PHYSICAL REVIEW D VOLUME 22, NUMBER 4 15 AUGUST 1980

Can we undo quantum measurements?

Asher Peres*
Center for Theoretical Physics, The University of Texas at Austin, Austin, Texas 78712
(Received 1 October 1979; revised manuscript received 14 April 1980)

The Schrödinger equation cannot convert a pure state into a mixture (just as Newton's equations cannot display irreversibility). However, to observe phase relationships between macroscopically distinguishable states, one has to measure very peculiar operators. An example, constructed explicitly, shows that the *classical* analog of such an operator cannot be measured, because to do so would violate classical irreversibility. This result justifies von Neumann's measurement theory, without any hypothesis on the role of the observer.

The measurement process in quantum physics was analyzed long ago by von Neumann[1] who showed that it could formally be described as the transformation of a pure state $\Psi=\sum c_n\phi_n$ into a mixture $\rho=\sum |c_n|^2 P_n$. Here, the ϕ_n are eigenstates of the dynamical variable being measured, and the P_n are the corresponding projection operators.

This irreversible transformation, commonly called the "collapse of the wave packet," cannot follow from the Schrödinger equation, since the latter generates a unitary mapping of the Hilbert space of states. In fact, the coupling of the eigenstates of the measured system to those of the measuring apparatus is a perfectly reversible process[2-4] as long as we are willing to measure correlations between the two. For these reasons, von Neumann's theory has been considered unsatisfactory, or at least incomplete.

There have been several attempts[5-7] to prove von Neumann's conjecture by supplementing quantum theory with superselection rules forbidding the measurement of operators of a certain type (those which connect macroscopically different states of the apparatus). The purpose of this paper is to show that systems with many degrees of freedom are indeed subject to such superselection rules. A general proof of this assertion would be very difficult, but the following model is typical enough to convey belief in the result.

Consider a macroscopic apparatus designed to measure the z component of the spin of an electron. This apparatus has a pointer (center-of-mass coordinate q, conjugate momentum p) initially localized around $q=0$. The pointer is to move through a macroscopic distance L to the right or the left depending on whether $s_z=\frac{1}{2}$ or $-\frac{1}{2}$. This can be achieved by the coupling $H=2V(t)s_z p$, where $V(t)$ is a large velocity, so large indeed that we can neglect all the other terms in the Hamiltonian during the brief duration of the coupling.[8]

Before the measurement, the state of the electron is $\binom{\alpha}{\beta}$ and that of the apparatus is $\Psi(q,q_2,q_3,\ldots,q_N)$ where $q_2,q_3,\ldots,q_N$ are the other, "irrelevant," degrees of freedom. Naturally, N is a very large number, say 10^{23}. (It would be more realistic to assume a density matrix instead of the pure state ψ, but this refinement is not needed at the present stage.)

After completion of the coupling, the combined state is

$$\begin{pmatrix}\alpha\\0\end{pmatrix}e^{-iLp}\psi+\begin{pmatrix}0\\\beta\end{pmatrix}e^{iLp}\psi\,, \qquad (1)$$

where $L=\int V\,dt$. Since ψ is peaked around $q=0$, $e^{\pm iLp}\psi$ is peaked around $q=\mp L$. Thus, the sign of q is correlated to that of s_z and

$$\langle s_z\rangle=\langle\tfrac{1}{2}\,\mathrm{sign}(q)\rangle=\tfrac{1}{2}(|\alpha|^2-|\beta|^2)\,. \qquad (2)$$

We have performed what von Neumann calls a measurement. (As we shall soon see, a better word would be "premeasurement.")

The question is whether this process is reversible and, in particular, whether the relative phase α/β is still observable. At the present stage, it is, as can be seen by measuring the expectation values of the operators,

$$A_1=s_x\cos 2Lp+s_y\sin 2Lp \qquad (3a)$$

and

$$A_2=s_x\sin 2Lp-s_y\cos 2Lp\,. \qquad (3b)$$

[To measure A_1 and A_2 we divide the electrons in two identical but disjoint ensembles. After each electron passage through the apparatus, we first measure p (modulo π/L) then the component of $\vec{s}$ in the direction of $\tan^{-1}(2Lp)$ or $\cot^{-1}(-2Lp)$. Note that the eigenvalues of A_1 and A_2 are $\pm\frac{1}{2}$.]

These operators can conveniently be combined as

$$A=A_1+iA_2=\begin{pmatrix}0&0\\1&0\end{pmatrix}e^{2iLp}\,. \qquad (4)$$

A straightforward calculation yields $\langle A\rangle=\alpha\beta^*$

which, together with Eq. (2), gives α and β separately, up to a common phase.

However, if we wait some time, the state (1) will evolve into

$$e^{-itH}\left[\begin{pmatrix}\alpha\\0\end{pmatrix}e^{-iLp}\psi+\begin{pmatrix}0\\\beta\end{pmatrix}e^{iLp}\psi\right], \tag{5}$$

where H is the Hamiltonian of the electron and apparatus. Assuming for simplicity that the two spin states have the same energy, we obtain

$$\langle A\rangle=\alpha\beta^*\int(e^{-itH}e^{iLp}\psi)^*e^{2iLp}e^{-itH}e^{-iLp}\psi d^Nq\,. \tag{6}$$

But $e^{iLp}e^{-itH}e^{-iLp}$ is simply $e^{-itH(q+L)}$ (i.e., H with the q coordinate shifted by L) and (6) can be written as

$$\langle A\rangle=\alpha\beta^*\langle e^{itH(q-L)}e^{-itH(q+L)}\rangle\,. \tag{7}$$

The coefficient of $\alpha\beta^*$ would still be one if H did not depend on q, but there is no reason to expect that. As the pointer moves with respect to some fixed scale on the apparatus, its energy may vary somewhat from place to place and the coefficient of $\alpha\beta^*$ may be less than one in absolute value. For small t we get

$$|\langle e^{itH(q-L)}e^{-itH(q+L)}\rangle|^2=1-t^2\delta H^2+\cdots, \tag{8a}$$

where

$$\delta H^2=\langle[H(q-L)-H(q+L)-\langle H(q-L)-H(q+L)\rangle]^2\rangle\,. \tag{8b}$$

Moreover, if the other degrees of freedom of the apparatus are in a mixed state, this coefficient will quickly fall to zero,[9] because of the randomness of the phases. The time needed to erase $\langle A\rangle$ is of the order of $1/\delta H$. It is therefore inversely proportional to the strength of the coupling of the macroscopic degree of freedom q, used for the measurement, with the *other* degrees of freedom of the apparatus. In the present model, this time could be of the order of the size of the pointer divided by the speed of sound (a few microseconds).

This neat distinction between the reversible premeasurement—Eq. (1)—and the ensuing irreversible process is admittedly unrealistic in most instances. In practice, a macroscopic apparatus has almost always an amplification mechanism based on a metastable initial state[10] and irreversibility appears at the very outset of the process. However, the amplification requirement is not essential and it obscures the true nature of the irreversibility of quantum measurements, which is explained below. (The reversal of an idealized premeasurement is illustrated in Fig. 1.)

The above discussion of Eq. (7)—or some similar argument[10,11]—is usually considered as a proof that the relative phase of the two branches of Eq. (5) is "lost" after some finite time. However, such arguments are not convincing, because Eq. (5) represents a pure state (what else could it be?) and this can be shown by measuring the expectation value of *another* operator, namely,

$$A'=e^{-itH}Ae^{itH}\,. \tag{9}$$

Indeed, we trivially have $\langle A'\rangle=\alpha\beta^*$, since the $e^{\pm itH}$ factors in A' cancel those of the wave functions.

However, the operator A' has very peculiar properties. (It is not of course the Heisenberg picture of A, the latter being $e^{itH}Ae^{-itH}$. In fact, we are always working in the Schrödinger picture.) This A' operator is *explicitly time dependent and is also a constant of the motion.*

To verify that it is a constant of the motion, it is enough to observe that its matrix elements between any two Schrödinger states are constant, or simply to go to the Heisenberg picture, where A'_H looks like A_S, without any time dependence.

Now, these explicitly time-dependent constants of the motion are very familiar in classical mechanics. For example, for a free particle, $q-tp/m$ is such a constant. Its meaning simply is the initial value of q. For an harmonic oscillator, such a constant would be $\tan^{-1}(m\omega q/p)-\omega t$, the meaning of which is the initial value of the phase. In general, for a system with N degrees of freedom, there are $2N$ constants of the motion, a few of which may be explicitly time independent (the total energy, momentum, etc.), but almost all

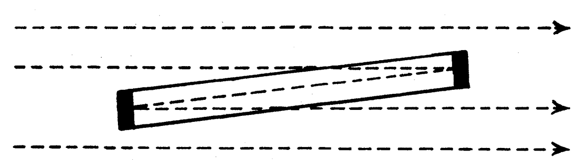

FIG. 1. Idealized premeasurement, using the recoil of a rigid double mirror. If a particle is reflected from the first mirror, a correlation is established between the momentum of the particle and that of the instrument (this is the premeasurement). That correlation is then reversibly destroyed when the particle is reflected from the second mirror. (Note that if we wish to complete the measurement and to observe the recoil of the double mirror *between* the two reflections, the latter must be prepared with $\Delta p\ll h/\lambda$. If this device is part of a double-slit experiment, it allows to determine through which slit the particle passed only at the expense of destroying the interference pattern, because $\Delta q\gg\lambda$. But if we forego observing the recoil, the interference pattern is restored because the *same* Δq is added and subtracted at both reflections.)

of which include the time explicitly. Their physical meaning is to give the $2N$ initial positions and momenta as explicit functions of the positions and momenta at some future time t. The structure of these constants of the motion is of course hopelessly complicated for large N and finite t. It leaves us no choice but to replace Newtonian mechanics by statistical mechanics. It is our inability to make use of these constants of the motion which is the cause of irreversibility.

In the present case, we must measure, instead of A given by Eq. (4),

$$A' = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} e^{2iLp'},$$

where $p' = e^{-itH} p e^{itH}$ is the value of p immediately after the premeasurement, expressed as a function of p and $2N-1$ other variables at a later time t. In classical physics, we would say that this is so complicated that only a "Maxwell demon" can measure all these variables and then compute p' (assuming H is known). In quantum physics, the task is even more difficult because the $2N$ variables do not commute. Therefore, the Maxwell demon must contrive a *single* measurement[12] for p', which is an incredibly complicated function of $2N$ noncommuting variables and of t.

In other words, we see that not every self-adjoint operator corresponds to an observable, simply because not every classical dynamical variable is observable. It is the inobservability of these operators which makes pure states appear as mixtures and causes the irreversibility of quantum measurements.

In conclusion, let us summarize the assumptions used in the derivation of this result. First, we note that the macroscopic degree of freedom used for the measurement—here, the center of mass of the pointer—is not completely isolated from the other degrees of freedom of the apparatus.[13] (We could, of course, have treated these other degrees of freedom as an external reservoir, but then our result would have been trivial. It is essential that our system be a *closed* one.)

The second assumption is the impossibility of measuring the classical analog of the operator p'. (There is also a tacit assumption that if a classical measurement is impossible, the same is true for the corresponding quantum measurement.) Here, it may be objected that as long as the number of degrees of freedom is finite, it is not impossible to measure p', only very difficult. In principle, a measurement of p' should always be possible at the cost of a great increase of entropy of the rest of the world.[14] From this point of view, as long as we are able to pay the price,[15] we definitely *can* undo quantum measurements,[4] except in the unattainable mathematical limit of an infinite apparatus.[16] However, if we admit that a finite system may appear irreversible (if the time needed for a Poincaré recurrence is longer than the Universe lifetime), the present paper shows how the irreversibility of quantum measurements is rooted in the familiar classical irreversibility.

I am very grateful to J. A. Wheeler for the warm hospitality of the University of Texas and to E. C. G. Sudarshan for many stimulating discussions. The final version of this paper has benefited from comments by J. S. Bell (CERN) and A. Ron (Technion). This work was supported in part by the Center for Theoretical Physics, The University of Texas at Austin, Austin, Texas 78712, and also by NSF Grant No. PHY78-26592.

APPENDIX: A MORE REALISTIC MODEL

The simple model discussed above involves an explicitly time-dependent coupling $V(t)$, supposedly switched on and off by an external agent. This may give the impression that we are dealing with an open (i.e., incompletely described) system, for which the transformation of a pure state into a mixture would be trivial.

In a real-life Stern-Gerlach experiment, this time dependence is of course due to the translational degree of freedom of the electron, which was arbitrarily ignored in our model. A more realistic description of what happens follows.

We write the complete Hamiltonian as

$$H = H_a + H_e + 2Vs_z p u(x_2 - x)u(x - x_1), \qquad \text{(A1)}$$

where H_a refers to the apparatus, H_e to the free electron (mass m, momentum k, position x), V is a coupling *constant*, u is the unit step function, and x_1 and x_2 are the entrance and exit points of the electron as it passes through the apparatus. The pointer is assumed massive enough so that its velocity p/M is negligible when the electron is outside the apparatus. When it is inside, the pointer velocity is $\pm V$.

The measurement process can be described as a scattering of the electron and the apparatus. Before the "collision," the electron has momentum k. When it reaches the apparatus, it meets an energy barrier of height $\pm Vp$ and thickness $x_2 - x_1$. Inside the barrier, its momentum is $k' = (k^2 \pm 2mVp)^{1/2} \simeq k \pm mVp/k$, where we have assumed that $k^2 \gg 2mVp$, so that most electrons are transmitted (a reflected electron would mean an unsuccessful experiment). The outgoing electron still has momentum k, but has been subject to a phase shift $(k'-k)(x_2 - x_1) = \pm \tau Vp$ where $\tau = m(x_2 - x_1)/k$ is the classical time of passage through the apparatus. We now identify $L = \tau V$ and the final state of the combined system is

$$e^{ikx}\left[\begin{pmatrix}\alpha\\0\end{pmatrix}e^{-iLp}\psi+\begin{pmatrix}0\\\beta\end{pmatrix}e^{iLp}\psi\right], \tag{A2}$$

whence the discussion proceeds as before.

However, several remarks are in order. First, we have treated p as a constant during the collision, i.e., we assumed that $\dot{p}=\partial H_a/\partial q=0$. This is of course incompatible with a nontrivial δH [see Eq. (8b)]. However, we can make the change in p arbitrarily small by increasing V and k (keeping their ratio constant, so that L remains unchanged). This does not affect $\dot{p}$, but makes τ arbitrarily small. The condition is easily seen to be $\tau \ll 1/\delta H$, i.e., the premeasurement must be very brief, compared to the time required to make the measurement irreversible.

To avoid a possible misunderstanding, it must be emphasized that δH is *much smaller* than the energy uncertainty $\Delta H=(\langle H^2\rangle-\langle H\rangle^2)^{1/2}$. Indeed there must be *many* different energy eigenstates involved to make the measurement possible.[17,18] In particular, the incoming electron must have $\Delta H \gg \delta H$ because the two branches of the outgoing electron will not interfere if $\delta H > \Delta H$ (that is, we would need an operator much more complicated than A to display their interference[19]).

The above remark is closely related to overall energy conservation. We have assumed hitherto that the outgoing electron had the same energy as the incoming one. This cannot, of course, be rigorously exact if $H(q-L)\neq H(q+L)$. A more correct treatment follows.

First, assume that initially the apparatus is in an eigenstate of energy E_0 and that the electron too is monochromatic with energy $K=k^2/2m$. Then obviously there is no irreversibility since the operator e^{-itH} in Eq. (A2) becomes a phase factor $\exp[-it(E_0+K)]$ and $\langle A\rangle$ is constant. The energy picked up or released by the electron exactly compensates the energy difference in the final state of the apparatus. It is thus important to understand why $|\langle A\rangle|$ may decrease if we have a superposition (or mixture), rather than an energy eigenstate.

Even if $|E_0\rangle$ is an eigenstate of H_a, the states $e^{\pm iLp}|E_0\rangle$ usually are not. We can write

$$e^{\pm iLp}|E_0\rangle=\int c_\pm(E_0,E,K)|E\rangle dE\,,$$

where the coefficients $c_\pm$ depend also on K, because $L=m|x_2-x_1|/k$. By virtue of energy conservation, the scattering process can therefore be written as

$$|E_0,K\rangle\to\int c_\pm(E_0,E,K)|E,K+E_0-E\rangle dE\,,$$

where the $\pm$ subscripts refer to $s_z=\pm\frac{1}{2}$.

Now let the electron be initially in a superposition $\int g(K)|K\rangle dK$ (the apparatus may still be initially in an energy eigenstate[10]). The outgoing states become

$$\int g(K)c_\pm(E_0,E,K)|E,K+E_0-E\rangle dE\,dK\,.$$

In order to compute $\langle A\rangle$, we first note that

$$\langle E',K'+E_0-E'|e^{2iLp}|E,K+E_0-E\rangle =\langle E'|e^{2iLp}|E\rangle\delta(K'-E'-K+E)\,,$$

so that

$$\langle A\rangle=\alpha\beta^*\int g(K)g^*(K')c_+(E_0,E,K)c_-^*(E_0,E',K') \times\langle E'|e^{2iLp}|E\rangle\delta(K'-E'-K+E)dE\,dE'dK\,dK'$$

We now make two essential physical assumptions. One is that $\Delta K\ll K$ (otherwise, L is ill-defined) so that in $c_\pm$ we can replace K by its average value K_0. The second one is that $\delta H\ll\Delta K$, i.e., $c_\pm(E_0,E,K_0)$ is very small unless $|E-E_0|\ll\Delta K$ (as explained above, $|K_{\rm out}-K_{\rm in}|$ must be much smaller than ΔK to allow the two "branches" of the electron to interfere). We can therefore replace $g^*(K+E'-E)$ by $g^*(K)$. Integration over K and K' gives

$$\langle A\rangle=\alpha\beta^*\int c_+(E_0,E,K_0)c_-^*(E_0,E',K_0) \times\langle E'|e^{2iLp}|E\rangle dE\,dE'\,.$$

We see that the electron energy no longer appears in the formula (except as an average). The result looks *as if* the final state of the apparatus were $\int c_\pm(E_0,E,K_0)|E\rangle dE$. Therefore, the energy shift of the electron cannot prevent $\langle A\rangle$ from having a nontrivial time dependence, due to the factor $e^{-i(E-E')t}$ in $\langle E'|e^{2iLp}|E\rangle$.

---

*On sabbatical leave from Technion-Israel Institute of Technology, Haifa.

[1]J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955).
[2]P. A. Moldauer, Phys. Rev. D 5, 1078 (1972).
[3]A. Peres, Am. J. Phys. 42, 886 (1974).
[4]J. S. Bell, Helv. Phys. Acta 48, 93 (1975).
[5]H. Wakita, Prog. Theor. Phys. 23, 32 (1960).
[6]A. Komar, Phys. Rev. 133, B542 (1964).
[7]E. C. G. Sudarshan, Pramana 6, 117 (1976); T. N. Sherry and E. C. G. Sudarshan, Phys. Rev. D 18, 4580 (1978); 20, 857 (1979).
[8]A more realistic model with no explicitly time-depen-

dent coupling is discussed in the Appendix.

[9]It is not necessary that $\langle A \rangle$ be "rigorously" zero because it is an *average* value, not an eigenvalue. The eigenvalues are $\pm\frac{1}{2}$ and the probable error on $\langle A \rangle$ after n measurements is $\sim 1/\sqrt{n}$. Now n cannot exceed e^{2000}, say, because of cosmological limitations. Therefore, after 1000 lifetimes $\langle A \rangle$ is indistinguishable from zero *as a matter of principle*, not only as an approximation.

[10]A. Daneri, A. Loinger, and G. M. Prosperi, Nucl. Phys. 33, 297 (1962); Nuovo Cimento 44B, 119 (1966). These authors assume that the initial state of the apparatus is metastable. This is often true in practice but, as shown in the present paper and in particular in the Appendix, this assumption is not necessary.

[11]K. Hepp, Helv. Phys. Acta 45, 237 (1972).

[12]For example, it is not the same thing to measure $s_x + s_y$ (eigenvalues $\pm 1/\sqrt{2}$) or to measure s_x and s_y separately and sum the results.

[13]A macroscopic superconducting current is *not* acceptable as a measuring instrument, unless it is coupled to some monitoring device with numerous degrees of freedom.

[14]Here again, we may encounter cosmological limitations [A. Peres and N. Rosen, Phys. Rev. 135, B1486 (1964)]. If we assume that the Universe is finite, there must be an upper limit to the amount of entropy which may be generated in it. It is then plausible that if N is large, irreversibility sets in after a finite time t. However, I do not wish to enter deeper into this subject: The purpose of this paper was not to explain classical irreversibility, but only to clarify its relationship to the quantum measurement problem.

[15]This is the only place where the human observer has any role. He decides (perhaps subjectively) which experiments are feasible. For a fuller discussion, see A. Peres, Found. Phys. (to be published).

[16]B. Whitten-Wolfe and G. G. Emch, Helv. Phys. Acta 49, 45 (1976).

[17]E. P. Wigner, Z. Phys. 133, 101 (1952).

[18]H. Araki and M. M. Yanase, Phys. Rev. 120, 622 (1960); M. M. Yanase, *ibid.* 123, 666 (1961).

[19]If the apparatus or part of it is in thermal equilibrium, we have $\Delta H \simeq kT$ and in most realistic situations $\tau\Delta H \gg 1$. This does not hamper the measurement as long as $\tau\delta H \ll 1$.