

How the Result of a Measurement of a
Component of the Spin of a Spin- $\frac{1}{2}$
Particle Can Turn Out to be 100

Aharonov, Albert and Vaidman

April 4, 1988 PRL

Abstract

We have found that the usual measuring procedure for preselected and postselected ensembles of quantum systems gives unusual results. Under some natural conditions of weakness of the measurement, its result consistently defines a new kind of value for a quantum variable, which we call the weak value. A description of the measurement of the weak value of a component of a spin for an ensemble of preselected and postselected spin- $\frac{1}{2}$ particles is presented.

This paper will describe an experiment which measures a spin component of a spin- $\frac{1}{2}$, particle and yields a result which is far from the range of “allowed” values. We shall start with a brief description of the standard measuring procedure. Considering measurements on an ensemble of preselected and postselected systems, we shall define a new concept: *a weak value of a quantum variable*. And, finally, we shall describe the measurement of the weak value on the example of a spin- $\frac{1}{2}$, particle.

In quantum theory, the result of a measurement of a variable A which has discrete eigenvalues a_i must necessarily be one of those values. The Hamiltonian of the standard measurement procedure[1,2] is

$$H = -g(t)qA \quad (1)$$

where $g(t)$ is a normalized function with a compact support near the time of measurement, and q is a canonical variable of the measuring device with a conjugate momentum π . The initial state of the measuring device in the ideal case has to be such that π is well defined. After the interaction (1) we can ascertain the value of A from the final value of $\pi : A = \delta\pi$.

As a reasonable approximation for a real situation, we may take the initial state of the measuring device as a Gaussian in the q (and consequently also in the π) representation. For this case, the Hamiltonian (1) leads to the transformation

$$e^{-i \int H dt} e^{-\pi^2/4(\Delta\pi)^2} \sum_i \alpha_i |A = a_i\rangle = \sum_i \alpha_i e^{-(\pi-a_i)^2/4(\Delta\pi)^2} |A = a_i\rangle \quad (2)$$

where $\sum_i \alpha_i |A = a_i\rangle$ is the initial state of our system. If the spread of the π distribution $\Delta\pi$ is much smaller than the differences between the a_i , then,

after the interaction, we shall be left with the mixture of Gaussians located around a_i correlated with different eigenstates of A . A measurement of π will then indicate the value of A .

In the opposite limit, where $\Delta\pi$ is much bigger than all a_i , the final probability distribution will be again close to a Gaussian with the spread $\Delta\pi$. The center of the Gaussian will be at the mean value of A :

$$\langle A \rangle = \sum_i |\alpha_i|^2 a_i$$

One measurement like this will give no information because $\Delta\pi \gg \langle A \rangle$; but we can make this same measurement on each member of an ensemble of N particles prepared in the same state, and that will reduce the relevant uncertainty by the factor $1/\sqrt{N}$, while the mean value of the average will remain $\langle A \rangle$. By enlarging the number N of particles in the ensemble, we can make the measurement of $\langle A \rangle$ with any desired precision.

The outcome of the measurement is the average of the obtained values π of the measuring devices. As we explained earlier, it will yield, for a sufficiently large ensemble, the value $\langle A \rangle$. We now raise the question: Can we change the above outcome by taking into account the values of π of only a part of the original ensemble, performing a particular postselection? We may, of course, achieve this rather trivially, by selecting only measuring devices with large values of π which we can always find, since the original distribution of π has nonvanishing tails. But suppose we allow only postselection performed on the particles themselves; how then can we maximize the outcome for the average of π ? It might appear at first that the best method for this will be to select all particles for which the final state corresponds to the eigenvalue a_{max} . But this is not the case. Surprisingly, we found that by making other postselections we can obtain much bigger outcomes.

Indeed, we shall now show that the above measurements (with large $\Delta\pi$, when applied to preselected and postselected ensembles, may yield new values which lie outside the “allowed” range, i.e., outside the interval $[\min(a_i), \max(a_i)]$. The procedure of the measurement is as follows. We start with a large ensemble of particles prepared in the same initial state. Every particle interacts with a *separate* measuring device, and then the measurement which selects the final state is performed. Finally, we take into account only the “readings” of the measuring devices corresponding to the postselected particles.

Let us consider an ensemble of particles with an initial state $|\psi_{in}\rangle$ and a final state $|\psi_f\rangle$. At a time in between we switch on the interaction (1) where the initial state of each of the measuring devices is

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

After the postselection the state of the measuring device (up to a normalization factor) is

$$\langle\psi_f|e^{-i\int H dt}|\psi_{in}\rangle \exp\left(-\frac{q^2}{4\Delta^2}\right) \cong \langle\psi_f|\psi_{in}\rangle \exp\left(iq\frac{\langle\psi_f|A|\psi_{in}\rangle}{\langle\psi_f|\psi_{in}\rangle}\right) \exp\left(-\frac{q^2}{4\Delta^2}\right) \quad (3)$$

This formula is valid if the spread Δ is sufficiently small[3]:

$$\Delta \ll \max_n \frac{|\langle\psi_f|\psi_{in}\rangle|}{(|\langle\psi_f|A^n|\psi_{in}\rangle|)^{1/n}} \quad (4)$$

In the π representation, the state of the measuring device is approximately

$$\exp\left[-\Delta^2\left(\pi - \frac{\langle\psi_f|A|\psi_{in}\rangle}{\langle\psi_f|\psi_{in}\rangle}\right)^2\right] \quad (5)$$

The standard interpretation of this outcome of the measuring device is that the measured value of A is

$$\frac{\langle\psi_f|A|\psi_{in}\rangle}{\langle\psi_f|\psi_{in}\rangle}$$

This is the *weak value* of A for this preselected and postselected ensemble:

$$A_w \equiv \frac{\langle\psi_f|A|\psi_{in}\rangle}{\langle\psi_f|\psi_{in}\rangle} \quad (6)$$

The uncertainty of π for each of the measuring devices much bigger than the measured value; i.e.,

$$\Delta\pi = \frac{1}{2\Delta} \gg A_w$$

[see Eq. (4)]. However, for an ensemble of N such devices, the uncertainty of the average of π is decreased by the factor of $1/\sqrt{N}$. Therefore, if N is sufficiently large, then $(1/\sqrt{N})\Delta\pi \ll A_w$ and the value of A_w can be

ascertained with arbitrary accuracy. As we see from the definition (6). A_w is not bounded by the minimal and the maximal eigenvalues of A [4].

One may wonder how a superposition of shifts, all smaller than a_{max} leads to a shift which is much larger than a_{max} . We shall demonstrate this by proving a mathematical identity which corresponds to the type of experiment considered in the present note (see also Aharonov et al.[5]).

For all functions $f(q)$ which tend to zero faster than exponential, $f(q) < \exp(-\ln \alpha^2|q|)$, the following equality can be made valid with any precision by our taking N large enough[6]:

$$\sum_{-N}^N c_n e^{iqn/N} f(q) \cong e^{i\alpha q} f(q) \quad (7)$$

where

$$c_n = \frac{N!}{n!(N-n)!} \left(\frac{1-\alpha^2}{4}\right)^N \left(\frac{1+\alpha}{1-\alpha}\right)^{n/2}$$

The Fourier transform of this equation exhibits the property described above: For function $\tilde{f}(\pi)$ [the Fourier transform of $f(q)$] the superposition of shifts, which are all smaller than 1, is equivalent to a shift by the arbitrarily large value α , i.e.,

$$\sum_{-N}^N c_n \tilde{f}(\pi - n/N) \cong \tilde{f}(\pi - \alpha) \quad (8)$$

We shall now describe an experiment that measures the weak value of the z component of a spin- $\frac{1}{2}$ particle and yields an arbitrarily large result for it. A version of this experiment can, we believe, be performed in the laboratory.

We start with a beam of particles moving in the y direction with a well-defined velocity. The particles are initially well localized in the xz plane and have their spins pointed in a direction $\hat{\xi}$. We choose $\hat{\xi}$ in the xz plane with an angle α between $\hat{\xi}$ and \hat{x} (Fig. 1).

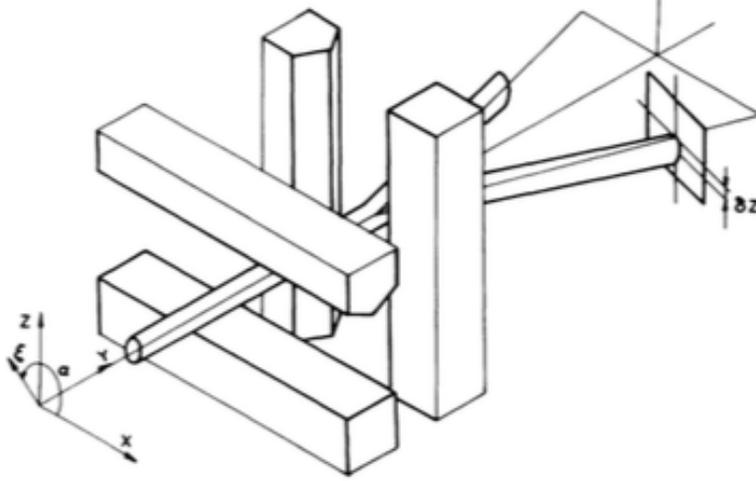


Figure 1: The experimental device for measurement of the weak value of σ_z . The beam of particles with the spin pointed in the direction $\hat{\xi}$ passes through an inhomogeneous (in the z direction) weak magnetic field and is split by the strong magnet with an inhomogeneous field in the x direction. The beam $\sigma_x = 1$, comes toward the screen and the deflection of the spot in the z direction is proportional to the weak value of σ_z : $\sigma_{z_w} = (\delta z p_0 \mu / l) (\partial B_z / \partial z)^{-1}$.

The prepared beam comes through a Stern-Gerlach device which measures the spin weakly in the z direction. The prepared beam comes through a Stern-Gerlach device which measures the spin weakly in the z direction. The requirement of weakness is fulfilled by our making the gradient of the magnetic field sufficiently small. The motion of the beam changes, therefore, only slightly. This weak measurement causes the spatial part of the wave function to change into a mixture of two slightly shifted functions in the p_z representation, correlated to the two values of σ_z . We then pass the particles through another, normal, Stern-Gerlach device which splits them into two beams corresponding to the two values of σ_x . We keep only the beam with $\sigma_x = 1$, which continues to move freely towards a screen placed in front of it. The screen is placed sufficiently far so that the displacement in the \hat{z} direction due to the average momentum p_z acquired during the above weak interaction will be larger than the initial uncertainty Δz . On the screen we shall obtain a wide spot whose displacement in the direction \hat{z} is measured.

This displacement will yield the weak value of σ_z :

$$\sigma_{z,w} = \frac{\langle \uparrow_x | \sigma_z | \uparrow_\xi \rangle}{\langle \uparrow_x | \uparrow_\xi \rangle} = \tan \frac{1}{2} \alpha \quad (9)$$

A brief mathematical description of this experiment follows. The particles have mass m , magnetic moment μ , an average momentum p_0 in the y direction. Their initial state is

$$|\psi_{in}\rangle = \Delta^{-3/2} (2\pi)^{-3/4} e^{-x^2/4\Delta^2} e^{-z^2/4\Delta^2} e^{-y^2/4\Delta^2} e^{-ip_0 y} (\cos(\alpha/2) |\uparrow_x\rangle + \sin(\alpha/2) |\downarrow_x\rangle) \quad (10)$$

The Hamiltonian of the weak interaction is

$$H_1 = \quad (11)$$

where $g(y - y_i)$ has a compact support at the location of the weak Stern-Gerlach device, which is arranged such that $(m/p_0) \int g(y) dy = 1$. It is, indeed, essentially a Hamiltonian of the von Neumann type [see Eq. (1)]. Since $y \cong (p_0/m)t$, $g(y - y_i)$ is effectively a function of time (we arrange the momentum in the y direction such that $p_0 \gg \Delta p_y = 1/2\Delta$; the canonical variable q of Eq. (1) is, here, $\mu(\partial B_z/\partial z)z$. The change in the momentum in the z direction during the interaction is $\delta p_z = \mu(\partial B_z/\partial z)\sigma_z$. For our ensemble of preselected and postselected particles, we shall see that $\delta p_z = \mu(\partial B_z/\partial z) \tan(\alpha/2)$. The necessary requirement of the weakness of the interaction [see Eq. (4)] is

$$\mu \left| \frac{\partial B_z}{\partial z} \right| \max \left[\left| \tan \frac{\alpha}{2} \right|, 1 \right] \ll \Delta p_z = \frac{1}{2\Delta} \quad (12)$$

The Hamiltonian of the second Stern-Gerlach device, which selects the particles with $\sigma_x = 1$, is

$$H_2 = -\mu \frac{\partial B_x}{\partial x} x \sigma_x g(y - y_2) \quad (13)$$

The requirement for the splitting of the beam is $\mu |\partial B_x/\partial x| \gg \Delta p_x = 1/2\Delta$. If the angle α between the directions of spin in the initial and the final states is close to π , then the requirement has to be stronger, i.e.,

$$\exp \left(-2\Delta^2 \mu^2 \left| \frac{\partial B_x}{\partial x} \right|^2 \right) \ll \cot^2(\alpha/2)$$

The direction of the beam with $\sigma_x = 1$ is $\mu(\partial B_x/\partial x)\hat{x} + p_0\hat{y}$. The screen is placed at the distance l such that

$$\frac{l\mu}{p_0} \left| \frac{\partial B_x}{\partial x} \right| \tan \frac{\alpha}{2} \gg \Delta$$

The wave function of the particle in the z representation, before the collapse on the screen, is approximately

$$\exp \left[-\Delta^2 \left(\frac{p_0}{l} \right)^2 \left(z - \frac{l\mu}{p_0} \frac{\partial B_z}{\partial z} \tan \frac{\alpha}{2} \right)^2 \right] \quad (14)$$

The measured value of the spin component corresponding to this wave function is $\tan(\alpha/2)$. A rough estimate for the number N of particles in the initial beam necessary for our obtaining accuracy of $1/M$ is $N \cong M^3 / \cos^2(\alpha/2)$.

In the above description we have not, so far, taken into account the influence of $\partial B_x / \partial x$ and $\partial B_y / \partial y$ which cannot both vanish since $\text{div} \mathbf{B}$. In a standard Stern-Gerlach experiment, one takes care of this problem by adding a large constant magnetic field in the direction of the measured component of the spin. It is interesting to note that the same method can be used in our weak measurement. The strong magnetic field in the z direction will, of course, rotate the direction of the spin and, therefore, during the interaction (11) we have to calculate the weak value sandwiched not between the states $|\uparrow_\xi\rangle$ and $|\uparrow_x i\rangle$ [see Eq. (6)], but between the rotated states. It turns out, however, that if we arrange the magnetic field in the z direction, such that it rotates the spin during the interaction by $2n\pi$, then the weak value of σ_z , during the whole period of the interaction does not change.

Another striking aspect of this experiment becomes evident when we consider it as a device for measuring a small gradient of the magnetic field $\partial B_z / \partial z$. Our choosing α close to π yields a tremendous amplification. Indeed, any weak coupling with the spin of the particles from our preselected and postselected ensemble will be amplified by the factor $[\cos(\alpha/2)]^{-1}$.

This research was supported in part by the National Science Foundation under Grant No. PHY 8408265.

[1] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer-Verlag, Berlin, 1932) [English translation: *Mathematical Foundations of Quantum Mechanics* (Princeton Univ. Press, Princeton, N3, 1955)]

[2] This Hamiltonian (1) is the effective Hamiltonian Stern-Gerlach measuring device as will be explained later [see Eq. (11)].

[3] If condition (4) is valid, then

$$\begin{aligned}
& \langle \psi_f | e^{-i \int H dt} | \psi_{in} \rangle e^{-q^2/4\Delta^2} \\
&= \sum_{n=0}^{\infty} \frac{(iq)^n}{n!} \langle \psi_f | A^n | \psi_{in} \rangle e^{-q^2/4\Delta^2} \\
&\cong \langle \psi_f | \psi_{in} \rangle \left(1 + \frac{iq \langle \psi_f | A | \psi_{in} \rangle}{\langle \psi_f | \psi_{in} \rangle} \right) e^{-q^2/4\Delta^2} \\
&\cong \langle \psi_f | \psi_{in} \rangle \exp \left(iq \frac{\langle \psi_f | A | \psi_{in} \rangle}{\langle \psi_f | \psi_{in} \rangle} \right) e^{-q^2/4\Delta^2}
\end{aligned}$$

[4] A_w may even be a complex number. The measurement of π gives only the real part $\text{Re}A_w$. In our case, in which the initial state of the measuring device is a Gaussian, we can also find the imaginary part $\text{Im}A_w$ by measuring the canonical variable q itself. Indeed, in the q representation the state of the measuring device is

$$e^{iq\text{Re}A_w} \exp \left(-\frac{1}{4\Delta^2} (q^2 + 2\Delta^2 \text{Im}A_w)^2 \right)$$

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

[6] The coefficients c_n , are derived from the binomial expansion. We can rewrite the left-hand side of (7), the part which multiplies $f(q)$ as

$$\begin{aligned}
& \sum_{n=-N}^N \frac{N!}{n!(N-n)!} \left(\frac{1-\alpha^2}{4} \right)^N \left(\frac{1+\alpha}{1-\alpha} \right)^{n/2} e^{iqn/N} \\
&= \left(\frac{1+\alpha}{2} e^{iq/2N} + \frac{1-\alpha}{2} e^{-iq/2N} \right)^{2N} \\
&= [\cos(q/2N) + i\alpha \sin(q/2N)]^{2N} \\
&= [1 - (1-\alpha^2) \sin^2(q/2N)]^N e^{i2N\phi}
\end{aligned}$$

where $\tan \phi = \alpha \tan(q/2N)$. For $f(q) < \exp(-\ln \alpha^2 |q|)$, we obtain

$$\begin{aligned}
& \lim_{N \rightarrow \infty} \sum_{n=-N}^N c_n e^{iq/2N} f(q) \\
&= \lim_{N \rightarrow \infty} [1 - (1-\alpha^2) \sin^2(q/2N)]^N e^{i2N\phi} f(q) \\
&= e^{i\alpha q} f(q)
\end{aligned}$$

This proves the statement in the text (7).