3

Generalized Measurement

3.1 Measuring the Polarization of a Single Photon

We now begin our discussion of generalized measurements, where the outcome of the measurement apparatus is not perfectly correlated with the quantum state of the system being measured. For this purpose, it is instructive to introduce another motivating experiment. We turn to the subject of optics, which has been a workhorse for the field of quantum science. It is fitting to discuss one of the oldest known optical effects, first noted to occur in calcite crystals. "Iceland Spar," naturally occurring calcite ($CaCO_3$), was described by Erasmus Bartholinus at the University of Copenhagen in 1669. Although legend tells of Vikings using the "sunstone" (thought to be Iceland Spar) to navigate the seas by the partial polarization of the sky, we have no definitive evidence for this much earlier discovery of optical polarization.

Bartholinus performed experiments on the curious properties of this crystal. In gazing through it, a double image would appear. We now understand that the crystal splits the light entering it into two distinct images, corresponding to two different polarizations of the light. While we can describe the physics of optical polarization perfectly well with electromagnetic wave theory, we transition to the quantum level by considering a single photon entering such a crystal. It is a curious fact of nature that the polarization properties of a single photon correspond nearly exactly to those of a classical beam of light from the sun or a flashlight, the main difference being that while the intensity of a light beam passing through a polarizer is continuous in its intensity as the polarizer is turned, a single photon, when detected afterwards, has only two choices: it is found to be a single photon in one of the two possible polarizations. Consequently, the physics of a single photon split by a calcite crystal into one of two polarizations is very similar to the physics of the Stern–Gerlach device deflecting a silver atom up or down, depending on the quantum state of its valence electron's spin. Indeed the quantum mechanics of a

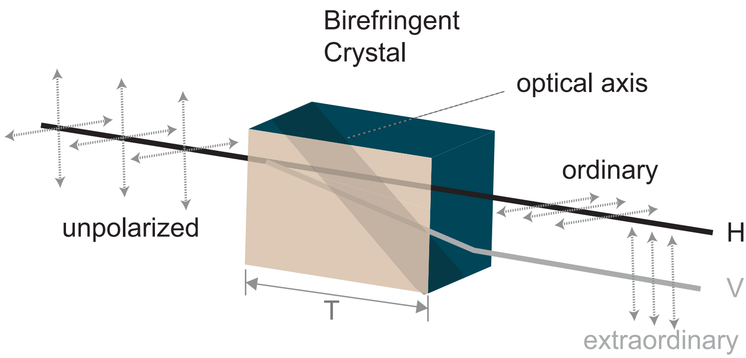


Figure 3.1 In a birefringent crystal, such as calcite, incoming radiation is transmitted unaltered if the polarization is parallel to the plane containing the optical axis and is called the ordinary. The orthogonal polarization is deflected and is called the extraordinary ray.

spin-1/2 object can be mapped on the quantum mechanics of the polarization of a single photon.

The basic physics of this effect comes from the crystalline structure of the material defining an optical axis, coming from the anisotropy of the material (see Fig. 3.1). Light is typically split into two rays: an ordinary ray, and an extraordinary ray. These correspond to two different indices of refraction associated with the two polarizations of light (n_o, n_e) resulting in different amounts of deflection of the optical ray. For a crystal of a given thickness, the optical path difference d between the two rays is the optical birefringence (defined as the difference between the two indices of refraction) times the thickness, T. That is, $d = T(n_e - n_0)$.

Just as for a spin-1/2 system we can define the quantum state of the single photon we have in mind by $|H\rangle$ and $|V\rangle$, corresponding to the horizontal and vertical polarization of light, relative to a reference, such as the optical axis of the crystal. Thus, any polarization state may be written as

$$|\psi\rangle = a|H\rangle + b|V\rangle, \tag{3.1}$$

where a, b are complex coefficients. If the coefficients are real, this state corresponds to linear polarization, but if one of the coefficients is imaginary, the state corresponds to circular or elliptical polarization.

We now consider several experiments where the thickness of the crystal is becoming thinner and thinner. At some point, we notice that the beams each have some finite transverse width that comes into play (see Fig. 3.2). This width cannot be made arbitrarily thin for a beam of light, or else we run afoul of the uncertainty

https://doi.org/10.1017/9781009103909.004 Published online by Cambridge University Press

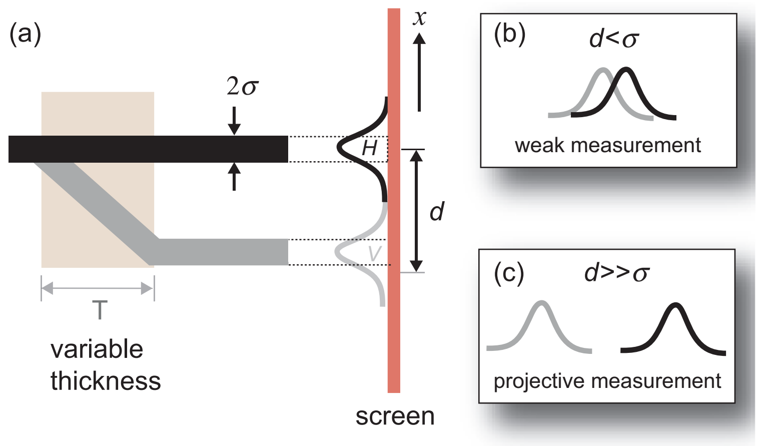


Figure 3.2 (a) After traversing a birefringent crystal, the ordinary and extraordinary rays are separated by a distance d relative to their beam waist σ, and the crystal thus performs a measurement of the polarization of the incoming light. (b) In the limit $d < \sigma$, the histograms corresponding to the different polarization states overlap, and the measurement is said to be "weak" since a single measurement outcome cannot unambiguously distinguish these states. (c) Conversely, when $d \gg \sigma$, the histograms are well separated and the measurement is said to be "projective."

principle. We consider a transverse wavefunction of the photon to be $\phi(x)$. It is usual to consider a Gaussian profile, as this is commonly found for optical states,

$$\phi(x) = \left(\frac{1}{2\pi\sigma^2}\right)^{1/4} \exp\left(-\frac{x^2}{4\sigma^2}\right). \tag{3.2}$$

Here σ is the standard deviation of the position of the photon detected in the beam, with an average of $x = 0$. It quantifies the expected scatter we expect on the final detection screen for an ensemble of measurements. Defining the origin of the x axis as the center of the ordinary beam, we can use the superposition principle to predict the state of the photon before the detection screen. Other details of the physics, such as the frequency of the photon, its temporal mode form, its speed, and so on, are suppressed for simplicity. Only the transverse degree of freedom and polarization are important to discuss here. We see that starting in the initial state

$$|\Psi\rangle = |\psi\rangle \otimes |\phi\rangle, \tag{3.3}$$

where $|\phi\rangle$ is the ket corresponding to the transverse (orbital) degree of freedom, we can find the final state before the screen by shifting the transverse state $|\phi\rangle$ to the left by an amount d. Let us then distinguish the transverse state corresponding

https://doi.org/10.1017/9781009103909.004 Published online by Cambridge University Press

to the horizontal polarization, $|\phi_H\rangle$, from the transverse state corresponding to the vertical polarization, $|\phi_V\rangle$, the latter being shifted to the left by an amount d. That is, $\langle x|\phi_H\rangle = \phi(x)$, while $\langle x|\phi_V\rangle = \phi(x+d)$. The superposition principle then dictates that the final state of the photon before the screen is the coherent superposition of the two possibilities,

$$|\Psi'\rangle = a|H\rangle|\phi_H\rangle + b|V\rangle|\phi_V\rangle. \tag{3.4}$$

Notice that if the separation d is much larger than the width of the profile σ, then the two distributions have no region of overlap, resulting in a nearly ideal measurement of optical polarization. That is, by simply detecting where on the screen the photon arrived, we can state confidently what the polarization of the photon is. They are perfectly correlated, so a position measurement stands in for a polarization measurement, made, for example, with a piece of Polaroid. However, for $d \leq \sigma$, the two profiles overlap, so a position measurement becomes imperfectly correlated with polarization. For example, if a photon lands to the far right of the screen, then we are pretty confident it is an H-polarized photon. However, suppose the photon lands halfway between the peaks of the two distributions. Then from position alone, we have no way of inferring whether the photon is H or V polarized.

3.2 Measuring Polarization with Position

We will now use the results of the previous chapter to work out the details of the example system introduced earlier. We wish to predict where the single photon will land on the screen. Furthermore, suppose we poke a small hole in the screen and examine only those photons that make it through the hole. What is their polarization, and does it depend on where we poke the hole in the screen?

The photon will only make it though the hole at position x if we are certain it is there. Consequently, this corresponds to a position measurement. From Section 2.4, we should then describe this situation with a position projection operator, $\hat{\Pi}_x = |x\rangle\langle x|$. Since the measurement is of position only, the action on the polarization degree of freedom is only the identity operator. The probability density in this situation is the continuous variable generalization of Eq. (2.11),

$$p(x) = \langle\Psi|\hat{\Pi}_x|\Psi\rangle, \tag{3.5}$$

and the actual probability of traversing the slit is related to this width. Taking a slit size Δx, the probability to traverse the slit is given by $P \approx p(x)\Delta x$, for suitably small Δx. Noting that $\langle x|\phi_j\rangle = \phi_j(x)$ for $j = H, V$, we can expand this expression to

$$p(x) = [\langle H|a^*\phi_H^*(x) + \langle V|b^*\phi_V^*(x)][a\phi_H(x)|H\rangle + b\phi_V(x)|V\rangle], \tag{3.6}$$
$$= |a|^2|\phi_H(x)|^2 + |b|^2|\phi_V(x)|^2. \tag{3.7}$$

https://doi.org/10.1017/9781009103909.004 Published online by Cambridge University Press

This last result has a simple interpretation: we can view the probability of passing the slit as being a sum of independent events where, given the photon is in either H of V, the probability of passing the slit is either $|\phi_H(x)|^2\Delta x$ or $|\phi_V(x)|^2\Delta x$. The net probability is then simply the weighted probability of passing the slit, where the weighting coefficients are the probability for the photon to be in either polarization H or V, that is, either $|a|^2$ or $|b|^2$. We will return to this interpretation later on.

3.3 Polarization State Update

We are now in the position to ask our follow-up question: given that the photon made it through the slit centered at position x, what is its polarization?

We can find this because, according to the discussion around Eq. (2.14), the postmeasurement state is simply $|x\rangle$ for the meter. However, what about the polarization state? This can be found from

$$\begin{aligned} |\Psi''\rangle &= \hat{\Pi}_x|\Psi'\rangle, && (3.8)\\ &= (a\phi_H(x)|H\rangle + b\phi_V(x)|V\rangle)|x\rangle. && (3.9) \end{aligned}$$

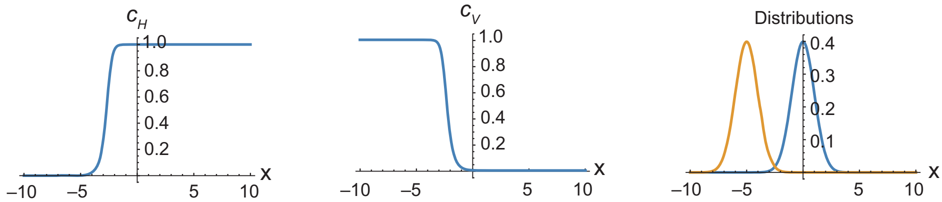
The state of the polarization is not normalized, since we have selected on this specific event. Consequently, it must be renormalized to get a physical state, accounting for the fact the photon made it through. Substituting the original and shifted functions $\phi_{H,V}(x)$, we find

$$|\psi''\rangle = \frac{a\phi(x)|H\rangle + b\phi(x+d)|V\rangle}{\sqrt{|a|^2|\phi_H(x)|^2 + |b|^2|\phi_V(x)|^2}}. \tag{3.10}$$

As you will show in Exercise 3.7.1, the postmeasurement state of the photon can be written as $|\psi''\rangle = c_H(x)|H\rangle + c_V(x)|V\rangle$, with new coefficients that are functions of x and the other parameters of the system. This effect is identical to a rotation of the polarization of the photon, conditioned on a photon traversing the slit located at position x.

It is of interest to plot these results to get a better intuition of the result of measurement in this generalized situation. We see in Figs. 3.3 and 3.4 the coefficients of the H and V polarization state in panels (a) and (b) respectively, plotted versus the position of the slit x. Panel (c) shows the two different distributions made by vertically polarized photons (the one shifted left) corresponding to the extraordinary ray and horizontal photons (the one centered at $x = 0$), corresponding to the ordinary ray. The difference between the figures corresponds to choosing the parameter $d = 5, \sigma = 1$ for Fig. 3.3, whereas we choose $d = 0.5, \sigma = 1$ for Fig. 3.4. In both cases, position is measured in units of σ.

https://doi.org/10.1017/9781009103909.004 Published online by Cambridge University Press



(a) The coefficient c_H plotted vs. position of the slit.

(b) The coefficient c_V plotted vs. position of the slit.

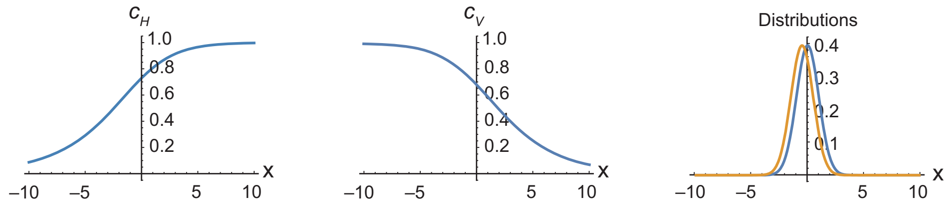
(c) The two distributions of V (left) and H (right). They are well separated.

Figure 3.3 Plots for the case $d = 5$, $\sigma = 1$, where the distributions are well separated. Position is measured in units of σ. $a = b = 1/\sqrt{2}$.

In the case of Fig. 3.3 we can more readily understand what is going on: the distributions are well separated, so when a photon is detected on the left side, it must correspond to a vertically polarized photon, whereas if it is detected on the right side, it must correspond to a horizontally polarized photon. We see there is a continuous (but rather abrupt) transition at the point between the peaks ($x_t = -2.5$), but we also notice that from the distributions shown, these values almost never occur. Consequently we have a good polarization measurement simply from looking at the position of the photon. Indeed, in such a situation, it is common to put a threshold at x_t and call anything to the right to be horizontal, and anything to the left vertical. This is not foolproof, however, because of the very slight overlap between the distributions – these lead to false positives and false negatives. However, in this case, such events are less than 1 percent.

The case of Fig. 3.4 is much more challenging to understand at first glance. We also see that there is a transition between the coefficients of the vertical and horizontal states as the position of the slit is moved across the screen. However, the transition is much more smooth. The reason for this can be seen from the mostly overlapping distributions. In this case, the position of the photon brings very little information about which polarization state the photon is in. Only in the far limits of the position being very negative or very positive can we make a confident statement about which polarization state the system is in. However, these events are very rare. The common events are those where we have very little confidence of which state the polarization is in. This is the “grey region” where we have no certainty. Consequently, the state of the photon’s polarization is quite similar to the initial polarization state. Indeed, we notice at the point where we have absolutely no knowledge at all – $x = x_t = -0.25$, the postmeasurement state is exactly the same as the premeasurement state!

https://doi.org/10.1017/9781009103909.004 Published online by Cambridge University Press



(a) The coefficient c_H plotted vs. position of the slit.

(b) The coefficient c_V plotted vs. position of the slit.

(c) The two distributions of V (left) and H (right). They are overlapping.

Figure 3.4 Plots of the state-coefficients and measurement distributions versus position for the case $d = 0.5$, $\sigma = 1$, where the distributions are well separated. Position is measured in units of σ. $a = b = 1/\sqrt{2}$.

Abstracting Away the Position of the Photon

In the previous analysis, we can simplify these considerations tremendously if we would be able to abstract away the position degree of freedom, and focus only on the polarization of the photon. Notice we can reexpress the probability density of getting the photon position (now simply called a "measurement result"), Eq. (3.7), as the following:

$$p(x) = [\langle H|a^* + \langle V|b^*]\hat{\Omega}_x^\dagger \hat{\Omega}_x [a|H\rangle + b|V\rangle], \tag{3.11}$$

where we have defined a new set of operators $\hat{\Omega}_x$ which operates in the Hilbert space of the photon's polarization. It is indexed by the result of position measurement, x. We can express this operator as a matrix in the basis defined by $|H\rangle, |V\rangle$ as

$$\hat{\Omega}_x = \begin{pmatrix} \phi_H(x) & 0 \\ 0 & \phi_V(x) \end{pmatrix}, \qquad \hat{\Omega}_x^\dagger \hat{\Omega}_x = \begin{pmatrix} |\phi_H(x)|^2 & 0 \\ 0 & |\phi_V(x)|^2 \end{pmatrix}. \tag{3.12}$$

Consequently, our probability rule for finding the measurement result x takes on a very simple form,

$$p(x) = \langle \psi | \hat{\Omega}_x^\dagger \hat{\Omega}_x | \psi \rangle. \tag{3.13}$$

Similarly, we can find the postmeasurement polarization state (3.10) as

$$|\psi'\rangle = \frac{\hat{\Omega}_x|\psi\rangle}{||\hat{\Omega}_x|\psi\rangle||}, \tag{3.14}$$

https://doi.org/10.1017/9781009103909.004 Published online by Cambridge University Press

where $||\dots||$ denotes the norm of the state. It is also important to note the probability density of result x is the square of the denominator of (3.14).

3.4 Mathematics of Generalized Measurement

The mathematical description concerning generalized measurements is an extension of the above physical example. We introduce a set of operators, called Kraus operators, in honor of Karl Kraus (1938–88), also known simply as measurement operators. Let us call this set $\{\hat{\Omega}_j\}$, where $j = 1, \dots, M$. The operators act on the system Hilbert space alone, but there is no necessary relationship between the number of these operators and the dimension of the Hilbert space we are considering. Let us also define the operator $\hat{E}_j = \hat{\Omega}_j^\dagger \hat{\Omega}_j$. $\hat{E}_j$ is a completely positive operator; its Hermitian nature guarantees its eigenvalues are real, and they must be nonnegative. The set of $\{\hat{E}_j\}$ obeys a further constraint; they must sum to the identity operator,

$$\sum_j \hat{E}_j = \mathbf{1}. \tag{3.15}$$

This constraint is analogous to the condition on projection operators, Eq. (2.10). For this reason, the set of the E operators is sometimes called a positive operator-valued measure (POVM) as opposed to a projector-value measure (PVM), although we will strive to avoid this poor choice of name.

As we found in the previous section, the label j denotes different measurement outcomes, so the probability of the outcome j is given by

$$P_j = \langle\psi|\hat{E}_j|\psi\rangle = \langle\psi|\hat{\Omega}_j^\dagger \hat{\Omega}_j|\psi\rangle = ||\hat{\Omega}_j|\psi\rangle||^2. \tag{3.16}$$

The action of these operators on the quantum state may be expressed as the conditional state, given that result j has occurred,

$$|\psi_j\rangle = \frac{\hat{\Omega}_j|\psi\rangle}{||\hat{\Omega}_j|\psi\rangle||} = \frac{\hat{\Omega}_j|\psi\rangle}{\sqrt{P_j}}. \tag{3.17}$$

Therefore, every outcome has (in general) a different quantum state to which the initial state is mapped. It is important to note that neither the operators $\hat{\Omega}_j$ nor $\hat{E}_j$ are in general projection operators. That is,

$$\hat{\Omega}_j \hat{\Omega}_k \neq \hat{\Omega}_j \delta_{jk}, \quad \hat{E}_j \hat{E}_k \neq \hat{E}_j \delta_{jk}, \tag{3.18}$$

and therefore they do not describe the usual textbook projections. This will be quite important in what is to come. We stress that from the point of view of the system, the change of state (3.17) is a coherent but nonunitary process. It may be interpreted as partial collapse of the wavefunction.

The operators $\hat{\Omega}_j$ can be represented in another way in terms of the $\hat{E}_j$ operator as

https://doi.org/10.1017/9781009103909.004 Published online by Cambridge University Press

$$\hat{\Omega}_j = \hat{U}_j\sqrt{\hat{E}_j}, \tag{3.19}$$

where $\hat{U}_j$ is some unitary operator. In Exercise 3.7.4, you will show this decomposition is consistent with the definition of $\hat{E}_j$.

The motivation for this decomposition is that the outcome probabilities only depend on the $\hat{E}_j$ operator. Consequently, we may view the state disturbance (3.17) as a "minimal" amount of quantum disturbance demanded by quantum mechanics, together with a unitary rotation on the state.

Constructing the Measurement Operators

While the general principles of quantum measurement turn out to be fairly simple in the end, there is no general way to construct the measurement operators. One can reformulate this conundrum as the following: what physical procedure should I implement in order to realize a set of measurement operators $\{\hat{\Omega}_j\}$? There is no unique answer to this question because we have abstracted away the measurement device. Nevertheless, we can follow the path of our motivating example and give some general procedure to follow. We consider a system together with an auxiliary system we call a meter. Let us describe them with the quantum state $|\Psi\rangle = |\psi\rangle_S \otimes |\phi\rangle_M$, where the system is in the first register, and the meter is in the second. Now allow them to interact with each other, which we model with an entangling unitary operation $\hat{U}_{SM}$, acting on both system and meter. Now project the meter with a complete set of meter projectors described by $\hat{\Pi}_j = |j\rangle_M {}_M\langle j|$, so the probability of each result is

$$P_j = \langle\Psi|\hat{U}_{SM}^\dagger|j\rangle_M {}_M\langle j|\hat{U}_{SM}|\Psi\rangle. \tag{3.20}$$

The (unnormalized) system state, postmeasurement, is described by

$$|\psi_j\rangle_S = {}_M\langle j|\hat{U}_{SM}|\psi\rangle_S \otimes |\phi\rangle_M = ({}_M\langle j|\hat{U}_{SM}|\phi\rangle_M)|\psi\rangle_S. \tag{3.21}$$

Therefore, we can associate $\hat{\Omega}_j = {}_M\langle j|\hat{U}_{SM}|\phi\rangle_M$. Although this looks like a matrix element, recall that $\hat{U}_{SM}$ acts on both system and meter, and consequently $\hat{\Omega}_j$ is still an operator acting on the system. From this result, we can then reverse engineer the appropriate meter, interaction unitary, and choice of measurements on the meter to generate the desired set of measurement operators.

3.5 von Neumann's Model: An Example with a Qubit and Free Particle Meter

While the preceding formulation looks rather abstract, let us illustrate the general result with a classic example: von Neumann's mode of measurement (sometimes called "premeasurement") [von Neumann (1955)]. We consider a qubit (two-level system such as a spin-1/2 object) as the system, and a free particle in one dimension as the meter. This is slightly generalized from the preceding section since the meter is a continuous variable, but we can use the results of Section 2.4 to fill in any gaps.

Define the initial system and meter state as $|\psi\rangle_S$ and $|\phi\rangle_M$ as in the previous section. We give the free particle its own Hamiltonian that will determine its dynamics, $\hat{H}_M = \hat{p}^2/(2m)$, and the same for the system. Here, $\hat{p}$ is the momentum operator of the meter. We take $\hat{H}_S = \epsilon\hat{\sigma}_z/2$, where σ_z is a Pauli operator for the z coordinate of the spin, and ϵ is the energy splitting between the ground and excited states. Now we allow them to physically interact via an interaction Hamiltonian. In order to illustrate the physics in a very simple way, we let the interaction Hamiltonian be time dependent, so we can clearly separate the pre- and postinteraction phase of the problem. While generally the dynamics of time-dependent Hamiltonians involves time-ordered exponentials and Dyson series, we can further simplify the problem by making the interaction impulsive in time, where a strong, but very short interaction is allowed to exist. We take the extreme of this limit and consider a Dirac delta function in time,

$$\hat{H}_{SM} = g\delta(t)\hat{\sigma}_z \otimes \hat{p}, \tag{3.22}$$

as the interaction Hamiltonian, where the hats stress the operator nature of these observables, involving both system and meter degrees of freedom.

Dividing the problem into time intervals – before the interaction, the interaction at $t = 0$, and after the interaction – allows for the construction of a piecewise solution. From the initial time t_0 to the time of the interaction, both systems have separable, independent dynamics. We can find the state just before the interaction with standard methods to give

$$|\tilde{\psi}\rangle_S = e^{-i\hat{H}_S(t-t_0)/\hbar}|\psi\rangle_S, \quad |\tilde{\phi}\rangle_M = e^{-i\hat{H}_M(t-t_0)/\hbar}|\phi\rangle_M, \tag{3.23}$$

At time $t = 0$, the interaction yields a joint unitary operator

$$\hat{U}_{SM} = \mathcal{T}\exp\left[-i\int^t dt'\hat{H}_{SM}(t')/\hbar\right] = \exp[-ig\hat{\sigma}_z \otimes \hat{p}/\hbar]. \tag{3.24}$$

The time-ordering symbol $\mathcal{T}$ does not play a role, because everything happens at $t = 0$. The time integral in the exponential removes the Dirac-delta function, giving only a unitary interaction that entangles the two degrees of freedom. The strength of the interaction is entirely encoded into the parameter g. Therefore, the state immediately following the interaction is given by

$$|\tilde{\Psi}\rangle = \exp[-ig\hat{\sigma}_z \otimes \hat{p}/\hbar]|\tilde{\psi}\rangle_S \otimes |\tilde{\phi}\rangle_M. \tag{3.25}$$

To proceed further, we recall that the displacement (or translation) operator is defined as

$$\hat{D}(\delta x) = e^{-i\delta x \hat{p}/\hbar}, \tag{3.26}$$

where δx is a position. The action of this operator on position eigenkets is its defining action,

$$\hat{D}(\delta x)|x\rangle = |x + \delta x\rangle. \tag{3.27}$$

Consequently, we can interpret the entangling unitary operator (3.24) as a conditional displacement of the state of the meter, depending on the qubit state. Let us take the qubit system state to be $|\tilde{\psi}\rangle_S = a|+\rangle_S + b|-\rangle_S$, where we choose the basis $|\pm\rangle$ that diagonalizes the system Hamiltonian, $\hat{\sigma}_z|\pm\rangle_S = \pm|\pm\rangle_S$.

Any function of an operator commutes with that operator, so the action of the exponential in Eq. (3.25) is diagonalized in the $|\pm\rangle$ basis on the system. Consequently, the state (3.25) simplifies to

$$|\tilde{\Psi}\rangle = a \exp[-ig\hat{p}/\hbar]|+\rangle_S \otimes |\tilde{\phi}\rangle_M + b \exp[ig\hat{p}/\hbar]|-\rangle_S \otimes |\tilde{\phi}\rangle_M. \tag{3.28}$$

Let us now consider a measurement to see where the meter particle is located. We describe this physics following the prescriptions of Chapter 2 and implement the meter projector $\hat{\Pi}_x = |x\rangle_{MM}\langle x|$, while the system is left untouched. Thus the operation on the system degrees of freedom is simply the identity. The projection on the system reveals some result, x, for the position. Therefore, the meter degree of freedom collapses to state $|x\rangle_M$. The postmeasurement system state, as well as the probability of finding result x, can be found by following the prescription of Subsection 3.4 to find the measurement operators

$$\hat{\Omega}_x = |+\rangle_{SS}\langle+|\, \tilde{\phi}(x-g) + |-\rangle_{SS}\langle-|\, \tilde{\phi}(x+g), \tag{3.29}$$

where the position-space wavefunction of the meter is defined as $_M\langle x|\tilde{\phi}\rangle_M = \tilde{\phi}(x)$.

We notice that similar to our motivating physical experiment – the polarization change by the calcite crystal – there is a shift of the meter. Rather than before, where the extraordinary ray was shifted while the normal ray was not, in this example, both meter states are moved in a symmetric way: the $+$ system state moves the meter to the right, while the $-$ system state moves the meter to the left. We have to add these effects coherently, so the final (unnormalized) system state is

$$|\psi'\rangle_S = a\tilde{\phi}(x-g)|+\rangle_S + b\tilde{\phi}(x+g)|-\rangle_S. \tag{3.30}$$

Therefore, we see that, conditioned on the value of x, the postmeasurement state varies smoothly between $|+\rangle_S$ and $|-\rangle_S$. We can consider this a partial wavefunction collapse, depending on how strongly correlated the system and meter are. We will explore this connection in greater depth in the next chapter.

It is interesting to consider variations on this scenario. For example, let us allow the system and meter to change in time further after the interaction for a time t. How will this change the preceding story? Similar to the beginning of this section, we must apply the separable unitary to the now entangled state,

$$|\tilde{\Psi}'\rangle = e^{-i\hat{H}_S t/\hbar} \otimes e^{-i\hat{H}_M t/\hbar}|\tilde{\Psi}\rangle. \tag{3.31}$$

Suppose now a measurement of the meter's position is made. How will the measurement operators change? As you will prove in Exercise 3.7.7, the resulting operators are given by

$$\begin{aligned}\hat{\Omega}'_x = {} & |+\rangle_{SS}\langle+|\, e^{-i\epsilon t/(2\hbar)}\langle x|e^{-i\hat{p}^2 t/(2m\hbar) - ig\hat{p}/\hbar}|\tilde{\phi}\rangle \\ & + |-\rangle_{SS}\langle-|\, e^{i\epsilon t/(2\hbar)}\langle x|e^{-i\hat{p}^2 t/(2m\hbar) + ig\hat{p}/\hbar}|\tilde{\phi}\rangle.\end{aligned} \tag{3.32}$$

Remarkably, the preceding result is still diagonal in the $\hat{\sigma}_z$ eigenbasis of the system. This property comes from the fact that the system Hamiltonian and the interaction Hamiltonian commute,

$$[\hat{H}_S, \hat{H}_{SM}] = 0. \tag{3.33}$$

In the literature, such a measurement is sometimes called a *quantum nondemolition measurement*, for historical reasons. The intuition why such a measurement is desirable is that, if one's goal is to measure the system in a textbook way in the energy eigenbasis (by making g much larger than the spread in the meter wavefunction, for example), then the measurement should not rotate the state during the measurement process. Otherwise, this would spoil what one is trying to do.

3.6 Generalization to Mixed States

At this point in the text, it is helpful to generalize our results to the situation when one has incomplete quantum information. This situation can come from ignorance of certain events in the world and is described with mixed states, corresponding to a density matrix or density operator. Readers unfamiliar with this description of quantum systems, or desiring a refresher, should stop here and read Appendix B, and then return.

If we have a description of the state as a mixed state $\hat{\rho}$, how do we generalize the statistical predictions of quantum mechanics? For projective measurements, in Appendix B, we already showed that the outcome probability for result j of a projection operator set $\hat{\Pi}_j$ is given by

$$P_j = \mathrm{Tr}[\hat{\rho}\hat{\Pi}_j], \tag{3.34}$$

while the postmeasurement state $\hat{\rho}'$ is simply $\hat{\Pi}_j$. To extend these results to generalized measurements, we recall we can decompose any mixed state as

$$\hat{\rho} = \sum_j p_j |\psi_j\rangle\langle\psi_j|, \tag{3.35}$$

for some states $|\psi_j\rangle$ and weights $0 \leq p_j \leq 1$. Recalling results (3.16) and (3.17), we can rewrite the probability expression for some result k as

$$P_k = \mathrm{Tr}[\hat{\Omega}_k|\psi\rangle\langle\psi|\hat{\Omega}_k^\dagger], \tag{3.36}$$

where Tr is the trace of the operator. Replacing $|\psi\rangle \to |\psi_j\rangle$, we can generalize the result to any mixed state, using the rule of total probability,

$$P_k = \sum_j p_j \mathrm{Tr}[\hat{\Omega}_k|\psi_j\rangle\langle\psi_j|\hat{\Omega}_k^\dagger] = \mathrm{Tr}[\hat{\Omega}_k\hat{\rho}\hat{\Omega}_k^\dagger] = \mathrm{Tr}[\hat{\Omega}_k^\dagger\hat{\Omega}_k\hat{\rho}], \tag{3.37}$$

where we have used the cyclic property of the trace in the last equality. The important difference to Eq. (3.34) is that, while $\hat{\Pi}_j$ is a projection operator, the operator $\hat{E}_k = \hat{\Omega}_k^\dagger\hat{\Omega}_k$ is only a positive, Hermitian operator.

The measurement disturbance can be found in a similar way. Equation (3.17) can be invoked, as applied to unnormalized states; and also, using its adjoint form, we have

$$|\psi_j\rangle\langle\psi_j| \to \hat{\Omega}_k|\psi_j\rangle\langle\psi_j|\hat{\Omega}_k^\dagger. \tag{3.38}$$

We can now apply the same rule to an incoherent superposition of many such states to find

$$\hat{\rho} \to \hat{\rho}' = \sum_j p_j(\hat{\Omega}_k|\psi_j\rangle\langle\psi_j|\hat{\Omega}_k^\dagger) = \hat{\Omega}_k\hat{\rho}\hat{\Omega}_k^\dagger. \tag{3.39}$$

As in the pure state case, we must renormalize the state such that $\mathrm{Tr}[\hat{\rho}'] = 1$ (see Appendix B). This is accomplished simply by dividing by the trace of $\hat{\rho}'$. We thus have a map of normalized states to normalized states, conditioned on result k,

$$\hat{\rho} \to \hat{\rho}' = \frac{\hat{\Omega}_k\hat{\rho}\hat{\Omega}_k^\dagger}{\mathrm{Tr}[\hat{\Omega}_k^\dagger\hat{\Omega}_k\hat{\rho}]}, \tag{3.40}$$

where we used the cyclic property of the trace in the denominator. Notice the denominator is exactly the same as the outcome probability (3.37).

While we have used the generalized measurement formalism to derive these results, it is instructive to rederive them from the point of view of expanding the system to include a meter as we did when deriving an explicit form of the measurement operators. Consider a mixed system state and a pure meter state $|\phi\rangle_M$ for simplicity, initially separable. Now let them interact with a joint interaction unitary operator $\hat{U}_{SM}$ resulting in the state

$$\hat{\rho}_{SM} = \hat{\rho}_S \otimes |\phi\rangle_{MM}\langle\phi| \rightarrow \hat{\rho}'_{SM} = \hat{U}_{SM}\hat{\rho}_S \otimes |\phi\rangle_{MM}\langle\phi|\hat{U}^{\dagger}_{SM}. \tag{3.41}$$

Applying a projection operator $\hat{\Pi}_k$ on the meter degrees of freedom (and identity on the system) gives the prediction of the probability of result k (3.34),

$$P_k = \mathrm{Tr}[\hat{\rho}'_{SM}\hat{\Pi}_k] = \mathrm{Tr}_S\mathrm{Tr}_M[\hat{U}_{SM}\hat{\rho}_S \otimes |\phi\rangle_{MM}\langle\phi|\hat{U}^{\dagger}_{SM}\hat{\Pi}_k], \tag{3.42}$$

where $\mathrm{Tr}_{S,M}$ refers to tracing over either the system or the meter part of the Hilbert space. The trace of the meter is carried out by choosing the basis corresponding to the projection operators on the meter to find

$$P_k = \mathrm{Tr}_S\left[\sum_l {}_M\langle l|\hat{U}_{SM}\hat{\rho}_S \otimes |\phi\rangle_{MM}\langle\phi|\hat{U}^{\dagger}_{SM}\hat{\Pi}_k|l\rangle_M\right], \tag{3.43}$$

$$= \mathrm{Tr}_S\left[{}_M\langle k|\hat{U}_{SM}\hat{\rho}_S \otimes |\phi\rangle_{MM}\langle\phi|\hat{U}^{\dagger}_{SM}|k\rangle_M\right], \tag{3.44}$$

$$= \mathrm{Tr}_S\left[({}_M\langle k|\hat{U}_{SM}|\phi\rangle_M)\hat{\rho}_S({}_M\langle\phi|\hat{U}^{\dagger}_{SM}|k\rangle_M)\right]. \tag{3.45}$$

Recalling the result (3.21), from the discussion in Section 3.4, we see that the result (3.37) is recovered with the identification $\hat{\Omega}_k = {}_M\langle k|\hat{U}_{SM}|\phi\rangle_M$. In Exercise 3.7.10, you will find the postmeasurement state of the system and meter, completing the preceding discussion.

The Mixed State Case for the von Neumann Measurement Model

We now illustrate the previous section by applying the results to the von Neumann measurement model of Section 3.5. Recall that we found, given the detection of the free particle at position x on the final screen, the measurement operators are given by Eq. (3.29), so they are diagonal in the $|+\rangle, |-\rangle$ basis.

For a qubit, the density matrix can be written in the $|+\rangle, |-\rangle$ basis as

$$\hat{\rho} = \begin{pmatrix} \rho_{++} & \rho_{+-} \\ \rho_{-+} & \rho_{--} \end{pmatrix}. \tag{3.46}$$

The positivity of the density tells us that $0 \leq \rho_{++}, \rho_{--} \leq 1$, while its normalization indicates that $\mathrm{Tr}\hat{\rho} = \rho_{++} + \rho_{--} = 1$. The fact that $\hat{\rho}$ is Hermitian further constrains $\rho_{+-} = \rho^*_{-+}$. Generalizing the results of the previous section to continuous variables, using the results of Section 2.4, the outcome probability, Eq. (3.37), gives us the probability density of finding the free particle at position x,

$$p(x) = \mathrm{Tr}[\hat{\Omega}^{\dagger}_x\hat{\Omega}_x\hat{\rho}] = \rho_{++}|\tilde{\phi}(x-g)|^2 + \rho_{--}|\tilde{\phi}(x+g)|^2. \tag{3.47}$$

https://doi.org/10.1017/9781009103909.004 Published online by Cambridge University Press

Thus, only the diagonal density matrix elements enter this expression. Similarly, the unnormalized state of the qubit can be found from Eq. (3.39),

$$\hat{\rho}' \propto \begin{pmatrix} |\tilde{\phi}(x-g)|^2 \rho_{++} & \tilde{\phi}(x-g)^* \tilde{\phi}(x+g) \rho_{+-} \\ \tilde{\phi}(x+g)^* \tilde{\phi}(x-g) \rho_{-+} & |\tilde{\phi}(x+g)|^2 \rho_{--} \end{pmatrix}. \tag{3.48}$$

These results generalize those from Section 3.5. To normalize this density operator, we should divide by the sum of the diagonal elements. Thus we see that the diagonal matrix elements are weighted by the absolute square of the meter wavefunctions, shifted by $\pm g$, while the off-diagonal matrix elements are weighted by $\tilde{\phi}(x+g)^* \tilde{\phi}(x-g)$, or its complex conjugate.

3.7 Quantum Bayesian Point of View

Let us return to the main results of Section 3.6, given in Eqs. (3.37) and (3.40). Further insight into the meaning of these equations can be had by using the decomposition (3.19), recalling that $\hat{E}_k = \hat{\Omega}_k^\dagger \hat{\Omega}_k$. We can then write the probability of result k and the conditional state change assignment $\hat{\rho}^{(k)}$ as

$$P_k = \mathrm{Tr}(\hat{E}_k \hat{\rho}), \quad \hat{\rho} \to \hat{\rho}^{(k)} = \frac{\hat{U}_k \sqrt{\hat{E}_k} \hat{\rho} \sqrt{\hat{E}_k} \hat{U}_k^\dagger}{P_k}. \tag{3.49}$$

We will now consider the simplest case of "plain" measurement, where no additional effective unitary operation is taking place, $\hat{U}_k = \mathbf{1}$ for all k. We consider the situation when the operators $\hat{E}_k$ are diagonalized all in the same basis so

$$\hat{E}_k |j\rangle = p(k|j) |j\rangle, \tag{3.50}$$

for all k. Here, the eigenvalues $p(k|j)$ have the natural interpretation as the conditional probability of finding result k, given that the system is in state $|j\rangle$. We can see this explicitly by finding the probability of result k after preparing the system in each one of its N states labeled by j. The total probability of result k occurring is then given by

$$P_k = \sum_j p(k|j) \rho_{jj}. \tag{3.51}$$

This is nothing more than the law of total probability described in Appendix A for classical probability theory, where we interpret the diagonal density matrix elements to be the probability of finding the system in state $|j\rangle$.

The diagonal matrix elements of the new density matrix also take on a simple form:

$$\langle j|\hat{\rho}^{(k)}|j\rangle = \rho_{jj}^{(k)} = \frac{p(k|j)\rho_{jj}}{P_k}. \tag{3.52}$$

This part of the state assignment may also be interpreted with classical probability theory as Bayes' rule (see Appendix A). The diagonal density matrix elements, interpreted as the probability to occupy state $|j\rangle$, are reassigned based on new information coming in the form of the result k, with the aid of the conditional probabilities $p(k|j)$.

The "quantum" part of the update rule can therefore be found in how the coherences change, given by

$$\rho_{ij}^{(k)} = \frac{\langle i|\sqrt{\hat{E}_k}\hat{\rho}\sqrt{\hat{E}_k}|j\rangle}{P_k} = \frac{\sqrt{p(k|i)p(k|j)}\,\rho_{ij}}{P_k}, \tag{3.53}$$

for $i \neq j$. This connection to Bayes' rule was noted in the book of Gardiner et al. (2004), and in the papers of Alexander Korotkov and collaborators (Korotkov, 1999, 2001; Jordan and Korotkov, 2006). Further elucidation of these results will be given in the next chapter.

Unselective Measurements and Decoherence

While we have so far focused on *selective measurements* (i.e. when we are interested in the conditional state disturbance corresponding to a single outcome), it is quite common to discuss *unselective* measurements, where either the meter state cannot be read, or it is read and forgotten, or averaged over. Such a situation becomes unavoidable in the case where the "meter" consists of many degrees of freedom, such as in a macroscopic environment. In this case, we can investigate the averaged quantum dynamics over all possible outcomes. We define the ensemble-averaged density matrix $\bar{\hat{\rho}}$ as

$$\bar{\hat{\rho}} = \sum_k P_k \hat{\rho}^{(k)} = \sum_k \hat{\Omega}_k \hat{\rho} \hat{\Omega}_k^\dagger. \tag{3.54}$$

Sometimes this mapping between $\hat{\rho} \to \bar{\hat{\rho}}$ is called a *quantum channel*. In the simple case discussed previously, where $\hat{U}_k = \mathbf{1}$, and $\hat{E}_k$ are all diagonalized in the same basis $\{|i\rangle\}$, we have the diagonal density matrix elements given by

$$\bar{\rho}_{ii} = \sum_k p(k|i)\rho_{ii} = \rho_{ii}. \tag{3.55}$$

The last equality follows from the conditional probability distributions $p(k|i)$ being normalized for every i. Therefore, the diagonal matrix elements are left unchanged by the averaging over results. On the other hand, the off-diagonal elements behave as

$$\bar{\rho}_{ij} = \sum_k \sqrt{p(k|i)p(k|j)}\,\rho_{ij}. \tag{3.56}$$

Here, we see that the off-diagonal matrix elements are multiplied by an overall factor

$$C_{ij} = \sum_k \sqrt{p(k|i)p(k|j)}. \tag{3.57}$$

Recalling the interpretation of $p(k|i)$ as the conditional probability of result k, given state i, the coefficient C_{ij} is the Bhattacharyya coefficient between distributions $p(k|i)$ and $p(k|j)$ (Bhattacharyya, 1943), which is closely related to one of the Rényi divergences (Van Erven and Harremos, 2014; Averin and Sukhorukov, 2005). The Bhattacharyya distance D_{ij} is related to the Bhattacharyya coefficient as $D_{ij} = -\ln C_{ij}$. This coefficient is a measure of the similarity of the two distributions. We note that $0 \leq C_{ij} \leq 1$, with the larger distance between the distributions corresponding to smaller values of C_{ij}. Thus, for $p(k|i) \neq p(k|j)$, the coherence of the density matrix element is suppressed. This process is called *decoherence*. It is important to stress that the loss of coherence comes from the loss of information and incoherently averaging together different density matrices corresponding to different realizations of the meter or environment.

Exercises

Exercise 3.7.1 Show that the postmeasurement state (3.10) of the photon can be written as $|\psi''\rangle = c_H(x)|H\rangle + c_V(x)|V\rangle$, where the new coefficients may be written as

$$c_H(x) = \frac{a}{\sqrt{|a|^2 + |b|^2 \exp(-xd/\sigma^2 - d^2/(2\sigma^2))}}, \tag{3.58}$$

$$c_V(x) = \frac{b \exp(-xd/(2\sigma^2) - d^2/(4\sigma^2))}{\sqrt{|a|^2 + |b|^2 \exp(-xd/\sigma^2 - d^2/(2\sigma^2))}}. \tag{3.59}$$

Exercise 3.7.2 Consider a meter that is another qubit (2-state system). The system is prepared in a general state, $|\psi\rangle_S = a|H\rangle_S + b|V\rangle_S$. The meter is prepared in the state $|\phi\rangle_M = \gamma|H\rangle_M + \bar{\gamma}|V\rangle_M$. Here we have used the symbols H, V instead of $0, 1$ to suggest polarization states of single photons. The physical interaction between system and meter maps the initially separable state to $|\psi\rangle_S \otimes |\phi\rangle_M$ to the state

$$|\Psi\rangle = (a\gamma|H\rangle_S + b\bar{\gamma}|V\rangle_S) \otimes |H\rangle_M + (a\bar{\gamma}|H\rangle_S + b\gamma|V\rangle_S) \otimes |V\rangle_M. \tag{3.60}$$

The meter photon is measured in the $H - V$ basis. Follow the procedure outlined in Section 3.4 to find the measurement operators appropriate for the system. This measurement was carried out by Pryde et al. (2005).

https://doi.org/10.1017/9781009103909.004 Published online by Cambridge University Press

Exercise 3.7.3 Verify the result (3.14) is correct by inserting the explicit form (3.12) for the preceding equation and check to see that you recover Eq. (3.10).

Exercise 3.7.4 Show that the decomposition (3.19) is consistent with $\hat{E}_j = \hat{\Omega}_j^\dagger \hat{\Omega}_j$.

Exercise 3.7.5 Confirm for yourself that the measurement operator given in Eq. (3.29) is correct.

Exercise 3.7.6 Consider the measurement operators (3.29). Consider the case where the initial meter wavefunction is Gaussian with width σ. Work out the simplified form of the postmeasurement system state that is normalized.

Exercise 3.7.7 Show that the measurement operators (3.32) for adding dynamics to the meter and system are correct.

Exercise 3.7.8 Consider the measurement operators (3.32). Consider the case where the initial meter wavefunction is Gaussian with width σ. Work out the simplified form of the measurement operator coefficients explicitly. Hint: compute the meter matrix elements by entering the momentum basis, and then Fourier transforming back to the position basis at the end.

Exercise 3.7.9 What would happen if the initial system Hamiltonian for Section 3.5 was changed to $\hat{H}_S = \Delta\hat{\sigma}_x/2$ instead? Would the measurement still be quantum nondemolition? Work out the new measurement operators after waiting some time t.

Exercise 3.7.10 Following the discussion in Section 3.6, find the postmeasurement state of the system and meter and show the system state is of the form (3.40), where the meter state is a projector on state $|k\rangle_M$.

Exercise 3.7.11 What would happen if the coupling Hamiltonian for Section 3.5 coupled to the meter's position operator, instead of its momentum operator? That should correspond to a splitting of the meter's state in momentum space, so subsequent free evolution in time would separate the wavepackets corresponding to different eigenstates of the system. Thus, the amount of time elapsed should also control the strength of the measurement. Confirm this is true quantitatively.

Exercise 3.7.12 Consider a fully mixed polarization state of light that enters a calcite crystal. Following a slit at position x, for $d = 0.5$ and $\sigma = 1$, find the purity of the quantum polarization state of a photon that passes the slit. Plot it as a function of x. Does this make sense?

Exercise 3.7.13 Consider the von Neumann model for when the meter is prepared in a Gaussian state of width σ. For an interaction that shifts the meter wavefunction by a distance of $\pm g$ depending on the qubit state, followed by immediate measurement, calculate the Bhattacharyya distance as a function of g and the meter width σ. For what value of g/σ will the ensemble-averaged coherence of the quantum state be reduced by a factor of 100? Of 1000?