$$= \sum_i p_i |i^A\rangle\langle i^A| \tag{2.210}$$

$$= \rho^A. \tag{2.211}$$

Thus $|AR\rangle$ is a purification of ρ^A.

Notice the close relationship of the Schmidt decomposition to purification: the procedure used to purify a mixed state of system A is to define a pure state whose Schmidt basis for system A is just the basis in which the mixed state is diagonal, with the Schmidt coefficients being the square root of the eigenvalues of the density operator being purified.

In this section we've explained two tools for studying composite quantum systems, the Schmidt decomposition and purifications. These tools will be indispensable to the study of quantum computation and quantum information, especially quantum information, which is the subject of Part III of this book.

Exercise 2.79: Consider a composite system consisting of two qubits. Find the Schmidt decompositions of the states

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}}; \quad \frac{|00\rangle + |01\rangle + |10\rangle + |11\rangle}{2}; \text{ and } \quad \frac{|00\rangle + |01\rangle + |10\rangle}{\sqrt{3}}. \tag{2.212}$$

Exercise 2.80: Suppose $|\psi\rangle$ and $|\varphi\rangle$ are two pure states of a composite quantum system with components A and B, with identical Schmidt coefficients. Show that there are unitary transformations U on system A and V on system B such that $|\psi\rangle = (U \otimes V)|\varphi\rangle$.

Exercise 2.81: (Freedom in purifications) Let $|AR_1\rangle$ and $|AR_2\rangle$ be two purifications of a state ρ^A to a composite system AR. Prove that there exists a unitary transformation U_R acting on system R such that $|AR_1\rangle = (I_A \otimes U_R)|AR_2\rangle$.

Exercise 2.82: Suppose $\{p_i, |\psi_i\rangle\}$ is an ensemble of states generating a density matrix $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ for a quantum system A. Introduce a system R with orthonormal basis $|i\rangle$.

(1) Show that $\sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle$ is a purification of ρ.

(2) Suppose we measure R in the basis $|i\rangle$, obtaining outcome i. With what probability do we obtain the result i, and what is the corresponding state of system A?

(3) Let $|AR\rangle$ be *any* purification of ρ to the system AR. Show that there exists an orthonormal basis $|i\rangle$ in which R can be measured such that the corresponding post-measurement state for system A is $|\psi_i\rangle$ with probability p_i.

2.6 EPR and the Bell inequality

Anybody who is not shocked by quantum theory has not understood it.
– Niels Bohr

I recall that during one walk Einstein suddenly stopped, turned to me and asked whether I really believed that the moon exists only when I look at it. The rest of this walk was devoted to a discussion of what a physicist should mean by the term 'to exist'.
– Abraham Pais

...quantum phenomena do not occur in a Hilbert space, they occur in a laboratory.
– Asher Peres

...what is proved by impossibility proofs is lack of imagination.
– John Bell

This chapter has focused on introducing the tools and mathematics of quantum mechanics. As these techniques are applied in the following chapters of this book, an important recurring theme is the unusual, *non-classical* properties of quantum mechanics. But what exactly is the difference between quantum mechanics and the classical world? Understanding this difference is vital in learning how to perform information processing tasks that are difficult or impossible with classical physics. This section concludes the chapter with a discussion of the Bell inequality, a compelling example of an essential difference between quantum and classical physics.

When we speak of an object such as a person or a book, we assume that the physical properties of that object have an existence independent of observation. That is, measurements merely act to *reveal* such physical properties. For example, a tennis ball has as one of its physical properties its *position*, which we typically measure using light scattered from the surface of the ball. As quantum mechanics was being developed in the 1920s and 1930s a strange point of view arose that differs markedly from the classical view. As described earlier in the chapter, according to quantum mechanics, an unobserved particle does not possess physical properties that exist independent of observation. Rather, such physical properties arise as a consequence of measurements performed upon the system. For example, according to quantum mechanics a qubit does not possess definite properties of 'spin in the z direction, σ_z', and 'spin in the x direction, σ_x', each of which can be revealed by performing the appropriate measurement. Rather, quantum mechanics gives a set of rules which specify, given the state vector, the probabilities for the possible measurement outcomes when the observable σ_z is measured, or when the observable σ_x is measured.

Many physicists rejected this new view of Nature. The most prominent objector was Albert Einstein. In the famous 'EPR paper', co-authored with Nathan Rosen and Boris Podolsky, Einstein proposed a thought experiment which, he believed, demonstrated that quantum mechanics is not a complete theory of Nature.

The essence of the EPR argument is as follows. EPR were interested in what they termed 'elements of reality'. Their belief was that any such element of reality *must* be represented in any complete physical theory. The goal of the argument was to show that quantum mechanics is not a complete physical theory, by identifying elements of reality that were not included in quantum mechanics. The way they attempted to do this was by introducing what they claimed was a *sufficient condition* for a physical property to be an element of reality, namely, that it be possible to predict with certainty the value that property will have, immediately before measurement.

Box 2.7: Anti-correlations in the EPR experiment

Suppose we prepare the two qubit state

$$|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}, \tag{2.213}$$

a state sometimes known as the *spin singlet* for historical reasons. It is not difficult to show that this state is an entangled state of the two qubit system. Suppose we perform a measurement of spin along the $\vec{v}$ axis on both qubits, that is, we measure the observable $\vec{v} \cdot \vec{\sigma}$ (defined in Equation (2.116) on page 90) on each qubit, getting a result of $+1$ or -1 for each qubit. It turns out that no matter what choice of $\vec{v}$ we make, the results of the two measurements are always opposite to one another. That is, if the measurement on the first qubit yields $+1$, then the measurement on the second qubit will yield -1, and vice versa. It is as though the second qubit knows the result of the measurement on the first, no matter how the first qubit is measured. To see why this is true, suppose $|a\rangle$ and $|b\rangle$ are the eigenstates of $\vec{v} \cdot \vec{\sigma}$. Then there exist complex numbers $\alpha, \beta, \gamma, \delta$ such that

$$|0\rangle = \alpha|a\rangle + \beta|b\rangle \tag{2.214}$$

$$|1\rangle = \gamma|a\rangle + \delta|b\rangle. \tag{2.215}$$

Substituting we obtain

$$\frac{|01\rangle - |10\rangle}{\sqrt{2}} = (\alpha\delta - \beta\gamma)\frac{|ab\rangle - |ba\rangle}{\sqrt{2}}. \tag{2.216}$$

But $\alpha\delta - \beta\gamma$ is the determinant of the unitary matrix $\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$, and thus is equal to a phase factor $e^{i\theta}$ for some real θ. Thus

$$\frac{|01\rangle - |10\rangle}{\sqrt{2}} = \frac{|ab\rangle - |ba\rangle}{\sqrt{2}}, \tag{2.217}$$

up to an unobservable global phase factor. As a result, if a measurement of $\vec{v} \cdot \vec{\sigma}$ is performed on both qubits, then we can see that a result of $+1$ (-1) on the first qubit implies a result of -1 $(+1)$ on the second qubit.

Consider, for example, an entangled pair of qubits belonging to Alice and Bob, respectively:

$$\frac{|01\rangle - |10\rangle}{\sqrt{2}}. \tag{2.218}$$

Suppose Alice and Bob are a long way away from one another. Alice performs a measurement of spin along the $\vec{v}$ axis, that is, she measures the observable $\vec{v} \cdot \vec{\sigma}$ (defined in Equation (2.116) on page 90). Suppose Alice receives the result $+1$. Then a simple quantum mechanical calculation, given in Box 2.7, shows that she can predict with certainty that Bob will measure -1 on his qubit if he also measures spin along the $\vec{v}$ axis. Similarly, if Alice measured -1, then she can predict with certainty that Bob will measure $+1$ on his qubit. Because it is always possible for Alice to predict the value of the measurement result recorded when Bob's qubit is measured in the $\vec{v}$ direction, that physical property must correspond to an element of reality, by the EPR criterion, and should be represented in any complete physical theory. However, standard quantum mechanics, as we have presented it, merely tells one how to calculate the probabilities of the respective measurement outcomes if $\vec{v} \cdot \vec{\sigma}$ is measured. Standard quantum mechanics certainly does not include any fundamental element intended to represent the value of $\vec{v} \cdot \vec{\sigma}$, for all unit vectors $\vec{v}$.

The goal of EPR was to show that quantum mechanics is incomplete, by demonstrating that quantum mechanics lacked some essential 'element of reality', by their criterion. They hoped to force a return to a more classical view of the world, one in which systems could be ascribed properties which existed independently of measurements performed on those systems. Unfortunately for EPR, most physicists did not accept the above reasoning as convincing. The attempt to impose on Nature *by fiat* properties which she must obey seems a most peculiar way of studying her laws.

Indeed, Nature has had the last laugh on EPR. Nearly thirty years after the EPR paper was published, an *experimental test* was proposed that could be used to check whether or not the picture of the world which EPR were hoping to force a return to is valid or not. It turns out that Nature *experimentally invalidates* that point of view, while agreeing with quantum mechanics.

The key to this experimental invalidation is a result known as *Bell's inequality*. Bell's inequality is *not* a result about quantum mechanics, so the first thing we need to do is momentarily *forget* all our knowledge of quantum mechanics. To obtain Bell's inequality, we're going to do a thought experiment, which we will analyze using our common sense notions of how the world works – the sort of notions Einstein and his collaborators thought Nature ought to obey. After we have done the common sense analysis, we will perform a quantum mechanical analysis which we can show *is not consistent with the common sense analysis*. Nature can then be asked, by means of a real experiment, to decide between our common sense notions of how the world works, and quantum mechanics.

Imagine we perform the following experiment, illustrated in Figure 2.4. Charlie prepares two particles. It doesn't matter how he prepares the particles, just that he is capable of repeating the experimental procedure which he uses. Once he has performed the preparation, he sends one particle to Alice, and the second particle to Bob.

Once Alice receives her particle, she performs a measurement on it. Imagine that she has available two different measurement apparatuses, so she could choose to do one of two different measurements. These measurements are of physical properties which we shall label P_Q and P_R, respectively. Alice doesn't know in advance which measurement she will choose to perform. Rather, when she receives the particle she flips a coin or uses some other random method to decide which measurement to perform. We suppose for simplicity that the measurements can each have one of two outcomes, $+1$ or -1. Suppose Alice's particle has a value Q for the property P_Q. Q is assumed to be an *objective property* of Alice's particle, which is merely revealed by the measurement, much as we imagine the position of a tennis ball to be revealed by the particles of light being scattered off it. Similarly, let R denote the value revealed by a measurement of the property P_R.

Similarly, suppose that Bob is capable of measuring one of two properties, P_S or P_T, once again revealing an objectively existing value S or T for the property, each taking value +1 or -1. Bob does not decide beforehand which property he will measure, but waits until he has received the particle and then chooses randomly. The timing of the experiment is arranged so that Alice and Bob do their measurements *at the same time* (or, to use the more precise language of relativity, in a causally disconnected manner). Therefore, the measurement which Alice performs cannot disturb the result of Bob's measurement (or vice versa), since physical influences cannot propagate faster than light.

Figure 2.4. Schematic experimental setup for the Bell inequalities. Alice can choose to measure either Q or R, and Bob chooses to measure either S or T. They perform their measurements simultaneously. Alice and Bob are assumed to be far enough apart that performing a measurement on one system can not have any effect on the result of measurements on the other.

We are going to do some simple algebra with the quantity $QS + RS + RT - QT$. Notice that

$$QS + RS + RT - QT = (Q + R)S + (R - Q)T. \tag{2.219}$$

Because $R, Q = \pm 1$ it follows that either $(Q + R)S = 0$ or $(R - Q)T = 0$. In either case, it is easy to see from (2.219) that $QS + RS + RT - QT = \pm 2$. Suppose next that $p(q, r, s, t)$ is the probability that, before the measurements are performed, the system is in a state where $Q = q, R = r, S = s$, and $T = t$. These probabilities may depend on how Charlie performs his preparation, and on experimental noise. Letting $\mathbf{E}(\cdot)$ denote the mean value of a quantity, we have

$$\begin{aligned}
\mathbf{E}(QS + RS + RT - QT) &= \sum_{qrst} p(q, r, s, t)(qs + rs + rt - qt) && (2.220)\\
&\le \sum_{qrst} p(q, r, s, t) \times 2 && (2.221)\\
&= 2. && (2.222)
\end{aligned}$$

Also,

$$\begin{aligned}
\mathbf{E}(QS + RS + RT - QT) &= \sum_{qrst} p(q, r, s, t)qs + \sum_{qrst} p(q, r, s, t)rs \\
&\quad + \sum_{qrst} p(q, r, s, t)rt - \sum_{qrst} p(q, r, s, t)qt && (2.223)\\
&= \mathbf{E}(QS) + \mathbf{E}(RS) + \mathbf{E}(RT) - \mathbf{E}(QT). && (2.224)
\end{aligned}$$

Comparing (2.222) and (2.224) we obtain the *Bell inequality*,

$$\mathbf{E}(QS) + \mathbf{E}(RS) + \mathbf{E}(RT) - \mathbf{E}(QT) \le 2. \tag{2.225}$$

This result is also often known as the *CHSH inequality* after the initials of its four discoverers. It is part of a larger set of inequalities known generically as Bell inequalities, since the first was found by John Bell.

By repeating the experiment many times, Alice and Bob can determine each quantity on the left hand side of the Bell inequality. For example, after finishing a set of experiments, Alice and Bob get together to analyze their data. They look at all the experiments where Alice measured P_Q and Bob measured P_S. By multiplying the results of their experiments together, they get a sample of values for QS. By averaging over this sample, they can estimate $\mathbf{E}(QS)$ to an accuracy only limited by the number of experiments which they perform. Similarly, they can estimate all the other quantities on the left hand side of the Bell inequality, and thus check to see whether it is obeyed in a real experiment.

It's time to put some quantum mechanics back in the picture. Imagine we perform the following quantum mechanical experiment. Charlie prepares a quantum system of two qubits in the state

$$|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \tag{2.226}$$

He passes the first qubit to Alice, and the second qubit to Bob. They perform measurements of the following observables:

$$Q = Z_1 \qquad S = \frac{-Z_2 - X_2}{\sqrt{2}} \tag{2.227}$$

$$R = X_1 \qquad T = \frac{Z_2 - X_2}{\sqrt{2}}. \tag{2.228}$$

Simple calculations show that the average values for these observables, written in the quantum mechanical $\langle \cdot \rangle$ notation, are:

$$\langle QS \rangle = \frac{1}{\sqrt{2}};\ \langle RS \rangle = \frac{1}{\sqrt{2}};\ \langle RT \rangle = \frac{1}{\sqrt{2}};\ \langle QT \rangle = -\frac{1}{\sqrt{2}}. \tag{2.229}$$

Thus,

$$\langle QS \rangle + \langle RS \rangle + \langle RT \rangle - \langle QT \rangle = 2\sqrt{2}. \tag{2.230}$$

Hold on! We learned back in (2.225) that the average value of QS plus the average value of RS plus the average value of RT minus the average value of QT can never exceed two. Yet here, quantum mechanics predicts that this sum of averages yields $2\sqrt{2}$!

Fortunately, we can ask Nature to resolve the apparent paradox for us. Clever experiments using photons – particles of light – have been done to check the prediction (2.230) of quantum mechanics versus the Bell inequality (2.225) which we were led to by our common sense reasoning. The details of the experiments are outside the scope of the book, but the results were resoundingly in favor of the quantum mechanical prediction. The Bell inequality (2.225) is *not* obeyed by Nature.

What does this mean? It means that one or more of the assumptions that went into the derivation of the Bell inequality must be incorrect. Vast tomes have been written analyzing the various forms in which this type of argument can be made, and analyzing the subtly different assumptions which must be made to reach Bell-like inequalities. Here we merely summarize the main points.

There are two assumptions made in the proof of (2.225) which are questionable:

(1) The assumption that the physical properties P_Q, P_R, P_S, P_T have definite values Q, R, S, T which exist independent of observation. This is sometimes known as the assumption of *realism*.
(2) The assumption that Alice performing her measurement does not influence the result of Bob's measurement. This is sometimes known as the assumption of *locality*.

These two assumptions together are known as the assumptions of *local realism*. They are certainly intuitively plausible assumptions about how the world works, and they fit our everyday experience. Yet the Bell inequalities show that at least one of these assumptions is not correct.

What can we learn from Bell's inequality? For physicists, the most important lesson is that their deeply held commonsense intuitions about how the world works are wrong. The world is *not* locally realistic. Most physicists take the point of view that it is the assumption of realism which needs to be dropped from our worldview in quantum mechanics, although others have argued that the assumption of locality should be dropped instead. Regardless, Bell's inequality together with substantial experimental evidence now points to the conclusion that either or both of locality and realism must be dropped from our view of the world if we are to develop a good intuitive understanding of quantum mechanics.

What lessons can the fields of quantum computation and quantum information learn from Bell's inequality? Historically the most useful lesson has perhaps also been the most vague: there is something profoundly 'up' with entangled states like the EPR state. A lot of mileage in quantum computation and, especially, quantum information, has come from asking the simple question: 'what would some entanglement buy me in this problem?' As we saw in teleportation and superdense coding, and as we will see repeatedly later in the book, by throwing some entanglement into a problem we open up a new world of possibilities unimaginable with classical information. The bigger picture is that Bell's inequality teaches us that entanglement is a fundamentally new resource in the world that goes essentially *beyond* classical resources; iron to the classical world's bronze age. A major task of quantum computation and quantum information is to exploit this new resource to do information processing tasks impossible or much more difficult with classical resources.

Problem 2.1: (Functions of the Pauli matrices) Let $f(\cdot)$ be any function from complex numbers to complex numbers. Let $\vec{n}$ be a normalized vector in three dimensions, and let θ be real. Show that

$$f\left(\theta \vec{n} \cdot \vec{\sigma}\right) = \frac{f(\theta) + f(-\theta)}{2} I + \frac{f(\theta) - f(-\theta)}{2} \vec{n} \cdot \vec{\sigma}. \tag{2.231}$$

Problem 2.2: (Properties of the Schmidt number) Suppose $|\psi\rangle$ is a pure state of a composite system with components A and B.

(1) Prove that the Schmidt number of $|\psi\rangle$ is equal to the rank of the reduced density matrix $\rho_A \equiv \mathrm{tr}_B(|\psi\rangle\langle\psi|)$. (Note that the rank of a Hermitian operator is equal to the dimension of its support.)
(2) Suppose $|\psi\rangle = \sum_j |\alpha_j\rangle|\beta_j\rangle$ is a representation for $|\psi\rangle$, where $|\alpha_j\rangle$ and $|\beta_j\rangle$ are (un-normalized) states for systems A and B, respectively. Prove that the

number of terms in such a decomposition is greater than or equal to the Schmidt number of $|\psi\rangle$, $\mathrm{Sch}(\psi)$.

(3) Suppose $|\psi\rangle = \alpha|\varphi\rangle + \beta|\gamma\rangle$. Prove that

$$\mathrm{Sch}(\psi) \geq |\mathrm{Sch}(\varphi) - \mathrm{Sch}(\gamma)|\,. \tag{2.232}$$

Problem 2.3: (Tsirelson's inequality) Suppose $Q = \vec{q} \cdot \vec{\sigma}, R = \vec{r} \cdot \vec{\sigma}, S = \vec{s} \cdot \vec{\sigma}, T = \vec{t} \cdot \vec{\sigma}$, where $\vec{q}, \vec{r}, \vec{s}$ and $\vec{t}$ are real unit vectors in three dimensions. Show that

$$(Q \otimes S + R \otimes S + R \otimes T - Q \otimes T)^2 = 4I + [Q, R] \otimes [S, T]. \tag{2.233}$$

Use this result to prove that

$$\langle Q \otimes S\rangle + \langle R \otimes S\rangle + \langle R \otimes T\rangle - \langle Q \otimes T\rangle \leq 2\sqrt{2}, \tag{2.234}$$

so the violation of the Bell inequality found in Equation (2.230) is the maximum possible in quantum mechanics.

History and further reading

There are an enormous number of books on linear algebra at levels ranging from High School through to Graduate School. Perhaps our favorites are the two volume set by Horn and Johnson[HJ85, HJ91], which cover an extensive range of topics in an accessible manner. Other useful references include Marcus and Minc[MM92], and Bhatia[Bha97]. Good introductions to linear algebra include Halmos[Hal58], Perlis[Per52], and Strang[Str76].

There are many excellent books on quantum mechanics. Unfortunately, most of these books focus on topics of tangential interest to quantum information and computation. Perhaps the most relevant in the existing literature is Peres' superb book[Per93]. Beside an extremely clear exposition of elementary quantum mechanics, Peres gives an extensive discussion of the Bell inequalities and related results. Good introductory level texts include Sakurai's book[Sak95], Volume III of the superb series by Feynman, Leighton, and Sands[FLS65a], and the two volume work by Cohen-Tannoudji, Diu and Laloë[CTDL77a, CTDL77b]. All three of these works are somewhat closer in spirit to quantum computation and quantum information than are most other quantum mechanics texts, although the great bulk of each is still taken up by applications far removed from quantum computation and quantum information. As a result, none of these texts need be read in detail by someone interested in learning about quantum computation and quantum information. However, any one of these texts may prove handy as a reference, especially when reading articles by physicists. References for the history of quantum mechanics may be found at the end of Chapter 1.

Many texts on quantum mechanics deal only with projective measurements. For applications to quantum computing and quantum information it is more convenient – and, we believe, easier for novices – to start with the general description of measurements, of which projective measurements can be regarded as a special case. Of course, ultimately, as we have shown, the two approaches are equivalent. The theory of generalized measurements which we have employed was developed between the 1940s and 1970s. Much of the history can be distilled from the book of Kraus[Kra83]. Interesting discussion of quantum measurements may be found in Section 2.2 of Gardiner[Gar91], and in the book by Braginsky and Khahili[BK92]. The POVM measurement for distinguishing non-orthogonal states described in Section 2.2.6 is due to Peres[Per88]. The extension described in Exercise 2.64 appeared in Duan and Guo[DG98].

Superdense coding was invented by Bennett and Wiesner[BW92]. An experiment implementing a variant of superdense coding using entangled photon pairs was performed by Mattle, Weinfurter, Kwiat, and Zeilinger[MWKZ96].

The density operator formalism was introduced independently by Landau[Lan27] and by von Neumann[von27]. The unitary freedom in the ensemble for density matrices, Theorem 2.6, was first pointed out by Schrödinger[Sch36], and was later rediscovered and extended by Jaynes[Jay57] and by Hughston, Jozsa and Wootters[HJW93]. The result of Exercise 2.73 is from the paper by Jaynes, and the results of Exercises 2.81 and 2.82 appear in the paper by Hughston, Jozsa and Wootters. The class of probability distributions which may appear in a density matrix decomposition for a given density matrix has been studied by Uhlmann[Uhl70] and by Nielsen[Nie99b]. Schmidt's eponymous decomposition appeared in[Sch06]. The result of Exercise 2.77 was noted by Peres[Per95].

The EPR thought experiment is due to Einstein, Podolsky and Rosen[EPR35], and was recast in essentially the form we have given here by Bohm[Boh51]. It is sometimes misleadingly referred to as the EPR 'paradox'. The Bell inequality is named in honour of Bell[Bel64], who first derived inequalities of this type. The form we have presented is due to Clauser, Horne, Shimony, and Holt[CHSH69], and is often known as the CHSH inequality. This inequality was derived independently by Bell, who did not publish the result.

Part 3 of Problem 2.2 is due to Thapliyal (private communication). Tsirelson's inequality is due to Tsirelson[Tsi80].