Weak Measurement in Quantum Mechanics

Author: Mile Vrbica
Mentor: Prof. Anton Ramšak

April 2020

Abstract

This seminar focuses on weak measurement with post-selection that introduces the so-called weak value, which is an experimentally accessible complex quantity that possesses non-intuitive behavior. The general measurement process and with an example is described. Properties of the weak value and its uses are discussed. The experimental support of the theory is provided and its applications are reviewed.

Contents

1 Introduction

In this seminar we shall examine the theory of quantum measurement. First we will describe the process of measuring a quantum state in detail. We will then introduce the concept of *weak measurement* which provides less information about the wave but has other benefits. In the realm of weak measurement we will observe some bizarre results.

It is very important to adjust our expectations about measurement theory. Even though we will describe the measurement process in more detail than merely stating the projection postulate, the fundamental measurement problem of quantum mechanics still remains. In the scope of this theory we cannot explain the projective nature of measurements, the Born rule or the wave function collapse. There is still a gap between the quantum world and our classical experience that is addressed by interpretations of quantum mechanics. We will not be dealing with this problematics in this seminar as we will be focused on describing interaction between the observed quantum system and the measurement device. This way we will be able to study the effects of measurement on the observed system, regulate the strength of interaction and still procure the necessary measurement statistics.

We will see in chapter 4 that the motivation behind weak measurements is not merely in the curiosity of a hopeless quantum enthusiast but in a powerful experimental application.

2 Quantum Measurement Theory

2.1 The Setup

We will study the measurement process as described by Bohm [1]. Let $\mathcal{S}$ be the observed quantum system and $\mathcal{M}$ the measurement device. Let the measurement device be dependent on a canonical variable q and lets denote its conjugate momentum p. It is common to call q the pointer as per von Neumann's nomenclature [2]. We can imagine the measurement device having a pointer and a gauge. The concept of pointers should be taken abstractly, as we only demand the pointer variable to be an observable that is directly accessible via classical measurement (e.g. position, not spin). The measurement device doesn't need to be an actual physical apparatus, it is rather a concept to help us grasp the measurement theory, and depends on the experimental setup. As we will see it is actually unnecessary for the measurement device wavefunction to be localized in q for a good measurement.

Our goal is to introduce correlations between the measured system and the apparatus in such way that we can deduce the sought after properties of the measured system through the apparatus only. Ideally we want the probability distribution of the apparatus in q space (or alternatively p space) to be a sum of well localized distributions around values of q (or p) that correspond to different eigenvalues of the measured observable. This way a pointer observed at a certain position will give us knowledge with large certainty about the post-measurement state of the quantum system and the corresponding eigenvalue of the observable.

We assume that the states of the observed system and measurement device are not entangled prior to interaction. Let $H_{\mathcal{S}}$ be the Hamiltonian of the measured system, $H_{\mathcal{M}}$ be the Hamiltonian of the measuring device and $H_{\text{int}}(t)$ be the interaction Hamiltonian which is non-zero only for a finite time. Before the measurement the systems are independent and they each evolve according to their own Hamiltonian. When we couple the systems and let them interact at $t = 0$ the total Hamiltonian acting on the composite system is $\mathcal{S} \otimes \mathcal{M}$ is

$$H = H_{\mathcal{S}} + H_{\mathcal{M}} + H_{\text{int}}(t). \tag{1}$$

We will now make a harmless assumption that the time of interaction is sufficiently small that Hamiltonians will alter the state of the system very little. We therefore neglect them and only deal with the interaction Hamiltonian.

2.2 The Interaction

Let A be the operator that acts on $\mathcal{S}$ and corresponds to the observable we are trying to measure. We let the systems interact in the time interval $[0, T]$ and introduce real normalized function $g(t)$ which is non-zero only during the time of interaction. For the H_S and H_M to be negligible one wants for T to be small, i.e. for the measurement to be impulsive. There are two main approaches to writing the measurement Hamiltonian - one requires us to measure q to get good measurement data, the other requires us to measure p. We will focus on the latter as it will be relevant when studying the Stern-Gerlach experiment in chapter 2.4. An appropriate interaction Hamiltonian has the form

$$H_{\text{int}} = -\mu g(t) A q. \tag{2}$$

Here μ is a constant factor that regulates the strength of the interaction. We could have, within reason, used analytic functions of the operators instead of using them bare, but since measuring an analytical function of a hermitian operator is just another hermitian operator we can just make them be the measured observables and save the ink. The unitary time evolution operator evaluated at any $t > T$ takes the form

$$U(T) = e^{-i\int_0^T H_{\text{int}} dt} = e^{i\mu A q}. \tag{3}$$

Appropriate units are assumed. Let us have an initial state of the system $|\psi\rangle = \sum_i c_i \chi(p) |a_i\rangle$, where $|a_i\rangle$ are the eigenstates of A with eigenvalues a_i and $\chi(p)$ the measurement device wavefunction in p representation. The state of the system after the interaction will be

$$U(T)|\psi\rangle = e^{i\mu A q} \sum_i c_i \chi(p) |a_i\rangle = \sum_i c_i e^{i\mu a_i q} \chi(p) |a_i\rangle = \sum_i c_i \chi(p - \mu a_i) |a_i\rangle . \tag{4}$$

Note that a_i is a scalar while q is still an operator. We remember that, since q is the canonical conjugate of p, e^{ikq} will generate a shift of the wavefunction in p space for k.

The effect of the interaction is now obvious. The wavefunction of the device gets displaced differently for each different eigenvalue of the observable and its amplitude is proportional to the probability amplitude of the corresponding component of the measured state.

It is of great importance to be aware that after such interaction we still need to perform strong measurement on the measuring device to get information about q or p. According to the postulates of quantum mechanics, such measurement projects the initial state (in our case after the interaction) onto an eigenspace of the observable that is randomly chosen obeying the Born rule. With our setup we only transferred the projective measurement process from the quantum system itself to the measurement device. We will see the usefulness of this approach in the following chapters.

Right after we measure p via classical means, the p will be well defined and the observed quantum system will be affected depending on the strength of correlation between the systems. In the Copenhagen interpretation one would say that the state has collapsed. This is irrelevant for our discussion, for reasons that will become apparent we are only interested in probability distributions.

Note that if we chose to put p into the Hamiltonian instead of q we would get displacement of the wavefunction in q-space rather than p-space.

2.3 Weak Measurement

A sensible choice for $\chi(p)$ is a Gaussian wavepacket. If the absolute differences between displacements $\mu|a_i - a_j|$ are much larger than the uncertainty of the wavepacket σ_p then the measurement is a *strong* one and correlation between $\mathcal{S}$ and $\mathcal{M}$ is strong. We can with large certainty know that a value of p measured around μa_i will in fact correspond to post-measurement state $|a_i\rangle$. On the contrary, if the overlap between neighbouring Gaussians is large, the measurement is called a *weak* one. Such a measurement can be achieved by making the initial σ_p large or by making the interaction weaker by making μ smaller. When performing a weak measurement we can never get as much information about the system as when performing a strong measurement. Suppose we observe the momentum p to have value of μa_i. Due to large overlap between neighbouring peaks we cannot be sure whether this measurement means truly indicates that the observed system is in post-measurement state $|a_i\rangle$. However, not all is bleak. Even with arbitrarily large σ_p we can determine the expectation value of A with arbitrary precision, provided we preform enough measurements on a large ensemble of identically prepared systems. This is due to the fact that uncertainty falls with $1/\sqrt{N}$ where N is the number of measurements. By the definition expectation value is completely independent of σ_p. Because the interaction is weak its disturbance to the observed system is much lesser than with strong measurement and herein lies one of the uses of the weak measurement.

2.4 The Stern-Gerlach experiment

We will apply the described formalism to the Stern-Gerlach experiment. The setting is usual: neutral particles with spin-1/2 fly into a inhomogeneous magnetic field which splits the up and down components of the wavepacket. Here our observed quantum system $\mathcal{S}$ corresponds to spin-1/2 Hilbert space with basis $|\uparrow\rangle, |\downarrow\rangle$, and our "measurement device" $\mathcal{M}$ corresponds to the spatial wave function of the particle. The two are initially uncorrelated. While this might seem odd at first, the formalism works and we will see that it makes sense.

We Taylor expand the magnetic field to the first order and neglect x and y components. The interaction Hamiltonian is

$$H_{\text{int}} = -\mu\sigma_z(1 + \alpha z), \tag{5}$$

where the constants are appropriately chosen. We can quickly see the similarity with eq. 2. The measured observable is σ_z and the pointer observable q is z. Values a_i are the eigenvalues of σ_z, 1 and -1. We assume that the particle is localized well enough that we can, without much error, assume it flies through the apparatus for a time T.

For simplicity we assume the spin of the incoming particle doesn't have a component in the y-direction. If we parametrize the incoming particle's spin with the polar angle θ then according to eq. 4, the post-interaction state will be

$$|\psi'\rangle = e^{i\mu}\chi(p - \mu\alpha)\cos\frac{\theta}{2}|\uparrow\rangle + e^{-i\mu}\chi(p + \mu\alpha)\sin\frac{\theta}{2}|\downarrow\rangle. \tag{6}$$

We see that the up component is given a momentum boost upwards and the down component is given a momentum boost downwards.

It might be surprising that it is momentum and not position that is being displaced. The resolution is that we only use measurement of position as means of measuring momentum. After the interaction the wave function is subjected to its kinetic Hamiltonian and a displacement in momentum will give boost the packet in position space as per Ehrenfest's theorem. We must therefore leave the particle to evolve on its own until we are satisfied with the separation of both up and down components in *position space*.

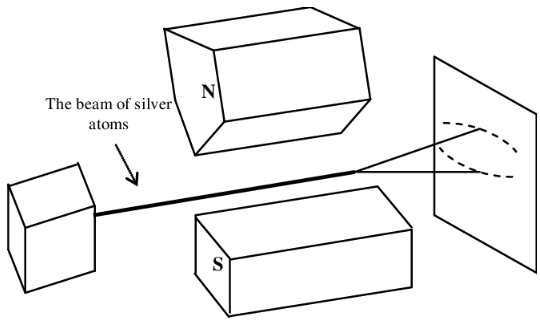

Figure 1: The Stern-Gerlach experiment scheme showing the centers of both deflected beams. The particles receive a boost of momentum in the Stern-Gerlach device that then deflects each component accordingly in position space. The screen must be distant enough for the packets to become sufficiently separated, taking their position uncertainty in account. Figure copied from [3].

3 Weak Measurement with Post-Selection

In this chapter we will describe some bizarre behavior surrounding weak measurements as it was first derived by Aharonov, Albert and Vaidman [4]. It turns out that with appropriate setup, it seems that the expectation value of, for example, spin-1/2 can be made arbitrarily large. For such results we must include a second measurement apparatus. We need to somehow be able to perform a *post-selection* on the state after the initial interaction. In formal language, we only want to observe the component of the post-interaction state that belongs to a certain subspace of the complete Hilbert space. This is realized by performing a strong measurement on a quantum system that has weakly interacted beforehand, thus projecting the state onto the subspace of our choice.

3.1 The Weak Value

We want to observe the certain components of the post-interaction state, so we perform a dot product with a chosen $|\phi_f\rangle$ from the $\mathcal{S}$ space.

$$\chi_f(p) = \langle\phi_f| e^{i\mu Aq} |\phi\rangle \chi(p) \tag{7}$$

When $\chi(p)$ is a packet curve with large σ_p (which is the case with weak measurement), we can linearize the exponent

$$\chi_f(p) \approx \langle\phi_f|\phi\rangle (1 + i\mu q A_w)\chi(p). \tag{8}$$

We introduced the *weak value* A_w of A

$$A_w = \frac{\langle\phi_f| A |\phi\rangle}{\langle\phi_f|\phi\rangle}. \tag{9}$$

Aware that we're still doing only a first order approximation, we can turn the linear function back into an exponential.

$$\chi_f(p) \approx \langle\phi_f|\phi\rangle e^{i\mu q A_w}\chi(p) = \langle\phi_f|\phi\rangle \chi(p - \mu A_w) \tag{10}$$

We emphasize again that this only holds true for sufficiently weak interactions. The uncertainty σ_p has to be larger than displacement μA_w.

Note that the weak value can be very big and even complex. Both real and imaginary part of the weak value can be obtained experimentally, since one roughly results in pointer momentum displacement and the other in pointer position displacement [5]. In the first order, the weak value (or more precisely, its real component) will be the expectation value of p, provided we measure only the accordingly post-selected states. It is of considerable interest that, in principle, the weak value is not bound even when the observable's spectrum is.

3.2 The Stern-Gerlach in-depth

We have derived that using post-selection, extreme expectation values can be achieved. The intuition behind this lies on very weak grounds. To remedy this, we will do a study on a series of two Stern-Gerlach devices without resorting to linearizing the time-evolution operator. Let the first Stern-Gerlach device have the magnetic field gradient in z direction and let the particle travel through it in y direction. We add a second

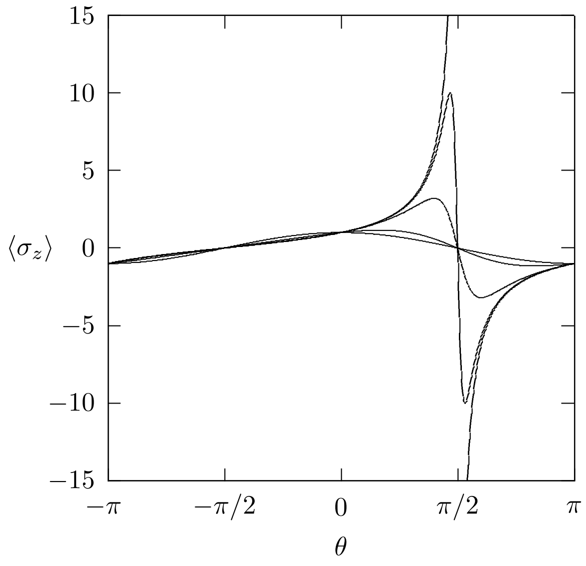

Figure 2: $\langle\sigma_z\rangle$ for $\Omega \in \{0, 0.5, 0.95, 0.995, 1\}$

Stern-Gerlach device which deflects particles in x direction for post-selection. Instead of modelling it like the first one, we will assume this one performs strong measurements. Lets choose to project particles on the direction of $|\sigma'\rangle$, which denotes a spin that points in $-x$ direction. We are looking for $\chi'(p) = \langle\sigma'|\psi'\rangle$ where we remember the post-interaction state $|\psi'\rangle$ from eq. 6

$$\chi'(p) \propto e^{i\mu} \cos\frac{\theta}{2}\chi(p-\mu\alpha) - e^{-i\mu}\sin\frac{\theta}{2}\chi(p+\mu\alpha). \tag{11}$$

Note that $\chi'(p)$ is not a normalized amplitude. If we assume is an even function $\chi(p)$ with mean 0 we can write the expression for the mean value of $\chi'(p)$,

$$\langle p\rangle = \mu\alpha\frac{\cos\theta}{1-\Omega\sin\theta}, \tag{12}$$

where Ω denotes the overlap between the 'up' and 'down' components multiplied by $\cos\mu$

$$\Omega = \cos\mu \int \chi^*(p+\mu\alpha)\chi(p-\mu\alpha)dp \tag{13}$$

and is assumed to be real. Ω is a measure of strength of interaction, as it roughly tells us how large the uncertainty is compared to the offset.

We remember that the displacement in p space is connected with the measured value of the σ_z via factor $\mu\alpha$, so that it is sensible to say that, considering only the post-selected particles, $\langle p\rangle = \mu\alpha\,\langle\sigma_z\rangle$. In the limit of $\Omega \to 0$ the measurement is a strong one and $\langle\sigma_z\rangle$ is confined within $[-1, 1]$ as one would intuitively expect. In the limit of $\Omega \to 1$ the measurement is a weak one and allows for outrageous expectation values. But even in the weak limit there is a relatively narrow interval of θ that allows for very large values (see Figure 2) It can be seen in a few lines of calculation that $\langle\sigma_z\rangle$ in the weak limit coincides with the weak value as defined in eq. 9

$$\sigma_{z,w} = \frac{\langle\sigma'|\,\sigma_z\,|\sigma\rangle}{\langle\sigma'|\sigma\rangle} = \frac{\cos\theta}{1-\sin\theta}. \tag{14}$$

The source behind this odd behavior is the normalization in the denominator. We have chosen not to work with the whole wave function but rather just a part of it. To interpret this part probabilistically we have to normalize it. The extreme values appear when θ is around $\pi/2$ i.e. when $\langle\sigma'|\sigma\rangle$ is close to 0. It is when there is a very small probability of particles deflecting in our chosen direction where unusual behavior is observed.

We have neglected the effects of spreading on the particle's journey to the screen as well as when passing through the device. While the exact distribution is certainly altered, the concept remain and so does the behavior of the weak value.

4 Experimental Verification and Applications

Some highly significant and diverse uses have been found for weak value in experimental settings. On one side of the spectrum the concept using the complex nature of weak value can be utilized to measure the full wave function directly, avoiding tomography. Usually one would repeatedly have to measure a very large array of observables and reconstruct the wave function from the probability distribution. Using weak values, one can access the wave function directly. Consider weakly measuring the observable $|x\rangle\langle x|$ on the state $|\psi\rangle$ and post-selecting the momentum eigenstates $|p\rangle$. According to eq. 9 the weak value of such an observable is

$$x_w = \frac{\langle p|x\rangle \langle x|\psi\rangle}{\langle p|\psi\rangle} = \frac{e^{ipx}\psi(x)}{\phi(p)}. \tag{15}$$

$\psi(x)$ and $\phi(p)$ are the wave functions in position and momentum space respectively. We see that one can, with only very minor post processing, access $\psi(x)$ in strong contrast to tomographical approach. This principle has been experimentally realized in various settings, for example determining the wave function of a photon [6, 7], a 27-dimensional orbital-angular-momentum state vector [8] and has even been generalized to mixed states [9]. Needless to say this method has great potential in the field of quantum computation. Weak value showcases its complex prowess also in measuring expectation values of non-Hermitian operators [10].

As we can see from eq. 12 the weak value of the measured obseravble is multiplied by a factor that depends on the magnetic field gradient. Because the weak value can be made to be very large, a very small magnetic field gradient can result in a very large offset on the actual screen. The process can therefore be used to amplify very subtle evolution parameters so that we aren't limited by small resolution of the actual measuring apparatus. This has been indeed experimentally utilized in many settings to measure small angular, temporal, frequency, velocity and temperature shifts as well as beam deflection [11].

We shall take a look at a simple example of an experimentally realizable setup described in [11]. Consider a Gaussian laser beam that first goes through a collimating lens and then through a half-wave and a quarter-wave plate. In Figure 3 a) the beam goes directly through a polariser that performs post-selection. A CCD is used to measure the position dependant beam intensity. In Figure 3 b) a birefringant crystal is placed between the plates and the polariser. The crystal spatially separates the horizontal and vertical components of the incoming beam. An appropriate lens is after the crystal to image the transverse position on the output face of the crystal onto the CCD. For measuring the weak value of polarisation the crystal should offset the beams only very slightly so that they can still interfere with each other. With the right choice of the orientations of all the elements in the setup a distribution that reflects the real part of the weak value can be detected on the CCD. A similar setup can be made to measure the imaginary part of the weak value.

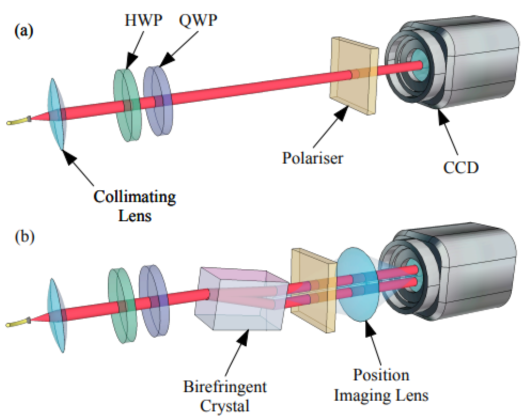

Figure 3: A simple setup for measuring the weak value of polarisation. Figure copied from [11].

5 Conclusion

We have (within our scope) defined the concept of a measurement and coupled it with post selection to come to physically significant results. We have described an experimentally plausible setup in detail and studied the unusual results in detail. Above all it should be clear that the concept of weak value is not simply a bit of quantum-mechanical curiosity but is a non-trivial tool that plays an increasingly significant role in experimental setups. Besides the practical use, there is also much theoretical interest in the implications of the weak value [12, 13].

References

[1] D. Bohm. *Quantum Theory*. Dover Books on Physics. Dover Publications, 2012.

[2] J. von Neumann, R.T. Beyer, and N.A. Wheeler. *Mathematical Foundations of Quantum Mechanics: New Edition*. Princeton University Press, 2018.

[3] Didiş Nilüfer and Şakir Erkoç. History of science for science courses: "spin" example from physics. *Latin-American Journal of Physics Education*, 01 2009.

[4] Yakir Aharonov, David Z. Albert, and Lev Vaidman. How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. *Phys. Rev. Lett.*, 60:1351–1354, Apr 1988.

[5] Richard Jozsa. Complex weak values in quantum measurement. *Physical Review A*, 76, 06 2007.

[6] Jeff S. Lundeen, Brandon Sutherland, Aabid Patel, Corey Stewart, and Charles Bamber. Direct measurement of the quantum wavefunction. *Nature*, 474(7350):188–191, 2011.

[7] Charles Bamber and Jeff S. Lundeen. Observing dirac's classical phase space analog to the quantum state. *Phys. Rev. Lett.*, 112:070405, Feb 2014.

[8] Mehul Malik, Mohammad Mirhosseini, Martin P. J. Lavery, Jonathan Leach, Miles J. Padgett, and Robert W. Boyd. Direct measurement of a 27-dimensional orbital-angular-momentum state vector. *Nature Communications*, 5(1):3115, 2014.

[9] Jeff S. Lundeen and Charles Bamber. Procedure for direct measurement of general quantum states using weak measurement. *Phys. Rev. Lett.*, 108:070402, Feb 2012.

[10] Gaurav Nirala, Surya Narayan Sahoo, Arun K. Pati, and Urbasi Sinha. Measuring average of non-hermitian operator with weak value in a mach-zehnder interferometer. *Phys. Rev. A*, 99:022111, Feb 2019.

[11] Justin Dressel, Mehul Malik, Filippo Miatto, Andrew Jordan, and Robert Boyd. Colloquium: Understanding quantum weak values: Basics and applications. *Review of Modern Physics*, 86:307, 03 2014.

[12] Yakir Aharonov and Lev Vaidman. Properties of a quantum system during the time interval between two measurements. *Phys. Rev. A*, 41:11–20, Jan 1990.

[13] Lev Vaidman. Weak-measurement elements of reality. *Foundations of Physics*, 26(7):895–906, 1996.