Chapter 7
Decoherence

In this chapter, we ask what happens when a quantum system is in contact with its environment. To describe this, we need a more general concept of the quantum state, namely the density matrix. We study the density matrix, and how it describes both classical and quantum uncertainty. This also naturally leads to a description of quantum systems at a non-zero temperature. We conclude this chapter with a short discussion on how entropy arises in quantum mechanics.

7.1 Classical and Quantum Uncertainty

In Chap. 1 we considered three simple experiments that went right to the heart of quantum mechanics. In the first two experiments we studied how a very weak laser triggered photodetectors, and we found that light comes in indivisible packets. This led us to hypothesise that light consists of particles, called photons. In the third experiment we set up a Mach–Zehnder interferometer that ensured the photon ended up in detector D_2, and never in detector D_1. When we use high-intensity laser light, this behaviour is completely explained by the classical wave theory of light, and is called interference. We then arrived at the uncomfortable conclusion that a single particle (the photon) can exhibit interference. To explore this further, we modified the third experiment to add quantum non-demolition (QND) detectors in the paths of the interferometer. This served to tell us which path the photons took inside the interferometer. Consequently, the interference pattern was destroyed: the photons no

---

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-3-319-92207-2_7) contains supplementary material, which is available to authorized users.

P. Kok, *A First Introduction to Quantum Physics*,
Undergraduate Lecture Notes in Physics,
https://doi.org/10.1007/978-3-319-92207-2_7

longer end up only in detector D_2, but randomly trigger D_1 or D_2. We concluded that photons are neither particles, nor waves, and make up a new type of physical objects called quanta with characteristics of both.

In this section, we will study this experiment a little further to establish what kind of knowledge we can have about quanta. We will arrive at the notion that there are two kinds of uncertainty: classical and quantum uncertainty. To see this, consider the experiment where we measure the path taken by the photon. In this case we lose the interference and find photons hitting both detector D_1 and detector D_2. We can imagine that this experiment is some setup on a bench, with lasers, beam splitters, mirrors, and detectors that are merrily clicking away. Suppose that the outcomes of the photodetectors D_1 and D_2 as well as the which-path QND detectors are recorded in two data files on a computer, one for the photodetectors, and one for the QND detectors. After a while of taking data we are getting a bit thirsty. Since our presence is not required for the correct running of our automated experiment, we decide to leave the lab briefly for a cup of coffee.

When we come back, we discover that disaster has struck! One of the data files on the hard drive of our computer is corrupted. The file with recorded data for D_1 and D_2 is fine, and we see that the photons indeed kept going randomly to D_1 and D_2. However, the data file for the QND detectors has been corrupted and can no longer be read out. This means that we no longer have the information which path the photons took. Nevertheless, as the data file of detectors D_1 and D_2 showed, we did not recover the interference (detector D_1 kept firing, as well as detector D_2). This is completely in agreement with common sense, because if we deliberately destroy the which-path information ourselves we do not expect to change events that we already observed. That would wreak havoc with causality!

However, this thought experiment has interesting implications. In one experiment we do not measure the path of the photon and obtain interference in the detector outcomes. We have no knowledge about the path of the photon. In the other experiment we lose the information about the path of the photon, and consequently have no knowledge about the path either. However, we have lost the interference in the detector outcomes. Since the two experiments have different implications for the interference in detectors D_1 and D_2, these must be two different kinds of uncertainty. The case where we lose the information about the path we call classical uncertainty, because we know that the photon has taken a well-defined path, and we are just uncertain as to which it was. On the other hand, when we do not make a path measurement in the first place, we cannot say that the photon took either one of the two possible paths. The path of the photon is indeterminate. This is quantum uncertainty.

We can ask the question whether it is important that a measurement was made. What if the data file was not corrupted, but instead the outcomes weren't send to the computer in the first place, perhaps due to a faulty cable? Maybe the electronics in the measurement apparatus was not working properly. It is easy to see that we can push the fault in the measurement apparatus closer and closer to the key aspect of what makes the which-path detector a measurement, namely the interaction of the apparatus with the photon. If there is an interaction of the photon with another system,

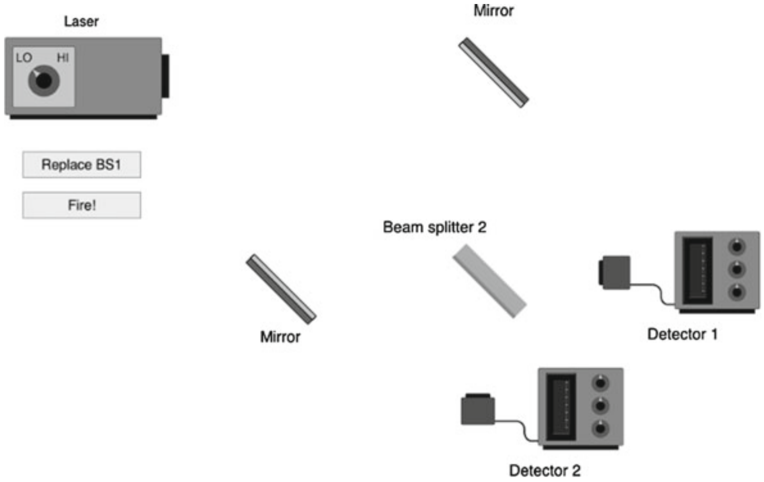

Fig. 7.1 A Mach–Zehnder interferometer with a removable beam splitter. The interactive figure is available online (see supplementary material 1)

and the occurrence of that interaction is stored *anywhere*, from a classical data file to the quantum state of the environment, then we lose the destructive interference in D_1. The loss of interference in the detectors indicates that something has changed in the state of the photon as a result of the interaction with the measurement apparatus. This change is called *decoherence*.

7.2 The Density Matrix

In the previous section we have seen that the interaction of a photon with a measurement apparatus or with the environment will cause the disappearance of interference, and the uncertainty about the path of the photon in the Mach–Zehnder interferometer is no longer quantum but classical. In this section we will see how we can describe classical uncertainties mathematically in terms of an object called the density matrix, and later in this chapter we will show how to describe the situation of the previous section.

We introduce classical uncertainty in our description of the state of a photon in the following way: Every time we send a single photon into a Mach–Zehnder interferometer, we toss a coin on whether or not we remove the first beam splitter BS1. When the beam splitter is present, detector D_2 will always click, as before. When the beam splitter is not present, the photon will always be in the upper arm of the interferometer (technically it is no longer a Mach–Zehnder interferometer), and the second beam splitter BS2 will make the photon end up in detectors D_1 and D_2, each with probability $1/2$. This is shown in Fig. 7.1.

The uncertainty in the measurement outcomes in the absence of BS1 is a quantum uncertainty, since prior to the detection the photon is in the superposition of moving towards D_1 and moving towards D_2. Mathematically, this state is

$$|\neg\text{BS1}\rangle = \frac{1}{\sqrt{2}}|D_1\rangle + \frac{1}{\sqrt{2}}|D_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \tag{7.1}$$

where the symbol $\neg$ means "not". On the other hand, when BS1 is present, there is no uncertainty about which detector the photon will trigger:

$$|\text{BS1}\rangle = |D_2\rangle = \begin{pmatrix}1\\0\end{pmatrix}. \tag{7.2}$$

The probability that detector D_2 will be triggered by the photon is different in both cases, and we can calculate this as

$$\begin{aligned}\Pr(D_2|\text{BS1}) &= |\langle D_2|\text{BS1}\rangle|^2 = 1\\ \Pr(D_2|\neg\text{BS1}) &= |\langle D_2|\neg\text{BS1}\rangle|^2 = \frac{1}{2}.\end{aligned} \tag{7.3}$$

The vertical bar in $\Pr(\cdot|\cdot)$ should be read as "given that", so $\Pr(a|b)$ is the probability that a happens, given that b occurred. This is a so-called conditional probability.

So far, we have looked only at the detection probability in D_2 for each separate arrangement of the interferometer. However, what if we do not know whether BS1 was removed or not, but we only know that a (fair) coin was used to determine if BS1 is present? In that case, we have a classical uncertainty about the setup. The classical rules of probability then dictate that we add the probabilities $\Pr(D_2|\text{BS1})$ and $\Pr(D_2|\neg\text{BS1})$, weighed by the probabilities of BS1 being present or absent:

$$\Pr(D_2) = \frac{1}{2}\Pr(D_2|\text{BS1}) + \frac{1}{2}\Pr(D_2|\neg\text{BS1}), \tag{7.4}$$

where we used that

$$\Pr(\text{BS1}) = \Pr(\neg\text{BS1}) = \frac{1}{2}. \tag{7.5}$$

Note that these probabilities are not conditioned; they are the straight-up probabilities of whether the first beam splitter is there or not.

We can rewrite Eq. (7.4) by substituting the expressions of the modulo-squared scalar products in Eq. (7.3), and we arrive at

$$
\begin{aligned}
\Pr(D_2) &= \frac{1}{2}|\langle D_2|\mathrm{BS1}\rangle|^2 + \frac{1}{2}|\langle D_2|\neg\mathrm{BS1}\rangle|^2 \\
&= \frac{1}{2}\langle D_2|\mathrm{BS1}\rangle\langle \mathrm{BS1}|D_2\rangle + \frac{1}{2}\langle D_2|\neg\mathrm{BS1}\rangle\langle\neg\mathrm{BS1}|D_2\rangle \\
&= \langle D_2|\left(\frac{1}{2}|\mathrm{BS1}\rangle\langle\mathrm{BS1}| + \frac{1}{2}|\neg\mathrm{BS1}\rangle\langle\neg\mathrm{BS1}|\right)|D_2\rangle \\
&= \langle D_2|\rho|D_2\rangle\,.
\end{aligned} \tag{7.6}
$$

where

$$\rho = \frac{1}{2}|\mathrm{BS1}\rangle\langle\mathrm{BS1}| + \frac{1}{2}|\neg\mathrm{BS1}\rangle\langle\neg\mathrm{BS1}| \tag{7.7}$$

is called the density matrix, or density operator. You see that it plays the role of the quantum state of the system in the calculation of the probability of the measurement outcome D_2, and that it includes the classical uncertainty in the form of the factors 1/2. We can write ρ in matrix notation by using

$$|\mathrm{BS1}\rangle\langle\mathrm{BS1}| = \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}, \tag{7.8}$$

and

$$|\neg\mathrm{BS1}\rangle\langle\neg\mathrm{BS1}| = \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1 & 1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}. \tag{7.9}$$

The density matrix is then

$$
\begin{aligned}
\rho &= \frac{1}{2}|\mathrm{BS1}\rangle\langle\mathrm{BS1}| + \frac{1}{2}|\neg\mathrm{BS1}\rangle\langle\neg\mathrm{BS1}| \\
&= \frac{1}{2}\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} + \frac{1}{2}\times\frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix} = \frac{1}{4}\begin{pmatrix}3 & 1\\1 & 1\end{pmatrix}.
\end{aligned} \tag{7.10}
$$

The probability of finding the photon in detector D_2 when we are uncertain about the presence or absence of BS1 is then

$$\Pr(D_2) = \langle D_2|\rho|D_2\rangle = \frac{1}{4}\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}3 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \frac{1}{4}\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}3\\1\end{pmatrix} = \frac{3}{4}. \tag{7.11}$$

Similarly, we can calculate the probability of finding the photon in detector D_1 using the same density matrix:

$$\Pr(D_1) = \langle D_1|\rho|D_1\rangle = \frac{1}{4}\begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}3 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \frac{1}{4}\begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{4}. \tag{7.12}$$

A quick check reveals that the probability of finding the photon in either D_1 or D_2 is

$$\frac{1}{4} + \frac{3}{4} = 1\,, \tag{7.13}$$

as it should be, since there are no other possibilities for the photon to behave (and we assumed lossless detectors). Note that while the density matrix describes the state of a quantum system, it is no longer a vector. This is because vectors do not have the flexibility to encode both the quantum and the classical uncertainty without going to state spaces of higher dimension.

Perhaps the coin we used to decide whether to include BS1 was not fair. We can assign probabilities that are different from 1/2, for example p_1 and p_2, with $p_1 + p_2 = 1$. If for brevity we write the two different photon states as $|\psi_1\rangle$ and $|\psi_2\rangle$, the density matrix becomes

$$\rho = p_1|\psi_1\rangle\langle\psi_1| + p_2|\psi_2\rangle\langle\psi_2|\,, \tag{7.14}$$

where $|\psi_1\rangle$ and $|\psi_2\rangle$ do not have to be orthogonal (for example, $|\text{BS1}\rangle$ and $|\neg\text{BS1}\rangle$ are not orthogonal; you should verify this!).

In the extreme case where $p_1 = 1$ and $p_2 = 0$, there is no classical uncertainty, and all probabilistic measurement outcomes are due to quantum uncertainty. There is a quick way to test whether a density matrix has classical uncertainty. Calculate the square of ρ :

$$\begin{aligned}\rho^2 &= (p_1|\psi_1\rangle\langle\psi_1| + p_2|\psi_2\rangle\langle\psi_2|)^2 \\ &= p_1^2|\psi_1\rangle\langle\psi_1| + p_2^2|\psi_2\rangle\langle\psi_2| + p_1p_2\left(\langle\psi_1|\psi_2\rangle|\psi_1\rangle\langle\psi_2| + \langle\psi_2|\psi_1\rangle|\psi_2\rangle\langle\psi_1|\right).\end{aligned} \tag{7.15}$$

In the special case where $p_1 = 1$ and $p_2 = 0$ we see that

$$\rho^2 = \rho\,. \tag{7.16}$$

In Chap. 5 we learned that this is the defining characteristic of a projection operator. So if there is no classical uncertainty, the quantum state of the system can be described by a projection operator. The converse is also true: if a quantum system is described by a projection operator, then there is no classical uncertainty in the system, and all probabilistic measurement outcomes are due to quantum uncertainty. When $\rho^2 = \rho$, we call the state of the system pure, and when $\rho^2 \neq \rho$ we call the state mixed:

$$\begin{aligned}\rho^2 = \rho \;&\to\; \rho \text{ is a pure state,} \\ \rho^2 \neq \rho \;&\to\; \rho \text{ is a mixed state.}\end{aligned} \tag{7.17}$$

The terminology "mixed" always refers to classical uncertainty, while the term "superposition" refers to quantum uncertainty. You should never call a superposition a "mixture" of possibilities, because that will lead to confusion.

Some quantum states are more mixed than others, in the sense that they contain a greater classical uncertainty. We can quantify this by calculating the trace of the square of the density matrix, which is the sum over the diagonals. For the density matrix in Eq. (7.10) we calculate

$$\rho^2 = \frac{1}{16}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} = \frac{1}{16}\begin{pmatrix} 10 & 4 \\ 4 & 2 \end{pmatrix} . \tag{7.18}$$

The trace of ρ^2 is then given by

$$\mathrm{Tr}\left(\rho^2\right) = \frac{10}{16} + \frac{2}{16} = \frac{12}{16} = \frac{3}{4} . \tag{7.19}$$

This is called the *purity* of the state ρ. By contrast, the pure state for the photon coming out of the interferometer without BS1 can be written as the density matrix

$$\rho_{\neg\mathrm{BS1}} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} . \tag{7.20}$$

Moreover, since there is no classical uncertainty about the absence of BS1 the quantum state is pure and $\rho_{\neg\mathrm{BS1}} = \rho^2_{\neg\mathrm{BS1}}$. The trace of the diagonal elements of $\rho^2_{\neg\mathrm{BS1}}$ is therefore

$$\frac{1}{2} + \frac{1}{2} = 1 . \tag{7.21}$$

Pure states have a purity of 1.

There are also maximally mixed states. These are quantum states with maximum classical uncertainty, where we know nothing about the path of the photon. We can construct such a state by sending photons at detectors D_1 and D_2 based on the random outcome of a fair coin. The density matrix for this state is

$$\rho_{\max} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \tag{7.22}$$

The purity of this state is 1/2. It cannot be lower for systems with two orthogonal states.

Note that the trace of ρ is always equal to 1. This is because the eigenvalues of ρ are the probabilities of finding the measurement outcomes associated with the eigenvectors of ρ. In addition, since probabilities are always real numbers between zero and one, density matrices have real eigenvalues. That makes them Hermitian matrices, for which $\rho^\dagger = \rho$. Also, the expectation value of ρ with respect to any quantum state must be positive, since it is the probability of finding that state in a measurement of the quantum system. This property of the density matrix is called positivity. Every density matrix in quantum mechanics has the following properties:

1. $\mathrm{Tr}\,(\rho) = 1$ (normalisation),
2. $\rho^\dagger = \rho$ (Hermiticity),
3. $\langle\psi|\rho|\psi\rangle \geq 0$ for any $|\psi\rangle$ (positivity).

Any matrix that satisfies these requirements is a valid density matrix.

So far we have shown how we can calculate the probabilities of measurement outcomes given a density matrix ρ. However, we also need to know how we can calculate the expectation value for an observable if our system is in the state ρ. Let our density matrix be given by a generalisation for Eq. (7.14):

$$\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j| \,, \tag{7.23}$$

where j runs over all possible states $|\psi_j\rangle$ that are prepared with probability p_j. The expectation value of an operator A with respect to one of the states $|\psi_j\rangle$ is given by

$$\langle A\rangle_j = \langle\psi_j|A|\psi_j\rangle \,. \tag{7.24}$$

When the preparation procedure is a probability distribution over all $|\psi_j\rangle$, the expectation value of A must be the weighted average over all $|\psi_j\rangle$. We therefore obtain

$$\langle A\rangle = \sum_j p_j \langle A\rangle_j = \sum_j p_j \langle\psi_j|A|\psi_j\rangle \,. \tag{7.25}$$

We can write this in a more suitable way for our purposes by using the expression in Eq. (5.80) in terms of the projector $P_j = |\psi_j\rangle\langle\psi_j|$:

$$\langle\psi_j|A|\psi_j\rangle = \mathrm{Tr}\big[A P_j\big] = \mathrm{Tr}[A|\psi_j\rangle\langle\psi_j|] \,. \tag{7.26}$$

We can then express the expectation value of A with respect to the mixed state ρ as

$$\langle A\rangle = \sum_j p_j \mathrm{Tr}[A|\psi_j\rangle\langle\psi_j|] = \mathrm{Tr}\left[A\left(\sum_j p_j |\psi_j\rangle\langle\psi_j|\right)\right] = \mathrm{Tr}[A\rho] \,. \tag{7.27}$$

This is a very compact expression that is easy to calculate: we need only multiply two matrices and calculate its trace, given by the sum over the diagonal elements. Other expressions, such as the variance of A, follow directly from this form.

7.3 Interactions with the Environment

Now that we know how to describe classical uncertainty in a quantum state, let's explore how this arises in the case where we place a quantum non-demolition detector in the arms of a Mach–Zehnder interferometer. Since the key mechanism of the loss of interference of the photon is the interaction with the QND detector, we want to study this some more. We can construct an interaction Hamiltonian that relates the path of the photon to the state of the QND detector, and calculate the time evolution of the composite system of the photon and the QND detectors. However, this can be quite a lengthy procedure that requires a mathematical sophistication that we don't have yet. Each different interaction will also have a different interaction Hamiltonian, so when we calculate the time evolution for one situation (such as our photon and the QND detectors), we still cannot say much about the time evolution of other systems interacting with detectors or the environment. Luckily, we do not have to do all this if we are not interested in the details of the evolution. Instead, we can define very general rules for the behaviour of detectors and work from there. This is what we will do in this section.

Consider again the photon in the Mach–Zehnder interferometer (with BS1 present). As we have seen, inside the interferometer the photon is in the state

$$|\text{photon}\rangle = \frac{|\text{path 1}\rangle + |\text{path 2}\rangle}{\sqrt{2}} . \tag{7.28}$$

As the photon interacts with the QND detector, it has to change the state of the detector, otherwise we cannot use the device as a detector in any sense of a measurement (since we are monitoring the state of the detector). We can define such states without knowing anything about the details of the detectors, because all we need is that the states are distinguishable, that is, orthonormal. Furthermore, we can consider the two QND detectors in each path as a single physical device with three relevant orthonormal states:

$$\begin{aligned} |Q_0\rangle &: \quad \text{the ``ready'' state,} \\ |Q_1\rangle &: \quad \text{the ``photon is in path 1'' state,} \\ |Q_2\rangle &: \quad \text{the ``photon is in path 2'' state.} \end{aligned} \tag{7.29}$$

The detector typically has many more states, but these are the ones we are interested in. The composite system of the photon and the QND detector is then initially in the state $|\text{photon}\rangle|Q_0\rangle$.

If the interaction of the photon with the detector is to produce a useful readout, the evolution of the composite system must obey the following rules:

$$\begin{aligned} |\text{path 1}\rangle|Q_0\rangle &\to |\text{path 1}\rangle|Q_1\rangle \\ |\text{path 2}\rangle|Q_0\rangle &\to |\text{path 2}\rangle|Q_2\rangle , \end{aligned} \tag{7.30}$$

that is, the photon in path 1 changes the state of the detector from the "ready" state to the state that indicates that the photon is in path 1, and similarly for path 2, as required. Also, the detector does not change the path of the photon. This is all we need from a device that we may reasonably call a detector.

Now let's see how this affects the state of the photon in the Mach–Zehnder interferometer:

$$
\begin{aligned}
|\text{photon}\rangle|Q_0\rangle &= \left(\frac{|\text{path 1}\rangle + |\text{path 2}\rangle}{\sqrt{2}}\right)|Q_0\rangle \\
&= \frac{1}{\sqrt{2}}\left(|\text{path 1}\rangle|Q_0\rangle + |\text{path 2}\rangle|Q_0\rangle\right) \\
&\to \frac{|\text{path 1}\rangle|Q_1\rangle + |\text{path 2}\rangle|Q_2\rangle}{\sqrt{2}}\,. \qquad (7.31)
\end{aligned}
$$

If we take a careful look at the composite state of the photon and the detector after the detection event, you see that the photon and the QND detector are now entangled. We can no longer write the separate states for the photon and for the detector. A moment's thought will convince you that entanglement must occur every time the photon is not initially in the state $|\text{path 1}\rangle$ or $|\text{path 2}\rangle$.

If we now apply the second beam splitter in the Mach–Zehnder interferometer to the photon, we obtain the following transformation:

$$|\text{path 1}\rangle \to \frac{|D_1\rangle + |D_2\rangle}{\sqrt{2}}\,, \qquad (7.32)$$

and

$$|\text{path 2}\rangle \to \frac{-|D_1\rangle + |D_2\rangle}{\sqrt{2}}\,. \qquad (7.33)$$

This transforms (7.23) into

$$\frac{1}{2}\left(|D_1\rangle|Q_1\rangle + |D_2\rangle|Q_1\rangle - |D_1\rangle|Q_2\rangle + D_2\rangle|Q_2\rangle\right)\,. \qquad (7.34)$$

The terms including $|\mathsf{D}_1\rangle$ no longer cancel, due to the presence of $|Q_1\rangle$ and $|Q_2\rangle$. This means that detector D_1 now also fires. There is no longer destructive interference that keeps D_1 dark. When $|Q_1\rangle = |Q_2\rangle$ the QND detector does not extract any information about the path of the photon, and we obtain the cancellation of the terms including $|D_1\rangle$.

To illustrate this further, we consider another example. A quantum particle is in a superposition of two orthogonal states $|\phi_a\rangle$ and $|\phi_b\rangle$ with amplitudes a and b:

$$|\psi\rangle = a|\phi_a\rangle + b|\phi_b\rangle\,. \qquad (7.35)$$

In terms of the density matrix, this can be written as

$$\rho_{\text{particle}} = \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix}. \qquad (7.36)$$

The particle interacts with the environment in such a way that it leaves a trace of its presence in the form of a changed state of the environment. The undisturbed state of the environment is $|E_0\rangle$, while the disturbed environmental states due to the states of the particle $|\phi_a\rangle$ and $|\phi_b\rangle$ are $|E_a\rangle$ and $|E_b\rangle$. The state of the particle and the environment after the interaction is then

$$|\psi\rangle|E_0\rangle \;\rightarrow\; a|\phi_a\rangle|E_a\rangle + b|\phi_b\rangle|E_b\rangle\,. \tag{7.37}$$

When $|E_a\rangle$ and $|E_b\rangle$ are different, this is again an entangled state.

So far, we have considered only the state of the composite system of the photon and the QND detector, or the particle and its environment. However, we are really interested in the state of the photon or the particle by itself. This means that we must somehow lose the description of the detector or the environment from our description.

We know that in general the state must be given by a density matrix, rather than a pure state, since classical uncertainty is involved. So we start by writing the state of the particle and the environment in Eq. (7.37) in the form of a density matrix

$$\rho = (a|\phi_a\rangle|E_a\rangle \,+\, b|\phi_b\rangle|E_b\rangle)(a^*\langle\phi_a|\langle E_a| \,+\, b^*\langle\phi_b|\langle E_b|) \tag{7.38}$$

that we can write out in individual terms as

$$\begin{aligned}\rho =& |a|^2|\phi_a\rangle\langle\phi_a| \otimes |E_a\rangle\langle E_a| + ab^*|\phi_a\rangle\langle\phi_b| \otimes |E_a\rangle\langle E_b| \\ &+ a^*b|\phi_b\rangle\langle\phi_a| \otimes |E_b\rangle\langle E_a| + |b|^2|\phi_b\rangle\langle\phi_b| \otimes |E_b\rangle\langle E_b|\,,\end{aligned} \tag{7.39}$$

where we used the symbol $\otimes$ to make a clear distinction between operators $|\phi_a\rangle\langle\phi_b|$ on the particle and operators $|E_a\rangle\langle E_b|$ on the environment.

One way to "forget" about the environment is to measure it and discard the measurement outcome. In our examples there are three possible measurement outcomes, namely E_0, E_a, and E_b. The probabilities of these outcomes are given by

$$p_a = \langle E_a|\rho_E|E_a\rangle \quad \text{and} \quad p_b = \langle E_b|\rho_E|E_b\rangle\,, \tag{7.40}$$

where ρ_E is the density operator for the environment. However, we do not have ρ_E, we have the density matrix ρ of both the particle and the environment. But we can still take the expectation value of ρ with respect to $|E_a\rangle$ and $|E_b\rangle$. The result is not a probability, since we retain the kets and bras of the particle:

$$\begin{aligned}\langle E_a|\rho|E_a\rangle =& |a|^2|\phi_a\rangle\langle\phi_a|\langle E_a|E_a\rangle\langle E_a|E_a\rangle + ab^*|\phi_a\rangle\langle\phi_b|\langle E_a|E_a\rangle\langle E_b|E_a\rangle \\ &+ a^*b|\phi_b\rangle\langle\phi_a|\langle E_a|E_b\rangle\langle E_a|E_a\rangle + |b|^2|\phi_b\rangle\langle\phi_b|\langle E_a|E_b\rangle\langle E_b|E_a\rangle.\end{aligned} \tag{7.41}$$

Since $\langle E_a|E_a\rangle = 1$ and $\langle E_a|E_b\rangle = \langle E_b|E_a\rangle = 0$, we find that

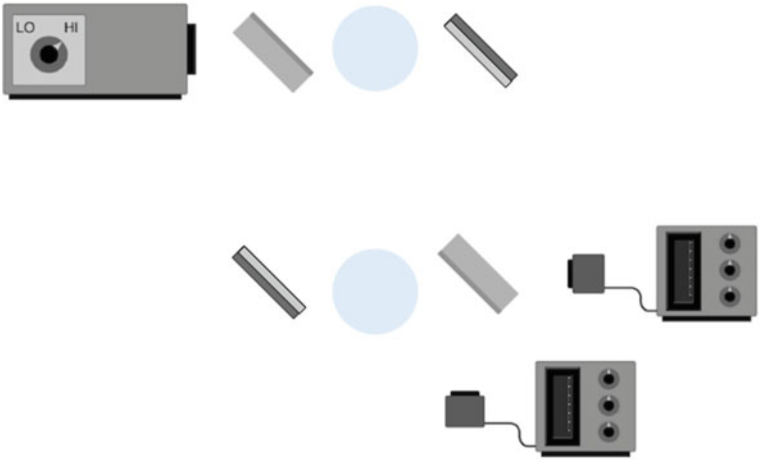

Fig. 7.2 A Mach–Zehnder interferometer where the photon interacts with the environment. The interactive figure is available online (see supplementary material 2)

$$\langle E_a|\rho|E_a\rangle = |a|^2|\phi_a\rangle\langle\phi_a| \,. \tag{7.42}$$

Similarly, we find that $\langle E_0|\rho|E_0\rangle = 0$, and

$$\langle E_b|\rho|E_b\rangle = |b|^2|\phi_b\rangle\langle\phi_b| \,. \tag{7.43}$$

Therefore, the density operator that describes the particle is given by

$$\rho_{\text{particle}} = |a|^2|\phi_a\rangle\langle\phi_a| + |b|^2|\phi_b\rangle\langle\phi_b| = \begin{pmatrix} |a|^2 & 0 \\ 0 & |b|^2 \end{pmatrix} , \tag{7.44}$$

where the trace of ρ_{particle} is given by $|a|^2 + |b|^2 = 1$, as required. The procedure we used to reduce the state of the composite system to the state of a subsystem is called the partial trace, or the "trace over the environment". Since the off-diagonal terms in Eq. (7.44) are zero, the quantum uncertainty between $|\psi_a\rangle$ and $|\psi_b\rangle$ present in Eq. (7.36) has vanished. Only classical uncertainty about the states $|\psi_a\rangle$ and $|\psi_b\rangle$ remains. The interaction of a photon in a Mach–Zehnder interferometer interacting with the environment is shown in Fig. 7.2.

When the interaction with the environment is weak, we retain some quantum uncertainty in the state of the particle. Suppose that the interaction is such that the environment does not evolve to orthogonal states, but that the state of the environment is perturbed only slightly:

$$\begin{aligned} |\phi_a\rangle|E_0\rangle &\rightarrow |\phi_a\rangle\left(\sqrt{1-|\epsilon|^2}\,|E_0\rangle + \epsilon|E_a\rangle\right) \\ |\phi_b\rangle|E_0\rangle &\rightarrow |\phi_b\rangle\left(\sqrt{1-|\epsilon|^2}\,|E_0\rangle + \epsilon|E_b\rangle\right) , \end{aligned} \tag{7.45}$$

where ϵ is some (potentially complex) parameter that depends on the actual interaction strength between the particle and the environment. This means that the environment does not gain complete information about the path of the particle. The state $|\psi\rangle = a|\phi_a\rangle + b|\phi_b\rangle$ after the interaction is then given by

$$|\psi\rangle|E_0\rangle \;\to\; \sqrt{1-|\epsilon|^2}\;|\psi\rangle|E_0\rangle + \epsilon\,(a|\phi_a\rangle|E_a\rangle + b|\phi_b\rangle|E_b\rangle)\,. \tag{7.46}$$

When $\epsilon = 0$ the particle does not interact with the environment at all, and the state of the particle and the environment remains separable. For any other value of ϵ the composite system exhibits entanglement. This example also allows us to model phenomenologically a quantum system that decoheres over time, by including a time dependence t in ϵ. Since $0 \leq |\epsilon| \leq 1$, we can choose that for late times $\epsilon \to 1$. One such function is

$$\epsilon(t) = \sqrt{1 - e^{-\gamma t}}\,. \tag{7.47}$$

This describes a particle that slowly loses its coherence. The parameter γ has units of inverse time, and it is the decoherence rate of the particle.

We can apply the partial trace over the environment to the composite state in Eq. (7.46), which gives for the density operator of the particle

$$\rho_{\text{particle}} = \begin{pmatrix} |a|^2 & (1-|\epsilon|^2)\,ab^* \\ (1-|\epsilon|^2)\,a^*b & |b|^2 \end{pmatrix}. \tag{7.48}$$

When we substitute the form of ϵ given by the decoherence model in Eq. (7.47) we obtain

$$\rho_{\text{particle}} = \begin{pmatrix} |a|^2 & ab^*\,e^{-\gamma t} \\ a^*b\,e^{-\gamma t} & |b|^2 \end{pmatrix}. \tag{7.49}$$

This demonstrates the exponential suppression of the off-diagonal terms of the density matrix that is associated with decoherence. The form of ϵ must be derived from a detailed model of the interaction Hamiltonian between the particle and the environment. This requires a lot more knowledge about modelling the interaction with the environment than we have developed here, and it is currently an area of active research.

Suppose that the state $|\phi_a\rangle = |g\rangle$ is the ground state on an atom and the state $|\phi_b\rangle = |e\rangle$ is the excited state. The constant (weak) interaction of the atom with the environment is increasingly likely to record the energy state of the atom, and this results in increasingly suppressed off-diagonal terms in the density matrix as quantum uncertainty gives way to classical uncertainty. This is a problem that designers of quantum computers must overcome to build a working machine, and typically the qubits (which may be atoms) can be kept sufficiently free of decoherence for only a short amount of time.

7.4 Quantum Systems at Finite Temperature

Another important physical application for density operators is when a physical system is held at a finite temperature instead of zero temperature. In such situations, the system is kept into thermal contact with a bath at temperature T, which means that there is some kind of interaction between the system and the bath. Typically, the thermal bath is a very large system, and will therefore be difficult—if not impossible—to characterise completely. Fortunately, when lots of particles in a bath interact with a quantum system, their interactions will average, and this can be described by referring only to the density matrix of the system itself. In effect, we are looking at the state of the system after the thermal bath (and its interaction with our system) has been traced out.

A thermal bath will inject energy into our quantum system that is proportional to T. This has the effect of populating the higher energy levels. At the same time, as we have seen in the previous section, ignoring the quantum state of the system that interacts with our quantum system, in this case the thermal bath, destroys the coherences of our system. We therefore expect that when we let our system come to thermal equilibrium with a bath, the resulting state should look something like

$$\rho(T) = \sum_n p_n(T)|E_n\rangle\langle E_n| \, , \tag{7.50}$$

where $|E_n\rangle$ are the energy eigenstates, and $p_n(T)$ is the probability of finding the system in energy state $|E_n\rangle$ given the temperature T. Our task is now to find these probabilities.

When our system is held at temperature T, there will be some chance that upon a measurement it will be found in energy state $|E_n\rangle$. However, the larger the energy E_n, the smaller the probability that our system is in the state $|E_n\rangle$. We relate temperature and energy by Boltzmann's constant $k_\mathrm{B} = 1.38 \times 10^{-23}$ J K^{-1}, such that $k_\mathrm{B}T$ has units of energy. The probability that the system is in state $|E_n\rangle$ when the energy of the bath $k_\mathrm{B}T$ is smaller than E_n is exponentially suppressed, because the random fluctuations of the bath must all coincide to push the system in the higher energy state. So we can write

$$p_n(T) \propto \exp\left(-\frac{E_n}{k_\mathrm{B}T}\right) , \tag{7.51}$$

where the $\propto$ symbol denotes "proportional to". This is called the Boltzmann factor. We will find it convenient to define $\beta = (k_\mathrm{B}T)^{-1}$ in what follows. When we substitute this expression into the thermal density matrix, we obtain

$$\tilde{\rho}(T) = \sum_n e^{-\beta E_n}|E_n\rangle\langle E_n| \, , \tag{7.52}$$

which is not a normalised density matrix (hence the tilde on ρ). Since the energy eigenstates are orthonormal states, the normalised thermal density matrix becomes simply

$$\rho(T) = \frac{\sum_n e^{-\beta E_n} |E_n\rangle\langle E_n|}{\sum_n e^{-\beta E_n}} \,. \tag{7.53}$$

This is the quantum state of a system in thermal equilibrium with a bath at temperature T. Note that the nature of the interaction between the system and the bath is not part of the description of $\rho(T)$, which is very convenient!

We can write Eq. (7.53) in a much more elegant way using the fact that the energy eigenstates and eigenvalues together form the spectral decomposition of the Hamiltonian:

$$H = \sum_n E_n |E_n\rangle\langle E_n| \,. \tag{7.54}$$

We can express the exponential function $e^{-\beta E_n}$ in a power series

$$e^{-\beta E_n} = \sum_{k=0}^{\infty} \frac{(-\beta)^k E_n^k}{k!} \,. \tag{7.55}$$

Substituting this into Eq. (7.52), we obtain

$$\begin{aligned}
\sum_n e^{-\beta E_n} |E_n\rangle\langle E_n| &= \sum_n \sum_{k=0}^{\infty} \frac{(-\beta)^k E_n^k}{k!} |E_n\rangle\langle E_n| \\
&= \sum_{k=0}^{\infty} \frac{(-\beta)^k}{k!} \left(\sum_n E_n^k |E_n\rangle\langle E_n| \right) \\
&= \sum_{k=0}^{\infty} \frac{(-\beta)^k}{k!} \left(\sum_n E_n |E_n\rangle\langle E_n| \right)^k \\
&= \sum_{k=0}^{\infty} \frac{(-\beta)^k}{k!} H^k = \exp(-\beta H) \,.
\end{aligned} \tag{7.56}$$

The step from the second to the third line is allowed because the energy eigenstates are orthonormal: $\langle E_n | E_m \rangle = \delta_{nm}$, with $\delta_{nm} = 0$ if $n \neq m$ and $\delta_{nm} = 1$ if $n = m$. We can apply a very similar argument to prove that

$$\sum_n e^{-\beta E_n} = \mathrm{Tr}\left[e^{-\beta H} \right] . \tag{7.57}$$

The normalised thermal density matrix then becomes

$$\rho(\beta) = \frac{e^{-\beta H}}{\mathrm{Tr}\left[e^{-\beta H}\right]} = \frac{1}{Z} e^{-\beta H} \,, \tag{7.58}$$

where we defined the so-called *partition function* $Z = \mathrm{Tr}\left[e^{-\beta H}\right]$. This is an important quantity in statistical physics, where it is used to calculate a variety of thermodynamical quantities.

As an example, we can calculate the average energy in our thermal system. According to Eq. (7.27), the expectation value of the Hamiltonian is given by

$$\langle H \rangle = \mathrm{Tr}[H\rho(\beta)] \,, \tag{7.59}$$

which can now be written as

$$\begin{aligned} \langle H \rangle &= \mathrm{Tr}[H\rho(\beta)] = \mathrm{Tr}\left[\frac{He^{-\beta H}}{Z}\right] = \frac{1}{Z}\mathrm{Tr}\left[He^{-\beta H}\right] \\ &= \frac{1}{Z}\mathrm{Tr}\left[-\frac{d}{d\beta}e^{-\beta H}\right] = -\frac{1}{Z}\frac{d}{d\beta}\mathrm{Tr}\left[e^{-\beta H}\right] \\ &= -\frac{1}{Z}\frac{dZ}{d\beta} = -\frac{d\ln Z}{d\beta} \,. \end{aligned} \tag{7.60}$$

So we can find the average energy straight from the partition function Z by a simple derivative with respect to β. Note that Z ultimately depends only on the energy eigenvalues.

In statistical physics, we are typically interested in a macroscopically large number of particles. In principle we can describe this using the density matrix for a large number of systems. The Hamiltonian must then be replaced with the many-particle Hamiltonian. This involves numerous technical steps that are beyond the scope of this book, e.g., what to do when we have identical particles. The resulting quantum theory predicts behaviour such as superconductivity, superfluidity, and Bose–Einstein condensation, and is an active area of current research.

7.5 ❧ Entropy and Landauer's Principle

An important concept in physics is the *entropy* of a system. It is one of the fundamental concepts in thermodynamics, together with energy, temperature, heat and work. Loosely speaking, it is a measure of the disorder of a system. The second law of thermodynamics says that the entropy in an isolated system can never decrease, and this has been related to the arrow of time. In this section we show how the entropy of a quantum system can be obtained from the density matrix.

Ludwig Boltzmann was the first person to think really deeply about the meaning of entropy at the microscopic (or atomic) scale. Boltzmann realised that the entropy is given by the number of ways in which you can arrange the atoms in a gas (the micro-states) without changing the macroscopic quantities, such as the temperature and the pressure (the macro-states). Let us denote by W the total number of micro-states that give rise to the same macro-state. A typical macroscopic amount of gas has an enormous number of molecules (approximately 6×10^{23} molecules per mole of gas), and the total number of ways these molecules can be distributed is therefore extremely large. To make the numbers manageable, we take the logarithm of W. The entropy S of a gas is therefore defined as

$$S = k_\mathrm{B} \log W \,, \tag{7.61}$$

where we typically take the natural log, defined by

$$\log e^a = a \,. \tag{7.62}$$

Later we will use the logarithm with base 2, defined as $\log_2 2^a = a$, which means that the entropy is measured in bits. We will also work in units where $k_\mathrm{B} = 1$, which amounts to a redefinition of the temperature scale.

Notice that we can think of the entropy as a lack of information, or uncertainty, about the microscopic state. For an ideal gas we can assume that all micro-states are equally likely, and the probability p_i for the system to be in the state i is then given by

$$p_i = \frac{1}{W} \qquad \text{and} \qquad \sum_{i=1}^{W} p_i = 1. \tag{7.63}$$

We can manipulate Boltzmann's equation to obtain

$$\begin{aligned} S = \log W &= \sum_i p_i \log W = \sum_i p_i \log \frac{1}{p_i} \\ &= -\sum_i p_i \log p_i \,, \end{aligned} \tag{7.64}$$

where $k_\mathrm{B} = 1$, and we used $\log a^b = b \log a$ in the last line. This is the formula for the entropy that we will use here. It is important to note that this formula is true also when the p_i are different, which is what we will be using shortly. This expression for the entropy was derived by Claude Shannon in the context of information theory. It was later pointed out to him that this is in fact the same as the Boltzmann entropy $S = k_\mathrm{B} \log W$ for an ideal gas when all the probabilities are the same.

The uncertainty about the microscopic state is a classical uncertainty, and we can interpret the probabilities p_i as the eigenvalues of the density matrix that describes the system:

$$\rho = \begin{pmatrix} p_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & p_W \end{pmatrix} . \tag{7.65}$$

Since W can be huge, this may be a very large matrix. When ρ is diagonal, we can write

$$\log \rho = \begin{pmatrix} \log p_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \log p_W \end{pmatrix} , \tag{7.66}$$

and multiplying this by ρ using matrix multiplication yields

$$\rho \log \rho = \begin{pmatrix} p_1 \log p_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & p_W \log p_W \end{pmatrix} . \tag{7.67}$$

The entropy is then the sum of the diagonal elements of this matrix, with an overall minus sign. This is just the trace of the matrix, so we can write

$$S = -\mathrm{Tr}\left(\rho \log \rho\right) , \tag{7.68}$$

which is the general relationship between the entropy of a system and its density matrix. It also holds when ρ is not a diagonal matrix, but then we must be more careful in taking the logarithm. The expression for S remins true also when W is small.

A curious consequence of this relationship between the entropy and the density matrix of a system is that erasing information will increase the entropy of a system. To see this, consider again the photon in a Mach–Zehnder interferometer with QND detectors that tell us which path inside the interferometer the photon took. After the measurement, we know that the photon is in, say, the upper path:

$$|\text{photon}\rangle = |\text{upper}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} . \tag{7.69}$$

The density matrix of this pure state is given by

$$\rho = |\text{upper}\rangle\langle\text{upper}| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} . \tag{7.70}$$

The eigenvalues of this density matrix are 1 and 0, which means that the entropy of this density matrix is given by

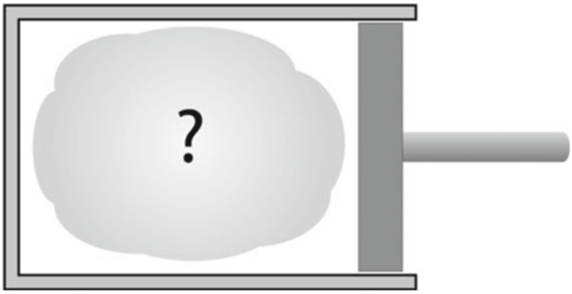

Fig. 7.3 Locating a particle in a box. The interactive figure is available online (see supplementary material 3)

$$S = -0 \log 0 - 1 \log 1 = 0 + 0 = 0\,, \tag{7.71}$$

since the logarithm of 1 is zero, and the logarithm of 0 is multiplied by zero (in other words, we define $0^0 = 1$). Therefore, the pure state $|\text{upper}\rangle$ has zero entropy. This argument can be used to show that *all pure states have zero entropy*.

When we erase the information about the photon path, we become completely uncertain (classically) about the path, and the density matrix must be the maximum uncertainty state:

$$\rho = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}. \tag{7.72}$$

The eigenvalues of this matrix are both 1/2, and the entropy becomes

$$S = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1\,. \tag{7.73}$$

In other words, erasing the path information (one bit: "upper" or "lower") of the photon increases the entropy by 1 bit.

We can use this calculation to solve an interesting paradox: Suppose that we construct a container with a single molecule or atom. One of the walls of the container is a piston that moves in and out of the container without friction (see Fig. 7.3). We can then create an engine as follows: At first, we do not know where the molecule in the container is, so we make a measurement. If we find the molecule in the half closest to the piston we wait a little bit longer, before we measure again. If we find the molecule in the half furthest away from the piston, we quickly move the piston into the container (to the halfway point). We can do this without doing any work, because the molecule is not there to oppose the motion of the piston. We then let the molecule push the piston outwards, which amounts to work done by the molecule on the piston. The energy for this work is supplied by the heat of the molecule (its kinetic energy), and we therefore convert heat into work. The piston returns to its initial position, so we can repeat this experiment until all the heat from the molecule is converted into work. This is called Szilárd's engine (Szilard 1929).

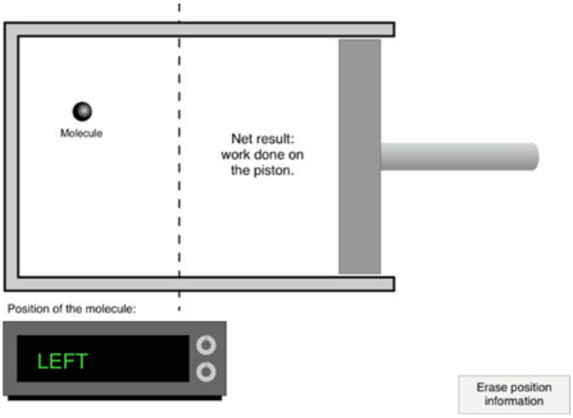

Fig. 7.4 Erasing information will increase the entropy. The interactive figure is available online (see supplementary material 4)

However, this means that the heat of the molecule goes to zero after many rounds of the experiment, and its entropy goes to zero (the molecule ends up in a pure state). Meanwhile, the work done by the molecule can be supplied to a task that does not increase entropy (such as pushing an stationary object up a hill). So the overall entropy seems to decrease, and this violates the second law of thermodynamics!

How do we solve this paradox? We have glossed over one crucial aspect of the thought experiment: after we measure the position of the molecule, we make the decision whether to pull in the piston or not. To make this decision we must record the measurement outcome. This can be done on a computer or in our head, but somewhere in the physical world the information about the molecule's position must be stored if we are to base a decision on it. Therefore, when we repeat this experiment a large number of times, we will build up a long string of measurement outcomes: "left", "right", "left", "left", "right", and so on. On the other hand, for the container and piston to operate in a proper cycle, we must bring the whole engine back to it's initial state. This means that we need to erase the memory of the measurement outcomes.

However, we have seen that the erasure of the path information of a photon leads to an entropy increase of $S = 1$ bit. Similarly, the erasure of the measurement outcome "left" or "right" must be stored in a physical bit whose quantum state can be denoted by the pure state $|\text{left}\rangle$ or $|\text{right}\rangle$. This state has zero entropy. To erase this bit we need to bring the memory to a state of maximum uncertainty, as in Eq. (7.72). We saw that this state has an entropy of $S = 1$.

This increase in entropy turns out to be exactly enough to save the second law of thermodynamics! In general, we can say that the erasure of one bit of information must fundamentally lead to an increase in entropy:

$$S = 1. \tag{7.74}$$

This is known as Landauer's principle (Landauer 1961). It is a surprising connection between information theory and heat (see Fig. 7.4).

Szilárd's engine is closely related to Maxwell's demon, which is a hypothetical being that sits at the gate in a separation between two gasses. By letting through fast molecules in one direction, and slow molecules in the other, the demon can create a temperature difference between the two gasses without doing work or injecting energy. This is also an apparent violation of the second law of thermodynamics…

Exercises

1. Calculate the purity of the state

$$\rho = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} .$$

   Is it a pure state or a mixed state?

2. Two qubits, 1 and 2, are prepared in the initial states

$$|\psi\rangle_1 = \frac{1}{\sqrt{2}}\left(|0\rangle_1 + |1\rangle_1\right) \qquad \text{and} \qquad |\phi\rangle_2 = |0\rangle_2 ,$$

   and the interaction between the two qubits is described by the Hamiltonian

$$H = \hbar g |1\rangle_1\langle 1| \otimes \left(|1\rangle_2\langle 0| + |0\rangle_2\langle 1|\right) .$$

   (a) Calculate the state of the joint two-qubit system after an interaction time T.
   (b) What are the quantum states of the *individual* qubits after they have interacted for a time T?
   (c) Calculate the entropy $S(\rho) = -\mathrm{Tr}\left[\rho \log_2 \rho\right]$ of the individual qubit states ρ.
   (d) For what value of T are the two qubits maximally entangled?

3. (a) A spin-$\frac{1}{2}$ particle is in the state $\rho = \frac{1}{3}|\uparrow\rangle\langle\uparrow| + \frac{2}{3}|\downarrow\rangle\langle\downarrow|$. Calculate the purity of the spin, given by $\mathrm{Tr}\left[\rho^2\right]$.
   (b) The spin evolves in time according to the Hamiltonian

$$H = \hbar\omega X \qquad \text{with} \qquad X = |\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow| .$$

   Calculate the state $\rho(t)$ at time t.
   (c) Sketch the probability of finding the measurement outcome "↑" in a measurement of the spin as a function of time.
   (d) How will the purity of the spin change when all spin coherence is lost?

4. Quantum teleportation with three entangled qubits.

   (a) A so-called GHZ state for three qubits can be written as

$$|\text{GHZ}\rangle = \frac{|000\rangle + |111\rangle}{\sqrt{2}} .$$

After losing the third qubit, what is the state of the remaining two qubits? Calculate the entropy $S(\rho) = -\text{Tr}\big[\rho \log_2 \rho\big]$ of this state.

(b) A so-called W state for three qubits can be written as

$$|\text{W}\rangle = \frac{|001\rangle + |010\rangle + |100\rangle}{\sqrt{3}} .$$

After losing the third qubit, what is the state of the remaining two qubits? Calculate the entropy of this state.

(c) The first qubit of the GHZ state is used as part of the entanglement channel for teleporting a fourth qubit in the state $\alpha|0\rangle + \beta|1\rangle$. Calculate the remaining two-qubit state after applying the teleportation protocol. Assuming perfect equipment, can the remaining two qubits retrieve the original qubit with certainty?

(d) Instead of the GHZ state, use the W state for the teleportation protocol in part (c). Can the remaining two qubits retrieve the original qubit with certainty?

5. The density matrix.

   (a) Show that $\frac{1}{2}|0\rangle + \frac{1}{2}|+\rangle$ is not a properly normalized state.
   (b) Show that $\text{Tr}[\rho] = 1$, and then prove that any density operator has unit trace and is Hermitian.
   (c) Show that density operators are *convex*, i.e., that $\rho = w_1\rho_1 + w_2\rho_2$ with $w_1 + w_2 = 1$ $(w_1, w_2 \geq 0)$, and ρ_1, ρ_2 again density operators.

6. Two qubits, labeled A and B, are prepared in the entangled state $|\Psi\rangle = \frac{3}{5}|00\rangle + \frac{4}{5}|11\rangle$.

   (a) Show that the correlations in the $\{|0\rangle, |1\rangle\}$ basis are perfect, but the correlations in the $\{|+\rangle, |-\rangle\}$ basis are not (think first carefully how you would define a correlation). The states $|+\rangle$ and $|-\rangle$ are defined by

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

   (b) Calculate the entanglement entropy $\mathcal{E}$ of $|\Psi\rangle$. This is a measure of the amount of entanglement in $|\Psi\rangle$, and is given by

   $$\mathcal{E} = S(\text{Tr}[[]B]|\Psi\rangle\langle\Psi|) \qquad \text{with} \qquad S(\rho) = -\text{Tr}\big[\rho \log_2 \rho\big] .$$

   (c) We teleport a state $\alpha|0\rangle + \beta|1\rangle$ using the entangled state $|\Psi\rangle$. What is the best and worst average teleportation fidelity $\langle F\rangle$, using otherwise ideal components? The fidelity F is defined by $F = |\langle\psi_{\text{ideal}}|\psi_{\text{teleported}}\rangle|^2$.

7. The *partial* transpose of a two-qubit state is calculated by writing the density matrix ρ as

$$\rho = \sum_{jklm} \rho_{jk,lm} |j,k\rangle\langle l,m| \qquad \text{with} \qquad j,k,l,m = 0,\ 1$$

and swapping k and m in the kets and bras (the full transpose would be swapping both k and m *and* j and l). If the eigenvalues of the resulting matrix include negative values, then the original state was entangled. Calculate the partial transpose of the states given in Exercise 1 of Chap. 6 (making the identification $|0\rangle \leftrightarrow |\uparrow\rangle$ and $|1\rangle \leftrightarrow |\downarrow\rangle$) and determine their eigenvalues. Which states are entangled? Does it agree with your previous answer?

8. An electron with spin 1/2 is prepared in a state

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & e^{-i\omega t - \gamma t} \\ e^{i\omega t - \gamma t} & 1 \end{pmatrix},$$

where ω and γ are real positive numbers, and t denotes time. Sketch the time evolution of the state in the Bloch sphere. What is the physical interpretation of ω and γ?

9. Any single-qubit mixed state can be written in terms of the Pauli matrices as

$$\rho = \frac{\mathbb{I} + \mathbf{r}\cdot\boldsymbol{\sigma}}{2},$$

where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ and $\mathbf{r} = (r_x, r_y, r_z)$, $|\mathbf{r}| \leq 1$. Moreover, any single qubit observable can be expressed as

$$A = a_0\mathbb{I} + \mathbf{a}\cdot\boldsymbol{\sigma},$$

with a_0 and $\mathbf{a} = (a_x, a_y, a_z)$ real numbers. Show that the expectation value of A with respect to ρ can be written as

$$\langle A\rangle = \frac{1}{2}a_0 + \frac{1}{2}\mathbf{a}\cdot\mathbf{r}.$$

10. The Boltzmann entropy of a system is given by the partition function and the average energy $\langle E\rangle$ as

$$S = k_{\rm B}\left(\ln Z + \beta\langle E\rangle\right).$$

Show that the Shannon entropy in the second line of Eq. (7.64) for a thermal state corresponds to the Boltzmann entropy when we choose $k_{\rm B} = 1$.

11. A two-level atom has energy separation $\hbar\omega$ between the ground state and the excited state. The atom is brought into thermal equilibrium with a bath at temperature T.

   (a) Find the thermal state of the atom.
   (b) Calculate the average energy and the entropy of the atom.
   (c) Let's assume that we prepare the atom in the excited state, and we do not let it thermalise with the bath. What would be the effective temperature of the atom? You can use the expression for the Boltzmann entropy in Exercise 10.

12. Does a thermal state evolve over time?
13. Prove Eq. (7.57).
14. Express the variance of an observable A in terms of the trace and the density operator, similar to Eq. (7.27).

References

R. Landauer, Irreversibility and heat generation in the computing process. IBM J. Res. Dev. **5**, 183 (1961)

L. Szilard, Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen. Zeitschrift für Physik **53**, 840 (1929)