to be:

$$00 : |\psi\rangle \rightarrow \frac{|00\rangle + |11\rangle}{\sqrt{2}} \tag{2.134}$$

$$01 : |\psi\rangle \rightarrow \frac{|00\rangle - |11\rangle}{\sqrt{2}} \tag{2.135}$$

$$10 : |\psi\rangle \rightarrow \frac{|10\rangle + |01\rangle}{\sqrt{2}} \tag{2.136}$$

$$11 : |\psi\rangle \rightarrow \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \tag{2.137}$$

As we noted in Section 1.3.6, these four states are known as the *Bell basis*, *Bell states*, or *EPR pairs*, in honor of several of the pioneers who first appreciated the novelty of entanglement. Notice that the Bell states form an orthonormal basis, and can therefore be distinguished by an appropriate quantum measurement. If Alice sends her qubit to Bob, giving Bob possession of both qubits, then by doing a measurement in the Bell basis Bob can determine which of the four possible bit strings Alice sent.

Summarizing, Alice, interacting with only a single qubit, is able to transmit two bits of information to Bob. Of course, two qubits are involved in the protocol, but Alice never need interact with the second qubit. Classically, the task Alice accomplishes would have been impossible had she only transmitted a single classical bit, as we will show in Chapter 12. Furthermore, this remarkable superdense coding protocol has received partial verification in the laboratory. (See 'History and further reading' for references to the experimental verification.) In later chapters we will see many other examples, some of them much more spectacular than superdense coding, of quantum mechanics being harnessed to perform information processing tasks. However, a key point can already be seen in this beautiful example: information is physical, and surprising physical theories such as quantum mechanics may predict surprising information processing abilities.

Exercise 2.69: Verify that the Bell basis forms an orthonormal basis for the two qubit state space.

Exercise 2.70: Suppose E is any positive operator acting on Alice's qubit. Show that $\langle\psi|E \otimes I|\psi\rangle$ *takes the same value* when $|\psi\rangle$ is any of the four Bell states. Suppose some malevolent third party ('Eve') intercepts Alice's qubit on the way to Bob in the superdense coding protocol. Can Eve infer anything about which of the four possible bit strings $00, 01, 10, 11$ Alice is trying to send? If so, how, or if not, why not?

2.4 The density operator

We have formulated quantum mechanics using the language of state vectors. An alternate formulation is possible using a tool known as the *density operator* or *density matrix*. This alternate formulation is mathematically equivalent to the state vector approach, but it provides a much more convenient language for thinking about some commonly encountered scenarios in quantum mechanics. The next three sections describe the density operator formulation of quantum mechanics. Section 2.4.1 introduces the density operator using the concept of an ensemble of quantum states. Section 2.4.2 develops some general properties of the density operator. Finally, Section 2.4.3 describes an application where the density operator really shines – as a tool for the description of *individual subsystems* of a composite quantum system.

2.4.1 Ensembles of quantum states

The density operator language provides a convenient means for describing quantum systems whose state is not completely known. More precisely, suppose a quantum system is in one of a number of states $|\psi_i\rangle$, where i is an index, with respective probabilities p_i. We shall call $\{p_i, |\psi_i\rangle\}$ an *ensemble of pure states*. The density operator for the system is defined by the equation

$$\rho \equiv \sum_i p_i |\psi_i\rangle\langle\psi_i|. \tag{2.138}$$

The density operator is often known as the *density matrix*; we will use the two terms interchangeably. It turns out that all the postulates of quantum mechanics can be reformulated in terms of the density operator language. The purpose of this section and the next is to explain how to perform this reformulation, and explain when it is useful. Whether one uses the density operator language or the state vector language is a matter of taste, since both give the same results; however it is sometimes much easier to approach problems from one point of view rather than the other.

Suppose, for example, that the evolution of a closed quantum system is described by the unitary operator U. If the system was initially in the state $|\psi_i\rangle$ with probability p_i then after the evolution has occurred the system will be in the state $U|\psi_i\rangle$ with probability p_i. Thus, the evolution of the density operator is described by the equation

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \xrightarrow{U} \sum_i p_i U|\psi_i\rangle\langle\psi_i|U^\dagger = U\rho U^\dagger. \tag{2.139}$$

Measurements are also easily described in the density operator language. Suppose we perform a measurement described by measurement operators M_m. If the initial state was $|\psi_i\rangle$, then the probability of getting result m is

$$p(m|i) = \langle\psi_i|M_m^\dagger M_m|\psi_i\rangle = \mathrm{tr}(M_m^\dagger M_m|\psi_i\rangle\langle\psi_i|), \tag{2.140}$$

where we have used Equation (2.61) to obtain the last equality. By the law of total probability (see Appendix 1 for an explanation of this and other elementary notions of probability theory) the probability of obtaining result m is

$$\begin{aligned} p(m) &= \sum_i p(m|i) p_i & (2.141) \\ &= \sum_i p_i \mathrm{tr}(M_m^\dagger M_m |\psi_i\rangle\langle\psi_i|) & (2.142) \\ &= \mathrm{tr}(M_m^\dagger M_m \rho). & (2.143) \end{aligned}$$

What is the density operator of the system after obtaining the measurement result m? If the initial state was $|\psi_i\rangle$ then the state after obtaining the result m is

$$|\psi_i^m\rangle = \frac{M_m|\psi_i\rangle}{\sqrt{\langle\psi_i|M_m^\dagger M_m|\psi_i\rangle}}. \tag{2.144}$$

Thus, after a measurement which yields the result m we have an ensemble of states $|\psi_i^m\rangle$ with respective probabilities $p(i|m)$. The corresponding density operator ρ_m is therefore

$$\rho_m = \sum_i p(i|m)|\psi_i^m\rangle\langle\psi_i^m| = \sum_i p(i|m)\frac{M_m|\psi_i\rangle\langle\psi_i|M_m^\dagger}{\langle\psi_i|M_m^\dagger M_m|\psi_i\rangle}. \tag{2.145}$$

But by elementary probability theory, $p(i|m) = p(m,i)/p(m) = p(m|i)p_i/p(m)$. Substituting from (2.143) and (2.140) we obtain

$$\rho_m = \sum_i p_i \frac{M_m|\psi_i\rangle\langle\psi_i|M_m^\dagger}{\mathrm{tr}(M_m^\dagger M_m\rho)} \tag{2.146}$$

$$= \frac{M_m\rho M_m^\dagger}{\mathrm{tr}(M_m^\dagger M_m\rho)}. \tag{2.147}$$

What we have shown is that the basic postulates of quantum mechanics related to unitary evolution and measurement can be rephrased in the language of density operators. In the next section we complete this rephrasing by giving an intrinsic characterization of the density operator that does not rely on the idea of a state vector.

Before doing so, however, it is useful to introduce some more language, and one more fact about the density operator. First, the language. A quantum system whose state $|\psi\rangle$ is known exactly is said to be in a *pure state*. In this case the density operator is simply $\rho = |\psi\rangle\langle\psi|$. Otherwise, ρ is in a *mixed state*; it is said to be a *mixture* of the different pure states in the ensemble for ρ. In the exercises you will be asked to demonstrate a simple criterion for determining whether a state is pure or mixed: a pure state satisfies $\mathrm{tr}(\rho^2) = 1$, while a mixed state satisfies $\mathrm{tr}(\rho^2) < 1$. A few words of warning about the nomenclature: sometimes people use the term 'mixed state' as a catch-all to include both pure and mixed quantum states. The origin for this usage seems to be that it implies that the writer is not necessarily *assuming* that a state is pure. Second, the term 'pure state' is often used in reference to a state vector $|\psi\rangle$, to distinguish it from a density operator ρ.

Finally, imagine a quantum system is prepared in the state ρ_i with probability p_i. It is not difficult to convince yourself that the system may be described by the density matrix $\sum_i p_i\rho_i$. A proof of this is to suppose that ρ_i arises from some ensemble $\{p_{ij}, |\psi_{ij}\rangle\}$ (note that i is fixed) of pure states, so the probability for being in the state $|\psi_{ij}\rangle$ is p_ip_{ij}. The density matrix for the system is thus

$$\rho = \sum_{ij} p_ip_{ij}|\psi_{ij}\rangle\langle\psi_{ij}| \tag{2.148}$$

$$= \sum_i p_i\rho_i, \tag{2.149}$$

where we have used the definition $\rho_i = \sum_j p_{ij}|\psi_{ij}\rangle\langle\psi_{ij}|$. We say that ρ is a *mixture* of the states ρ_i with probabilities p_i. This concept of a mixture comes up repeatedly in the analysis of problems like quantum noise, where the effect of the noise is to introduce ignorance into our knowledge of the quantum state. A simple example is provided by the measurement scenario described above. Imagine that, for some reason, our record of the result m of the measurement was lost. We would have a quantum system in the state ρ_m with probability $p(m)$, but would no longer know the actual value of m. The state of such a quantum system would therefore be described by the density operator

$$\rho = \sum_m p(m)\rho_m \tag{2.150}$$

$$= \sum_m \mathrm{tr}(M_m^\dagger M_m \rho)\frac{M_m \rho M_m^\dagger}{\mathrm{tr}(M_m^\dagger M_m \rho)} \tag{2.151}$$

$$= \sum_m M_m \rho M_m^\dagger, \tag{2.152}$$

a nice compact formula which may be used as the starting point for analysis of further operations on the system.

2.4.2 General properties of the density operator

The density operator was introduced as a means of describing ensembles of quantum states. In this section we move away from this description to develop an intrinsic characterization of density operators that does not rely on an ensemble interpretation. This allows us to complete the program of giving a description of quantum mechanics that does not take as its foundation the state vector. We also take the opportunity to develop numerous other elementary properties of the density operator.

The class of operators that are density operators are characterized by the following useful theorem:

Theorem 2.5: (**Characterization of density operators**) An operator ρ is the density operator associated to some ensemble $\{p_i, |\psi_i\rangle\}$ if and only if it satisfies the conditions:

(1) **(Trace condition)** ρ has trace equal to one.
(2) **(Positivity condition)** ρ is a positive operator.

Proof

Suppose $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ is a density operator. Then

$$\mathrm{tr}(\rho) = \sum_i p_i \mathrm{tr}(|\psi_i\rangle\langle\psi_i|) = \sum_i p_i = 1, \tag{2.153}$$

so the trace condition $\mathrm{tr}(\rho) = 1$ is satisfied. Suppose $|\varphi\rangle$ is an arbitrary vector in state space. Then

$$\langle\varphi|\rho|\varphi\rangle = \sum_i p_i \langle\varphi|\psi_i\rangle\langle\psi_i|\varphi\rangle \tag{2.154}$$

$$= \sum_i p_i |\langle\varphi|\psi_i\rangle|^2 \tag{2.155}$$

$$\geq 0, \tag{2.156}$$

so the positivity condition is satisfied.

Conversely, suppose ρ is any operator satisfying the trace and positivity conditions. Since ρ is positive, it must have a spectral decomposition

$$\rho = \sum_j \lambda_j |j\rangle\langle j|, \tag{2.157}$$

where the vectors $|j\rangle$ are orthogonal, and λ_j are real, non-negative eigenvalues of ρ.

From the trace condition we see that $\sum_j \lambda_j = 1$. Therefore, a system in state $|j\rangle$ with probability λ_j will have density operator ρ. That is, the ensemble $\{\lambda_j, |j\rangle\}$ is an ensemble of states giving rise to the density operator ρ. $\square$

This theorem provides a characterization of density operators that is intrinsic to the operator itself: we can *define* a density operator to be a positive operator ρ which has trace equal to one. Making this definition allows us to reformulate the postulates of quantum mechanics in the density operator picture. For ease of reference we state all the reformulated postulates here:

> **Postulate 1**: Associated to any isolated physical system is a complex vector space with inner product (that is, a Hilbert space) known as the *state space* of the system. The system is completely described by its *density operator*, which is a positive operator ρ with trace one, acting on the state space of the system. If a quantum system is in the state ρ_i with probability p_i, then the density operator for the system is $\sum_i p_i \rho_i$.
>
> **Postulate 2**: The evolution of a *closed* quantum system is described by a *unitary transformation*. That is, the state ρ of the system at time t_1 is related to the state ρ' of the system at time t_2 by a unitary operator U which depends only on the times t_1 and t_2,
>
> $$\rho' = U\rho U^\dagger. \tag{2.158}$$
>
> **Postulate 3**: Quantum measurements are described by a collection $\{M_m\}$ of *measurement operators*. These are operators acting on the state space of the system being measured. The index m refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is ρ immediately before the measurement then the probability that result m occurs is given by
>
> $$p(m) = \mathrm{tr}(M_m^\dagger M_m \rho), \tag{2.159}$$
>
> and the state of the system after the measurement is
>
> $$\frac{M_m \rho M_m^\dagger}{\mathrm{tr}(M_m^\dagger M_m \rho)}. \tag{2.160}$$
>
> The measurement operators satisfy the *completeness equation*,
>
> $$\sum_m M_m^\dagger M_m = I. \tag{2.161}$$
>
> **Postulate 4**: The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Moreover, if we have systems numbered 1 through n, and system number i is prepared in the state ρ_i, then the joint state of the total system is $\rho_1 \otimes \rho_2 \otimes \ldots \rho_n$.

These reformulations of the fundamental postulates of quantum mechanics in terms of the density operator are, of course, mathematically equivalent to the description in terms of the state vector. Nevertheless, as a way of thinking about quantum mechanics, the density operator approach really shines for two applications: the description of quantum systems whose state is not known, and the description of subsystems of a composite quantum system, as will be described in the next section. For the remainder of this section we flesh out the properties of the density matrix in more detail.

Exercise 2.71: (Criterion to decide if a state is mixed or pure) Let ρ be a density operator. Show that $\mathrm{tr}(\rho^2) \leq 1$, with equality if and only if ρ is a pure state.

It is a tempting (and surprisingly common) fallacy to suppose that the eigenvalues and eigenvectors of a density matrix have some special significance with regard to the ensemble of quantum states represented by that density matrix. For example, one might suppose that a quantum system with density matrix

$$\rho = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1| \,. \tag{2.162}$$

must be in the state $|0\rangle$ with probability $3/4$ and in the state $|1\rangle$ with probability $1/4$. However, this is not necessarily the case. Suppose we define

$$|a\rangle \equiv \sqrt{\frac{3}{4}}|0\rangle + \sqrt{\frac{1}{4}}|1\rangle \tag{2.163}$$

$$|b\rangle \equiv \sqrt{\frac{3}{4}}|0\rangle - \sqrt{\frac{1}{4}}|1\rangle, \tag{2.164}$$

and the quantum system is prepared in the state $|a\rangle$ with probability $1/2$ and in the state $|b\rangle$ with probability $1/2$. Then it is easily checked that the corresponding density matrix is

$$\rho = \frac{1}{2}|a\rangle\langle a| + \frac{1}{2}|b\rangle\langle b| = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1|. \tag{2.165}$$

That is, these two *different* ensembles of quantum states give rise to the *same* density matrix. In general, the eigenvectors and eigenvalues of a density matrix just indicate *one* of many possible ensembles that may give rise to a specific density matrix, and there is no reason to suppose it is an especially privileged ensemble.

A natural question to ask in the light of this discussion is what class of ensembles does give rise to a particular density matrix? The solution to this problem, which we now give, has surprisingly many applications in quantum computation and quantum information, notably in the understanding of quantum noise and quantum error-correction (Chapters 8 and 10). For the solution it is convenient to make use of vectors $|\tilde{\psi}_i\rangle$ which may not be normalized to unit length. We say the set $|\tilde{\psi}_i\rangle$ *generates* the operator $\rho \equiv \sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|$, and thus the connection to the usual ensemble picture of density operators is expressed by the equation $|\tilde{\psi}_i\rangle = \sqrt{p_i}|\psi_i\rangle$. When do two sets of vectors, $|\tilde{\psi}_i\rangle$ and $|\tilde{\varphi}_j\rangle$ generate the same operator ρ? The solution to this problem will enable us to answer the question of what ensembles give rise to a given density matrix.

Theorem 2.6: **(Unitary freedom in the ensemble for density matrices)** The sets $|\tilde{\psi}_i\rangle$ and $|\tilde{\varphi}_j\rangle$ generate the same density matrix if and only if

$$|\tilde{\psi}_i\rangle = \sum_j u_{ij}|\tilde{\varphi}_j\rangle \,, \tag{2.166}$$

where u_{ij} is a unitary matrix of complex numbers, with indices i and j, and we 'pad' whichever set of vectors $|\tilde{\psi}_i\rangle$ or $|\tilde{\varphi}_j\rangle$ is smaller with additional vectors 0 so that the two sets have the same number of elements.

As a consequence of the theorem, note that $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i| = \sum_j q_j|\varphi_j\rangle\langle\varphi_j|$ for *normalized* states $|\psi_i\rangle, |\varphi_j\rangle$ and probability distributions p_i and q_j if and only if

$$\sqrt{p_i}|\psi_i\rangle = \sum_j u_{ij}\sqrt{q_j}|\varphi_j\rangle, \tag{2.167}$$

for some unitary matrix u_{ij}, and we may pad the smaller ensemble with entries having probability zero in order to make the two ensembles the same size. Thus, Theorem 2.6 characterizes the freedom in ensembles $\{p_i, |\psi_i\rangle\}$ giving rise to a given density matrix ρ. Indeed, it is easily checked that our earlier example of a density matrix with two different decompositions, (2.162), arises as a special case of this general result. Let's turn now to the proof of the theorem.

Proof
Suppose $|\tilde{\psi}_i\rangle = \sum_j u_{ij}|\tilde{\varphi}_j\rangle$ for some unitary u_{ij}. Then

$$\sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i| = \sum_{ijk} u_{ij}u_{ik}^*|\tilde{\varphi}_j\rangle\langle\tilde{\varphi}_k| \tag{2.168}$$

$$= \sum_{jk}\left(\sum_i u_{ki}^\dagger u_{ij}\right)|\tilde{\varphi}_j\rangle\langle\tilde{\varphi}_k| \tag{2.169}$$

$$= \sum_{jk}\delta_{kj}|\tilde{\varphi}_j\rangle\langle\tilde{\varphi}_k| \tag{2.170}$$

$$= \sum_j |\tilde{\varphi}_j\rangle\langle\tilde{\varphi}_j|, \tag{2.171}$$

which shows that $|\tilde{\psi}_i\rangle$ and $|\tilde{\varphi}_j\rangle$ generate the same operator.

Conversely, suppose

$$A = \sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i| = \sum_j |\tilde{\varphi}_j\rangle\langle\tilde{\varphi}_j|\,. \tag{2.172}$$

Let $A = \sum_k \lambda_k|k\rangle\langle k|$ be a decomposition for A such that the states $|k\rangle$ are orthonormal, and the λ_k are strictly positive. Our strategy is to relate the states $|\tilde{\psi}_i\rangle$ to the states $|\tilde{k}\rangle \equiv \sqrt{\lambda_k}|k\rangle$, and similarly relate the states $|\tilde{\varphi}_j\rangle$ to the states $|\tilde{k}\rangle$. Combining the two relations will give the result. Let $|\psi\rangle$ be any vector orthonormal to the space spanned by the $|\tilde{k}\rangle$, so $\langle\psi|\tilde{k}\rangle\langle\tilde{k}|\psi\rangle = 0$ for all k, and thus we see that

$$0 = \langle\psi|A|\psi\rangle = \sum_i \langle\psi|\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|\psi\rangle = \sum_i |\langle\psi|\tilde{\psi}_i\rangle|^2. \tag{2.173}$$

Thus $\langle\psi|\tilde{\psi}_i\rangle = 0$ for all i and all $|\psi\rangle$ orthonormal to the space spanned by the $|\tilde{k}\rangle$. It follows that each $|\tilde{\psi}_i\rangle$ can be expressed as a linear combination of the $|\tilde{k}\rangle$, $|\tilde{\psi}_i\rangle = \sum_k c_{ik}|\tilde{k}\rangle$. Since $A = \sum_k |\tilde{k}\rangle\langle\tilde{k}| = \sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|$ we see that

$$\sum_k |\tilde{k}\rangle\langle\tilde{k}| = \sum_{kl}\left(\sum_i c_{ik}c_{il}^*\right)|\tilde{k}\rangle\langle\tilde{l}|. \tag{2.174}$$

The operators $|\tilde{k}\rangle\langle\tilde{l}|$ are easily seen to be linearly independent, and thus it must be that

$\sum_i c_{ik} c_{il}^* = \delta_{kl}$. This ensures that we may append extra columns to c to obtain a unitary matrix v such that $|\tilde{\psi}_i\rangle = \sum_k v_{ik}|\tilde{k}\rangle$, where we have appended zero vectors to the list of $|\tilde{k}\rangle$. Similarly, we can find a unitary matrix w such that $|\tilde{\varphi}_j\rangle = \sum_k w_{jk}|\tilde{k}\rangle$. Thus $|\tilde{\psi}_i\rangle = \sum_j u_{ij}|\tilde{\varphi}_j\rangle$, where $u = vw^\dagger$ is unitary. □

Exercise 2.72: (Bloch sphere for mixed states) The Bloch sphere picture for pure states of a single qubit was introduced in Section 1.2. This description has an important generalization to mixed states as follows.

(1) Show that an arbitrary density matrix for a mixed state qubit may be written as

$$\rho = \frac{I + \vec{r} \cdot \vec{\sigma}}{2}, \tag{2.175}$$

where $\vec{r}$ is a real three-dimensional vector such that $\|\vec{r}\| \leq 1$. This vector is known as the *Bloch vector* for the state ρ.

(2) What is the Bloch vector representation for the state $\rho = I/2$?

(3) Show that a state ρ is pure if and only if $\|\vec{r}\| = 1$.

(4) Show that for pure states the description of the Bloch vector we have given coincides with that in Section 1.2.

Exercise 2.73: Let ρ be a density operator. A *minimal ensemble* for ρ is an ensemble $\{p_i, |\psi_i\rangle\}$ containing a number of elements equal to the rank of ρ. Let $|\psi\rangle$ be any state in the support of ρ. (The *support* of a Hermitian operator A is the vector space spanned by the eigenvectors of A with non-zero eigenvalues.) Show that there is a minimal ensemble for ρ that contains $|\psi\rangle$, and moreover that in any such ensemble $|\psi\rangle$ must appear with probability

$$p_i = \frac{1}{\langle\psi_i|\rho^{-1}|\psi_i\rangle}, \tag{2.176}$$

where ρ^{-1} is defined to be the inverse of ρ, when ρ is considered as an operator acting only on the support of ρ. (This definition removes the problem that ρ may not have an inverse.)

2.4.3 The reduced density operator

Perhaps the deepest application of the density operator is as a descriptive tool for *subsystems* of a composite quantum system. Such a description is provided by the *reduced density operator*, which is the subject of this section. The reduced density operator is so useful as to be virtually indispensable in the analysis of composite quantum systems.

Suppose we have physical systems A and B, whose state is described by a density operator ρ^{AB}. The reduced density operator for system A is defined by

$$\rho^A \equiv \mathrm{tr}_B(\rho^{AB}), \tag{2.177}$$

where tr_B is a map of operators known as the *partial trace* over system B. The partial trace is defined by

$$\mathrm{tr}_B\left(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|\right) \equiv |a_1\rangle\langle a_2|\,\mathrm{tr}(|b_1\rangle\langle b_2|), \tag{2.178}$$

where $|a_1\rangle$ and $|a_2\rangle$ are any two vectors in the state space of A, and $|b_1\rangle$ and $|b_2\rangle$ are any two vectors in the state space of B. The trace operation appearing on the right hand side is the usual trace operation for system B, so $\mathrm{tr}(|b_1\rangle\langle b_2|) = \langle b_2|b_1\rangle$. We have defined the partial trace operation only on a special subclass of operators on AB; the specification is completed by requiring in addition to Equation (2.178) that the partial trace be linear in its input.

It is not obvious that the reduced density operator for system A is in any sense a description for the state of system A. The physical justification for making this identification is that the reduced density operator provides the correct measurement statistics for measurements made on system A. This is explained in more detail in Box 2.6 on page 107. The following simple example calculations may also help understand the reduced density operator. First, suppose a quantum system is in the product state $\rho^{AB} = \rho \otimes \sigma$, where ρ is a density operator for system A, and σ is a density operator for system B. Then

$$\rho^A = \mathrm{tr}_B(\rho \otimes \sigma) = \rho\,\mathrm{tr}(\sigma) = \rho, \tag{2.184}$$

which is the result we intuitively expect. Similarly, $\rho^B = \sigma$ for this state. A less trivial example is the Bell state $(|00\rangle + |11\rangle)/\sqrt{2}$. This has density operator

$$\begin{aligned}
\rho &= \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)\left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) &(2.185)\\
&= \frac{|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|}{2}. &(2.186)
\end{aligned}$$

Tracing out the second qubit, we find the reduced density operator of the first qubit,

$$\begin{aligned}
\rho^1 &= \mathrm{tr}_2(\rho) &(2.187)\\
&= \frac{\mathrm{tr}_2(|00\rangle\langle 00|) + \mathrm{tr}_2(|11\rangle\langle 00|) + \mathrm{tr}_2(|00\rangle\langle 11|) + \mathrm{tr}_2(|11\rangle\langle 11|)}{2} &(2.188)\\
&= \frac{|0\rangle\langle 0|\langle 0|0\rangle + |1\rangle\langle 0|\langle 0|1\rangle + |0\rangle\langle 1|\langle 1|0\rangle + |1\rangle\langle 1|\langle 1|1\rangle}{2} &(2.189)\\
&= \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} &(2.190)\\
&= \frac{I}{2}. &(2.191)
\end{aligned}$$

Notice that this state is a *mixed state*, since $\mathrm{tr}((I/2)^2) = 1/2 < 1$. This is quite a remarkable result. The state of the joint system of two qubits is a pure state, that is, it is known *exactly*; however, the first qubit is in a mixed state, that is, a state about which we apparently do not have maximal knowledge. This strange property, that the joint state of a system can be completely known, yet a subsystem be in mixed states, is another hallmark of quantum entanglement.

Exercise 2.74: Suppose a composite of systems A and B is in the state $|a\rangle|b\rangle$, where $|a\rangle$ is a pure state of system A, and $|b\rangle$ is a pure state of system B. Show that the reduced density operator of system A alone is a pure state.

Exercise 2.75: For each of the four Bell states, find the reduced density operator for each qubit.

Quantum teleportation and the reduced density operator

A useful application of the reduced density operator is to the analysis of quantum teleportation. Recall from Section 1.3.7 that quantum teleportation is a procedure for sending

Box 2.6: Why the partial trace?

Why is the partial trace used to describe part of a larger quantum system? The reason for doing this is because the partial trace operation is the *unique* operation which gives rise to the correct description of *observable* quantities for subsystems of a composite system, in the following sense.

Suppose M is any observable on system A, and we have some measuring device which is capable of realizing measurements of M. Let $\tilde{M}$ denote the corresponding observable for the same measurement, performed on the composite system AB. Our immediate goal is to argue that $\tilde{M}$ is necessarily equal to $M \otimes I_B$. Note that if the system AB is prepared in the state $|m\rangle|\psi\rangle$, where $|m\rangle$ is an eigenstate of M with eigenvalue m, and $|\psi\rangle$ is any state of B, then the measuring device must yield the result m for the measurement, with probability one. Thus, if P_m is the projector onto the m eigenspace of the observable M, then the corresponding projector for $\tilde{M}$ is $P_m \otimes I_B$. We therefore have

$$\tilde{M} = \sum_m m P_m \otimes I_B = M \otimes I_B. \tag{2.179}$$

The next step is to show that the partial trace procedure gives the correct measurement statistics for observations on part of a system. Suppose we perform a measurement on system A described by the observable M. Physical consistency requires that any prescription for associating a 'state', ρ^A, to system A, must have the property that measurement averages be the same whether computed via ρ^A or ρ^{AB},

$$\mathrm{tr}(M\rho^A) = \mathrm{tr}(\tilde{M}\rho^{AB}) = \mathrm{tr}((M \otimes I_B)\rho^{AB}). \tag{2.180}$$

This equation is certainly satisfied if we choose $\rho^A \equiv \mathrm{tr}_B(\rho^{AB})$. In fact, the partial trace turns out to be the *unique* function having this property. To see this uniqueness property, let $f(\cdot)$ be any map of density operators on AB to density operators on A such that

$$\mathrm{tr}(Mf(\rho^{AB})) = \mathrm{tr}((M \otimes I_B)\rho^{AB}), \tag{2.181}$$

for all observables M. Let M_i be an orthonormal basis of operators for the space of Hermitian operators with respect to the Hilbert–Schmidt inner product $(X, Y) \equiv \mathrm{tr}(XY)$ (compare Exercise 2.39 on page 76). Then expanding $f(\rho^{AB})$ in this basis gives

$$f(\rho^{AB}) = \sum_i M_i \mathrm{tr}(M_i f(\rho^{AB})) \tag{2.182}$$

$$= \sum_i M_i \mathrm{tr}((M_i \otimes I_B)\rho^{AB}). \tag{2.183}$$

It follows that f is uniquely determined by Equation (2.180). Moreover, the partial trace satisfies (2.180), so it is the unique function having this property.

quantum information from Alice to Bob, given that Alice and Bob share an EPR pair, and have a classical communications channel.

At first sight it appears as though teleportation can be used to do faster than light communication, a big no-no according to the theory of relativity. We surmised in Section 1.3.7 that what prevents faster than light communication is the need for Alice to communicate her measurement result to Bob. The reduced density operator allows us to make this rigorous.

Recall that immediately before Alice makes her measurement the quantum state of the three qubits is (Equation (1.32)):

$$\begin{aligned} |\psi_2\rangle = \frac{1}{2}\Big[&|00\rangle\big(\alpha|0\rangle + \beta|1\rangle\big) + |01\rangle\big(\alpha|1\rangle + \beta|0\rangle\big) \\ &+|10\rangle\big(\alpha|0\rangle - \beta|1\rangle\big) + |11\rangle\big(\alpha|1\rangle - \beta|0\rangle\big)\Big]. \end{aligned} \tag{2.192}$$

Measuring in Alice's computational basis, the state of the system after the measurement is:

$$|00\rangle\Big[\alpha|0\rangle + \beta|1\rangle\Big] \text{ with probability } \frac{1}{4} \tag{2.193}$$

$$|01\rangle\Big[\alpha|1\rangle + \beta|0\rangle\Big] \text{ with probability } \frac{1}{4} \tag{2.194}$$

$$|10\rangle\Big[\alpha|0\rangle - \beta|1\rangle\Big] \text{ with probability } \frac{1}{4} \tag{2.195}$$

$$|11\rangle\Big[\alpha|1\rangle - \beta|0\rangle\Big] \text{ with probability } \frac{1}{4}. \tag{2.196}$$

The density operator of the system is thus

$$\begin{aligned} \rho = \frac{1}{4}\Big[&|00\rangle\langle 00|(\alpha|0\rangle + \beta|1\rangle)(\alpha^*\langle 0| + \beta^*\langle 1|) + |01\rangle\langle 01|(\alpha|1\rangle + \beta|0\rangle)(\alpha^*\langle 1| + \beta^*\langle 0|) \\ &+|10\rangle\langle 10|(\alpha|0\rangle - \beta|1\rangle)(\alpha^*\langle 0| - \beta^*\langle 1|) + |11\rangle\langle 11|(\alpha|1\rangle - \beta|0\rangle)(\alpha^*\langle 1| - \beta^*\langle 0|)\Big]. \end{aligned} \tag{2.197}$$

Tracing out Alice's system, we see that the reduced density operator of Bob's system is

$$\begin{aligned} \rho^B &= \frac{1}{4}\Big[(\alpha|0\rangle + \beta|1\rangle)(\alpha^*\langle 0| + \beta^*\langle 1|) + (\alpha|1\rangle + \beta|0\rangle)(\alpha^*\langle 1| + \beta^*\langle 0|) \\ &\quad +(\alpha|0\rangle - \beta|1\rangle)(\alpha^*\langle 0| - \beta^*\langle 1|) + (\alpha|1\rangle - \beta|0\rangle)(\alpha^*\langle 1| - \beta^*\langle 0|)\Big] && (2.198) \\ &= \frac{2(|\alpha|^2 + |\beta|^2)|0\rangle\langle 0| + 2(|\alpha|^2 + |\beta|^2)|1\rangle\langle 1|}{4} && (2.199) \\ &= \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} && (2.200) \\ &= \frac{I}{2}, && (2.201) \end{aligned}$$

where we have used the completeness relation in the last line. Thus, the state of Bob's system *after* Alice has performed the measurement but *before* Bob has learned the measurement result is $I/2$. This state has no dependence upon the state $|\psi\rangle$ being teleported, and thus any measurements performed by Bob will contain no information about $|\psi\rangle$, thus preventing Alice from using teleportation to transmit information to Bob faster than light.