an important *Lemma* ([193][11]): Let $|\phi\rangle$ and $|\psi\rangle$ be two state-vectors in a Hilbert space $\mathcal{H}$ of dimension at least 3, such that for a given system state $\langle P(|\psi\rangle)\rangle = 1$ and $\langle P(|\phi\rangle)\rangle = 0$. Then $|\phi\rangle$ and $|\psi\rangle$ cannot be arbitrarily close to each other. In particular, $|| \, |\phi\rangle - |\psi\rangle \, || > \frac{1}{2}$.

The central result of Gleason is the following.

Theorem ([193]): All probability measures that can be defined on the lattice of quantum propositions from the quantum statistical operators, that is all quantum probabilities, are of the form $p(P_i) = \text{tr}(\rho P_i)$, for some statistical operator ρ on Hilbert space $\mathcal{H}$, for all $\mathcal{H}$ of dimension greater than two.

Gleason's theorem shows that every probability measure over the set of projectors arises from a quantum state ρ on the Hilbert space of the system. The trace measure assigns to each projector the dimension of its range, which can then be normalized by the dimension of the pertinent (finite-dimensional) Hilbert space. It is thus obtainable by considering ρ to be the maximally mixed state on the space (see [371], Section 1.5). Gleason's theorem shows

---

[11] This version of the lemma is that given by Bell [24].

that the only natural generalization of Kolmogorovian probability functions of the type needed for quantum mechanics is just that appearing in Hilbert-space formulation. The values corresponding to orthogonal projectors thus obey a rule of the type introduced by Born [67] and explicated by Pauli.

The relationship between these results and the putative dispersion-free states, for which projectors take expectation values of only either 0 or 1 under the above mapping, can be understood as follows [24]. The condition $\sum_i \langle P(|\phi_i\rangle)\rangle = 1$ implies that both 0 and 1 occur because (1) there are no other possible values for satisfying the condition and (2) neither alone suffices. But, then, there must be arbitrarily close pairs $|\psi\rangle, |\phi\rangle$ having different expectation values, 0 and 1 respectively; however, such pairs cannot be arbitrarily close, by the above lemma. Therefore, there can be *no* dispersion-free states providing quantum statistics. Accordingly, no hidden variables that parameterize dispersion-free probability measures exist for systems with Hilbert spaces of dimension greater than 2 [24]; because it provides the probability measures definable on the lattice of quantum propositions corresponding to the quantum projectors, the set of quantum states is *complete.* In the context of quantum information theory, this is related to the no-cloning theorem, discussed in Section 4.2; dispersion-free states would enable perfect quantum cloning.

To see more explicitly how this considerations bear on the class of local hidden-variables theories, following Kochen–Specker, consider the complete set of Hermitian self-adjoint operators for the entire set of quantum states of a system with a Hilbert space of dimension greater than two, with the very natural constraint that the algebraic relations of these operators must be reflected in the assigned values and the assignment of real numbers to the operators of quantum mechanics that might be taken to represent the values of the corresponding properties of that system. The Kochen–Specker theorem shows that such an assignment cannot be found for a finite sublattice of quantum propositions [277]. Consider the *value function*, v_ψ connecting an observable O to a value of a physical magnitude O when a system is in a state ψ, as mentioned at the outset of this chapter. It is natural to define $F(\mathrm{O})$, the value associated with $F(O)$ for all functions F, where the mapping from values of O to O is one-to-one and onto. One can then imagine taking $v_\psi(F(\mathrm{O})) = F(v_\psi(\mathrm{O}))$, which has the consequence that v_ψ is additive and multiplicative on operators that commute, the latter itself having the consequence that $v_\psi(\mathbb{I})$ for all states ψ so long as there is at least one magnitude O for which $v_\psi(O) = 0$ (*cf.* [247], pp. 191-192). Another consequence of this multiplicativity is that $v_\psi(\mathrm{P}_i)$ must be either 0 or 1 for all propositions P_i, which have corresponding projectors P_i. Thus, if one considers a resolution of the identity into a set of projectors $\{P_i\}$, that is, this set is such that $\sum_i P_i = \mathbb{I}$ in an interpretation of quantum properties where *one and only one* of the corresponding magnitudes P_i can take the value 1, one finds that, except for an overly restricted class of properties, no such function exists.

It is worth noting in passing that, because the calculations involved in Gleason's proof require the dispersion-free states to provide relationships between experiments that cannot, as a matter of principle, be made *simultaneously*, as noted by Bell ([22], Ch. 1), it does not preclude an essentially empty hidden-variables theory of the kind that could be produced for *any* theory, that is, if *no* constraints are placed on the relationships between observables ([277]); it remains possible for a set of observables and states assigning probabilities to Borel sets of values of the observables to have a set of hidden variables that are sufficient, by simply considering real functions, one for each observable such that probabilities associated with the Borel sets, for each state—for each observable, a probability given by the hidden variable distribution given by the inverse of the associated real function will suffice.

2.4 Problems with Measurement in Quantum Mechanics

Many have been uncomfortable with the fundamental role of measurement in quantum mechanics. For example, one of Bell's greatest concerns was that it introduces *conceptual imprecision* into physics.

> "If the theory is to apply to anything but idealized laboratory operations, are we not obliged to admit that more or less 'measurement-like' processes are going on more or less all the time more or less everywhere? Is there then ever then a moment when there is no jumping and the Schrödinger equation applies? The concept of 'measurement' becomes so fuzzy that it is quite surprising to have it appearing in physical theory *at the most fundamental level*. . . does not any *analysis* of measurement require concepts more *fundamental* than measurement? And should not the fundamental theory be about these more fundamental concepts?" ([24], pp. 117-118)

Instead of the term *measurement*, Bell strongly preferred that the term *experiment* be used, because

> "in fact the [former] word has had such a damaging effect on the discussion, that I think it should now be banned altogether in quantum mechanics. . . the latter word is altogether less misleading." ([25], p. 20)

This concern can be related to the concern about the prominence of observation. Some have accepted the centrality of measurement to quantum mechanics despite the fact that it is seen as a key shortcoming of the theory. Nonetheless, even Wigner, one of the more radical in approach to achieving consistency between measurements and the remainder of quantum theory due to his explicit appeal to observation, understood the discomfort it causes.

"The principal difficulty [of quantum mechanics] is that it elevates
the measurement, that is the observation of a quantity, to the basic
concept of the theory... it seems dangerous to consider the act of
observation, a human act, as the basic one for a theory of inanimate
objects. It is, nevertheless, at least in my opinion, an unavoidable
conclusion. If it is accepted, we have considered the act of observa-
tion, a mental act, as the primitive concept of physics..." ([502])

That quantum measurements are not dynamical defined in a way consistent
with the Schrödinger equation but are used to explicate how observables come
to have definite values after not having had them, was well understood by
Wigner, who acknowledged that by accepting measurement as fundamental
to the theory "it may well be said that we explain a riddle by a mystery"
([502], p. 1).

Given that quantum mechanics remains a fundamental theory, something
the results of the previous section support, and that theory ought to be
properly connected with practice, confronting this difficulty is unavoidable,
as Wigner noted, even if his own solution to it may be rejected. The stan-
dard approach to measurement in quantum theory is to consider measuring
instruments to be physical objects that can in principle be realized in the
world, say in a laboratory environment, rather than being mere conceptual
crutches. This is particularly important to bear in mind given the prominence
of thought experiments in the arguments pertaining to the foundations of the
theory, which would lose considerable force were they to be absolutely dis-
tanced from practice. Indeed, because quantum mechanics is one of our most
fundamental theories and, perhaps, the most empirically successful one in his-
tory, it must describe measurement whether it does so alone or together with
other equally fundamental elements of physics, if it is to be considered truly
complete, although, as with indeterminism, there also remains some question
regarding the uniqueness of quantum mechanics in this regard (*cf.* [389], p.
55). As Krips has pointed out, it is

"[not] necessary that [quantum theory] by itself generates [models
for the process of measurement], any more than we require a de-
scription of the functioning of devices for measuring charge, say, to
be entirely given from within electromagnetism (although arguably if
[quantum theory] is to be 'complete' then it must provide such mod-
els)... [Nonetheless,] one would expect that [quantum theory] places
certain restrictions on how to measure... (just as the laws of electro-
magnetism do...). These turn out to be quite strong..." ([281], p.
106)

In quantum mechanics, as in classical physics, measurements must be per-
formed for one to find the values of physical magnitudes of a system, which
may have been prepared in an incompletely known state. Measurements re-
quire the physical coupling of some apparatus to the system that is the object

of measurement. In quantum mechanics, however, due to (at least one version of the) the Heisenberg relations, it cannot simply be assumed that measurements have no effect on the physical state or its description, as can be consistently assumed in classical mechanics. Recall that the distinctive element of quantum theory most clearly setting it apart from classical statistical mechanics is the superposition principle regarding quantum states, which implies the Heisenberg relations. A measurement result is given as the registration of an appropriate physical magnitude, a *pointer observable.* Pointer-magnitude values can differ from the those of the measurement memory register so long as there is a well defined *pointer function* serving to bring the elements of the two sets of values into one-to-one correspondence. In this way, quantum mechanics specifies the connection between the subject (observer) and observed phenomena to which Einstein referred. However, the Schrödinger dynamics typically fails to specify the result obtained.

The ongoing interest in the interpretation of quantum mechanics is strongly related to the degree of concern over the difficulty of providing a detailed account of measurement, which is far greater than, for example, providing one for classical measurements. Standard quantum mechanics assumes a formal relationship between physical magnitudes and eigenstates, namely, that a system magnitude is attributed a definite value *when and only when* the system is in a state that is an eigenstate-vector of the operator corresponding to that magnitude. Otherwise, at least on the Basic interpretation of the theory, no such precise value need be attributed to the physical magnitude of the system. This formal relationship, already mentioned in Chapter 1, known as the *eigenvalue–eigenstate link* [371], makes explicit the meaning of Postulate I of the standard formalism (*cf.* Appendix). Although some interpretations *do* deny this link, those that do inherit other problems as a result, as shown in the following chapter.[12]

The connection between measurement results and physical magnitudes is made explicit through the *calibration postulate*, namely, the requirement that if a system is in an eigenstate of an operator corresponding to a physical magnitude then a measurement of this magnitude leads *with certainty* to an outcome, indicating that the system is in this eigenstate at the moment the measurement ends [319]. A quantum state is thus typically characterized by its *preparation*—for example, as described by Schrödinger in his summary quoted in the introduction of this chapter—which may be performed in the same way as a measurement and may determine the system's quantum state, allowing the prediction of its future quantum state. In this way, one can view probabilities specified by the quantum state as having an implicitly conditional character. However, the conditional character of quantum probabilities does not automatically render them epistemic in nature. The state may (or may not) predict the outcomes of future measurements with perfect accuracy, depending on whether or not the two projectors corresponding to the prepa-

---

[12] See also, for example, [137], p. 22.

ration and the measurement commute; in the former case, they may, whereas in the latter they don't.

The unusual role apparently played by measurement in quantum mechanics, by comparison with classical mechanics, when straightforwardly applying the theory to even the simplest measurement situation, has been one of the motives for the explicit interpretation of the quantum formalism, as evidenced by the Bohr–Einstein debate. Indeed, there is considered to be a *measurement problem* in quantum theory because, for example, when one takes the Schrödinger evolution to describe the closed system constituted by the measuring apparatus and the system under measurement, including their environments when appropriate, absurdities appear due to the unitary character of this evolution.[13] Such a description predicts a number of different but equally valid measurement outcomes if one assumes measurement outcomes to exist whenever the appropriate one-to-one correlation of measuring apparatus states and object system states occurs and the superposition principle is enforced.[14] The measurement problem is also referred to as the problem of 'the reduction of the wave packet' to the post-measurement eigenstate or of 'the actualization of potentialities,' that is, of the appearance of individual measurement outcomes (eigenvalues). The significance of this under a given interpretation of quantum mechanics ultimately relates to some reliance on the eigenvalue–eigenstate link or on causality.

Comparing the standard quantum predictions with the observed results, one can fairly easily see why such a description of measurement fails (*cf., e.g.,* [87], [371]). Consider a measuring apparatus initially in an eigenstate $|p_0\rangle$ of its pointer magnitude. Take the system measured for a magnitude of interest having corresponding Hermitian operator O with discrete non-degenerate eigenvalues $\{o_j\}$ to be in any eigenstate $|o_i\rangle$ before the measurement process begins. Finally, assume that the latter remains unchanged during the measurement process, so that the physical magnitude measured is that before as well as that after measurement has finished. Measurement should then result in composite-system state transformations $|\Psi_j^{(i)}\rangle \equiv |p_0\rangle|o_j\rangle \rightarrow |\Psi_j^{(f)}\rangle \equiv |p_j\rangle|o_j\rangle$, for each value of j that is a possible measurement outcome. However, the system being measured must also be capable of being measured were it instead initially *not* in an eigenstate of O but instead the state $\sum_j a_j|o_j\rangle$, because that is also a state allowed by the superposition principle. Because the temporal evolution operator acting on the composite system of the apparatus subsystem and the measured subsystem—assuming they form a closed system—is unitary, it acts linearly on state-vectors and the formal description of the measurement process involves a transformation of the form $|\Psi\rangle \equiv |p_0\rangle \sum_j a_j|o_j\rangle \rightarrow |\Psi'\rangle \equiv \sum_j a_j|p_j\rangle|o_j\rangle$.

---

[13] This is option discussed in detail in the account of the Collapse-Free interpretation of quantum mechanics given in Chapter 3.

[14] This has been seen as both a weakness and a strength in the case of the Collapse-Free approach, *cf.* 3.4.

Thus, in the statistical operator description, $\rho(0) \to \rho'(t) = U(t)\rho(0)\, U(t)^{\dagger}$, where

$$\rho = P(|\Psi\rangle) \ , \quad \rho' = P\Big(\sum_j a_j |p_j\rangle |o_j\rangle \Big) \ ; \tag{2.1}$$

there is a transition from a pure state to a pure state because unitary transformations preserve purity. However, what one requires of measurement is that the transformation of the overall system result in a final state of the form

$$\rho^{(f)} = \sum_j |a_j|^2 P\big(|\Psi_j^{(f)}\rangle\big) \ , \tag{2.2}$$

that is, a mixed state describing a collection of distinct states with probabilities $|a_j|^2$, where a specific definite outcome is obtained for each measurement. Because the composite system evolves into a coherent superposition involving several distinct measuring system states for the ensemble, rather than just one, this does not constitute a good measurement description. Indeed, *any* unitary evolution, including that given by the Schrödinger equation, predicts that measurement of the quantity O yields neither a definite outcome nor an appropriate mixture. Furthermore, including the pure environmental state in the description makes no difference in this regard.

This failure presents difficulties under a number of interpretations of quantum mechanics. On the Basic interpretation, the Schrödinger cat thought experiment and the Wigner friend experiment, both discussed below, strikingly illustrate the measurement problem, which was to an extent anticipated by the EPR thought experiment. Furthermore, when not introducing state collapse during measurement, again if the above sort of correlation between measuring apparatus and system is taken to constitute a successful measurement as is naturally assumed, for example, in the absence of the introduction of some additional psychophysical process *à la* Wigner, an exponential number of differing sequences of outcomes will arise when *sequences* of measurements are performed.

2.5 Elements of Quantum Measurement Theory

The quantum measurement problem is the result of considering the measurement process to take place in the same way as does any other quantum physical process, as discussed by Bell in the quote provided in the previous section. A natural way of avoiding the problem while continuing to treat measurement devices as quantum systems is to make a distinction between measurements and other physical processes. Measurement has been viewed via two 'stages.'

The first stage is that of the system–apparatus quantum correlation discussed in the previous section, sometimes called *pre-measurement*, which, although necessary, is in itself insufficient for a successful measurement, as just seen. The second stage is a change of state of the measured system, the stage

in which a measurement outcome appears. Von Neumann explained the need for a two-stage description as follows.

> "Why then do we need the special process **1**. for the measurement? This reason is this: In the measurement we cannot observe the system **S** by itself, but must rather investigate the system **S+M**, in order to obtain (numerically) its interaction with the measuring apparatus **M**. The theory of measurement is a statement concerning **S+M**, and should describe how the state of **S** is related to certain properties of the state of **M** (namely, the positions of a pointer, since the observer reads these). Moreover, it is rather arbitrary whether or not one includes the observer in **M**, and replaces the relation between the **S** state and the pointer positions in **M** by the relations of this state and the chemical changes in the observer's eye or even the brain (*i.e.* to that which he has 'seen' or 'perceived')." ([477], Section V.2)

In von Neumann's formal explication of this process, the first stage takes place entirely in accordance with the usual unitary Schrödinger equation (process **2**.), whereas the second stage involves a non-unitary change of state (process **1**.); the apparatus includes the physical system from which measurement outcome data can be obtained, subject to the awareness of the measuring agent.

The discontinuous change of quantum state in the second stage, associated with the measurement outcome coming to be known, accords with the *the repeatability hypothesis*, that an immediate repetition of the measurement with certainty yields the same result as the initial measurement. This second stage was given a specific form, now known as the (original) *projection postulate*, used by Dirac, Heisenberg, Pauli, and others beginning in the late 1920s. The name *projection postulate* for the prescription was first given by Henry Margenau in 1958 [310]; before then, it was typically referred to as a *quantum jump*. Dirac described quantum jumping and corresponding state change, as follows. "[A] measurement always causes the system to jump into an eigenstate of the dynamical variable being measured" [139]. That is, there is a "quantum jump" to the state corresponding to the outcome.[15] This description of state change is sometimes referred to as the *collapse of the wave-function*, as if it were a lawful dynamical process rather than an instantaneous indeterministic change.

One typically finds differences in the descriptions of measurement among interpretations of the theory. For example, on the Copenhagen interpretation, discussed in Section 3.3 below, as articulated by Bohr in particular, the measurement apparatus and process need not necessarily be given an explicit quantum mechanical description as above but must only be capable of description in *classical* terms (*cf.* [322]). Heisenberg described the occurrence of a

---

[15] The projection postulate is a central element of the Basic interpretation of quantum mechanics, *cf.* Section 3.2.

quantum jump in a way that emphasizes the subjective character of the observation associated with (selective) measurement. "[O]bservation itself changes the probability function discontinuously; it selects of all the possible events the actual one that has taken place" ([219], p. 54). Fellow *Copenhagener* Pauli viewed the need for a random projection as arising naturally from the fact that the interaction between the measuring system and the measured system is "in many respects intrinsically uncontrollable," in particular in relation to the idea, rejected by Heisenberg in response to Bohr's cautionary comments, that the uncertainty relations might be similarly understood, as discussed in Section 1.2 [338].

On the Collapse-Free interpretation of the theory, the distinction between the above two sorts of ("relative state") change is not fundamental. Instead, it is understood (in most versions) to result from the emergence of different observational perspectives. On it, measurement is entirely described by the Schrödinger evolution acting in the tensor-product space for the composite system formed by the measured system, the measuring system, and the entire environment of the two, within a superposition of states containing correlations between all of these systems formally similar to that of Equation 2.6; ultimately, one considers this to involve a unique 'wave-function of the universe,' an elaborate superposition state that never 'collapses.' Classical phenomena are also sometimes viewed as naturally emerging in it due to the pervasive effect of quantum state decoherence. Those portions of the resulting ramifying set of situations in which different sequences of measurement outcomes are obtained by measurers are assumed in some way to be inaccessible to each other [319, 410]. This interpretation thus represents an attempt to avoid the quantum measurement problem at the level of the entire universe rather than invoking physical state projection at the fundamental level. The Collapse-Free treatment was first carefully investigated by Everett with the support of John A. Wheeler [135], although it was eventually rejected by the latter. The various versions of the resulting Collapse-Free interpretation and its measurement-related problems are discussed in Section 3.5.

In von Neumann's statistical operator formulation, the change of state during measurement of a quantity, typically one not commuting with that of the preparation, is expressed explicitly by the rule that, when subject to measurement, a quantum system initially in a pure state $P(|\psi\rangle)$ evolves, typically non-unitarily, into a mixed state ρ':

$$P(|\psi\rangle) \longrightarrow \rho' = \sum_i \big(\langle\psi_i|P(|\psi\rangle)|\psi_i\rangle\big)P(|\psi_i\rangle) \ , \qquad (2.3)$$

where the projectors $P(|\psi_i\rangle)$ onto the eigenvectors $|\psi_i\rangle$ sum to the identity. The weights $\langle\psi_i|P(|\psi\rangle)|\psi_i\rangle = w_k$ sum to unity and correspond to probabilities that the values of the physical magnitude being measured are found to be those of the subensembles corresponding to the projectors $P(|\psi_i\rangle)$, after the system was prepared in the state $|\psi\rangle$; see Figure 2.1. The state is thus no longer coherent in the *measurement basis* $\{|\psi_i\rangle\}$.
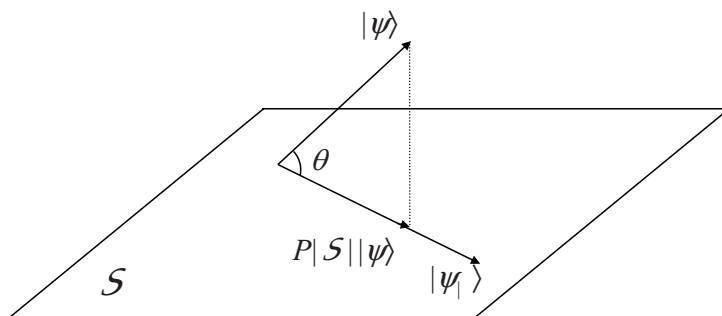
Fig. 2.1. Projection of a quantum state-vector $|\psi\rangle$ into a vector subspace $\mathcal{S}$ by a projector $P(\mathcal{S})$. The projection of $|\psi\rangle$ onto a ray corresponding to $|\psi_m\rangle$, with which it makes an angle θ, is shown here; the probability for this transition to occur is $\cos^2\theta$. A von Neumann measurement corresponds to the set of possible such projections onto a complete orthogonal set rays of the Hilbert space being measured. In a Neumann-Lüders measurement, the projections, then denoted by P_k rather than $P(|\psi_k\rangle)$, are onto subspaces that are not necessarily rays as shown here.

Projectors corresponding to any pair of possible outcomes i and j are related by $P(|\psi_i\rangle)P(|\psi_j\rangle) = \delta_{ij}P(|\psi_i\rangle)$; for each possible outcome, the same subspace would projected to were the measurement immediately repeated, in accordance with the repeatability hypothesis. In an arbitrary basis $\{|\alpha_i\rangle\}$, the matrix elements of the final statistical operator are

$$[\rho']_{ij} = \sum_k \langle\alpha_i|\psi_k\rangle\langle\psi_k|\psi\rangle\langle\psi|\psi_k\rangle\langle\psi_k|\alpha_j\rangle = \sum_k w_k[\rho'_k]_{ij} \ , \qquad (2.4)$$

where $w_k = |\langle\psi_k|\psi\rangle|^2$ and $[\rho'_k]_{ij} = \langle\alpha_i|P(|\psi_k\rangle)|\alpha_j\rangle$, exhibiting the fact that, in general, von Neumann's Process 1 takes pure states to mixtures described by the weights w_k. That is, these probabilities can be understood as standard probabilities. Whenever the state is measured in a basis different from the one in which it was prepared, that is, in general, Process 1 gives rise to a non-unitary change in state and is irreversible [477]. When a subensemble, corresponding to a given value of k, constituting a proportion w_k of the total normalized ideal ensemble, is then also *selected*—in the case of an individual system by its being actualized, which happens with probability w_k—one has $P(|\psi\rangle) \to P_k|\psi\rangle$, the result generally having norm $w_k \neq 1$. The resulting ensemble state must then also be renormalized through division by w_k, and the state may be describable by the statistical operator $P(|\psi_k\rangle)$. Process 1 together with pure subensemble selection constitutes a maximal selective measurement.

By contrast, during the continuous Schrödinger evolution,

$$\rho(0) \rightarrow \rho'(t) = U(t)\rho(0) \ U(t)^\dagger \ , \tag{2.5}$$

where $U(t)$ is the unitary operator describing that 'automatic' temporal evolution, the purity of the statistical operator always remains unchanged. As von Neumann emphasized, when the interaction between a measurement apparatus and a system being measured is analyzed, the unitarity of the Schrödinger equation describing the state evolution of the composite system formed by these two systems provides consistency between alternative descriptions of system behavior in which the division (or *Heisenberg Schnitt*, or *cut*) between measuring system and measured system is chosen differently.

Schrödinger vigorously objected to the introduction of Process 1, due to its evident inconsistency with the usual Process 2, in that

> "any *measurement* suspends the law that otherwise governing the continuous time-dependence of the ψ-function and brings about in it a quite different change, not governed by any laws, but rather dictated by the result of the measurement. But laws differing from the usual ones cannot apply during a measurement, for objectively viewed it is a natural process like any other, and it cannot interrupt the orderly course of natural events." ([394])

Therefore, Schrödinger argued, 'wave-function collapse' could not be a dynamical process; his objection is interpretational and distinguishes Schrödinger's particular views as to how quantum mechanics ought ultimately to be articulated, although he, like Einstein, never arrived at a full interpretation of quantum mechanics with which he was satisfied (*cf.* [49]).

A selective measurement of an observable is said to be *maximal* (or *complete*) when it provides fully distinct values for the quantity measured, so that no further knowledge of its preparation can be obtained by further measurement of the observable. For such measurements, the projection as described by von Neumann's Process 1 in its original form is entirely satisfactory. However, if instead the measurement performed is capable of discriminating only sets of values, the measurement is said to be *non-maximal*; in that case, it provides incomplete knowledge of the observable. Consider, for example, the measurement of a spin-1 system (or *qutrit*), which is a quantum system possessing a trivalent observable O with eigenvalues $o_i = -1, 0, 1$; a maximal measurement will have three possible outcomes, one for each of the possible values. By contrast, a measurement with only two outcomes, say "$-$" for system observable values -1 or 0, and "$+$" for system observable value $+1$, is a non-maximal measurement. An example situation wherein the latter is realized is a (imperfect) Stern–Gerlach type apparatus acting on a spin-1 system such that a particle with z-spin $+1\hbar$ enters a distinct spatial beam downstream from the magnet but particles with spins $0\hbar$ or $-1\hbar$ are not separated and enter a common, second beam.

If the outcome for a measurement of an observable corresponding to a subspace of finite dimension *greater than one*, then the von Neumann projection postulate, including subensemble selection and renormalization, prescribes the process

$$\rho \longrightarrow \rho' = P_k , \qquad (2.6)$$

where the projector is written P_k, rather than $P(|\psi_k\rangle)$, because the projection is made onto a subspace of dimension greater than one. However, this rule dictates a system state after measurement that is *independent* of the details of the state before the measurement, beyond those pertinent to the measurement outcome itself, as can be seen mathematically in that state ρ associated with its prepration does not appear in the description of the resulting state ρ'. Thus, the von Neumann prescription fails to maintain the distinction between initially pure states and initially mixed states, and so fails to preserve coherence of pure states in non-maximal measurements, which is not necessarily lost because state decoherence does not always occur in non-maximal measurements. Nonetheless, the original projection rule can be easily and naturally adjusted to characterize more precisely measurements of physical magnitudes whose eigenvectors have degenerate, that is, non-unique eigenvalues. The prescription for describing the state-change as a result of a selective quantum measurement that maintains the distinction between the two sorts of preparation in such cases is the *Lüders projection* (*Lüders rule*),

$$\rho \longrightarrow \rho' = \frac{P_k \rho P_k}{\text{tr}(\rho P_k)}. \qquad (2.7)$$

By contrast to the original projection rule, the Lüders prescription of state after measurement clearly *is* dependent on the state of the system before measurement; in particular, the values of successive measurements under this rule will coincide when another measurement is made between successive measurements of O of an observable *compatible with O* in the sense that the corresponding operators commute [273, 305]. That is, under the above rule, first introduced by Gerhard Lüders, if one prepares two ensembles of systems in the state ρ, the first being measured for some observable that is compatible with O and the second having O measured *first*, the relative frequencies of the values of the compatible observable for those two cases are the same and yield pure subensembles from pure ensembles. Furthermore, this is the *only* projection rule for which this is true [432], as shown in the next section. This rule is now commonly considered to be the appropriate general description of precise measurements in the standard theory. It can also be viewed as a consequence of Feynman's rules for computing quantum probability amplitudes based on the indistinguishability of processes [427], discussed in Section 3.8.

An important aspect of the approach to measurement of von Neumann is that physical theories are required to accommodate a physical correlate to subjective perception. This is the principle of *psychophysical parallelism*:

"it is a fundamental requirement of the scientific viewpoint—the so-called principle of the psycho-physical parallelism—that it must be possible so to describe the extraphysical process of the subjective perception as if it were in reality in the physical world—*i.e.*, to assign to its parts equivalent physical processes in the objective environment, in ordinary space." ([477], p. 418-419)

He considered chains of physical systems connecting human observers with physical phenomena, as in the situation of Schrödinger's cat discussed below, concluding that one "must always divide the world into two parts, the one being the observed system, the other the observer" (Heisenberg's *Schnitt*), despite the fact that, mathematically, this inclusion can be shown to be arbitrary. That is, in quantum mechanics, one ultimately must be able to consider a bipartite decomposition of composite system involved in a measurement into *observed system* and *observer system*, even if the former lies within a human observer's body (*cf.* [477], p. 419).

In relation to this principle, von Neumann stated that

"[Bohr] was the first to point out that the dual description [of measurement into the usual evolution and state projection] which is necessitated by the quantum mechanical description of nature is fully justified by the physical nature of things that it may be connected with the principle of the psycho-physical parallelism." ([477], p. 420)

This requires only that the *abstraktes Ich* ('abstract I') be associated with one component of the division and that the object of its attention be associated with the other. In the case of Schrödinger's cat, if cats have consciousness of the relevant sort humans have for observation, a conscious cat would certainly be aware that it is alive whenever it is so. If one is concerned primarily with preserving the appearances, then there being two distinct state-evolution processes in the standard formulation of quantum mechanics is not a grave concern; Process 1 comes into play only when and as soon as the observer's interaction with the observed occurs and is consistent with its observations.

However, there remains the issue that "the principle of the psycho-physical parallelism is violated, so long as it is not shown that the boundary between the observed system and the observer can be displaced arbitrarily..." ([477], p. 421). To address this, von Neumann also explicitly demonstrated that changing the boundary between the above two portions of the world has no affect on quantum mechanical predictions, that is, that "the boundary between the two parts is arbitrary to a very large extent" ([477], p. 420), although he did not demonstrate that the stage of the progressing chain of measurement subprocesses from the system to the observer's brain at which the application of (the measurement) Process 1 occurs *in all cases* does not create inconsistency in quantum predictions. In any event, there is no evidence that the superposition or projection of 'states of consciousness' can occur.

This point was considered by Wigner, who argued that to be adequate quantum mechanics must at least provide a prescription for the precise *physical* circumstances in which Process 1 should take place [499], if the quantum measurement problem is truly to be resolved. Wigner identified six views regarding the relevance of Process 1 to this problem: Everett's ("there is no need to assume a reduction"), Fock's ("measuring instruments must be described classically"), Ludwig's ("quantum mechanics does not apply to macroscopic systems"), the London—Bauer elaboration [302] following von Neumann's above (adding a collapse postulate), the view that quantum mechanics never describes individual events (which Wigner considers ultimately solipsistic on the basis of the Friend thought experiment discussed in Section 2.7), and Zeh's ("the state of isolation [of the measuring instrument and measured system] is very difficult to maintain... their state-vector... will soon go over into a mixture")—*cf.* [502], p. 1. As mentioned above, Wigner ultimately held the view that the *act of observation* by the conscious agent making a measurement plays a fundamental role in inducing a wave-funtion projection, through a process now know as the *von Neumann–London–Bauer collapse*.

The quantum-logical approach discussed in Section 2.1, assuming state projection takes place when a measurement outcome is registered by a conscious agent, has been used in the attempt to solve the measurement problem, in the sense of explaining how measurements can have objectively definite outcomes and still accord with quantum equations of motion. The idea is that this problem will not arise if one assumes value-definiteness at *all* times rather than only when the quantum system is in an eigenstate of the observable of interest: if both the measuring apparatus and the measured system have definite values for all physical magnitudes, then the outcomes would naturally arise despite the existence of the joint superposition of Equation 2.1. However, as already noted, the value-definiteness thesis is highly problematic.

Some have argued that the measurement problem is artificial, that is, "it is often held that the restrictions placed on the measurement process from within [quantum theory] are *too* strong in that they impose paradoxical requirements" ([281], p. 107). However, as Shimony has pointed out,

> "[it] has often been claimed that [the measurement problem] is specious, arising from a narrow or inadequate representation of the measurement process... however... the problem is a fundamental anomaly, which cannot be lightly dismissed. Some proposals for solving this problem [postulate] nonlinear or stochastic modifications of quantum mechanics... Our conclusion is that the main conceptual innovations of quantum mechanics [objective indefiniteness, objective chance, objective probability, potentiality, entanglement, and quantum non-locality] are probably embedded permanently in physical theory, but that some further radical innovation will probably have to be made." ([410], p. 373)

Whatever the nature of this innovation, it appears increasingly likely that it will be related to quantum entanglement and have some connection to information theory.[16] Because the following section is somewhat technical, one may prefer on a first reading to proceed directly to Section 2.7.

2.6 Advances in Quantum Measurement Theory

An important result obtained by considering quantum mechanics from the perspective of logic is the demonstration, mentioned above, that Lüders' rule for the state after quantum measurement is the *only one* that prescribes the correct generalization of conditional probability to the quantum realm, in the sense that it provides the classical rule for conditional probability when the two operators related to the pertinent events commute [86, 244, 432].

To see how the Lüders rule is naturally picked out, consider a generalized probability function q on the set of subspaces of Hilbert space $\mathcal{H}$. Because any such function is additive over orthogonal subspaces, it is defined by its probability assignments to the one-dimensional subspaces of $\mathcal{H}$. In addition, Gleason's theorem shows that q is provided via a density operator, ρ_q. Finally, if such a q assigns the value 1 for a projector Q, then it assigns the value 0 to projectors onto rays in the *complement* of the subspace onto which Q projects so that $\rho_q|v\rangle = \mathbf{0}$ for all (not necessarily normalized) vectors along such rays. For any such ray $\bar{v}$, $q(\bar{v}) = \mathrm{tr}(\rho_q|v\rangle\langle v|) = \mathrm{tr}\big(\rho_q\big(Q+(I-Q)\big)|v\rangle\langle v|\big) = \mathrm{tr}(\rho_q Q|v\rangle\langle v|)$, for all $|v\rangle \in \bar{v}$, because of the linearity of the trace and because $(I-Q)|v\rangle = \mathbf{0}$ by the definition of $\bar{v}$. Thus, $q(\bar{v}) = |a|^2 q(|u\rangle\langle u|)$ for some $a \in \mathbb{C}$, where $|u\rangle$ is the normalized vector along the ray, so that *any* q such that $\rho_q = 1$ is fully specified by the values it assigns to vectors in the ray onto which Q projects. Hence, for any generalized probability function p on the subsets of $\mathcal{H}$, there is a unique q on those subsets such that for all subspaces of the subspace onto which Q projects one has $q(P) = p(P)/p(Q)$ and in turn q is uniquely represented by ρ_q. Note then that

$$\frac{QPQ}{\mathrm{tr}(QPQ)} \tag{2.8}$$

is a statistical operator, so that one can write

$$\rho_q = \frac{QPQ}{\mathrm{tr}(QPQ)} \; . \tag{2.9}$$

Therefore, Lüders' rule provides the *unique* generalized probability function q such that, for all projectors into subspaces of the space onto which Q projects, $q(P) = p(P)/p(q)$. This uniqueness result bolsters standard quantum measurement theory, despite the presence of the quantum measurement problem

---

[16] One sophisticated move along these lines is discussed later in Section 3.8.

which might be addressed by an innovative interpretation of the quantum formalism.

One can further advance the description of processes within quantum mechanics along similar lines by continuing to elaborate and generalize the standard theory of quantum measurement, keeping in mind the basic requirements of probability theory. Consider for a moment quantum measurement in a broad mathematical setting where the spectrum of the Hermitian operator O representing a physical magnitude may be continuous, so that a measurement might find its value within a Borel set $\Delta \in \mathbb{R}$ and leave the state of the system with support (O, Δ) with respect to O. A projector $P_O(\Delta)$ from the spectral decomposition of O might describe the quantum mechanically maximally specified state of such a system. Although there are conceptual difficulties with such a description, unsharp measurements do provide a well-defined way of proceeding [95, 426].

For simplicity, let us again consider measurements in the more straightforward case of discrete spectra. *Unsharp measurements* are the class of quantum operations that are described by (normalized) positive-operator-valued measures (POVMs) [118]. Given a nonempty set S and a σ-algebra Σ of its subsets X_m, a *positive-operator-valued measure E* is a collection of operators $\{E(X_m)\}$ satisfying the following conditions.[17]

(i) *Positivity*: $E(X_m) \geq E(\emptyset)$, for all $X_m \in \Sigma$.

(ii) *Additivity*: for all countable sequences of disjoint sets X_m in Σ,

$$E(\cup_m X_m) = \sum_m E(X_m) \ . \tag{2.10}$$

(iii) *Completeness*: $E(S) = \mathbb{I}$.

If the *value space* (S, Σ) of a POVM E is a subspace of the real Borel space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, then E provides a unique Hermitian operator on $\mathcal{H}$, namely $\int_{\mathbb{R}} \text{Id } dE$, where Id is the identity map. The positive operators $E(X_m)$ in the range of a POVM are referred to as *effects*, the expectation values of which provide the quantum probabilities.

Each quantum state ρ induces an expectation functional on $\mathcal{L}(\mathcal{H})$, the space of linear operators on the Hilbert space $\mathcal{H}$. This provides well defined probabilities because the effects are bounded by $\mathbb{O}$ and $\mathbb{I}$, so that the ranges of the effect spectra are restricted to lie in the closed unit interval, due to the positivity and normalization of the POVM; the operator ordering $\leq$ on

---

[17] A *Borel σ-algebra* is the σ-algebra generated by the open intervals (or the closed intervals) on a topological space—for example, in $\mathbb{R}$—which are the *Borel sets*. The set S is often a standard measurable space, that is, a Borel subset of a complete separable metric space. Because such spaces of each cardinality are isomorphic, they are all measure-theoretically equivalent to Borel subsets of the real line, $\mathbb{R}$. The sequences here are taken to converge in the weak operator topology on $\mathcal{L}(\mathcal{H})$ [95].

the effects has the zero operator and identity as its upper and lower bounds. Only if the Hilbert space in question is $\mathbb{C}$ do the effects constitute a lattice, as described in the Appendix A.5; a complementation $\perp$ defined by $E \equiv 1 - E$ exists that satisfies $(E^{\perp})^{\perp} = E$ and reverses the operator order but is not an orthocomplementation, so that law of the excluded middle does not hold, however.[18] POVMs are thus the natural correspondents of standard probability measures, described in Section 2.2, in the operator space of quantum mechanics.

The probability of outcome m upon a generalized measurement of a pure state $P(|\psi\rangle)$ is given by

$$p(m) = \langle\psi|E(X_m)|\psi\rangle = \mathrm{tr}\big((|\psi\rangle\langle\psi|)E(X_m)\big) \ . \tag{2.11}$$

When the state is mixed, this probability is given by

$$p(m) = \mathrm{tr}\big(\rho E(X_m)\big) \tag{2.12}$$

(*cf.* Section 2.2). The output of a POVM measurement of the initial state is exhibited by post-measurement states and corresponding outcome probabilities $p(m)$. The post-measurement state ρ'_m of a system initially described by a statistical operator ρ under a POVM $\{E(X_m)\}$ is often taken to be

$$\rho'_m = \frac{M_m \rho M_m^{\dagger}}{\mathrm{tr}\big(M_m \rho M_m^{\dagger}\big)} \ , \tag{2.13}$$

where each of the $E(X_m)$ can be written $M_m^{\dagger}M_m$, M_m being called a *measurement decomposition operator* (*cf.* [324]); in the special case that the M_m are projectors, this expression coincides with the Lüders–von Neumann measurement rule given by Equation 2.7—this can be seen by recalling that projectors are Hermitian and idempotent.

When, and only when, the measurement operators M_m are *projectors*—so that the POVM is a *projection-valued measure* (PVM)—are they *identical* to decomposition operators $E(X_m)$, in which case they are also multiplicative, that is, $E(X_m \cap X_n) = E(X_m)E(X_n)$ for all countable subsets of the corresponding set—equivalently, $E(X_m)^2 = E(X_m)$. When providing positive outcomes, POVM elements allow one to eliminate quantum states from consideration as describing the measured system.[19]

---

[18] Recall, as noted above in the first section of this chapter, the projection operators *do* form an orthocomplement lattice with just this order and complement. The lattice of projection operators $\bar{L}(\mathcal{H})$ has the *sharp properties* as its elements.

[19] An example of a POVM used in this way is the following [32]. Given the two projectors $P(\neg|\phi\rangle) \equiv \mathbb{I}-P(|\phi\rangle)$ and $P(\neg|\phi'\rangle) \equiv \mathbb{I}-P(|\phi'\rangle)$, where $\langle\phi|\phi'\rangle = \sin 2\theta$, one can construct a POVM $\{E_m\}$ with the elements $E_1 = P(\neg|\phi\rangle)/(1+|\langle\phi|\phi'\rangle|)$, $E_2 = P(\neg|\phi'\rangle)/(1 + |\langle\phi|\phi'\rangle|)$, $E_3 = \mathbb{I} - (E_1 + E_2)$. POVM measurements using $\{E_1, E_2, E_3\}$, for example, are more efficient for quantum key distribution and quantum eavesdropping than traditional measurements described by the projectors $\{P(\neg|\phi\rangle), P(\neg|\phi'\rangle)\}$. Similarly, POVMs sometimes allow quantum state tomography to be performed with improved efficiency.

The effects form a convex subset of $\mathcal{L}(\mathcal{H})$, the extremal elements of this subset being the projection operators. A collection of effects is said to be *coexistent* if the union of their ranges is contained within the range of a POVM. Any two quantum observables E_1 and E_2 are representable as PVMs on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ exactly when $[E_1, E_2] = \mathbb{O}$, following from results of von Neumann for Hermitian operators. For POVMs, however, commutativity remains sufficient but is not necessary for coexistence [95]. The use of POVMs, that is, unsharp measurements thus allows one to circumvent the restriction of commutativity on measurements of noncommuting observables by including *unsharp properties*. A *regular effect* is an effect with spectrum both above and below $\frac{1}{2}$. One can define *properties* in general by the following set of conditions, given an effect A.

(i) There exists a property $A^\perp$;

(ii) There exist states ρ and ρ' such that both $\mathrm{tr}(A\rho) > \frac{1}{2}$ and $\mathrm{tr}(A\rho') > \frac{1}{2}$;

(iii) If A is regular, for any effect B below A and $A^\perp$, $2B \le A + A^\perp = \mathbb{I}$.
   (This renders $^\perp$ an orthocomplementation for the regular effects.)

The set of properties $\mathcal{E}_p(\mathcal{H}) = \{A \in \mathcal{E}(\mathcal{H}) | A \not\le \frac{1}{2}\mathbb{I}, A \not\ge \frac{1}{2}\mathbb{I}\} \cup \{\mathbb{O}, \mathbb{I}\}$ satisfies these conditions. The set of *unsharp properties* is then $\mathcal{E}_u(\mathcal{H}) \equiv \mathcal{E}(\mathcal{H})_p/L(\mathcal{H})$. A POVM is an *unsharp observable* if there exists an unsharp property in its range [95]. Coexistent observables are those that can be measured simultaneously in a common measurement arrangement; when two observables are coexistent, there exists an observable the statistics of which contain those of both observables, known as the *joint observable*—typically, the two observables are recoverable as marginals of a joint distribution on the product of the corresponding two outcome spaces.

The technical advances in the theory of quantum measurement described in this section are formidable. However, beyond the first of the above, their effect on the quantum mechanical world view remains unclear in the absence of an interpretation of quantum mechanics that involves a novel theory of measurement that exploits them. In any event, the importance of measurement in quantum mechanics was recognized early on in the history of the theory; quantum measurement theory makes predictions that threaten basic elements of the traditional understanding of the physical world. We now consider two thought experiments illustrating this.

2.7 Schrödinger's Cat and Wigner's Friend

Two challenging thought experiments relating to the question of the sharpness of measurements and the experiences of observers were provided by Schrödinger and Wigner. Introduced in the decades following the formalization of quantum theory, they helped illustrate the surprising character of its predictions and continue to do so today. They suggest more than ever that

radical innovations are needed to come to terms with the implications of the theory. As will be seen in the following chapter, the interpretations of quantum mechanics that have been offered since the introduction of these thought experiments continue to leave much to be desired, assuming that quantum mechanics remains one of our fundamental theories. As a result, these experiments remain important examples with respect to which one can compare new interpretations and consider the relevance of various physical effects, such as state decoherence. In particular and foremost among the thought experiments that followed the EPR experiment considered in the last chapter are those of Schrödinger's cat and Wigner's friend; the latter is essentially an extension of the former by the inclusion of an additional human observer.

Quantum mechanics allows one to conceive of situations clearly in contradiction with common sense, resulting from the application of the superposition principle in the macroscopic realm. Such an application is desirable given that it has the potential to serve as our most fundamental mechanical theory when proper care is taken in the relativistic case. Schrödinger conceived a number of physical examples, including what he called the "ridiculous case" of the now infamous cat in a box that appears both alive and dead, in which all matter is described quantum mechanically, in order to point out as starkly as possible the difficulties posed by standard quantum theory within a realist world view, such as that of common sense. Indeed, one of the characteristics of standard quantum mechanics is that both the measurement apparatus and the system being measured are treated exactly the same way as any other quantum mechanical object is treated.[20]

Before considering these two thought experiments in detail, it is helpful to have in mind at least one more recent example, beyond EPR's early reality criterion, of how *realism*, which is discussed further in the next chapter, has been related to quantum mechanics. A clear such example has been provided by Krips, who spelled out principles cohering with a realist world view in terms of the elements Q of the set of physical quantities as follows. "(Det Q) Physical quantities always have determinate values, (Pass Q) The measured value of Q in [state] S at [time] t = the value possessed by Q in S at t, and (NDQ) The value possessed by the measured quantity just after measurement = the value registered by the measurement, *i.e.* the measured value" ([281], Index of principles). This amounts to the conjunction of the value-definiteness thesis, a principle of faithful measurement, and the repeatability hypothesis. Although, as will be seen later, this involves a very strong form of realism, it is the sort of realism that Schrödinger appears to have had in mind.

---

[20] Note, however, that the Copenhagen interpretation can finesse this point by appealing to classical mechanics as an equally fundamental theory at the macroscopic scale; arguably, Bohr equivocated in this regard in that in one debate with Einstein he applied the uncertainty principle to at least one macroscopic measuring apparatus (*cf.* Section 3.3).

Roger Penrose has explained the difficulty posed by these experiments, particularly for the realist, as follows.

"Why, then, is there such reluctance to accept the state-vector as describing an actual physical reality?...'Schrödinger's cat' (or some essential equivalent such as 'Wigner's friend,' etc.) It seems hard to believe in any actual 'reality' at the level of a cat which requires that such states of death and life can co-exist. The situation is, of course, more puzzling for Wigner's friend... According to *strict quantum mechanics*, it seems to me, ψ_t—representing some complex linear combination of a live cat and a dead one—should still have to provide an objective description of a reality [in the box]." ([344], p. 132)

Schrödinger himself had at one point hoped that the complex-square of the state amplitude of the wave-function could provide a complete direct physical description of quantum systems in ordinary three-dimensional space. On such an understanding of the wave-function, physical magnitudes in quantum mechanics are "smeared out," something which might be viewed as unobjectionable at scales below and up to the atomic scale (*cf.* [49], pp. 1-2). However, he found such a description objectionable at the truly macroscopic scale, such as that of a cat, because it contradicts experience.

Schrödinger's experiment is the following.

"A cat is penned up in a steel chamber, along with the following diabolical device (which must be secured against direct interference by the cat): in a Geiger counter there is a tiny bit of radioactive substance, *so* small, that *perhaps* in the course of one hour one of the atoms decays, but also, with equal probability, perhaps none; if it happens, the counter tube discharges and through a relay releases a hammer which shatters a small flask of hydrocyanic acid. If one has left this entire system to itself for an hour, one would say that the cat still lives *if* meanwhile no atom has decayed. The first atomic decay would have poisoned it. The ψ-function of the entire system would express this by having in it the living and the dead cat (pardon the expression) mixed or smeared out in equal parts." ([394])

In such an arrangement, under the standard quantum mechanical description of measurement, the answer to the question of whether the cat is alive or dead can be viewed as *indefinite* until the box is opened by an experimenter and the cat is observed. In von Neumann's conception, at the point of observation the observed system is the object of the attention of the *abstraktes Ich*, the essence of the observer, occuring in parallel with a state projection that accords with what is observed ([477], p. 421); until then, one might, at best, only 'approximately' (in some sense [14]) characterize the cat's vital status. The purpose of the example is to point out the apparent inadequacy of the quantum-mechanical characterization of this situation as expressed in his summary quoted at the end of the introduction to this chapter: Any cat is al-

ways either alive or dead, whereas quantum mechanics (apparently) does not describe the situation in that way, an apparently clear failure of the theory in relation to the measurement problem described in Section 2.4.

Although this thought experiment is easily visualized, there are reasons for doubting its efficacy. For example, one can question the very idea of applying quantum mechanics to entire functioning *biological* organisms, all questions of vitalism aside, as Heisenberg did.

> "Logically, it may be that the difference between the two statements: 'The cell is alive' or 'the cell is dead' cannot be replaced by a quantum theoretical statement about the state (certainly a mixture of many states) of the system." ([434])

It is unclear that 'alive' and 'dead' can be immediately reduced to well-defined mechanical quantities. If this theoretical step cannot be made, the example is a non-starter. Moreover, no rigorous experiment measuring such a biological quantum superposition has been performed.[21] This appears to be difficult precisely because it is unclear what *quantum physical* magnitude one would measure. Physicalists are responsible for providing a careful account of how this should be done in practice. If this example is to be taken seriously as an objection to the completeness of quantum theory such an account is clearly required. Nonetheless, the thought experiment retains considerable force in that the vital signs of any organism are clearly physical. Furthermore, Einstein provided a simpler and less easy to criticize example of the same basic sort which avoids the question of applying quantum theory directly to biology in this way: He asked simply whether it makes sense that his *bed* would jump into a definite state only when he or another observer enters the bedroom and looks at the bed. Einstein's example avoids the difficulties presented by involving the vital state of the cat as a physical observable but without the shock value of the consideration of the death of a potential pet.[22]

The Schrödinger cat experiment has been more effectively criticized on the basis that it assumes that the collection of systems from the radioactive sample to the observer who opens the box constitutes a genuinely *closed* system, which it is clearly not in that, for example, the cat also must breathe air to live (*cf.* [391]); Heisenberg noted that the state of this collection will, *de facto*, be a mixed one. However, one can in principle straightforwardly include *all* pertinent objects in any way interacting with these systems without changing its fundamental implication, assuming again that cat biology and all the necessary air molecules are in the domain of quantum theory. Because the entangled state of the composite system in the 'chain' of interactions from the

---

[21] However, work of Zeilinger and co-workers demonstrating the interference of C_{60} molecules, with the distant goal of similarly using *viruses*, shows that the avenue for the experimental investigation of such examples remains open [9].

[22] This example was offered in a personal communication with Putnam, recounted in [367].

radionuclide to the cat and its environment involves a 'fuzzy' micro-physical object, the quantum-mechanical description of the cat will be 'fuzzy' as well. The resulting state of all objects in this chain of interaction will involve the (at least pairwise) entanglement of systems as they interact. This contradicts common sense and what is observed in the natural world, at least at the level of *cats*, in that a heart monitor on a cat records it as either alive or dead and not *both* and not in a superposition of two such states, not to mention one of the cat together with all the physical elements of the chain of measurement.

A further element of paradox arising from the quantum mechanical description of this measurement situation, in Schrödinger's view, is the following.

> "It is rather discomforting that the theory should allow a system to be steered or piloted into one or the other type of state at the experimenter's mercy in spite of his having no access to it," ([394], p. 556)

as is the case when one considers the effect of measurement on entangled systems, as strikingly pointed out by delayed-choice versions of the 'quantum eraser' experiment discussed in Chapter 1. That is, the state of the cat in this example appears to be subject to the choice of measurement of one looking in the box. Nonetheless, von Neumann, in his theory of measurement, provided a demonstration that the changing of boundary between the observing and observed portions of the universe has no essential effect on quantum mechanical predictions and places no requirement to the effect that the application of (the measurement) Process 1 occurs at a specific point in the process of measurement, except to the extent required by psychophysical parallelism.[23]

Wigner argued that an objective prescription for the physical stage at which Process 1 takes place, which is not provided by quantum theory, is needed for a complete characterization of measurement [499]. If the cat's consciousness of its situation were considered part of a self-measurement on its part, this would certainly pertain. More than a quarter century after the introduction of Schrödinger's cat experiment, Wigner offered another thought experiment to probe this question, along the following lines. Consider a system S that flashes when in one state, ψ_1, and does not flash if in an orthogonal one, ψ_2. The friend, F, observing the flash will have corresponding states χ_i, in respective cases. Thus, the bipartite composite system SF will have corresponding states $\psi_i \otimes \chi_i$. If the system, much like Schrödinger's cat, is in a superposition state $\alpha\psi_1 + \beta\psi_2$, the linearity of quantum mechanics dictates that the state of SF must be $\alpha\psi_1 \otimes \chi_1 + \beta\psi_2 \otimes \chi_2$. The probability of the friend seeing the flash will accordingly be $|\alpha|^2$, and that of not seeing the flash will be $|\beta|^2$. In order to provide a good answer to Wigner's query to the friend about his observation, the friend must receive a measurement result when observing S that accords with his answer.

---

[23] To make this connection more specific, a prescription for coordinating psychological or subjective time with physical time might be required.

Wigner noted that, "So long as I maintain my privileged position as ultimate observer," there will be no logical inconsistency with von Neumann's theory of measurement [499]. However, after the experiment, if

> "I ask my friend, 'What did you feel about the flash before I asked you?' He will answer, 'I told you already, I did [did not] see a flash,' as the case may be. In other words, the question whether he did or did not see a flash was already decided in his mind, before I asked him. If we accept this, we are driven to the conclusion that the proper wave function immediately after the interaction of the friend and object was already either $\psi_1 \times \chi_1$ or $\psi_2 \times \chi_2$ and not the linear combination $\alpha(\psi_1 \times \chi_1) + \beta(\psi_2 \times \chi_2)$. This is a contradiction because the state described by the wave function $\alpha(\psi_1 \times \chi_1) + \beta(\psi_2 \times \chi_2)$ describes a state that has properties which neither $\psi_1 \times \chi_1$, nor $\psi_2 \times \chi_2$ has. If we substitute for 'friend' some simple physical apparatus, such as an atom which may or may not be excited by the light-flash, this difference has observable effects and *there is no doubt that $\alpha(\psi_1 \times \chi_1) + \beta(\psi_2 \times \chi_2)$ describes the properties of the joint system correctly,* [whereas] *the assumption that the wave function is either $\psi_1 \times \chi_1$ or $\psi_2 \times \chi_2$ does not.* If the atom is replaced by a conscious being the wave function, $\alpha(\psi_1 \times \chi_1) + \beta(\psi_2 \times \chi_2) \dots$ appears absurd because it implies that my friend was in a state of suspended animation before he answered my question." ([499], p. 293)

That is, if the friend is to have a specific result before being asked, the joint state cannot have been the superposition state in the basis defined by the definite results. From this, Wigner concluded that the result must have become determinate at the moment the friend made his measurement *as a result of his being conscious* and that (i) "it follows that the being with a consciousness must have a different role in quantum mechanics than the inanimate measuring device" and (ii) "the quantum mechanical equation of motion cannot be linear" ([499], p. 294). He argued that the alternative denies the friend consciousness and is tantamount to a form of solipsism. Furthermore, as he put it,

> "In von Neumann's view...the content of [the observer's] mind is not obtainable by means of the laws of the theory. It is either that these laws do not apply to the functioning of the mind (whatever that word means) or that the conscious content of the mind is not uniquely given by its state vector, *i.e.* by the quantity which quantum mechanics uses for the description of all objects." ([502], pp. 122-123)

Thus, Wigner believed that the problem is to be solved by assenting to a special role for consciousness in quantum theory. However, most physicists view this solution as even more objectionable than the original problem, exactly because of its explicit appeal at the fundamental physical level to processes that are not fully physical.