Introduction to Weak Measurements and Weak Values

Boaz Tamir [1] *& Eliahu Cohen* [2]

[1] *Faculty of Interdisciplinary Studies, S.T.S. program, Bar-Ilan University, Ramat-Gan, Israel. E-mail: boaz.tamir@mail.huji.ac.il*
[2] *School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel. E-mail: eliahuco@post.tau.ac.il*

We present a short review of the theory of weak measurement. This should serve as a map for the theory and an easy way to get familiar with the main results, problems and paradoxes raised by the theory.
Quanta 2013; 2: 7–17.

1 Introduction

Weak measurement is increasingly acknowledged as one of the most promising research tools in quantum mechanics. Tasks believed to be self contradictory by nature such as 'determining a particle's state *between* two measurements' prove to be perfectly possible with the aid of this technique. Similarly for the theory within which weak measurement has been conceived, namely the *two-state vector formalism*. Within this framework several new questions can be raised, for which weak measurement turns out again to be the most suitable tool for seeking answers. In the following we very briefly present some of the techniques, problems and paradoxes arising from the theory.

Weak measurements can reveal some information about the amplitudes of a quantum state without collaps-

ing the state into eigenvectors. This is done by a weak coupling between the measurement device and the system. Weak measurements generalize ordinary quantum (projective) measurements: following the weak coupling the state vector is not collapsed but biased by a small angle, and the measurement device does not show a clear eigenvalue, but a superposition of several values. One question is immediately raised: what type of information can weak measurement reveal?

While trying to answer the above question we realize that weak measurement theory presents a plethora of quantum strange phenomena, not yet completely understood. In the paper we shall present and discuss some of the paradoxes and peculiarities of the theory.

2 Weak measurements: the one-state vector formalism

We start with a short review of ordinary strong measurement, then we define the notion of weak measurement as a generalization of strong measurement. Next we describe some properties of weak measurement.

2.1 Strong quantum measurements

Let $|\psi\rangle$ be a pure vector state. Consider the standard formalism of quantum measurement theory. We differentiate between the outcome of the measurement m, the probability to get that measurement outcome $p(m)$, and the state

of the system after the measurement $|\psi_m\rangle$ [1–4].

We can describe the measurement process using the formalism of vector operators. Let $\hat{M}_m$ be a set of linear operators satisfying $\sum \hat{M}_m^\dagger \hat{M}_m = \hat{I}$ then we will say that the set defines a quantum measurement where the probability is defined by

$$p(m) = \langle \psi | \hat{M}_m^\dagger \hat{M}_m | \psi \rangle \tag{1}$$

and the vector state (non normalized) after the measurement is

$$|\psi_m\rangle = \hat{M}_m |\psi\rangle \tag{2}$$

If $\hat{M}_m$ are also projective operators i.e. $\hat{M}_m^\dagger \hat{M}_m = \hat{M}_m$ then we say that the measurement is projective.

Example 1

Consider the state

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \tag{3}$$

where $|0\rangle$ and $|1\rangle$ are the eigenvectors of the Pauli spin matrix in the z-direction

$$\hat{S}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4}$$

Measuring the state using the projectors $\hat{M}_0 = |0\rangle\langle 0|$, $\hat{M}_1 = |1\rangle\langle 1|$ gives the result 0 or 1 respectively with probability $\frac{1}{2}$. Following the measurement, the state of the system is $|0\rangle$ or $|1\rangle$ respectively.

In some cases it is enough to know the probabilities of the measurement outcome e.g. when the final state is not so important. In such cases it is enough to demand the existence of a set of positive operators $\hat{E}_m$ such that

$$\sum_m \hat{E}_m = \hat{I} \tag{5}$$

where the probability to get m is

$$p(m) = \langle \psi | \hat{E}_m | \psi \rangle \tag{6}$$

This is known as positive operator-valued measurement (POVM). As a general scheme for implementing a POVM one can couple the system with an ancilla which represents the environment, evolve the whole system and ancilla using a unitary operator, and then measure the ancilla space using a projective measurement. This produces a series of positive operators on the original system. The technique is frequently used in quantum operations and is also known as the operator-sum representation (see also [4, chapter 8]).

2.2 Weak measurements: definition

There are many motivations for generalizing quantum projective measurements into weak measurements: maintaining the initial state while gradually accumulating information [5], determining the system's state in-between two strong measurements [6] and revealing unusual weak values [7] are just three examples, others will be described below.

Weak measurement should be treated as a generalization of quantum strong measurement. In weak measurement theory both the system and the measurement device are quantum systems [8]. Weak measurement consists of two steps. In the first step we weakly couple the quantum measurement device to the quantum system. In the second step we strongly measure the measurement device. The collapsed state of the measurement device is referred to as the outcome of the weak measurement process. For a measurement to be weak, the standard deviation of the measurement outcome should be larger than the difference between the eigenvalues of the system. We will now describe this process in details. The procedure resembles the strong measurement process described in [1], however here we use a very weak entanglement between the system and the measurement device (see [8]).

Let $|\phi_d\rangle$ denote the wave function of the measurement device. When represented in the position basis it will be written as

$$|\phi\rangle = |\phi_d\rangle = \int_x \phi(x)|x\rangle dx \tag{7}$$

where x is the position variable of the measuring needle. Let $\hat{X}_d$ be the position operator such that $\hat{X}_d|x\rangle = x|x\rangle$ (here, we use $\hat{X}$ to distinguish the operator $\hat{X}$ from its eigenvector $|x\rangle$ and eigenvalue x, the same for $\hat{P}$, the subscript d is used for measuring device). We will also assume that $\phi(x)$ behaves normally around 0 with some variance $\Delta = \sigma^2$:

$$\phi(x) = (2\pi\sigma^2)^{-\frac{1}{4}} e^{-x^2/4\sigma^2} \tag{8}$$

We will later (strongly) measure $|\phi_d\rangle$ i.e. collapse the device's needle to get a value which is the weak measurement's outcome.

Let S denote our system to be measured. Suppose $\hat{A}$ is an Hermitian operator on the system S. Suppose $\hat{A}$ has N eigenvectors $|a_j\rangle$ such that $\hat{A}|a_j\rangle = a_j|a_j\rangle$.

Consider the general state vector $|\psi\rangle$ expressed in the eigenbasis of $\hat{A}$:

$$|\psi\rangle = \sum_j \alpha_j |a_j\rangle \tag{9}$$

Consider the interaction Hamiltonian $\hat{H}_{\text{int}}$ [8, chapter 7]

$$\hat{H} = \hat{H}_{\text{int}} = g(t)\hat{A} \otimes \hat{P}_d \tag{10}$$

Here $g(t)$ is a coupling impulse function satisfying

$$\int_0^T g(t)dt = 1 \qquad (11)$$

where T is the coupling time and $\hat{P}_d$ is the operator conjugate to $\hat{X}_d$ such that $[\hat{X}_d, \hat{P}_d] = \iota\hbar$.

We shall start the measuring process with the vector

$$|\psi\rangle \otimes |\phi(x)\rangle \qquad (12)$$

in the tensor space of the two systems. Then we apply the Hamiltonian

$$e^{-\iota\hat{H}t/\hbar}|\psi\rangle \otimes |\phi(x)\rangle \qquad (13)$$

It is easy to see that on each of the vectors $|a_j\rangle \otimes |\phi(x)\rangle$ the Hamiltonian $\hat{H}$ takes $\hat{X}_d$ to $\hat{X}_d + a_j$, (Heisenberg evolution)

$$\hat{X}_d(T) - \hat{X}_d(0) = \int_0^T dt \frac{\partial \hat{X}_d}{\partial t}$$
$$= \int_0^T \frac{\iota}{\hbar}[\hat{H}, \hat{X}_d]dt = a_j \qquad (14)$$

(see [9, section 8.4]). The corresponding transformation of the coordinates of the wave function is

$$e^{-\iota\hat{H}T/\hbar}|\psi\rangle \otimes |\phi\rangle = \sum_j \alpha_j|a_j\rangle \otimes |\phi(x - a_j)\rangle \qquad (15)$$

The above wave functions ϕ have high variance and therefore they overlap each other. The higher the variance the weaker the measurement process. If these normal wave functions do not overlap then the measurement is strong. Therefore we can control the process by the choice of the variance.

Example 2

Take $\hat{A} = \frac{\hbar}{2}\hat{S}_z$ and let $\hat{H}_{\text{int}} = \hat{A} \otimes \hat{P}_d$. Write $|\psi\rangle$ in the basis of eigenvectors of $\hat{S}_z$ i.e. $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$.

Following the weak coupling the system and the measuring device are entangled

$$\int_x [\alpha|0\rangle \otimes \phi(x - \hbar/2) + \beta|1\rangle \otimes \phi(x + \hbar/2)]|x\rangle dx \qquad (16)$$

where the above functions ϕ are two normal functions overlapping each other, due to their high variance.

2.3 Biasing the initial vector

In Example 2 we can write the entangled system and the measurement device as

$$\int_x [e^{-\frac{(x-\hbar/2)^2}{4\sigma^2}}\alpha|0\rangle \otimes |x\rangle + e^{-\frac{(x+\hbar/2)^2}{4\sigma^2}}\beta|1\rangle \otimes |x\rangle]dx \qquad (17)$$

We will now strongly measure the needle of the measuring device. Suppose the needle collapses to the vector $|x_0\rangle$, then our system is now in the state

$$[e^{-\frac{(x_0-\hbar/2)^2}{4\sigma^2}}\alpha|0\rangle + e^{-\frac{(x_0+\hbar/2)^2}{4\sigma^2}}\beta|1\rangle] \otimes |x_0\rangle \qquad (18)$$

without the integration. The eigenvalue x_0 could be anywhere around 0 or 1, or even further away especially if σ is big enough i.e. the measurement is very weak.

If the needle collapses to a value x_0 around 1 this means that the amplitude to post-select $|0\rangle$ is a little higher than the amplitude to post-select $|1\rangle$ and vice versa. So the collapse of the needle biases the system's vector. However if σ is very big with respect to the difference between the eigenvalues of $\hat{S}_z$ then the bias will be very small and the outcome system's vector will be very similar to the original vector.

To sum up, the system's vector is being biased a little in a direction that corresponds to the needle's outcome value. This is a generalization of strong measurements. On the one hand the information we get, that is the value of the needle, is very vague, and on the other hand the system's vector does not collapse but is being biased a little. If the measurement is getting stronger we will have a clearer value, very close to one of the eigenvalues of the system and an almost collapse of the system's vector i.e. a strong bias in the direction that corresponds to the needle's value. Note that there is always a correspondence between the value of the needle and the direction of the bias. This is the reason we can look at the weak measurement process as a generalization of the strong measurement one.

2.4 Computing the average of all eigenvalues

Let $\hat{A}$ be an Hermitian operator on the system S. Suppose $|a_j\rangle$ are eigenvectors for $\hat{A}$ with eigenvalues a_j. We will show that using weak measurements it is easy to compute the average of all eigenvalues [10].

Weakly couple the state $|\psi\rangle = \sum_j \alpha_j|a_j\rangle$ to a measuring device which has a normal distribution centered around 0. Our system is now described by

$$|\psi\rangle = (2\sigma^2\pi)^{-\frac{1}{4}} \sum_j \alpha_j|a_j\rangle e^{-(x-a_j)^2/4\sigma^2} \qquad (19)$$

Looking at the needle, the probability density to get x is

$$p(x) = (2\sigma^2\pi)^{-\frac{1}{2}} \sum_j |\alpha_j|^2 e^{-(x-a_j)^2/2\sigma^2} \qquad (20)$$

which is a multinormal distribution with many modes. Now, if σ is big enough with respect to the variance of eigenvalues, it is simple to show that

$$p(x) \approx (2\sigma^2\pi)^{-\frac{1}{2}} e^{-(x-\sum_j |\alpha_j|^2 a_j)^2/2\sigma^2} \qquad (21)$$

where the above normal distribution is centered around $\sum_j |\alpha_j|^2 a_j$ which is the average over all eigenvalues (use the fact that $e^{y/\sigma^2} \approx 1 + y/\sigma^2$ for $\sigma \gg \sqrt{y}$). We can then use the distribution of the needle to sample an estimate for the average or the sum of all eigenvalues.

2.5 The uncertainty relation

The argument for the uncertainty principle follows a discussion in [8]. Given a large ensemble of identical particles, suppose we measure part of them using an observable $\hat{X}$ and the other part with a conjugate observable $\hat{P}$. We will show that the variances ΔX and ΔP satisfy the inequality $\Delta X \Delta P > \hbar$ (see [4, p.89] and also subsection 3.9).

Suppose we want to weakly measure the position x of a particle. We shall use a probe particle as a measuring device. Consider

$$\hat{H}_{\text{int}} = -g(t)\hat{X} \otimes \hat{X}_d \qquad (22)$$

where X_d is the position coordinate of the probe particle. Hence the variance ΔP_d of the probe's momentum sets a lower bound for the variance ΔX of the system

$$\Delta X \geq \Delta P_d \qquad (23)$$

Alternatively, suppose we want to weakly measure the momentum p of the particle. Consider

$$\hat{H}_{\text{int}} = g(t)\hat{P} \otimes \hat{P}_d \qquad (24)$$

Hence the variance ΔX_d of the coordinate of the probe sets a lower bound for the variance ΔP of the momentum of the system

$$\Delta P \geq \Delta X_d \qquad (25)$$

Now given a large ensemble of identical particle, some will be weakly measured by $\hat{X}$, the others by $\hat{P}$. By the above two inequalities

$$\Delta P \Delta X \geq \Delta P_d \Delta X_d \qquad (26)$$

Hence the uncertainty principle for the system is deduced from the uncertainty for the probes. However the probes are being measured by strong measurements and clearly satisfy the uncertainty principle.

2.6 Physical realizations

We end this section with a few remarks regarding the realization of weak measurements. Many practical experiments such as [11–15] have been performed over the years to implement weak measurements. Although the setups are different, most of them share a common basis: inserting a small bias to the wave function (in position,

polarization, frequency etc.) which hardly affects its time evolution. When this bias is later detected, it reveals some information regarding the wave-function and enables to calculate the weak-value in retrospect. In [12] for example, the weak measurement is accomplished with a thin piece of birefringent calcite that slightly changes the polarization of the photons passing through by introducing a phase shift between the ordinary and extraordinary components of polarization. This phase shift depends on their transverse momentum and therefore upon dissolving the rays using a beam displacer and detecting them with a CCD camera their trajectories can be reconstructed.

Up to now, the experiments verify all theoretical predictions, but nevertheless, much effort is invested in order to produce more accurate and broader experimental setups.

3 Weak measurement and post-selection: the two-state vector formalism

In this section we define the notion of weak measurement with post-selection. Introducing post-selection into the theory of weak measurement results in many strange, interesting and sometimes puzzling phenomena such as huge, negative and even complex weak values, super-oscillations, amplification, negative probabilities etc. We start with some definitions.

3.1 Definitions

Let $\hat{A}$ be an Hermitian operator on the system S, let $|\psi_{\text{in}}\rangle$ and $|\psi_{\text{fin}}\rangle$ denote two state vectors in the Hilbert space of S, and $|\phi(x)\rangle$ the state vector of the needle. Suppose we weakly measure an ensemble of particles prepared in the state $|\psi_{\text{in}}\rangle$ using the Hamiltonian $\hat{H} = \hat{A} \otimes \hat{P}_d$. We will now post select the final vector $|\psi_{\text{fin}}\rangle$. The amplitude to get $|\psi_{\text{fin}}\rangle$ will be

$$\langle \psi_{\text{fin}}|e^{-\iota \hat{H} T/\hbar}|\psi_{\text{in}}\rangle \qquad (27)$$

It means that we start with several copies of our system, all are weakly measured by $\hat{H}$. We then select only those that are found to be in the direction of $|\psi_{\text{fin}}\rangle$ [8, 16, 17]. Practically and computationally post-selection could be viewed as the result of a strong measurement. If $\{|\psi_{\text{fin},i}\rangle\}_i$ is a basis for the Hilbert space of S, where $|\psi_{\text{fin},0}\rangle = |\psi_{\text{fin}}\rangle$ then consider the two operators

$$\hat{P}_1 = |\psi_{\text{fin}}\rangle\langle\psi_{\text{fin}}| \otimes \hat{I}_d \qquad (28)$$

$$\hat{P}_0 = \sum_{i \neq 0} |\psi_{\text{fin},i}\rangle\langle\psi_{\text{fin},i}| \otimes \hat{I}_d \qquad (29)$$

We will use the above two operators to strongly measure $|\psi_w\rangle = e^{-\iota \hat{H} T/\hbar} |\psi_{\text{in}}\rangle \otimes |\phi(x)\rangle$. We will only look at the vector states for the outcome 1, (see subsection 2.1)

$$|\psi_{\text{fin}}\rangle\langle\psi_{\text{fin}}|e^{-\iota \hat{H} T/\hbar}|\psi_{\text{in}}\rangle \otimes |\phi(x)\rangle \qquad (30)$$

We assume now that $\hat{P}_d$ is distributed around 0 with low variance. Therefore $\hat{X}_d$ will have high variance and the measurement will be weak. Hence the above vector is

$$|\psi_{\text{fin}}\rangle\langle\psi_{\text{fin}}|(1 - \iota \hat{A} \otimes \hat{P}_d T/\hbar)|\psi_{\text{in}}\rangle \otimes |\phi(x)\rangle$$
$$= |\psi_{\text{fin}}\rangle \otimes \langle\psi_{\text{fin}}|\psi_{\text{in}}\rangle(1 - \iota\langle \hat{A}\rangle_w \hat{P}_d T/\hbar)|\phi(x)\rangle$$
$$= |\psi_{\text{fin}}\rangle \otimes \langle\psi_{\text{fin}}|\psi_{\text{in}}\rangle e^{-\iota\langle \hat{A}\rangle_w \hat{P}_d T/\hbar}|\phi(x)\rangle \qquad (31)$$

where [18]:

$$\langle \hat{A}\rangle_w = \frac{\langle\psi_{\text{fin}}|\hat{A}|\psi_{\text{in}}\rangle}{\langle\psi_{\text{fin}}|\psi_{\text{in}}\rangle} \qquad (32)$$

We can now compute $p(1)$, i.e. the probability to get the outcome 1, by the post-selection due to strong measurement

$$p(1) = \langle\psi_w|\hat{P}_1^\dagger \hat{P}_1|\psi_w\rangle = |\langle\psi_{\text{fin}}|\psi_{\text{in}}\rangle|^2 \qquad (33)$$

Note that the normal probability distribution we get is centered around the weak value $\langle \hat{A}\rangle_w$. Note also that $|\langle\psi_{\text{fin}}|\psi_{\text{in}}\rangle|^2$ multiplies the normal distribution. This factor could be very small if $|\psi_{\text{fin}}\rangle$ and $|\psi_{\text{in}}\rangle$ are almost orthogonal and therefore the post-selection process could require the use of many particles.

Following the weak coupling, the tensor is described by Equation 15. The post-selection procedure reshuffles the Gaussian functions using new amplitudes that depend on $|\psi_{\text{in}}\rangle$ and $|\psi_{\text{fin}}\rangle$. This is the reason for the fact that all those exponents gather around a new value when we apply the post-selection.

3.2 Strange weak values

If we use post-selection it is easy to see that the needle's distribution could be centered at a value which is far from any of the eigenvalues of the operator $\hat{A}$ [8].

Example 3

Suppose $\hat{A} = \frac{\hbar}{2}\hat{S}_z$, $|0\rangle$ and $|1\rangle$ are the eigenvectors of $\hat{S}_z$. Let

$$|\psi_{\text{in}}\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \qquad (34)$$

$$|\psi_{\text{fin}}\rangle = \frac{1}{2}\left(2 + \sqrt{2}\right)^{\frac{1}{2}}|0\rangle + \frac{1}{2}\left(2 - \sqrt{2}\right)^{\frac{1}{2}}|1\rangle \qquad (35)$$

Then $\langle \hat{A}\rangle_w = (1 + \sqrt{2})\frac{\hbar}{2}$ (see [8, chapter 16]), which is more than twice the largest eigenvalue of $\hat{A}$.

In Example 3 the strange value of the needle was in the range of the variance of the needle, i.e. close to the eigenvalues of $\hat{A}$. However it is possible to construct an example where the strange value goes very far. From Equation 32 it can be noted that if $|\psi_{\text{fin}}\rangle$ and $|\psi_{\text{in}}\rangle$ are almost orthogonal this value could be large. In [19] a weak measurement of the spin of a particle gave a value of $100\frac{\hbar}{2}$.

The following example is a version of an example introduced in [20]. It shows that by post selecting a very rare vector we could get a very strange weak value. In other words these strange weak values appear with very low probability.

Consider $e^{-\iota \hat{S}_z \otimes \hat{L}/\hbar}$ where the momentum operator $\hat{L}$ represents a small shift of the needle's wave function. Let $|\psi_{\text{in}}\rangle = \alpha|0\rangle + \beta|1\rangle$. Then consider the equation of operators on the needle's space

$$e^{-\iota \hat{S}_z \otimes \hat{L}/\hbar}|\psi_{\text{in}}\rangle = \alpha|0\rangle e^{-\iota \hat{L}/\hbar} + \beta|1\rangle e^{\iota \hat{L}/\hbar} \qquad (36)$$

We shall now post select two vectors $\hat{U}|0\rangle$ and $\hat{U}|1\rangle$, where

$$\hat{U} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \qquad (37)$$

First compute (as operator equation on the needle's space)

$$\hat{U}^\dagger e^{-\iota \hat{S}_z \otimes \hat{L}/\hbar}|\psi_{\text{in}}\rangle = \frac{\alpha}{\sqrt{2}}(|0\rangle + |1\rangle)e^{-\iota \hat{L}/\hbar}$$
$$+ \frac{\beta}{\sqrt{2}}(|1\rangle - |0\rangle)e^{\iota \hat{L}/\hbar}$$
$$= \frac{1}{\sqrt{2}}(\alpha e^{-\iota \hat{L}/\hbar} - \beta e^{\iota \hat{L}/\hbar})|0\rangle$$
$$+ \frac{1}{\sqrt{2}}(\alpha e^{-\iota \hat{L}/\hbar} + \beta e^{\iota \hat{L}/\hbar})|1\rangle \qquad (38)$$

If we now post-select $|0\rangle$ the momentum operator on the needle is

$$(1 - \iota \hat{L}_0/\hbar) \qquad (39)$$

where

$$\hat{L}_0 = \frac{\alpha + \beta}{\alpha - \beta}\hat{L} \qquad (40)$$

Else, if we post-select $|1\rangle$ the momentum operator on the needle is

$$(1 - \iota \hat{L}_1/\hbar) \qquad (41)$$

where

$$\hat{L}_1 = \frac{\alpha - \beta}{\alpha + \beta}\hat{L} \qquad (42)$$

Suppose α is very close to β. If we post select $\hat{U}|0\rangle$ then $\hat{L}_0$ could be high above the eigenvalues of $\hat{S}_z$, however this post-selection will have a very low probability $|\alpha - \beta|^2$. Alternatively, the probability to post select $\hat{U}|1\rangle$ is high $|\alpha + \beta|^2$, but $\hat{L}_1$ is close to 0. To sum up, we could have a very low probability to get a distribution function (for the needle) which is shifted very far from the eigenvalues of the measured operator. The post-selection process sums the shifted normal wave function with some weights. Such sums could accumulate around a 'strange' value.

3.3 Super-oscillations

Consider the ensemble of particles

$$|\psi_{\text{in}}\rangle = 2^{-\frac{N}{2}} \bigotimes_{i=1}^{N} (|0_i\rangle + |1_i\rangle) \qquad (43)$$

$$|\psi_{\text{fin}}\rangle = \bigotimes_{i=1}^{N} (\alpha|0_i\rangle - \beta|1_i\rangle) \qquad (44)$$

where $|\alpha|^2 + |\beta|^2 = 1$. Let $\hat{A} = \sum \hat{S}_z/N$, then following the post-selection it is easy to see that the wave function of the needle is [8]:

$$2^{-\frac{N}{2}} \left(\alpha e^{-\iota \hat{P}_d/N\hbar} + \beta e^{\iota \hat{P}_d/N\hbar} \right)^N |\phi(x)\rangle \qquad (45)$$

If we apply these operators on the function $|\phi(x)\rangle$ one by one, we can see a superposition of shifted functions which are also weighted by polynomials in α and β. This resembles a random walk. In general, such expressions give rise to super-oscillations [21, 22].

Consider the following function

$$f(x) = \left(\frac{1+a}{2} e^{2\pi \iota x/N} + \frac{1-a}{2} e^{-2\pi \iota x/N} \right)^N \qquad (46)$$

where $a > 1$. If we take a very small x, $(x < \sqrt{N})$:

$$f(x) \approx e^{-2\pi \iota a x} \qquad (47)$$

The peculiar thing is that originally the expression for $f(x)$ contained only high wavelengths $(N > 1)$, however its global expression for small x has a wavelength $1/a$ which could be much smaller. Such super-oscillations do exist, however they last for a very small length $(x < \sqrt{N})$ since their length is exponentially proportional to their energy [22].

Super-oscillations in weak measurement theory could exist if $\hat{P}_d$ is small, and indeed we demand that $\hat{P}_d$ is distributed around 0 with a very low variance (weak interaction). Then the global shift of the needle which is the weak value corresponds to some super-oscillation.

3.4 The interpretation of post-selection

There is a question about the meaning of the weak value that we get following the post-selection [23]. Could it be that the measurement depends on a post selected vector i.e. a future event? Before the post-selection any measurement outcome looks random in between the eigenvalues of $\hat{A}$. Following the post-selection these results (more accurately small part of the results) seems to accumulate around a value. It looks as if this value suggests a physical entity, however the information concerning this entity is presented to us only *a posteriori*.

In [8] and [16] it was claimed that the correct way to interpret those weak values is in terms of a time symmetric theory. Quantum theory should be thought of as a time symmetric theory. A pre selected state is no more physical than a post selected state. Any physical theory should consist of an initial condition, an evolutionary process and a final condition. The evolutionary process could be looked at as a process that evolves the initial condition forward in time, but also as a process that evolves the final condition backwards in time. Looking at the quantum process as most physicists do, from the initial condition forward in time, inserts an artificial time direction into the theory, a direction which does not exists in its own in quantum theory. The evolution process of the quantum theory follows the Aharonov–Bergmann–Lebowitz formula [24]: given a non degenerate operator $\hat{A}$ with eigenvectors $|a_i\rangle$ and eigenvalues a_i, the probability to measure the eigenvalue a_j is

$$p(a_j) = \frac{|\langle\psi_{\text{fin}}|a_j\rangle\langle a_j|\psi_{\text{in}}\rangle|^2}{\sum_i |\langle\psi_{\text{fin}}|a_i\rangle\langle a_i|\psi_{\text{in}}\rangle|^2} \qquad (48)$$

It is easy to write the known S-matrices using such a formula while keeping the initial and final states. Therefore the Aharonov–Bergmann–Lebowitz formula generalizes the computation of amplitudes in the ordinary pre selected time antisymmetric quantum theory.

Note that the probability to get some value j could be 1 even if $|\psi_{\text{in}}\rangle$ is not an eigenvector of $\hat{A}$. In the denominator it could indeed be that all of the summands besides the jth cancel and therefore the quotient is 1 (see Example 4). A quick look at Equation 1 shows that this could not happen for strong projective measurements.

Whenever the Aharonov–Bergmann–Lebowitz formula produces an eigenvalue with certainty we can interpret it as an element of reality. This value is certain under the initial and final conditions that we picked. A related discussion regarding the different definitions for 'elements of reality' appears in [10].

3.5 Complex weak values

The weak value $\langle\hat{A}\rangle_w$ is a matrix coefficient and could have complex values. How do complex values effect the needle's wave function?

Theorem [Jozsa] [25]: Consider $\hat{H}_{\text{int}} = g(t)\hat{A} \otimes \hat{P}_d$ and let $|\alpha\rangle$ be the pointer's wave function following the post-selection

$$|\alpha\rangle = \langle\psi_{\text{fin}}|\psi_{\text{in}}\rangle e^{-\iota\langle\hat{A}\rangle_w \hat{P}_d}|\phi\rangle \qquad (49)$$

If $\langle \hat{A} \rangle_w = a + \iota b$, $\hat{M}$ any observable, then

$$\langle \hat{M} \rangle_f = \langle \hat{M} \rangle_i + \iota g a \langle \hat{P}_d \hat{M} - \hat{M} \hat{P}_d \rangle_i$$
$$+ g b (\langle \hat{P}_d \hat{M} + \hat{M} \hat{P}_d \rangle_i - 2 \langle \hat{P}_d \rangle_i \langle \hat{M} \rangle_i) \quad (50)$$

where

$$\langle \hat{M} \rangle_i = \frac{\langle \phi | \hat{M} | \phi \rangle}{\langle \phi | \phi \rangle} \quad (51)$$

$$\langle \hat{M} \rangle_f = \frac{\langle \alpha | \hat{M} | \alpha \rangle}{\langle \alpha | \alpha \rangle} \quad (52)$$

Proof: The proof is a simple use of the definition of $|\alpha\rangle$. The computation of $\langle \hat{M} \rangle_f$ is up to $O(g)$ where g is the coupling impulse function (Equation 11). Note also that $e^{-\iota \langle \hat{A} \rangle_w \hat{P}_d}$ is non-unitary (has complex arguments) and $\langle \alpha | \alpha \rangle = 1 + 2 g b \langle \hat{P}_d \rangle_i$.

Corollary: If $\langle \hat{A} \rangle_w$ is imaginary, $\langle \hat{A} \rangle_w = \iota b$, then

$$\langle \hat{P}_d \rangle_f = \langle \hat{P}_d \rangle_i + 2 g b \mathrm{Var}(\hat{P}_d) \quad (53)$$

If $\langle \hat{A} \rangle_w$ is real, $\langle \hat{A} \rangle_w = a$, then

$$\langle \hat{X}_d \rangle_f = \langle \hat{X}_d \rangle_i + g a \quad (54)$$

So real values of $\langle \hat{A} \rangle_w$ shift the needle's expectation value of $\hat{X}_d$ and imaginary values of $\langle \hat{A} \rangle_w$ shift the needle's expectation value of $\hat{P}_d$.

3.6 Amplification

We can use weak measurements for an amplification process [26–28]. If $|\psi_{\mathrm{fin}}\rangle$ is almost orthogonal to $|\psi_{\mathrm{in}}\rangle$ then $|\langle \hat{A} \rangle_w|$ could be as large as we want. We could use this to amplify signals. There is a trade off between the amplification and the complexity of the post-selection process. When the two functions are almost orthogonal it would be hard to post select $|\psi_{\mathrm{fin}}\rangle$, however having post selected $|\psi_{\mathrm{fin}}\rangle$ the amplification factor is high. In [27] weak measurements were used to amplify a small momentum perturbation in a Sagnac interferometer. In the following we will briefly review the argument.

Let the initial vector be

$$|\psi_{\mathrm{in}}\rangle = \frac{1}{\sqrt{2}} \left(e^{-\iota \eta/2} | \circlearrowleft \rangle + \iota e^{\iota \eta/2} | \circlearrowright \rangle \right) \quad (55)$$

where η is some phase differentiating between the two beams. Let the final vector be

$$|\psi_{\mathrm{fin}}\rangle = \frac{1}{\sqrt{2}} \left(\iota | \circlearrowright \rangle + | \circlearrowleft \rangle \right) \quad (56)$$

Suppose we add a small amount of momentum k to the device. We want to measure its effect on the phase difference between the two beams. We can take any observable $\hat{A}$ that differentiates between the two paths and measure $k\hat{A}$:

$$\hat{A} = | \circlearrowright \rangle \langle \circlearrowright | - | \circlearrowleft \rangle \langle \circlearrowleft | \quad (57)$$

The interaction Hamiltonian is

$$\hat{H}_{\mathrm{int}} = g k \hat{A} \otimes \hat{X}_d \quad (58)$$

where the wave function of the needle will be written in the momentum representation.

It is easy to compute

$$\langle \hat{A} \rangle_w = \iota k \cot(\eta/2) \quad (59)$$

$$\langle \psi_{\mathrm{fin}} | \psi_{\mathrm{in}} \rangle = \sin^2(\eta/2) \quad (60)$$

Since $\langle \hat{A} \rangle_w$ is imaginary, we will look at the shift of $\langle \hat{X}_d \rangle$. By the above theorem of Jozsa (for the Hamiltonian Equation 58 where $\hat{X}_d$ is used instead of $\hat{P}_d$):

$$\langle \hat{X}_d \rangle_f = \langle \hat{X}_d \rangle_i + 2 g k \cot(\eta/2) \mathrm{Var}(\hat{X}_d) \quad (61)$$

Hence for a very low η the value of $\langle \hat{X}_d \rangle_i$ will be shifted away. Note however that for such small angles it will also be hard to post-select since $\langle \psi_{\mathrm{fin}} | \psi_{\mathrm{in}} \rangle = \sin^2(\eta/2)$.

Compare now the shift of the weak value of $\hat{X}_d$ with the phase shift we usually measure on a Sagnac interferometer as functions of the small momentum k added to the system. This will be the amplification factor of the signal. It looks as if this factor could be very large and even unlimited. In [28] it was shown that if we use the Jozsa expansion to all orders of g we can see the bound on the amplification factor.

3.7 Violation of the product rule

It is easy to see that a sum of weak values is a weak value of the sum of the measurement operators. However such a theorem is not true for the product of weak measurements. In standard quantum theory if a state is an eigenvector for two operators then the product of the eigenvalues is the eigenvalue of the product of the operators. But as we saw above weak values could be certain even if the initial state was not an eigenvector. The following example clarifies this fact.

Example 4: The shell game

Let

$$|\psi_{\mathrm{in}}\rangle = \frac{1}{\sqrt{3}} (|A\rangle + |B\rangle + |C\rangle) \quad (62)$$

$$|\psi_{\mathrm{fin}}\rangle = \frac{1}{\sqrt{3}} (|A\rangle + |B\rangle - |C\rangle) \quad (63)$$

Let $\hat{\mathbb{P}}_A = |A\rangle\langle A|$, $\hat{\mathbb{P}}_B = |B\rangle\langle B|$ and $\hat{\mathbb{P}}_C = |C\rangle\langle C|$. Each of the above projection operators has eigenvalues 0 or

1. These operators are degenerate and therefore to compute probabilities we use a version of the Aharonov–Bergmann–Lebowitz formula for degenerate operators. We conclude that measuring $\hat{\mathbb{P}}_A$ yields 1 with probability 1 [8]. The same is true for $\hat{\mathbb{P}}_B$: measuring $\hat{\mathbb{P}}_B$ yields 1 with probability 1. However measuring the product $\hat{\mathbb{P}}_A\hat{\mathbb{P}}_B$ yields 0 with probability 1. Hence the product rule in not true in the two-state vector formalism.

3.8 Negative probabilities

Consider a projective operator $\hat{\mathbb{P}}$. Its eigenvalues are 0 or 1, therefore its expectation value on a general state $|\psi\rangle$ could be anywhere between 0 and 1. We could think of such a value as the probability to be in the state defined by the projection.

Projection operators which are weakly measured could have other values. In Example 4 above $\langle\hat{\mathbb{P}}_C\rangle_w = -1$. What does it mean that the weak projection operator $\hat{\mathbb{P}}$ has a negative value? We could generalize the above argument saying that there is a negative probability to be projected into the state defined by $\hat{\mathbb{P}}$ [29], [30]. Let's dwell on this argument.

The initial wave function in Example 4 represents a superposition of three position states. Suppose we have a positive particle that could be in one of three positions, and the initial state is a superposition of those three states. Trying to verify if the particle is in position A is like measuring the projector $\hat{\mathbb{P}}_A$. The measuring device for each of the projectors could measure the transverse momentum of an electron while passing near the corresponding position

$$\hat{H}_{\text{int}} = \hat{X}_d \otimes \hat{\mathbb{P}}_A \tag{64}$$

If the positive particle is in one of the above positions then we expect a deflection of the electron towards that direction. However, computing the weak values of the projectors we get

$$\langle\hat{\mathbb{P}}_A\rangle_w = 1 \tag{65}$$
$$\langle\hat{\mathbb{P}}_B\rangle_w = 1 \tag{66}$$
$$\langle\hat{\mathbb{P}}_C\rangle_w = -1 \tag{67}$$

A negative value of $\hat{\mathbb{P}}_C$ corresponds to a deflection of the momentum wave function in the opposite direction, away from the particle. This looks very strange as if the electron is avoiding the positive charge. However experiments do show that the deflection is away from the positive particle [11]. This also looks as if the particle has changed its charge from positive to negative.

We could use this result to say that a negative probability could be interpreted as a positive probability for an opposite physical event.

Negative probability could also be used to explain paradoxical results that follow from the violation of the product rule. Consider the following example.

Example 5: The Cheshire cat experiment

A Cheshire cat experiment is a case where we can separate quantum variables by using projection operators. In this case we separate a cat from its grin [31].

Let $\hat{\mathbb{P}}_0 = |0\rangle\langle 0|$ denote a cat's grin, $\hat{\mathbb{P}}_1 = |1\rangle\langle 1|$ a cat's frown where $|0\rangle$ and $|1\rangle$ are the eigenvectors of the Pauli spin matrix $\hat{S}_z$. Let $\hat{\mathbb{P}}_A = |A\rangle\langle A|$ denote a position A, $\hat{\mathbb{P}}_B = |B\rangle\langle B|$ a position B. Let

$$|\psi_{\text{in}}\rangle = \frac{1}{2}(|0\rangle + |1\rangle)|A\rangle + \frac{1}{2}(|0\rangle - |1\rangle)|B\rangle \tag{68}$$

$$|\psi_{\text{fin}}\rangle = \frac{1}{2}(|0\rangle - |1\rangle)(|A\rangle + |B\rangle) \tag{69}$$

By simple computation we can see that $\langle\hat{\mathbb{P}}_A\rangle_w = 0$, $\langle\hat{\mathbb{P}}_B\rangle_w = 1$, so the cat is in position B. Also $\langle\hat{S}_z\rangle_w = 1$ so the cat is grinning. However $\langle\hat{S}_z\hat{\mathbb{P}}_B\rangle_w = 0$ and $\langle\hat{S}_z\hat{\mathbb{P}}_A\rangle_w = 1$ so the grin is in position A. Hence the cat is in position B and its grin is in position A. This strange result is an outcome of the violation of the product rule.

The Cheshire cat experiment could be interpreted using negative probabilities, i.e. negative values of projective operators. We compute the weak values of all the products of the above projectors

$$\langle\hat{\mathbb{P}}_A\hat{\mathbb{P}}_0\rangle_w = \frac{1}{2} \tag{70}$$

$$\langle\hat{\mathbb{P}}_A\hat{\mathbb{P}}_1\rangle_w = -\frac{1}{2} \tag{71}$$

$$\langle\hat{\mathbb{P}}_B\hat{\mathbb{P}}_0\rangle_w = \frac{1}{2} \tag{72}$$

$$\langle\hat{\mathbb{P}}_B\hat{\mathbb{P}}_1\rangle_w = \frac{1}{2} \tag{73}$$

Note that although some of the values are negative the cat's position and the grin's position have positive real values and therefore positive probability. Note that $\langle\hat{\mathbb{P}}_B\rangle_w = \langle\hat{\mathbb{P}}_B(\hat{\mathbb{P}}_0 + \hat{\mathbb{P}}_1)\rangle_w = 1$ and therefore the cat is in position B. Also $\langle(\hat{\mathbb{P}}_A + \hat{\mathbb{P}}_B)\hat{\mathbb{P}}_0\rangle_w = 1$ and therefore the cat is grinning. Moreover $\langle\hat{\mathbb{P}}_A\hat{\mathbb{P}}_0\rangle_w - \langle\hat{\mathbb{P}}_A\hat{\mathbb{P}}_1\rangle_w = 1$. This should be interpreted as the probability of a grin in position A, minus the probability of a frown in position A. If we interpret the probability of a frown as a negative probability of a grin then we can conclude that the grin is in position A.

Observe that although some of the probabilities above were negative, in each question having a real physical meaning the negative probabilities disappeared, and we were left with positive probability. Feynman used exactly

this property in his discussion on negative probabilities [29].

Negative probabilities were also used as an alternative approach to Bell's inequalities [32]. For a general review on extended probabilities see [33].

3.9 Paradoxes

There are several paradoxes that stem from weak measurements with post-selection. Some of them were already discussed above like the Cheshire cat paradox or the quantum shell game. There is also the faster than light paradox [34] and many others. Here we shall discuss a paradox that concerns the two slit experiment. We will show that it is possible to describe the path of the particle without disturbing the interference pattern [12, 13]. The pattern of the interference will be measured by a weak coupling while the path will be revealed *a posteriori*. The possibility of measuring the average trajectory without disturbing the interference does not in itself contradict the mathematical formulation of the Heisenberg principle. This formulation of the principle states that on a large ensemble of identical particles while measuring some of them with a variable $\hat{X}$ and the others with a variable $\hat{P}$ conjugate to $\hat{X}$, the product of the uncertainties satisfies $\Delta X \Delta P > \hbar$ [3, 4]. It is generally accepted that the correct interpretation of the Heisenberg principle is that the measurements of $\hat{X}$ do disturb the measurements of $\hat{P}$. Therefore we can say that weak measurement theory question the interpretation that the mere disturbance is the core of the principle.

Let $\hat{A}$ denote some weak measurement on a system S with initial sate $|\psi\rangle$. Let $|1\rangle, ..., |l\rangle$ denote an orthogonal basis for the system. Suppose $|\psi\rangle = \sum_k c_k |k\rangle$. Couple weakly the system to the measurement device. Now suppose we want to post select using the above basis. In fact, this is a strong projective measurement using the projectors $|k\rangle\langle k|$. We shall repeat the process N times for some large N, each time register the weak value

$$\langle \hat{A} \rangle_k = \frac{\langle k|\hat{A}|\psi\rangle}{\langle k|\psi\rangle} \qquad (74)$$

for the $|k\rangle$ that was post selected. Therefore we use all results of the strong post-selection. We will get N_k values of $\langle \hat{A} \rangle_k$ for each of the k's. Observe now that

$$\sum_k \frac{N_k}{N} \frac{\langle k|\hat{A}|\psi\rangle}{\langle k|\psi\rangle} \approx \sum_k |c_k|^2 \frac{\langle k|\hat{A}|\psi\rangle}{\langle k|\psi\rangle}$$

$$= \sum_k c_k^* \langle k|\hat{A}|\psi\rangle$$

$$= \langle \psi|\hat{A}|\psi\rangle \qquad (75)$$

We can therefore use the post-selected states to compute the average $\langle \psi|\hat{A}|\psi\rangle$ (see also subsection 2.4).

However the operator $\hat{A}$ and the measurement operator of the post-selection could represent two non-commuting observables. Consider for example the two slit experiment. Suppose we place a series of momentum detectors at the screen edge of the experiment. The detector can identify the transverse momentum of the particle and therefore identify the slit through which it passed. This in fact is a post-selection. We can identify *a posteriori* the path of the particle. Suppose now $\hat{A}$ enables a weak measurement of the projection of the position variable of the particle into a small area between the two slits and the screen (we have already used weak projectors in the sections above). The average of a projection operator could be considered as a density or amplitude. We could use the above process to compute $\langle \psi|\hat{A}|\psi\rangle$ which corresponds to the amplitude of the wave function of the particles in a small area. The measurement is weak and therefore does not interfere with the wave pattern. We can continuously change the position of $\hat{A}$, therefore $\hat{A}$ is like a 'weak screen' in between the two slits and the 'strong screen'. Therefore we can see the whole pattern of the interference and at the same time have a list of all trajectories which is a paradox.

It seems that by using the two-state vector formalism it is possible to know the average trajectory of the particle without disturbing the interference. Notice however that a trajectory is known *a posteriori* from the point of view of the one-state vector formalism [12].

Acknowledgment

References

[1] von Neumann J. Mathematical Foundations of Quantum Mechanics. Investigations In Physics, Beyer RT (translator), Princeton: Princeton University Press, 1955.

[2] Messiah A. Quantum Mechanics, Volume 1. North Holland Publishing Company, 1967. http://archive.org/details/QuantumMechanicsVolumeI

[3] Namiki M, Pascazio S, Nakazato H. Decoherence and Quantum Measurements. Singapore: World Scientific, 1997.

[4] Nielsen MA, Chuang IL. Quantum Computation and Quantum Information, 10th Anniversary edition. Cambridge: Cambridge University Press, 2011.

[5] Elitzur AC, Cohen E. The retrocausal nature of quantum measurement revealed by partial and weak measurements. AIP Conference Proceedings 2011; 1408 (1): 120-131. http://dx.doi.org/10.1063/1.3663720

[6] Aharonov Y, Cohen E, Grossman D, Elitzur AC. Can a future choice affect a past measurement's outcome? 2012; http://arxiv.org/abs/1206.6224

[7] Aharonov Y, Cohen E, Ben-Moshe S. Unusual interaction of a pre-and-post-selected particle. 2012; http://arxiv.org/abs/1208.3203

[8] Aharonov Y, Rohrlich D. Quantum Paradoxes: Quantum Theory for the Perplexed. Weinheim: Wiley-VCH, 2005.

[9] Peres A. Quantum Theory: Concepts and Methods. Fundamental Theories of Physics, Kluwer Academic Publishers, 2002.

[10] Vaidman L. Weak-measurement elements of reality. Foundations of Physics 1996; 26 (7): 895-906. http://arxiv.org/abs/quant-ph/9601005

[11] Resch KJ, Lundeen JS, Steinberg AM. Experimental realization of the quantum box problem. Physics Letters A 2004; 324 (2-3): 125-131. http://arxiv.org/abs/quant-ph/0310091

[12] Kocsis S, Braverman B, Ravets S, Stevens MJ, Mirin RP, Shalm LK, Steinberg AM. Observing the average trajectories of single photons in a two-slit interferometer. Science 2011; 332 (6034): 1170-1173. http://dx.doi.org/10.1126/science.1202218

[13] Rozema LA, Darabi A, Mahler DH, Hayat A, Soudagar Y, Steinberg AM. Violation of Heisenberg's measurement-disturbance relationship by weak measurements. Physical Review Letters 2012; 109 (10): 100404. http://dx.doi.org/10.1103/PhysRevLett.109.100404

[14] Pryde GJ, O'Brien JL, White AG, Ralph TC, Wiseman HM. Measurement of quantum weak values of photon polarization. Physical Review Letters 2005; 94 (22): 220405. http://arxiv.org/abs/quant-ph/0412204

[15] Götte JB, Dennis MR. Generalized shifts and weak values for polarization components of reflected light beams. New Journal of Physics 2012; 14 (7): 073016. http://dx.doi.org/10.1088/1367-2630/14/7/073016

[16] Aharonov Y, Vaidman L. The two-state vector formalism: an updated review. In: Time in Quantum Mechanics, vol.734, Muga JG, Mayato RS, Egusquiza ÍL (editors), Berlin, Heidelberg: Springer, 2007, pp.399-447. http://arxiv.org/abs/quant-ph/0105101

[17] Story JG, Ritchie NWM, Hulet RG. Weak measurements. Modern Physics Letters B 1991; 5 (26): 1713-1725. http://dx.doi.org/10.1142/S0217984991002069

[18] Vaidman L. Weak value and weak measurements. In: Compendium of Quantum Physics, Greenberger D, Hentschel K, Weinert F (editors), Berlin, Heidelberg: Springer, 2009, pp.840-842. http://arxiv.org/abs/0706.1348

[19] Aharonov Y, Albert DZ, Vaidman L. How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. Physical Review Letters 1988; 60 (14): 1351-1354. http://dx.doi.org/10.1103/PhysRevLett.60.1351

[20] Kempe J. Quantum random walks: An introductory overview. Contemporary Physics 2003; 44 (4): 307-327. http://arxiv.org/abs/quant-ph/0303081

[21] Ferreira PJSG, Kempf A. Superoscillations: Faster than the Nyquist rate. IEEE Transactions on Signal Processing 2006; 54 (10): 3732-3740. http://dx.doi.org/10.1109/TSP.2006.877642

[22] Aharonov Y, Colombo F, Sabadini I, Struppa DC, Tollaksen J. Some mathematical properties of superoscillations. Journal of Physics A 2011; 44 (36): 365304. http://dx.doi.org/10.1088/1751-8113/44/36/365304

[23] Vaidman L. Defending time-symmetrized quantum theory. 1996; http://arxiv.org/abs/quant-ph/9609007

[24] Aharonov Y, Bergmann PG, Lebowitz JL. Time symmetry in the quantum process of measurement. Physical Review B 1964; 134 (6): 1410-1416. http://dx.doi.org/10.1103/PhysRev.134.B1410

[25] Jozsa R. Complex weak values in quantum measurement. Physical Review A 2007; 76 (4): 044103. http://arxiv.org/abs/0706.4207

[26] Nishizawa A, Nakamura K, Fujimoto M-K. Weak-value amplification in a shot-noise-limited interferometer. Physical Review A 2012; 85 (6): 062108. http://arxiv.org/abs/1201.6039

[27] Dixon PB, Starling DJ, Jordan AN, Howell JC. Ultrasensitive beam deflection measurement via interferometric weak value amplification. Physical Review Letters 2009; 102 (17): 173601. http://arxiv.org/abs/0906.4828

[28] Koike T, Tanaka S. Limits on amplification by Aharonov-Albert-Vaidman weak measurement. Physical Review A 2011; 84 (6): 062106. http://arxiv.org/abs/1108.2050

[29] Feynman R. Simulating physics with computers. International Journal of Theoretical Physics 1982; 21 (6): 467-488-488. http://dx.doi.org/10.1007/BF02650179

[30] Zhu X, Wei Q, Liu Q, Wu S. How "negative probabilities" emerge in weak measurements? 2012; http://arxiv.org/abs/1212.6324

[31] Aharonov Y, Popescu S, Skrzypczyk P. Quantum Cheshire cats. 2012; http://arxiv.org/abs/1202.0631

[32] Cereceda JL. Local hidden-variable models and negative-probability measures. 2000; http://arxiv.org/abs/quant-ph/0010091

[33] Mückenheim W, Ludwig G, Dewdney C, Holland PR, Kyprianidis A, Vigier JP, Cufaro Petroni N, Bartlett MS, Jaynes ET. A review of extended probabilities. Physics Reports 1986; 133 (6): 337-401. http://dx.doi.org/10.1016/0370-1573(86)90110-9

[34] Rohrlich D, Aharonov Y. Cherenkov radiation of superluminal particles. Physical Review A 2002; 66 (4): 042102. http://arxiv.org/abs/quant-ph/0107025