Submitted in partial fullment of the requirements for the degree of Master of Science of Imperial College London

Weak Measurements

Author:

Angela Karanjai

Supervisor:

Prof. Terry Rudolph

September 24, 2013

Contents

1 Introduction

The Weak Measurement idea was first introduced by Y.Aharonov et al in 1988[2]. It was used to show how post selection could dramatically change the outcome of a measurement of an operator . For example, when measuring the spin of a spin 1/2 particle in the $\hat{z}$ direction, if only the measurement outcomes of the states, which are post selected to have spin +1/2 in the $\hat{x}$ direction are considered, then it is possible for the average outcome of the measurements of the $\hat{z}$ component of the spin to be far greater than 1/2.

A Weak Measurement is not a single measurement but in fact a series of measurements on a pre selected quantum state. Each measurement is "weak" due to a large uncertainty in the measuring device. Due to the uncertainty, the outcome of the measurement cannot be resolved in one measurement and hence a large number of measurements are carried out on separate copies of the pre selected quantum state in order to accurately determine the measurement outcome. However, only the measurement outcomes of the post selected states are used to calculate the average value of the measurement outcomes.

In this report, I will review Weak measurements particularly focussing on how one must interpret the results correctly. To this aim, I shall discuss its strange results and how they might be due to a misinterpretation. Finally I shall investigate the weak value of Gaussian quantum states using

ERL mechanics. Despite this result being limited to Gaussian quantum states, it aims to offer a more intuitive interpretation of the results of weak measurements.

2 Weak Measurements

In this chapter, I shall introduce the general Weak Measurement protocol. This will lead us to the definition of the Weak value of an operator $\langle \hat{A} \rangle_W$, which is effectively the measured quantity in a Weak Measurement.

2.1 The Weak Measurement Protocol

In order to perform the Weak Measurement, we need to start with a large ensemble of particles, all prepared in the same initial state $|\Psi_i\rangle$. Each particle will interact with a separate measuring device, which is in the state

$$|\Phi(q)\rangle = \int \frac{1}{\sqrt{\Delta}(2\pi)^{\frac{1}{4}}} e^{-\frac{q^2}{4\Delta^2}} |q\rangle \, dq, \tag{2.1}$$

where Δ is the standard deviation of the position of the device. In momentum space, this is a Gaussian with mean 0 with a standard deviation of $\frac{1}{2\Delta}$.

The interaction Hamiltonian is

$$H = \chi(t)\hat{p} \otimes \hat{A}, \tag{2.2}$$

where $\hat{p}$ is the momentum operator and $\hat{A}$ is the operator for quantity, we aim to measure.

There are two ways in which, a measurement can be made weak. The more obvious way is by having an impulsive measurement at time t_0

$$\chi(t) = \chi\delta(t - t_0), \tag{2.3}$$

in the limit of χ being small. However the more interesting way of making a weak measurement (which is the method used in the original paper [2]) is by requiring the initial momentum of the device to be ≈ 0. In order to ensure that the momentum is small, the uncertainty in the momentum has to be very small as well. Due to the Uncertainty Principle, this requires the uncertainty in position to be large. Hereafter, I shall use to the second condition as the limit of a Weak Measurement. In order to emphasise that the "weakness" of the measurement is not due to a weak coupling constant, let

$$\int_0^{t_0} \chi(t)\, dt = 1. \tag{2.4}$$

The particle and the device evolve under the interaction Hamiltonian into the following entangled state (Note that hereafter, we use units in which $\hbar = 1$)

$$\begin{aligned} e^{-i\hat{p}\otimes\hat{A}}|\Phi(q)\rangle \otimes |\Psi_i\rangle &= (\sum_j |a_j\rangle e^{-i\hat{p}a_j}\langle a_j|)|\Phi(q)\rangle|\Psi_i\rangle \\ &= \sum_j \alpha_j|\Phi(q - a_j)\rangle|a_j\rangle \end{aligned} \tag{2.5}$$

where $|a_j\rangle$ is the eigenvector of $\hat{A}$ with eigenvalue a_j and

$$|\Psi_i\rangle = \alpha_j|a_j\rangle. \tag{2.6}$$

(In the first step, the following result was used:

$$e^M = Ue^D U^\dagger, \tag{2.7}$$

where

$$M = UDU^\dagger \tag{2.8}$$

and U is a Unitary matrix.) Hence, we can determine $\hat{A}$ of the particle by measuring the shift in position of the device. However if

$$\Delta \gg a_j \forall j, \tag{2.9}$$

in other words, if the uncertainty in the device is much larger than the amount it is shifted by, then it is impossible to determine $\hat{A}$ of the particle from a single measurement. Instead, the average value $\langle \hat{A} \rangle$ of the ensemble of particles can be determined by repeating this measurement on each individual particle prepared in the same initial state. After interacting with the particle, the probability of a position measurement of the device being q is

$$P(q) = \sum_j |\alpha_j|^2 \frac{1}{\sqrt{\Delta}(2\pi)^{\frac{1}{2}}} e^{-\frac{(q-a_j)^2}{3\Delta^2}} \tag{2.10}$$

and since Δ is large,

$$\begin{aligned} P(q) &\approx \sum_j |\alpha_j|^2 \frac{1}{\sqrt{\Delta}(2\pi)^{\frac{1}{2}}} (1 - \frac{(q-a_j)^2}{3\Delta^2}) \\ &= \sum_j \frac{1}{\sqrt{\Delta}(2\pi)^{\frac{1}{2}}} (1 - \frac{(q-|\alpha_j|^2 a_j)^2}{3\Delta^2}) \\ &\approx \sum_j \frac{1}{\sqrt{\Delta}(2\pi)^{\frac{1}{2}}} e^{-\frac{(q-|\alpha_j|^2 a_j)^2}{3\Delta^2}} \\ &= \sum_j \frac{1}{\sqrt{\Delta}(2\pi)^{\frac{1}{2}}} e^{-\frac{(q-\langle \hat{A} \rangle)^2}{3\Delta^2}} \end{aligned} \tag{2.11}$$

Thus the probability distribution of the position of the device is a gaussian centred around $\langle \hat{A} \rangle$. Since for an ensemble of N such particles, the uncertainty in the average is reduced by $\frac{1}{\sqrt{N}}$, the average value $\langle \hat{A} \rangle$ of the system may be determined with arbitrary accuracy.

The reason this measurement procedure is termed "Weak" is because the quantum state of each particle does not change much as a result of the measurements. This is because after the measurement the state of the system is

$$|\Psi\rangle = \frac{1}{R} \sum_j \alpha_j \frac{1}{\sqrt{\Delta}(2\pi)^{\frac{1}{4}}} e^{-\frac{(q_0 - a_j)^2}{4\Delta^2}} |a_j\rangle, \tag{2.12}$$

where R is a normalisation constant and q_0 is the measurement outcome of the position of the device. And in the limit of Δ being very large, $|\Psi\rangle$ approaches $|\Psi_i\rangle$. Hence, the "weakness" of the measurement is due to the lack of information extracted from the particle.

2.2 The Weak Value

The Weak Measurement protocol has nothing strange or paradoxical about it. However when post selection is added to the weak measurement protocol, it yields some interesting results. These results will be discussed in *Chapter3*. In this section, I will describe how adding post selection to the protocol affects the measurement outcome, define the weak value and the limits under which the calculations are valid.

We now choose a state $|b\rangle$ such that

$$\hat{B}|b\rangle = b|b\rangle, \tag{2.13}$$

for some operator $\hat{B}$ that does not commute with $\hat{A}$, to be the post selected state. We perform a strong measurement on the particle(using a separate measuring device), after it has interacted with the weak measuring device , to collapse the particle into one of $\hat{B}'s$ eigenstates. Only the particles that collapse into the $|b\rangle$ state are then used to calculate the average value of $\hat{A}$. Hence only the post selected subset of the particles from the original ensemble are used to calculate the average value of $\hat{A}$.

After post selection the state of the particle and device is

$$\begin{aligned}
|\Omega\rangle &= \frac{1}{K}|b\rangle\langle b|e^{-i\hat{p}\hat{A}}|\Phi(q)\rangle\||\Psi_i\rangle \\
&\approx \frac{1}{K}|b\rangle\langle b|(1-i\hat{p}\hat{A})|\Phi(q)\rangle\||\Psi_i\rangle \\
&= \frac{1}{K}|b\rangle\langle b|\Psi_i\rangle(1-i\hat{p}\frac{\langle b|\hat{A}|\Psi_i\rangle}{\langle b|\Psi_i\rangle})|\Phi(q)\rangle \\
&\approx \frac{1}{K}\langle b|\Psi_i\rangle|b\rangle e^{-i\hat{p}\frac{\langle b|\hat{A}|\Psi_i\rangle}{\langle b|\Psi_i\rangle}}|\Phi(q)\rangle,
\end{aligned} \tag{2.14}$$

where K is the normalisation constant. Here, I would like to point out that $K \neq |\langle b|\Psi_i\rangle|$ if $\langle\hat{A}\rangle_W$ is complex. This is because, in that case $e^{-i\hat{p}\frac{\langle b|\hat{A}\Psi_i\rangle}{\langle b|\Psi_i\rangle}}$ is not unitary anymore and hence does not preserve the inner product. This fact is not immediately clear in most of the literature on this topic as in most cases, the normalisation constant is completely ignored and calculations are done on the unnormalised state of the device. However, I shall address this point in more depth in the next *Chapter*.

So, now the device measures a quantity called the Weak Value of $\hat{A}$

$$\boxed{\langle\hat{A}\rangle_W = \frac{\langle b|\hat{A}|\Psi_i\rangle}{\langle b|\Psi_i\rangle}} \tag{2.15}$$

Now we need to look carefully at the limits under which the approximations made in (2.14) are valid. The approximation made in going from step 1 to step 2 is only valid if

$$|p^n\langle b|\hat{A}^n|\Psi_i\rangle| \ll |\langle b|\Psi_i\rangle| \tag{2.16}$$

and

$$|p^n \langle b|\hat{A}^n|\Psi_i\rangle| \ll |\langle b|\hat{A}|\Psi_i\rangle| \tag{2.17}$$

for $n \geq 2$.

The approximation in going from step 4 to step 5 requires

$$|p\langle\hat{A}\rangle_W| \ll 1. \tag{2.18}$$

This makes condition (2.22) stronger than condition (2.16). Since p's spread around 0 is determined by Δ , equation (2.23) implies

$$\boxed{\Delta \gg |\frac{\langle\hat{A}\rangle_W}{2}|} \tag{2.19}$$

and condition (2.22) implies

$$\boxed{\Delta \gg \max_{n=2,3,4...} |\frac{1}{2}(\frac{\langle\hat{A}\rangle_W}{\langle b|\hat{A}^n|\Psi_i\rangle})^{\frac{1}{1-n}}|} \tag{2.20}$$

Note that the limits defined in the original weak measurement paper are incorrect [11].

We can now see that adding post selection to weak measurement, changes the value the probability distribution of the device shifts by from $\langle\hat{A}\rangle$ to $\langle\hat{A}\rangle_W$.

In the previous chapter, we had restricted the coupling constant to obey (2.4) in order to show an alternative method of making a measurement weak. However having a small coupling constant and solving to only include $O(\chi)$ terms will yield exactly the same results as having $p \sim 0$. One can see this from (2.14) :

$$
\begin{aligned}
|\Omega\rangle &= \frac{1}{K}|b\rangle\langle b|e^{-i\chi\hat{p}\hat{A}}|\Phi(q)\rangle\||\Psi_i\rangle \\
&\approx \frac{1}{K}|b\rangle\langle b|(1-i\chi\hat{p}\hat{A})|\Phi(q)\rangle\||\Psi_i\rangle \\
&= \frac{1}{K}|b\rangle\langle b|\Psi_i\rangle(1-i\chi\hat{p}\frac{\langle b|\hat{A}|\Psi_i\rangle}{\langle b|\Psi_i\rangle})|\Phi(q)\rangle \\
&\approx \frac{1}{K}\langle b|\Psi_i\rangle|b\rangle e^{-i\chi\hat{p}\frac{\langle b|\hat{A}|\Psi_i\rangle}{\langle b|\Psi_i\rangle}}|\Phi(q)\rangle.
\end{aligned}
\tag{2.21}
$$

The approximation made in going from step 1 to step 2 is only valid if

$$
|\chi p^n\langle b|\hat{A}^n|\Psi_i\rangle| \ll |\langle b|\hat{A}|\Psi_i\rangle| \tag{2.22}
$$

for $n \geq 2$.

The approximation in going from step 4 to step 5 requires

$$
|\chi p\langle\hat{A}\rangle_W| \ll 1. \tag{2.23}
$$

So after a weak measurement the state of the measuring device is $|\Phi(q - \chi\langle\hat{A}\rangle_W)\rangle$. When doing calculations, it is a matter of choice/ convenience as to which approximation is chosen to ensure weakness.

3 The Strange Weak Value

3.1 Complex values

In the previous chapter, I claimed that when the post selection process is incorporated into the weak measurement procedure, the measurement outcome of the device is the weak value. However, from (2.14), it is not immediately apparent, the effect of the weak value on the measuring device. Naively one might think that the Gaussian of the position of the measuring device is shifted by the weak value. This is, however not always the case. From the definition of the weak value

$$\langle \hat{A} \rangle_W = \frac{\langle \Psi_{PS} | \hat{A} | \Psi_i \rangle}{\langle \Psi_{PS} | \Psi_i \rangle} \tag{3.1}$$

one can see that despite $\hat{A}$ being hermitian, the weak value may be complex depending on the choice of the post selected state $|\Psi_{PS}\rangle$. In cases where the weak value has an imaginary part, $e^{-i\hat{p}\frac{\langle b|\hat{A}\Psi_i\rangle}{\langle b|\Psi_i\rangle}}$ is no longer unitary. So how does a complex weak value affect the state of the measuring device?

Theorem:[13] Let $\langle \hat{A} \rangle = a + ic$. After a weak measurement, the mean position of the measuring device of mass m is

$$\langle \hat{q} \rangle_f = \langle \hat{q} \rangle_i + \chi a + \chi c m \frac{dVar_q}{dt}\Big|_{t=t_0} \tag{3.2}$$

and the mean momentum of the device is

$$\langle \hat{p} \rangle_f = \langle \hat{p} \rangle_i + 2\chi c Var_p, \tag{3.3}$$

where

$$Var_q = \langle \Phi | q^2 | \Phi \rangle - \langle \Phi | q | \Phi \rangle^2, \tag{3.4}$$

$$Var_p = \langle \Phi | p^2 | \Phi \rangle - \langle \Phi | p | \Phi \rangle^2 \tag{3.5}$$

and t_0 is the time of the impulsive measurement interaction.

To work through the proof, it is convenient to assume the first limit of the weak measurement, where χ is small. Having said that the results are equivalent to if the weakness was achieved by restricting the momentum of the measuring device by condition (2.19),(2.20).

In order to derive (3.2), start by substituting

$$\hat{p} = -i\frac{\partial}{\partial q} \tag{3.6}$$

into the equation for the unnormalised final state of the device

$$|\alpha\rangle = \langle \Psi_{PS} | \Psi_i \rangle e^{-i\chi \langle \hat{A} \rangle_W \hat{p}} |\Phi\rangle \tag{3.7}$$

and multiplying it by its complex conjugate

$$\begin{aligned}\alpha(q)^*\alpha(q) &= [(1-\chi(a-ic)\frac{\partial}{\partial q})\Phi^*][(1-\chi(a+ic)\frac{\partial}{\partial q})\Phi] \\ &= \Phi^*\Phi - \chi a(\Phi^{*\prime}\Phi + \Phi^*\Phi^\prime) - i\chi c(\Phi^{*\prime}\Phi \ - \Phi^*\Phi^\prime)\end{aligned} \tag{3.8}$$

where Φ' is the derivative with respect to q. Note that only $O(\chi)$ terms have been retained.

Since $\Phi(q)$ might be complex we rewrite it as

$$\Phi(q) = Re^{iS} \tag{3.9}$$

.

Using

$$\Phi^{*\prime}\Phi + \Phi^*\Phi' = \frac{\partial\Phi^*\Phi)}{\partial q} \tag{3.10}$$

and

$$\begin{aligned} \Phi^{*\prime}\Phi - \Phi^*\Phi' &= Re^{-iS}\frac{\partial(Re^{iS})}{\partial q} - Re^{iS}\frac{\partial(Re^{-iS})}{\partial q} \\ &= 2iR^2\frac{\partial S}{\partial q}, \end{aligned} \tag{3.11}$$

we have

$$\alpha^*\alpha = \rho - \chi a\rho' + 2\chi c\rho S', \tag{3.12}$$

where $\rho = R^2$.

Since, $|\alpha\rangle$ is unnormalised, the expression for $\langle\hat{q}_f\rangle$ is

$$\begin{aligned} \langle\hat{q}\rangle_f &= \frac{\int \alpha^* q\alpha dq}{\int \alpha^*\alpha dq} \\ &= \langle\hat{q}\rangle_f + \chi a - \chi c\int (q - \langle\hat{q}\rangle_f)^2(\rho S')'. \end{aligned} \tag{3.13}$$

Substituting (3.9) into the time dependent Shrödinger equation at the time of the measurement impulse t_0

$$i\frac{\partial\Phi}{\partial t} = -\frac{1}{2m}\Phi\text{''} + V(q)\Phi \tag{3.14}$$

and solving the imaginary part yields

$$(\rho S')' = -m\frac{\partial \rho}{\partial t}|_{t=t_0}. \tag{3.15}$$

Putting this pack into (3.13) and solving the integral gives

$$\langle \hat{q} \rangle_f = \langle \hat{q} \rangle_i + \chi a + \chi c m \frac{dVar_q}{dt}|_{t=t_0}. \tag{3.16}$$

We can now see how the weak value affects the position gaussian of the device. It depends not only on whether the weak value has an imaginary part or not but also on the initial position gaussian of the device. If the device's initial state is real valued ($S = 0$), then the integral in (3.13) with coefficient c, i.e the imaginary part of the weak value, disappears. Hence the expectation value of the position is only shifted by the real part of the weal value.

In order to prove the change in the device's momentum, we shall work in the Heisenberg picture. Again, because $|\alpha\rangle$ is unnormalised, the expectation value for the momentum of the device after post selection is

$$\begin{aligned} \langle \hat{p} \rangle &= \frac{\langle \alpha|\hat{p}|\alpha\rangle}{\langle \alpha|\alpha\rangle} \\ &= \frac{\langle \alpha|\hat{p}|\alpha\rangle - i\chi\langle \hat{A}\rangle_W \langle \Phi|\hat{p}^2|\Phi\rangle + i\chi\langle \hat{A}\rangle_W^* \langle \Phi|\hat{p}^2|\Phi\rangle}{\langle \Phi|\Phi\rangle - i\chi\langle \hat{A}\rangle_W \langle \Phi|\hat{p}|\Phi\rangle + i\chi\langle \hat{A}\rangle_W^* \langle \Phi|\hat{p}|\Phi\rangle} \\ &= \langle \hat{p} \rangle_i + 2\chi c Var_p \end{aligned} \tag{3.17}$$

3.2 Great Expectations

At first glance, the weak value looks like a conditional expectation value as it is essentially the expectation value of a subset of the ensemble. However, this interpretation of the weak value runs into a big problem. In some cases, the weak value of an operator can be far greater than any of its eigenvalues. As the first paper of weak measurement highlighted in its title, the weak value of the spin of a spin $\frac{1}{2}$ can even be 100. In this section, I will discuss this particular claim, known as the AAV effect.

From the definition of $\langle\hat{A}\rangle_W$, one can see that it is not bounded by the $\hat{A}$'s eigenvalues. In fact the weak value can be made arbitrarily large as $\langle\Phi_{PS}|\Phi_i\rangle$ approaches 0. Now consider what happens to the device after it weakly interacts with a particle under the interaction hamiltonian. The wavegunction of each device is shifted by an eigenvalue of $\hat{A}$. The distribution of all the devices combined is a superposition of wavefunctions of each device. It therefore seems counterintuitive that for a post selected ensemble such a superposition of shifts can result in an overall shift that is much larger than the individual shifts. To understand this phenomenon better, let us work through the example used in the original paper [2],[11].

A beam of spin $\frac{1}{2}$ particles move in the $\hat{y}$ direction with their spins lying in the $\hat{x}-\hat{z}$ plane at an angle θ to the $\hat{y}$ axis. The spin state of the particles is

$$|\Psi_i\rangle = \frac{1}{\sqrt{2}}[(\cos\frac{\theta}{2} + \sin\frac{\theta}{2})|0\rangle + (\cos\frac{\theta}{2} - \sin\frac{\theta}{2})|1\rangle, \tag{3.18}$$

where $|0\rangle$ and $|1\rangle$ are the eigenstates of $\hat{z}$-component spin operator σ_z with eigenvalues +1 and -1 respectively.

The particles' velocities are well defined. The spatial wavefunction of the particles in the $\hat{z}$ direction is a Gaussian with standard deviation $\frac{1}{2\Delta}$. Therefore the momentum with which the beam is spreading in the $\hat{z}$ direction is Δ. The wavefunction of the beam in momentum space is

$$|\Phi\rangle = \int e^{-\frac{p_z^2}{4\Delta^2}} f(p_x, p_y)|p\rangle dp, \tag{3.19}$$

where the definition of the function $f(p_x, p_y)$ is unimportant. The $\hat{z}$-component of the spins of the particles are measured by passing the beam through a Stern-Gerlach device. The interaction Hamiltonian for this is

$$H = -\lambda\chi\hat{z}\hat{\sigma_z}, \tag{3.20}$$

where $\hat{z}$ measures the position of the particle along the $\hat{z}$ axis and $\lambda = \mu\frac{\partial B}{\partial z}$, in other words the particle's magnetic moment multiplied by the magnetic field gradient of the Stern-Gerlach device in the $\hat{z}$ direction.

To relate this scenario to our general discussion of weak measurements earlier; here, the the measuring "device" is the momentum wavefunction of the particle itself and the quantity to be weakly measured is $\lambda s i g \hat{m} a_z$ of the particle. We enforce conditions (2.19) and (2.20) on Δ to ensure weakness and χ obeys (2.4).

The beam is then passed through a second Stern Gerlach device, which strongly measures the $\hat{x}$ component spin of the particle. This splits the beam into two , one corresponding to +1 eigenstate of the $\hat{\sigma_x}$ operator and one to the -1 eigenstate. A screen is placed only in front of the beam corresponding to the +1 eigenstate

$$|\Psi_{PS}\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \tag{3.21}$$

The distribution of particle along the $\hat{z}$ axis is used to determine the momentum distribution along the $\hat{z}$ axis which would have shifted by the weak value of $\lambda\hat{\sigma_z}$.

$$|\Phi\rangle_f = \int e^{-\frac{(p_z - \langle\lambda\hat{\sigma_z}\rangle_W)^2}{4\Delta^2}} h(p_x, p_y)|p\rangle dp, \tag{3.22}$$

where the definition of the function $h(p_x, p_y)$ is unimportant.

The weak value of $\lambda\hat{\sigma_z}$ is

$$\begin{aligned}\langle\lambda\hat{\sigma_z}\rangle_W &= \frac{\langle\Psi_{PS}|\lambda\hat{\sigma_z}|\Psi_i\rangle}{\langle\Psi_{PS}|\Psi_i\rangle} \\ &= \frac{\lambda\sin\frac{\theta}{2}}{\cos\frac{\theta}{2}} \\ &= \lambda\tan\frac{\theta}{2}.\end{aligned} \tag{3.23}$$

Therefore,

$$|\Phi\rangle_f = \int e^{-\frac{(p_z-\tan\frac{\theta}{2})^2}{4\Delta^2}} h(p_x,p_y)|p\rangle dp. \tag{3.24}$$

And since, $tan\frac{\theta}{2}$ is always real, this leads to a shift of $\tan\frac{\theta}{2}$ in the mean of the $\hat{z}$ component momentum of the particle. One can see that as θ approaches π, this shift can become arbitrarily large. In this way, a measurement of a spin $\frac{1}{2}$ particle can turn out to be a 100 as claimed by [2]. This statement, however does not make any sense. Quantum mechanics tells us that this cannot be true. The mean spin of a set spin$\frac{1}{2}$ particles can never be a 100, regardless of how specifically the group of particles is chosen. Hence either, we are seeing some new physics here or the above claim is incorrect. We shall see in the next *Chapter* that this claim must be taken with caution as it might be due to the misinterpretation of the weak value. For now let us address a more immediate question of how a superposition of small shifts can lead to a very large shift.

It is true that, a weighted superposition of gaussians, each shifted by some small amount (eg: the eigenvalues of an operator) can never yield a shift which is larger than the biggest shift in the superposition. This seems obvious. However if the weights are not positive definite then this is not always true [11]. When the weights are not all positive definite, the individual gaussians cancel in a complicated way sometimes causing the resultant gaussian to be shifted by a very large amount. This is exactly what happens in the case of the AAV effect.

In general, after a weak measurement of an operator $\hat{A}$, the state of the device is (using initial state of device and Hamiltonian as defined in (3.20) and (6.41) respectively)

$$|\phi\rangle_f = \frac{1}{K}\sum_j \alpha_j \beta_j^* \int e^{-\frac{(p-a_j)^2}{4\Delta^2}} |p\rangle dp, \tag{3.25}$$

where $|a_j\rangle$ are the eigenstates of $\hat{A}$ and

$$|\Psi_i\rangle = \sum_j \alpha_j |a_j\rangle \tag{3.26}$$

and

$$|\Psi_{PS}\rangle = \sum_j \beta_j |a_j\rangle. \tag{3.27}$$

From (3.25), one can see that the weights of the individual Gaussians are not necessarily positive definite. To see what happens specifically in the case of the AAV effect as θ approaches π, let $\theta = \pi - 2\epsilon$ where $\epsilon \ll 1$. In this limit equation (3.25) yields

$$\phi_f(p) = \frac{1}{2K}[(1+\epsilon)e^{-\frac{(p-\lambda)^2}{4\Delta^2}} - (1-\epsilon)e^{-\frac{(p+\lambda)^2}{4\Delta^2}}] \tag{3.28}$$

The combination of the two terms each peaked at $p = \pm\lambda$ leads to an approximate gaussian peaked at $p = \frac{\lambda}{\epsilon}$. So we can see that as ϵ approaches 0, the Gaussian peaks at an arbitrarily large value.This effect is normally not encountered when calculating expectation values as the weights are always positive definite.

There is however one constraint on how large the weak value can get. As shown in the condition in(2.19), the weak value can never be greater than the uncertainty Δ, as in this case, the approximations made in calculating the weak value break down. Despite this fact, in the literature, it is sometimes claimed that the weak value is greater than the uncertainty Δ [1].

4 What is the Weak Value?

In this *Chapter*, I shall offer a very simple explanation to what the weak value is and how one should interpret its surprising features. But first I shall start with how it is interpreted by the authors of the original paper.

4.1 Interpreting the Weak Value using the Two-State Vector Formalism (TVSF)

The weak value is often given physical meaning using the Two-State Vector Formalism (TVSF). In the TVSF, two quantum states are used to describe a quantum system at any particular time. The advocates of TVSF argue that an artificial arrow of time is inserted into Quantum theory, that should not be there [7]. They argue that Quantum theory must be a time symmetric theory as reflected by the time symmetry in Shrödinger's and Heisenberg's equations. They believe that the time - asymmetry is introduced into Quantum mechanics by the theory of measurement[6]. To illustrate their point, I cite an example from [6]: Consider an ensemble of spin $\frac{1}{2}$ particles which are in the +1 eigenstate of $\hat{\sigma_x}$ at time t. We are able to predict that the probability of finding $\hat{\sigma_y} = 1$ immediately after time t is a $\frac{1}{2}$. Hence we can say with certainty, that given a large enough ensemble, exactly half the particles would be have $\hat{\sigma_y} = 1$. However we cannot predict the probability of $\hat{\sigma_y} = 1$ before time t to also be $\frac{1}{2}$. The ensemble could have all been prepared in the state $\hat{\sigma_y} = 1$, or in the state $\hat{\sigma_y} = -1$. Hence, if we do not know what the state prior to time t was, then we cannot determine with certainty what proportion of the ensemble would have been in $\hat{\sigma_y} = 1$ regardless of the size of the ensemble. The difference here is that we do not assume the existence of a future state but we do assume the existence of a past state. It is argued that our assumption that the past exists and the future does not at a particular time t introduces an artificial arrow of time into Quantum theory that is not intrinsic to the theory. For a classical system, having an initial boundary condition is equivalent to a final boundary condition as the same outcomes are produced by evolving the system forward in time from the initial boundary condition or backwards in time from the final boundary conditions. However this is not true in Quantum mechanics and hence, it is argued that in order to achieve a time symmetry one must impose both initial and final conditions on a quantum system. The way to do this is by pre selecting and post selecting the ensemble so that in the time interval between the pre and post selection, the situation is symmetric with respect to time. TVSF describes the time interval between two measurements by the pre selected state evolving forward in time and the post selected state evolving backwards in time. TVSF says that at a particular time t between the time of pre selection t_i and post selection t_f, a system ought to be described by two wavefunctions $|\Psi_1\rangle$ evolving forward in time and $\langle\Psi_2|$ evolving backward in time.

$$|\Psi_1\rangle = e^{-i\int_{t_i}^{t} Hd\tau}|a\rangle \tag{4.1}$$

$$\langle\Psi_2| = \langle b|e^{-i\int_{t}^{t_f} Hd\tau} \tag{4.2}$$

where $|a\rangle$ and $|b\rangle$ are the pre and post selected states respectively. In this framework, the weak value of an operator $\hat{O}$ is considered the expectation value of $\hat{O}$ for the given initial and final conditions i.e for the given pre and post selected ensemble in the weak measurement limit.

For a device in the state

$$|\Phi(q)\rangle = \int \frac{1}{\sqrt{\Delta}(2\pi)^{\frac{1}{4}}} e^{-\frac{q^2}{4\Delta^2}} |q\rangle \, dq, \tag{4.3}$$

assuming the free Hamiltonian is 0 and the interaction Hamiltonian is

$$H = \chi(t)\hat{p}\hat{O}, \tag{4.4}$$

where $\chi(t)$ obeys (2.4); it's interaction with a particle in the state $\langle\Psi_{PS}||\Psi_i\rangle$ shifts its position distribution by the weak value.

$$\begin{aligned} |\Phi(q)'\rangle &\approx \frac{1}{K}\langle\Psi_{PS}|\Psi_i\rangle e^{-i\hat{p}\frac{\langle\Psi_{PS}|\hat{O}|\Psi_i\rangle}{\langle\Psi_{PS}|\Psi_i\rangle}} |\Phi(q)\rangle \\ &\approx \int \frac{1}{\sqrt{\Delta}(2\pi)^{\frac{1}{4}}} e^{-\frac{(q-Re(\langle\hat{A}\rangle)^2}{4\Delta^2}} |q\rangle \, dq. \end{aligned} \tag{4.5}$$

(Note here for simplicity, it is assumed the change due to the imaginary part of the weak value to be small)

I propose a simple way of checking this interpretation of the weak value. If one does not assume that the free Hamiltonian acting on the particle is 0 but rather small compared to the interaction Hamiltonian, then if post selection is carried out after a sufficiently long time such that $e^{-i\int_t^{t_f} H_{free} d\tau}$ is not trivial; then the evolution of $\langle\Psi_{PS}|$ should also be non-trivial. In this case, the shift in the position distribution of the measuring device would no longer be $\langle\Psi_{PS}|\hat{A}|\Psi_i\rangle$.

If

$$\langle \Psi_{PS}| e^{-i\int_t^{t_f} H_{free} d\tau} = \langle \Psi'| \tag{4.6}$$

then the device is shifted by the weak value $\langle \Psi'|\hat{A}|\Psi_i\rangle$. However, if we assume the the free Hamiltonian acting on the device is approximately 0, then, the state of the device does not evolve significantly in time. Hence, its probability distribution should remain unchanged regardless of the post selection time. This implies that the shift in its probability distribution does not depend on the time of post selection. This is a contradiction. It would therefore be easy to check the validity of this interpretation by seeing if the altering the length of time between the weak measurement and post selection would change the weak value.

It is also suggested that instead of thinking of the expectation value of variable as a statistical average of its eigenvalues, one can think of the expectation value as a quantum average of weak values [4].

$$\begin{aligned} \langle \Psi|\hat{O}|\Psi\rangle &= \sum_j |\langle b_j|\Psi\rangle|^2 \frac{\langle b_j \hat{O}|\Psi\rangle}{\langle b_j|\Psi\rangle} \\ &= \sum_j P(b_j|\Psi_i)\langle \hat{O}\rangle_{W,b_j} \end{aligned} \tag{4.7}$$

where $|b_j\rangle$ is the eigenvector of the post selection operator B. The argument is that since traditionally the expectation value is only calculated using an initial condition (pre selected state), it is the weighted sum of the expectation of a pre and post selected system for all possible post selected states.

Even though the TVSF attempts to offer an interpretation of the weak value, their success in explaining why they can in some cases lie outside the range of eigenvalues is debatable if not unsatisfying. In the literature, there is very little said for what the “eccentric” weak values mean, that is the weak values that lie outside the range on eigenvalues of the operator like for the AAV effect. It is simply justified using the mathematical fact that complex amplitudes of superposition are responsible for this effect, as discussed in section 3.2 and that the stranger the value, the rarer it is. In other words, there is a very low probability of the state collapsing to the post selected states that are correspond to the eccentric weak values. The TVSF interpretation of the weak value is also quite radical in a sense as it considers it a physical property.

“*...the weak value of A, (which) is to be regarded as a definite physical property of an unperturbed quantum system in the time interval between two complete measurements.*”-Y.Aharonov [4].

This implies that a physical property of a system is allowed to lie outside the spectrum of eigenvalues of its operator. This requires a complete re-evaluation of the Quantum theory. In the next section, I will offer a simple explanation as to why I think the weak value cannot be given the status it has been given in this formalism and more importantly show that the spin of a spin $\frac{1}{2}$ particle can not be 100.

4.2 Entangled

For a moment, forget about weak measurements or any measurement at all and just consider entanglement. Assume, the states of two objects, object G and object H are entangled. They become entangled by evolving under a unitary operator U. Now assume there are N copies of these entangled pair of objects. G and H are entangled in such a way that the probability distribution of the position of G being q ,given that the measurement of variable V_H of H has outcome v is a gaussian centred around a particular value $zeta$.

To reiterate,

$$Pr(Q_G = q|V_H = v) = ce^{-\frac{(q-\zeta)^2}{2b^2}} \tag{4.8}$$

Note that, G is not "feeling" anything from H, it is typical entanglement, where the probabilities of certain measurement outcomes of G and H are correlated. One cannot immediately claim that G is in any way measuring any variable of H.

Now, only if G and H were entangled in a such a way that the probability of the position of G being q is centred around a value, which happens to be the expectation value of a variable X of H, then the probability distribution of the position of G can be used to find this expectation value. Only now can we say that G is used to measure $\langle X \rangle$ of H and can be considered a measuring device in any sense. A measuring device, is a device that gives you information on the quantity it is measuring, which in this case is $\langle X \rangle$ of H. Now if one only considers the pairs in which the measurement of variable V_H of H has outcome v, then the probability distribution of the position of G is given by (4.8), where the gaussian is centred around value ζ. For this specific subset of the entangled pairs, G can no longer be considered a measuring device because it is not measuring anything! We cannot deduce from it the expectation of X, hence it is no longer a measuring device used to measure $\langle X \rangle$ of H. Only if ζ happened to be a specific quantity of H, would we consider G a measuring device specifically for measuring that particular quantity of H. Otherwise G and H are merely two entangled particles and nothing more.

It is now easy to see how this argument lends itself to the situation of weak measurements. The object referred to as the "measuring device" is a measuring device that can only be used for measure the expectation value of variable $\hat{A}$ of the particles. This is because, its probability distribution happens to be peaked at the value $\langle \hat{A} \rangle$. However by only using a post selected subset of devices, we alter its probability distribution so it is no longer peaked around value $\langle \hat{A} \rangle$. Hence it is no longer a measuring device for measuring $\langle \hat{A} \rangle$ of the particle. It is now nothing more than an entangled particle.

One might now ask, then what is the expectation of $\hat{A}$ of $|\Psi_i\rangle$ given that a measurement of the operator $\hat{B}$ on it has yielded the value b. However there is a fundamental error in asking this question. In asking this question, one is confusing the state of the particle with the individual particle itself. The state of the individual particle changes every time, we extract any information about it. Even in the limit of weak measurements, the individual measurements are only weak because, we are extracting almost no information about that individual particle by measuring it. The reason this is perhaps confusing is because weak measurements creates the illusion that $\hat{A}$ of the state $|\Psi_i\rangle$ is first measured, after which it evolves almost in the same state to further have $\hat{B}$ measured. However, this is clearly not true. As emphasised earlier the only quantity being measured is the $\langle \hat{A} \rangle$ of the ensemble by exploiting its entanglement with the "device". We are not measuring $\hat{A}$ of the individual particles in state $|\Psi_i\rangle$. If we did it would most definitely collapse into an eigenstate of $\hat{A}$ and would no longer have anything to do with the state $|\Psi_i\rangle$.

So a more sensible question to ask would be, what is the expectation value of $\hat{A}$ for the specific pre and post selected subset of particles. We can calculate this easily. Let the size of the pre selected ensemble be N_p.

Then the number of states that collapse to $|a_j\rangle$ is

$$|\langle\Psi_i|a_j\rangle|^2 N_p. \tag{4.9}$$

The number of $|a_j\rangle$ states that then collapse to the post selected state $|b\rangle$ is

$$|\langle b|a_j\rangle|^2|\langle\Psi_i|a_j\rangle|^2 N_p. \tag{4.10}$$

The size of the pre and post selected ensemble is therefore

$$\sum_j |\langle b|a_j\rangle|^2|\langle\Psi_i|a_j\rangle|^2 N_p. \tag{4.11}$$

Hence the probability of $A = a_j$ for particles in the post and pre selected ensemble are

$$P(A = a_j|pre, postselection) = \frac{|\langle b|a_j\rangle|^2|\langle\Psi_i|a_j\rangle|^2}{\sum_k |\langle b|a_k\rangle|^2|\langle\Psi_i|a_k\rangle|^2}. \tag{4.12}$$

Finally the estimate of the pre and post selected ensemble of particles is

$$\langle\hat{A}\rangle_{pre,post-selected} = \sum_j \frac{|\langle b|a_j\rangle|^2|\langle\Psi_i|a_j\rangle|^2}{\sum_k |\langle b|a_k\rangle|^2|\langle\Psi_i|a_k\rangle|^2} a_j \tag{4.13}$$

where a_j is the j^{th} eigenvalue of $\hat{A}$. Equation (4.12) is known as the Aharonov-Bergmann-Lebowitz (ABL) formula, which was first introduced in [3] in order to calculate probabilities for pre and post selected ensembles. This expression is the same irrespective of whether one measured $\hat{A}$ first or $\hat{B}$ showing that their is no time asymmetry in calculating this conditional expectation value. Note that $\langle\hat{A}\rangle_{pre,post-selected} \neq \langle\hat{A}\rangle_W$ and reassuringly, can never be greater than eigenvalues of $\hat{A}$ as the weights are positive definite and hence the average spin of spin $\frac{1}{2}$ particles can never be a 100.

Here, I should perhaps clarify that, I am not saying that the mathematics in calculating the weak value is incorrect or that it is not experimentally realisable. Of course if one carried out, the weak measurement + post selection protocol experimentally, they would indeed have the distribution of their "measuring" variable peak at the weak value. However, this does not make the weak value a meaningful physical property of the particle. Moreover, interpreting it as an expectation value of the variable for the pre and post selected ensemble is not justified. The debate of what is physical or unphysical is a somewhat slippery slope and unrelated to the point I am making here. The point I am making here is that the weak value of $\hat{A}$ has as much physical meaning as an arbitrary function $\hat{A}$.

In conclusion, by only using the post selected ensemble to calculate the expectation of $\hat{A}$, from the probability distribution of the "devices", one is not measuring any characteristic of the particles or ensemble of any physical meaning. $\langle \hat{A} \rangle_W$ is merely an artefact of the correlated probability distributions of the particle and the "device" due to their entanglement.

5 The Weak Value solves Quantum Paradox?

In this chapter, I will illustrate, furthur the general interpretation of the the weak value by showing how its use claims to solve a counter- factual paradox.

5.1 The Cheshire Cat

In [5], an experiment is suggested, to show that physical properties of an object can be separated from the object itself. The experiment claims to separate the polarisation of a photon from the photon so that at a particular time, the photon could be in one place with its polarisation in another.

The pre selected state is

$$|\Psi_i\rangle = \frac{1}{\sqrt{2}}(|L\rangle + |R\rangle)|H\rangle, \tag{5.1}$$

where $|L\rangle$ and $|R\rangle$ are states corresponding to the photon being in the left and right arms of the interferometer respectively. $|H\rangle$ refers to the horizontal polarisation of the photon. This is achieved by passing a horizontally polarized photon through a 50:50 beamsplitter (BS1) as shown in figure 5.1. The detectors in the left and right arms perform measurements

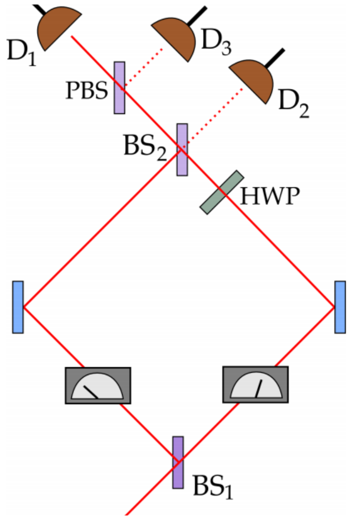

Figure 5.1: Measurement setup (copied from [5]).

on the photon. We shall come back to the type of measurement performed shortly. The photon then travels towards the second beamsplitter (BS2). Just before the BS2, there is a half wave plate (HWP) on the right arm, which changes the polarisation of a vertically polarised photon to a horizontally polarised photon.

$$HWP : |V\rangle \longleftrightarrow |H\rangle. \tag{5.2}$$

The second beam splitter performs the following function:

$$BS2 : \begin{cases} \frac{1}{\sqrt{2}}(|L\rangle + |R\rangle) \longrightarrow |L'\rangle \\ \frac{1}{\sqrt{2}}(|L\rangle - |R\rangle) \longrightarrow |R'\rangle \end{cases}$$

where $|L'\rangle$ carries on towards to polarising beam splitter (PBS) and $|R'\rangle$

travels towards detector D_2. The PBS performs the following function:

$$PBS : \begin{cases} |L'\rangle|H\rangle \longrightarrow |L''\rangle \\ |L'\rangle|V\rangle \longrightarrow |R''\rangle \end{cases}$$

where $|L''\rangle$ carries on towards D_1 and $|R''\rangle$ travels towards detector D_3.

Hence if D_1 clicks then we are certain that the state just before the HWP was

$$|\Psi_{PS}\rangle = \frac{1}{\sqrt{2}}(|L\rangle|H\rangle + |R\rangle|V\rangle). \tag{5.5}$$

This is the post selected state. So we only take into account measurements when D_1 clicks.

Firstly let us set the detectors in each arm to check which arm the photon passes though. The corresponding projection operators for this are

$$\Pi_L = |L+\rangle\langle L+| + |L-\rangle\langle L-| \tag{5.6}$$

$$\Pi_R = |R+\rangle\langle R+| + |R-\rangle\langle R-| \tag{5.7}$$

where $|+\rangle$ and $|-\rangle$ are the eigenstates of the Pauli X operator σ_x and correspond to the diagonal polarisation of the photon.($|H\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle and |V\rangle = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle)$.

Now if we look at the post selected state, one can see that it is orthogonal to

$$|\Psi_R\rangle = \frac{1}{\sqrt{2}}(|R+\rangle + |R-\rangle) \tag{5.8}$$

which is the state of collapse if the photon is detected in the right arm. Hence for the post selected ensemble, the probability of the measuring device finding the photon in the right arm is 0.

However if the measuring device checks for the polarisation of the photon corresponding to the $\sigma_x$in the right arm, then projectors for the measurements are

$$\Pi_{R+} = |R+\rangle\langle R+| \tag{5.9}$$

$$\Pi_{R-} = |R-\rangle\langle R-| \tag{5.10}$$

Hence after the measurement the photon will either be in state $|R+\rangle$ or $|R-\rangle$ or $\frac{1}{\sqrt{2}}(|L+\rangle + |L-\rangle$. And neither $|R+\rangle$ or $|R-\rangle$ are orthogonal to the post selected state, hence the probability of finding these outcomes is non zero for the pre and post selected ensemble. Therefore the paradox lies in the fact that the probability of detecting the photon in the right arm is zero for the post selected ensemble but the probability of detecting its polarisation is the right arm is not zero.

This is a counter-factual type of paradox as the contradiction disappears when one makes the measurements. For example when a measurement of polarisation is the right arm detects a non zero polarisation, then the photon too is found in the right arm. Hence one is not able to isolate the polarisation of the photon from the photon by doing strong measurements. It is suggested that performing weak measurements could however achieve this Cheshire cat like result [5]. Here we will focus on the theoretical results. For a detailed description of the experimental set of such weak measurements, I refer the reader to [10].

Since the expectation value of a projector is the probability of the outcome corresponding to the projector, the authors suggest that one could measure the weak value of the projectors to ascertain the probabilities of the corresponding outcomes. The weak value of finding the photon in the left arm is

$$\langle \Pi_L \rangle_W = \frac{\langle \Psi_{PS} | \Pi_L | \Psi_i \rangle}{\langle \Psi_{PS} | \Psi_i \rangle} = 1 \tag{5.11}$$

and the weak value of finding the photon in the right arm is

$$\langle \Pi_R \rangle_W = \frac{\langle \Psi_{PS} | \Pi_R | \Psi_i \rangle}{\langle \Psi_{PS} | \Psi_i \rangle} = 0. \tag{5.12}$$

The weak value of $\hat{\sigma_x}$ in the right arm corresponds to the weak value of

$$\hat{\sigma_x^R} = (|R\rangle\langle R|) \otimes (|+\rangle\langle +| - |-\rangle\langle -|) \tag{5.13}$$

and the weak value of $\hat{\sigma_x^R}$ is

$$\langle \hat{\sigma_x^R} \rangle_W = \frac{\langle \Psi_{PS} | \hat{\sigma_x^R} | \Psi_i \rangle}{\langle \Psi_{PS} | \Psi_i \rangle} = 1 \tag{5.14}$$

Since these measurements are weak, the measurement leaves the state of the photon unaffected and therefore, the weak measurement of the polarisation of the photon in the right arm can give a non trivial result whilst the weak value of the projection operator corresponding to the photon being in the right arm is 0. According to [5],[12], these results show that the polarisation of the photon can be disembodied from the the photon itself. This conclusion can only be drawn if the weak value of an operator is interpreted exactly like its expectation value under conditions of weakness and pre and post selection. However this kind of a straightforward interpretation is not justified. Especially equating the weak value of a projection operator to a probability can be very problematic. One can see, by looking at the expression for the weak value of a projector, that when

$$\langle \Psi_{PS} | \Pi | \Psi_i \rangle > \langle \Psi_{PS} | \Psi_i \rangle, \tag{5.15}$$

the weak value can be greater than 1. Also if

$$\langle \Psi_{PS}|\Psi_i\rangle < 0, \quad \langle \Psi_{PS}|\Pi|\Psi_i\rangle > 0 \tag{5.16}$$

then the weak value is negative. Thus interpreting these weak values as probabilities, as the authors of [5] have done, would mean getting probabilities greater than 1 and even negative probabilities. However this does not make sense as it goes against the very definition of probability. In fact, I see this as proof that this kind of interpretation of the weak value is wrong. In the paper [5], it was a convenient choice of post selected state that lead to weak values that could be sensibly interpreted as probabilities. I will now illustrate by choosing a different post selected state, the strange "probabilities". Let the post selected state of the photon be

$$|\Psi'_{PS}\langle = \frac{1}{\sqrt{3}}|LH\rangle + \frac{2}{\sqrt{3}}|RH\rangle \tag{5.17}$$

For the same pre selected state the weak values for the left and right projectors are

$$\begin{aligned} \langle \Pi_L\rangle_W &= \frac{\langle \Psi'_{PS}|\Pi_L|\Psi_i\rangle}{\langle \Psi_{PS'}|\Psi_i\rangle} \\ &= \frac{1}{\sqrt{6}}(\frac{1}{\sqrt{6}} - \frac{1}{\sqrt{3}})^{-1} \\ &\approx -2.4 \end{aligned} \tag{5.18}$$

$$
\begin{aligned}
\langle \Pi_R \rangle_W &= \frac{\langle \Psi'_{PS} | \Pi_R | \Psi_i \rangle}{\langle \Psi_{PS'} | \Psi_i \rangle} \\
&= -\frac{1}{\sqrt{3}} (\frac{1}{\sqrt{6}} - \frac{1}{\sqrt{3}})^{-1} \qquad (5.19) \\
&\approx 3.4
\end{aligned}
$$

Therefore, one can see that if we accept the interpretation of weak values of projectors as probabilities, then unless we post select our state carefully, we get probabilities that are greater than 1 and also negative. And because the claim of finding the Cheshire cat is only valid under this interpretation, it is most likely that we have not found the Cheshire cat.

6 Weak Value in a hidden variable theory

Here I shall investigating weak measurements using a semi-classical hidden variable theory called epistemically-restricted Liouville mechanics (ERL mechanics). The aim is to understand the interpretation of the weak value within a hidden variable theory of quantum mechanics.

6.1 Introduction to Wigner functions and Weyl transformations

Wigner functions introduced by Wigner in 1932 [15] maps a Quantum wave function to a probability distribution is phase space. It should however be thought of as a quasi-probability distribution as some wave functions result in Wigner functions that are negative in some regions of phase space. Here we shall consider only wave functions that have a non-negative Wigner function everywhere is phase space. A pure state that obeys this condition if and only if it is described by a wave function whose magnitude is Gaussian in configuration space i.e a wavefunction of the form

$$\Psi(q) = e^{\frac{1}{2}(aq^2+2bq+c)} \tag{6.1}$$

where a,b and c are complex numbers with $Re(a) > 0$ [14]. A sub-theory of Quantum mechanics about this specific subset of Quantum states is know as Gaussian Quantum theory. Since ERL mechanics is a hidden variable theory for Gaussian Quantum states, we shall consider, the Wigner function as proper probabilistic distribution within this theory. In this section, I fill very briefly introduce Wigner functions by including just key definitions and results. For a more detailed introduction, I refer the reader to [9].

The Wigner function is defined as

$$W(p,q) = \frac{1}{2\pi}\int e^{-ipy}\Psi(q+\frac{y}{2})\Psi^*(q+\frac{y}{2})dy \tag{6.2}$$

or alternatively as

$$W(p,q) = \frac{1}{2\pi}\int e^{-ipu}\tilde{\Psi}(p+\frac{u}{2})\tilde{\Psi}^*(q+\frac{u}{2})du \tag{6.3}$$

where $\Psi(q)$ is the position space wavefunction and $\tilde{\Psi}(p)$ is the momentum space wavefunction. The following results justify the Wigner function as a quasi-probabilistic distribution in phase space:

$$\int W(q,p)dp = \Psi(q)\Psi^*(q) \tag{6.4}$$

$$\int W(q,p)dq = \tilde{\Psi}(p)\tilde{\Psi}^*(p) \tag{6.5}$$

$$\int\int W_a(q,p)W_b(q,p)dqdp = \frac{1}{2\pi}|\langle\Psi_a|\Psi_b\rangle|^2. \tag{6.6}$$

The Weyl transform is used to transform an operator $\hat{A}$ from configuration space to phase-space. It is defined as

$$A(q,p) = \int e^{-ipy} \langle q + \frac{y}{2} | \hat{A} | \langle q - \frac{y}{2} \rangle dy \tag{6.7}$$

where $\hat{A}$ is basis of the position eigenstates. Equivalently, if $\hat{A}$ is in basis of the momentum eigenstates, then

$$A(q,p) = \int e^{-ipu} \langle p + \frac{u}{2} | \hat{A} | \langle p - \frac{u}{2} \rangle du. \tag{6.8}$$

If $\hat{A}$ is only a function of $\hat{q}$ i.e $\hat{A} = f(\hat{q})$, then

$$\begin{aligned} A &= \int e^{-ipy} \langle q + \frac{y}{2} | (\hat{q}) | \langle q - \frac{y}{2} \rangle dy \\ &=^{-ipy} f(q - \frac{y}{2}) \delta(y) dy \\ &= f(q) \end{aligned} \tag{6.9}$$

Similarly If $\hat{A}$ is only a function of $\hat{p}$ i.e $\hat{A} = f(\hat{p})$, then $A = f(p)$. The expectation of $\hat{A}$ can be calculated as

$$\langle \hat{A} \rangle = \int \int W(q,p) A(q,p) dq dp. \tag{6.10}$$

The wavefunction can be recovered from the Wigner function unto a phase factor by

$$\Psi(q) = \frac{1}{K} \int W(\frac{q}{2}, p) e^{ipq} dp \tag{6.11}$$

where $|K|$ can be determined by the normalisation of the wavefunction.

6.2 ERL mechanics

In this section, I will introduce ERL mechanics focussing on specific results which we will require to investigate weak measurements. For proofs of the results and a full description of this theory, I refer the reader to the original paper [8]. The main purpose of ERL mechanics is to show that it is possible to derive Gaussian Quantum mechanics from Liouville mechanics by adding an epistemic restriction. In other words, by applying a restriction on our knowledge of a classical system, which is described by a probability distribution in phase space, we are able to recover Gaussian Quantum theory.

6.2.1 Quantum mechanics

In Quantum mechanics, the commutations relation between the position and momentum of a system is be summarised by the expression

$$[\hat{z}_i, \hat{z}_j] = i\hbar\zeta_{ij} \tag{6.12}$$

where $\zeta_{ij} = \delta_{i,j+1} - \delta_{i+1,j}$, $\hat{z}_{2i-1} = \hat{q}_i$, that is the i^{th} position of the system and $\hat{z}_{2i} = \hat{p}_i$, which is its conjugate momentum.(Note just in this section, we are no longer using units where $\hbar = 1$ for the purposes of introducing the Classical uncertainty principle) For a state described by density matrix ρ, the vector of the means of its canonical operators is

$$\langle\hat{z}_i\rangle_i = Tr(\rho\hat{z}_i) \tag{6.13}$$

and it covariance matrix is

$$\gamma_{ij} = Tr(\rho(\hat{z}_i - \langle\hat{z}_i\rangle_i)(\hat{z}_j - \langle\hat{z}_i\rangle_j)) - i\hbar\zeta_{ij}. \tag{6.14}$$

The Quantum uncertainty principle can be expressed in terms of the covariance matrix as

$$\gamma + i\hbar\zeta \geq 0 \tag{6.15}$$

which is preserved by all unitary transformations.

6.2.2 Liouville mechanics

Classically, the relation between phase space can be summarised by an equation similar to its Quantum counterpart

$$\{z_i, z_j\} = i\hbar\zeta_{ij} \tag{6.16}$$

where the poisson brackets have replaced the commutator. Instead of a density matrix, a system is described by a probability density distribution in phase space $\mu(\vec{z})$ which belongs to the set $L_+(M)$

$$L_+(M) = \{\mu(\vec{z}) s.t. \mu : M \to R, \mu(\vec{z}) \geq 0, |\mu| = 1\} \tag{6.17}$$

where M is the phase space manifold i.e $\vec{z} \in M$ where the norm of a function is defined as

$$|\mu(\vec{z})| = \int_M d\vec{z}\mu(\vec{z}). \tag{6.18}$$

The covariance matrix of the system described by μ is

$$\gamma_{ij} = 2\langle(z_i - \langle z_i\rangle_i)(z_j - \langle z_i\rangle_j))\rangle. \tag{6.19}$$

The only constraint imposed on the covariance matrix in Liouville mechanics is

$$\gamma \geq 0. \tag{6.20}$$

Thus to recover the uncertainty principle, one must put in another restriction as an axiom. The restriction that achieves this result is called the Epistemic restriction and is defined as follows:

Epistemic Restriction [8]- A phase space distribution μ can describe an observer's state of knowledge if and only if it satisfies

1. The classical uncertainty principle:

$$\gamma + i\lambda\zeta \geq 0 \tag{6.21}$$

2. The maximum entropy principle: For a particular covariance matrix, the distribution should maximise the entropy

$$S(\mu) = -\int_M \mu(\vec{z}) log\mu(\vec{z}) d\vec{z}. \tag{6.22}$$

Setting $\lambda = \hbar$ recovers the Quantum uncertainty principle and the maximum entropy principle is necessary for bijective mapping, that is so that the number of possible phase space distributions does not exceed the number of quantum states for a given covariance matrix.

The distributions satisfying the epistemic restriction can be shown to be of the form [8]

$$\mu(\vec{z}) = \frac{1}{(2\pi)^n det\gamma^{1/2}} e^{1/2(\vec{z}-\langle\vec{z}\rangle)^T \gamma^{-1}(\vec{z}-\langle\vec{z}\rangle)}. \tag{6.23}$$

The classical uncertainty principle and the maximum entropy principle are only preserved under linear symplectic transformations. A symplectic transformation is defined as a square matrix A such that $A \dagger \zeta A = \zeta$ and transform the phase space vectors as

$$\vec{z} \rightarrow A \dagger \vec{z} \tag{6.24}$$

and a linear symplectic transformation is a the subset that acts lineally on the canonical coordinates. These are the only type of transformations allowed within ERL mechanics to ensure an epistemic state is always transformed to another valid epistemic state. A feature of these transformations is that the uncertainty about a system cannot decrease as a result of a symplectic transformation.

In Quantum mechanics, two quantities of a system, G and H can be jointly known only if they commute i.e.

$$[\hat{G}, \hat{H}] = 0 \tag{6.25}$$

but in ERL mechanics, G and H can be jointly known only if

$$\{\hat{G}, \hat{H}\} = 0 \tag{6.26}$$

where the Poisson bracket is defined as

$$\{\hat{G}, \hat{H}\} = \sum i(\frac{\partial G}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial G}{\partial p_i}\frac{\partial H}{\partial q_i}) \tag{6.27}$$

This is not true for deducing the state of a system in the past by post selection. For example, if the initial state of the system is one where we have complete certainty about its position q_s and then measure its momentum with outcome p_s, then we can deduce that before the measurement the system's position was q_s and momentum was p_s. In this way we know both the position and momentum of a system at a particular time in the past even though p and q do not commute with the poisson bracket. It is however not possible to jointly know the value of the variables that do not commute with the Poisson bracket for the present or be able to predict their values with certainty in the future. This is because of the disturbance induced by the process of measurement. The disturbance, caused by a measurement is of two forms. The first is the disturbance to the epistemic state, that is that the state collapses to the outcome of the measurement from a probability distribution over its values. The second disturbance is a physical disturbance due to the measuring device. This is a disturbance to the ontic state of the system. For example, in order to measure the position of a system, you couple it to the device under a Hamiltonian

$$H = \chi(t) q_s p_d \tag{6.28}$$

where q_s is the position of the system and p_d and $\int_0^T \chi(t) = \chi$ is the momentum of the device. From Hamiltonian mechanics we know that the equation of motion for the position of the device is

$$\begin{aligned} \frac{dq_d}{dt} &= \frac{\partial H}{\partial p_d} \\ \Rightarrow q_d &= q_{d,int} + \chi q_s \end{aligned} \tag{6.29}$$

where $q_{d,int}$ is the initial position of the device. So if we know the initial position of the device, then we can measure the position of the system. However, having complete certainty about the initial position of the device implies that we have no knowledge of its initial momentum. From the equation of motion of the system

$$\begin{aligned} \frac{dp_s}{dt} &= -\frac{\partial H}{\partial q_s} \\ \Rightarrow p_s &= p_{s,int} - \chi p_d \end{aligned} \tag{6.30}$$

one can see that the momentum of the system changes due to the momentum of the device, but since we have no knowledge of the value of p_d, we have no knowledge of the shift in momentum of the system. Hence, in the process of measuring the position of the system, we lose all knowledge of its momentum.

In ERL mechanics, all dynamics of physical systems are classical, that is all their ontic states behave classically. It is only the observer's lack of knowledge that gives the illusion of states behaving like quantum states. It has been shown that ERL mechanics is operationally equivalent to Gaussian quantum mechanics[8]. This means that any phenomenon that is possible in Gaussian quantum mechanics is reproducible using ERL mechanics. With this guarantee in mind, we will now proceed to find the weak value within ERL mechanics.

6.3 The weak value in ERL mechanics

In this section we will weakly measure the position of a system and then post select using the operator

$$A = \cos\theta q + \sin\theta p. \tag{6.31}$$

Let the state of the measuring device be described by the following covariance matrix (in this section we return to using units where $\lambda = \hbar = 1$):

$$\gamma_d = \begin{pmatrix} 2\sigma^2 & 0 \\ 0 & (2\sigma)^{-2} \end{pmatrix}$$

A state described by this covariance matrix is a minimum uncertainty state as $det(\gamma - i\zeta) = 0$. The probability distribution for this covariance matrix that obeys the maximum entropy principle is

$$\mu(q_d, p_d) = G_{a,\sigma}(q_d) G_{f,1/(2\sigma)}(p_d) \tag{6.33}$$

where

$$G_{a,\sigma}(q_d) = \frac{1}{2\sqrt{\pi}\sigma} e^{-\frac{(q_d - a)^2}{4\sigma^2}} \tag{6.34}$$

and

$$G_{f,1/(2\sigma)}(p_d) = \frac{1}{2\sqrt{\pi}(2\sigma) - 1} e^{-\frac{(p_d - f)^2}{4(2\sigma)^- 2}}. \tag{6.35}$$

The system is a state where we know its momentum and hence have no knowledge of its position. This is described by

$$\nu(q_s, p_s) = \delta(p_s - p_{s,int}) \tag{6.36}$$

The interaction Hamiltonian is

$$H = \chi(t) q_s p_d \tag{6.37}$$

where χ obeys

$$\int_0^T \chi(t)\, dt = 1. \tag{6.38}$$

If the position of the device before the measurement is $q_{d,int}$, then after the measurement, it will be shifted to $q_{d,int} + q_{s,int}$, i.e. by the position of the system. This is because

$$\frac{dq_d}{dt} = \frac{\partial H}{\partial p_d} = \chi(t) q_s. \tag{6.39}$$

The momentum of the device will remain unchanged as

$$\frac{dp_d}{dt} = -\frac{\partial H}{\partial q_d} = 0 \tag{6.40}$$

Hence the epistemic state of the device is now described by the probability distribution

$$\begin{aligned} \mu'(q_d, p_d) &= \int\int \mu(q_d - q_s, p_d)\nu(q_s, p_s)dq_s dp_s \\ &= \int \mu(q_d - q_s, p_d)dq_s \end{aligned} \tag{6.41}$$

The momentum of the system after the measurement is $p_{s,int} - p_{d,int}$ where $p_{d,int}$ is the momentum of the measuring device. This is because

$$\frac{dp_s}{dt} = -\frac{\partial H}{\partial q_s} = -\chi(t)p_d \tag{6.42}$$

and obviously its position is unchanged as it is the quantity we are trying to measure. Therefore the probability distribution of the system is now

$$\begin{aligned} \nu'(q_s, p_s) &= \int\int \nu(q_s, p_s + p_d)\mu(q_d, p_d)dq_d dp_d \\ &= \int \nu(q_s, p_s + p_d)G_{f,1/(2\sigma)}(p_d)dp_d \\ &= \int \delta(p_d - (p_{s,int} - p_s))G_{f,1/(2\sigma)}(p_d)dp_d \\ &= G_{f,1/(2\sigma)}(-p_s + p_{s,int}) \end{aligned} \tag{6.43}$$

In the limit where the uncertainty in p_d approaches 0

$$\lim_{(2\sigma)-1\to 0} = G_{f,1/(2\sigma)}(-p_s + p_{s,int}) = \delta(p_s - p_{s,int} + f) \tag{6.44}$$

Thus if $f = 0$

$$\begin{aligned} \nu'(q_s, p_s) &= \delta(p_s - p_{s,int}) \\ &= \nu(q_s, p_d) \end{aligned} \tag{6.45}$$

Hence if the condition for weakness is satisfied, that is the momentum of the measuring device is sharply peaked around 0, then the epistemic state of the system is unchanged. However this implies that the uncertainty in the position of the device $\sigma \to \inf$, which means we cannot measure the position of the system in one measurement. Hence we need to measure the position of a large ensemble of particles to determine the expectation of the position of the particle.

We post select states where $\cos\theta q_s + \sin\theta p_s = b$. We can use equation (6.41) to determine the probability distribution of the post selected ensemble of devices.

$$\begin{aligned} \mu''(q_d, p_d) &= \int \mu(q_d - (b - \sin\theta p_s)/\cos\theta, p_d)\nu'((b - \sin\theta p_s/)\cos\theta, p_s)dp_s \\ &= \int \mu(q_d - (b - \sin\theta p_s)/\cos\theta, p_d)\delta(p_s - p_{s,int})dp_s \\ &= \mu(q_d - (b - \sin\theta p_{s,int})/\cos\theta, p_d) \end{aligned} \tag{6.46}$$

We can therefore see that, the measurement causes the probability distribution of the position of the device to be shifted by the value $(b - \sin\theta p_{s,int})/\cos\theta$. One can see that this is because, the weak measurement does not disturb the initial state of the system leaving it almost unchanged. This means that even after the measurement, the measuring device still has its original momentum $p_{s,int}$ and since we only post select states where $\cos\theta q_s + \sin\theta p_s = b$, the position of the state is $(b - \sin\theta p_{s,int})/\cos\theta$, which is what the position of the device is shifted by.

Now let us calculate the Weak value for the equivalent quantum system. The weak value of the system is

$$\begin{aligned}\langle\hat{q}\rangle_W &= \frac{\langle b|\hat{q}|\Psi\rangle}{\langle b|\Psi\rangle} \\ &= \frac{\int q\langle b|q\rangle\langle q|\Psi\rangle dq}{\int\langle b|q\rangle\langle q|\Psi\rangle dq}\end{aligned} \tag{6.47}$$

where Ψ is the initial state of the particle which can be found,using (6.11) to be $|p_{s,int}\rangle$ and and $\langle b|$ if the eigenstate of operator $\hat{B} = \cos\theta\hat{q} + \sin\theta\hat{p}$ with eigenvalue b. The overlap $\langle b|q\rangle$ is

$$\langle b|q\rangle = \frac{1}{2\pi\sin\theta i}^{1/2} e^{-i(bq/\sin\theta - \cot\theta q^2/2)} \tag{6.48}$$

This is found by solving the following differential equation for $\Phi(q)$ which is the wave function of $|b\rangle$ in position space, that is $\hat{B}\Phi(q) = b\Phi(q)$.

$$\cos\theta q - i\sin\theta\frac{d\Phi}{dq} = b\Phi \tag{6.49}$$

The overlap $\langle q|\Psi\rangle = e^{-ip_{s,int}q}$ as $\Psi = |p_{s,int}\rangle$. Putting the overlaps back into (6.47) gives us the weak value of the position of the system, which is

$$\langle\hat{q}\rangle = (b - \sin\theta p_{s,int})/\cos\theta. \tag{6.50}$$

This is exactly the value the probability distribution of the position of the device is shifted by. Hence in this case, the weak value of the system does give a conditional expectation value. This is perhaps obvious for a hidden variable model, as the particle has a well defined position at all times which causes a proportional shift in the device. However this cannot immediately be extended to answer most of the surprising features of the weak value. This result is true only for a specific subset of quantum states.

7 Conclusion

From this report, it is clear that the interpreting the the weak value is problematic. It's interpretation is dependant on whether one considers the quantum state, the ontic state of a system or adopts a hidden variable interpretation where the quantum state is a manifestation of the observer's lack of knowledge. For example if one adopts the view of the quantum state being epistemic, then each particle's position produces a small shift in each device and hence after post selection, the probability distribution of the devices will indeed be shifted by the mean of the position of the particle. However if one adopts the ontic view of the quantum state, then it can be argued that if the particle is not in a position eigenstate, then the particle does not have a defined position value. In this case the weak value is little more than a manifestation of the entanglement between the particles and the device.

Whereas with other functions in quantum mechanics, the interpretation of quantum mechanics itself does not govern its proper use, we have seen that this is not the case with the weak value. We have seen how the using the weak value as truly a conditional expectation value, leads to some surprising effects such as the average spin of spin 1/2 particles being a 100 and probabilities that negative and larger than 1. This to some can be seen as evidence that such an interpretation of the weak value is incorrect. However, I think, the reason for the debate over the interpretation of the weak value despite it giving strange results is because the quantum theory is so strange in many ways, it is difficult to detect when a result is too strange to be real. As the axioms of quantum theory are not held sacred, one cannot immediately disprove a result by showing it contradicts an axiom of Quantum mechanics. If we did hold Quantum axioms sacred, then the fact that the weak value of spin of spin $\frac{1}{2}$ particles could be 100 would be proof that the weak value cannot be interpreted as an expectation value of a physical property of a particle under any circumstances. It can be argued that because we truly understand Quantum mechanics so little, it is difficult to understand how to interpret a quantity correctly within the theory.

On the other hand, if the interpretation of the weak value as a condition expectation violated any mathematical axioms, some may be more persuaded to believe that this interpretation is incorrect. This would then in turn be very strong evidence for that the quantum state is not epistemic. One might say that, interpreting the weak value as a conditional epectation already does violate a mathematical axiom by yielding negative probabilities or probabilities greater than 1. However, the concept of negative probabilities is not new in physics, which shows that this might not be enough to convince everyone that the weak value cannot be interpreted as the conditional value. I would like to conclude by emphasising that we can interpret the weak value as an expectation value if and only if we do not see the quantum state as the ontic state of the system, that is for particle's characteristic to be defined at anytime regardless of whether it is measured or not.

Reference

[1] Yakir Aharnov, David Z. Albert, A Casher, and Lev Vaidman. Surprising quantum effects. *Physical Letters A*, 124(4-5):199–203, 1987.

[2] Yakir Aharnov, David Z. Albert, and Lev Vaidman. How the result of a measurement on a component of the spin of a spin 1/2 particle can turn out to be a 100. *Physical Review Letters*, 60(14), 1988.

[3] Yakir Aharnov, Peter G. Bergmann, and Joel J. Lebowitz. The symmetry in the quantum process of measurement. *Physical Review*, 134(6B), 1964.

[4] Yakir Aharnov and Alonso Botero. Quantum averages of weak values. *Physical Review A*, 72(052111), 2005.

[5] Yakir Aharnov, Sandu Popescu, and Paul Skrzypczyk. Quantum cheshire cats. *arXiv:1202.0631v1*, 2012.

[6] Yakir Aharnov and Lev Vaidman. Properties of a quantum system during the time interval between two measurements. *Physical Review A*, 41(1), 1990.

[7] Yakir Aharnov and Lev Vaidman. The two-state vector formalism: an updated review. *arXiv:quant-ph/0105101v2*, 2007.

[8] Stephen Bartlett, Terry Rudolph, and Robert W. Spekkens. Reconstruction of gaussian quantum mechanics from liouville mechanics with an epistemic restriction. *arxiv:1111.5057v2*, 2012.

[9] William B. Case. Wigner functions and weyl transformations for pedestrians. *American Journal of Physics*, 76(10):937, 2008.

[10] Antonio Di Lorenzo. Hunting for the quantum cheshire cat. *arxiv:1205.3755v1*.

[11] I.M. Duck, P.M Stevenson, and E.C.G. Sudarshan. The sense in which a "weak measurement" of a spin 1/2 particle's spin component yields a value of 100. *Physical Review D*, 40(6):2112, 1989.

[12] Yelena Guryanova, Nicholas Brunner, and Sandu Popescu. The complete quantum cheshire cat. *arXiv:1203.4215v1*, 2012.

[13] Richard Jozsa. Complex weak values in quantum measurement. *Physical Review A*, 76:044103, 2007.

[14] Hudson R. L. When is the wigner quasi-probability density nonnegative. *Reports on Mathematical Physics*, 6(2):249–252, 1974.

[15] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Physical Review*, 40(749-759), 1932.