9. Quantum Foundations

What is the essence of quantum mechanics? What makes the quantum world truly different from the classical one? Is it the discrete spectrum of energy levels? Or the inherent lack of determinism?

The purpose of this chapter is to go back to basics in an attempt to answer this question. For the most part, we will not be interested in the dynamics of quantum systems (although Section 9.5 is an exception). Instead, we will look at the framework of quantum mechanics in an attempt to get a better understanding of what we mean by a "state", and what we mean by a "measurement".

9.1 Entanglement

> "I would not call that *one* but rather *the* characteristic trace of quantum mechanics, the one that enforces its entire departure from classical lines of thought"
>
> *Erwin Schrödinger on entanglement*

The differences between the classical and quantum worlds are highlighted most emphatically when we look at a property called *entanglement*. This section and, indeed, much of this chapter will be focussed on building the tools necessary to understand the surprising features of entangled quantum states.

Entanglement is a property of two or more quantum systems. Here we consider two systems, with associated Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. The Hilbert space of the combined system is then $\mathcal{H}_1 \otimes \mathcal{H}_2$. A state of this combined system is said to be *entangled* if it cannot be written in the form

$$|\Psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \tag{9.1}$$

For example, suppose we have two particles, each of which can have one of two states. This is called a *qubit*. We take a basis of this Hilbert space to be the spin in the z-direction, with eigenstates spin up $|\uparrow\rangle$ or spin down $|\downarrow\rangle$. Then the state

$$|\Psi\rangle = |\uparrow\rangle \otimes |\downarrow\rangle$$

is not entangled. In contrast, the state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle)$$

is entangled. In fact, this is the most famous of all entangled states and is usually known as an *EPR pair*, after Einstein, Podolsky and Rosen. Note that this state is a sum over states of the form (9.1) and cannot be written in a simpler form; this is what makes it entangled. In what follows, we'll simplify our notation and drop the $\otimes$ symbol, so the EPR pair is written as

$$|EPR\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle) \tag{9.2}$$

To illustrate the concept of entanglement, we could just as easily have chosen the states $|\Psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle + |\downarrow\rangle|\uparrow\rangle)$ or $|\Psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\uparrow\rangle + |\downarrow\rangle|\downarrow\rangle)$. Both of these are also entangled. However, just because a state is written as a sum of terms of the form (9.1) does not necessarily mean that it's entangled. Consider, for example,

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle + |\downarrow\rangle|\downarrow\rangle)$$

This can also be written as $|\Psi\rangle = |\rightarrow\rangle|\downarrow\rangle$ where $|\rightarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$ and so this state is not entangled. We'll provide a way to check whether or not a state is entangled in Section 9.3.3.

9.1.1 The Einstein, Podolsky, Rosen "Paradox"

In 1935, Einstein, Podolsky and Rosen tried to use the property of entanglement to argue that quantum mechanics is incomplete. Ultimately, this attempt failed, revealing instead the jarring differences between quantum mechanics and our classical worldview.

Here is the EPR argument. We prepare two particles in the state (9.2) and subsequently separate these particles by a large distance. There is a tradition in this field, imported from the world of cryptography, to refer to experimenters as Alice and Bob and it would be churlish of me to deny you this joy. So Alice and Bob sit in distant locations, each carrying one of the spins of the EPR pair. Let's say Alice chooses to measure her spin in the z-direction. There are two options: she either finds spin up $|\uparrow\rangle$ or spin down $|\downarrow\rangle$ and, according to the rules of quantum mechanics, each of these happens with probability 50%. Similarly, Bob can measure the spin of the second particle and also finds spin up or spin down, again with probability 50%.

However, the measurements of Alice and Bob are not uncorrelated. If Alice measures the first particle to have spin up, then the EPR pair (9.2) collapses to $|\uparrow\rangle|\downarrow\rangle$, which means that Bob *must* measure the spin of the second particle to have spin down. It would appear, regardless of how far apart they are, the measurement of Alice determines the measurement of Bob: whatever Alice sees, Bob always sees the opposite. Viewed in the usual framework of quantum mechanics, these correlations arise because of a "collapse of the wavefunction" which happens instantaneously.

Now, for any theoretical physicist — and for Einstein in particular — the word "instantaneous" should ring alarm bells. It appears to be in conflict with special relativity and, although we have not yet made any attempt to reconcile quantum mechanics with special relativity, it would be worrying if they are incompatible on such a fundamental level.

The first thing to say is that there is no direct conflict with locality, in the sense that there is no way to use these correlations to transmit information faster than light. Alice and Bob cannot use their entangled pair to send signals to each other: if Bob measures spin down then he has no way of knowing whether this happened because he collapsed the wavefunction, or if it happened because Alice has already made a measurement and found spin up. Nonetheless, the correlations that arise *appear* to be non-local and this might lead to a sense of unease.

There is, of course, a much more mundane explanation for the kinds of correlations that arise from EPR pairs. Suppose that I take off my shoes and give one each to Alice and Bob, but only after I've sealed them in boxes. I send them off to distant parts of the Universe where they open the boxes to discover which of my shoes they've been carrying across the cosmos. If Alice is lucky, she finds that she has my left shoe. (It is a little advertised fact that Alice has only one leg.) Bob, of course, must then have my right shoe. But there is nothing miraculous or non-local in all of this. The parity of the shoe was determined from the beginning; any uncertainty Alice and Bob had over which shoe they were carrying was due only to their ignorance, and my skill at hiding shoes in boxes.

This brings us to the argument of EPR. The instantaneous collapse of the wavefunction in quantum mechanics is silly and apparently non-local. It would be much more sensible if the correlations in the spins could be explained in the same way as the correlations in shoes. But if this is so, then quantum mechanics must be incomplete because the state (9.2) doesn't provide a full explanation of the state of the system. Instead, the outcome of any measurement should be determined by some property of the spins that is not encoded in the quantum state (9.2), some extra piece of information that was there from the beginning and says what the result of any measurement will give. This hypothetical extra piece of information is usually referred to as a *hidden variable*. It was advocated by Einstein and friends as a way of restoring some common sense to the world of quantum mechanics, one that fits more naturally with our ideas of locality.

There's no reason that we should have access to these hidden variables. They could be lying beyond our reach, an inaccessible deterministic world which we can never see. In this picture, our ignorance of these hidden variables is where the probability of quantum mechanics comes from, and the uncertainties of quantum mechanics are then no different from the uncertainties that arise in the weather or in the casino. They are due, entirely, to lack of knowledge. This wonderfully comforting vision of the Universe is sometimes called *local realism*. It is, as we will now show, hopelessly naive.

9.1.2 Bell's Inequality

The hypothetical hidden variables that determine the measurements of spin must be somewhat more subtle than those that determine the measurement of my shoes. This is because there's nothing to stop Alice and Bob measuring the spin in directions other than the z-axis.

Suppose, for example, that both choose to measure the spin in the x-direction. The eigenstates for a single spin are

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) \quad , \quad |\leftarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle)$$

with eigenvalues $+\hbar/2$ and $-\hbar/2$ respectively. We can write the EPR pair (9.2) as

$$|EPR\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle) = \frac{1}{\sqrt{2}}(|\leftarrow\rangle|\rightarrow\rangle - |\rightarrow\rangle|\leftarrow\rangle)$$

So we again find correlations if the spins are measured along the x-axis: whenever Alice finds spin $+\hbar/2$, then Bob finds spin $-\hbar/2$ and vice-versa. Any hidden variable has to account for this too. Indeed, the hypothetical hidden variables have to account for the measurement of the spin along any choice of axis. This will prove to be their downfall.

A Review of Spin

Before we proceed, let's first review a few facts about how we measure the spin along different axes. An operator that measures spin along the direction $\mathbf{a} = (\sin\theta, 0, \cos\theta)$ is

$$\boldsymbol{\sigma} \cdot \mathbf{a} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

Below we'll denote this matrix as $\boldsymbol{\sigma} \cdot \mathbf{a} = \sigma_\theta$. It has eigenvectors

$$|\theta_+\rangle = \cos\frac{\theta}{2}|\uparrow\rangle + \sin\frac{\theta}{2}|\downarrow\rangle \quad \text{and} \quad |\theta_-\rangle = -\sin\frac{\theta}{2}|\uparrow\rangle + \cos\frac{\theta}{2}|\downarrow\rangle$$

From this, we learn that if we prepare a state in, say, $|\downarrow\rangle$, then the probability $P(\theta_\pm)$ of measuring either spin $+$ or spin $-$ along the vector $\mathbf{a}$ is

$$P(\theta_+) = \sin^2\frac{\theta}{2} \quad \text{and} \quad P(\theta_-) = \cos^2\frac{\theta}{2}$$

From the form of the eigenstates $|\theta_\pm\rangle$, we see that the EPR pair can be written as

$$|EPR\rangle = \frac{1}{\sqrt{2}}(|\theta_+\rangle|\theta_-\rangle - |\theta_-\rangle|\theta_+\rangle) \tag{9.3}$$

for any θ. This means that, as long as Alice and Bob both choose to measure the spin along the same direction $\mathbf{a}$, then their results will always be perfectly anti-correlated: when one measures spin $+$ the other is guaranteed to measure spin $-$. This is a special property of the EPR pair that is not shared by other entangled states. It follows from some group theory: under addition of angular momentum $\frac{1}{2}\otimes\frac{1}{2} = 0 \oplus 1$, and the EPR state is the rotationally invariant singlet.

What's Wrong with Hidden Variables

Suppose, now that Alice measures the spin along the z-axis, and Bob measures the spin along the $\mathbf{a}$ axis. If Alice measures spin $|\uparrow\rangle$, then we know that Bob has spin $|\downarrow\rangle$, so whether he measures spin $+$ or $-$ is determined by the probabilities above. We'll write this as $P(\sigma_z^A, \sigma_\theta^B)$ where σ^A denotes the spin measured by Alice and σ^B the spin measured by Bob. There are four possibilities are

$$\begin{aligned} P(\sigma_z^A = +, \sigma_\theta^B = +) &= \frac{1}{2}\sin^2\frac{\theta}{2} \quad , \quad P(\sigma_z^A = +, \sigma_\theta^B = -) = \frac{1}{2}\cos^2\frac{\theta}{2} \\ P(\sigma_z^A = -, \sigma_\theta^B = +) &= \frac{1}{2}\cos^2\frac{\theta}{2} \quad , \quad P(\sigma_z^A = -, \sigma_\theta^B = -) = \frac{1}{2}\sin^2\frac{\theta}{2} \end{aligned} \tag{9.4}$$

Note, in particular, that if $\theta = 0$ so that Alice and Bob measure the spin along the same axis, then we revert to our previous perfect anti-correlation.

It is not difficult to account for these results in a hidden variables theory. Each of the particles carries with them two labels s_z and s_θ which have values $+1$ or -1 and determine the result of a spin measurement along the z-axis and $\mathbf{a}$ axis respectively. The perfect anti-correlation means that the value of each spin for Bob's particle must be the opposite of Alice's. We write $s_z^B = -s_z^A$ and $s_\theta^B = -s_\theta^A$. We then only need to talk about the probability distribution $p(s_z^A, s_\theta^A)$ for the spins of Alice's particles. To reproduce the predictions (9.4), we must take these to be

$$\begin{aligned} P(s_z^A = +, s_\theta^A = -) &= \frac{1}{2}\sin^2\frac{\theta}{2} \quad , \quad P(s_z^A = +, s_\theta^A = +) = \frac{1}{2}\cos^2\frac{\theta}{2} \\ P(s_z^A = -, s_\theta^A = -) &= \frac{1}{2}\cos^2\frac{\theta}{2} \quad , \quad P(s_z^A = -, s_\theta^A = +) = \frac{1}{2}\sin^2\frac{\theta}{2} \end{aligned} \tag{9.5}$$

Mathematically this is straightforward: the probability distributions are, after all, essentially the same as those in (9.4). But physically we've done something a little slippery. We've said that whenever Bob measures his spin σ_θ to be, say, $+1$ then this determines the spin of Alice's particle to be $s_\theta = -1$ *even though* Alice didn't measure the spin in the direction $\mathbf{a}$. In this way, we've managed to assign labels to Alice's particle corresponding to spin in two different directions. But this is against the spirit of quantum mechanics because these operators for spins in different directions don't commute. Indeed, we will now see that the spirit of quantum mechanics will come back and bite us.

The trouble comes when we throw a third possible measurement into the mix. Suppose that Alice and Bob are given a choice. Each can measure the spin along the z-axis, along the $\mathbf{a} = (\sin\theta, 0, \cos\theta)$ axis or along the $\mathbf{b} = (\sin\phi, 0, \cos\phi)$ axis. Now each particle must be assigned a hidden variable that determines the choice of each of these measurements. So Alice's particle comes with s_z^A, s_θ^A and s_ϕ^A, each of which can take value ± 1. The probabilities of the different choices are governed by some distribution $p(s_z^A, s_\theta^A, s_\phi^A)$. We will now show that no such distribution exists that can reproduce the results of measurements of the EPR pair.

Let's assume that such a distribution does exist. This implies certain relations between the probability distributions $P(s_i^A, s_j^A)$. For example, by summing over the variables which weren't measured, we find

$$\begin{aligned} P(s_\theta^A = +, s_\phi^A = -) &= p(++-) + p(-+-) \\ &\leq [p(++-) + p(+++)] + [p(-+-) + p(---)] \\ &= P(s_z^A = +, s_\theta^A = +) \, + \, P(s_z^A = -, s_\phi^A = -) \end{aligned}$$

But we know what each of these distributions $P(s^A, s^A)$ must be: they are given by (9.5). This then gives the *Bell inequality*

$$\sin^2\frac{\theta - \phi}{2} \leq \cos^2\frac{\theta}{2} + \cos^2\frac{\phi}{2} \tag{9.6}$$

where the left-hand side follows from the rotational invariance (9.3) of the EPR state.

There's a problem with the Bell inequality (9.6): it's simply not true for all values of θ and ϕ! Suppose, for example, that we take $\theta = 3\pi/2$ and $\phi = 3\pi/4$. Then

$$\sin^2\frac{3\pi}{8} - \cos^2\frac{3\pi}{8} = -\cos\frac{3\pi}{4} = \frac{1}{\sqrt{2}}$$

Meanwhile

$$\cos^2 \frac{3\pi}{4} = \frac{1}{2}$$

Obviously $1/2 < 1/\sqrt{2}$. These values violate the Bell inequality.

The Bell inequality (9.6) was derived under the assumption that there was some hidden variable underlying quantum mechanics. Its violation tells us that this is simply not possible. Of course, physics is an experimental science and we can ask whether or not the Bell inequalities are violated in Nature. They are. The experiment was first done in the early 1980s by Aspect and has been repeated many times since, with different groups trying to finesse the experiments in order to close off increasingly preposterous loopholes that philosophers claim to have discovered in the argument.

The original EPR argument was an attempt to show that locality, together with common sense, imply that there should be hidden variables underlying quantum mechanics. Nature, however, disagrees. Indeed, the Bell inequalities turn the EPR argument completely on its head. If you want to keep locality, then you're obliged to give up common sense which, here, means a view of the world in which particles carry the properties that are measured. In contrast, if you want to keep common sense, you will have to give up locality. Such a loophole arises because the derivation of Bell's inequality assumed that a measurement on one particle does not affect the probability distribution of the other. Given that the two particles can be separated by arbitrarily large distances, any such effect must be superluminal and, hence, non-local. Therefore, the best one can say is that Bell's argument forbids *local* hidden variable theories.

Most physicists cherish locality over common sense. In particular, all of our most successful laws of physics are written in the language of Quantum Field Theory, which is the framework that combines quantum mechanics with local dynamics. With locality sitting firmly at the heart of physics, it is very difficult to see role for any kind of hidden variables.

It is sometimes said that the correlations inherent in EPR-type pairs are *non-local*. I don't think this is a particularly helpful way to characterise these correlations because, as we have seen, there is no way to use them to signal faster than light. Nonetheless, it is true that the correlations that arise in quantum mechanics cannot arise in any local classical model of reality. But the key lesson to take from this is not that our Universe is non-local; it is instead that our Universe is non-classical.

9.1.3 CHSH Inequality

The essence of Bell's inequality can be distilled to a simpler form, due to Clauser, Horne, Shimony and Holt.

We stick with the general framework where both Alice and Bob are each sent a two-state quantum system. Alice can choose to measure one of two quantum observables, A_1 or A_2. Similarly, Bob can choose to measure B_1 or B_2. Each of these observables has two possible eigenvalues, $a_i \pm 1$ and $b_i = \pm 1$.

We require that

$$[A_i, B_j] = 0 \qquad i, j = 1, 2 \tag{9.7}$$

This is the statement that Alice and Bob can happily perform their measurements without interfering with the other. In particular, this is where the assumption of locality comes in: if Alice and Bob are spacelike separated then (9.7) must hold. In contrast, we will make no such assumption about $[A_1, A_2]$ or $[B_1, B_2]$.

We're going to look at the expectation value of the observable

$$C = (A_1 + A_2)B_1 + (A_1 - A_2)B_2 \tag{9.8}$$

We do this first in a hidden variable theory, and next in the quantum theory. We'll see that a hidden variable theory places a stricter range on the allowed values of the expectation value $\langle C \rangle$. To see this, we make the seemingly innocuous assumption that the system possesses well-defined values for a_i and b_i. In this case, we write

$$C_{\text{h.v.}} = (a_1 + a_2)b_1 + (a_1 - a_2)b_2 \tag{9.9}$$

But since $a_i = \pm 1$, then there are two possibilities

- $a_1 + a_2 = 0 \;\Rightarrow\; a_1 - a_2 = \pm 2$
- $a_1 - a_2 = 0 \;\Rightarrow a_1 + a_2 = \pm 2$

In either case, $C_{\text{h.v.}} = \pm 2 b_i$ for some b_i. Since b_i can only take values ± 1, we have $|\langle b_i \rangle| \leq 1$, and so

$$-2 \leq \langle C_{\text{h.v.}} \rangle \leq 2 \tag{9.10}$$

This is the CHSH inequality. It is entirely analogous to the Bell inequality (9.6).

What about in quantum theory? Now we don't admit to a_1 and a_2 having simultaneous meaning, so we're not allowed to write (9.9). Instead, we have to manipulate (9.8) as an operator equation. Because the eigenvalues are ± 1, we must have $A_1^2 = A_2^2 = B_1^2 = B_2^2 = \mathbf{1}$, the identity operator. After a little algebra, we find

$$\begin{aligned} C^2 &= 4\mathbf{1} - A_1A_2B_1B_2 + A_2A_1B_1B_2 + A_1A_2B_2B_1 - A_2A_1B_2B_1 \\ &= 4\mathbf{1} - [A_1, A_2][B_1, B_2] \end{aligned}$$

Now $|\langle [A_1, A_2]\rangle| \leq |\langle A_1A_2\rangle| + |\langle A_2A_1\rangle| \leq 2$, since each operator has eigenvalue ± 1. From this we learn that in the quantum theory,

$$\langle C^2\rangle \leq 8$$

Since $\langle C^2\rangle \geq \langle C\rangle^2$, we find that the range of values in quantum mechanics to be

$$-2\sqrt{2} \leq \langle C\rangle \leq 2\sqrt{2}$$

This is referred to as the *Cirel'son bound*. Clearly the range of values allowed by quantum mechanics exceeds that allowed by hidden variables theories (9.10).

It remains for us to exhibit states and operators which violate the CHSH bound. For this, we can return to our spin model. From (9.4), we know that

$$\langle EPR|\,\sigma_z^A \otimes \sigma_\theta^B\,|EPR\rangle = \sin^2\frac{\theta}{2} - \cos^2\frac{\theta}{2} = -\cos\theta$$

This means that if we take the four operators A_2, B_1, A_1 and B_2 to be spin operators, aligned in the (x, y) at successive angles of $45°$. (i.e. A_2 has $\theta = 0$, B_1 has $\theta = \frac{\pi}{4}$, A_1 has $\theta = \frac{\pi}{2}$ and B_2 has $\theta = \frac{3\pi}{4}$) then

$$\langle A_1B_1\rangle = \langle A_1B_1\rangle = \langle A_1B_1\rangle = -\frac{1}{\sqrt{2}} \quad \text{and} \quad \langle A_2B_2\rangle = +\frac{1}{\sqrt{2}}$$

and we see that

$$\langle C\rangle = -2\sqrt{2}$$

saturating the Cirel'son bound.

9.1.4 Entanglement Between Three Particles

If we consider the case of three particles rather than two, then there is even sharper contradiction between the predictions of quantum mechanics and those of hidden variables theories. As before, we'll take each particle to carry one of two states, with a basis given by spins $|\uparrow\rangle$ and $|\downarrow\rangle$, measured in the z-direction.

Consider the entangled three-particle state

$$|GHZ\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\uparrow\rangle|\uparrow\rangle - |\downarrow\rangle|\downarrow\rangle|\downarrow\rangle)$$

named after Greenberger, Horne and Zeilinger. These three particles are sent to our three intrepid scientists, each waiting patiently in far-flung corners of the galaxy. Each of these scientists makes one of two measurements: they either measure the spin in the x-direction, or they measure the spin in the y-direction. Obviously, each experiment gives them the result $+1$ or -1.

The state $|GHZ\rangle$ will result in correlations between the different measurements. Suppose, for example, that two of the scientists measure σ_x and the other measures σ_y. It is simple to check that

$$\sigma_x^A \otimes \sigma_x^B \otimes \sigma_y^C|GHZ\rangle = \sigma_x^A \otimes \sigma_y^B \otimes \sigma_x^C|GHZ\rangle = \sigma_y^A \otimes \sigma_x^B \otimes \sigma_x^C|GHZ\rangle = +|GHZ\rangle$$

In other words, the product of the scientist's three measurements always equals $+1$.

It's tempting to follow the hidden variables paradigm and assign a spin s_x and s_y to each of these three particles. Let's suppose we do so. Then the result above means that

$$s_x^A s_x^B s_y^C = s_x^A s_y^B s_x^C = s_y^A s_x^B s_x^C = +1 \tag{9.11}$$

But from this knowledge we can make a simple prediction. If we multiply all of these results together, we get

$$(s_x^A s_x^B s_x^C)^2\, s_y^A s_y^B s_y^C = +1 \quad \Rightarrow \quad s_y^A s_y^B s_y^C = +1 \tag{9.12}$$

where the implication follows from the fact that the spin variables can only take values ± 1. The hidden variables tell us that whenever the correlations (9.11) hold, the correlation (9.12) must also hold.

Let's now look at what quantum mechanics tells us. Rather happily, $|GHZ\rangle$ happens to be an eigenstate of $\sigma_y^A \otimes \sigma_y^B \otimes \sigma_y^C$. But we have

$$\sigma_y^A \otimes \sigma_y^B \otimes \sigma_y^C\,|GHZ\rangle = -|GHZ\rangle$$

In other words, the product of these three measurements must give -1. This is in stark contradiction to the hidden variables result (9.12). Once again we see that local hidden variables are incapable of reproducing the results of quantum mechanics.

If Only We Hadn't Made Counterfactual Arguments...

In both the Bell and GHZ arguments, the mistake in assigning hidden variables can be traced to our use of *counterfactuals*. This is the idea that we can say what would have happened had we made different choices.

Suppose, for example, that Alice chooses to measure σ_z to be $+1$ in an EPR state. Then Bob can be absolutely certain that he will find σ_z to be -1 should he choose to measure it. But even that certainty doesn't give him the right to assign $s_z^B = -1$ unless he actually goes ahead and measures it. This is because he may want to measure spin along some other axis, σ_θ^B, and assuming that both properties exist will lead us to the wrong conclusion as we've seen above. The punchline is that you don't get to make counterfactual arguments based on what would have happened: only arguments based on what actually did happen.

9.1.5 The Kochen-Specker Theorem

The Kochen-Specker theorem provides yet another way to restrict putative hidden-variables theories. Here is the statement:

Consider a set of N Hermitian operators A_i acting on $\mathcal{H}$. Typically some of these operators will commute with each other, while others will not. Any subset of operators which mutually commute will be called *compatible*.

In an attempt to build a hidden variables theory, all observables A_i are assigned a value $a_i \in \mathbf{R}$. We will require that whenever A, B and $C \in \{A_i\}$ are compatible then the following properties should hold

- If $C = A + B$ then $c = a + b$.
- If $C = AB$ then $c = ab$.

These seem like sensible requirements. Indeed, in quantum mechanics we know that if $[A, B] = 0$ then the expectation values obey the relations above and, moreover, there are states where we can assign definite values to A, B and therefore to $A + B$ and to AB. We will not impose any such requirements if $[A, B] \neq 0$.

As innocuous as these requirements may seem, the Kochen-Specker theorem states that in Hilbert spaces $\mathcal{H}$ with dimension $\dim(\mathcal{H}) \geq 3$, there are sets of operators $\{A_i\}$ for which it is not possible to assign values a_i with these properties. Note that this isn't a statement about a specific state in the Hilbert space; it's a stronger statement that there is no consistent values that can possibly be assigned to operators.

The issue is that a given operator, say A, can be compatible with many different operators. So, for example, it may appear in the compatible set (A, B, C) and also in (A, D, E) and should take the same value a in both. Meanwhile, B may appear in a different compatible set and so on. The proofs of the Kochen-Specker theorem involve exhibiting a bunch of operators which cannot be consistently assigned values.

The original proof of the Kocken-Specker theorem is notoriously fiddly, involving a set of $N = 117$ different projection operators in a $\dim(\mathcal{H}) = 3$ dimensional Hilbert space[11]. Simpler versions of the proof with $\dim(\mathcal{H}) = 3$ now exist, although we won't present them here.

There is, however, a particularly straightforward proof that involves $N = 18$ operators in a $\dim(\mathcal{H}) = 4$ dimensional Hilbert space. We start by considering the following 18 vectors $\psi_i \in \mathbf{C}^4$,

$$\begin{array}{cccc} \psi_1 = (0,0,0,1) \ , & \psi_2 = (0,0,1,0) \ , & \psi_3 = (1,1,0,0) \ , & \psi_4 = (1,-1,0,0) \\ \psi_5 = (0,1,0,0) \ , & \psi_6 = (1,0,1,0) \ , & \psi_7 = (1,0,-1,0) \ , & \psi_8 = (1,-1,1,-1) \\ \psi_9 = (1,-1,-1,1) \ , & \psi_{10} = (0,0,1,1) \ , & \psi_{11} = (1,1,1,1) \ , & \psi_{12} = (0,1,0,-1) \\ \psi_{13} = (1,0,0,1) \ , & \psi_{14} = (1,0,0,-1) \ , & \psi_{15} = (0,1,-1,0) \ , & \psi_{16} = (1,1,-1,1) \\ & \psi_{17} = (1,1,1,-1) \ , & \psi_{18} = (-1,1,1,1) & \end{array}$$

From each of these, we can build a projection operator

$$P_i = \frac{|\psi_i\rangle\langle\psi_i|}{\langle\psi_i|\psi_i\rangle}$$

Since the projector operators can only take eigenvalues 0 or 1, we want to assign a value $p_i = 0$ or $p_i = 1$ to each projection operator P_i.

Of course, most of these projection operators do not commute with each other. However, there are subsets of four such operators which mutually commute and sum to give the identity operator. For example,

$$P_1 + P_2 + P_3 + P_4 = \mathbf{1}_4$$

In this case, the requirements of the Kocken-Specker theorem tell us that one of these operators must have value $p = 1$ and the other three must have value $p = 0$.

[11]More details can be found at https://plato.stanford.edu/entries/kochen-specker/.

Now comes the twist. We can, in fact, construct nine such subsets of four operators. These are listed in the columns of the following table:

P_1	P_1	P_8	P_8	P_2	P_9	P_{16}	P_{16}	P_{17}
P_2	P_5	P_9	P_{11}	P_5	P_{11}	P_{17}	P_{18}	P_{18}
P_3	P_6	P_3	P_7	P_{13}	P_{14}	P_4	P_6	P_{13}
P_4	P_7	P_{10}	P_{12}	P_{14}	P_{15}	P_{10}	P_{12}	P_{15}

This table has the nice property that each P_i appears in exactly two different columns. Now the task is clear: assign values $p_i = 0, 1$ to each P_i such that each column has a single $p = 1$ and three $p = 0$. It is best to sit down and try to do this. And then try again. By the time you've tried for the third time, it should be increasingly clear that no consistent assignment of values p_i is possible. This is the Kochen-Specker theorem.

9.2 Entanglement is a Resource

In the previous section, we used entangled states to reveal how quantum mechanics differs from our older, classical framework. In this section, we will view entanglement somewhat differently. It is a precious commodity that allows us to achieve things that classical physicists cannot.

9.2.1 The CHSH Game

To illustrate the advantage that entanglement brings, we start by describing a game. It's not a particularly fun game. It's designed purely as a point of principle to show that entanglement can be useful.

The game is one of cooperation between two players – Alice and Bob of course – who cannot communicate with each other, but can prepare a strategy beforehand. Alice and Bob are both given an envelope. Inside each envelope is either a red card or a blue card. This means that there are four possibilities for their cards: red/red, red/blue, blue/red or blue/blue.

After seeing their card, Alice and Bob have to decide whether to say "turtles" or to say "cucumber". This is, I think you will agree, a silly game. The rules are as follows:

- Alice and Bob win if both cards are red and they said different words.
- Alice and Bob win if at least one card was blue and they said the same word.
- Otherwise, they lose.

What's their best strategy? First suppose that Alice and Bob are classical losers and have no help from quantum mechanics. It's not hard to convince yourself that their best strategy is just to say "cucumber" every time, regardless of the colour of their card. They only lose if both cards turn out to be red. Otherwise they win. This means that they win 75% of the time.

Suppose, however, that Alice and Bob have spent many decades developing coherent qubits. This pioneering technology resulted in them being kidnapped by a rival government who then, for reasons hard to fathom, subjected them to this stupid game. Can their discoveries help them get out of a bind? Thankfully, the answer is yes. Although, arguably, not so much that it's worth all the trouble.

To do better, Alice and Bob must share a number of EPR pairs, one for each time that the game is played. Here is their gameplan. Whenever Alice's card is blue, she measures A_1; whenever it is red she measures A_2. Whenever these measurements give $+1$ she says "turtles"; whenever it is -1 she says "cucumber". Bob does something similar: B_1 when blue, B_2 when red; "turtles" when $+1$, "cucumber" when -1.

Suppose that both cards are blue. Then they win if A_1 and B_1 give the same result and lose otherwise. In other words, they win if the measurement gives $A_1B_1 = +1$ and lose when $A_1B_1 = -1$. This means

$$P(\text{win}) - P(\text{lose}) = \langle A_1B_1 \rangle$$

In contrast, if both cards are red then they lose if A_2 and B_2 give the same measurement and win otherwise, so that

$$P(\text{win}) - P(\text{lose}) = -\langle A_2B_2 \rangle$$

Since each combination of cards arises with probability $p = \frac{1}{4}$, the total probability is

$$P(\text{win}) - P(\text{lose}) = \frac{1}{4}\langle A_1B_1 + A_1B_2 + A_2B_1 - A_2B_2 \rangle$$

But we've seen this before: it's precisely the combination of operators (9.8) that arose in the CHSH proof of the Bell inequality. We can immediately import our answer from there to learn that

$$P(\text{win}) - P(\text{lose}) \leq \frac{1}{\sqrt{2}}$$

We saw previously that we can find operators which saturate this inequality. Since $P(\text{win}) + P(\text{lose}) = 1$, there's a choice of measurements A_i and B_i — essentially spin measurements which differ by 45° — which ensures a win rate of

$$P(\text{win}) = \frac{1}{2}\left(\frac{1}{\sqrt{2}} + 1\right) \approx 0.854$$

This beats our best classical strategy of 75%.

Having the ability to win at this particular game is unlikely to change the world. Obviously the game was cooked up by starting from the CHSH inequality and working backwards in an attempt to translate Bell's inequality into something approximating a game. But it does reveal an important point: the correlations in entangled states can be used to do things that wouldn't otherwise be possible. If we can harness this ability to perform tasks that we actually care about, then we might genuinely be able to change the world. This is the subject of quantum information. Here we give a couple of simple examples that move in this direction.

9.2.2 Dense Coding

For our first application, Alice wants to send Bob some classical information, which means she wants to tell him "yes" or "no" to a series of questions. This is encoded in a classical bit as values 0 and 1.

However, Alice is fancy. She has qubits at her disposal and can send these to Bob. We'd like to know if she can use this quantum technology to aid in sending her classical information.

First note that Alice doesn't lose anything by sending qubits rather than classical bits. (Apart, of course, from the hundreds of millions of dollars in R&D that it took to get them in the first place.) She could always encode the classical value 0 as $|\uparrow\rangle$ and 1 as $|\downarrow\rangle$ and, provided Bob is told in advance to measure σ_z, the qubit contains the same information as a classical bit. But this does seem like a waste of resources.

Is it possible to do better and transmit more than one classical bit in a single qubit? The answer is no: a single qubit carries the same amount of information as a classical bit. However, this changes if Alice's qubit is actually part of an entangled pair that she shares with Bob. In this case, she can encode two classical bits of information in a single qubit. This is known as *dense coding.*

To achieve this feat, Alice first performs an operation on her spin. We'll introduce some new notation for this state that will become useful in the following section: we call the EPR pair

$$|EPR\rangle = |\chi^-\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle)$$

Alice then has four options:

- She does nothing. Obviously, the entangled pair remains in the state $|\chi^-\rangle$.
- Alice acts with σ_x. This changes the state to $-|\phi^-\rangle$ where

$$|\phi^-\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\uparrow\rangle - |\downarrow\rangle|\downarrow\rangle)$$

- Alice acts with σ_y. This changes the state to $-i|\phi^+\rangle$ where

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\uparrow\rangle + |\downarrow\rangle|\downarrow\rangle)$$

- Alice acts with σ_z. This changes the state to $|\chi^+\rangle$.

$$|\chi^+\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle + |\downarrow\rangle|\uparrow\rangle)$$

The upshot of this procedure is that the entangled pair sits in one of four different states

$$|\phi^\pm\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\uparrow\rangle \pm |\downarrow\rangle|\downarrow\rangle) \quad \text{or} \quad |\chi^\pm\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle \pm |\downarrow\rangle|\uparrow\rangle) \qquad (9.13)$$

Alice now sends her qubit to Bob, so Bob has access to the whole entangled state. Since the four different states are orthogonal, it must be possible to distinguish them by performing some measurements. Indeed, the measurements Bob needs to make are

$$\sigma_x \otimes \sigma_x \quad \text{and} \quad \sigma_z \otimes \sigma_z$$

These two operators commute. This means that, while we don't get to know the values of both s_x and s_z of, say, the first spin, it does make sense to talk about the products of the spins of the two qubits in both directions. It's simple to check that the four possible states above are eigenstates of these two operators

$$\sigma_x \otimes \sigma_x|\phi^\pm\rangle = \pm|\phi^\pm\rangle \quad \text{and} \quad \sigma_x \otimes \sigma_x|\chi^\pm\rangle = \pm|\chi^\pm\rangle \qquad (9.14)$$

$$\sigma_z \otimes \sigma_z|\phi^\pm\rangle = +|\phi^\pm\rangle \quad \text{and} \quad \sigma_z \otimes \sigma_z|\chi^\pm\rangle = -|\chi^\pm\rangle$$

So, for example, if Bob measures $\sigma_x \otimes \sigma_x = +1$ and $\sigma_z \otimes \sigma_z = -1$, then he knows that he's in possession of state $|\chi^+\rangle$. Bob then knows which of the four operations Alice performed. In this way she has communicated two classical bits of information through the exchange of a single qubit.

Admittedly, two qubits were needed for this to fly: one which was exchanged and one which was in Bob's possession all along. In fact, in Section 9.3.2, we'll show that entanglement between spins can only be created if the two spins were brought together at some point in the past. So, from this point of view, Alice actually exchanged two qubits with Bob, the first long ago when they shared the EPR pair, and the second when the message was sent. Nonetheless, there's still something surprising about dense coding. The original EPR pair contained no hint of the message that Alice wanted to send; indeed, it could have been created long before she knew what that message was. Nor was there any information in the single qubit that Alice sent to Bob. Anyone intercepting it along the way would be no wiser. It's only when this qubit is brought together with Bob's that the information becomes accessible.

9.2.3 Quantum Teleportation

Our next application has a sexy sounding name: quantum teleportation. To put it in context, we first need a result that tells us what we cannot do in quantum mechanics.

The No Cloning Theorem

The no cloning theorem says that it is impossible to copy a state in quantum mechanics.

Here's the game. Someone gives you a state $|\psi\rangle$, but doesn't tell you what that state is. Now, you can determine some property of the state but any measurement that you make will alter the state. This means that you can't then go back and ask different questions about the initial state.

Our inability to know everything about a state is one of the key tenets of quantum mechanics. But there's an obvious way around it. Suppose that we could just copy the initial state many times. Then we could ask different questions on each of the replicas and, in this way, build up a fuller picture of the original state. The no cloning theorem forbids this.

To prove the theorem, we really only need to set up the question. We start with a state $|\psi\rangle \in \mathcal{H}_A$. Suppose that we prepare a separate system in a blank state $|0\rangle \in \mathcal{H}_B$. To create a copy of the initial state, we would like to evolve the system so that

$$|\text{In}(\psi)\rangle = |\psi\rangle \otimes |0\rangle \longrightarrow |\text{Out}(\psi)\rangle = |\psi\rangle \otimes |\psi\rangle$$

But this can't happen through any Hamiltonian evolution because it is not a unitary operation. To see this, consider two different states $|\psi_1\rangle$ and $|\psi_2\rangle$. We have

$$\langle \text{In}(\psi_1)|\text{In}(\psi_2)\rangle = \langle \psi_1|\psi_2\rangle \quad \text{while} \quad \langle \text{Out}(\psi_1)|\text{Out}(\psi_2)\rangle = \langle \psi_1|\psi_2\rangle^2$$

We might try to wriggle out of this conclusion by allowing for some other stuff in the Hilbert space which can change in any way it likes. This means that we now have three Hilbert spaces and are looking an evolution of the form

$$|\psi\rangle \otimes |0\rangle \otimes |\alpha(0)\rangle \longrightarrow |\psi\rangle \otimes |\psi\rangle \otimes |\alpha(\psi)\rangle$$

By linearity, if such an evolution exists it must map

$$(|\phi\rangle + |\psi\rangle) \otimes |0\rangle \otimes |\alpha(0)\rangle \longrightarrow |\phi\rangle \otimes |\phi\rangle \otimes |\alpha(\phi)\rangle + |\psi\rangle \otimes |\psi\rangle \otimes |\alpha(\psi)\rangle \tag{9.15}$$

But this isn't what we wanted! The map is supposed to take

$$\begin{aligned}(|\phi\rangle + |\psi\rangle) \otimes |0\rangle \otimes |\alpha(0)\rangle &\longrightarrow (|\phi\rangle + |\psi\rangle) \otimes (|\phi\rangle + |\psi\rangle) \otimes |\alpha(\psi+\phi)\rangle \\ &= \; (|\phi\rangle|\phi\rangle + |\psi\rangle|\phi\rangle + |\phi\rangle|\psi\rangle + |\psi\rangle|\psi\rangle) \otimes |\alpha(\psi+\phi)\rangle\end{aligned}$$

where, in the last line, we dropped the $\otimes$ between the first two Hilbert spaces. The state that we get (9.15) is not the state that we want. This concludes our proof of the no cloning theorem.

Back to Teleportation

With the no cloning theorem as background, we can now turn to the idea of quantum teleportation. Alice is given a qubit in state $|\psi\rangle$. The challenge is to communicate this state to Bob.

There are two limitations. First, Alice doesn't get to simply put the qubit in the mail. That's no longer the game. Instead, she must describe the qubit to Bob using classical information: i.e. bits, not qubits. Note that we're now playing by different rules from the previous section. In "dense coding" we wanted to send classical information using qubits. Here we want to send quantum information using classical bits.

Now this sounds like teleportation must be impossible. As we've seen, Alice has no way of figuring out what state $|\psi\rangle$ she has. If she doesn't know the state, how on earth is she going to communicate it to Bob? Well, magically, there is way. For this to work, Alice and Bob must also share an EPR pair. We will see that they can sacrifice the entanglement in this EPR pair to allow Bob to reproduce the state $|\psi\rangle$.

First, Alice. She has two qubits: the one we want to transfer, $|\psi\rangle$, together with the her half of the pair $|EPR\rangle$. She makes the following measurements:

$$\sigma_x \otimes \sigma_x \quad \text{and} \quad \sigma_z \otimes \sigma_z$$

where, in each case, the first operator acts on $|\psi\rangle$ and the second on her half of $|EPR\rangle$.

As we saw in the previous section, these are commuting operators, each with eigenvalues ± 1. This means that there are four different outcomes to Alice's experiment and the state will be projected onto the eigenstates $|\phi^\pm\rangle$ or $|\chi^\pm\rangle$ defined in (9.13). The different possible outcomes of the measurement were given in (9.14).

Let's see what becomes of the full state after Alice's measurements. We write the unknown qubit $|\psi\rangle$ as

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle \tag{9.16}$$

with $|\alpha|^2 + |\beta|^2 = 1$. Then the full state of three qubits – two owned by Alice and one by Bob – is

$$\begin{aligned}
|\psi\rangle \otimes |EPR\rangle &= \frac{1}{\sqrt{2}}\Big(\alpha|\uparrow\rangle|\uparrow\rangle|\downarrow\rangle - \alpha|\uparrow\rangle|\downarrow\rangle|\uparrow\rangle + \beta|\downarrow\rangle|\uparrow\rangle|\downarrow\rangle - \beta|\downarrow\rangle|\downarrow\rangle|\uparrow\rangle\Big) \\
&= \frac{1}{2}\Big(\alpha(|\phi^+\rangle + |\phi^-\rangle)|\downarrow\rangle - \alpha(|\chi^+\rangle + |\chi^-\rangle)|\uparrow\rangle \\
&\qquad +\beta(|\chi^+\rangle - |\chi^-\rangle)|\downarrow\rangle - \beta(|\phi^+\rangle - |\phi^-\rangle)|\uparrow\rangle\Big) \\
&= \frac{1}{2}\Big(|\phi^+\rangle(-\beta|\uparrow\rangle + \alpha|\downarrow\rangle) + |\phi^-\rangle(\beta|\uparrow\rangle + \alpha|\downarrow\rangle) \\
&\qquad +|\chi^+\rangle(-\alpha|\uparrow\rangle + \beta|\downarrow\rangle) - |\chi^-\rangle(\alpha|\uparrow\rangle + \beta|\downarrow\rangle)\Big)
\end{aligned}$$

When Alice makes her measurement, the wavefunction collapses onto one of the four eigenstates $|\phi^\pm\rangle$ or $|\chi^\pm\rangle$. But we see that Bob's state — the final one in the wavefunction above — has taken the form of a linear superposition of $|\uparrow\rangle$ and $|\downarrow\rangle$, with the same coefficients α and β that characterised the initial state $|\psi\rangle$ in (9.16). Now, in most of these cases, Bob's state isn't exactly the same as $|\psi\rangle$, but that's easily fixed if Bob acts with a unitary operator. All Alice has to do is tell Bob which of the four states she measured and this will be sufficient for Bob to know how he has to act. Let's look at each in turn.

- If Alice measures $|\phi^+\rangle$ then Bob should operate on his qubit with σ_y to get

  $$\sigma_y(-\beta|\uparrow\rangle + \alpha|\downarrow\rangle) = i\beta|\downarrow\rangle + i\alpha|\uparrow\rangle = i|\psi\rangle$$

  which, up to a known phase, is Alice's initial state.

- If Alice measures $|\phi^-\rangle$ then Bob should operate on his qubit with σ_x,

  $$\sigma_x(\beta|\uparrow\rangle + \alpha|\downarrow\rangle) = \beta|\downarrow\rangle + \alpha|\uparrow\rangle = |\psi\rangle$$

- If Alice measures $|\chi^+\rangle$ then Bob should operate on his qubit with σ_z,

$$\sigma_x(\beta|\uparrow\rangle + \alpha|\downarrow\rangle) = \beta|\downarrow\rangle + \alpha|\uparrow\rangle = |\psi\rangle$$

- If Alice measures $|\phi^+\rangle$, Bob can put his feet up and do nothing. He already has $-|\psi\rangle$ sitting in front of him.

We see that if Alice sends Bob two bits of information — enough to specify which of the four states she measured — then Bob can ensure that he gets state $|\psi\rangle$. Note that this transfer occurred with neither Alice nor Bob knowing what the state $|\psi\rangle$ actually is. But Bob can be sure that he has it.

9.2.4 Quantum Key Distribution

If you want to share a secret, it's best to have a code. Here is an example of an unbreakable code. Alice and Bob want to send a message consisting of n classical bits, a string of 0's and 1's. To do so securely, they must share, in advance, a *private key*. This is a string of classical bits that is the same length as the message. Alice simply adds the key to the message bitwise ($0 + 0 = 1 + 1 = 0$ and $0 + 1 = 1 + 0 = 1$) before sending it to Bob who, upon receiving it, subtracts the key to reveal the message. Any third party eavesdropper – traditionally called Eve – who intercepts the transmission is none the wiser.

The weakness of this approach is that, to be totally secure, Alice and Bob, should use a different key for each message that they want to send. If they fail to do this then Eve can use some knowledge about the underlying message (e.g. it's actually written in German and contains information about U-boat movements in the Atlantic) to detect correlations in the transmissions and, ultimately, crack the code. This means that Alice and Bob must have a large supply of private keys and be sure that Eve does not have access to them. This is where quantum mechanics can be useful.

BB84

BB84 is a quantum protocol for generating a secure private key. It's named after its inventors, Bennett and Brassard, who suggested this approach in 1984.

The idea is remarkably simple. Alice takes a series of qubits. For each, she chooses to measure the spin either in the z-direction, or in the x-direction. This leaves her with a qubit in one of four possible states: $|\uparrow\rangle$, $|\downarrow\rangle$, $|\rightarrow\rangle$ or $|\leftarrow\rangle$. Alice then sends this qubit to Bob. He has no idea which measurement Alice made, so he makes a random decision to measure the spin in either the z-direction or the x-direction. About half the time he will make the same measurement as Alice, the other half he will make a different measurement.

Having performed these experiments, Alice and Bob then announce publicly which spin measurements they made. Whenever they measured the spin in different directions, they simply discard their results. Whenever they measured the spin in the same direction, the measurements must agree. This becomes their private key.

The whole purpose of generating a private key is that it must be private. For example, the keys for the enigma machine — as shown in the picture — were sent out monthly. If you were lucky enough to capture this book, you could break the codes for the next month. How can Alice and Bob be certain that their key hasn't been intercepted by Eve?

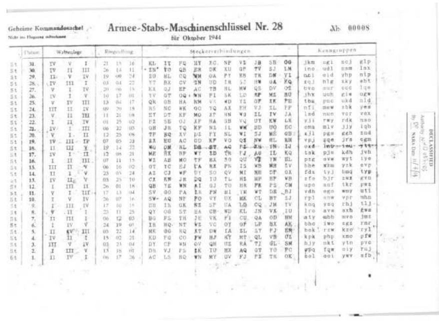
Armee-Stabs-Maschinenschlüssel Nr. 28
für Oktober 1944

Figure 88:

This is where the laws of quantum physics come to the rescue. First, the no-cloning theorem ensures that Eve has no way of copying the qubit if she intercepts it. Nor does she have any way of determining its state. Even if she knows the game that Alice and Bob are playing, the best that she can do is to measure the spin in either the z-direction or the x-direction, before sending it on to Bob. Half the time, she will make the same measurement as Alice and leave the state unchanged. But the other half, she will change the state and so change the possible results that Bob finds in his measurements. To guard against this possibility, Alice and Bob can simply choose to publicly announce a subset of the results of their correlated measurements. If they don't perfectly agree, then they know that someone has tampered with the transmission.

The BB84 protocol doesn't make any use of quantum entanglement. There is, however, a minor variation where entanglement plays a role. In this scenario, Alice prepares a succession of entangled pairs in, say, the state

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\uparrow\rangle + |\downarrow\rangle|\downarrow\rangle)$$

She then sends the second spin to Bob. When the two of them both have their spins, they can follow the BB84 rules to generate the key. The slight advantage of this approach is that Alice doesn't have to record her measurements before sending them to Bob. This protects her from the possibility that someone breaks into her lab and takes sneaky photographs of her measurement results. Of course, one might wonder if the extra resources involved in generating coherent entangled states might not be put to better use in, for example, buying a decent safe.

The moral behind quantum key distribution is clear: quantum information is more secure than classical information because no one, whether friend or enemy, can be sure what quantum state they've been given.

9.3 Density Matrices

In Section 9.1, we've made a big deal out the fact that quantum correlations cannot be captured by classical probability distributions. In the classical world, uncertainty is due to ignorance: the more you know, the better your predictions. In the quantum world, the uncertainty is inherent and can't be eliminated by gaining more knowledge.

There are situations in the quantum world where we have to deal with both kinds of uncertainties. There are at least two contexts in which this arises. One possibility is ignorance: we simply don't know for sure what quantum state our system lies in. Another possibility is that we have many quantum states — an ensemble — and they don't all lie in the same state, but rather in a mixture of different states. In either context, we use the same mathematical formalism.

Suppose that we don't know which of the states $|\psi_i\rangle$ describes our system. These states need not be orthogonal – just different. To parameterise our ignorance, we assign classical probabilities p_i to each of these states. The expectation value of any operator A is given by

$$\langle A\rangle = \sum_i p_i \langle\psi_i|A|\psi_i\rangle \tag{9.17}$$

This expression includes both classical uncertainty (in the p_i) and quantum uncertainty (in the $\langle\psi_i|A|\psi_i\rangle$).

Such a state is described by an operator known as the *density matrix*.

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \tag{9.18}$$

Clearly, this is a sum of projections onto the spaces spanned by $|\psi_i\rangle$, weighted with the probabilities p_i. The expectation value (9.17) of any operator can now be written simply as

$$\langle A\rangle = \mathrm{Tr}(\rho A)$$

where the trace is over all states in the Hilbert space.

Pure States vs Mixed States

Previously, we thought that the state of a quantum system is described by a normalised vector in the Hilbert space. The density matrix is a generalisation of this idea to incorporate classical probabilities. If we're back in the previous situation, where we know for sure that the system is described by a specific state $|\psi\rangle$, then the density matrix is simply the projection operator

$$\rho = |\psi\rangle\langle\psi|$$

In this case, we say that we have a *pure state*. If the density matrix cannot be written in this form then we say that we have a *mixed state*. Note that a pure state has the property that

$$\rho^2 = \rho$$

Regardless of whether a state is pure or mixed, the density matrix encodes all our information about the state and allows us to compute the expected outcome of any measurement. Note that the density matrix does not contain information about the phases of the states $|\psi_i\rangle$ since these have no bearing on any physical measurement.

Properties of the Density Matrix

The density matrix (9.18) has the following properties

- It is self-adjoint: $\rho = \rho^\dagger$
- It has unit trace: $\mathrm{Tr}\rho = 1$. This property is equivalent to the normalisation of a probability distribution, so that $\sum_i p_i = 1$.
- It is positive: $\langle\phi|\rho|\phi\rangle \geq 0$ for all $|\phi\rangle \in \mathcal{H}$. This property, which strictly speaking should be called "non-negative", is equivalent to the requirement that $p_i \geq 0$. As shorthand, we sometimes write the positivity requirement simply as $\rho \geq 0$.

Furthermore, any operator ρ which satisfies these three properties can be viewed as a density matrix for a quantum system. To see this, we can look at the eigenvectors of ρ, given by

$$\rho|\phi_n\rangle = p_n|\phi_n\rangle$$

where, here, p_n is simply the corresponding eigenvalue. Because $\rho = \rho^\dagger$, we know that $p_n \in \mathbf{R}$. The second two properties above then tell us that $\sum_n p_n = 1$ and $p_n \geq 0$.

This is all we need to interpret p_n as a probability distribution. We can then write ρ as

$$\rho = \sum_n p_n |\phi_n\rangle\langle\phi_n| \tag{9.19}$$

This way of writing the density matrix is a special case of (9.18). It's special because the $|\phi_n\rangle$ are eigenvectors of a Hermitian matrix and, hence, orthogonal. In contrast, the vector $|\psi_i\rangle$ in (9.18) are not necessarily orthonormal. However, although the expression (9.19) is special, there's nothing special about ρ itself: any density matrix can be written in this form. We'll come back to this idea below when we discuss specific examples.

An Example: Statistical Mechanics

There are many places in physics where it pays to think of probability distributions over ensembles of states. One prominent example is what happens for systems at finite temperature T. This is the subject of Statistical Mechanics.

Recall that the Boltzmann distribution tells us that the probability p_n that we sit in an energy eigenstate $|n\rangle$ is given by

$$p_n = \frac{e^{-\beta E_n}}{Z} \quad \text{ where } \beta = \frac{1}{k_B T} \text{ and } Z = \sum_n e^{-\beta E_n}$$

where k_B is the Boltzmann constant. It is straightforward to construct an density matrix corresponding to this ensemble. It is given by

$$\rho = \frac{e^{-\beta H}}{Z} \tag{9.20}$$

where H is the Hamiltonian. Similarly, the partition function is given by

$$Z = \operatorname{Tr} e^{-\beta H}$$

It is then straightforward to reformulate much of statistical mechanics in this language. For example, the average energy of a system is $\langle E\rangle = \operatorname{Tr}(\rho H)$.

In these lectures, we won't necessarily be interested in the kind of macroscopic systems that arise in statistical physics. Instead, we'll build some rather different intuition for the meaning of the density matrix.

Time Evolution

Recall that in the Schrödinger picture, any state evolves as

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle \quad \text{with} \quad U(t) = e^{-iHt/\hbar}$$

From this we learn that the density matrix evolves as

$$\rho(t) = U(t)\rho(0)U^\dagger(t)$$

Differentiating with respect to t gives us a differential equation governing time evolution,

$$\frac{\partial\rho}{\partial t} = -\frac{i}{\hbar}[H,\rho] \tag{9.21}$$

This is the *Liouville equation.* Or, more accurately, it is the quantum version of the Liouville equation which we met in the Classical Dynamics lectures where it governs the evolution of probability distributions on phase space.

Note that any density operator which depends only on the Hamiltonian H is independent of time. The Boltzmann distribution (9.20) is the prime example.

9.3.1 The Bloch Sphere

As an example, let's return to our favourite two-state system. If we measure spin along the z-axis, then the two eigenstates are $|\uparrow\rangle$ and $|\downarrow\rangle$.

Suppose that we know for sure that we're in state $|\uparrow\rangle$. Then, obviously,

$$\rho = |\uparrow\rangle\langle\uparrow|$$

If however, there's probability $p = \frac{1}{2}$ that we're in state $|\uparrow\rangle$ and, correspondingly, probability $1 - p = \frac{1}{2}$ that we're in state $|\downarrow\rangle$, then

$$\rho = \frac{1}{2}|\uparrow\rangle\langle\uparrow| + \frac{1}{2}|\downarrow\rangle\langle\downarrow| = \frac{1}{2}\mathbf{1} \tag{9.22}$$

This is the state of maximum ignorance, something we will quantify below in Section 9.3.3. In particular, the average value for the spin along any axis always vanishes: $\langle\boldsymbol{\sigma}\rangle = \mathrm{Tr}(\rho\boldsymbol{\sigma}) = 0$.

Let's now consider other spin states. Consider the spin measured along the x-axis. Suppose that there's probability $p = \frac{1}{2}$ that we're in state $|\rightarrow\rangle$ and probability $1-p = \frac{1}{2}$ that we're in state $|\leftarrow\rangle$, then

$$\rho = \frac{1}{2}\Big[|\rightarrow\rangle\langle\rightarrow| + |\leftarrow\rangle\langle\leftarrow|\Big] = \frac{1}{2}\mathbf{1} \tag{9.23}$$

Once again, we find a state of maximum ignorance. This highlights an important fact: given a density matrix ρ, there is no unique way to decompose in the form (9.18).

As a final example, there is nothing to stop us taking an ensemble of non-orthogonal states. So we could be in state $|\uparrow\rangle$ with probability $p = \frac{1}{2}$ and in state $|\rightarrow\rangle$ with probability $p = \frac{1}{2}$. The resulting density matrix is

$$\begin{aligned}\rho &= \frac{1}{2}|\uparrow\rangle\langle\uparrow| + \frac{1}{2}|\rightarrow\rangle\langle\rightarrow| \\ &= \frac{1}{2}|\uparrow\rangle\langle\uparrow| + \frac{1}{4}(|\uparrow\rangle + |\downarrow\rangle)(\langle\uparrow| + \langle\downarrow|) \\ &= \frac{1}{4}\mathbf{1} + \frac{1}{2}|\uparrow\rangle\langle\uparrow| + \frac{1}{4}|\uparrow\rangle\langle\downarrow| + \frac{1}{4}|\downarrow\rangle\langle\uparrow|\end{aligned}$$

We haven't written this density matrix in the form (9.19), although its not difficult to do so. Nonetheless, it's simple to check that it obeys the three conditions above. We find $\langle\sigma^1\rangle = \langle\sigma^3\rangle = 1/2$ and $\langle\sigma^2\rangle = 0$.

Let's now look at the most general density matrix for a two-state system. The most general Hermitian 2×2 matrix can be expanded in terms of $\mathbf{1}$ and the Pauli matrices σ^i. Since $\text{Tr}\mathbf{1} = 2$ and $\text{Tr}\sigma^i = 0$, the requirement that $\text{Tr}\rho = 1$ means that we can write

$$\rho = \frac{1}{2}\left(\mathbf{1} + \mathbf{a}\cdot\boldsymbol{\sigma}\right) \tag{9.24}$$

for some 3-vector $\mathbf{a}$. All that's left is to require that this matrix has positive eigenvalues. The sum of the two eigenvalues is given by $\text{Tr}\rho = 1$, so at least one of them must be positive. The product of the eigenvalues is given by $\det\rho$. It's simple to compute

$$\det\rho = \frac{1}{4}(1 - \mathbf{a}\cdot\mathbf{a})$$

The two eigenvalues are both non-negative if $\det\rho \geq 0$. We learn that (9.24) defines a density matrix for a two-state system if

$$|\mathbf{a}| \leq 1$$

This is the interior of a 3-sphere which should be called the *Bloch Ball*. Unfortunately the names are a little mixed-up and this interior is sometimes referred to as the *Bloch Sphere*. The interior of the ball, with $|\mathbf{a}| < 1$, describes mixed states. The surface of the ball with $|\mathbf{a}| = 1$ — which should really be called the Bloch Sphere — describes pure states.

For both mixed and pure states, the direction $\mathbf{a}$ is referred to as the *polarisation* of the spin. For $\mathbf{a} \neq 0$, there will be a preference for the measurements of spin in the direction $\mathbf{a}\cdot\boldsymbol{\sigma}$. In contrast, when $\mathbf{a} = 0$, the state is said to be unpolarised. We met two examples of this above.

The Ambiguity of Preparation

The are typically many different interpretations of a density matrix. We've seen an example above, where two different probability distributions over states (9.22) and (9.23) both give rise to the same density matrix. It's sometimes said that these density matrices are *prepared* differently, but describe the same state.

More generally, suppose that the system is described by density matrix ρ_1 with some probability λ and density matrix ρ_2 with some probability $(1-\lambda)$. The expectation value of any operator is determined by the density matrix

$$\rho(\lambda) = \lambda\rho_1 + (1-\lambda)\rho_2$$

Indeed, nearly all density operators can be expressed as the sum of other density operators in an infinite number of different ways.

There is an exception to this. If the density matrix ρ actually describes a pure state then it cannot be expressed as the sum of two other states.

9.3.2 Entanglement Revisited

The density matrix has a close connection to the ideas of entanglement that we met in earlier sections. Suppose that our Hilbert space decomposes into two subspaces,

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$$

This is sometimes referred to as a *bipartite decomposition* of the Hilbert space. It really means that $\mathcal{H}_A$ and $\mathcal{H}_B$ describe two different physical systems. In what follows, it will be useful to think of these systems as far separated, so that they don't interact with each other. Nonetheless, as we've seen in Section 9.1, quantum states can be entangled between these two systems, giving rise to correlations between measurements.

Let's consider things from Alice's perspective. She only has access to the system described by $\mathcal{H}_A$. This means that she gets to perform measurements associated to operators of the form

$$\mathcal{O} = A \otimes \mathbf{1}$$

If the state of the full system is described by the density matrix ρ_{AB}, then measurements Alice makes will have expectation value

$$\langle A \rangle = \mathrm{Tr}_{\mathcal{H}_A}\mathrm{Tr}_{\mathcal{H}_B}\Big((A \otimes \mathbf{1})\rho_{AB}\Big) \equiv \mathrm{Tr}_{\mathcal{H}_A}(A\,\rho_A)$$

where we've defined

$$\rho_A = \mathrm{Tr}_{\mathcal{H}_B}\, \rho_{AB}$$

This is called the *reduced density matrix.* It is related to the full density matrix by taking the *partial trace* over the Hilbert space $\mathcal{H}_B$. We see that, from Alice's perspective, the part of the system that she has access to is described by the density matrix ρ_A.

Suppose that the full system ρ_{AB} lies in a pure state. This means that it takes the form

$$|\Psi\rangle = \sum_{i,j} \alpha_{ij} |\phi_i\rangle \otimes |\tilde{\phi}_j\rangle \tag{9.25}$$

where we've introduced a basis $|\phi_i\rangle$ for $\mathcal{H}_A$ and $|\tilde{\phi}_j\rangle$ for $\mathcal{H}_B$. (These two Hilbert spaces need not have the same dimension.). Note that, in general, this is an example of an entangled state.

The density matrix for the full system is

$$\rho_{AB} = |\Psi\rangle\langle\Psi| = \sum_{i,j,k,l} \alpha_{ij}\alpha^{\star}_{kl}\, |\phi_i\rangle \otimes |\tilde{\phi}_j\rangle\langle\phi_k| \otimes \langle\tilde{\phi}_l|$$

Taking the partial trace then gives the reduced density matrix

$$\rho_A = \sum_{ik} \beta_{ik} |\phi_i\rangle\langle\phi_k| \qquad \text{with } \beta_{ik} = \sum_j \alpha_{ij}\alpha^{\star}_{kj}$$

But this is the density matrix for a mixed state. This means that even if the full system is in a pure state, as far Alice is concerned it effectively lies in a mixed state. This illustrates how the probabilities p_i can arise from our lack of knowledge of other parts of the system. However, the presence of entanglement in the original state means that even ignorance about physics in far flung places forces us to deal with a mixed state.

In fact, this approach allows us define entanglement between two subsystems, something that we avoided doing in the previous sections. The state $|\Psi\rangle$ is said to be *entangled* only if the reduced density matrix $\rho_A = \mathrm{Tr}_{\mathcal{H}_B}|\Psi\rangle\langle\Psi|$ describes a mixed state. Otherwise $|\Psi\rangle$ is said to be *separable.* We will quantify the amount of entanglement in a system in Section 9.3.3 using the concept of entropy.

EPR Pairs Revisisted

Let's return to our favourite example of entanglement between two qubits. The EPR state is

$$|EPR\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$$

where, in an attempt to stop us going boggle-eyed in later equations, we're using notation such that $|\uparrow\rangle|\uparrow\rangle \equiv |\uparrow\uparrow\rangle$. The associated density matrix is

$$\rho_{EPR} = \frac{1}{2}\Big(|\uparrow\downarrow\rangle\langle\uparrow\downarrow| + |\downarrow\uparrow\rangle\langle\downarrow\uparrow| - |\uparrow\downarrow\rangle\langle\downarrow\uparrow| - |\downarrow\uparrow\rangle\langle\uparrow\downarrow|\Big) \tag{9.26}$$

We now take the trace over Bob's spin to get the reduced density matrix for Alice,

$$\rho_A = \mathrm{Tr}_{\mathcal{H}_B}\rho_{EPR} = \frac{1}{2}\Big(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|\Big) = \frac{1}{2}\mathbf{1} \tag{9.27}$$

Everything that Alice can measure on her own is captured in ρ_A, which is the state of maximum ignorance. We see that although the total density matrix knows about the correlations, there's no way that Alice can know about this on her own.

To illustrate this, suppose that Bob performs a measurement on his spin. This projects the EPR pair into state $|\uparrow\downarrow\rangle$ with probability $p = \frac{1}{2}$ and into state $|\downarrow\uparrow\rangle$ with probability $p = \frac{1}{2}$. Bob, of course, knows which of these states the system has collapsed to. However, if we don't know the outcome of this measurement then we should describe the system in terms of the mixed state

$$\rho_{\rm mixed} = \frac{1}{2}|\uparrow\downarrow\rangle\langle\uparrow\downarrow| + \frac{1}{2}|\downarrow\uparrow\rangle\langle\downarrow\uparrow|$$

This differs from the EPR density matrix (9.26). However, if we take the trace over Bob's degrees of freedom then we find that Alice's reduced density matrix ρ_A is once again given by (9.27). This is the statement that nothing changes for Alice when Bob performs a measurement. We can also repeat this exercise when Bob performs a measurement in a different spin direction. Once again, we find that ρ_A is given by (9.27). All of this is telling us something that we already knew: we cannot use the non-local correlations inherent in quantum states to transmit information in a non-local fashion.

Schmidt Decomposition

Consider a pure state $|\Psi\rangle$ in $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Given a set of basis $|\phi_i\rangle$ and $|\tilde{\phi}_j\rangle$, we can always decompose the state in the form (9.25). Moreover, it turns out that there is a preferred choice of basis states. The resulting expression is known as the *Schmidt decomposition*.

First, let's define a canonical basis for $\mathcal{H}_A$. As we've seen above, we can take the partial trace over $\mathcal{H}_B$ to find derive the reduced density matrix ρ_A. We'll choose $|\phi_i\rangle$ to be the eigenvectors of ρ_A, as in (9.19). We can then write

$$\rho_A = \sum_i p_i |\phi_i\rangle\langle\phi_i| \tag{9.28}$$

Our next task is to construct a suitable basis for $\mathcal{H}_B$. We could, of course, choose the basis of ρ_B and, in fact, ultimately this is what we'll end up doing. But in order to illustrate a rather nice property of this decomposition, we'll get there in a slightly roundabout way. Given a decomposition of the form (9.25), we define the vectors

$$|\chi_i\rangle = \sum_i \alpha_{ij} |\tilde{\phi}_j\rangle \;\in\; \mathcal{H}_B$$

Note that nothing guarantees that the vectors $|\chi_i\rangle$ normalised, and nothing guarantees that they are orthogonal. For now, their only purpose is to allow us to write the state (9.25) as

$$|\Psi\rangle = \sum_i |\phi_i\rangle \otimes |\chi_i\rangle$$

Now let's compute ρ_A from this state. We have

$$\rho_A = \sum_{i,j} \mathrm{Tr}_{\,\mathcal{H}_B} |\phi_i\rangle \otimes |\chi_i\rangle\langle\phi_j| \otimes \langle\chi_j| = \sum_{i,j} \langle\chi_i|\chi_j\rangle\; |\phi_i\rangle\langle\phi_j|$$

But we know that this reduced density matrix takes the form (9.28). This means that the overlap of the $|\chi_i\rangle$ vectors must be

$$\langle\chi_i|\chi_j\rangle = p_i\delta_{ij}$$

We learn that these vectors aren't normalised but, perhaps surprisingly, they are orthogonal. It's then straightforward to define a basis of vectors by

$$|\tilde{\chi}_i\rangle = \frac{1}{\sqrt{p_i}}|\chi_i\rangle$$

Only those vectors with $p_i \neq 0$ actually appear so we don't have to worry about about dividing by zero here. The upshot of this is that we can write any pure bipartite state in the canonical decomposition

$$|\Psi\rangle = \sum_i \sqrt{p_i}\, |\phi_i\rangle \otimes |\tilde{\chi}_i\rangle \tag{9.29}$$

This is the *Schmidt decomposition*. Note that there is a nice symmetry between the reduced density matrices ρ_A and ρ_B. They are, respectively,

$$\rho_A = \sum_i p_i |\phi_i\rangle\langle\phi_i| \quad , \quad \rho_B = \sum_i p_i |\tilde{\chi}_i\rangle\langle\tilde{\chi}_i|$$

We see that the basis $|\tilde{\chi}_i\rangle$ are the eigenvectors of ρ_B, even though this wasn't how we initially constructed them. Further, the probabilities p_i are eigenvalues of both ρ_A and ρ_B. In particular if, say, $\dim\mathcal{H}_B > \dim\mathcal{H}_A$ then there must be some states in $\mathcal{H}_A$ that do not appear in the Schmidt decomposition (9.29).

If the probabilities p_i are distinct then the Schmidt decomposition is unique. In contrast, if ρ_A has degenerate eigenvalues then there is some ambiguity in the Schmidt decomposition, as we get to decide which of the degenerate eigenvectors in $\mathcal{H}_A$ pairs with their counterpart in $\mathcal{H}_B$.

The *Schmidt rank* R is the number of non-zero eigenvalues p_i in the decomposition (9.29). If $R = 1$ then the state takes the form

$$|\Psi\rangle = |\phi\rangle \otimes |\tilde{\chi}\rangle$$

and is separable. If $R > 1$, the state is entangled.

Finally, let's go back to Alice and Bob. Each gets to act on their subsystem by transforming the state they have to any other. This means that, between them, they get to act with unitary operators on $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ of the form

$$U = U_A \otimes U_B$$

However, the state $|\Psi\rangle$ and the state $U|\Psi\rangle$ have the same Schmidt rank. This is important. It tells us that we cannot change the amount of entanglement by local operators which act only on part of the Hilbert space. To create entanglement, we need to act with operators which rotate $\mathcal{H}_A$ into $\mathcal{H}_B$. In other words, there has to be some interaction between the two parts of the subsystem. Entanglement can only be created by bringing the two subsystems together.

Purification

There is a simple corollary to our discussion above. For any density matrix ρ describing a state in a Hilbert space $\mathcal{H}_A$, one can always find a pure state $|\Psi\rangle$ in a larger Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ such that $\rho = \mathrm{Tr}_{\mathcal{H}_B}|\Psi\rangle\langle\Psi|$. This process is referred to as *purification* of the state.

Everything that we need to show this is in our derivation above. We write the density matrix in the orthonormal basis (9.28). We then introduce the enlarged Hilbert space $\mathcal{H}_B$ whose dimension is that same as the number of non-zero p_i in (9.28). The Schmidt decomposition (9.29) then provides an example of a purification of ρ.

9.3.3 Entropy

Given a classical probability distribution $\{p_i\}$, the *entropy* is defined by

$$S = -\sum_i p_i \log p_i \tag{9.30}$$

where log is the natural logarithm. In information theory, this is called the *Shannon entropy*. In physics, this quantity is usually multiplied by the Boltzmann constant k_B and is called the *Gibbs entropy*. It plays an important role in the lectures on Statistical Physics.

The entropy is a measure of the uncertainty encoded in the probability distribution. For example, if there's no uncertainty because, say $p_1 = 1$ while all other $p_i = 0$, then we have $S = 0$. In contrast, if there are N possibilities the entropy is maximised when we have no idea which is most likely, meaning that $p_i = 1/N$ for each. In this case $S = \log N$.

For a quantum state described by a density matrix ρ, we defined the entropy to be

$$S(\rho) = -\mathrm{Tr}\,\rho \log \rho \tag{9.31}$$

This is the *von Neumann entropy* (because entropy really needs more names attached to it). If we're dealing with a reduced density matrix, that came from taking a partial trace of a pure state of a larger system, then S is referred to as the *entanglement entropy*. In all cases, we're simply going to call it the entropy.

When the density matrix is expanded in an orthonormal basis,

$$\rho = \sum_i p_i |\phi_i\rangle\langle\phi_i|$$

then the definition (9.31) coincides with the earlier definition (9.30).

A pure state has $p_i = 1$ for some $|\phi_i\rangle$, and so has vanishing entropy. But $S \neq 0$ for any mixed state.

The entropy has a number of properties, some of which are easier to prove than others. First the properties that are straightforward to show:

- Positivity: $S(\rho) \geq 0$.
- Minimum: $S(\rho) = 0$ if and only if ρ is a pure state.
- Maximum: If the probabilities are non-vanishing on an N dimensional Hilbert space $\mathcal{H}_N$, then the entropy takes its maximum value $S = \log N$ when $\rho = \frac{1}{N}\mathbf{1}$ on $\mathcal{H}_N$.
- Concavity: If $\sum \lambda_i = 1$, then

$$S(\sum_i \lambda_i \rho_i) \geq \sum_i \lambda_i S(\rho_i)$$

  This tells us that if we are more ignorant about the make-up of our state, then the entropy increases.

The entropy obeys a number of further properties. Two which are particularly important are:

- Subadditivity: If $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ then

$$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B) \tag{9.32}$$

  with equality only if the two systems are uncorrelated, so that $\rho_{AB} = \rho_A \otimes \rho_B$. Subadditivity tells us that the entropy of the whole is less than the sum of its parts. This result fairly straightforward to prove, although we won't do so here.

- Strong Subadditivity: If $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ then

$$S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC})$$

  This result is famously tricky to prove. It's perhaps best thought of by thinking of AB and BC as two systems which overlap on B. Then strong subadditivity says that the total entropy of the two parts is not less than the total entropy together with the entropy of their overlap.

9.4 Measurement

The act of measurement is one of the more mysterious aspects of quantum mechanics. It is here that we appear to abandon unitary evolution in favour of the abrupt collapse of the wavefunction, and it is here that we must embrace the indeterministic nature of the quantum world. In this section, we'll take a closer look at what we mean by measurement.

9.4.1 Projective Measurements

We start by recalling what we learned in previous courses. An *observable* in quantum mechanics is a Hermitian operator $\mathcal{O}$. We can decompose this in a *spectral representation*, meaning we write

$$\mathcal{O} = \sum_m \lambda_m P_m \tag{9.33}$$

where λ_m are the eigenvalues of $\mathcal{O}$ and P_m are the projectors onto the corresponding eigenspaces. The projection operators obey $P_m = P_m^\dagger$. The eigenspaces are necessarily orthogonal, meaning

$$P_m P_n = P_m \delta_{mn} \tag{9.34}$$

Moreover, the eigenvectors span the entire Hilbert space, so we also have

$$\sum_m P_m = \mathbf{1} \tag{9.35}$$

Given a state $|\psi\rangle$, the result of a measurement in quantum mechanics is dictated by two, further axioms. The first says that a measurement of the operator $\mathcal{O}$ returns the result λ_m with probability

$$p(m) = \langle\psi|P_m|\psi\rangle \tag{9.36}$$

This is the *Born rule*.

The axiom states that after the measurement, the system no longer sits in the state $|\psi\rangle$. Instead, the act of measurement has disturbed the state, leaving it in the new state

$$|\psi\rangle \mapsto \frac{P_m|\psi\rangle}{\sqrt{p(m)}} \tag{9.37}$$

where the $\sqrt{p(m)}$ in the denominator is there to ensure that the resulting state is correctly normalised. The non-unitary evolution captured by (9.37) is the infamous *collapse of the wavefunction*.

There are a couple of simple generalisations of the above formalism. First, suppose that we start with a mixed state, described by a density matrix ρ. Then the Born rule (9.36) and collapse (9.37) are replaced by

$$p(m) = \mathrm{Tr}(\rho P_m) \quad \text{and} \quad \rho \mapsto \frac{P_m \rho P_m}{p(m)} \tag{9.38}$$

Note, in particular, that the resulting density matrix still has unit trace, as it must to describe a state.

As an alternative scenario, suppose that we don't know the outcome of the measurement. In this case, the collapse of the wavefunction turns an initial state $|\psi\rangle$ into a mixed state, described by the density matrix

$$|\psi\rangle \mapsto \sum_m p(m) \frac{P_m|\psi\rangle\langle\psi|P_m}{p(m)} = \sum_m P_m|\psi\rangle\langle\psi|P_m \tag{9.39}$$

If we don't gain any knowledge after our quantum system interacts with the measuring apparatus, this is the correct description of the resulting state.

We can rephrase this discussion without making reference to the original operator $\mathcal{O}$. We say that a *measurement* consists of presenting a quantum state with a complete set of orthogonal projectors $\{P_m\}$. These obey (9.34) and (9.35). We ask the system "Which of these are you described by?" and the system responds by picking one. This is referred to as a *projective measurement*.

In this way of stating things, the projection operators take centre stage. The answer to a projective measurement is sufficient to tell us the value of any physical observable $\mathcal{O}$ whose spectral decomposition (9.33) is in terms of the projection operators $\{P_m\}$ which we measured. In this way, the answer to a projective measurement can only furnish us with information about commuting observables, since these have spectral representations in terms of the same set of projection operators.

Gleason's Theorem

Where does the Born rule come from? Usually in quantum mechanics, it is simply proffered as a postulate, one that agrees with experiment. Nonetheless, it is the rule that underlies the non-deterministic nature of quantum mechanics and given this is such a departure from classical mechanics, it seems worth exploring in more detail.

There have been many attempts to derive the Born rule from something simpler, none of them very convincing. But there is a mathematical theorem which gives some comfort. This is *Gleason's theorem*, which we state here without proof. The theorem says that for any Hilbert space $\mathcal{H}$ of dimension $\dim\mathcal{H} \geq 3$, the only consistent way of assigning probabilities $p(m)$ to all projection operators P_m acting on $\mathcal{H}$ is through the map

$$p(m) = \text{Tr}(\rho P_m)$$

for some self-adjoint, positive operator ρ with unit trace. Gleason's theorem doesn't tell us why we're obliged to introduce probabilities associated to projection operators. But it does tell us that if we want to go down that path then the only possible way to proceed is to introduce a density matrix ρ and invoke the Born rule.

9.4.2 Generalised Measurements

There are circumstances where it is useful to go beyond the framework of projective measurements. Obviously, we're not going to violate any tenets of quantum mechanics, and we won't be able to determine the values of observables that don't commute. Nonetheless, focussing only on projection operators can be too restrictive.

A *generalised measurement* consists of presenting a quantum state with a compete set of Hermitian, positive operators $\{E_m\}$ and asking: "Which of these are you described by?". As before, the system will respond by picking one.

We will require that the operators E_m satisfy the following three properties:

- Hermitian: $E_m = E_m^\dagger$
- Complete: $\sum_m E_m = \mathbf{1}$
- Positive: $\langle\psi|E_m|\psi\rangle \geq 0$ for all states $|\psi\rangle$.

These are all true for projection operators $\{P_m\}$ and the projective measurements described above are a special case. But the requirements here are weaker. In particular, in contrast to projective measurements, the number of E_m in the set can be larger than the dimension of the Hilbert space. A set of operators $\{E_m\}$ obeying these three conditions is called a *positive operator-valued measure*, or *POVM* for short.

Given a quantum state $|\psi\rangle$, we will define the probability of finding the answer E_m to our generalised measurement to be

$$p(m) = \langle\psi|E_m|\psi\rangle$$

Alternatively, if we are given a density matrix ρ, the probability of finding the answer E_m is

$$p(m) = \mathrm{Tr}\,(\rho E_m) \tag{9.40}$$

At the moment we will take the above rules as a definition, a generalisation of the usual Born rule. Note, however, that the completeness and positivity requirements above ensure that $p(m)$ define a good probability distribution. Shortly we will see how this follows from the more familiar projective measurements.

An Example: State Determination

Before we place generalised measurements in a more familiar setting, let's first see how they are may be useful. Suppose that someone hands you a qubit and tells you that it's either $|\uparrow\rangle$ or it's $|\rightarrow\rangle = (|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}$. How can you find out which state you've been given?

The standard rules of quantum mechanics ensure that there's no way to distinguish two non-orthogonal states with absolute certainty. Nonetheless, we can see how well we can do. Let's start with projective measurements. We can consider the set

$$P_1 = |\uparrow\rangle\langle\uparrow| \quad , \quad P_2 = |\downarrow\rangle\langle\downarrow|$$

If the result of the measurement is P_1 then we can't say anything. If, however, the result of the measurement is P_2 then we must have been handed the state $|\rightarrow\rangle$ because the other state obeys $P_2|\uparrow\rangle = 0$ and so has vanishing probability of giving the answer P_2. This means that if we're handed a succession of states $|\uparrow\rangle$ and $|\rightarrow\rangle$, each with equal probability, then we can use projective measurements to correctly identify which one we have 25% of the time.

Generalised measurements allow us to do better. Consider now the set of operators

$$E_1 = \frac{1}{2}|\downarrow\rangle\langle\downarrow| \quad , \quad E_2 = \frac{1}{2}|\leftarrow\rangle\langle\leftarrow| \quad , \quad E_3 = \mathbf{1} - E_1 - E_2 \tag{9.41}$$

If the result of the measurement is E_1, then we must have been handed the state $|\rightarrow\rangle$. But, by the same argument, if the result of the state is E_2 then we must have been handed the state $|\uparrow\rangle$. Finally, if the result is E_3 then we've got no way of knowing which state we were handed. The upshot is that if we're handed a succession of states $|\uparrow\rangle$ and $|\rightarrow\rangle$, we can use generalised measurements to correctly identify which one we have 50% of the time.

Generalised Measurements are Projective Measurements in Disguise

The generalised measurements are not quite as novel as they first appear. They can always be realised as projective measurements in disguise, where the disguise in question involves some hidden, larger Hilbert space.

Let's first consider our POVM (9.41). Suppose that when we were handed the states $|\uparrow\rangle$ and $|\rightarrow\rangle$, they were actually the first in a pair of qubits, whose full states were given by

$$|\Psi_1\rangle = |\uparrow\rangle \otimes |\uparrow\rangle \quad \text{and} \quad |\Psi_2\rangle = |\rightarrow\rangle \otimes |\downarrow\rangle \tag{9.42}$$

Now these states are orthogonal to each other and, therefore, distinguishable.

We will suppose that the density matrix in the full Hilbert space is separable, meaning $\rho = \rho_1 \otimes \rho_2$. Someone – say, Alice – who has access to both spins can perform projective measurements in the full four-dimensional Hilbert space, with the resulting probabilities

$$p(m) = \mathrm{Tr}_{\mathcal{H}_1}\mathrm{Tr}_{\mathcal{H}_2}(\rho P_m)$$

What about Bob, who has access only to the first spin? Written in terms of operators acting on the first qubit, we have

$$p(m) = \mathrm{Tr}_{\mathcal{H}_1}(\rho_1 E_m) \quad \text{where} \quad E_m = \mathrm{Tr}_{\mathcal{H}_2}(\rho_2 P_m) \tag{9.43}$$

Here the operators E_m form a POVM on $\mathcal{H}_1$, the Hilbert space of the first qubit. Both positivity and completeness follow from the properties of the density matrix ρ_2 and the projection operators P_m. For example, completeness comes from

$$\sum_m E_m = \mathrm{Tr}_{\mathcal{H}_2}(\rho_2 \sum_m P_m) = \mathrm{Tr}_{\mathcal{H}_2}(\mathbf{1}_2 \otimes \rho_2) = \mathbf{1}_2$$

We learn that the formalism of generalised measurements allows Bob to reproduce any information that pertains only to the first spin. This is sensible because the original density matrix $\rho = \rho_1 \otimes \rho_2$ was separable, which means that there will be no hidden correlations between the two spins that Alice has access to, but Bob does not.

There are different ways to arrive at the particular POVM (9.41). For example, we could consider the situation where we have maximal ignorance about the second spin, so $\rho_2 = \frac{1}{2}\mathbf{1}_2$. Then we can then consider the projectors

$$P_1 = |\Psi_1\rangle\langle\Psi_1| \quad , \quad P_2 = |\Psi_2\rangle\langle\Psi_2| \quad , \quad P_3 = \mathbf{1}_4 - P_1 - P_2$$

In this case, the POVM defined by (9.43) coincides with (9.41).

It should be clear that the construction leading to the POVM (9.43) holds more generally than our two-state system. A projective measurement in any Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$ reduces to a POVM when taken on separable density matrices. In fact that converse is also true: any POVM can be realised by projection operators acting on a larger Hilbert space. This follows from a fairly simple result in linear algebra known as *Naimark's dilatation theorem* (sometimes transliterated from the Russian as *Neumark's theorem*.)

9.4.3 The Fate of the State

The projective measurements that we met in Section 9.4.1 have two ingredients. The first is the probability that a given result occurs; the second is the fate of the state after the measurement

$$p(m) = \mathrm{Tr}(\rho P_m) \quad \text{and} \quad \rho \mapsto \frac{P_m \rho P_m}{p(m)} \tag{9.44}$$

For our generalised measurements, we have explained how the probabilities are replaced by $p(m) = \mathrm{Tr}(\rho E_m)$. But what happens to the state after the measurement?

We could try to take inspiration from thinking about generalised measurements in terms of projection operators in an enlarged Hilbert space. We know that

$$\rho = \rho_1 \otimes \rho_2 \mapsto \frac{P_m \rho P_m}{p(m)} \quad \Rightarrow \quad \rho_1 \mapsto \mathrm{Tr}_{\mathcal{H}_2} \frac{P_m \rho P_m}{p(m)}$$

But there's no simple way of writing this in terms of the elements of the POVM $E_m = \mathrm{Tr}_{\mathcal{H}_2}(\rho_2 P_m)$. And this is for good reason: the POVM does not include enough information to tell us the fate of the state.

Instead, we have to define a "square-root" of E_m. This is an operator M_m such that

$$\sum_m M_m^\dagger M_m = E_m \tag{9.45}$$

The M_m need not be Hermitian. Furthermore, these operators are not uniquely determined by (9.45): any unitary transformation $M_m \to U M_m$ still obeys (9.45). The completeness of the POVM means that they obey

$$\sum_m M_m^\dagger M_m = \mathbf{1}$$

The choice of M_m is the extra information we need to specify the state after a generalised measurement. If we perform a generalised measurement and find the answer E_m, then the state becomes

$$\rho \mapsto \frac{M_m \rho M_m^\dagger}{p(m)} \tag{9.46}$$

This new density matrix is Hermitian and has unit trace, as it must.

A full generalised measurement – one in which both the probabilities and the end state are known – is specified by the set of operators $\{M_m\}$, such that $E_m = M_m^\dagger M_m$ form a POVM. The generalised measurement reduces to the projective measurement of Section 9.4.1 only when M_m are orthogonal projection operators.

Finally, note that if we make a measurement, but don't know the result, then the resulting density matrix is not given by (9.46), but instead by

$$\rho \mapsto \sum_m M_m \rho M_m^\dagger \tag{9.47}$$

This generalises our result (9.39) for projective measurements.

Repeated Measurements

The special class of projective measurements enjoys some nice properties that are not shared by their generalised counterparts. Perhaps the most prominent is what happens upon repeated measurements.

For projective measurements, if we get a result P_m the first time round, then any subsequent measurement is guaranteed to give the same result. This result is familiar from our earlier courses on quantum mechanics: if you measure the spin of a particle to be up then, as long as the particle is left alone, its spin will continue to be up next time round.

This property doesn't hold for generalised measurements. Returning to our POVM (9.41), a measurement of E_1 in the first round does not preclude a measurement of E_2 or E_1 the next time round.

An Example: Detecting a Photon

The idea of generalised measurement is useful even when our POVM consists of projection operators. A standard example is the detection of a photon. Before the measurement takes place, either the photon exists $|1\rangle$ or it doesn't $|0\rangle$.

A projective measurement (9.44) would tell us that if we detect a photon, then it's there to detect again on our next measurement. But that's not what happens. Typically when we detect a photon, the photon doesn't live to tell the tale. Instead, it is destroyed in the process. This means that whether a photon is seen or not, the end result is always the same: no photon $|0\rangle$. In terms of our new generalised measurements, this can be simply described by the operators

$$M_1 = |0\rangle\langle 0| \quad \text{and} \quad M_2 = |0\rangle\langle 1|$$

which corresponds to the POVM

$$E_1 = M_1^\dagger M_1 = |0\rangle\langle 0| \quad \text{and} \quad E_2 = M_2^\dagger M_2 = |1\rangle\langle 1|$$

In this case, the POVM consists of projection operators. But the collapse of the wavefunction (9.46) differs from the usual projective measurement. Regardless of the outcome of the initial experiment, if you now try to repeat it the photon will not be there.

9.5 Open Systems

In this section we will again consider situations where the full Hilbert space decomposes into two parts: $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$. However, we will no longer think of these subspaces as the far-separated homes of Alice and Bob. Instead, $\mathcal{H}_S$ will denote the system that we want to study, and $\mathcal{H}_E$ will denote the surrounding environment.

Here the *environment* is typically a vast Hilbert space which we have no way of understanding completely. In this sense, it plays a similar role to the thermal baths that we introduce in statistical physics. When performing an experiment on a quantum system, much of the challenge is trying to shield it from the environment. However, in many cases this is not possible and there will be coupling between $\mathcal{H}_S$ and $\mathcal{H}_E$. We then say that $\mathcal{H}_S$ is an *open system*. The purpose of this section is to understand how such open quantum systems behave.

9.5.1 Quantum Maps

We will assume that the combined system+environment is described by a pure state $|\Psi\rangle$. We've seen in Section 9.3.2 that, after tracing over $\mathcal{H}_E$, the system we care about is typically described by a reduced density matrix

$$\rho = \mathrm{Tr}_{\mathcal{H}_E} |\Psi\rangle\langle\Psi|$$

We would like to understand how this density matrix evolves.

The state $|\Psi\rangle$ evolves by a unitary operator $U(t)$ acting on the full Hilbert space $\mathcal{H}$. The story that we are about to tell only works if, at time $t = 0$, the two systems lie in a separable state,

$$|\Psi_0\rangle = |\psi\rangle \otimes |\chi\rangle \tag{9.48}$$

This means that the original density matrix $\rho_0 = |\psi\rangle\langle\psi|$ describes a pure state on $\mathcal{H}_S$. We now look at how this density matrix evolves. We have

$$\rho(t) = \mathrm{Tr}_{\mathcal{H}_E}\, U(t)|\Psi_0\rangle\langle\Psi_0|U^\dagger(t) = \sum_m \langle m|U(t)|\Psi_0\rangle\langle\Psi_0|U^\dagger(t)|m\rangle$$

with $|m\rangle$ a complete basis for $\mathcal{H}_E$. This encourages us to define a set of operators on $\mathcal{H}_S$, given by

$$M_m(t) = \langle m|U(t)|\chi\rangle = \mathrm{Tr}_{\mathcal{H}_E}\Big(U(t)|\chi\rangle\langle m|\Big) \tag{9.49}$$

The unitarity of $U(t)$ translates into a completeness condition on the operators $M_m(t)$,

$$\sum_m M_m^\dagger(t) M_m(t) = \sum_m \langle\chi|U^\dagger(t)|m\rangle\langle m|U(t)|\chi\rangle = \mathbf{1}$$

We see that the original density matrix on $\mathcal{H}_S$ evolves as

$$\rho(t) = \sum_m M_m(t)\, \rho_0\, M_m^\dagger(t) \tag{9.50}$$

In general, this will describe the evolution from a pure state to a mixed state. This evolution is not, in general, reversible.

A quick comment: this evolution takes the same general form as the measurement process (9.47), at least if we don't gain any knowledge about the result of the measurement. This is not coincidence. A measuring apparatus is a macroscopic system that becomes entangled with the quantum state. In this sense, it plays a similar role to the environment in the discussion above.

In contrast, if we read off the result of a measurement, then the resulting state is described by (9.46); this does not take the form (9.50).

Kraus Representation Theorem

Above, we have derived the evolution (9.50) in a rather simple example. However, it turns out that this form has more general applicability. Consider a density operator $\mathcal{H}_S$ which evolves by the map

$$\rho \;\mapsto\; \mathcal{L}[\rho]$$

Such a map is sometimes called a *superoperator* (because it maps operators to operators, rather than states to states). We will require some special properties of our map, most of which are inherited from the properties of the density matrices listed in Section 9.3

- Linearity: $\mathcal{L}[a\rho_1 + b\rho_2] = a\mathcal{L}[\rho_1] + b\mathcal{L}[\rho_2]$.
- Hermiticity Preserving: $\rho = \rho^\dagger \;\Rightarrow\; \mathcal{L}[\rho] = \mathcal{L}[\rho]^\dagger$.
- Trace Preserving: $\mathrm{Tr}\,\mathcal{L}[\rho] = \mathrm{Tr}\,\rho$.
- Complete Positivity. This one requires some explanation. It is natural to insist that the map is positive, so that $\mathcal{L}[\rho] \geq 0$ whenever $\rho \geq 0$. However, this is not sufficient. Instead, we require the stronger statement that the map $\mathcal{L} \otimes \mathbf{1}_E$ is positive on any extension of the Hilbert space $\mathcal{H}_S$ to $\mathcal{H}_S \otimes \mathcal{H}_E$. This is the statement of complete positivity. It ensures that the map $\mathcal{L} \otimes \mathbf{1}_E$ will take a valid density matrix on the composite system to another density matrix.

A superoperator obeying these conditions is called a *trace preserving, completely positive (TPCP) map*, with the first two conditions taken for granted. In the quantum information community, this map is referred to as a *quantum channel*.

The *Kraus representation theorem* (which we do not prove here) states that any quantum map, obeying the four conditions above, can be written as

$$\mathcal{L}[\rho] = \sum_m M_m \rho M_m^\dagger \qquad \text{with} \quad \sum_m M_m^\dagger M_m = \mathbf{1} \tag{9.51}$$

In this framework, the M_m are called *Kraus operators*. They are not unique. The number of Kraus operators in the quantum map does not exceed $\dim(\mathcal{H}_S)^2$.

You might wonder why the collapse of the wavefunction (9.46) fails to take the Kraus form (9.51). It is because the map is not linear: the probability $p(m)$ which normalises the resulting density matrix itself depends on ρ through (9.40).

9.5.2 Decoherence

In this section, we explore some simple examples of quantum maps. We'll use these toy models to highlight some important and general features that emerge when quantum systems interact with an environment.

Phase-Damping

We will take the quantum system $\mathcal{H}_S$ to be our trusty qubit. Meanwhile, we will model the environment $\mathcal{H}_E$ by a three-state system, spanned by $|0\rangle$, $|1\rangle$ and $|2\rangle$. Consider the following unitary evolution

$$\begin{aligned} U|\uparrow\rangle \otimes |0\rangle &= |\uparrow\rangle \otimes (\sqrt{1-p}\,|0\rangle + \sqrt{p}\,|1\rangle) \\ U|\downarrow\rangle \otimes |0\rangle &= |\downarrow\rangle \otimes (\sqrt{1-p}\,|0\rangle + \sqrt{p}\,|2\rangle) \end{aligned} \tag{9.52}$$

This means that our qubit interacts with the environment with probability p, changing the initial state $|0\rangle$ into either $|1\rangle$ or $|2\rangle$ depending on the state of the qubit. Note, however, that the state of the qubit is unchanged by this interaction. So this model describes a system in which the energies needed to change the qubit are substantially larger than those needed to change the environment.

If you want a specific picture in mind, you could think of the qubit as a simplified model for a heavy dust particle which, in this case, can only sit in one of two positions $|\uparrow\rangle$ or $|\downarrow\rangle$. The environment could be a background bath of photons which scatter off this dust particle with probability p.

The Kraus operators for this quantum map are easily calculated. Using (9.49), they are given by

$$\begin{aligned} M_0 &= \langle 0|U|0\rangle = \sqrt{1-p}\,\mathbf{1} \\ M_1 &= \langle 1|U|0\rangle = \sqrt{p}\,|\uparrow\rangle\langle\uparrow| \\ M_2 &= \langle 2|U|0\rangle = \sqrt{p}\,|\downarrow\rangle\langle\downarrow| \end{aligned} \tag{9.53}$$

which can be checked to obey the required completeness condition $\sum_m M_m^\dagger M_m = \mathbf{1}$. The state of the qubit, described by a density matrix ρ, then evolves as

$$\begin{aligned} \rho \;\mapsto\; \mathcal{L}[\rho] = \sum_m M_m^\dagger \rho M_m &= (1-p)\,\rho + p|0\rangle\langle 0|\rho|0\rangle\langle 0| + p|1\rangle\langle 1|\rho|1\rangle\langle 1| \\ &= (1-\frac{1}{2}p)\rho + \frac{1}{2}p\sigma^3\rho\sigma^3 \end{aligned}$$

We can see the essence of this quantum map if we write the density matrix in terms of components

$$\rho = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} \;\mapsto\; \begin{pmatrix} \rho_{00} & (1-p)\rho_{01} \\ (1-p)\rho_{10} & \rho_{11} \end{pmatrix}$$

We learn that the off-diagonal components are suppressed by the evolution. It is these off-diagonal elements which encode possible superpositions of $|\uparrow\rangle$ and $|\downarrow\rangle$. The interactions with the environment — or, more precisely, the resulting entanglement with the environment — means that these off-diagonal elements are reduced under time evolution. This process is known as *decoherence*; it is the evolution of a pure state into a mixed state through interactions with the environment.

We can get a better sense of this if we look at successive maps. This is a little subtle because it's not obvious when we can apply successive Kraus operators (9.53). We will discuss this in more detail in Section 9.5.3, but for now we simply look at what happens.

We define the probability of scattering per unit time to be Γ. Then, in time δt, we have $p = \Gamma\delta t \ll 1$. After a total time $t = N\delta t$, the off-diagonal terms in the density matrix are suppressed by

$$(1-p)^N = (1-\Gamma t/N)^N \approx e^{-\Gamma t} \tag{9.54}$$

Suppose that we initially prepare our qubit in a state

$$|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle \qquad |\alpha|^2 + |\beta|^2 = 1$$

Then after time t, the density matrix becomes

$$\rho(t) = \begin{pmatrix} |\alpha|^2 & \alpha\beta^\star e^{-\Gamma t} \\ \alpha^\star \beta e^{-\Gamma t} & |\beta|^2 \end{pmatrix}$$

We see that these off-diagonal components decay exponentially quickly, with the system ultimately settling down into a mixed state. The choice of preferred basis $|\uparrow\rangle$, $|\downarrow\rangle$ can be traced to the form of the original interaction (9.52)

To flesh this out a little, let's return to our interpretation of this model in terms of a heavy dust particle which can sit in one of two positions, $|\uparrow\rangle = |x_+\rangle$ or $|\downarrow\rangle = |x_-\rangle$. We may, of course, choose to place this particle in a superposition

$$|\psi\rangle = \alpha|x_+\rangle + \beta|x_-\rangle$$

and hope to measure this superposition in some way. This, of course, is what happens in the double-slit experiment. However, decoherence makes this difficult. Indeed, if the particle takes time $t \gg \Gamma^{-1}$ to traverse the double slit experiment then all the hint of the superposition will be washed out. Furthermore, Γ^{-1} is typically a very short timescale; it is the rate at which a single photon scatters off the particle. This can be much much shorter than the rate at which the classical properties of the particle – say its energy – are affected by the photons.

There is one final important lesson to take away from this model. It explains why the decoherence occurs in the position basis $|x_\pm\rangle$ rather than say, $(|x_+\rangle \pm |x_-\rangle)/\sqrt{2}$. This is because the interactions (9.52) are *local*.

The locality of interactions is one of the key features of all physical laws; indeed, it underlies the idea of quantum field theory. Combined with decoherence, this explains why we only see our favourite pets in the state $|\text{alive}\rangle$ or $|\text{dead}\rangle$. Interactions with the environment mean that it is overwhelmingly unlikely to observe Schrödinger's cat in the state $|\Psi\rangle = (|\text{alive}\rangle \pm |\text{dead}\rangle)/\sqrt{2}$.

Amplitude Damping

Our second example will not give us further insight into decoherence, but instead provides a simple model for the decay of an excited atom. (A more detailed look at the dynamics underling this can be found in Section 8.4.3.) Consider a two-state atomic system. If the atom is in the ground state $|\uparrow\rangle$ then nothing happens, but if atom is in the excited state $|\downarrow\rangle$ then it decays with probability p emitting a photon in the process, so that the environment changes from $|0\rangle$ to $|1\rangle$. This is captured by the unitary evolution

$$\begin{aligned} U|\uparrow\rangle \otimes |0\rangle &= |\uparrow\rangle \otimes |0\rangle \\ U|\downarrow\rangle \otimes |0\rangle &= \sqrt{1-p}\,|\downarrow\rangle \otimes |0\rangle + \sqrt{p}\,|\uparrow\rangle \otimes |1\rangle) \end{aligned}$$

The resulting Kraus operators are

$$M_0 = \langle 0|U|0\rangle = |\uparrow\rangle\langle\uparrow| + \sqrt{1-p}|\downarrow\rangle\langle\downarrow| \quad , \quad M_1 = \langle 1|U|0\rangle = \sqrt{p}\,|\uparrow\rangle\langle\downarrow|$$

This time the quantum map is given by

$$\rho = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} \mapsto \begin{pmatrix} \rho_{00} + p\rho_{11} & \sqrt{1-p}\,\rho_{01} \\ \sqrt{1-p}\,\rho_{10} & (1-p)\rho_{11} \end{pmatrix}$$

If, as previously, we can think about performing this map successive time, with the probability for decay p related to the lifetime Γ^{-1} of the excited state through (9.54) then we find the time-dependent density matrix given by

$$\rho(t) = \begin{pmatrix} \rho_{00} + (1-e^{-\Gamma t})\rho_{11} & e^{-\Gamma t/2}\rho_{01} \\ e^{-\Gamma t/2}\rho_{10} & e^{-\Gamma t}\rho_{11} \end{pmatrix}$$

Interestingly, as $t \to \infty$, the system ends up in the pure state $|\uparrow\rangle$, regardless of whatever superposition or mixed state it started in. On the one hand this is not surprising: it is simply the statement that if we wait long enough the atom will surely have decayed. Nonetheless, it does provide a simple example in which quantum maps can take a mixed state to a pure state.

9.5.3 The Lindblad Equation

Usually in physics, the most powerful way to describe the evolution of a system is through a differential equation. For a closed quantum system in a pure state, the relevant equation is the Schrödinger equation. For a closed quantum system in a mixed state, it is the Liouville equation (9.21)

$$\hbar\,\frac{\partial\rho}{\partial t} = -i[H,\rho]$$

where the density matrix ρ is an operator on $\mathcal{H}_S$. Here we would like to derive the analogous equation for an open quantum system, where $\mathcal{H}_S$ is also coupled to an environment $\mathcal{H}_E$.

It is not at all clear that such an equation will exist. Knowledge of the density matrix ρ on $\mathcal{H}_S$ at some time will not, in general, be sufficient to tell you how the density matrix will behave in the future. The problem is not just that the environment can affect our system — that, after all is what we're trying to model. The problem is more one of memory.

As time progresses, the system changes the environment. Our concern is that these changes accumulate, so that the environment starts to affect the system in different ways. In this sense, the environment can act as a memory, where the state of the system in the future depends not only on the present state, but on its entire history. These kind of situations are complicated.

We've already seen a hint of this in our earlier work. Recall that when we first looked at quantum maps, we assumed that the initial state (9.48) was separable, with no correlations between $\mathcal{H}_S$ and $\mathcal{H}_E$. Had we included these correlations, we would not have found such a simple, linear quantum map. Yet, such correlations inevitably build with time, meaning that we should be careful about performing successive quantum maps. This is a manifestation of the memory of the environment.

To make progress, we will restrict ourselves to situations where this memory does not last. We will consider the environment to be vast, similar to the heat reservoirs that we use in statistical mechanics. We assume that correlations between the system and the environment are lost over a certain time scale. We will denote this time scale by τ, and seek an equation which dictates the dynamics of ρ on timescales $t \gg \tau$.

Our starting point is the quantum map (9.50),

$$\rho(t+\delta t) = \sum_m M_m(t+\delta t)\,\rho(t)\,M_m^\dagger(t+\delta t) \tag{9.55}$$

We will take δt to be small, as if we were dealing with usual calculus of infinitesimals. But we should bear in mind that really we want $\delta t \gg \tau$. For this equation to hold, we must have one Kraus operator — say M_0 — to take the form $M_0 = \mathbf{1} + \mathcal{O}(\delta t)$. The remaining operators should be $M_m \sim \mathcal{O}(\sqrt{\delta t})$. We write

$$M_0 = \mathbf{1} + \frac{1}{\hbar}(K - iH)\delta t \quad , \quad M_m = \frac{1}{\sqrt{\hbar}} L_m \sqrt{\delta t} \quad m = 1, 2, \ldots$$

where both H and K are chosen to be Hermitian matrices. These Kraus operators must obey the completeness relation (9.51),

$$\sum_{m=0} M_m^\dagger M_m = \mathbf{1} \quad \Rightarrow \quad 2K + \sum_{m=1} L_m^\dagger L_m = \mathcal{O}(\delta t^2)$$

We therefore write

$$K = -\frac{1}{2}\sum_{m=1} L_m^\dagger L_m$$

Plugging these expressions into the quantum map (9.55), and keeping only terms of order δt, we get our final result

$$\hbar\,\frac{\partial\rho}{\partial t} = -i[H,\rho] + \sum_{m=1}\left[L_m\rho L_m^\dagger - \frac{1}{2}L_m^\dagger L_m\rho - \frac{1}{2}\rho L_m^\dagger L_m\right]$$

This is the *Lindblad equation.* It should be thought of as a quantum version of the Fokker-Planck equation that is described in the lectures in Kinetic Theory. We see that the evolution is governed not just by the Hamiltonian H, but also by further *Lindblad operators* L_m which capture the interaction with the environment. The presence of the final two terms ensures that $d(\mathrm{Tr}\,\rho)/dt = 0$, as it should for a density matrix.

The Increase of Entropy

Something particularly nice happens when the Lindblad operators are Hermitian, so $L_m = L_m^\dagger$. In this case, the entropy increases. The von Neumann entropy is defined as (9.31)

$$S(\rho) = -\mathrm{Tr}\,\rho\log\rho$$

Its change in time is given by

$$\frac{dS}{dt} = -\mathrm{Tr}\left(\frac{\partial\rho}{\partial t}(1+\log\rho)\right) = -\mathrm{Tr}\left(\frac{\partial\rho}{\partial t}\log\rho\right)$$

Inserting the Lindblad equation, we see that the first term vanishes, courtesy of the fact that $[\rho,\log\rho]=0$. We're left with

$$\hbar\frac{dS}{dt} = \sum_m \mathrm{Tr}\left[(L_m\rho L_m - L_m L_m\rho)\log\rho\right]$$

To proceed, we decompose the density matrix ρ in terms of its eigenvectors

$$\rho = \sum_i p_i|\phi_i\rangle\langle\phi_i|$$

and take the trace by summing over the complete basis $|\phi_i\rangle$. We have

$$\begin{aligned}
\hbar\frac{dS}{dt} &= \sum_m\sum_i\langle\phi_i|(L_m\rho L_m - L_m L_m\rho)|\phi_i\rangle\log p_i \\
&= \sum_m\sum_{i,j}\langle\phi_i|L_m|\phi_j\rangle\langle\phi_j|L_m|\phi_i\rangle(p_j-p_i)\log p_i \\
&= \frac{1}{2}\sum_m\sum_{i,j}|\langle\phi_i|L_m|\phi_j\rangle|^2(p_j-p_i)(\log p_i - \log p_j)
\end{aligned}$$

where, in going to the final line, we took advantage of the anti-symmetric properties of the middle line under the exchange of i and j. However, the expression $(x-y)(\log x - \log y)$ is positive for all values of x and y. (This same fact was needed in the proof of the H-theorem which is the classical analog of the result we're deriving here.) We learn that

$$\hbar \frac{dS}{dt} \geq 0$$