

## Chapter 4: Quantum States

In the previous chapters we built the mathematical setting of quantum theory: complex Hilbert spaces and the operators acting on them. We now ask a central question:

> What mathematical object represents the state of a quantum system?

At first, the answer seems simple: a quantum state is a vector in a Hilbert space. This is true for the most idealized kind of quantum state, called a pure state. But quantum information theory also needs a more flexible object: the density matrix, also called a density operator. Density matrices describe both pure states and mixed states, and they are the right language for measurement probabilities, noisy systems, subsystems of entangled systems, and eventually POVMs and Naimark dilation.

The purpose of this chapter is to make quantum states mathematically precise. We will study:

- pure states,
- mixed states,
- density matrices,
- the trace,
- the partial trace,
- purification,
- and the operational meaning of a state as a rule for producing probabilities.

The density-operator formalism is standard in quantum information because it treats uncertainty, entanglement, and subsystems in one unified language (Nielsen and Chuang, 2010; Watrous, 2018).

---

### 4.1 What should a quantum state do?

Before defining quantum states formally, let us ask what a state is supposed to accomplish.

In classical probability, if a coin is known to be fair, we describe it by the probability distribution

$$P(\text{heads}) = \frac{1}{2}, \quad P(\text{tails}) = \frac{1}{2}.$$

The state of our knowledge is a rule that assigns probabilities to possible observations.

Quantum theory is similar in spirit but different in structure. A quantum state should allow us to compute probabilities for possible measurement outcomes. However, unlike in classical probability, not all measurements can be thought of as simply revealing a pre-existing value. Different measurements may be incompatible, and the act of measurement can disturb the system. This is one reason the mathematical formalism of quantum states is richer than ordinary probability distributions (von Neumann, 1955; Nielsen and Chuang, 2010).

For now, we will use the following guiding idea:

> A quantum state is a mathematical object that determines the probabilities of all possible measurement outcomes.

This idea will become especially important when we reach POVMs. Naimark dilation is ultimately about generalized measurements, so we need a precise way to say what probabilities those measurements produce.

---

## 4.2 Pure states: state vectors and rays

Let  $\mathcal{H}$  be a finite-dimensional complex Hilbert space. A pure state vector is a unit vector

$$|\psi\rangle \in \mathcal{H}$$

satisfying

$$\langle\psi|\psi\rangle = 1.$$

The notation  $|\psi\rangle$  is called Dirac ket notation. It is just a convenient notation for a vector in a complex Hilbert space.

For example, the standard Hilbert space of a qubit is

$$\mathbb{C}^2.$$

The standard basis vectors are written

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

A general pure qubit state has the form

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle,$$

where  $\alpha, \beta \in \mathbb{C}$  and

$$|\alpha|^2 + |\beta|^2 = 1.$$

The normalization condition ensures that the probabilities predicted by the state sum to 1.

### Example: a superposition state

The vector

$$|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$$

is a pure qubit state because

$$\langle + | + \rangle = \frac{1}{2} + \frac{1}{2} = 1.$$

If this state is measured in the computational basis  $\{|0\rangle, |1\rangle\}$ , then the Born rule gives

$$P(0) = |\langle 0 | + \rangle|^2 = \frac{1}{2}, \quad P(1) = |\langle 1 | + \rangle|^2 = \frac{1}{2}.$$

The Born rule, in its standard form, assigns probabilities by squared inner-product magnitudes. This rule is one of the basic postulates of quantum mechanics and is used throughout quantum information theory (Nielsen and Chuang, 2010).

---

### 4.3 Global phase: why vectors are slightly redundant

There is a subtle point. The vectors

$$|\psi\rangle \quad \text{and} \quad e^{i\theta}|\psi\rangle$$

represent the same physical pure state for any real number  $\theta$ .

The complex number  $e^{i\theta}$  has absolute value 1. Multiplying a state vector by such a number changes its global phase.

Why does this not change the physical state? Suppose we compute the probability of obtaining a basis vector  $|\phi\rangle$ . Then

$$|\langle\phi|e^{i\theta}\psi\rangle|^2 = |e^{i\theta}\langle\phi|\psi\rangle|^2 = |e^{i\theta}|^2|\langle\phi|\psi\rangle|^2 = |\langle\phi|\psi\rangle|^2.$$

So all measurement probabilities remain the same.

Thus, strictly speaking, a pure state is not a unit vector itself, but a ray: the collection of all unit vectors that differ only by a global phase. In quantum information, however, we often use unit vectors as representatives of pure states because the notation is convenient.

---

### 4.4 From state vectors to rank-one projectors

For Naimark dilation and POVMs, it is more useful to represent pure states as operators.

Given a unit vector  $|\psi\rangle$ , define

$$|\psi\rangle\langle\psi|.$$

This is the operator that acts on a vector  $|v\rangle$  by

$$|\psi\rangle\langle\psi|v\rangle = |\psi\rangle\langle\psi|v\rangle.$$

Here  $\langle\psi|v\rangle$  is a scalar, so the result is that scalar times  $|\psi\rangle$ .

The operator

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

is the orthogonal projection onto the one-dimensional subspace spanned by  $|\psi\rangle$ . It is called a rank-one projector.

### Example

Let

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

Then

$$\langle\psi| = (\bar{\alpha} \quad \bar{\beta}),$$

and

$$|\psi\rangle\langle\psi| = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} (\bar{\alpha} \quad \bar{\beta}) = \begin{pmatrix} |\alpha|^2 & \alpha\bar{\beta} \\ \beta\bar{\alpha} & |\beta|^2 \end{pmatrix}.$$

For the state

$$|+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

we get

$$|+\rangle\langle+| = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

This matrix contains more than just the probabilities  $\frac{1}{2}, \frac{1}{2}$ . It also contains off-diagonal entries. These off-diagonal entries encode coherence, the phase-sensitive information that distinguishes a quantum superposition from an ordinary classical mixture.

---

## 4.5 The trace

To define density matrices properly, we need the trace.

For an  $n \times n$  matrix  $A$ , the trace is the sum of its diagonal entries:

$$\text{Tr}(A) = A_{11} + A_{22} + \cdots + A_{nn}.$$

For example,

$$A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$$

has trace

$$\text{Tr}(A) = 2 + 3 = 5.$$

The trace has several important properties.

First, it is linear:

$$\text{Tr}(A + B) = \text{Tr}(A) + \text{Tr}(B),$$

and

$$\text{Tr}(cA) = c \text{Tr}(A).$$

Second, for square matrices  $A$  and  $B$  of the same size,

$$\text{Tr}(AB) = \text{Tr}(BA).$$

This is called the cyclic property of the trace. More generally, for products of several matrices,

$$\text{Tr}(ABC) = \text{Tr}(BCA) = \text{Tr}(CAB),$$

as long as the matrix products are defined.

Third, if  $A$  is positive semidefinite, then

$$\text{Tr}(A) \geq 0.$$

This is because a positive semidefinite operator has nonnegative eigenvalues, and the trace equals the sum of the eigenvalues, counted with multiplicity.

For a pure state projector,

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

we have

$$\text{Tr}(\rho) = 1.$$

Indeed, if  $|\psi\rangle$  is a unit vector, then the projection onto its span has eigenvalue 1 in the direction  $|\psi\rangle$  and eigenvalue 0 on the orthogonal directions.

The trace is essential because it allows us to write quantum probabilities in a compact operator form.

---

## 4.6 Density matrices

A density matrix, or density operator, on a finite-dimensional Hilbert space  $\square$  is an operator  $\rho$  satisfying two conditions:

1.  $\rho$  is positive semidefinite:

$$\rho \geq 0.$$

2.  $\rho$  has trace 1:

$$\text{Tr}(\rho) = 1.$$

These two conditions are not arbitrary. They are exactly what we need for probabilities.

Positivity ensures that probabilities are never negative. Trace 1 ensures that total probability is normalized.

A pure state  $|\psi\rangle$  corresponds to the density matrix

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

This operator is positive semidefinite and has trace 1, so it is a density matrix.

But density matrices also describe situations where the state vector is not known exactly or where the system is part of a larger entangled system.

---

## 4.7 Mixed states as probabilistic ensembles

Suppose a source prepares one of several pure states

$$|\psi_1\rangle, \dots, |\psi_m\rangle$$

with probabilities

$$p_1, \dots, p_m,$$

where

$$p_i \geq 0, \quad \sum_{i=1}^m p_i = 1.$$

Then the corresponding density matrix is

$$\rho = \sum_{i=1}^m p_i |\psi_i\rangle\langle\psi_i|.$$

Such a state is called a mixed state if it cannot be represented by a single pure state projector.

**Example: a fair classical mixture of  $|0\rangle$  and  $|1\rangle$**

Suppose a qubit is prepared as  $|0\rangle$  with probability  $\frac{1}{2}$  and  $|1\rangle$  with probability  $\frac{1}{2}$ . Then

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|.$$

In matrix form,

$$|0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad |1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Therefore,

$$\rho = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} = \frac{I}{2}.$$

This state is called the maximally mixed qubit state. It represents a qubit with no preferred direction in Hilbert space.

Compare this with the pure state

$$|+\rangle\langle +| = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Both states give equal probabilities for outcomes 0 and 1 if measured in the computational basis. But they are not the same state. Their difference appears if we measure in another basis, such as the  $|+\rangle, |-\rangle$  basis.

For  $|+\rangle\langle +|$ , the outcome  $+$  occurs with probability 1.

For  $I/2$ , the outcome  $+$  occurs with probability  $\frac{1}{2}$ .

Thus, a superposition is not the same thing as a classical mixture.

---

## 4.8 Every density matrix is a mixture of pure states

In finite dimensions, every density matrix can be written as a probabilistic mixture of pure states.

Because  $\rho$  is positive semidefinite and self-adjoint, the spectral theorem gives an orthonormal basis of eigenvectors

$$|e_1\rangle, \dots, |e_n\rangle$$

and nonnegative eigenvalues

$$\lambda_1, \dots, \lambda_n$$

such that

$$\rho = \sum_{j=1}^n \lambda_j |e_j\rangle\langle e_j|.$$

Since

$$\text{Tr}(\rho) = 1,$$

we have

$$\sum_{j=1}^n \lambda_j = 1.$$

Therefore the eigenvalues form a probability distribution. So  $\rho$  is a mixture of the orthonormal pure states  $|e_j\rangle$ , prepared with probabilities  $\lambda_j$ .

This decomposition is called the spectral decomposition of  $\rho$ .

However, a density matrix may have many different ensemble decompositions. That is, the same  $\rho$  can often be written as a mixture of pure states in more than one way. Operationally, if two preparation procedures lead to the same density matrix, then no measurement performed only on that system can distinguish between those procedures (Nielsen and Chuang, 2010; Watrous, 2018).

### **Example: two different ensembles for the same density matrix**

We already saw that

$$\frac{I}{2} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|.$$

But also,

$$\frac{I}{2} = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|,$$

where

$$|-\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle.$$

Indeed,

$$|+\rangle\langle +| = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

and

$$|-\rangle\langle -| = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

So

$$\frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -| = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{I}{2}.$$

The same density matrix can arise from different preparation stories.

---

## 4.9 Pure versus mixed density matrices

How can we tell whether a density matrix represents a pure state?

A density matrix  $\rho$  is pure exactly when

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

for some unit vector  $|\psi\rangle$ . Equivalently,

$$\rho^2 = \rho.$$

This says that  $\rho$  is a projection.

Another useful test is:

$$\rho \text{ is pure} \iff \text{Tr}(\rho^2) = 1.$$

For a mixed state,

$$\text{Tr}(\rho^2) < 1.$$

The quantity

$$\text{Tr}(\rho^2)$$

is often called the purity of  $\rho$ .

### Why this test works

Let

$$\rho = \sum_{j=1}^n \lambda_j |e_j\rangle\langle e_j|$$

be the spectral decomposition of  $\rho$ . Then

$$\rho^2 = \sum_{j=1}^n \lambda_j^2 |e_j\rangle\langle e_j|.$$

Therefore,

$$\text{Tr}(\rho^2) = \sum_{j=1}^n \lambda_j^2.$$

Since the  $\lambda_j$ 's are nonnegative and sum to 1, we have

$$\sum_{j=1}^n \lambda_j^2 \leq 1.$$

Equality occurs exactly when one eigenvalue is 1 and all the others are 0. That means  $\rho$  has rank 1, so it is a pure state projector.

### Example

For

$$\rho = |+\rangle\langle+|,$$

we have

$$\rho^2 = \rho,$$

so

$$\text{Tr}(\rho^2) = \text{Tr}(\rho) = 1.$$

For

$$\rho = \frac{I}{2},$$

we have

$$\rho^2 = \frac{I}{4},$$

so

$$\text{Tr}(\rho^2) = \text{Tr}\left(\frac{I}{4}\right) = \frac{1}{2}.$$

Thus  $I/2$  is mixed.

---

## 4.10 The Born rule in trace form

Let us now connect density matrices to probabilities.

Suppose the system is in the pure state  $|\psi\rangle$ , and we ask whether the system lies in a subspace represented by an orthogonal projection  $P$ . The Born rule says that the probability of the “yes” outcome is

$$\langle\psi|P|\psi\rangle.$$

Using the density matrix

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

we can rewrite this as

$$\langle\psi|P|\psi\rangle = \text{Tr}(P\rho).$$

So the probability is

$$p = \text{Tr}(P\rho).$$

This formula also works for mixed states. If

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|,$$

then

$$\text{Tr}(P\rho) = \text{Tr} \left( P \sum_i p_i |\psi_i\rangle\langle\psi_i| \right).$$

By linearity of trace,

$$\text{Tr}(P\rho) = \sum_i p_i \text{Tr}(P|\psi_i\rangle\langle\psi_i|) = \sum_i p_i \langle\psi_i|P|\psi_i\rangle.$$

This is exactly the average probability over the ensemble.

Thus, in density-matrix language:

$$\text{Probability of outcome associated with } P = \text{Tr}(P\rho).$$

Later, when we define POVMs, the projection  $P$  will be replaced by a more general positive operator  $E$ , and the probability rule will become

$$p = \text{Tr}(E\rho).$$

This trace formula is one reason density matrices are the natural language for Naimark dilation.

---

## 4.11 Density matrices as probability-generating objects

It is tempting to think of a density matrix as a hidden description of what the system “really is.” For quantum information, a more productive viewpoint is operational:

> A density matrix is a compact rule for assigning probabilities to measurement outcomes.

If two physical preparation procedures produce the same density matrix  $\rho$ , then all measurements performed on that system give the same probability distributions. Conversely, in finite-dimensional quantum theory, the probabilities of sufficiently many measurements determine  $\rho$ . This is the basis of quantum state tomography, where one reconstructs an unknown state from measurement statistics (Nielsen and Chuang, 2010).

## Example: probability from a qubit density matrix

Let

$$\rho = \begin{pmatrix} 3/4 & 0 \\ 0 & 1/4 \end{pmatrix}.$$

If we measure in the computational basis, the projectors are

$$P_0 = |0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

and

$$P_1 = |1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then

$$p(0) = \text{Tr}(P_0\rho) = \text{Tr} \begin{pmatrix} 3/4 & 0 \\ 0 & 0 \end{pmatrix} = \frac{3}{4}.$$

Similarly,

$$p(1) = \text{Tr}(P_1\rho) = \frac{1}{4}.$$

The matrix  $\rho$  is not itself a list of measurement outcomes. It is a machine for producing probability distributions once a measurement is specified.

---

## 4.12 Composite systems

Quantum information often studies systems made from smaller systems. If system A has Hilbert space  $\mathcal{H}_A$ , and system B has Hilbert space  $\mathcal{H}_B$ , then the combined system has Hilbert space

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

A pure state of the combined system is a unit vector

$$|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B.$$

A density matrix of the combined system is a positive semidefinite operator

$$\rho_{AB}$$

on  $\mathbb{C}^A \otimes \mathbb{C}^B$  with

$$\text{Tr}(\rho_{AB}) = 1.$$

Some joint states are simple product states. For example, if system A is in state  $\rho_A$  and system B is in state  $\rho_B$ , then the combined state is

$$\rho_{AB} = \rho_A \otimes \rho_B.$$

If both are pure,

$$\rho_A = |\psi\rangle\langle\psi|, \quad \rho_B = |\phi\rangle\langle\phi|,$$

then

$$\rho_{AB} = |\psi\rangle\langle\psi| \otimes |\phi\rangle\langle\phi| = |\psi\rangle|\phi\rangle\langle\psi|\langle\phi|.$$

We usually write

$$|\psi\rangle|\phi\rangle$$

as shorthand for

$$|\psi\rangle \otimes |\phi\rangle.$$

But not every joint state is a product state. Some states are entangled, meaning that the combined system cannot be described as each subsystem simply having its own independent pure state. Entanglement is one of the central resources of quantum information theory (Nielsen and Chuang, 2010).

---

### 4.13 The partial trace: the state of a subsystem

Suppose we have a joint density matrix

$$\rho_{AB}$$

on

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

Often, we only have access to system A. We then need a way to extract the density matrix of system A alone.

This operation is called the partial trace over B, written

$$\rho_A = \text{Tr}_B(\rho_{AB}).$$

The partial trace is the mathematical operation that discards one subsystem.

#### Defining the partial trace

Let

$$\{|b_1\rangle, \dots, |b_m\rangle\}$$

be an orthonormal basis for  $\square(B)$ . Then

$$\text{Tr}_B(\rho_{AB}) = \sum_{j=1}^m (I_A \otimes \langle b_j |) \rho_{AB} (I_A \otimes |b_j\rangle).$$

The result is an operator on  $\square(A)$ .

Although this formula uses a basis for  $\mathcal{H}(B)$ , the resulting operator does not depend on which orthonormal basis is chosen. The partial trace is characterized by the identity

$$\text{Tr}(M_A \text{Tr}_B(\rho_{AB})) = \text{Tr}((M_A \otimes I_B)\rho_{AB})$$

for every operator  $M(A)$  on  $\mathcal{H}(A)$ . This identity expresses the operational meaning of the partial trace: measuring only system A gives the same probabilities whether we use the full state  $\rho_{AB}$  or the reduced state  $\rho(A)$  (Watrous, 2018).

The state

$$\rho_A = \text{Tr}_B(\rho_{AB})$$

is called the reduced density matrix of system A.

---

#### 4.14 Example: tracing out one qubit from a Bell state

Consider the two-qubit Bell state

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle).$$

Its density matrix is

$$\rho_{AB} = |\Phi^+\rangle\langle\Phi^+|.$$

Expanding,

$$\rho_{AB} = \frac{1}{2} (|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|).$$

Now trace out system B.

Use the rule

$$\text{Tr}_B(|a\rangle|b\rangle\langle a'|\langle b'|) = |a\rangle\langle a'| \langle b'|b\rangle.$$

Then

$$\text{Tr}_B(|00\rangle\langle 00|) = |0\rangle\langle 0|\langle 0|0\rangle = |0\rangle\langle 0|.$$

Also,

$$\text{Tr}_B(|00\rangle\langle 11|) = |0\rangle\langle 1|\langle 1|0\rangle = 0.$$

Similarly,

$$\text{Tr}_B(|11\rangle\langle 00|) = 0,$$

and

$$\text{Tr}_B(|11\rangle\langle 11|) = |1\rangle\langle 1|.$$

Therefore,

$$\rho_A = \text{Tr}_B(\rho_{AB}) = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{I}{2}.$$

This is a crucial lesson:

> A subsystem of a pure entangled state can be mixed.

The full two-qubit state  $|\Phi^+\rangle$  is pure, but each individual qubit has reduced state  $I/2$ . This phenomenon is central to quantum information theory and is one reason density matrices are unavoidable.

---

## 4.15 Purification

We have just seen that a mixed state can arise by ignoring part of a larger pure state. Purification says that this is not an accident.

A purification of a density matrix  $\rho(A)$  on  $\square(A)$  is a pure state

$$|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_R$$

on a larger Hilbert space such that

$$\text{Tr}_R(|\Psi\rangle\langle\Psi|) = \rho_A.$$

The additional system R is sometimes called a reference system or an ancilla. The word ancilla means an auxiliary system introduced to help describe or implement a process.

Purification is a standard and powerful tool in quantum information. It allows us to treat mixed states as parts of larger pure states, which is often mathematically simpler (Nielsen and Chuang, 2010; Watrous, 2018).

### Constructing a purification

Let  $\rho_A$  have spectral decomposition

$$\rho_A = \sum_{j=1}^r \lambda_j |e_j\rangle\langle e_j|,$$

where

$$\lambda_j > 0, \quad \sum_{j=1}^r \lambda_j = 1,$$

and  $r$  is the rank of  $\rho_A$ .

Choose a reference Hilbert space  $\mathcal{H}_R$  with orthonormal vectors

$$|j\rangle_R, \quad j = 1, \dots, r.$$

Define

$$|\Psi\rangle = \sum_{j=1}^r \sqrt{\lambda_j} |e_j\rangle_A \otimes |j\rangle_R.$$

First check that  $|\Psi\rangle$  is normalized:

$$\langle\Psi|\Psi\rangle = \sum_{j=1}^r \lambda_j = 1.$$

Now compute its density matrix:

$$|\Psi\rangle\langle\Psi| = \sum_{j,k=1}^r \sqrt{\lambda_j\lambda_k} |e_j\rangle\langle e_k| \otimes |j\rangle\langle k|.$$

Taking the partial trace over R,

$$\text{Tr}_R(|j\rangle\langle k|) = \langle k|j\rangle = \delta_{jk},$$

where  $\delta_{jk}$  is 1 if  $j=k$  and 0 otherwise. Therefore,

$$\text{Tr}_R(|\Psi\rangle\langle\Psi|) = \sum_{j=1}^r \lambda_j |e_j\rangle\langle e_j| = \rho_A.$$

So  $|\Psi\rangle$  is a purification of  $\rho_A$ .

---

## 4.16 Example: purifying the maximally mixed qubit

Let

$$\rho_A = \frac{I}{2}.$$

Its spectral decomposition is

$$\rho_A = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|.$$

A purification is

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle).$$

As shown earlier,

$$\text{Tr}_B(|\Phi^+\rangle\langle\Phi^+|) = \frac{I}{2}.$$

Thus, the maximally mixed qubit can be understood as one half of a pure entangled two-qubit state.

This does not mean that every physical maximally mixed qubit is secretly entangled with another system in a unique way. Rather, purification says that any density matrix can be represented as the reduced state of some larger pure state. The purification is generally not unique.

---

## 4.17 Why purification matters for dilation

Purification prepares us for the main theme of this book: explaining complicated-looking quantum objects by embedding them into larger, simpler-looking objects.

Naimark dilation will say something similar about measurements:

> A generalized measurement on a system can be represented as a projective measurement on a larger system.

Purification says something similar about states:

> A mixed state on a system can be represented as a pure state on a larger system.

Both ideas follow the same philosophical pattern:

1. Start with an object on a smaller Hilbert space.
2. Introduce a larger Hilbert space.
3. Represent the original object as what remains after compressing, restricting, or tracing out part of the larger description.

This pattern is one of the major organizing principles of quantum information theory.

---

## 4.18 States and transformations: a brief preview

Although this chapter focuses on states, quantum information also studies transformations of states.

A closed quantum system evolves by a unitary operator  $U$ . If the initial pure state is

$$|\psi\rangle,$$

then the final pure state is

$$U|\psi\rangle.$$

In density-matrix form, the transformation is

$$\rho \mapsto U\rho U^*.$$

This formula works for pure and mixed states.

For example, if

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|,$$

then

$$U\rho U^* = \sum_i p_i U|\psi_i\rangle\langle\psi_i|U^*.$$

So each pure component evolves by  $U$ , and the probabilities remain the same.

Later, when we discuss measurement operators, quantum instruments, and Stinespring dilation, we will need more general transformations called completely positive maps. For now, the important point is that density matrices are flexible enough to describe both states and their transformations.

## 4.19 Summary

A pure quantum state can be represented by a unit vector  $|\psi\rangle$ , but vectors differing by a global phase represent the same physical state. The operator form of a pure state is the rank-one projector

$$|\psi\rangle\langle\psi|.$$

A general quantum state is represented by a density matrix  $\rho$ , which satisfies

$$\rho \geq 0, \quad \text{Tr}(\rho) = 1.$$

Mixed states describe probabilistic ensembles and subsystems of larger entangled systems. Every finite-dimensional density matrix has a spectral decomposition

$$\rho = \sum_j \lambda_j |e_j\rangle\langle e_j|,$$

where the  $\lambda_j$ 's form a probability distribution.

The trace formula

$$p = \text{Tr}(P\rho)$$

expresses measurement probabilities for projective measurements, and it will later generalize to POVMs as

$$p = \text{Tr}(E\rho).$$

For composite systems, the partial trace

$$\rho_A = \text{Tr}_B(\rho_{AB})$$

gives the reduced state of a subsystem. A mixed state can always be purified: it can be represented as part of a larger pure state.

These ideas are essential for Naimark dilation. To understand generalized measurements as ordinary measurements on a larger space, we must first understand how quantum theory uses larger spaces to represent states, subsystems, and probability rules.

---

## 4.20 Exercises

### Exercise 4.1: Checking density matrices

For each matrix below, determine whether it is a valid density matrix.

$$\rho_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \rho_2 = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}, \quad \rho_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Remember to check positivity and trace 1.

### Exercise 4.2: Pure or mixed?

Compute  $\text{Tr}(\rho^2)$  for

$$\rho = \begin{pmatrix} 3/4 & 0 \\ 0 & 1/4 \end{pmatrix}.$$

Is  $\rho$  pure or mixed?

### Exercise 4.3: Two ensembles, one state

Show directly that

$$\frac{1}{2}|+\rangle\langle+| + \frac{1}{2}|-\rangle\langle-| = \frac{I}{2}.$$

Then explain in words why this means that the same density matrix can correspond to different preparation procedures.

### Exercise 4.4: Partial trace practice

Let

$$|\Psi\rangle = \sqrt{\frac{3}{4}}|00\rangle + \sqrt{\frac{1}{4}}|11\rangle$$

# Document information

## Chapter 4: Quantum States

---

<b>Project</b>	Naimark Dilation from First Principles
<b>Document</b>	Document 1.8
<b>Author</b>	mujirin
<b>Verifier</b>	Not verified
<b>Downloaded</b>	July 04, 2026 19:17 KST
<b>Status</b>	Working
<b>Document link</b>	<a href="https://www.theorytrace.com/projects/naimark-dilation-from-first-principles/documents/-chapter-4-quantum-states/">https://www.theorytrace.com/projects/naimark-dilation-from-first-principles/documents/-chapter-4-quantum-states/</a>