4

Weak Measurement

The physics of weak measurement is a special case of the previous chapter where the measurement strength is taken to zero as a limit. The limit is useful because the measurement back-action goes to zero; but as we will see later in this chapter, the amount of information also limits to zero. By taking many realizations of the weak measurement process from identically prepared initial conditions, information about the system of interest may be gradually extracted in such a way that the quantum state disturbance per measurement is negligible.

4.1 The Limit of a Very Weak Stern–Gerlach Magnet

Let us consider a motivating experiment, where we take a Stern–Gerlach device with a very weak magnetic field gradient, or a piece of calcite crystal that is just a sliver, as was done in the experiment of Ritchie et al. (1991b) (see Fig. 4.1).

In this limit, the splitting of the atomic or optical beam corresponding to the different system state is very small, so the ability to distinguish the system is nearly nonexistent. In the limit $d \ll \sigma$, we Taylor expand the measurement operators of the calcite crystal case (3.12) to first order in d to find

$$\hat{\Omega}_x = \phi(x)\hat{\Pi}_H + \phi(x+d)\hat{\Pi}_V \approx \phi(x)\mathbf{1} + d\phi'(x)\hat{\Pi}_V, \tag{4.1}$$

where $\phi'(x)$ indicates the spatial derivative, and we have used the projection operator property $\hat{\Pi}_H + \hat{\Pi}_V = \mathbf{1}$.

In this limit, we can determine the probability density to find the photon at position x, given an initial polarization state $|\psi\rangle = a|H\rangle + b|V\rangle$ as

$$p(x) = \langle\psi|\hat{\Omega}_x^\dagger\hat{\Omega}_x|\psi\rangle \approx |\phi(x)|^2 + 2d\,\mathrm{Re}[\phi'(x)\phi(x)^*]|b|^2 + \hat{\mathcal{O}}(d^2). \tag{4.2}$$

As you will show in Exercise 4.5.2, this distribution is normalized to first order in d. Notice that if $d = 0$, the polarization has no role to play, but for small d, there

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

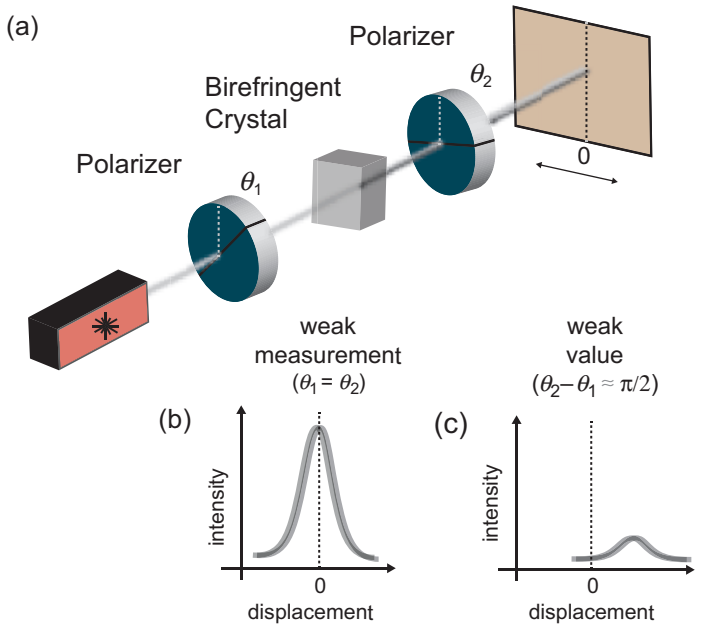

Figure 4.1 (a) In the experiment of Ritchie, Story, and Hulet (Ritchie et al., 1991b), a laser source is used to produce a collimated beam that is passed through a pair of polarizers with transmission angles θ_1 and θ_2. In between these polarizers is a birefringent crystal, which generates a small deflection in one polarization direction. (b) Without the final polarizer, a weak measurement is performed and the output intensity is the combination of two, very slightly shifted Gaussian profiles, depending on the initial polarization. (c) If the two polarizers are tuned to near orthogonality, the two Gaussian profiles destructively interfere, reducing the overall intensity but causing a sizable shift of the pattern, readily revealing the weak value associated with this choice of input and output states.

is a slight shift of the probability distribution to find the photon at position x, that involves the population of the vertical polarization, $|b|^2$.

The new polarization state, assuming the photon passes a thin slit at position x, is given by

$$|\tilde{\psi}\rangle = \frac{\hat{\Omega}_x|\psi\rangle}{||\hat{\Omega}_x|\psi\rangle||} \propto \phi(x)|\psi\rangle + d\phi'(x)b|V\rangle \propto |\psi\rangle + d(\ln\phi(x))'b|V\rangle, \qquad (4.3)$$

where, in the last equality, we have replaced $\phi'(x)/\phi(x) = (\ln\phi(x))'$. In Eq. (4.3), we have kept the state unnormalized to show the main effect – the state is still almost the same as $|\psi\rangle$, but with a small correction in the direction of the vertical polarization. One interesting feature of this result is that the amount of shift depends on the position x. For example, if we take a Gaussian meter state,

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

$\phi(x) \propto \exp(-x^2/(4\sigma^2))$, then $(\ln\phi(x))' = -x/(2\sigma^2)$. This gives a prefactor of $-dxb/(2\sigma^2)$ in front of the $|V\rangle$ state of Eq. (4.3). On one hand, we notice that if for some reason, x were to take on a large value, this correction could be, in fact, quite large. On the other hand, from the probability distribution (4.2), we see that such values of x are highly improbable. Nevertheless, these values are possible and underpin the effect of the *weak value* we will discuss later in this chapter.

Now let us generalize this result, but considering an interaction strength g between a system and meter. In this case, we can express the $\hat{E}_j$ system operator in the limit of very small g as

$$\hat{E}_j = p_j \mathbf{1} + g\Gamma_j \hat{\mathcal{O}}, \tag{4.4}$$

where p_j is the meter probability of finding result j in absence of the system, Γ_j is a coefficient associated with the detector, while $\hat{\mathcal{O}}$ is a system operator. Such a measurement is defined as a *weak measurement* of system operator $\hat{\mathcal{O}}$. For such a decomposition, we can find the probability of the meter result j as

$$P_j = \mathrm{Tr}(\hat{\rho}\hat{E}_j) = p_j + g\Gamma_j\langle\hat{\mathcal{O}}\rangle, \tag{4.5}$$

where $\langle\hat{\mathcal{O}}\rangle$ is the expectation value of the system operator $\hat{\mathcal{O}}$ in the state $\hat{\rho}$. Notice that, in the previous chapter, the general theory of measurement lost connection with the concept of measuring a system observable. Here, in the weak measurement limit, we regain it. Knowledge of the detector parameters p_j, Γ_j and the coupling strength g allows us to infer the expected system observable $\langle\mathcal{O}\rangle$ given sufficient trials in identically prepared initial conditions.

4.2 Information–Disturbance Trade-off

We have seen in the previous section that, in the weak measurement limit, there is typically little disturbance to the quantum state, but also that typically only a little information can be extracted per measurement. Let us now quantify that relationship. Returning to the simple case of plain (or QND) measurement of Section 3.7, where the operators E_k are all diagonalized in the same basis, (3.50), then the probability of result k (3.51) contains the accessible information about the quantum state. Let us view this situation as the probability of distinguishing which of the j states the system resides in. The information acquired from the measurement can be quantified by the statistical overlap of the distributions

$$C_{ij} = \sum_k \sqrt{p(k|i)p(k|j)}. \tag{4.6}$$

We have already encountered this quantity as the Bhattacharyya coefficient. A *quantum limited detector* corresponds to when the information acquisition equals

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

the ensemble-averaged degree of decoherence, defined as $d_{ij} = -\ln(\sum_k P_k \rho_{ij}^{(k)}/\rho_{ij})$ of the coherence of states i and j. More generally, we have the inequality

$$D_{ij} \leq d_{ij}, \tag{4.7}$$

where again $D_{ij} = -\ln C_{ij}$ is the Bhattacharyya distance. This inequality implies that not all possible information about the quantum state is usually revealed to the detector, and information can be lost in a variety of ways. These include not measuring in the optimal basis, not using the optimal meter states, and more practically, loss, inefficiency of the detector, and other sources of technical noise, such as detector dark counts, thermal noise, and the like. Such effects either increase decoherence or decrease distinguishability of the different quantum states. A more general situation can occur with even quantum limited detectors, where the decomposition (3.19) is considered with the unitary term U_j reintroduced. In this case, an ensemble average will still result in decoherence, but only because of the different unitaries applied for different members of the ensemble. On any given run, perfect coherence is retained in the quantum system. In fact, the effect of the outcome-dependent unitary can be eliminated with the use of suitable coherent feedback, a subject we will return to in Chapter 10.

Case Study: Selective Tunneling and a Null-Result Measurement

Better intuition about these results can be obtained by studying some physical examples. In this section, we take the case of selective tunneling, which gives a natural strength parameter for the measurement (see Fig. 4.2). This experiment was carried out in the group of John Martinis (Katz et al., 2006). We give a slightly simplified description of the experiment below.

The phase qubit is defined as a superconducting loop, interrupted by a Josephson junction, and controlled by an external magnetic flux Φ_e. The phase across the junction ϕ obeys an equation of motion

$$C\Phi_0^2 \ddot{\phi} + E_J \sin\phi = \Phi_0 I, \tag{4.8}$$

where C is the capacitance of the junction, E_J is the Josephson energy, and $\Phi_0 = h/(2e)$ is the magnetic flux quantum. The external current controlling the junction I is here replaced by the external flux control via the loop inductance $I = \Phi_e/L$. This dynamical equation may be interpreted as describing a fictitious particle of mass $C\Phi_0^2$ moving in a potential well given by

$$V(\phi) = -\frac{\Phi_0 \Phi_e}{L}\phi - E_J \cos\phi. \tag{4.9}$$

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

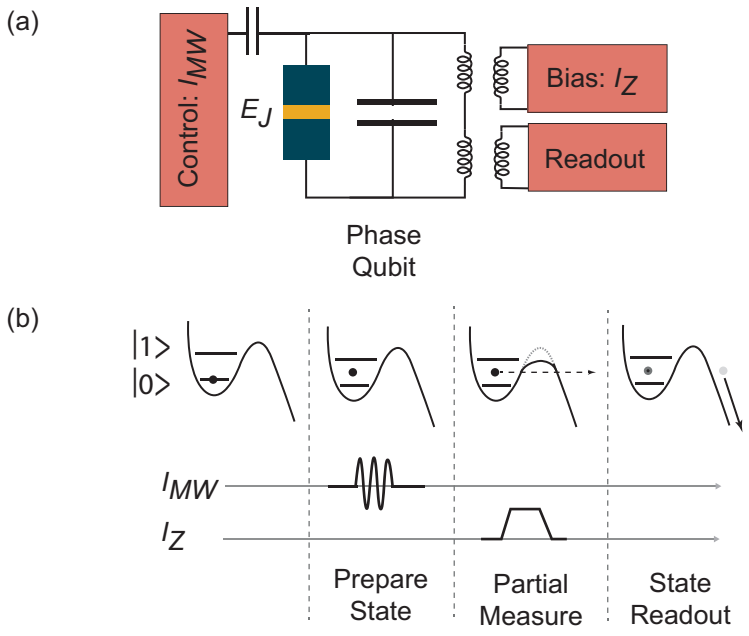

Figure 4.2 (a) A phase qubit generally consists of a capacitively shunted Josephson junction that is biased with a static current I_Z to reduce the number of quantized levels in the potential energy surface to a few. In this particular realization, the current bias is applied through a so-called flux transformer; a similar structure is also used to sense a voltage that develops when the junction switches to the dissipative state. A microwave signal I_{MW} is used to drive transitions between the quantum levels. The basic pulse sequence utilized in the experiment carried out by Katz et al. (2006) is shown in (b): a microwave pulse is used to create a desired quantum superposition of the ground and excited states; the tunneling barrier is lowered for a short duration to perform a partial measurement; the state of the system is inferred by associating the presence of a voltage pulse with occupancy of the excited state, originating from state-selective quantum tunneling out of the potential well.

Such a potential that oscillates while having a slope is often called a "tilted washboard" potential. When the tilt of the washboard gets too large, the local minima disappear, and the potential has no stable solutions.

The phase qubit is arranged so that the applied flux is just under the critical amount, so shallow minima exist that are 2π periodic. A quantum treatment of the circuit imposes the canonical commutation relation on the charge and flux degrees of freedom, $[\hat{\phi}, \hat{Q}] = i\hbar$, so the quantum physics of the circuit gives rise to quantized energy levels, coherent superpositions, and the rest of the quantum phenomena. Among those quantum phenomena is *quantum tunneling*, where a particle can escape from a metastable potential well, even if its energy is less than the

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

barrier height. This phenomenon is very important for this section. The tunneling rate for escaping the potential well takes the form

$$\Gamma = \omega_a e^{-S_c}, \tag{4.10}$$

where ω_a is the attempt frequency, while $S_c(E)$ is the classical action of the escape path in an inverted potential, given an energy E. We consider the simple case where the potential well is sufficiently shallow that there are only two energy levels, named $|g\rangle$ with energy E_g and $|e\rangle$ with energy E_e, for ground and excited state. The higher energy of the excited state (see Fig. 4.2) implies that the associated classical action of the escape path is smaller than the same for the ground state. The exponential dependence of the tunneling rate on this action leads us to the approximation that the ground state escape rate can be set to 0 for the timescale of the experiment we will consider, so only the excited state $|e\rangle$ will have a nonzero escape rate Γ.

We are interested in the tunneling events as a state-selective event – the upper level can tunnel out, while the lower level cannot. The tunneling event can be detected in the following way: if the particle tunnels out of the well, it has sufficient energy to run down the washboard without getting trapped in any metastable states. This results in the phase ϕ increasing rapidly in time, which gives rise to an additional magnetic flux through the ring. This can be detected with a sensitive magnetometer, which will register an event when the tunneling occurs. In the experiment of Katz et al. (2006), this was a "DC SQUID," also a ring, but with two Josephson junctions and connected to leads with a current source that gives rise to interference.

Let us analyze the measurement physics of this situation. Consider an arbitrary initial state of the phase qubit – it can be prepared by applying an electromagnetic signal, oscillating at the frequency of $(E_e - E_g)/\hbar$, to drive a "Rabi oscillation" of variable time to make any state on demand. Setting the clock to $t = 0$ when the initial state is prepared, let us wait some time t and see if the detector registers a tunneling event (named d) or not (named $\bar{d}$). If the system is prepared in the excited state, the event d (or $\bar{d}$) will occur with probability $p_d = 1 - \exp(-\Gamma t)$ and $p_{\bar{d}} = \exp(-\Gamma t)$, which can be found by solving a simple rate equation, $\dot{p}_{\bar{d}} = -\Gamma p_{\bar{d}}$. The lower level will not decay on the timescales of the experiment. That information is sufficient to construct the operators $\hat{E}_d$ and $\hat{E}_{\bar{d}}$ corresponding to the two possibilities, after waiting time t,

$$\hat{E}_d = [1 - \exp(-\Gamma t)]|e\rangle\langle e|, \quad \hat{E}_{\bar{d}} = \exp(-\Gamma t)|e\rangle\langle e| + |g\rangle\langle g|. \tag{4.11}$$

We can now find the probabilities of finding either a no-tunneling event or a tunneling event if the system is prepared in the state $\hat{\rho}$, with matrix elements ρ_{ij}, where $i,j = e,g$. Applying the formalism of the previous chapter, we find

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

$$P_d = \mathrm{Tr}(\hat{\rho}\hat{E}_d) = \rho_{ee}(1 - e^{-\Gamma t}), \quad P_{\bar{d}} = \mathrm{Tr}(\hat{\rho}\hat{E}_{\bar{d}}) = \rho_{gg} + \rho_{ee}e^{-\Gamma t}. \tag{4.12}$$

As is clear from the preceding formulas, the distribution has positive probabilities and is normalized. At time $t = 0$, the detector brings no information, so we always register a null result. However, after waiting a time $t > \Gamma^{-1}$, then we begin to have better distinguishability between the states $|e\rangle$ and $|g\rangle$: the detection event is associated with the system being in the excited state, while the null event is associated with the system being in the ground state. We stress an important observation: the no-result measurement (when the detector sits quietly, doing nothing) also brings information. This is reminiscent of the incident of the "dog that did not bark," told by Sir Arthur Conan Doyle in one of his short stories, "The Adventure of Silver Blaze" (1894). The following exchange occurs between Gregory, the Scotland Yard detective, and Sherlock Holmes:

- **Gregory:** Is there any other point to which you would wish to draw my attention?
- **Holmes:** To the curious incident of the dog in the night-time.
- **Gregory:** The dog did nothing in the night-time.
- **Holmes:** That was the curious incident.

Just like the dog *not* barking brought the new information needed to solve the mystery, the detector *not* clicking brings new information and indeed, as we shall see, results in a change of system state.

To find the quantum state dynamics of the system, we must account for the measurement dynamics, as well as the Hamiltonian. Fortunately, both of these effects are diagonal in the $|g\rangle, |e\rangle$ basis (a QND measurement), so we can combine them together. Later on the book, we will encounter examples of the unitary dynamics and the measurement dynamics that interfere with each other. Let us first consider the case where the system tunnels, resulting in a detection event. This corresponds to a *destructive* event, because once this occurs, the state no longer occupies the metastable well, and the system no longer exists. If such an event occurs, the system must be reset in order to start over. This phenomenon is similar to photo-detection. In that case the detection of the photon also goes together with destroying it, so further experiments can no longer take place. We can say, though, that before the phase qubit tunneled and was destroyed, it was in its excited state. However, the no-tunneling event is more interesting. One might be tempted to simply project the phase qubit onto the ground state, but this would be incorrect. We can apply the formalism of the last chapter to find the measurement operator,

$$\hat{\Omega}_{\bar{d}} = e^{i\omega t/2}|g\rangle\langle g| + e^{-i\omega t/2 - \Gamma t/2}|e\rangle\langle e|, \tag{4.13}$$

where the frequency $\omega = (E_e - E_g)/\hbar$ is introduced to account for the system Hamiltonian. The post–null measurement state can now be calculated,

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

$$\hat{\rho}' = \frac{\hat{\Omega}_{\bar{d}}\hat{\rho}\hat{\Omega}_{\bar{d}}^{\dagger}}{P_{\bar{d}}} = \frac{1}{\rho_{gg} + e^{-\Gamma t}\rho_{ee}} \begin{pmatrix} \rho_{gg} & \rho_{ge}e^{i\omega t - \Gamma t/2} \\ \rho_{eg}e^{-i\omega t - \Gamma t/2} & \rho_{ee}e^{-\Gamma t} \end{pmatrix}. \tag{4.14}$$

In the last expression, the density matrix is expressed in the $|g\rangle, |e\rangle$ basis. It is important to stress that, given the no-tunneling event, the prediction for the conditional state (4.14) is deterministic. We note that the new conditional density matrix is properly normalized. We can now associate the state disturbance with the information gain. For time $t = 0$, we have gained no information, but we also see that $\hat{\rho}'(t = 0) = \hat{\rho}$. In the opposite limit, for $\Gamma t \gg 1$, the postmeasurement state (corresponding to no-tunneling) is simply a projection operator on the ground state, corresponding to statistical certainty that the qubit must be in its ground state. We see there is a continuous transition between these two extremes controlled by the amount of time waited for an event that never comes.

These state dynamics may be interpreted as a partial collapse of the wavefunction, controlled by the measurement strength Γt. The fact that the system does not collapse to the ground state immediately comes from the fact that there are two possible explanations for the absence of the detection event: either the qubit is in its ground state the entire time, or it is in its excited state, but we simply have not given it enough time to tunnel out of the barrier. It is important to note that if the system begins in a coherent superposition of the ground and excited states that is pure, it stays in a pure state throughout the gradual collapse process.

4.3 Weak Value

Now that we understand weak measurements, we can combine them together with more weak measurements, or with projective (or strong) measurements. An interesting application of weak measurements is the *weak value*, a concept introduced by Aharonov et al. (1988). We motivate the effect by considering the experiment of Ritchie et al. (1991a) in the group of R. Hulet, where a thin birefingent crystal (quartz, in this case) is followed by a polarizer, to select a given polarization of one's choosing. We continue the treatment of Section 4.1, and rather than place a screen with a pinhole, the experiment is terminated with a polarizer followed by a detecting screen.

Let us first revisit the case with no final polarizer. The thin quartz crystal weakly measures $\hat{\Pi}_V$ as before. Recalling the probability of finding the photon at position (4.2), the probability density for a photon to land at position x is given by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \left(1 - \frac{xd|b|^2}{\sigma^2}\right), \tag{4.15}$$

where we have assumed a Gaussian meter wavefunction with width σ. We note the distribution is normalized, with a mean of $\langle x \rangle = -d|b|^2$, so the mean is shifted by

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

anywhere between $[-d, 0]$, which we can associate with d times the eigenvalues of the $-\hat{\Pi}_V$ projection operator.

Now, insert the polarizer. Suppose the initial polarization state is $|\psi\rangle = a|H\rangle + b|V\rangle$, as before, and the polarizer is tilted to allow only the polarization $|\tilde{\psi}\rangle = a'|H\rangle + b'|V\rangle$ to pass through, where a', b' are different coefficients. Suppose the measurement is weak ($d \ll \sigma$), so the probability of passing the polarization is approximately given by Malus' law, $P_{pass} = |\langle\tilde{\psi}|\psi\rangle|^2$, given by the squared cosine of the relative angle of the photon polarization and the polarizer. To find where on the screen the photon is likely to land, we can return to our initial discussion of this effect, (3.8), and project the meter state onto position $|x\rangle$, while we project the system state with the projection operator $\hat{\Pi}_{\psi'} = |\tilde{\psi}\rangle\langle\tilde{\psi}|$. This procedure leaves the meter state as

$$\phi(x) = \phi_0(x)\langle\tilde{\psi}|\psi\rangle + d\phi_0'(x)\langle\tilde{\psi}|V\rangle\langle V|\psi\rangle, \tag{4.16}$$

where we denote $\phi_0(x)$ as the original Gaussian meter state. Taking all states and wavefunctions to have real values for simplicity, we can simplify the probability density for where the photon will land as

$$p_{pol}(x) = |\langle\tilde{\psi}|\psi\rangle|^2 \phi_0^2 \left(1 - \frac{xd}{\sigma^2}\frac{\langle\tilde{\psi}|V\rangle\langle V|\psi\rangle}{\langle\tilde{\psi}|\psi\rangle}\right). \tag{4.17}$$

The first term, $|\langle\tilde{\psi}|\psi\rangle|^2$, we have already interpreted as the probability for the photon to pass the polarizer. Accordingly, the probability for the photon to land anywhere on the screen is given by

$$\int_{-\infty}^{\infty} dx\, p_{pol}(x) = |\langle\tilde{\psi}|\psi\rangle|^2, \tag{4.18}$$

the same as if there were no spatial mode. However, we can ask a more probing question: given that the photon passes the polarizer, where will it land? We can find the mean position as

$$\langle x\rangle_{ps} = \frac{\int_{-\infty}^{\infty} dx\, x p_{pol}(x)}{\int_{-\infty}^{\infty} dx\, p_{pol}(x)} = -d\frac{\langle\tilde{\psi}|\hat{\Pi}_V|\psi\rangle}{\langle\tilde{\psi}|\psi\rangle}, \tag{4.19}$$

where the variance is unchanged: σ^2. We see the presence of the polarizer has shifted the center of the distribution from $-d|b|^2 = -d|\langle V|\psi\rangle|^2$ to a new position, $-d(\hat{\Pi}_V)_w$, where we define the weak value of the operator $\hat{\Pi}_V$ as

$$(\Pi_V)_w = \frac{\langle\tilde{\psi}|\hat{\Pi}_V|\psi\rangle}{\langle\tilde{\psi}|\psi\rangle}. \tag{4.20}$$

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

This new position shift controlled by the weak value involves all the ingredients: the initial state $|\psi\rangle$, the final state $|\tilde{\psi}\rangle$, and the operator that is being weakly measured, $\hat{\Pi}_V$.

Notice that some strange things are happening with this new shift. When the initial polarization state of the photon and the final state of the polarizer become orthogonal, $\langle\tilde{\psi}|\psi\rangle \rightarrow 0$, corresponding to $aa' + bb' \rightarrow 0$, the weak value can diverge since we can arrange that bb' remains finite. Therefore, while without the polarizer, the beam shift is bounded by the eigenvalues of the operator that is being weakly measured (scaled by the weak measurement strength), with the polarizer, the shift can be made arbitrarily larger than the eigenvalue bounds! When applied to a similar effect in the Stern–Gerlach device, Aharonov, Albert, and Vaidman called this effect "How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100."

Let us understand in greater detail how this interesting effect can arise. Recall that the conditional shift in the polarization shift of the photon, provided it passed through a slit at position x, is given by Eq. (4.3). Notice that, while usually the polarization of the photon is only weakly rotated by the weak measurement, occasionally very large values of position occur (although they are very unlikely). When $|x| > \sigma^2/d$, this leads to the rotation of the polarization by a large amount. We can interpret that to mean that although the coupling is weak, a rare event can make a large change in the system state. The weak value is similar to the reverse of this effect. When we demand that the photon be in a nearly orthogonal state to the polarization it began in, only very large values of x are capable of doing this, so their average must be anomalously large. Notice also that the divergence of the weak value coincides with the probability of the photon passing through the polarizer going to zero. This implies an important trade-off between how large the weak value can be, and how frequently the anomalous shift can occur. We will return to this point later in the chapter.

Weak Value Derivation within the von Neumann Model

Consider the von Neumann model with $H_{int} = g\delta(t)\hat{p} \otimes \hat{A}$, where $\hat{p}$ is the meter momentum, and $\hat{A}$ is a general system operator. Here we consider a finite-dimensional Hilbert space. Consequently, the unitary generating the entanglement is given by

$$\hat{U}_{SM} = e^{-ig\hat{p}\hat{A}/\hbar}. \tag{4.21}$$

Let the system and meter state be $|\psi\rangle \otimes |\phi\rangle$, and we take the weak measurement limit so that g is much less than the width of the meter states. If the eigenstates and eigenvalues of A are defined as

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

$$\hat{A}|a_i\rangle = a_i|a_i\rangle, \tag{4.22}$$

then expressing the meter in the position basis $|x\rangle$, we find the system/meter wavefunction takes the form

$$\langle x|\hat{U}_{SM}|\psi\rangle\phi\rangle = \sum_i c_i|a_i\rangle\phi(x - ga_i). \tag{4.23}$$

Here, we used the fact that the unitary is a displacement operator in the position basis, and that the entire state is a coherent superposition of varying degrees of spatial shift, depending on the eigenvalues of A. For large values of g such that there is no longer any overlap of the states of the meter, an x measurement is sufficient to determine the eigenvalues of A. This is because the eigenvalues can be associated to the well-separated peaks of the meter, for which the coefficients c_i control the relative weights of the states.

We now consider when the system is measured projectively with the operator $\hat{\Pi}_{\tilde{\psi}} = |\tilde{\psi}\rangle\langle\tilde{\psi}|$. We focus on the outcome corresponding to this state. Operationally, this is called *postselection* – we simply wait until this event, of all the possible events, occurs, and examine what happens in that case. We then have a meter state that corresponds to

$$\begin{aligned}
\phi(x) &= \langle\tilde{\psi}|\langle x|e^{-ig\hat{p}\hat{A}}|\psi\rangle|\phi\rangle && (4.24)\\
&\approx \langle\tilde{\psi}|\langle x|(1 - ig\hat{p}\hat{A})|\psi\rangle|\phi\rangle && (4.25)\\
&= \langle\tilde{\psi}|\psi\rangle\phi(x) - ig\langle\tilde{\psi}|\hat{A}|\psi\rangle\langle x|\hat{p}|\phi\rangle && (4.26)\\
&\approx \langle\tilde{\psi}|\psi\rangle\langle x|e^{-ig\hat{p}A_w}|\phi\rangle && (4.27)\\
&= \langle\tilde{\psi}|\psi\rangle\phi(x - gA_w). && (4.28)
\end{aligned}$$

In line (4.25), we have Taylor expanded to first order in g. In line (4.27), we have factored out the quantity $\langle\tilde{\psi}|\psi\rangle$ and re-exponentiated. In doing so the weak value of the operator A is defined as

$$A_w = \frac{\langle\tilde{\psi}|\hat{A}|\psi\rangle}{\langle\tilde{\psi}|\psi\rangle}. \tag{4.29}$$

The interpretation of the result (4.28) is that the meter wavefunction is shifted by the quantity gA_w, while the amplitude is reduced by the overlap between initial and final states.

The weak value (4.29) has several interesting properties. One must "preselect" and "postselect" the initial and final states. We have already noted that it can exceed its eigenvalue range. The weak value is symmetric with respect to exchange of these two states, which is related to time-reversal symmetry of the weak value in the weak measurement limit. The weak value can be a complex number. Physically,

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

this is related to shifting the meter in the position basis (for the real weak value) and/or the momentum basis (for the imaginary weak value).

Different interpretations of the weak value exist in the literature. The most common one is that it represents the value the operator takes in a pre- and postselected ensemble, acting as a generalization of an eigenvalue to this situation. There are a variety of interesting phenomena to support this view, such as the influence of the system on any weakly coupled system acting like the system operator takes the weak value – however, this remains an interpretation of the physics.

4.4 Weak Value Amplification

In this section, we discuss an application of the weak value to an interesting problem in the science of making precise measurements. The basic idea is to note that we can turn our perspective around to take the physics of the weak value for granted and focus on the estimation of the measurement parameter g, taken as an unknown parameter, by a known weak value. Applied to the physics of the Stern–Gerlach device, Aharonov, Albert, and Vaidman point out: "Another striking aspect of this experiment becomes evident when we consider it as a device for measuring a small gradient of the magnetic field" (Aharonov et al., 1988). The basic idea is that without the postselection, the average of the meter's position is bounded between the eigenvalues of the measured operator, while with the postselection, the shift of the meter, by the weak value, can greatly exceed the eigenvalue range. Consequently, because the meter reading is much larger with the postselection, by dividing the meter reading by the weak value, a more precise determination of the parameter g can be made. However, there is a trade-off for the weak value, mentioned earlier: a larger weak value comes at the expense of a smaller postselection fraction, reducing the amount of data. As the weak value diverges, the amount of data goes to zero!

This idea was put to use in two influential experiments to make high-precision measurements of optical parameters. While the experiments were done with coherent states of light, and also can be explained with electromagnetic wave interference, we adopt the quantum formalism that inspired them, which carries over to the single-photon experiments. Hosten and Kwiat were trying to measure a tiny optical effect predicted to exist: the optical spin Hall effect, an effect analogous to the spin Hall effect in condensed matter physics, when an electric field is applied to a semiconductor, a dissipationless spin-dependent current perpendicular to the field can be generated (Hosten and Kwiat, 2008). In the optical case, a modified form of spin-orbit interaction can exist, resulting in a polarization-dependent spatial shift of the photon, where the role of the electric field is played by a change in the index of refraction. This effect occurs even in the most mundane case of light

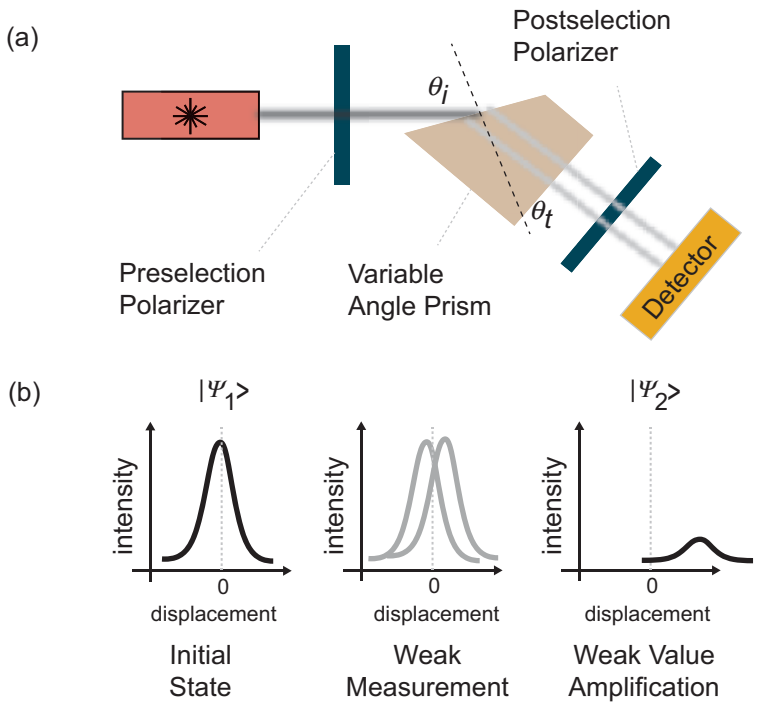

Figure 4.3 (a) In the weak value amplification experiment by Hosten and Kwiat, a light beam splits into two displaced beams when incident on a glass prism on account of the spin Hall effect for light (Hosten and Kwiat, 2008). The variable angle prism consists of two round wedge BL7 prisms joined with index matching fluid, allowing the two surfaces to be oriented such that light beams exit the optic at normal incidence without any further perturbation. Polarizers before and after the prisms allow for pre- and post-selection of the quantum state being measured, resulting in a weak value amplification of the tiny displacement between the two spin Hall shifted beams, as illustrated in (b).

passing from air to glass. In order to detect this effect, weak value amplification methods were used. As shown in Fig. 4.3, initially polarized light was incident on an air–glass interface, resulting in a slight displacement of the two polarization components. This effect is different from birefringence, which corresponds to different amounts of bending. By placing a polarizer after the interface, with the orientation nearly orthogonal to the initial polarization of the light, the deflection was greatly magnified. Using an optical detector sensitive to the location of the optical beam resulted in an accurate measurement of the size of the effect, with the precision of an angstrom.

In a second experiment by Dixon et al. (2009), a method was devised to make a very sensitive measurement of optical beam deflection, as shown in Fig. 4.4.

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

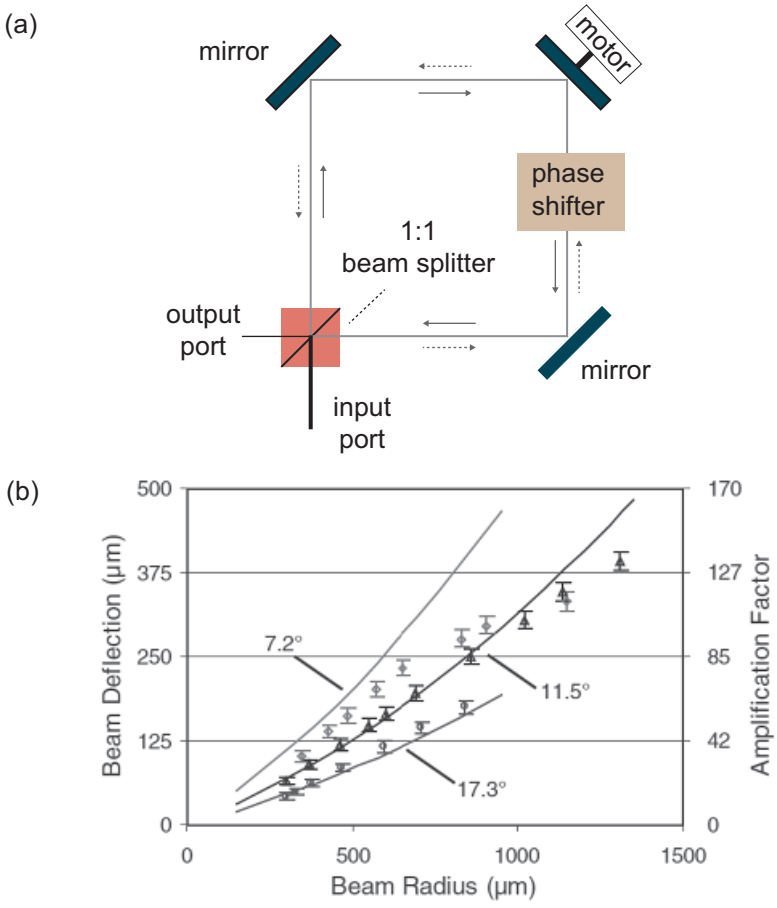

Figure 4.4 (a) Dixon, Starling, Jordan, and Howell conceived of and implemented a general scheme for weak value amplification in the optical domain using a Sagnac interferometer (Dixon et al., 2009). In this device, incident light is split at the input port with a 1:1 beam slitter, with the two components traversing a square in opposite directions. A motorized mirror can induce a deflection, changing the effective path length for the two light beams and resulting in a signal at the output port that is perfectly dark for identical paths. A tunable phase shift can be imposed using a continuously variable wave-plate such as a Soleil–Babinet compensator and other elements to postselect the quantum state being measured. (b) Data showing the measured beam deflection versus beam width for different settings of the variable wave-plate, and thus choice of quantum states. The curves were all taken for an unamplified deflection of 2.95 μm. Adapted with permission from the American Physical Society.

A Sagnac interferometer was constructed by causing two optical paths to make a cycle in a square geometry and then interfere on a beam splitter. By inserting polarization controlling optical elements, a relative phase shift ϕ was induced

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

between the clockwise and counterclockwise path. By attaching a piezoelectric actuator to one of the interferometer mirrors, a tiny angular deflection of the mirror can be introduced. The central question the experiment sought to answer was how small an angular tilt can be made and still be detected? With the use of the 50/50 beam splitter and the relative phase shift elements, an effective postselection state $|\psi\rangle = (ie^{i\phi/2}|CC\rangle + e^{-i\phi/2}|C\rangle/\sqrt{2}$ was created, where $|C\rangle, |CC\rangle$ refers to the clockwise and counterclockwise directions, respectively. The weak tilt can be modeled with a unitary operator $\hat{U} = e^{-ik\hat{x}\hat{A}}$, where $\hat{x}$ is the position operator of the transverse degree of freedom, $\hat{A} = |CC\rangle\langle CC| - |C\rangle\langle C|$ is a which-path operator, and k is the effective momentum kick the mirror gives to the optical beam, resulting in a deflection of the path. Indeed, the form of $\hat{U}$ indicates a displacement in the momentum of the meter (chosen to be Gaussian with width σ), resulting in a deflection rather than a displacement. By detecting the light coming out the other port of the beam splitter (the dark port in the limit of small ϕ), the effective postselection state is $\tilde{\psi} = (|CC\rangle + i|C\rangle)/\sqrt{2}$. This arrangement produces a postselection probability $|\langle\tilde{\psi}|\psi\rangle|^2 = \sin^2(\phi/2)$, and an entirely imaginary weak value,

$$A_w = -i\cot\frac{\phi}{2}. \tag{4.30}$$

The imaginary nature of the weak value puts the displacement back into the position space, resulting in a shift of the meter, given by

$$\langle\hat{x}\rangle = \frac{4k\sigma^2}{\phi}, \tag{4.31}$$

where we consider the case $k\sigma \ll \phi \ll 1$. We can view the preceding result as amplifying the mirror momentum kick k by a factor of $1/\phi$. By using a position-sensitive detector for the displacement of this weak optical beam, a very precise measurement of parameter k can be measured. Indeed, this first experiment obtained a precision of the mirror angle down to around 500 frad, with a precision of travel of the piezoelectric actuator to 20 fm. This angular precision is equivalent to a hair's width at the moon's distance (Steinberg, 2010).

Signal-to-Noise Ratio

We can make a more quantitative analysis of this estimation procedure by introducing the concept of signal-to-noise ratio, a very useful concept that experimentalists spend a great deal of time with. In this way, we can discuss the estimation precision without entering into a detailed discussion of the field of quantum metrology, which would take us too far away from the subject of our focus. The idea is that the measured signal, r, is proportional to the desired quantity that is to be measured, g, but is masked by some noise that could be from a variety of sources: thermal

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

fluctuations, instrumental noise, shot noise of electrons or photons, $1/f$ noise from the electronics of the detector, and so on. We will consider the simplest type of noise for this discussion, "white noise," which is additive and uncorrelated with itself from moment to moment. Thus, the measurement result can be written as

$$r = cg + \xi, \tag{4.32}$$

where c is a constant of proportionality, related to the experimental setup, and ξ is a random variable that models the noise. The value of the random variable is different from run to run and stands in for some unpredictability in the measurement. We assume the noise term has zero mean and has a finite variance, $\langle \xi^2 \rangle = \Sigma^2$, where the angle brackets here represent a statistical average over many realizations of the random process. Consequently, there is statistical uncertainty in the estimate of g, given by $g = r/c \pm \Sigma/c$. The uncertainty fixes how accurately we can estimate the parameter g. The signal-to-noise ratio (SNR) is defined as the ratio of the mean to the standard deviation of the measured signal,

$$\mathcal{R} = \frac{cg}{\Sigma}, \tag{4.33}$$

which indicates the relative contribution of the signal, versus the noise in the measurement result. The larger $\mathcal{R}$ is, the more confident we are about the value of g. A common way to increase the SNR, assuming the value of g does not change, or at least changes slowly, is to take many measurements and average the results together. Let a time series of measurement results be $\{r_1, r_2, \ldots r_N\}$. We can then find the average as

$$\bar{r} = \frac{1}{N}\sum_{i=1}^{N} r_i = cg + \frac{1}{N}\sum_{i=1}^{N} \xi_i. \tag{4.34}$$

We calculate the variance of $\bar{r}$ under the assumption of uncorrelated noise, so $\langle \xi_i \xi_j \rangle = \sigma^2 \delta_{ij}$. Thus,

$$\langle \bar{r}^2 \rangle - \langle \bar{r} \rangle^2 = \frac{1}{N^2}\sum_{i=1,j=1}^{N} \langle \xi_i \xi_j \rangle = \frac{1}{N^2}\sum_{i=1,j=1}^{N} \sigma^2 \delta_{ij} = \frac{\sigma^2}{N}. \tag{4.35}$$

Averaging the signal increases the signal-to-noise ratio by a factor of $\sqrt{N}$. The central limit theorem indicates this result is more general, because adding together N iterations of the same random variable gives an average whose distribution becomes Gaussian under a wide range of conditions, so only the variance needs to be specified.

We can now apply these considerations to the weak value estimation of the parameter g. Under the weak value amplification process, where the operator $\hat{A}$

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

is weakly measured, the signal is boosted as $c = a_{\max} \to A_w$, so the largest eigenvalue (under the best case conditions) is replaced by the weak value, $A_w \gg a_{\max}$. The standard deviation σ is interpreted as the spatial width of the meter, which is unchanged after the postselection. However, the postselection only occurs for a small fraction of all the events when the weak value is large, so the number of events is reduced as $N \to N|\langle\tilde{\psi}|\psi\rangle|^2$. Consequently, the SNR is given by

$$\mathcal{R} \to \frac{gA_w}{\sigma}\sqrt{N|\langle\tilde{\psi}|\psi\rangle|^2} = \frac{g}{\sigma}\sqrt{N}\langle\tilde{\psi}|A|\psi\rangle. \tag{4.36}$$

Here, we take $\langle\tilde{\psi}|A|\psi\rangle$ to be real for simplicity. The important point in the preceding formula is that the divergence of the weak value exactly compensates the postselection loss, so the SNR is a wash for a quantum-limited system (Starling et al., 2009), up to the difference between $a_{\max}$ and $\langle\tilde{\psi}|A|\psi\rangle$, which are typically comparable. There is, however, an important difference: about the same SNR is obtained by using a very small fraction of the data that would have been obtained with all the data in the standard measurement. We can thus interpret this method as a kind of filtering procedure. This immediately gives a number of advantages to this technique:

- In an optical context, a small portion of the laser light beam can be shaved off to measure optical properties with the same precision as if the entire beam had been measured, but then using the rest of the beam in a subsequent experiment (Starling et al., 2010).
- In an optical context, it is typically easy to use more photons by increasing the power of a laser, but most laboratory detectors can accept only a finite amount of power before they begin to burn or saturate. Consequently, weak value amplification allows experiments to use much more optical power than can be accepted by the detector and still retain the full advantage of all the photons (Starling et al., 2009).
- The non-postselected events can sometimes be recycled, reinjecting the rejected events back into the experiment to further utilize finite resources to increase precision (Dressel et al., 2013; Lyons et al., 2015; Krafczyk et al., 2021).
- Very often there are situations where other noise sources present, or the experimental setup implementing the measurements, are important (Jordan et al., 2014), where "less is more." It is sometimes experimentally advantageous to have more signal, but fewer events (Magaña-Loaiza et al., 2014), such as in the presence of systematic noise (Pang et al., 2016), or to put more time between colored noise events (Feizpour et al., 2011; Sinclair et al., 2017).

Let us now give a simple illustration of this physics with the von Neumann measurement model. Consider a preselected state $|\psi\rangle = (|+\rangle + |-\rangle)/\sqrt{2}$, and

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

a postselected state $|\tilde{\psi}\rangle = \sin(\pi/4+\phi)|+\rangle - \cos(\pi/4+\phi)|-\rangle$. Let us weakly measure the operator $\hat{\sigma}_z$. In Exercise 4.5.17, you will show $(\hat{\sigma}_z)_w = \cot\phi$, with postselection probability $\sin^2\phi$. In this case, an initial Gaussian meter with standard deviation σ is shifted in position by the amount g if the system is prepared in an eigenvalue $\hat{\sigma}_z = 1$, but by the amount $g(\sigma_z)_w$ if it is pre- and postselected. Consequently, a measurement of the meter's position will give a much larger reading, improving the signal of an unknown g measurement. Applying our result (4.36) indicates that the SNR of this measurement is given by

$$\mathcal{R} = \frac{g(\hat{\sigma}_z)_w\sqrt{N|\langle\tilde{\psi}|\psi\rangle|^2}}{\sigma} = \frac{g\cos\phi\sqrt{N}}{\sigma}. \tag{4.37}$$

Consequently, as $\phi \to 0$ (while keeping g even smaller), the SNR recovers the SNR with all the data from the eigenstate measurement. The observant student may have a bright idea at this point: perhaps we can do *even better* if we also use the non-postselected data to further improve our precision. In Exercise 4.5.18 you will show that in fact this is not the case, and that any well-designed weak-value amplification experiment has negligible information in the rejected events, effectively concentrating all information into the rare postselected events.

The previous section only scratches the surface about weak value amplification. There are now hundreds of experiments applying variants of this method to all manner of estimation problems. An early review can be found in Dressel et al. (2014). We mention here a few highlights of exciting recent work in this area of research. The group of Aephraim Steinberg demonstrated weak-value amplification of the nonlinear effect of a single photon (Hallaji et al., 2017), where one photon effectively acted like many photons. Weak value amplification has also been demonstrated in integrated optical platforms in the group of Jaime Cardenas (Song et al., 2021). In that work, miniaturized multimode interferometers were designed on-chip to sensitively measure phase shifts induced by a heater as well as measure a change in optical frequency (with the help of a ring resonator) down to a precision of 2 kHz. Finally, recent work has considered the advantages of postselected metrology in quantum physics more generally (Lupu-Gladstein et al., 2022), beyond the weak value paradigm. The parameter estimation protocol has come to be the primary application of weak values.

4.5 Generalized Eigenvalues for Any Measurement Type

In the treatment of generalized measurements of Chapter 3, we learned about new types of measurement, how to calculate their probabilities, and the resulting change of the system state. However, we lost an important notion of what we were measuring. That is, the link to the system's observable property became

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

obscured. In the case of weak measurement, as discussed in the previous section, there is a clear interpretation of weakly measuring a system observable. We now discuss how to generalize this notion to arbitrary types of measurement by introducing the concept of *generalized eigenvalues* of an observable, labeled $\{\kappa_j\}$, where $j = 1, \ldots, M$, associated with the possible outcomes of a generalized measurement. The assignment of these generalized eigenvalues depends on the kind of generalized measurement to be made. These thus depend on the measurement context, so these values are sometimes called "contextual values" (Dressel et al., 2010).

In order to assign these values, given a set of measurement operators $\{\hat{\Omega}_j\}$, we can express the expectation value of the operator $\hat{\mathcal{O}}$ (2.19) as

$$\langle \hat{\mathcal{O}} \rangle = \sum_{i=1}^{N} \lambda_i P_i = \sum_{j=1}^{M} \kappa_j \mathcal{P}_j, \tag{4.38}$$

where we replace $\lambda_i \to \kappa_j$ and the probability of the projective measurement result $P_i = \mathrm{Tr}[\hat{\rho}\hat{\Pi}_j]$ with the probability of the generalized measurement, $\mathcal{P}_j = \mathrm{Tr}[\hat{\rho}\hat{E}_j]$. Note that $M \neq N$ in general. This assignment should be true for all possible states $\hat{\rho}$, so we have then an operator equation,

$$\hat{\mathcal{O}} = \sum_{i=1}^{N} \lambda_i \hat{\Pi}_i = \sum_{j=1}^{M} \kappa_j \hat{E}_j. \tag{4.39}$$

It is clear from this expression that the generalized eigenvalues must be real numbers. Once a valid assignment of the generalized eigenvalues is made, we note that the expectation of any moment of the operator $\hat{\mathcal{O}}$ (2.20) can be found as

$$\langle \hat{\mathcal{O}}^n \rangle = \sum_{i=1}^{N} \lambda_i^n P_i = \sum_{j_1, j_2, \ldots, j_n = 1}^{M} \kappa_{j_1} \kappa_{j_2} \ldots \kappa_{j_n} \mathcal{P}_{j_1, j_2, \ldots, j_n}, \tag{4.40}$$

where we have defined

$$\mathcal{P}_{j_1, j_2, \ldots, j_n} = \mathrm{Tr}[\hat{E}_{j_1} \hat{E}_{j_2} \ldots \hat{E}_{j_n} \hat{\rho}], \tag{4.41}$$

which may be interpreted as the probability of finding results $j_1, j_2, \ldots, j_n$ in a sequence of n generalized measurements. We have made the simplifying assumption that the observable $\hat{\mathcal{O}}$ and all operators $\hat{E}_j$ commute.

Before a general discussion of how to satisfy our main equation of this section, Eq. (4.39), we first give a simple example, based on Exercise 3.7.2. This problem concerns coupling a two-level system to a two-level meter, realized as the polarization of a single photon. If you solved the problem correctly, you found that the measurement operators satisfy $\hat{E}_{1,2} = (1/2)\mathbf{1} \pm g\hat{\sigma}_z$, where we label 1 and 2 as

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

when the meter photon registers the results H and V, respectively. In this case, we have the operator equation (4.39) given by

$$\hat{\mathcal{O}} = \sum_{i=1}^{N} \lambda_i \hat{\Pi}_i = \frac{\kappa_1 + \kappa_2}{2}\mathbf{1} + g\frac{\kappa_1 - \kappa_2}{2}\hat{\sigma}_z. \tag{4.42}$$

From this expression, it is clear that we can reconstruct operators in the span of $\{\mathbf{1}, \hat{\sigma}_z\}$. Let us choose the operator $\hat{\mathcal{O}} = \hat{\sigma}_z = |H\rangle\langle H| - |V\rangle\langle V|$, to be definite, with eigenvalues $\lambda_{1,2} = \pm 1$ in the $H - V$ basis. The solution to Equation (4.42) is then uniquely given by

$$\kappa_1 = \frac{1}{g}, \quad \kappa_2 = -\frac{1}{g}. \tag{4.43}$$

It is interesting to note that as the measurement strength g limits to zero, $\kappa_{1,2}$ diverge to plus or minus infinity. For any value of $g < 1$, the generalized eigenvalues exceed the eigenvalue bounds ± 1, reminiscent of the weak value from the previous section. Nevertheless, despite this behavior, we can confirm that the desired behavior for the expectation value of $\hat{\mathcal{O}}$ is satisfied for a general pure state $|\psi\rangle = a|H\rangle + b|V\rangle$:

$$\langle\hat{\mathcal{O}}\rangle = \langle\psi|\hat{\sigma}_z|\psi\rangle = |a|^2 - |b|^2 = \sum_{j=1,2} \kappa_j \langle\psi|\hat{E}_j|\psi\rangle = \sum_{j=1,2} \kappa_j \mathcal{P}_j. \tag{4.44}$$

Indeed, we find that either way, we get the correct expression. To have a deeper understanding, we write the probabilities of the generalized measurement as $\mathcal{P}_j = 1/2 \pm g\langle\psi|\hat{\sigma}_z|\psi\rangle = 1/2 \pm g(|a|^2 - |b|^2)$. The fact that these probabilities have only a very weak dependence on the system properties $|a|^2 - |b|^2$ when g is small, implies that in order to extract the system dependence, we must effectively amplify it with the generalized eigenvalues. This is why they must diverge as g becomes small, and be opposite signed: The constant term must be canceled off, and the asymmetric terms must be amplified.

Recalling the defining equation for the generalized eigenvalues, Eq. (4.39), we see that, if $N = M$, the solution is unique, while if $N > M$ the system is overdetermined, while, if $N < M$, then the system is underdetermined. Physically, this dictates our ability to construct averages of system operators from the type of measurements and information at hand. For example, if we carry out a two-outcome generalized measurement on a position of a particle, then it is not too surprising that with that data, it is not possible to accurately calculate the particle's expected position, $\langle x\rangle$, in general. On the other hand, if you have access to an infinite number of possible outcomes weakly measuring a two-state system, there can be many ways to extract the desired information about that system. Nevertheless, in both the over- and underdetermined cases, there is a principled way to make a choice of the

assigned values, either to make the best available approximation in the overdetermined case or to minimize the norm of the solution in the underdetermined case. To illustrate why the later condition is desirable here, consider the following example.

Let us return to the von Neumann measurement model (3.5), where the measurement operators, indexed by the position of the meter registration, is given by

$$\hat{\Omega}(x) = \phi(x - g\hat{\sigma}_z), \tag{4.45}$$

where $\phi(x)$ is the meter's position wavefunction. It is shifted either left or right by g, depending on the state of the system. Let us suppose we are interested to find the expectation value of $\hat{\sigma}_z$ of the two-level system in whatever state it is in. The operator $\hat{E}(x)$ is given by $\hat{E}(x) = P_m(x - g\hat{\sigma}_z)$, where $P_m(x)$ is the meter's probability density. Let us consider two possible choices (of the many) to assign to $\kappa(x)$ to estimate $\hat{\sigma}_z$:

$$\kappa_A(x) = \frac{x}{g}, \quad \kappa_B(x) = \frac{P_m(x-g) - P_m(x+g)}{a-b}, \tag{4.46}$$

where a, b are constants, given by $a = \int dx P_m(x)^2, b = \int dx P_m(x-g)P_m(x+g)$. In Exercise 4.5.13, you will show that both choices satisfy the condition (4.39).

In terms of pros and cons of which one of these functions to use, notice that while the choice $\kappa_A(x)$ is simplest, and independent of even the type of meter wavefunction, its value continues to increase linearly as x grows, while the choice $\kappa_B(x)$ rapidly decays to zero once x extends beyond the typical width of the meter wavefunction. This difference plays an important role in experiments. When collecting data, many rounds of the experiment are done from the same initial conditions. In terms of the operational procedure, the experimentalist would register the events x_k where k indexes the round number, and the x axis would usually be binned to create statistics. The average would then be constructed by weighting the generalized eigenvalue assigned to that bin by the frequency of events in a given bin, and then summing. Once this procedure is understood, the advantage of choice κ_B over κ_A becomes clear: given finite statistics, the convergence of the weighted average will be much faster. Therefore, we can make an optimal choice for $\kappa(x)$, that is, independent of the state of the system, by choosing the assignment that minimizes the norm of κ_j, that is, $\sum_j \kappa_j^2$. This choice can be found by assuming that the operator $\hat{\mathcal{O}}$ and the operator $\hat{E}_j$ are all diagonal in the same basis. We can then express the operator equation (4.39) as a matrix equation,

$$\boldsymbol{\lambda} = \overline{\overline{\mathbf{F}}}\boldsymbol{\kappa}, \qquad F_{ij} = \mathrm{Tr}[\hat{\Pi}_i \hat{E}_j], \tag{4.47}$$

where the vectors $\boldsymbol{\kappa}$ and $\boldsymbol{\lambda}$ are column vectors of length M and N, respectively, with elements κ_j and λ_i. $\overline{\overline{\mathbf{F}}}$ is a real $N \times M$ matrix with the elements defined earlier.

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

We recognize the constitutive equation (4.47) as a linear system of equations for $\boldsymbol{\kappa}$. Solving this linear system in the underdetermined or overdetermined cases is a well-studied problem in linear algebra. If $M = N$ we can usually simply find the inverse of the matrix F^{-1} to solve the problem, $\boldsymbol{\kappa} = \overline{\overline{\mathbf{F}}}^{-1}\boldsymbol{\lambda}$. In the overdetermined case, a solution is generally impossible; however, we can find the "least-squares" solution – that is, the solution that minimizes the error, $||\boldsymbol{\lambda} - \overline{\overline{\mathbf{F}}}\boldsymbol{\kappa}||$, where $||\ldots||$ is the Euclidean norm. This solution is given by

$$\boldsymbol{\kappa}_{ls} = (\overline{\overline{\mathbf{F}^T}}\,\overline{\overline{\mathbf{F}}})^{-1}\,\overline{\overline{\mathbf{F}^T}}\boldsymbol{\lambda}, \tag{4.48}$$

also known as the pseudo-inverse of a left-invertible matrix. A similar pseudo-inverse can be constructed in the undetermined case. We quote the most general type of pseudo-inverse, also known as the Moore–Penrose pseudo-inverse, which uses the singular value decomposition. Any real matrix has a singular value decomposition, $\overline{\overline{\mathbf{F}}} = \overline{\overline{\mathbf{U}}}\,\overline{\overline{\Sigma}}\,\overline{\overline{\mathbf{V}^T}}$, where $\overline{\overline{\mathbf{U}}}$ is an $N \times N$ orthogonal matrix, $\overline{\overline{\Sigma}}$ is an $N \times M$ real matrix with nonnegative elements on its diagonal, and $\overline{\overline{\mathbf{V}}}$ is an $M \times M$ orthogonal matrix. The diagonal elements of the rectangular matrix $\overline{\overline{\Sigma}}$ are called the singular values. With this construction, the Moore–Penrose pseudo-inverse is given by

$$\overline{\overline{\mathbf{F}^+}} = \overline{\overline{\mathbf{V}}}\,\overline{\overline{\Sigma^+}}\,\overline{\overline{\mathbf{U}^T}}, \tag{4.49}$$

where the rectangular $M \times N$ matrix $\overline{\overline{\Sigma^+}}$ is defined by inverting all nonzero elements of $\overline{\overline{\Sigma^T}}$. The benefit of the solution $\boldsymbol{\kappa}_{mp} = \overline{\overline{\mathbf{F}^+}}\boldsymbol{\lambda}$ is that it satisfies our earlier desire of having the minimum norm. All other solutions to Eq. (4.47) can be written as $\boldsymbol{\kappa} = \boldsymbol{\kappa}_{mp} + \mathbf{n}$, where $\mathbf{n}$ is another vector in the null-space of F. This additional vector can only lengthen the norm of the solution. Further discussion of these solutions and examples can be found in Dressel et al. (2010) and Dressel and Jordan (2012a).

Conditioned Averages and the Recovery of the Weak Value

As an illustration of the application of the generalized eigenvalues, we now discuss how to recover the weak value formula in the weak measurement limit. We begin by considering a generalized measurement, followed by a projective measurement. Consider an initial state $|\psi\rangle$. We treat the pure state for simplicity, but our treatment is straightforward to generalize to the mixed state case. We consider the measurement of the system operator $\hat{\mathcal{O}}$. The generalized eigenvalues of the first measurement are taken to be $\{\kappa_j\}$, along with measurement operators $\{\hat{\Omega}_j\}$, and the final projection is described with projectors $\{\hat{\Pi}_k\}$. The effect of the two

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

measurements together is then $\hat{M}_{jk} = \hat{\Pi}_k \hat{\Omega}_j$. The probability of finding both results j and k in a measurement sequence is given by

$$P(j,k) = \langle \psi | \hat{\Omega}_j^\dagger \hat{\Pi}_k \hat{\Omega}_j | \psi \rangle. \tag{4.50}$$

Recall that the average of operator $\hat{\mathcal{O}}$ is given in Eq. (4.38) as a sum of the generalized eigenvalues weighted by the probability of the result. It is then natural in the measurement sequence to define the *conditioned average* of operator $\hat{\mathcal{O}}$, when postselecting on a particular final state, $\hat{\Pi}_f = |\tilde{\psi}\rangle\langle\tilde{\psi}|$, as follows,

$${}_{\tilde{\psi}}\langle \hat{\mathcal{O}} \rangle_{\psi} = \sum_j \kappa_j P(j|k=f). \tag{4.51}$$

Here, the probability $P(j|k=f)$ is the conditional probability of result j, given a final postselection on $k=f$. We can calculate this probability using Bayes' rule (see Appendix A):

$$P(j|k=f) = \frac{P(j,k=f)}{P_f}, \qquad P(f) = \sum_j P(j,k=f). \tag{4.52}$$

Before proceeding further, we notice a natural explanation for why it is now possible for the conditional average to exceed the eigenvalue bounds: the generalized eigenvalues can exceed the normal eigenvalue bounds, so if we weight them with a different distribution, it is possible for the result to also exceed the eigenvalue bounds. This is related to the effect that if we take the variance of the generalized eigenvalues, it is not given by the variance of the operator; see (4.40).

It is now fairly easy to see how the weak value emerges in the weak measurement limit. Recall that if we introduce a measurement strength g, then the form (4.4) holds in the weak measurement limit. We consider the simplest case of a QND measurement, where in the decomposition (3.19), we choose $U_j = \mathbf{1}$, or at least a U_j that limits to the identity to higher order in g (note that the effect of U_f can be included by unitarily rotating the postselection state). We then have

$$\hat{\Omega}_j = \sqrt{\hat{E}_j} = \sqrt{p_j \mathbf{1} + g\Gamma_j \hat{\mathcal{O}}} \approx \sqrt{p_j}\mathbf{1} + \frac{g\Gamma_j}{2\sqrt{p_j}}\hat{\mathcal{O}}, \tag{4.53}$$

where we exclude $p_j = 0$ for any j, for simplicity, and we drop higher-order terms in g. Let us then consider the denominator in Eqs. (4.51,4.52):

$$P_f = \sum_j \langle \psi | (\sqrt{p_j}\hat{\Pi}_f \sqrt{p_j} | \psi \rangle + O(g) \approx |\langle \tilde{\psi} | \psi \rangle|^2. \tag{4.54}$$

We see that so long as the overlap between the initial and final states is not zero, the leading-order term is independent of g. Let us now turn to the numerator N of Eq. (4.51):

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

$$N = \sum_j \kappa_j \left| \langle \tilde{\psi} | \sqrt{p_j} \mathbf{1} + \frac{g\Gamma_j}{2\sqrt{p_j}} \hat{\mathcal{O}} | \psi \rangle \right|^2 , \tag{4.55}$$

$$= \sum_j \kappa_j \left| \langle \tilde{\psi} | \psi \rangle \sqrt{p_j} + \frac{g\Gamma_j}{2\sqrt{p_j}} \langle \tilde{\psi} | \hat{\mathcal{O}} | \psi \rangle \right|^2 , \tag{4.56}$$

$$\approx \sum_j \kappa_j \left[p_j |\langle \tilde{\psi} | \psi \rangle|^2 + \frac{g\Gamma_j}{2} (\langle \tilde{\psi} | \psi \rangle^* \langle \tilde{\psi} | \hat{\mathcal{O}} | \psi \rangle + \langle \tilde{\psi} | \psi \rangle \langle \tilde{\psi} | \hat{\mathcal{O}} | \psi \rangle^*) \right], \tag{4.57}$$

where we drop higher-order terms in g. To further simplify this result, note that since $\sum_j \kappa_j \hat{E}_j = \hat{\mathcal{O}}$ by construction, and since $\hat{E}_j \approx p_j \mathbf{1} + g\Gamma_j \hat{\mathcal{O}}$ in the small g limit, we must have the relations

$$\sum_j \kappa_j p_j = 0, \qquad g \sum_j \kappa_j \Gamma_j = 1. \tag{4.58}$$

These sum rules then simplify the numerator (4.57) by eliminating the first term, and eliminating the $\Gamma_j \kappa_j$ factors in the second term. We then have a final result,

$$_{\tilde{\psi}}\langle \hat{\mathcal{O}} \rangle_{\psi} = \frac{\langle \tilde{\psi} | \psi \rangle^* \langle \tilde{\psi} | \hat{\mathcal{O}} | \psi \rangle + \langle \tilde{\psi} | \psi \rangle \langle \tilde{\psi} | \hat{\mathcal{O}} | \psi \rangle^*}{2 |\langle \tilde{\psi} | \psi \rangle|^2} = \mathrm{Re} \frac{\langle \tilde{\psi} | \hat{\mathcal{O}} | \psi \rangle}{\langle \tilde{\psi} | \psi \rangle}. \tag{4.59}$$

It is important to this derivation that the choice of generalized eigenvalues isolates the desired operator to be measured, and amplifies it, in order to get a finite result depending on the system operator as g limits to zero. We point out that very few assumptions have been made about the type of measurement to obtain the weak value in the small g limit, indicating a robustness, if not universality, of this result. A final point about our result is that it is the real part of the weak value that appears. In hindsight, this comes from our starting point (4.51) which stresses an operational interpretation of dealing with data from an experiment. Such data is always real, and since our generalized eigenvalues are also defined to be real, it is guaranteed to obtain a real number for any type of average of such numbers. The imaginary part of the weak value also has a physical interpretation as discussed earlier in this chapter. See Steinberg (1995), Aharonov and Botero (2005), Jozsa (2007), and Dressel and Jordan (2012b) for further reading.

Exercises

Exercise 4.5.1 Consider the von Neumann measurement model in the weak measurement limit, where the parameter g is very small. What system observable is being weakly measured? What function plays the role of Γ_j in Eq. (4.4)?

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

Exercise 4.5.2 Demonstrate that the distribution (4.2) remains normalized to $\mathcal{O}(d)$.

Exercise 4.5.3 A common model for "pure dephasing" is to add a random phase θ to a coherent superposition, and to then average the phase over a distribution we can take to be Gaussian. For a qubit, work out what the quantum channel for this process is. That is, find the appropriate operators $\hat{\Omega}_j$.

Exercise 4.5.4 For Exercise 3.7.2, study the weak measurement limit. What system operator is being weakly measured by the meter?

Exercise 4.5.5 For the von Neumann measurement model, calculate the Bhattacharyya coefficient for the measurement probabilities. Using a Gaussian meter wavefunction, check to see if this is a quantum limited detector by investigating the ensemble averaged decoherence. What happens if the meter wavefunction is a complex function?

Exercise 4.5.6 For Exercise 3.7.2, calculate the Bhattacharyya coefficient for the measurement probabilities. Check to see if this is a quantum limited detector by investigating the ensemble averaged decoherence.

Exercise 4.5.7 Consider a particle of mass m in a harmonic trap of frequency ω. Suppose you want to weakly measure its position with an impulsive interaction using a free particle meter. What interaction Hamiltonian should be designed? What about a weak momentum measurement? A weak measurement of the mechanical energy?

Exercise 4.5.8 Consider a metastable potential approximated as a cubic function. Use the WKB approximation to calculate the tunneling rate out of the well. Verify that the rate depends exponentially on the energy of the particle inside the well.

Exercise 4.5.9 Verify that another way of expressing the result (4.14) for the no-tunneling conditional evolution of the phase qubit is

$$\frac{\rho_{ee}(t)}{\rho_{gg}(t)} = e^{-\Gamma t}\frac{\rho_{ee}(0)}{\rho_{gg}(0)}, \qquad \frac{\rho_{eg}(t)}{\sqrt{\rho_{gg}(t)\rho_{ee}(t)}} = e^{i\omega t}\frac{\rho_{eg}(0)}{\sqrt{\rho_{gg}(0)\rho_{ee}(0)}}, \tag{4.60}$$

while maintaining the norm 1 and Hermitian nature of the density matrix.

Exercise 4.5.10 Verify that in Eq. (4.14), if the initial density matrix is pure, that is, $\hat{\rho} = |\psi\rangle\langle\psi|$ for some state $|\psi\rangle$, that the state remains pure during the continuous collapse.

Exercise 4.5.11 For the previous problem, write the initial pure state as $|\psi\rangle = \cos\theta_i|g\rangle + \sin\theta_i|e\rangle$. Find the continuous collapse rule for the null measurement case after time t. That is, find $\theta(t)$ that begins at θ_i and ends at $\theta = 0$ in the long time limit.

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press

Exercise 4.5.12 Following the discussion in Section 4.4, what are the generalized eigenvalues for the projection operators for H or V polarization, either $\hat{\Pi}_H$ or $\hat{\Pi}_V$?

Exercise 4.5.13 Verify that both choices, κ_A and κ_B in Eq. (4.46), obey the central equation (4.39), suitably generalized to continuous variables. Hint: Consider one matrix element at a time if you need to.

Exercise 4.5.14 Use the properties of the complete set of projections operators $\hat{\Pi}_i$ to show the matrix equation (4.47) is true.

Exercise 4.5.15 Suppose the two-level system is measured with a three-outcome measurement. The observable being measured is $\hat{\sigma}_z$. The measurement operators, expressed in the $\hat{\sigma}_z$ eigenbasis, are $\hat{E}_1 = \text{diag}(1/3, 1/2)$, $\hat{E}_2 = \text{diag}(1/3, 1/5)$, $\hat{E}_3 = \text{diag}(1/3, 3/10)$. Find the generalized eigenvalues of $\hat{\sigma}_z$ for this measurement that minimizes the norm of the solution.

Exercise 4.5.16 A three-level system is measured with a two-outcome measurement, described by the operators $\hat{E}_1 = \text{diag}(1/2 + g, 1/2 - g, 1/2 + w)$, $\hat{E}_2 = \text{diag}(1/2 - g, 1/2 + g, 1/2 - w)$, where $-1/2 < w, g < 1/2$. Find the generalized eigenvalue needed to measure a general observable in this basis of the form $\hat{\mathcal{O}} = \text{diag}(a, b, c)$, where a, b, c are real numbers, that has the minimum error.

Exercise 4.5.17 Show for the von Neumann model that for the choices in Section 4.5.1, the postselection probability and weak value are given by

$$|\langle\tilde{\psi}|\psi\rangle|^2 = \sin^2\phi, \qquad (\sigma_z)_w = \cot\phi. \tag{4.61}$$

Exercise 4.5.18 See if there is any information in the rejected events from a weak value amplification model of Section 4.5.1 by finding the SNR using the deflection signal from the other postselection state, $|\psi_\perp\rangle = \cos(\pi/4 + \phi)|+\rangle + \sin(\pi/4 + \phi)|-\rangle$ in the limit as $\phi \to 0$. This experiment was carried out in Viza et al. (2015).

https://doi.org/10.1017/9781009103909.005 Published online by Cambridge University Press