Properties of a quantum system during the time interval between two measurements

Yakir Aharonov and Lev Vaidman
*Department of Physics, University of South Carolina, Columbia, South Carolina 29208
and School of Physics and Astronomy, Tel-Aviv University, Ramat Aviv 69978, Israel*
(Received 6 February 1989)

A description of quantum systems at the time interval between two successive measurements is presented. Two wave functions, the first preselected by the initial measurement and the second postselected by the final measurement, describe quantum systems at a single time. It is shown how this approach leads to a new concept: a *weak value* of an observable. Weak values represent novel characteristics of quantum systems between two measurements. They are outcomes of a standard measurement procedure that fulfills certain requirements of "weakness." We call it *weak measurement.* Physical meaning, underlying mathematical structure, and prospects of practical usage of weak measurements are explored.

I. INTRODUCTION

Recently we have developed a description of quantum systems during the time interval between two measurements. This description uncovered several new aspects of quantum theory.[1–3]

In our approach we assign to a quantum system at a given time two wave functions (instead of one). In addition to the standard wave function we consider another wave function, evolving from the future toward the past. The two-wave-functions formalism was introduced by Aharonov, Bergman, and Lebowitz[4] in order to simplify the calculation of the probability of finding a given result in a measurement that is performed in an intermediate time between two other measurements.

The most important outcome of our approach is the possibility to define a new concept: the *weak value* of a quantum variable. It is a physical property of a quantum system between two measurements, i.e., a property of a system belonging to an ensemble that is both preselected and postselected. This property can manifest itself through a measurement that fulfills certain requirements of weakness. In fact, the effect of any interaction that is weak enough will depend on such weak values. The weak value of a variable may differ significantly from the eigenvalues of an associated operator. Due to this property weak measurements can be used as a novel amplification scheme.

The plan of this article is as follows. In Sec. II we discuss symmetry under time reversal in quantum theory. In Sec. III we show how the two-wave-functions approach leads to the idea of weak measurements. In Sec. IV we define the concept of a weak value for both preselected and postselected ensemble and prove that the weak value can be obtained in a suitable "weak" measurement. Section V considers an example of a spin-$\frac{1}{2}$ particle. In Sec. VI we consider the "weak value" for regular preselected ensemble and show that it is equal to the mean value of the corresponding operator. In Sec. VII we describe situations in which weak values can be detected in a single experiment. In the end of Sec. VII an explanation of this result is given in the framework of the standard formalism. In Sec. VIII we derive an unexpected mathematical identity underlying our results and we discuss a procedure that seemingly transmits (but actually does not) signals faster than light. A few comments concerning the feasibility of experimental tests of our predictions are given in Sec. IX. The conclusions are drawn and some remarks about directions of the future work are made in Sec. X.

II. TIME REVERSAL

In this section we discuss the description of a quantum system at the time interval between two measurements that is symmetric under time reversal.[4] First, let us discuss time assymmetry of the standard approach. In quantum theory the dynamical laws are time symmetric as are their classic counterparts, namely, Hamilton's equations of motion. The asymmetry enters through the theory of measurements. The "collapse" of a wave function which is a part of the measurement process is not (at least in the standard approach) time symmetric: the wave function existing before the measurement "collapses," in general, to a new wave function in accordance with the result of the measurement. In the standard approach it is not clear how we can restore time-reversal symmetry since there is no state evolving backwards in time. The following example will clarify the difference between the two time directions.

Suppose that we have an ensemble of spin-$\frac{1}{2}$ particles, which were found at time t, in the state $\sigma_x = 1$. We can predict that the probability of finding $\sigma_y = 1$ immediately after is $\frac{1}{2}$. However, we cannot ascertain by "retrodiction" that the probability of finding $\sigma_y = 1$ immediately before the time t is also $\frac{1}{2}$. It may happen that all particles in the ensemble were prepared in the state $\sigma_y = -1$, in which case no particles would be found with $\sigma_y = 1$, or it may be that all particles were prepared with $\sigma_y = 1$ and then they all yield $\sigma_y = 1$ again. We can still ascertain by retrodiction that the probability of finding $\sigma_x=1$ before the time t is 1; however, for the outcomes of measurements of other observables we have to take into account the state of the system before the time t. We assume that such a state existed, and if we do not know what it was, we cannot find the probability for given outcomes of most measurements.

With the measurement after time t this problem does not arise because we do not assume that there is a state (even unknown) coming from the future. So the difference between past and future is not an intrinsic property of quantum theory, but it is the feature of our approach to the arrow of time: at present we view the past as existing and future as nonexisting (yet).[5] However, if our task is a description of a quantum system between two successive measurements, then we know the boundary conditions in the future as well as in the past. (We assume that both measurements are complete.) Therefore for the intermediate time interval we have a complete symmetry under time reversal. The contribution to the description of the quantum system from the result of the initial measurement is the usual wave function evolving from the past toward the future, from the initial measurement to the final measurement. Because of the symmetry under time reversal, the contribution of the final measurement should be similar: the wave function evolving backwards in time from the final measurement to the initial measurement. This is our proposal: to describe a quantum system between two measurements by two wave functions evolving in opposite directions of time.

Let us consider a quantum system between the measurements of two variables A and B. At time t_1 an observable A was measured and a nondegenerate eigenvalue a was found, and at time t_2 B was measured and nondegenerate eigenvalue b was found. At the intermediate time t the system is described by the following two wave functions: a bra $|\Psi_1\rangle$ (wave function evolving toward the future) and a ket $\langle\Psi_2|$ (wave function evolving toward the past):

$$\begin{aligned} |\Psi_1\rangle &= \exp\left[-i\int_{t_1}^{t} H\,d\tau\right]|A=a\rangle\ , \\ \langle\Psi_2| &= \langle B=b|\exp\left[-i\int_{t}^{t_2} H\,d\tau\right]\ . \end{aligned} \tag{1}$$

These two wave functions $\langle\Psi_2|$ and $|\Psi_1\rangle$ are very useful for calculating probabilities of measurements at the time t. They were used for this purpose by Aharonov, Bergman, and Lebowitz[4] already in 1964. In the next section we show how the two wave functions description leads to an idea of weak measurements which, as it will be shown in Sec. IV, yield a new type of values: the weak values.

III. TWO NONCOMMUTING OBSERVABLES HAVE DEFINITE VALUES AT THE TIME PERIOD BETWEEN TWO MEASUREMENTS

Equation (1) shows how to take into account the time evolution of the wave functions describing a quantum system. Interesting phenomena appear even if the free Hamiltonian of the system is zero, and we shall consider, for simplicity, this case. Now, the description of the quantum system at the time t, $t_1<t<t_2$, by both $|\Psi_1\rangle=|A=a\rangle$ and $\langle\Psi_2|=\langle B=b|$ suggests that both $A=a$ and $B=b$ at that time. Obviously, if A were measured at time t, the result would be $A=a$, and if B were measured instead, the result would be $B=b$. The discussion above may be taken to imply that the value of $C\equiv A+B$ in the intermediate time should be $a+b$. But, for noncommuting variables A and B, the value $a+b$ may differ from any eigenvalue of C and, therefore, the measurement of C cannot yield the value $a+b$. The reason for this discrepancy is that both $A=a$ and $B=b$ are correct at the time t if only *one* of these two measurements were performed. If A and B were measured in between and the measurement of A occurred before the measurement of B, then clearly the results are $A=a$, $B=b$. However, if B were measured before A, then the outcomes of the measurements of A and B yield, in general, different results. If we perform the measurements of A and B simultaneously, then they again disturb each other and, consequently, the result of the measurement of $A+B$ is not $a+b$.

The failure of obtaining both properties $A=a$ and $B=b$ using the measurement of $C=A+B$ is not surprising. The measurement of C changes the situation and, therefore, we cannot anymore associate the wave functions $|\Psi_1\rangle$ and $\langle\Psi_2|$ with our system in the time period $[t_1,t_2]$. This suggests that we use measurements that do not change significantly the two wave functions above. We are thus led to consider a measuring procedure with "weakened" interaction which should yield $A=a$ and $B=b$ even if the measurements are performed in the "wrong" order, namely, B before A. But the same should be correct if the measurements are performed simultaneously and, therefore, the weakened measurement of $C=A+B$ must yield the "forbidden" value $a+b$.

In the next section we shall show that the outcome of such weakened procedure for the measurement of C, which we call *weak measurement,* will indeed be $a+b$. "Weakening" of the interaction will necessarily decrease the accuracy of a single measurement such that it will provide almost no information. In order that such measurements provide meaningful information they have to be performed on an ensemble of identical systems. The ensemble will, however, be of a special kind: any system belonging to this ensemble is both preselected and postselected.

IV. *WEAK VALUES* ARE THE OUTCOMES OF *WEAK MEASUREMENTS*

We begin this section by reviewing briefly the standard von Neumann[6] measuring procedure. The Hamiltonian describing the interaction with a measuring device is

$$H=-g(t)qA\ , \tag{2}$$

where $g(t)$ is a normalized function with compact support near the time of measurement, and q is a canonical variable of the measuring device with conjugate momentum p. After the interaction (2) is over, we can ascertain the value of A from the final value of p

$$A = p_f - p_{\text{in}} \equiv \delta p \ . \tag{3}$$

Any precise measurement of A necessarily disturbs in an uncontrollable manner the values of observables that fail to commute with A. This is due to the fact that a precise measurement of A requires that the value of p be precisely fixed prior to the time of measurement. Consequently, the uncertainty in q during the measurement interaction described in Eq. (2) (and hence the possible *strength* of that interaction as well) is arbitrary large. (This explains how the measurement of C of the example in the Sec. II changes the values of A and B and, consequently, does not yield the outcome $a+b$.)

Our proposal is to modify the von Neumann measuring procedure by weakening the interaction (2). This can be done by preparing an initial state of the measuring device for which the probability of finding a large q is sufficiently small. We shall now prove that such "weak measurement" of A performed on an ensemble of systems, which were preselected in a state $|\Psi_1\rangle$ and were postselected in a state $\langle\Psi_2|$, will yield an outcome which we call a *weak value* of A

$$A_w \equiv \frac{\langle\Psi_2|A|\Psi_1\rangle}{\langle\Psi_2|\Psi_1\rangle} \ . \tag{4}$$

To this end consider an ensemble of systems which are both preselected and postselected. All members of the ensemble are described by the same pair of wave functions $|\Psi_1\rangle$ and $\langle\Psi_2|$. We perform the same measurements on each individual system with a separate measuring device. The interaction Hamiltonians are

$$H_i = -g(t) q_i A_i \ , \tag{5}$$

where the index i refers to the ith system in the ensemble or ith measuring device. For convenience we take the initial state of each measuring device to be the Gaussian

$$\frac{1}{\sqrt{\Delta q}\,(2\pi)^{1/4}} \exp\left[-\frac{q_i^2}{4(\Delta q)^2}\right] .$$

We measure p_i for each measuring device after the interaction. Subsequently we perform the final, postselection measurements on the systems of our ensemble. We then collect the outcomes p_i only of those systems for which the final state turned out be $|\Psi_2\rangle$.

To simplify the following proof we note that changing the time ordering between the p_i measurements and the postselection measurements will not affect their outcomes. Indeed, after the measuring interaction is over, there is no further interaction between the systems of the ensemble and the corresponding measuring devices and, consequently, any action on one system will not affect the results of measurements performed on any of the other systems.

This sequence of events, where we measure p_i only after postselection, is much simpler to analyze. It also corresponds (as will be shown in Sec. IX) to a practical method for performing this type of measurements.

The state of each measuring device that has been postselected is given, up to normalization factor, by the following wave function (we omit the index i referring to each individual system):

$$\begin{aligned}
\langle\Psi_2|\exp\left[-i\int H\,dt\right]|\Psi_1\rangle \exp\left[-\frac{q^2}{4(\Delta q)^2}\right]
&= \langle\Psi_2|\exp(iqA)|\Psi_1\rangle \exp\left[-\frac{q^2}{4(\Delta q)^2}\right] \\
&= \sum_{n=0}^{\infty} \frac{(iq)^n}{n!} \langle\Psi_2|A^n|\Psi_1\rangle \exp\left[-\frac{q^2}{4(\Delta q)^2}\right] \\
&= \langle\Psi_2|\Psi_1\rangle \sum_{n=0}^{\infty} \frac{(iq)^n}{n!} (A^n)_w \exp\left[-\frac{q^2}{4(\Delta q)^2}\right] ,
\end{aligned} \tag{6}$$

where [as defined in Eq. (4)] $(A^n)_w \equiv \langle\Psi_2|A^n|\Psi_1\rangle / \langle\Psi_2|\Psi_1\rangle$. The last expression can be rewritten as the initial wave function of the measuring device multiplied by e^{iqA_w} plus a correction term which is negligible for small Δq

$$\begin{aligned}
\langle\Psi_2|\Psi_1\rangle \sum_{n=0}^{\infty} \frac{(iq)^n}{n!} (A^n)_w \exp\left[-\frac{q^2}{4(\Delta q)^2}\right]
&= \langle\Psi_2|\Psi_1\rangle \exp\left[iq\frac{\langle\Psi_2|A|\Psi_1\rangle}{\langle\Psi_2|\Psi_1\rangle}\right] \exp\left[-\frac{q^2}{4(\Delta q)^2}\right] \\
&\quad + \langle\Psi_2|\Psi_1\rangle \sum_{n=2}^{\infty} \frac{(iq)^n}{n!} [(A^n)_w - (A_w)^n] \\
&\quad \times \exp\left[-\frac{q^2}{4(\Delta q)^2}\right] .
\end{aligned} \tag{7}$$

We are interested in the p representation of the state of the measuring device. By taking Δq such that for all $n \geq 2$

$$(2\Delta q)^n \frac{\Gamma(n/2)}{(n-2)!} |(A^n)_w - (A_w)^n| \ll 1 \ , \tag{8}$$

we can neglect the contribution of the correction in the Fourier transform of (7) and, therefore, the final wave function of the measuring device in the p representation is to a good approximation

$$\exp\left[-(\Delta q)^2 \left[p - \frac{\langle\Psi_2|A|\Psi_1\rangle}{\langle\Psi_2|\Psi_1\rangle}\right]^2\right] . \tag{9}$$

The probability distribution of p is a Gaussian with spread $\Delta p = (2\Delta q)^{-1}$ centered at $p = \mathrm{Re}(A_w)$.

The weak value of A, A_w as defined by (4) may have, also, an imaginary part. This part affects the distribution of the canonical variable q. Indeed, in the q representation the state of the measuring device will turn out to be

$$\exp[iq\,\mathrm{Re}(A_w)] \exp\left[-\frac{[q + 2(\Delta q)^2 \mathrm{Im}(A_w)]^2}{4(\Delta q)^2}\right] . \tag{10}$$

Consequently, the probability distribution of q is a Gaussian with the same spread Δq centered at $q = -2(\Delta q)^2$ $\mathrm{Im}(A_w)$. The uncertainty in p and q will not allow us to deduce $\mathrm{Re}(A_w)$ or $\mathrm{Im}(A_w)$ from a single measurement. However, performing the measurement on an ensemble of

N systems will decrease the uncertainty of the outcome by a factor $1/\sqrt{N}$. Therefore, by taking N large enough $[(2\Delta q\sqrt{N})^{-1} \ll \mathrm{Re}(A_w), \mathrm{Im}(A_w)]$, we can measure the complex value of A_w with any desired precision.

The requirement (8) ensures that the outcome of the measurement is A_w defined by Eq. (4). In particular, if the initial or the final state is an eigenstate of A, then (8) is automatically satisfied. This must be so because the weak value is, in this particular case, also the usual "strong" value of A.

One can argue that a weak value is obtained after some mathematical manipulation on an ensemble and does not have a physical meaning. To emphasize the "reality" of a weak value we note that after the interactions (5) of a physical ensemble of identical systems with an ensemble of measuring devices (but before observing the measuring devices) there is a physical variable of the measuring devices that reflects the weak value of the measured variables. Indeed the observable $\sum p_i/N$ has a mean value which is equal to A_w, while its uncertainty $(2\Delta q\sqrt{N})^{-1}$ is negligible when the number of members in the ensemble is large.

The properties of weak values (4) imply that if $C=A+B$ then $C_w=A_w+B_w$. In the example of Sec. III we considered the weak measurement of $C\equiv A+B$ between the measurements of A and B. Since $|\Psi_1\rangle=|A=a\rangle$, $\langle\Psi_2|=\langle B=b|$, we have $A_w=a$, $B_w=b$ and, therefore, the weak value of C is indeed equal to $a+b$. The fundamental property of weak measurements, namely, that the disturbance caused by them may be neglected, manifests itself in this example. Since the interaction Hamiltonian of the weak measurement is sufficiently small, the weak measurement of B causes a negligible change in A and vice versa. Thus the results of the weak measurements of A and B remain unaltered even if we perform them simultaneously which is, in effect, the weak measurement of C.

V. WEAK VALUES OF SPIN COMPONENTS OF A SPIN-$\frac{1}{2}$ PARTICLE

The weak value of A can differ widely from any eigenvalue of A. In particular, the real part $\mathrm{Re}(A_w)$ can be much bigger (smaller) than the maximum (minimum) eigenvalue of A. To illustrate how weak measurements can yield forbidden values let us consider the following example. Let $A=\sigma_x$, $B=\sigma_\xi$ be the operators corresponding to the components of a spin-$\frac{1}{2}$ particle along the $\hat{\mathbf{x}}$ and $\hat{\boldsymbol{\xi}}$ unit vectors, and let α be the angle between $\hat{\mathbf{x}}$ any $\hat{\boldsymbol{\xi}}$ (see Fig. 1). Let the free Hamiltonian of the system be zero. The operator $C\equiv A+B$ is then proportional to σ_θ where $\hat{\boldsymbol{\theta}}$ is the unit vector bisecting the angle α

$$C=\sigma_x+\sigma_\xi=2\cos(\alpha/2)\sigma_\theta \ . \tag{11}$$

If a particle were initially in a state $|\sigma_x=1\rangle$ and were found in the end with $\sigma_\xi=1$, then the weak values of σ_x and σ_ξ at all intermediate times will also be 1. Combining the above results with Eq. (11), we find the surprising result that the weak value of σ_θ is

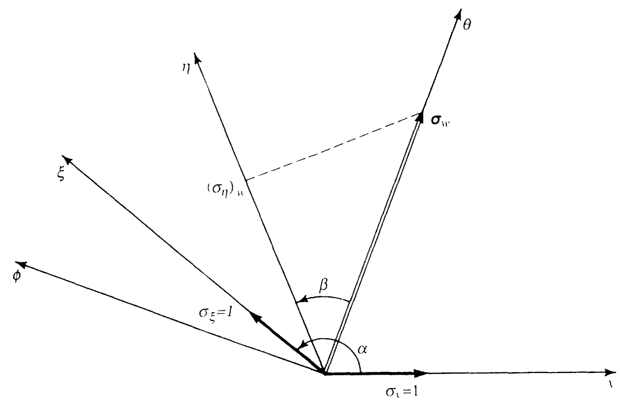

FIG. 1. Geometry of weak values of spin components of a spin-$\frac{1}{2}$ particle in the xy plane. The particle was prepared in the initial state $\sigma_x=1$ and was found in the final state $\sigma_\xi=1$. At the time between these two measurements the weak values of spin components in the xy plane are described by the weak vector $\boldsymbol{\sigma}_w$. Its size is $1/\cos(\alpha/2)$ and it points in the direction $\hat{\boldsymbol{\theta}}$. The weak value of a spin component in arbitrary direction $\hat{\boldsymbol{\eta}}$ is equal to the projection of the weak vector $\boldsymbol{\sigma}_w$ on this direction: $(\sigma_\eta)_w=\cos(\beta)/\cos(\alpha/2)$.

$$(\sigma_\theta)_w=\frac{(\sigma_x)_w+(\sigma_\xi)_w}{2\cos(\alpha/2)}=\frac{1}{\cos(\alpha/2)} \ . \tag{12}$$

The following simple geometrical picture illustrates how one may find the weak value of the spin in any direction given the above boundary conditions $\sigma_x=1$, $\sigma_\xi=1$ (see Fig. 1). We draw in the xy plane a vector of the size $1/\cos(\alpha/2)$ pointing in the direction $\hat{\boldsymbol{\theta}}$. Its projection on any axis in the xy plane yields the weak value of this component. This result can be derived easily from the property of "linearity" of weak value (i.e., $C=A+B$ implies $C_w=A_w+B_w$). The spin component in any direction $\hat{\boldsymbol{\eta}}$ in the xy plane can be decomposed along the orthogonal directions $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$ where

$$\hat{\boldsymbol{\theta}}\equiv\frac{\hat{\boldsymbol{\xi}}+\hat{\mathbf{x}}}{2\cos(\alpha/2)}, \quad \hat{\boldsymbol{\phi}}\equiv\frac{\hat{\boldsymbol{\xi}}-\hat{\mathbf{x}}}{2\cos(\alpha/2)} \ . \tag{13}$$

The decomposition is

$$\sigma_\eta=\sigma_\theta\cos(\beta)+\sigma_\phi\sin(\beta) \ . \tag{14}$$

In the case presently discussed $(\sigma_x=1, \sigma_\xi=1)$, however,

$$(\sigma_\phi)_w=\frac{(\sigma_\xi)_w-(\sigma_x)_w}{2\cos(\alpha/2)}=0 \tag{15}$$

and therefore

$$(\sigma_\eta)_w=(\sigma_\theta)_w\cos(\beta) \ . \tag{16}$$

The weak value of the component of spin in the the z direction turns out to be imaginary:

$$(\sigma_z)_w=\frac{\langle\sigma_\xi=1|\sigma_z|\sigma_x=1\rangle}{\langle\sigma_\xi=1|\sigma_x=1\rangle}=i\tan(\alpha/2) \ , \tag{17}$$

so the "weak vector" $\boldsymbol{\sigma}_w$ in three dimensions (the axes are $\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\phi}},\hat{\mathbf{z}}$) will be

$$\sigma_w = \left[\frac{1}{\cos(\alpha/2)}, \ 0, \ i \tan(\alpha/2) \right] . \tag{18}$$

The weak vector, apart from being complex valued, behaves similarly to a classical vector. This is the consequence of the linearity property of weak values.

VI. MEAN VALUE OF AN OPERATOR AS AN EXAMPLE OF A WEAK VALUE

Many real experiments performed on ensembles of identical systems fulfill the requirement of "weakness," i.e., each individual system practically does not change its state during the measurement. However, the forbidden outcomes have not been observed so far because the standard experiments are performed on ensembles of systems that are preselected only. As an example of such a weak measurement we may consider a compass needle near an ideal ferromagnet (consisting of one domain). The needle will point out the direction of the total spin of the above ensemble of identical systems, while its interaction with each individual spin of the ferromagnet is negligibly small.

We shall prove now that weak measurements of the type described above yield the *mean values* of the observed variables provided that only preselection measurements were performed. The weak value of Eq. (4) is not defined for an ensemble of systems that have been preselected only. We need a result of a future measurement as well. However, since weak measurements hardly disturb the initial wave function, we can predict, with probability arbitrarily close to 1 the result of a future verification experiment of the known (initial) state, even in the event that such weak measurements were performed during the intermediate period. Since future actions cannot affect results of experiments already performed, whether or not the verification measurement in the future in fact takes place will not change the outcome of the weak measurement at the present time. When no verification measurement is performed in the future, the above is the normal procedure for a measurement of a mean value of the ensemble. With the verification measurement added, the above procedure describes a proper measurement of weak value for the special situation when the two wave functions are the same. From Eq. (4) we see indeed that the weak value between two identical states is the same as the mean value for one of them

$$A_w = \frac{\langle \Psi_1 | A | \Psi_1 \rangle}{\langle \Psi_1 | \Psi_1 \rangle} = \langle \Psi_1 | A | \Psi_1 \rangle = \overline{A} . \tag{19}$$

The above measurement, with the compass needle serving as a measuring device, is indeed a weak measurement in the sense that it is almost a nondisturbing experiment, but it differs from the weak measurements described in previous sections. One distinction is that the systems in the ensemble considered here are not postselected. But there is also another difference: while in the previous cases separated measuring devices have been used for different members of the ensemble, here a *single* measuring device is used for all spins. The experiment with a compass needle illustrates the general situation where *one* measuring device measures an average property of an ensemble. We shall show now that in such a case the state of the ensemble is nearly an eigenstate of the "average" operator $\sum_{i=1}^{N} A_i / N$.

To prove the above statement we shall start with the following simple formula:[7]

$$A|\Psi\rangle = \overline{A}|\Psi\rangle + \Delta A |\Psi_\perp\rangle , \tag{20}$$

where $\Delta A \equiv [(\overline{A^2}) - (\overline{A})^2]^{1/2}$, while $|\Psi_\perp\rangle$ is a state orthogonal to $|\Psi\rangle$. For an ensemble of identical systems we obtain

$$\frac{1}{N}\sum_{i=1}^{N} A_i \prod_{i=1}^{N} |\Psi\rangle_i = \frac{1}{N}\sum_{i=1}^{N} (\overline{A}|\Psi\rangle_i + \Delta A |\Psi_\perp\rangle_i) \prod_{k\neq i}^{N} |\Psi\rangle_k$$
$$= \overline{A} \prod_{i=1}^{N} |\Psi\rangle_i + \frac{\Delta A}{N} \sum_{i=1}^{N} |\Psi_\perp\rangle_i \prod_{k\neq i}^{N} |\Psi\rangle_k . \tag{21}$$

In the limit of large N the second term on the right-hand side of (21) can be neglected because of the mutual orthogonality of the states in the sum, i.e.,

$$\left\| \frac{\Delta A}{N} \sum_{i=1}^{N} |\Psi_\perp\rangle_i \prod_{k\neq i}^{N} |\Psi\rangle_k \right\| = \frac{\Delta A}{\sqrt{N}} \underset{N\to\infty}{\longrightarrow} 0 . \tag{22}$$

Therefore the product of N identical states is, in the limit $N \to \infty$, an eigenstate of any average operator.

The present discussion is reminiscent of the previous idea of a nondisturbing measurement since the interaction Hamiltonian now is

$$H = -g(t) q \frac{1}{N} \sum_{i=1}^{N} A_i . \tag{23}$$

Its action on each separate system is weakened by the factor $1/N$.

The equation $\overline{A} = \langle \Psi | A | \Psi \rangle$ can be taken as a definition of a mean value for a single system. In general, an ensemble of identical systems is needed for measuring $\overline{A}$, however, if the uncertainty ΔA is small compared with δ, the desired precision of the measurement, then even a single measurement of A will yield an outcome close to $\overline{A}$. We shall see in the next section that a similar statement is valid also for weak values.

VII. WEAK MEASUREMENT PERFORMED ON A SINGLE QUANTUM SYSTEM

From our presentation so far one could infer that since one weak measurement because of its uncertainty, provides almost no information, the weak value can be measured only on an ensemble. However, this is not always the case. Even a single measurement may yield information about a weak value provided the "weak uncertainty"

$$\Delta A_w \equiv |(A^2)_w - (A_w)^2|^{1/2} \tag{24}$$

is small compared with the desired precision of the measurement, i.e., $\Delta A_w \ll \delta$. In this case the effective uncertainty of the outcome will be

$$\Delta p_f = [(\Delta p_{\rm in})^2 + (\Delta A_w)^2]^{1/2} . \tag{25}$$

In order for one measurement to yield information about A_w, we have to require, on the one hand, that the measurement is weak enough so that the equation $p_f \cong \mathrm{Re}(A_w)$ [which follows from Eq. (9)] is valid, and, on the other hand, that the measurement is strong enough so that a single outcome yields meaningful information. Thus, if the initial wave function is a Gaussian with $\Delta p_{\rm in} = (2\Delta q)^{-1}$, then Δq should fulfill the condition (8) while the final uncertainty should be small compared with the desired precision $\Delta p_f \ll \delta$.

Let us now demonstrate that there are situations in which A_w is far from the range of eigenvalues of A, and yet the above requirements are fulfilled. In such cases the measurement of A_w, even when performed on a single system, yields a forbidden value. The measurement is imprecise, but the uncertainty is much smaller than the desired precision. For the cases with weak values far from the range of eigenvalues, the probability of obtaining the desired result in the final measurement is extremely small and cannot be easily realized. Nevertheless, it is important conceptually to bolster our conviction that the weak values do indeed describe the reality of a quantum system. We shall describe now such an example.

Consider N spin-$\frac{1}{2}$ particles, all prepared in the state $\sigma_x = 1$ and all found later in the state $\sigma_\xi = 1$ (the directions $\hat{\mathbf{x}}$ and $\hat{\xi}$ are shown in Fig. 1). At an intermediate time t the operator corresponding to the θ component ($\hat{\theta}$ bisects the angle between $\hat{\mathbf{x}}$ and $\hat{\xi}$) of the total spin divided by N is measured. Its weak value is

$$\left[\frac{1}{N} \sum_{i=1}^{N} (\sigma_\theta)_i \right]_w = \frac{1}{\cos(\alpha/2)} , \tag{26}$$

while its weak uncertainty (24) is

$$\Delta \left[\frac{1}{N} \sum_{i=1}^{N} (\sigma_\theta)_i \right]_w = \frac{1}{\sqrt{N}} \left[\frac{1-\cos(\alpha/2)}{\cos(\alpha/2)} \right]^{1/2} . \tag{27}$$

The weakness requirement (8), which will ensure obtaining outcomes close to the weak value (26) for the above interaction, will be satisfied if the following restriction is imposed:

$$\Delta q \ll \sqrt{N} \cos(\alpha/2) . \tag{28}$$

where, again, Δq is the uncertainty of the canonical variable of the measuring device. For large N, Eq. (28) will be fulfilled if, for example,

$$\Delta q = N^{(1/2)-\epsilon} \cos(\alpha/2) , \tag{29}$$

where ϵ is some small positive number. Therefore

$$\Delta p = \frac{1}{2\Delta q} = \frac{N^{-(1/2)+\epsilon}}{2\cos(\alpha/2)} . \tag{30}$$

For N large enough, the uncertainty (25) is smaller than any desired precision δ.

This experiment on a system of N spin-$\frac{1}{2}$ particles is conceptually different from the experiment on an ensemble of N particles discussed in the previous sections. Here *one* weak measurement on *one* system yields the weak value. This outcome is indeed most surprising. For example, when the angle α is chosen to be close to π, the total spin of the above N spin-$\frac{1}{2}$ particles in the direction $\hat{\theta}$ turns out to be much larger than N [it is equal, in fact, to $N/\cos(\alpha/2)$]. Note, however, that this is a very rare event. The result of the experiment is taken into account only if the outcome of the final measurement happens to be $\sigma_\xi = 1$ for all particles. The probability for this event is approximately $[\cos^2(\alpha/2)]^N$. Thus, in order to obtain this surprising result, we have to perform the experiment approximately $[\cos^2(\alpha/2)]^{-N}$ times. Still, it should be emphasized that for those rare cases when the final measurement yields the desired result $[(\sigma_\xi)_i = 1]$, the above measurement shows what appears to be an absolutely impossible result.

We shall now show how the above result can be explained using the standard formalism, which we in no way dispute. In the single-wave-function formalism the above situation is described as follows. Before the weak measurement, the wave function of the measuring device in the p representation is chosen to be a Gaussian centered around zero with the variance given by Eq. (30). The interaction describing the measurement is

$$H = -g(t) q \frac{1}{N} \sum_{i=1}^{N} (\sigma_\theta)_i . \tag{31}$$

The resulting state of the measuring device after the above interaction is over will be a superposition of Gaussians that are shifted in correlation with the spin wave functions. These shifts range from -1 to 1 with steps of size $2/N$. The postselection, i.e., the measurement of the spin components σ_ξ of the particles, will leave the measuring device in a pure state again. This pure state is a particular superposition of Gaussian wave functions which turns out to be approximately a Gaussian located at $1/\cos(\alpha/2)$. A "paradox" happens: all the Gaussians interfere destructively in the region $[-1,1]$ and their tails create a Gaussian centered at a definite point $p = 1/\cos(\alpha/2)$ far away (see Fig. 2). We shall prove this result in the next section.

VIII. A MATHEMATICAL "PARADOX," CAN WE USE IT FOR SENDING SIGNALS FASTER THAN LIGHT?

Let us show that the "paradox" described in the previous section happens not only for Gaussians, but for a large family of functions $f(q)$ describing the initial state of the measuring devices in the q representation, provided they satisfy the condition

$$|f(q)| < a e^{-b|q|} . \tag{32}$$

where a is an arbitrary posistive constant and $b = \frac{1}{2}\tan^2(\alpha/2)$. Again, as we did before, we shall consider the measuring device after the postselection has been performed. Its wave function (up to normalization factor) is

$$|\Psi_{\rm MD}\rangle = \langle \Psi_f| \exp\left[-i\int H\,dt\right]|\Psi_{\rm in}\rangle f(q)$$
$$= \prod_{i=1}^{N} \langle \uparrow_\xi|_i \exp\left[i\int g(t) q \frac{1}{N}\sum_{j=1}^{N} (\sigma_j)_\theta dt\right] \prod_{k=1}^{N} |\uparrow_x\rangle_k f(q)$$
$$= \prod_{i=1}^{N} \langle \uparrow_\xi|_i \exp\left[i\frac{q}{N}(\sigma_i)_\theta\right]|\uparrow_x\rangle_i f(q) . \quad (33)$$

In order to facilitate the calculations it is convenient to represent $\langle \uparrow_\xi|_i$ and $|\uparrow_x\rangle_i$ in the σ_θ representation ($\hat{\theta}$ bisects the angle between $\hat{\mathbf{x}}$ and $\hat{\xi}$)

$$\prod_{i=1}^{N} \langle \uparrow_\xi|_i \exp\left[i\frac{q}{N}(\sigma_i)_\theta\right]|\uparrow_x\rangle_i f(q)$$
$$= \prod_{i=1}^{N} (-i)[\cos(\alpha/4)\langle\uparrow_\theta|_i - \sin(\alpha/4)\langle\downarrow_\theta|_i]\exp\left[i\frac{q}{N}(\sigma_i)_\theta\right][\cos(\alpha/4)|\uparrow_\theta\rangle_i + \sin(\alpha/4)|\downarrow_\theta\rangle_i]f(q)$$
$$= (-i)^N\left[\cos^2(\alpha/4)\exp\left[i\frac{q}{N}\right] - \sin^2(\alpha/4)\exp\left[-i\frac{q}{N}\right]\right]^N f(q) . \quad (34)$$

We shall use later the binomial expansion of this formula, but now we shall continue the calculation in another way:

$$(-i)^N\left[\cos^2(\alpha/4)\exp\left[i\frac{q}{N}\right] - \sin^2(\alpha/4)\exp\left[-i\frac{q}{N}\right]\right]^N f(q)$$
$$= (-i)^N\left[\cos(\alpha/2)\cos\left[\frac{q}{N}\right] + i\sin\left[\frac{q}{N}\right]\right]^N f(q)$$
$$= [-i\cos(\alpha/2)]^N\left[1+\tan^2(\alpha/2)\sin^2\left[\frac{q}{N}\right]\right]^{N/2} e^{i\phi} f(q) , \quad (35)$$

where

$$\phi = N\arctan\left[\frac{\tan(q/N)}{\cos(\alpha/2)}\right] .$$

Now, if $|q| < N^{(1/2)-\epsilon}$, where ϵ is a small positive number, we obtain

$$\lim_{N\to\infty}\left[1+\tan^2(\alpha/2)\sin^2\left[\frac{q}{N}\right]\right]^{N/2} = 1 ,$$
$$\lim_{N\to\infty}\phi = \lim_{N\to\infty} N\arctan\left[\frac{\tan(q/N)}{\cos(\alpha/2)}\right] = \frac{q}{\cos(\alpha/2)} . \quad (36)$$

Therefore, in the limit $N\to\infty$, the wave function of the measuring device (up to normalization factor) is

$$\lim_{N\to\infty}|\Psi_{\rm MD}\rangle = \exp\left[i\frac{q}{\cos(\alpha/2)}\right] f(q) . \quad (37)$$

In order to find wave function of the measuring device in the p representation we cannot just take the Fourier transform of Eq. (37). Instead we have first to take the Fourier integral of $|\Psi_{\rm MD}\rangle$, and then to take the limit $N\to\infty$. We shall then obtain the desired result, that is, that the shift of the wave function in the p representation is equal to $1/\cos(\alpha/2)$ provided that the contribution to the Fourier integral from the region $|q| > N^{(1/2)-\epsilon}$ may be neglected. The rough estimate of the upper bound of this contribution can be done in the following way. From Eq. (35) we see that, up to normalization factor, the contribution to the Fourier integral is smaller than

$$2\int_{N^{(1/2)-\epsilon}}^{\infty}\left[1+\tan^2(\alpha/2)\sin^2\left[\frac{q}{N}\right]\right]^{N/2} a e^{-bq} dq . \quad (38)$$

But,

$$\left[1+\tan^2(\alpha/2)\sin^2\left[\frac{q}{N}\right]\right]^{N/2}$$
$$< \exp\left[\tan^2(\alpha/2)\sin^2\left[\frac{q}{N}\right]\frac{N}{2}\right]$$
$$< \exp\left[\frac{1}{2}\tan^2(\alpha/2) q\right] ,$$

and therefore the contribution is smaller than

$$2a\int_{N^{(1/2)-\epsilon}}^{\infty}\exp\left\{\left[\frac{1}{2}\tan^2(\alpha/2) - b\right] q\right\} dq .$$

If $b > \frac{1}{2}\tan^2(\alpha/2)$ then this integral vanishes as $N \to \infty$. Thus the requirement (32) allows us to take the Fourier transform of Eq. (37).[8] This leads to the surprising result: the superposition of shifts (all smaller than 1) of the function $\tilde{f}(p)$ [the Fourier transform of $f(q)$] is equivalent to a shift by an arbitrarily large value $1/\cos(\alpha/2)$, i.e., (see Fig. 2)

$$\sum_{n=0}^{N} c_n \tilde{f}\left[p + \frac{2n-N}{N}\right] \cong \tilde{f}\left[p - \frac{1}{\cos(\alpha/2)}\right], \qquad (39)$$

where c_n was obtained from the binomial expansion of (34)

$$c_n = \left[\frac{\cos^2(\alpha/4)}{\cos(\alpha/2)}\right]^N \left[-\tan^2(\alpha/4)\right]^n \frac{N!}{n!(N-n)!}. \qquad (40)$$

It would appear that the above procedure may lead to a violation of a number of physical laws. For example, it may seem that we can use this method to send signals with superluminal velocity. To this end we take a system consisting of N spin-$\frac{1}{2}$ particles which is described by the Hamiltonian

$$H = c\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\sigma}_i \cdot \mathbf{p}_i, \qquad (41)$$

where c is the velocity of light. The velocity of the center of mass of the above system of particles is proportional to the total spin

$$\mathbf{V} = c\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\sigma}_i. \qquad (42)$$

The eigenvalues of the total spin are bounded by N, therefore the velocity of the system in any direction is limited by c. Let us now repeat our procedure for the weak measurement of a spin component. Initially we prepare the system with all spins pointed in the x direction and finally we postselect all spins that point in the ξ direction ($\hat{\mathbf{x}}$, $\hat{\boldsymbol{\theta}}$, and $\hat{\boldsymbol{\xi}}$ are specified in the Fig. 1). The initial state is also a product of momentum wave functions that fulfill the weakness condition (32). Consequently, the free Hamiltonian (41) plays the role of a weak measurement of the components of the total spin. For large N the above outcome $[(\sigma_\xi)_i = 1]$ for a final measurement is an extremely rare event, but again, theoretically, we can prepare enough systems so that it will happen at least once. For this rare case the weak value of the component of the total spin in the θ direction [the direction of the real part of the weak vector (18)] is $N/\cos(\alpha/2)$ and, consequently, the effective velocity of the system is [see Eq. (42)] $c/\cos(\alpha/2)$, which is bigger than the velocity of light. This is the velocity of the wave packet which represents our system. Since the form of the wave packet practically does not change, it seems that we can use this wave packet to send information faster than light.

The solution of the paradox is that the amplitude of the wave packet which represented the system initially could not be zero at the final location of the system. In fact, our procedure will not increase the probability of finding the system there because the probability of obtaining the required result for the final measurement is extremely small. We also cannot send the information using the shape of the wave packet which moves with effective superluminal velocity. Indeed the restriction on the wave function of the system that follows from the weakness requirement [similary to (32)] is

$$|f(p)| < \exp\left[-\frac{1}{2}\tan^2(\alpha/2)|p|\right].$$

Consequently, the Fourier transform $\tilde{f}(r)$ is an analytic function in a strip around the real axis. Therefore, even before our measurement procedure took place, there could be found in any open region the information about the shape of the wave function everywhere. However, the desired information about the form of the wave function in a given specific initial location of the wave function is, usually, hidden under the "noise." The above procedure, although it just recovers the existing information, is still very interesting. It does accomplish a very nontrivial task of recovering the information in a way that is not sensitive to the noise; only the signal is amplified.

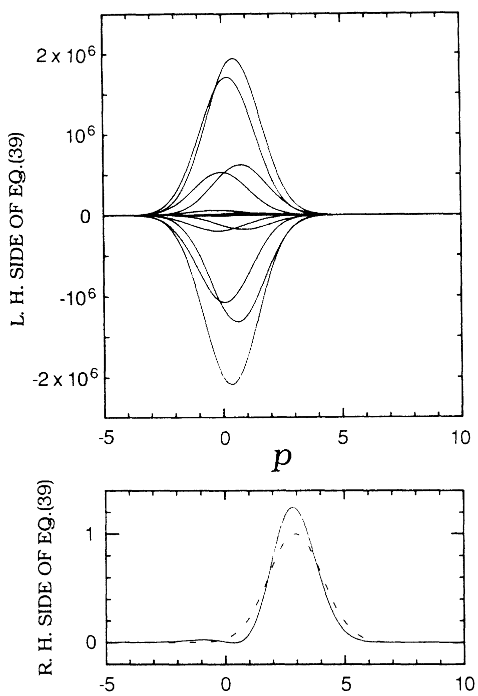

FIG. 2. A mathematical "paradox": superposition of Gaussians shifted by values between -1 and 1 equal to a Gaussian shifted by the value 3. The graph on the top shows all terms in the sum [left-hand side of Eq. (39)] for $N=15$, $\alpha=140°$, $\tilde{f}(p)=\exp[-p^2/4(0.8)^2]$. The solid line of the bottom graph shows the sum of the Gaussians itself, while the dashed line is the Gaussian shifted by $1/\cos(\alpha/2) \cong 3$ [the right-hand side of Eq. (39)].

IX. PRACTICAL REALIZATION OF WEAK MEASUREMENTS

Many of the measurements performed in the laboratory are indeed weak measurements. All thermodynamical variables are some averages of an extremely large number of microscopical systems. During a typical measurement of a thermodynamical quantity only a negligibly small fraction of the microsystems are disturbed.

Although weak measurements are performed regularly, they generally do not yield the surprising outcomes discussed above [such as the spin component of a spin-$\frac{1}{2}$ particle equaling 100 (Ref. 3)]. Surprising outcomes appear only when the measurements are performed on ensembles that are both preselected and postselected and, moreover, they happen only when the postselection is for an improbable outcome. However, weak experiments on an ensemble will yield strange outcomes even if we shall postselect an appropriate finite fraction of the initial preselected ensemble. The problem is to choose the "right" fraction. Before the postselection takes place there is no way to know which are the right systems, so we have to perform measurements on all of them and, after postselection, to take into account the results of measurements only for systems with the proper outcome of the final measurement. This procedure, which seems, at first, very difficult to perform, becomes feasible when the measuring device is the observed system itself. Some other (not the measured one) degree of freedom plays the role of a pointer of the measuring device. In a realistic experiment suggested by us before[3] the weak value of the spin component of a spin-$\frac{1}{2}$ particle is measured while the position of the particle serves as the pointer of the measuring device. This experiment is of the standard Stern-Gerlach type, modified to fulfill the requirements of weakness, with both preselection and postselection included. A clear discussion of experimental conditions for this experiment is given by Duck, Stevenson, and Sudarshan.[9]

It is possible that, as in the case of checking Bell inequalities, it is easier to perform an optical analog of the above proposed experiment. One of the possibilities is to replace the preselection and postselection spin measurements by filters of linear polarization, while the optical analog of a weak Stern-Gerlach measurement will be a small-angle prism made from optically active material.[10]

X. CONCLUSIONS

In this article a formalism that describes a quantum system at a given time using two wave functions evolving in opposite directions of time has been discussed. This description introduces new characteristics of quantum systems between two measurements: the weak values of physical variables. The weak values are measurable quantities. We can obtain them as an outcome of weak measurements performed on an ensemble of both preselected and postselected quantum systems. Although usually we can measure weak values with good accuracy only by performing an experiment on a large ensemble of identical systems, there are situations in which one weak measurement performed on a single system yields a weak value of a measured value with good precision.

The weak values of an observable can differ widely from the range of eigenvalues of the observable. The weakness requirement of the interaction which ensures obtaining weak values is not an extraordinary one. In fact, many standard experiments fulfill this weakness condition. No strange weak values have been observed so far because usually the experiments do not involve ensembles that are both preselected and postselected. It seems, however, that such experiments can be done. For example, a set of Stern-Gerlach devices or their optical analog can apparently suit this purpose.

The effective value of A for *any* (weak enough) interaction with the systems belonging to preselected and postselected ensembles is, in fact, the weak value A_w. This result does not contradict the standard theory, but its explanation in the standard approach is rather subtle and is due to peculiar quantum interference.

Weak measurements of this type hold a promise for important applications. When performed on an ensemble of systems that are both preselected and postselected they can effectively amplify or "tune" any physical variable to a certain (even forbidden) value.

ACKNOWLEDGMENTS

We wish to acknowledge valuable discussions with David Albert, Aharon Casher, Shmuel Nussinov, Abner Shimony, and Sandu Popescu; and we wish to thank Orly Alter and Oron Zachar for careful reading of our work. This work was supported in part by the National Science Foundation under Grant No. PHY-8807812.

---

[1] Y. Aharonov, D. Albert, A. Casher, and L. Vaidman, Phys. Lett. A **124**, 199 (1987).

[2] L. Vaidman, Y. Aharonov, and D. Albert, Phys. Rev. Lett. **58**, 1385 (1987).

[3] Y. Aharonov, D. Albert, and L. Vaidman, Phys. Rev. Lett. **60**, 1351 (1988).

[4] Y. Aharonov, P. G. Bergman, and J. L. Lebowitz, Phys. Rev. B **134**, 1410 (1964).

[5] In the classical theory this view does not lead to asymmetry under time reversal since the classical theory is deterministic.

[6] J. von Neumann, *Mathematische Gründerlagen der Quantemechanik* (Springer-Verlag, Berlin, 1932) [English translation: *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1983)].

[7] Although the proof of this simple formula takes two lines, we never saw it in the literature. We can always decompose $A|\Psi\rangle=\alpha|\Psi\rangle+\beta|\Psi_\perp\rangle$ with β real and non-negative. Then $\langle\Psi|A|\Psi\rangle=\langle\Psi|(\alpha|\Psi\rangle+\beta|\Psi_\perp\rangle)$ yields $\alpha=\overline{A}$, and $\langle\Psi|A^\dagger A|\Psi\rangle=(\alpha^*\langle\Psi|+\beta^*\langle\Psi_\perp|)(\alpha|\Psi\rangle+\beta|\Psi_\perp\rangle)$ yields $\beta=[(\overline{A^2})-(\overline{A})^2]^{1/2}\equiv\Delta A$.

[8] The limit defined by the condition (32), $b=\frac{1}{2}\tan^2(\alpha/2)$, is not the optimal one. A better limit can be found by selecting the

smallest value of b such that the integral (38) in the limit $N \rightarrow \infty$ vanishes.

[9]I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan, Phys. Rev. D **40**, 2112 (1989).

[10]Somewhat different optical realizations of weak experiments were suggested in Ref. 9 and, independently, by R. Y. Chiao (private communication).