On the Interpretation of Measurement in Quantum Theory

H. D. Zeh

Institut für Theoretische Physik, Universität Heidelberg, Heidelberg, Germany

It is demonstrated that neither the arguments leading to inconsistencies in the description of quantum-mechanical measurement nor those "explaining" the process of measurement by means of thermodynamical statistics are valid. Instead, it is argued that the probability interpretation is compatible with an objective interpretation of the wave function.

1. INTRODUCTION

The problem of measurement in quantum theory and the related problem of how to describe classical phenomena in the framework of quantum theory have received increased attention during recent years. The various contributions express very different viewpoints, and may roughly be classified as follows:

    1. Those emphasizing contradictions obtained when the process of measurement is itself described in terms of quantum theory.[1]

    2. Those claiming that measurement may well be explained by quantum theory in the sense that "quantum-mechanical noncausality" can be derived from statistical uncertainties inherent in the necessarily macroscopic apparatus of measurement.[2]

    3. Those introducing new physical concepts like hidden variables.[3]

    Suggestions of the third group are usually based on the first viewpoint, and are meaningful only if they lead to experimental consequences. These have not been confirmed so far.

A measurement in quantum theory is axiomatically described by means of a Hermitian operator. If the eigenstates of this operator are φ_n, and the state of the measured system S is $\varphi = \sum c_n \varphi_n$, then, according to the axiom, the result of the measurement will, with probability $|c_n|^2$, be the corresponding eigenvalue a_n represented physically by a "pointer position," i.e., by an appropriate state of the measuring device M. For the most frequent class of measurements, it is furthermore predicted that any following measurement can be described by assuming S to be in the state φ_n after the measurement.

When describing the process of measurement as a whole in the framework of quantum theory, it is assumed that the apparatus M can be described by a wave function ϕ_α, the state of the total system $M + S$ obeying the Schrödinger equation,

$$\psi(t) = e^{iHt} \phi_\alpha \sum_n c_n \varphi_n = \sum_{n,m,\beta} c_n U_{\alpha\beta}^{nm}(t) \phi_\beta \varphi_m \tag{1}$$

with $U_{\alpha\beta}^{nm}(0) = \delta_{nm} \delta_{\alpha\beta}$. As the state of a macroscopic apparatus can be determined only incompletely, there must be a large set of states $\{\phi\}_0$ compatible with the knowledge about M. If this set of states is assumed to be independent of the state of S before measurement, a condition on the coefficients $U_{\alpha\beta}^{nm}(t)$ can be derived from the requirement that the axiom of measurement be fulfilled in the case $c_n = \delta_{nn_0}$, i.e., $\varphi = \varphi_{n_0}$. The interaction must be of the von Neumann type[4]

$$U_{\alpha\beta}^{nm}(t) = \delta_{nm} u_{\alpha\beta}^n(t) \tag{2}$$

for all but a negligible measure of states of the set $\{\phi\}_0$, and for times t larger than the duration of the measurement. Furthermore, practically all states $\sum_\beta u_{\alpha\beta}^n(t) \phi_\beta$ must be members of a set $\{\phi\}_n$ corresponding to a "pointer position n" of M.

In the case of a general state φ, the final total state now takes the form

$$\psi(t) = \sum_{n,\beta} c_n u_{\alpha\beta}^n(t) \phi_\beta \varphi_n \tag{3}$$

It represents a superposition of different pointer positions. This result is said to be in contradiction to the axiom of measurement, because the latter states that the result of the measurement is one of the states $\sum_\beta u_{\alpha\beta}^n(t) \phi_\beta \varphi_n$. It is of course very unsatisfactory to assume that the laws of nature change according to whether or not a physical process is a measurement.

The difficulties arising when a macroscopic system is described by quantum theory can be seen more directly by applying the main axiom of quantum theory, i.e., the superposition principle. If there are two possible pointer positions $\{\phi\}_{n_1}$ and $\{\phi\}_{n_2}$, any superposition $c_1 \phi_{n_1} + c_2 \phi_{n_2}$ must be a possible state. As such superpositions have never been observed (see Wigner[1]) one should at least find dynamical causes for their nonoccurrence. Although recent work[5] has shown that dynamical stability conditions in the original sense of Schrödinger's[6] have a much wider field of applicability than previously expected, the process of measurement does not, because of the above arguments, belong to this class of phenomena.

2. CRITICISM OF STATISTICAL INTERPRETATIONS

Results apparently in contradiction to those of the preceding section have been derived in a series of papers[2] which try to make use of the uncertainties in the microscopic properties of the apparatus of measurement. The mathematical concept used in these theories is the density matrix formalism.

A simple example may illustrate such theories. If the density matrix describing M is $\sum_\alpha p_\alpha \phi_\alpha \phi_\alpha^*$, the total system is described by

$$\rho(0) = \sum_{\alpha,n,n'} p_\alpha c_n c_{n'}^* \phi_\alpha \phi_\alpha^* \varphi_n \varphi_{n'}^*, \tag{4}$$

For a von Neumann interaction, one obtains

$$\rho(t) = e^{iHt}\rho(0)\,e^{-iHt} = \sum_{\alpha n n' \beta \beta'} p_\alpha c_n c_{n'}^* u_{\alpha\beta}^n(t)\, u_{\alpha\beta'}^{*n'}(t)\, \phi_\beta \phi_{\beta'}^* \varphi_n \varphi_{n'}^*, \tag{5}$$

Provided the coefficients $u_{\alpha\beta}^n(t)$ possess arbitrarily distributed phases guaranteeing that

$$\sum_\alpha p_\alpha u_{\alpha\beta}^n(t)\, u_{\alpha\beta'}^{*n'}(t) \approx \delta_{nn'} q_{\beta\beta'}^n(t) \tag{6}$$

(the diagonality in $\beta\beta'$ is not needed), $\rho(t)$ becomes

$$\rho(t) \approx \sum_n |c_n|^2\, \varphi_n \varphi_n^* \sum_{\beta\beta'} q_{\beta\beta'}^n(t)\, \phi_\beta \phi_{\beta'}^*. \tag{7}$$

This density matrix describes exactly the situation postulated by the axiom of measurement.[4]

It is tempting to interpret this result by saying that the statistical uncertainty inherent in the macroscopic apparatus is transferred by means of the interaction to the system S. This means that the outcome of a measurement, i.e., the pointer position, should be exactly predictable if we knew the microscopic state of M. Equation (3) demonstrates that this interpretation is wrong.[1]

The contradiction between Eqs. (3) and (7) is—aside from the dubious nature of the statistical assumption—due to a circular argument. The density matrix formalism is itself based upon the axiom of measurement. In order to see this, consider the case of a set of states $\psi^{(i)} = \sum_n c_n^{(i)} \psi_n$ prepared with probabilities $p^{(i)}$. The probability of finding the eigenvalue a_n is then

$$w_n = \sum_i p^{(i)} |c_n^{(i)}|^2 = \mathrm{tr}\{P_n \rho\} \tag{8}$$

---

[1] The above example is not identical with any of the theories of Ref. 2. It does not, however, use any additional assumptions. As it leads to a contradiction, one of the assumptions used must be wrong. Some of these theories do not start with an ensemble for the initial state of the apparatus, but assume instead that the "pointer position" is represented by some time average. The latter is then transformed into an ensemble average by means of the ergodic theorem. Interpreted rigorously, these theories would prove that the pointer position fluctuates in time.

if P_n $\psi_n\psi_n{}^*$, and $\rho = \sum_i p^{(i)}\psi^{(i)}\psi^{(i)*}$. The states $\psi^{(i)}$ will in general not be linearly independent, although ρ may of course be expanded quadratically in terms of a complete orthogonal set. The reason for the usefulness of ρ is that, according to the axiom of measurement, all observable quantities can be expressed as linear-antilinear functionals of the wave function.

For example, the statistical ensemble consisting of equal probabilities of neutrons with spin up and spin down in the x direction cannot be distinguished by measurement from the analogous ensemble having the spins parallel or antiparallel to the y direction. Both ensembles, however, can be easily prepared by appropriate versions of the Stern–Gerlach experiment. One is justified in describing both ensembles by the same density matrix as long as the axiom of measurement is accepted. However, the density matrix formalism cannot be a complete description of the ensemble, as the ensemble cannot be rederived from the density matrix. The discrepancy between Eqs. (3) and (7) arises since, on the one hand, Eq. (3) must hold for all but a negligible number of members of the ensemble, whereas Eq. (7) is interpreted as describing an ensemble of states $\varphi_n \sum_\beta u^n_{\alpha\beta}(t) \phi_\beta$, i.e., each state being essentially different from (3). Only if the measurement axiom is accepted can these ensembles not be distinguished by subsequent observations.

The circularity is more obvious in some versions which avoid the density matrix formalism (e.g., Rosenfeld[2] who made repeated use of the probability interpretation although the latter is to be derived). In such cases, the circular argument is considered a "proof of consistency." This viewpoint cannot be accepted, as it would mean that the secondary observation of the pointer position (by a conscious observer or a second apparatus) is a measurement in the axiomatic sense. It corresponds to the interpretation of measurement due to Heisenberg and von Neumann[4] (claiming the arbitrariness of the position of the "*Heisenbergscher Schnitt*"), and does not require any contribution from thermodynamics. Bohm's analysis of the process of measurement,[7] however, shows the importance of the amplification of the result of a measurement up to the macroscopic scale, thus leading to a natural position of the "*Heisenbergscher Schnitt.*" (Relative phases between microscopically realized pointer positions could still be measured.)

The secondary (macroscopic) observation is significantly different from the primary (microscopic) one, for the physical situation between these two observations is described by the reduced wave function. The macroscopic observation can thus be performed in a reversible way, in contrast to the microscopic one, which must result in the reduction. It is implicitly assumed in applying the density matrix formalism that the macroscopic measurement is accompanied by a reduction of the wave function.

3. CONSEQUENCES OF A UNIVERSALLY VALID QUANTUM THEORY

The arguments presented so far were based on the assumption that a macroscopic system (the apparatus of measurement) can be described by a wave function ϕ. It appears that this assumption is not valid, for dynamical reasons:

If two systems are described in terms of basic states $\phi^{(1)}_{k_1}$ and $\phi^{(2)}_{k_2}$, the wave

function of the total system can be written as $\phi = \sum_{k_1 k_2} c_{k_1 k_2} \phi_{k_1}^{(1)} \phi_{k_2}^{(2)}$. The case where the subsystems are in definite states ($\phi = \phi^{(1)} \phi^{(2)}$) is therefore an exception. Any sufficiently effective interaction will induce correlations. The effectiveness may be measured by the ratios of the interaction matrix elements and the separation of the corresponding unperturbed energy levels. Macroscopic systems possess extremely dense energy spectra. The level distances, for example, of a rotator with moment of inertia 1 g cm² are of the order 10^{-42} eV, which value may be compared with the interaction between two electric dipoles of 1 e $\times$ cm at distance R, $e^2 \times$ cm²$/R^3 \approx 10^{-7} (\text{cm}/R)^3$ eV. It must be concluded that macroscopic systems are always strongly correlated in their microscopic states. They still do have uncorrelated macroscopic properties, however, if the summations over k_1 and k_2 are each essentially limited to macroscopically equivalent states.[8] Since the interactions between macroscopic systems are effective even at astronomical distances, the only "closed system" is the universe as a whole. The assumption of a closed system $M + S$ is hence unrealistic on a microscopic scale.

The arguments leading to Eq. (3) can be accepted only if the states ϕ_α are interpreted as those of the "remainder of the universe" including the apparatus of measurement, instead of those of the latter alone. It is of course very questionable to describe the universe by a wave function that obeys a Schrödinger equation. Otherwise, however, there is no inconsistency in measurement, as there is no theory. This assumption is referred to as that of "universal validity of quantum theory." It leads—as is demonstrated below—to some unusual consequences, but is able to avoid the discrepancies of quantum theory.

The nonexistence of the microscopic states of macroscopic subsystems of the universe leads to severe difficulties in the interpretation of observation or measurement in terms of information transfer between systems. In particular, since no microscopic state of an organism exists, the principle of "psychophysical parallelism"[4] does not apply.

In order to understand Eq. (3), the meaning of superpositions of macroscopically different states has to be investigated. Consider, for the moment, a right-handed sugar molecule with wave function φ_R. This is different from an eigenstate of its Hamiltonian H_S, $\varphi_R \pm \varphi_L$. In contrast to the analogous situation for an ammonia molecule, the tunneling time from φ_R to φ_L is much larger than the age of the universe. The interaction matrix element $\langle \varphi_R \mid H_S \mid \varphi_L \rangle$ is extremely small, as H_S can at most change the state of two particles. Assume now that an eigenstate $\varphi_R \pm \varphi_L$ had been prepared. The two components would then interact in different ways with their environment,

$$e^{iHt}\phi(\varphi_R \pm \varphi_L) \approx \phi^{(R)}(t) \varphi_R \pm \phi^{(L)}(t) \varphi_L \equiv \psi^{(R)}(t) \pm \psi^{(L)}(t) \qquad (9)$$

(Destruction of the sugar molecule is neglected, and excitations may be taken into ϕ.) With respect to the parity quantum number, the sugar molecule behaves like a macroscopic object—the energy difference between the eigenstates is extremely small. The two world components $\psi^{(R)}$ and $\psi^{(L)}$ will behave practically independently after they have been prepared, since $\langle \psi^{(R)} \mid H \mid \psi^{(L)} \rangle$ becomes even smaller with increasing time. There are no transitions between them any more. The "handedness" of the sugar is dynamically stable, whereas one component of the oriented ammonia molecule would emit a photon,

Such a dynamical decoupling of components is even more extreme if φ_R and φ_L represent two states of a pointer corresponding to different positions. Each state will now produce macroscopically correlated states: different images on the retina, different events in the brain, and different reactions of the observer. The different components represent two completely decoupled worlds. This decoupling describes exactly the "reduction of the wave function." As the "other" component cannot be observed any more, it serves only to save the consistency of quantum theory. Omitting this component is justified pragmatically, but leads to the discrepancies discussed above.

This interpretation, corresponding to a "localization of consciousness" not only in space and time, but also in certain Hilbert-space components, has been suggested by Everett[9] in connection with the quantization of general relativity, and called the "relative state interpretation" of quantum theory. It amounts to a reformulation of the "psychophysical parallelism" which has in any case become necessary as a consequence of the above discussion of dynamical correlations between states of macroscopic systems.[2] A theory of measurement must necessarily be empty if it does not have a substitute for psychophysical parallelism. Everett's relative state interpretation is ambiguous, however, since the dynamical stability conditions[3] are not considered. This ambiguity is present in the orthodox interpretation of quantum theory as well, where it has always been left to intuition which property of a system is measured "automatically" (e.g., handedness for the sugar, but parity for the ammonia molecule). The dynamical stability appears also to be the cause why microscopic oscillators are observed in energy eigenstates, whereas macroscopic ones occur in "coherent states."[5]

According to the twofold localization of consciousness, there are two kinds of subjectivity: The result of a measurement is subjective in that it depends on the world component of the observer; it is objective in the sense that all observers of this world component observe the same result. The question of whether the other components still "exist" after the measurement is as meaningless as asking about the existence of an object while it is not being observed. It *is* meaningful, however, to ask whether or not the assumption of this existence (i.e., of an objective world) leads to a contradiction.

The probability postulate of quantum theory can be formulated in the following way: Suppose a sequence of equivalent measurements have been performed, each creating an equivalent "branching of the universe." The observer can explain the results by assuming that his final branch has been "chosen randomly" if the components are weighted by their norm. The irreversibility connected with this branching is different from that due to thermodynamical statistics, and thus cannot be explained in terms of the latter. Instead, the effect of branching, i.e., measurement, should be of importance for the foundation of thermodynamics. It seems to be partly taken into account by using the density matrix formalism.[4]

---

[2] Another suggestion of Wigner's,[10] which postulates an active role of consciousness, would require corrections to the equations of motion.

[3] The importance of stability for organic systems has been emphazised by Elsasser.[11]

[4] This may indeed be the reason why the foundation of quantum-mechanical thermodynamics appears simpler than that of classical thermodynamics. Proofs of the master equation would, however, be circular again if the process of measurement and hence the density matrix formalism were themselves based on thermodynamics.

The famous paradox of Einstein, Rosen, and Podolski[12] is solved straightforwardly: A particle of vanishing spin is assumed to decay into two spin-$\frac{1}{2}$ particles. As a consequence, and according to the axiom of measurement, each particle possesses spin projections of equal probability with respect to any direction in space. After measuring the spin of one particle, however, the spin of the other one is determined. According to Einstein *et al.*, this cannot be true if quantum theory is complete, as there is no interaction with the second particle. The interpretation is that the measurement corresponds to the transformation

$$e^{iHt}\phi(\varphi_1^+\varphi_2^- - \varphi_1^-\varphi_2^+) = \varphi_1^+\varphi_2^-\phi^{(+)}(t) - \varphi_1^-\varphi_2^+\phi^{(-)}(t) \tag{10}$$

where $\phi^{(+)}$ and $\phi^{(-)}$ are dynamically decoupled after a short time. Hence, there is one world component in which the experimentalists observe φ_1^+ and φ_2^-, another one in which they observe φ_1^- and φ_2^+. As these components cannot "communicate," the result is in accord with the axiom of measurement.

This interpretation of measurement may also explain certain "superselection rules"[13] which state, for example, that superpositions of states with different charge cannot occur. It is very plausible that any measurement performed with such a system must necessarily also be a measurement of the charge. Superpositions of states with different charge therefore cannot be observed for similar reasons as those valid for superpositions of macroscopically different states: They cannot be dynamically stable because of the significantly different interaction of their components with their environment, in analogy to the different handedness components of a sugar molecule.

If experimental evidence verifies a spontaneous symmetry-breaking of the vacuum as predicted by many field theories[14] this would not prove an asymmetry of the world. One may formally construct invariant wave functions $\Psi = \int d\Omega \, U_\Omega \phi$ from symmetry-violating wave functions ϕ (as done for microscopic systems[15]). The former cannot be distinguished from its components $U_\Omega \phi$ if the relative state interpretation is accepted.

It appears that the objective interpretation of quantum theory does not contradict the probability interpretation. It has to be admitted, however, that the "relative state wave function" describes only part of the universe. There is no information on other components except for those which have been created by branching in the past. No estimate can therefore be made on the probability of an inverse branching process, i.e., the spontaneous occurrence of components by accidental overlap.

ACKNOWLEDGMENT

REFERENCES

1. E. P. Wigner, *Am. J. Phys.* **31**, 6 (1963); B. d'Espagnat, *Nuovo Cimento (Suppl.)* **4**, 828 (1966); T. Earman and A. Shimony, *Nuovo Cimento* **54B**, 332 (1968); J. M. Jauch, E. P. Wigner, and

M. M. Yanase, *Nuovo Cimento* **48B**, 144 (1967); G. Ludwig, in *Werner Heisenberg und die Physik unserer Zeit* (Braunschweig, 1961).

2. G. Ludwig, *Die Grundlagen der Quantenmechanik* (Berlin, 1954), p. 122 ff.; *Z. Physik* **135**, 483 (1953); A. Danieri, A. Loinger, and G. M. Prosperi, *Nucl. Phys.* **33**, 297 (1962); *Nuovo Cimento* **44B**, 119 (1966); L. Rosenfeld, *Progr. Theoret. Phys. (Suppl.)* p. 222 (1965); W. Weidlich, *Z. Physik* **205**, 199 (1967).

3. J. S. Bell, *Rev. Mod. Phys.* **38**, 447 (1966); D. Bohm and J. Bub, *Rev. Mod. Phys.* **38**, 453 (1966).

4. J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932) [English translation: Mathematical Foundations of Quantum Mechanics (Princeton University Press, Princeton, N. J., 1955)].

5. R. J. Glauber, *Phys. Rev.* **131**, 2766 (1963); P. Caruthers and M. M. Nieto, *Am. J. Phys.* **33**, 537 (1965); B. Jancovici and D. Schiff, *Nucl. Phys.* **58**, 678 (1964); C. L. Mehta and E. C. G. Sudarshan, *Phys. Rev.* **138**, B274 (1965).

6. E. Schrödinger, *Z. Physik* **14**, 664 (1926).

7. D. Bohm, *Quantum Theory* (Prentice-Hall, Englewood Cliffs, N.J., 1951).

8. J. M. Jauch, *Helv. Phys. Acta* **33**, 711 (1960).

9. H. Everett, *Rev. Mod. Phys.* **29**, 454 (1957); J. A. Wheeler, *Rev. Mod. Phys.* **29**, 463 (1957).

10. E. P. Wigner, in *The Scientist Speculates*, L. J. Good, ed. (Heinemann, London, 1962), p. 284.

11. W. M. Elsasser, *The Physical Foundation of Biology* (Pergamon Press, New York and London, 1958).

12. A. Einstein, N. Rosen, and B. Podolski, *Phys. Rev.* **47**, 777 (1935); D. Bohm and Y. Aharonov, *Phys. Rev.* **108**, 1070 (1957).

13. G. C. Wick, A. S. Wightman, and E. P. Wigner, *Phys. Rev.* **88**, 101 (1952); E. P. Wigner and M. M. Yanase, *Proc. Natl. Acad. Sci. (US)* **49**, 910 (1963).

14. W. Heisenberg, *Rev. Mod. Phys.* **29**, 269 (1957); Y. Nambu and G. Jona-Lasinio, *Phys. Rev.* **122**, 345 (1961).

15. H. D. Zeh, *Z. Physik* **202**, 38 (1967).