Physica Scripta. Vol. T76, 85–92, 1998

On the Two-State Vector Reformulation of Quantum Mechanics

Yakir Aharonov[a,b] and Lev Vaidman[a]

[a] School of Physics and Astronomy, Tel-Aviv University, Tel-Aviv, 69978 Israel
[b] Physics Department, University of South Carolina Columbia, South Carolina, 29208, U.S.A.

Received September 18, 1997; revised version received November 25, 1997; accepted December 4, 1997

PACS Ref: 03.65.Bz

Abstract

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 post-selected 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 measuring procedure that fulfills certain requirements of "weakness". We call it *weak measurement*. Physical meaning and underlying mathematical structure of weak measurements are explored.

1. Introduction

In this paper we discuss a new formulation of quantum mechanics (QM) which is time symmetric. The basic idea is that at any moment of time a quantum system can be described by two wave-functions, the usual one evolving from past to future as well as a second wave-function evolving from future to past. Although this new formulation reproduces all the predictions of standard QM – (that is, the unusual effects that we predict can be explained in the standard formulation of QM) – our formulation has several advantages: it provides a new conceptual framework which suggests possible modifications to QM and it allows one to ask and solve new interesting questions which are hidden in the old viewpoint.

In 1964 Aharonov, Bergmann and Lebowitz (ABL) [1] investigated the claim that fundamental time asymmetry was introduced by quantum measurement theory and its assumed collape of the wave-function. They found, however, that QM does not itself introduce an asymmetry but that the asymmetry is a result of the way in which the statistical ensemble is created. This apparently was a habit carried over from classical determinism which allowed predictions of future events based solely on a single complete set of initial conditions. ABL showed that if the ensembles are created time-symmetrically, by specifying both initial and final boundary conditions, then the intermediate probability distributions are also time-symmetric. Indeed, in QM boundary conditions specified at different times are not equivalent to and cannot in general be mapped to boundary conditions specified at one time. For example, suppose we know the state of a particle at time t_1 to be $|\Psi_1\rangle$ and we also know all the interactions which occur between t_1 and a later time t_2. In general, we cannot know with certainty the result of another measurement at t_2. Actually performing the measurement and obtaining an answer provides new information. ABL suggested that this could be described by a fundamentally new kind of ensemble. ABL did this by considering measurements performed on a quantum system between two other measurements, the results of which were given. They proposed describing the quantum system betweem two measurements by using two states: the usual one, evolving towards the future from the time of the first measurement, and a second state evolving backwards in time, from the time of the second measurement. If a system has been prepared at time t_1 in a state $|\Psi_1\rangle$ and is found at time t_2 in a $|\Psi_2\rangle$, then at time t, $t_1 < t < t_2$, the system is described by

$$\langle\Psi_2|U^\dagger(t, t_2) \quad \text{and} \quad U(t_1, t)|\Psi_1\rangle. \tag{1}$$

In order to obtain such a system we prepare an ensemble of systems in the state $|\Psi_1\rangle$, perform measurement of a desired variable using separate measuring devices for each system in the ensemble, and perform the post-selection measurement. If the outcome of the post-selection was not the desired result, we discard the system and corresponding measuring device. We look only on measuring devices corresponding to the systems post-selected in the state $\langle\Psi_2|$. $\langle\Psi_2|$ and $|\Psi_1\rangle$, therefore, define the pre- and post-selected sub-ensemble between the times t_1 and t_2. That is, we are not concerned with the probabilities of obtaining a given pre- and post-selected system, but rather with the properties of this particular sub-ensemble, assuming we are successful in preparing it. ABL considered these properties by investigating the conditional probabilities for obtaining different outcomes for an observable A at the intermediate time t. The probability for an outcome $A = a_i$ given just the initial state $|\Psi_1\rangle$ is $|\langle a_i|U(t, t_1)|\Psi_1\rangle|^2$. If $A = a_i$ is obtained at time t, then the state of the particle can be replaced by $A = a_i$ and the probability to obtain $\langle\Psi_2|$ at time t_2 given $A = a_i$ is: $|\langle\Psi_2|U(t_2, t)|a_i\rangle|^2$. Therefore, the probability to obtain $A = a_i$ followed by $\langle\Psi_2|$ is:

$$P(a_i|\langle\Psi_2|, |\Psi_1\rangle) = \frac{|\langle\Psi_2|U(t_2, t)|a_i\rangle|^2 |\langle a_i|U(t, t_1)|\Psi_1\rangle|^2}{\sum_j |\langle\Psi_2|U(t_2, t)|a_j\rangle|^2 |\langle a_j|U(t, t_1)|\Psi_1\rangle|^2}. \tag{2}$$

(The denominator is the normalization, which is the total probability to obtain $\langle\Psi_2|$ given that A was measured during the intermediate time.) Several features of this result are evident upon inspection: (1) the statistics of A depends on the particular pre- and post-selected ensemble, (2) the statistics for A does not depend on how the pre- and post-selected ensemble was obtained, and (3) once a complete set of observables is ascertained at t_1 and t_2, the statistics for intermediate times does not depend on measurements made before t_1 or after t_2. This formula can be rewritten to be manifestly time-symmetric by applying the time evolution operatore to $\langle \Psi_2 |$, instead of to the eigenstates of A, since $U(t_2, t) = \exp[-iH(t_2 - t)] = U^\dagger(t, t_2)$ and therefore $\langle \Psi_2 | U^\dagger(t_2, t) = [U(t_2, t) | \Psi_2\rangle]^\dagger$,

$$P(a_i | \langle \Psi_2 |, | \Psi_1 \rangle) = \frac{|\langle a_i | U(t, t_2) | \Psi_2 \rangle|^2 \, |\langle a_i | U(t, t_2) | \Psi_1 \rangle|^2}{\sum_j |\langle a_j | U(t, t_2) | \Psi_2 \rangle|^2 \, |\langle a_j | U(t, t_1) | \Psi_1 \rangle|^2}. \quad (3)$$

This form of the ABL probability is also better from a computational standpoint since it is easier to evolve a single final state $\langle \Psi_2 |$ backwards in time than it is to evolve forward many intermediate states $| a_i \rangle$.

The basic concepts of the two-state approach, weak measurement and weak values, were developed several years ago [2–4]. The weak value of any physical variable A in the time interval between pre-selection of the state $| \Psi_1 \rangle$ and post-selection of the state $| \Psi_2 \rangle$ is given by

$$A_w \equiv \frac{\langle \Psi_2 | A | \Psi_1 \rangle}{\langle \Psi_2 | \Psi_1 \rangle}. \quad (4)$$

Let us present the main idea by way of a simple example. We consider, at time t, a quantum system which was prepared at time t_1 in the state $| B = b \rangle$ and was found at time t_2 in the state $| C = c \rangle$, $t_1 < t < t_2$. The measurements at times t_1 and t_2 are complete measurements of, in general, non-commuting variables B and C. In this example, the free Hamiltonian is zero, and therefore, the first quantum state at time t is $| B = b \rangle$. In the two-state approach we characterize the system at time t by backwards-evolving state $\langle C = c |$ as well. Our motivation for including the future state is that we know that if a measurement of C has been performed at time t then the outcome is $C = c$ with probability 1. This intermediate measurement, however, destroys our knowledge that $B = b$, since the coupling of the measuring device to the variable C can change B. The idea of weak measurements is to make the coupling with the measuring device sufficiently weak so that the change in B can be neglected. In fact, we require that both quantum states do not change significantly, neither the usual one $| B = b \rangle$ evolving towards the future nor $\langle C = c |$ evolving backwards.

During the whole time interval between t_1 and t_2, both $B = b$ and $C = c$ are true (in some sense). But then, $B + C = b + c$ must also be true. The latter statement, however, might not have meaning in the standard quantum formalism because the sum of the eigenvalues $b + c$ might not be an eigenvalue of the operator $B + C$. An attempt to measure $B + C$ using a standard measuring procedure will lead to some change of the two quantum states and thus the outcome will not be $b + c$. A weak measurement, however, will yield $b + c$.

When the "strong" value of an observable is known with certainty, i.e., we know the outcome of an ideal (infinitely strong) measurement with probability 1, then the weak value is equal to the strong value. Let us analyze the example above. The strong value of B is b, its eigenvalue. The strong value of C is c, as we know from *retrodiction.* From the definition (4) immediately follows: $B_w = b$ and $C_w = c$. But weak values, unlike strong values are defined not just for B and C but for *all* operators. The strong value of the sum $B + C$ when $[B, C] \neq 0$ is not certain, but the weak value of the sum is: $(B + C)_w = b + c$.

2. Quantum measurements

In the standard approach to measurements in quantum theory, we measure observables which correspond to Hermitian operators. The latter have eigenvalues and a ("good") measurement must yield one of these eigenvalues. If the state of a quantum system is not an eigenstate of the measured operator, then one can predict only probabilities for different outcomes of the measurement. The state of the system invariably "collapses" to an outcome corresponding to one eigenvalue. A standard measurement of a variable A is modeled in the von Neumann theory of measurement [5] by a Hamiltonian

$$H = g(t) P A, \quad (5)$$

where P is a canonical momentum, conjugate to the pointer variable Q of the measuring device. The function $g(t)$ is nonzero only for a very short time interval corresponding to the measurement, and is normalized so that $\int g(t) dt = 1$. During the time of this impulsive measurement, the Hamiltonian (5) dominates the evolution of the measured system and the measuring device. Since $[A, H] = 0$, the variable A does not change during the measuring interaction. If initially the position Q of the measuring device is precisely defined, then in the Heisenberg representation, the pointer shifts its value by $Q_{\text{fin}} - Q_{\text{in}} = A$. This measurement of A disturbs any other observable B which does not commute with A, i.e. $\dot{B} = g(t) P [A, B]$ and since $[A, B] \neq 0$, B changes during the measurement. If Q_{in} is precisely fixed, then its conjugate variable P is completely uncertain, thus implying that any variable B that does not commute with A will be uncontrollably disturbed during the measurement of A. However, by allowing for inaccuracies in Q, thereby weakening the interaction, we can decrease the disturbance in non-commuting observables. These inaccuracies can be modeled by introducing a spread of Δ in the Gaussian distribution of Q, thereby producing errors in the measurement of A of order Δ, The initial state of the pointer variable is thus modeled by a Gaussian centered at zero:

$$\Phi_{\text{in}}(Q) = (\Delta^2 \pi)^{-1/4} \, e^{-Q^2/2\Delta^2}. \quad (6)$$

Therefore, if the initial state of the system is a superposition $| \Psi_1 \rangle = \sum \alpha_i | a_i \rangle$, then after the interaction (5) the state of the system and the measuring device is:

$$(\Delta^2 \pi)^{-1/4} \sum \alpha_i | a_i \rangle \, e^{-(Q - a_i)^2/2\Delta^2}. \quad (7)$$

If the separation between various eigenvalues a_i is much larger than the width of the Gaussian Δ, we obtain strict correlation between the values of the variable A and nearly orthogonal states of the measuring device. The measuring procedure continues with an amplification scheme which yields effective (or, according to some physicists, real) collapse to one of the pointer positions and the corresponding eigenstate $| a_i \rangle$. In this model the only possible outcomes of the measurement of the quantum variable A are the eigenvalues a_i. This fact perfectly matches the premise that the only values which can be associated with A are the a_i.

In the two-state vector formalism the system at time t in a pre- and post-selected ensemble is defined by two states, the usual one evolving from the time of the preparation and the state evolving backwards in time from the post-selection. We may neglect the free Hamiltonian if the time between the pre-selection and the post-selection is very short. Consider a system which has been pre-selected in a state $|\Psi_1\rangle$ and shortly afterwards post-selected in a state $|\Psi_2\rangle$. The weak value of any physical variable A in the time interval between the pre-selection and the post-selection is given by eq. (4). Let us show briefly how weak values emerge from a measuring procedure with a sufficiently weak interaction.

We consider a sequence of measurements: a pre-selection of $|\Psi_1\rangle$, a (weak) measurement interaction of the form of eq. (5), and a post-selection measurement finding the state $|\Psi_2\rangle$. The state of the measuring device after this sequence is given (up to normalization) by

$$\Phi(Q) = \langle\Psi_2|\,\mathrm{e}^{-\mathrm{i}PA}|\Psi_1\rangle\,\mathrm{e}^{-Q^2/2\Delta^2}. \quad (8)$$

After simple algebraic manipulation we can rewrite it (in the P-representation) as

$$\tilde{\Phi}(P) = \langle\Psi_2|\Psi_1\rangle\,\mathrm{e}^{-\mathrm{i}A_\mathrm{w}P}\,\mathrm{e}^{-\Delta^2P^2/2} + \langle\Psi_2|\Psi_1\rangle \sum_{n=2}^{\infty} \frac{(\mathrm{i}P)^n}{n!}\,[(A^n)_\mathrm{w} - (A_\mathrm{w})^n]\,\mathrm{e}^{-\Delta^2P^2/2}. \quad (9)$$

If Δ is sufficiently large, we can neglect the second term of (9) when we Fourier transform back to the Q-representation. Large Δ corresponds to weak measurement in the sense that the interaction Hamiltonian (5) is small. Thus, in the limit of weak measurement, the final state of the measuring device (in the Q-representation) is

$$\Phi(Q) = (\Delta^2\pi)^{-1/4}\,\mathrm{e}^{-(Q-A_\mathrm{w})^2/2\Delta^2}, \quad (10)$$

This state represents a measuring device pointing to the weak value, A_w.

Weak measurements on pre- and post-selected ensembles yield, instead of eigenvalues, a value which might lie far outside the range of the eigenvalues. This suggests that quantum observables may have a richer structure than suggested by standard QM. Although we have shown this result for a specific von Neumann model of measurements, the result is completely general: any coupling of a pre- and post-selected system to a variable A, provided the coupling is sufficiently weak, results in effective coupling to A_w. This weak coupling between a single system and the measuring device will not, in most cases, lead to a distinguishable shift of the pointer variable, but collecting the results of measurements on an ensemble of pre- and post-selected systems will yield the weak values of a measured variable to any desired precision.

When the strength of the coupling to the measuring device goes to zero, the outcomes of the measurement invariably yield the weak value. To be more precise, a measurement yields the real part of the weak value. Indeed, the weak value is, in general, a complex number, but its imaginary part will contribute only a (position dependent) phase to the wave function of the measuring device in the position representation of the pointer. Therefore, the imaginary part will not affect the probability distribution of the pointer position which is what we see in a usual measurement. However, the imaginary part of the weak value also has physical meaning. It expresses itself as a change in the conjugate momentum of the pointer variable [3].

3. An example: spin measurement

Let us consider a simple Stern-Gerlach experiment: measurement of a spin component of a spin-1/2 particle. We shall consider a particle prepared in the initial state spin "up" in the $\hat{x}$ direction and post-selected to be "up" in the $\hat{y}$ direction. At the intermediate time we measure, weakly, the spin component in the $\hat{\xi}$ direction which is bisector of $\hat{x}$ and $\hat{y}$, i.e., $\sigma_\xi = (\sigma_x + \sigma_y)/\sqrt{2}$. Thus $|\Psi_1\rangle = |{\uparrow_x}\rangle$, $|\Psi_2\rangle = |{\uparrow_y}\rangle$, and the weak value of σ_ξ in this case is:

$$(\sigma_\xi)_\mathrm{w} = \frac{\langle{\uparrow_y}|\sigma_\xi|{\uparrow_x}\rangle}{\langle{\uparrow_y}|{\uparrow_x}\rangle} = \frac{1}{\sqrt{2}}\,\frac{\langle{\uparrow_y}|(\sigma_x+\sigma_y)|{\uparrow_x}\rangle}{\langle{\uparrow_y}|{\uparrow_x}\rangle} = \sqrt{2}. \quad (11)$$

This value is, of course, "forbidden" in the standard interpretation where a spin component can obtain the (eigen)values ± 1 only.

An effective Hamiltonian for measuring σ_ξ is

$$H = g(t)P\sigma_\xi. \quad (12)$$

Writing the initial state of the particle in the σ_ξ representation, and assuming the initial state (6) for the measuring device, we obtain that after the measuring interaction the quantum state of the system and the pointer of the measuring device is

$$\cos(\pi/8)|{\uparrow_\xi}\rangle\,\mathrm{e}^{-(Q-1)^2/2\Delta^2} + \sin(\pi/8)|{\downarrow_\xi}\rangle\,\mathrm{e}^{-(Q+1)^2/2\Delta^2}. \quad (13)$$

The probability distribution of the pointer position, if it is observed now without post-selection, is the sum of the distributions for each spin value. It is, up to normalization,

$$\mathrm{prob}\,(Q) = \cos^2(\pi/8)\,\mathrm{e}^{-(Q-1)^2/\Delta^2} + \sin^2(\pi/8)\,\mathrm{e}^{-(Q+1)^2/\Delta^2}. \quad (14)$$

In the usual strong measurement $\Delta \ll 1$. In this case, as shown on Fig. 1(a), probability distribution of the pointer is localized around -1 and $+1$ and it is strongly correlated to the values of the spin, $\sigma_z = \pm 1$.

Weak measurements correspond to a Δ which is much larger than the range of the eigenvalues, i.e., $\Delta \gg 1$. Fig. 1(b) shows that the pointer distribution has a large uncertainty, but it is peaked between the eigenvalues, more precisely, at the expectation value $\langle{\uparrow_x}|\sigma_\xi|{\uparrow_x}\rangle = 1/\sqrt{2}$. An outcome of an individual measurement usually will not be close to this number, but it can be found from an ensemble of such measurements, see Fig. 1(c). Note, that we have not yet considered the post-selection.

In order to simplify the analysis of measurements on the pre- and post-selected ensemble, let us assume that we first make the post-selection of the spin of the particle and only then look at the pointer of the device that weakly measures σ_ξ. We must get the same result as if we first look at the outcome of the weak measurement, make the post-selection, and discard all readings of the weak measurement corresponding to the cases in which the result is not $\sigma_y = 1$. The post-selected state of the particle in the σ_ξ representation is $|{\uparrow_y}\rangle = \cos(\pi/8)|{\uparrow_\xi}\rangle - \sin(\pi/8)|{\downarrow_\xi}\rangle$. The state of the measuring device after the post-selection of the spin state is obtained by projection of (13) onto the post-selected spin state:

$$\Phi(Q) = \mathcal{N}(\cos^2(\pi/8)\,\mathrm{e}^{-(Q-1)^2/2\Delta^2} - \sin^2(\pi/8)\,\mathrm{e}^{-(Q+1)^2/2\Delta^2}), \quad (15)$$

where $\mathcal{N}$ is a normalization factor. The probability distribution of the pointer variable is given by

$$\mathrm{prob}\,(Q) = \mathcal{N}^2(\cos^2(\pi/8)\,\mathrm{e}^{-(Q-1)^2/2\Delta^2} - \sin^2(\pi/8)\,\mathrm{e}^{-(Q+1)^2/2\Delta^2})^2. \quad (16)$$

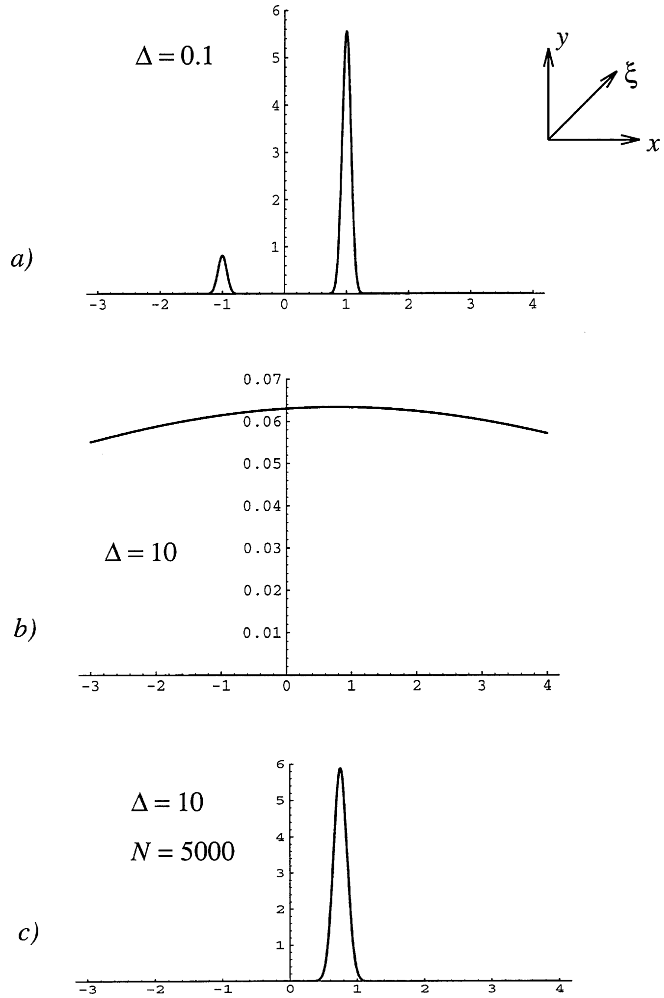

Fig. 1. Spin component measurement without post-selection. Probability distribution of the pointer variable for measurement of σ_ξ when the particle is pre-selected in the state $|\uparrow_x\rangle$. (a) Strong measurement, $\Delta = 0.1$. (b) Weak measurement, $\Delta = 10$. (c) Weak measurement on the ensemble of 5000 particles. The original width of the peak, 10, is reduced to $10/\sqrt{5000} \simeq 0.14$. In the strong measurement (a) the pointer is localized around the eigenvalues ± 1, while in the weak measurements (b) and (c) the peak is located in the expectation value $\langle\uparrow_x|\sigma_\xi|\uparrow_x\rangle = 1/\sqrt{2}$.

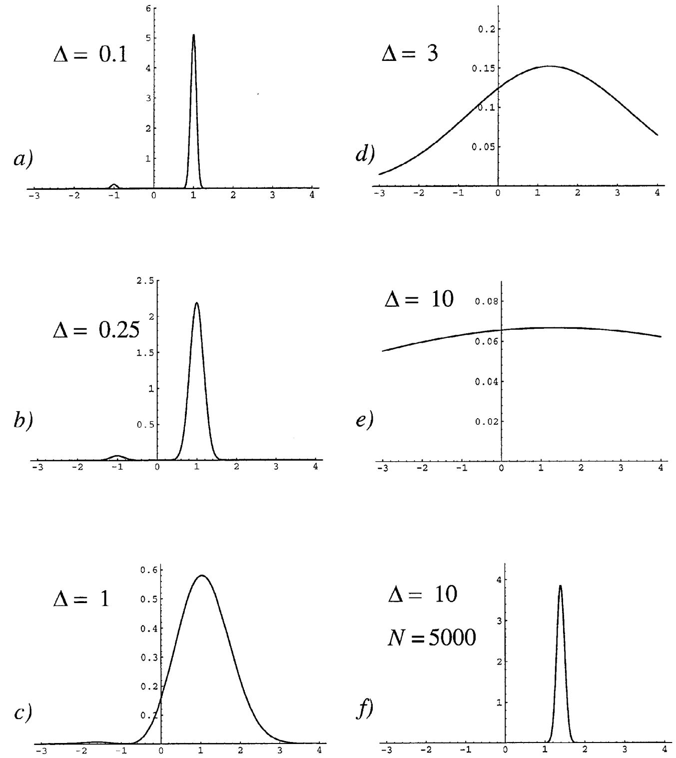

Fig. 2. Measurement on pre- and post-selected ensemble. Probability distribution of the pointer variable for measurement of σ_ξ when the particle is pre-selected in the state $|\uparrow_x\rangle$ and post-selected in the state $|\uparrow_y\rangle$. The strength of the measurement is parameterized by the width of the distribution Δ. (a) $\Delta = 0.1$; (b) $\Delta = 0.25$; (c) $\Delta = 1$; (d) $\Delta = 3$; (e) $\Delta = 10$. (f) Weak measurement on the ensemble of 5000 particles; the original width of the peak, $\Delta = 10$, is reduced to $10/\sqrt{5000} \simeq 0.14$. In the strong measurements (a)–(b) the pointer is localized around the eigenvalues ± 1, while in the weak measurements (d)–(f) the peak of the distribution is located in the weak value $(\sigma_\xi)_w = \langle\uparrow_y|\sigma_\xi|\uparrow_x\rangle/\langle\uparrow_y|\uparrow_x\rangle = \sqrt{2}$. The outcomes of the weak measurement on the ensemble of 5000 pre- and post-selected particles, (f), are clearly outside the range of the eigenvalues, $(-1, 1)$.

If the measuring interaction is strong, $\Delta \ll 1$, then the distribution is localized around the eigenvalues ± 1 (mostly around 1 since the pre- and post-selected probability to find $\sigma_\xi = 1$ is more than 85%, see Figs 2(a), (b). But when the strength of the coupling is weakened, i.e., Δ is increased, the distribution gradually changes to a single broad peak around $\sqrt{2}$, the weak value, see Figs 2(c)–(e).

The width of the peak is large and therefore each individual reading of the pointer usually will be far from $\sqrt{2}$. The physical meaning of the weak value can, in this case, be associated only with an ensemble of pre- and post-selected particles. The accuracy of defining the center of the distribution goes as $1/\sqrt{N}$, so increasing N, the number of particles in the ensemble, we can find the weak value with any desired precision, see Fig. 2(f).

In our example, the weak value of the spin component is $\sqrt{2}$, which is only slightly more than the maximal eigenvalue, 1. By appropriate choice of the pre- and post-selected states we can get pre- and post-selected ensembles with arbitrarily large weak value of a spin component. One of our first proposals [6] was to obtain $(\sigma_\xi)_w = 100$. In this case the post-selected state is nearly orthogonal to the pre-selected state and, therefore, the probability to obtain appropriate post-selection becomes very small. While in the case of $(\sigma_\xi)_w = \sqrt{2}$ the (pre- and) post-selected ensemble was just half of the pre-selected ensemble, in the case of $(\sigma_\xi)_w = 100$ the post-selected ensemble will be smaller than the original ensemble by the factor of $\sim 10^{-4}$.

4. Weak measurements on a single system

We have shown that weak measurements can yield very surprising values which are far from the range of the eigenvalues. However, the uncertainty of a single weak measurement (i.e., performed on a single system) in the above example is larger than the deviation from the range of the eigenvalues. Each single measurement separately yields almost no information and the weak value arises only from the statistical average on the ensemble. The weakness and the uncertainty of the measurement goes together. Weak measurement corresponds to small value of P in the Hamiltonian (5) and, therefore, the uncertainty in P has to be very small. This requires large Δ, the uncertainty of the pointer variable. Of course, we can construct measurement with large uncertainty which is not weak at all, for example, by preparing the measuring device in a mixed state instead of a Gaussian, but no precise measurement with weak coupling is possible. So, usually, a weak measurement on a single system will not yield the weak values with a good precision.

However. there are special cases when it is not so. Usual strength measurement on a single pre- and post-selected system can yield "unusual" (very different from the eigenvalues) weak value with a good precision. Good precision means that the uncertainty is much smaller than the deviation from the range of the eigenvalues.

Our example above was not such a case. The weak value $(\sigma_\xi)_w = \sqrt{2}$ is larger than the highest eigenvalue, 1, only by ~ 0.4, while the uncertainty, 1, is not sufficiently large for obtaining the peak of the distribution near the weak value, see Fig. 2(c). Let us modify our experiment in such a way that a single experiment will yield a meaningful surprising result. We consider a system of N spin-1/2 particles all prepared in the state $|\uparrow_x\rangle$ and post-selected in the state $|\uparrow_y\rangle$, i.e., $|\Psi_1\rangle = \prod_{i=1}^N |\uparrow_x\rangle_i$ and $|\Psi_2\rangle = \prod_{i=1}^N |\uparrow_y\rangle_i$. The variable which is measured at the intermediate time is $A \equiv (\sum_{i=1}^N (\sigma_i)_\xi)/N$. The operator A has $N+1$ eigenvalues equally spaced between -1 and $+1$, but the weak value of A is

$$A_w = \frac{\prod_{k=1}^N \langle\uparrow_y|_k \sum_{i=1}^N ((\sigma_i)_x + (\sigma_i)_y) \prod_{j=1}^N |\uparrow_x\rangle_j}{\sqrt{2}N(\langle\uparrow_y|\uparrow_x\rangle)^N} = \sqrt{2}. \quad (17)$$

The interaction Hamiltonian is

$$H = \frac{g(t)}{N} P \sum_{i=1}^{N} (\sigma_i)_\xi. \quad (18)$$

The initial state of the measuring device defines the precision of the measurement. When we take it to be the Gaussian (6), it is characterized by the width Δ. For a meaningful experiment we have to take Δ small. Small Δ corresponds to large uncertain P, but now, the strength of the coupling to each individual spin is reduced by the factor $1/N$. Therefore, for large N, both the forward-evolving state and the backward-evolving state are essentially not changed by the coupling to the measuring device. Thus, this single measurement yields the weak value. In Ref. [7] it is proven that if the measured observable is an average on a large set of systems, $A = (\sum_i^N A_i)/N$, then we can always construct a single, good-precision measurement of the weak value. Here let us present just numerical calculations of the probability distribution of the measuring device for N pre- and post-selected spin-1/2 particles. The state of the pointer after the post-selection for this case is

$$\mathcal{N} \sum_{i=1}^{N} (-1)^i (\cos^2(\pi/8))^{N-i} (\sin^2(\pi/8))^i \, e^{-(Q-[(2N-i)/N])^2/2\Delta^2}. \quad (19)$$

The proability distribution for the pointer variable Q is

$$\text{prob}\,(Q) = \mathcal{N}^2 \left(\sum_{i=1}^{N} (-1)^i (\cos^2(\pi/8))^{N-i} \times (\sin^2(\pi/8))^i \, e^{-(Q-[(2N-i)/N])^2/2\Delta^2} \right)^2. \quad (20)$$

The result for $N = 20$ and different values of Δ are presented in Fig. 3. We see that for $\Delta = 0.25$ and larger, the obtained results are very good: the final probability distribution of the pointer is peaked at the weak value, $((\sum_{i=1}^N (\sigma_i)_\xi)/N)_w = \sqrt{2}$. This distribution is very close to that of a measuring device measuring operator O on a system in an eigenstate $|O = \sqrt{2}\rangle$. For N large, the relative uncertainty can be decreased almost by a factor $1/\sqrt{N}$ without changing the fact that the peak of the distribution points to the weak value.

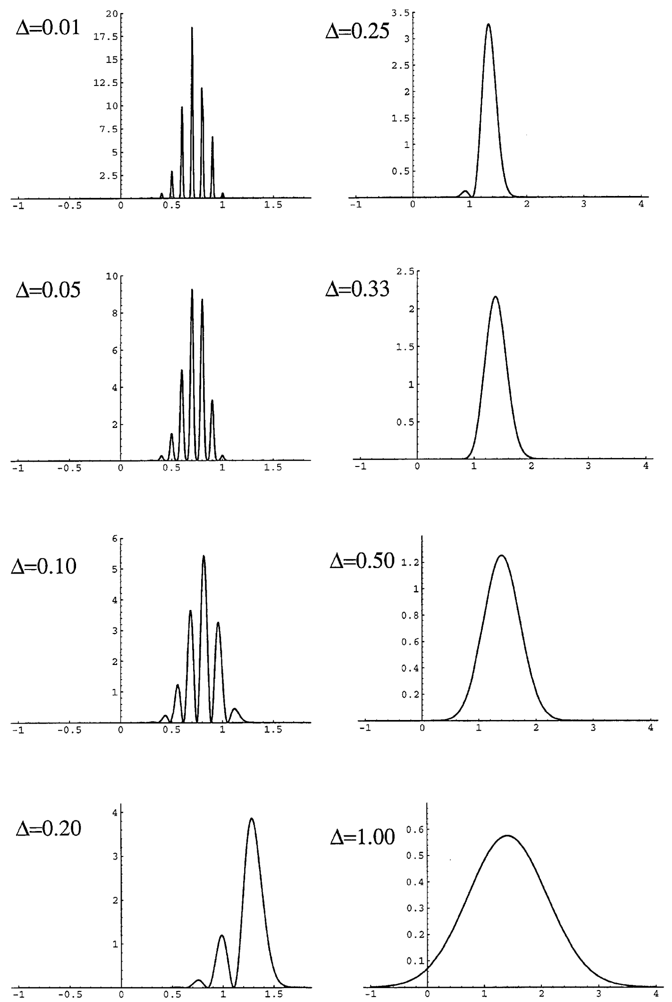

Fig. 3. Measurement on a single system. Probability distribution of the pointer variable for measurement of $A = (\sum_{i=1}^{20} (\sigma_i)_\xi)/20$ when the system of 20 spin-1/2 particles is pre-selected in the state $|\Psi_1\rangle = \prod_{i=1}^{20} |\uparrow_x\rangle_i$ and post-selected in the state $|\Psi_2\rangle = \prod_{i=1}^{20} |\uparrow_y\rangle_i$. While in the very strong measurements, $\Delta = 0.01$–0.05, the peaks of the distribution located at the eigenvalues, starting from $\Delta = 0.25$ there is essentially a single peak at the location of the weak value, $A_w = \sqrt{2}$.

Although our set of particles pre-selected in one state and post-selected in another state is considered as one system, it looks very much as an ensemble. In quantum theory, measurement of the sum does not necessarily yield the same result as the sum of the results of the separate measurements, so conceptually our measurement on the set of particles differs from the measurement on an ensemble of pre- and post-selected particles. However, in our example of weak measurements, the results are the same.

A less ambiguous case is the example considered in the first work on weak measurements [2]. In this work a single system of a large spin N is considered. The system is pre-selected in the state $|\Psi_1\rangle = |S_x = N\rangle$ and post-selected in the state $|\Psi_2\rangle = |S_y = N\rangle$. At an intermediate time the spin component S_ξ is weakly measured and again the "forbidden" value $\sqrt{2}N$ is obtained. The uncertainty has to be only slightly larger than $\sqrt{N}$. The probability distribution of the results is centered around $\sqrt{2}N$, and for large N it lies clearly outside the range of the eigenvalues, $(-N, N)$. Unruh [8] made computer calculations of the distribution of the pointer variable for this case and got results which are very similar to what is presented on Fig. 3.

An even more dramatic example is a measurement of the kinetic energy of a tunneling particle [9]. We consider a particle pre-selected in a bound state of a potential well which has negative potential near the origin and vanishing potential far from the origin; $|\Psi_1\rangle = |E = E_0\rangle$. Shortly later, the particle is post-selected to be far from the well, inside a classically forbidden tunneling region; this state can be characterized by vanishing potential $|\Psi_2\rangle = |U = 0\rangle$. At an intermediate time a measurement of the kinetic energy is performed. The weak value of the kinetic energy in this case is

$$K_w = \frac{\langle U = 0 | K | E = E_0 \rangle}{\langle U = 0 | E = E_0 \rangle} = \frac{\langle U = 0 | E - U | E = E_0 \rangle}{\langle U = 0 | E = E_0 \rangle} = E_0 . \quad (21)$$

The energy of the bound state, E_0, is negative, so the weak value of the kinetic energy is negative. In order to obtain this negative value the coupling to the measuring device need not be too weak. In fact, for any finite strength of the measurement we can choose the post-selected state sufficiently far from the well to ensure the negative value. Therefore, for appropriate post-selection, the normal *strong* measurement of a positive definite operator invariably yields a negative result! This weak value predicted by the two-state vector formalism demonstrates a remarkable consistency: the value obtained is exactly the value that we would expect a particle to have when the particle is characterized in the intermediate times by the two wave-functions, one in a ground state, and the other localized outside the well. Indeed, we obtain this result precisely when we post-select the particle far enough from the well that it could not have been kicked there as a result of the intermediate measurement. A peculiar interference effect of the pointer takes place: destructive interference in the whole "allowed" region and constructive interference of the tails in the "forbidden" negative region. The initial state of the measuring device $\Phi(Q)$, due to the measuring interaction and the post-selection, transforms into a superposition of shifted wave functions. The shifts are by the (possibly small) eigenvalues, but the superposition is approximately equal to the original wave function shifted by a (large and/or forbidden) weak value:

$$\sum_i c_i \Phi(Q - a_i) \simeq \Phi(Q - A_w). \quad (22)$$

The example of a single weak measurement on the system of 20 pre- and post-selected spin-1/2 which was considered above demonstrates this effect for a Gaussian wave function of the measuring device, but we have proved [7] that the "miraculous" interference (22) occurs not just for the Gaussians, but for a large class of functions. The only requirement is that their Fourier transform must be essentially bounded.

It is possible to use this idea for constructing a quantum time machine, a device which can make a cat out of a kitten in a minute [7, 10]. The superposition of quantum states shifted by small periods of time can yield a large shift in time; and it even can be a shift to the past.

These surprising, even paradoxical effects are really gedanken experiments. The reason is that, unlike weak measurements on an ensemble, these are extremely rare events. For yielding an unusual weak value, a single pre-selected system needs an extremely improbable outcome of the post-selection measurement. Let us compare this with a weak measurement on an ensemble. In order to get N particles in a pre- and post-selected ensemble which yield $(\sigma_\xi)_w = 100$, we need $\sim N10^4$ particles in the pre-selected ensemble. But, in order to get a single system of N particles yielding $(S_\xi)_w = 100N$, we need $\sim 10^{4N}$ systems of N pre-selected particles. In fact, the probability to obtain an unusual value by error is much larger than the probability to obtain the proper post-selected state. What makes these rare effects interesting is that there is a strong (although only one-way) correlation: for example, every time we find in the post-selection measurement the particle sufficiently far from the well, we know that the result of the kinetic energy is negative, and not just negative: it is equal to the weak value, $K_w = E_0$, with a good precision.

5. Experimental realizations of weak measurements

Realistic weak measurements (on an ensemble) involve preparation of a large pre-selection ensemble, coupling to the measuring devices of each element of the ensemble, post-selection measurement which, in all interesting cases, selects only a small fraction of the original ensemble, selection of corresponding measuring devices, and statistical analysis of their outcomes. In order to obtain good precision, this selected ensemble of the measuring devices has to be sufficiently large. Although there are significant technological developments in "marking" particles running in an experiment, clearly the most effective solution is that the particles themselves serve as measuring devices. The information about the measured variable is stored, after the weak measuring interaction, in their other degree of freedom. In this case, the post-selection of the particles in the required final state automatically yields selection of measuring devices. The requirement for the post-selection measurement is, then, that there is no coupling between the variable in which the result of the weak measurement is stored and the post-selection device.

An example of such a case is the Stern-Gerlach experiment where the shift in the momentum of a particle, translated into a spatial shift, yields the outcome of the spin measurement. Post-selection measurement of a spin component in a certain direction can be implemented by another (this time strong) Stern-Gerlach coupling which splits the beam of the particles. The beam corresponding to the desired value of the spin is then analyzed for the result of the weak measurement. The requirement of non-disturbance of the results of the weak measurement by post-selection can be fulfilled by arranging the shifts due to the two Stern-Gerlach devices to be orthogonal to each other. The details are spelled out in Ref. [6].

An analysis of realistic experiment which can yield large weak value Q_w appears in Ref. [11]. Duck, Stevenson and Sudarshan [12] proposed slightly different optical realization which uses birefringent plate instead of a prism. In this case the measured information is stored directly in the spatial shift of the beam without being generated by the shift in the momentum. Ritchie, Story and Hulet adopted this scheme and performed the first successful experiment measuring weak value of the polarization operator [13].

Their results are in very good agreement with theoretical predictions. They obtained weak values which are very far from the range of the eigenvalues, $(-1, 1)$, their highest reported result is $Q_w = 100$. The discrepancy between calculated and observed weak value was 1%. The RMS deviation from the mean of 16 trials was 4.7%. The width of the probability distribution was $\Delta = 1000$ and the number of pre- and post-selected photons was $N \sim 10^8$, so the theoretical and experimental uncertainties were of the same order of magnitude. Their other run, for which they showed experimental data on graphs (which fitted very nicely theoretical graphs), has the following characteristics: $Q_w = 31.6$, discrepancy with calculated value 4%, the RMS deviation 16%, $\Delta = 100$, $N \sim 10^5$.

Another system which is a good candidate for weak measurements, due to a well developed technology of preparation and selection of various quantum states, is a Rydberg two-level atom. Between the pre- and post-selection the atom can have weak coupling with a resonant field in a microwave cavity [14, 15].

There are many experiments measuring escape time of tunneling particles. Tunneling is a pre- and post-selection experiment: a particle is pre-selected inside the bounding potential and post-selected outside. Recently, Steinberg [16] suggested that many of these experiments are indeed weak measurements.

6. Conclusions

The two basic elements of our approach were investigated separately. The theory of "unsharp" measurements by Bush [17] has the element of weakness of the interaction. Popular today, the "consistent histories" approach [18], originated by Griffiths [19], includes the idea of pre- and post-selection. But it is the combination of the two which created our formalism. It allowed us to see peculiar features of quantum systems. The quantum time machine, the method of increasing sensitivity using post-selection, and other surprising phenomena were inaccessible within the framework of the standard formalism. Neither consistent histories nor unsharp measurements provided tools to see these effects, although they might be helpful for analyzing these phenomena [20]. The concept of weak values, however, is simple and universal. Weak values are defined for all variables and for all possible histories of quantum systems. They manifest themselves in all couplings which are sufficiently weak. We have also shown how certain systems in nature automatically perform post-selections and thereby manifest the weak values [21].

The formalism of weak measurement can also be helpful in describing existing peculiar effects. The controversy of superluminal motion of tunneling particles can be resolved by recognizing that the experiments showing superluminal motion are weak measurements [16]. We have shown how, under conditions of weak measurements, the post-selection leads to superluminal motion of light wave packets (Section VIII of Ref. [3]).

Among applications of the weak value concept is a proposal to study the back reaction of a quantum field on a particle-antiparticle pair created by the field [22]. The weak value of the field is considered between the (initial) vacuum state and the (final) state which includes the particle-antiparticle pair. The aim of this proposal is the analysis of particle creation by a black hole and the problem of what happens in the final stages of black hole evaporation [23, 24]. One key to this problem is the back reaction of the pair to the gravitational field that created it, and here the application of weak values signals the possibility of a major breakthrough.

Another application is to use the weak measurement as an amplification scheme for some parameter in the measuring device, rather than as a measurement of the system [3]. Indeed, when we consider known initial state $|\Psi_1\rangle$ and known final state $|\Psi_2\rangle$, the weak value (4) is known prior to the measurement, and our experiment yields no new information. But we can perform the weak measuring procedure when the strength of the weak coupling is not known. Then, from the result of the weak measurement we can find the strength of the coupling.

We would like to end the article with a more speculative proposal for generalizing QM from the perspective of the two-state vector theory. Up to now we have considered the possibility of assigning two boundary conditions due to selections made before and after a measurement. In standard QM, it is assumed that a wave-function for a system exists even if we do not perform a measurement and therefore do not know what it is. In the same way, it is feasible and even suggestive to consider an extension of QM to include both a wave-function coming from the past and a second wave-function coming from the future which are determined by two boundary conditions, rather than a measurement and selection. This proposal might solve the issue of the "collapse" of the wave-function in a new and more natural way.

Acknowledgements

We are grateful to Jeff Tollaksen who helped us to prepare this paper on the basis of the contribution of one of us (LV) to the Erice School on Fundamental Aspects of Quantum Theory [25]. This research was supported in part by grant PHY-9601280 of the National Science Foundation and by grant 614/95 of the Israel Science Foundation.

References

1. Aharonov, Y., Bergmann, P. G. and Lebowitz, J. L., Phys. Rev. **B134**, 1410 (1964).
2. Aharonov, Y., Albert, D., Casher, A. and Vaidman, L., Phys. Lett. **A124**, 199 (1987).
3. Aharonov, Y. and Vaidman, L., Phys. Rev. **A41**, 11 (1990).
4. Aharonov, Y. and Vaidman, L., J. Phys. **A24**, 2315 (1991).
5. von Neumann, J., "Mathematical Foundations of Quantum Theory" (Princeton University Press, New Jersey 1983).
6. Aharonov, Y., Albert, D. and Vaidman, L., Phys. Rev. Lett. **60**, 1351 (1988).
7. Aharonov, Y., Anandan, J., Popescu, S. and Vaidman, L., Phys. Rev. Lett. **64**, 2965 (1990).
8. Unruh, W. G., "Varieties of Quantum Measurements", Fundamental Problems of Quantum Theory, (Edited by D. Greenberger) NYAS, **755** (1994).
9. Aharonov, Y., Popescu, S., Rohrlich, D. and Vaidman, L., Phys. Rev. **A48**, 4084 (1993); Jap. J. App. Phys. **9**, 41 (1993).
10. Vaidman, L., Found. Phys. **21**, 947 (1991).
11. Knight, J. M. and Vaidman, L., Phys. Lett. **A143**, 357 (1990).
12. Duck, M., Stevenson, P. M. and Sudarshan, E. C. G., Phys. Rev. **D40**, 2112 (1989).

13. Ritchie, N. W. M., Story, J. G. and Hulet, R. G., Phys. Rev. Lett. **66**, 1107 (1991); Story, J. G., Ritchie, N. W. M. and Hulet, R. G., Mod. Phys. Lett. **B5**, 1713 (1991).
14. Aharonov, Y., Davidovich, L. and Zagury, N., Phys. Rev. **A48**, 1687 (1993).
15. Zagury, N. and de Toledo, Piza, A. F. R., Phys. Rev. **A50**, 2908 (1994).
16. Steinberg, A. M., Phys. Rev. Lett. **74**, 2405 (1995).
17. Bush, P., Phys. Rev. **D33**, 2253 (1986).
18. Gell-Mann, M. and Hartle, J. B., Phys. Rev. **D47**, 3345 (1993).
19. Griffiths, R. B., J. Stat. Phys. **70**, 2201 (1993).
20. Bush, P., Phys. Lett. **A130**, 323 (1988).
21. Aharonov, Y., Massar, S., Popescu, S., Tollaksen, J. and Vaidman, L., Phys. Rev. Lett. **77**, 983 (1996).
22. Brout, R., Massar, S., Popescu, S., Parentani, R. and Spindel, Ph., Phys. Rev. **D52**, 1119 (1995).
23. Englert, F., Masaar, S. and Parentani, R., Class. Quant. Grav. **11**, 2919 (1994).
24. Stephens, C. R., 't Hooft, G. and Whiting, B. F., Class. Quant. Grav. **11**, 621 (1994).
25. Vaidman, L., in "Advances in Quantum Phenomena" (Edited by E. Beltrametti and J. M. Levy-Leblond) NATO ASI Series B: Physics (Plenum Press, NY 1995) Vol. 347, pp. 357–373.