Quantum physics for non-physicists

Lídia del Rio[1], Lorenzo Laneve[2], and Giulia Carocari[2]

[1]Institute for Theoretical Physics, ETH Zurich
[2]Department of Computer Science, ETH Zurich
Fall semester 2020

This is a first formal course into quantum mechanics, through an information-theoretical approach, especially suited for students with backgrounds in computer science, mathematics or engineering. We start from the postulates of quantum theory and build up to the tools needed to study the behaviour of complex systems, from entangled spins to the hydrogen atom and nano heat engines.

This script is a work in progress. It is based on Lídia's lectures, and mostly written by Giulia and Lorenzo. Please send comments and questions to them at gcarocari@student.ethz.ch or llaneve@student.ethz.ch.

> "When you ask what are electrons and protons I ought to answer that this question is not a profitable one to ask and does not really have a meaning. The important thing about electrons and protons is not what they are but how they behave, how they move. I can describe the situation by comparing it to the game of chess. In chess, we have various chessmen, kings, knights, pawns and so on. If you ask what chessman is, the answer would be that it is a piece of wood, or a piece of ivory, or perhaps just a sign written on paper, or anything whatever. It does not matter. Each chessman has a characteristic way of moving and this is all that matters about it. The whole game of chess follows from this way of moving the various chessmen."
>
> Paul Dirac, *Paul Dirac,* Dirac Papers, 2/14/5 (FSU)

Lídia del Rio: lidia@phys.ethz.ch
Lorenzo Laneve: llaneve@student.ethz.ch
Giulia Carocari: gcarocari@student.ethz.ch

Preamble

As scientists, we are like players thrown into a new unknown game, trying to figure out the rules as we go along. So it was in the early XX century with the discovery of quantum physics: scientists would make experiments (or try plays) that didn't result in the expected outcomes, and adapt their hypotheses to explain the underlying rules, making comparisons to known games ("it is a little like wave dynamics here, a little like Poisson evolution there" is similar in spirit to "this part of the game reminds me of poker, but that mechanism is more like bridge"), and slowly letting go of their initial assumptions, until eventually people like Heisenberg and Schrödinger had the first big breakthroughs in capturing the rules of Nature at that scale. Now, if you were in that position and wanted to teach a new player about the game, how would you go about it? Would you spend weeks explaining the process of how you managed to figure out the rules, or would you try to explain the rules in a direct way, to allow the new person to start playing? Traditional quantum mechanics courses start with the historical perspective, which is precisely a version of "we didn't know the rules, tried that play, postulated, tested and abandoned all these hypotheses, and now we have this theory as to how the game works." Which would be fine if the story explained *why* the rules are as they are — but it doesn't, and as such we subject students to this long tale of how we got to our current understanding, without it enlightening why the world works the way it does. Here we try to first give you the rules of quantum mechanics, and explore what we can do with them, before we start to question them.

Course logistics

Schedule. Due to the pandemic, all the lectures and tutorials will be online. They will be streamed on Zoom, and later made publicly available on YouTube. Office hours are also online.

Date	**Time**	**Type**	**Zoom meeting id**
Tuesday	9:45 – 11:30	Lecture	935 7294 4577
Tuesday	15:00 – 17:00	Open office hours	917 2922 0926 (book by email)
Thursday	9:45 – 11:30	Tutorial	939 9366 1901
Thursday	11:45 – 12:30	Lecture	935 7294 4577

Resources. All course materials are available on Moodle. This includes the tablet handwritten notes, links to all the video recordings, and these lecture notes. If you don't have access to Moodle (but somehow got these notes) and would like to see the materials, email me.

Examination. There will be a final written exam, which will be two hours long. We will provide a sample exam before this.

Pre-requisites. You should be able to multiply matrices and have seen integrals before. Usually, first-year courses of linear algebra and calculus ensure this. If I use terms that you are not familiar with, don't hesitate to ask. Appendices give a quick recap of the main concepts we will use.

Overlap with other ETH courses. There is some overlap in content with Quantum Mechanics 1 and 2, but it is a different approach, and we will learn things in a different order. There is some overlap in approach with Quantum Information Theory, but after the first couple of weeks the topics covered will diverge: QIT focuses on more abstract things like transmitting

information, cryptography and teleportation, whereas we will study the physics of quantum mechanics, with applications to concrete systems. It is still worth taking both courses. Similarly, in the spring semester there's a course called Quantum Information Processing I, which is essentially an introduction to quantum computing — so it uses the same basic formalism, but the focus is very different. There is in parallel a course called QIP II which is about physical implementations of quantum computing, for which the present course is relevant. As two possible follow up courses to this one, in Spring we'll have Advanced Topics in QIT (focusing on quantum thermodynamics, quantum clocks, quantum reference frames and foundations) by Ralph Silva and myself, and Quantum Sensing and Metrology by Mischa Woods.

A note on Covid19. This is a strange and difficult year. Online lectures lose their novelty pretty quick, and studying from home can be isolating and grating. Please take it easy, and prioritize your physical and mental well-being. Quantum mechanics can wait.

Contents

Chapter 1
Introduction

This chapter introduces the formalism we need to describe quantum systems. For a more extensive discussion, we refer the reader to Chapter 3 (*States and observables*) of Schumacher and Westmoreland [1]. The chapter also recaps many useful notions and results from linear algebra. As regards continuous state spaces, you may want to look at Chapter 10 (*A particle in space*) of the same book, sections 10.1 and 10.2 in particular.

1.1 Hilbert spaces and the braket notation

Definition 1.1. *An inner product space $(\mathcal{X}, \cdot)$ over a complex field and inner product $\cdot$ is said to be a Hilbert space when the distance d induced by the inner product forms a complete metric space, i.e. a metric space where every Cauchy sequence has a limit contained in the space itself.*

The definition of a Hilbert space is important as it guarantees that any converging infinite or integral sum of elements in the space is still in the space. In quantum physics, a physical system is represented by a Hilbert space over complex field, and a (column) vector in this space represents a possible **state** of the system. A column vector is written in the form $|\psi\rangle$, denoted as **ket**. We define the transpose conjugate of $|\psi\rangle$ as

$$\langle\psi| = (|\psi\rangle)^\dagger$$

and this row vector is what we call a **bra**. The inner product of the Hilbert space between two vectors $|\phi\rangle$, $|\psi\rangle$ can then be expressed in the following way:

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

and this is the **bra-ket** notation for the inner product. A bra can be also seen as a function:

$$\begin{aligned}\langle\phi| : \mathcal{H} &\to \mathbb{C} \\ |\psi\rangle &\mapsto \langle\phi|\psi\rangle\end{aligned}$$

we will see later that this view is not unusual, and it is used to extract information from the ket in a very convenient way. It is also possible to define an outer product between states, called **ket-bra**:

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

Like in standard linear algebra, while the inner product is a scalar (in our case, a complex number), the outer product is an **operator**[1]. In this way we can easily construct transformations (or better, endomorphisms) on $\mathcal{H}$.

Discrete spaces and qubits. When we have to represent a system with a finite or countably infinite number of states, we use a Hilbert space spanned by a basis with a discrete number of vectors:

$$\mathcal{H} = \text{span}\{|x\rangle\}_x \quad \ni \quad |\psi\rangle = \sum_x \psi_x |x\rangle, \quad \psi_x \in \mathbb{C}\ \forall x, \quad \sum_x |\psi_x|^2 = 1.$$

[1] You may be more familiar with the term "matrix". The term operator denotes an extension of the concept of a matrix to vector spaces with infinite dimensions, which may be the case in quantum physics. The idea, however, remains the same and you can still imagine an operator as a matrix.

The last constraint is what we call **normalization**, i.e. vectors of the Hilbert space representing a state must have unitary norm (we will see later why this property is important, when we will talk about probability of outcomes). The special case where we only have two basis vectors is called **qubit**:

$$\mathcal{H} = \text{span}\{|0\rangle, |1\rangle\} \implies |\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad |\alpha|^2 + |\beta|^2 = 1$$

$\{|0\rangle, |1\rangle\}$ is usually called the **computational basis** of a qubit.

Infinite continuous dimensions. A continuous space is used when we need to deal with physical systems that involve continuous variables (e.g. position in space):

$$\mathcal{H} = \text{span}\{|x\rangle\}_{x\in\mathbb{R}} \quad \ni \quad |\psi\rangle = \int_{\mathbb{R}} \psi(x)|x\rangle dx, \quad \int_{\mathbb{R}} |\psi(x)|^2 dx = 1$$

In both cases $|\psi\rangle$ is an arbitrary state of the system which can be expressed in terms of a basis of the Hilbert space. From now on we will assume that all the bases we use are **orthonormal**, it will be clear later why this is important.

1.2 The wave function

We expressed a state $|\psi\rangle$ in terms of a basis of the Hilbert space of a system. Let us consider the continuous case (the discrete case is analogous):

$$\psi = \int_{\mathbb{R}} \psi(x)|x\rangle dx$$

where $\psi(x)$ is a function containing the components of the vector $|\psi\rangle$ with respect to the basis $\{|x\rangle\}$. It is called the **wave function**. Let us now compute the inner product $\langle x|\psi\rangle$:

$$\begin{aligned}
\langle x|\psi\rangle &= \langle x| \int_{\mathbb{R}} \psi(x')|x'\rangle dx' \\
&= \int_{\mathbb{R}} \psi(x')\langle x|x'\rangle dx' \\
&= \int_{\mathbb{R}} \psi(x')\delta(x - x')dx' \\
&= \psi(x)
\end{aligned}$$

where $\delta(x)$ is the **Dirac delta function**, and it follows from the fact that the basis $\{|x\rangle\}_x$ is orthonormal. A short introduction to the Dirac delta is given in Appendix A. Therefore, the inner product of a state with a basis element yields the value of the wave function with respect to that particular element, i.e. the projection of $|\psi\rangle$ onto $|x\rangle$.

1.3 Measurements

The wave function has an important physical meaning. Consider a basis $\{|x\rangle\}_x$ of the Hilbert space of a physical system: we would like to **measure** the system with respect to this basis. The first property of a quantum system is that, when we measure it with respect to a basis, the result will be an element of the basis. A particular element $|x\rangle$ is the result of the measurement with probability:

$$|\langle x|\psi\rangle|^2 = |\psi(x)|^2$$

where $|\psi\rangle$ is the state of the system at the moment of the measurement, and $\psi(x)$ is the corresponding wave function. Another important property is that, when the measurement happens, the state **collapses** to the measured basis element: if we measure a system and we read an element $|x\rangle$, then $|x\rangle$ will be the new state of the system.

As an analogy to better grasp this concept, consider throwing a dice and "measuring" the outcome by looking at it; if we cover it or look away for a while, we still expect to see the same number on top after checking the dice again.

1.4 Probability of outcomes in a measurement

For a concrete example, consider the position of an electron on a line (this will be our running example throughout this chapter). Let $\{|x\rangle\}_x$ be the position basis of the Hilbert space, i.e. it consists of states representing a point $x \in \mathbb{R}$, which is the position of the electron. The formalization of a measurement with respect to this basis induces a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, where $\Omega = \{x\}_{x\in\mathbb{R}}$ consists of all the elements of the basis, and the probability of each element is given by:

$$\mathbf{P}(x) = |\psi(x)|^2 dx$$

The differential operator dx appears because here we are considering a continuous space. In the case of a discrete system the induced probability space is discrete, and the probability can be defined accordingly. In any case, one can see now why we required a state to be a **normalized** vector.

Therefore, coming back to our electron on a line example, if we want to know the probability of measuring the position of the electron in a certain interval $[a, b]$, we can compute it as:

$$\mathbf{P}([a,b]) = \int_a^b \mathbf{P}(x)\, dx = \int_a^b |\psi(x)|^2 dx$$

If we define $P_{[a,b]}$ as the projection operator in the subspace spanned by $\{|x\rangle\}_{x\in[a,b]}$, we can rewrite the probability above in the following way:

$$\begin{aligned}
\langle\psi|P_{[a,b]}|\psi\rangle &= \langle\psi| \left(\int_a^b |x\rangle\langle x| dx \right) |\psi\rangle \\
&= \int_a^b \langle\psi|x\rangle\langle x|\psi\rangle dx \\
&= \int_a^b |\langle x|\psi\rangle|^2 dx \\
&= \int_a^b |\psi(x)|^2 dx \\
&= \mathbf{P}([a,b])
\end{aligned}$$

Therefore, it is sufficient to compute the inner product of $|\psi\rangle$ with a projection operator onto a subspace, and we get the probability that the state collapses into that subspace upon measurement.

At this point, it can be useful to introduce the following terminology:

Definition 1.2 (Superposition)**.** *Consider a system in a state $|\psi\rangle$ and let $\{|x\rangle\}_x$ be a basis of the corresponding Hilbert space. $|\psi\rangle$ is a **basis state** with respect to $\{|x\rangle\}_x$ if $|\psi\rangle = |x\rangle$ for some basis element $|x\rangle$. In any other case, $|\psi\rangle$ is said to be in a **superposition** of the elements of the basis $\{|x\rangle\}_x$.*

From what we saw above, one can imagine that measuring a basis state will yield a trivial probability space, where the probability of measuring the element of the basis equal to the state is 1. Keep in mind that both the induced probability space and the notion of basis state and superposition are **relative to the particular basis** we use for measurement: for every state $|\psi\rangle$ there is a basis in which $|\psi\rangle$ is a basis state and a basis in which $|\psi\rangle$ is in superposition.

Representing a measurement with projectors. It is often useful, especially when we talk about continuous systems, to not measure with respect to single basis elements, but to group elements of the measurement basis in **projector operators** like the one above. For example, if we only want to know whether a particle is on the left or on the right of a certain position L, we can simply divide the identity $\mathbb{1}$ into two projectors:

$$\mathbb{1} = P_{(-\infty,L)} + P_{(L,+\infty)}$$

Each of these projectors can be used to compute the collapse probability as above.

Moreover, it naturally follows that any projector P_A is idempotent:

$$\begin{aligned} P_A^2 &= \iint_{A^2} |x\rangle\langle x|x'\rangle\langle x'|dxdx' \\ &= \iint_{A^2} |x\rangle\delta(x-x')\langle x'|dxdx' \\ &= \int_A |x\rangle\langle x|dx = P_A \end{aligned}$$

Global and relative phase. Since the Hilbert space of a quantum system is complex, we would also like to understand why **phases** are important. We distinguish two cases: the first is called **global phase**, which is a phase $e^{i\phi}$ that multiplies the whole state of a system, and we now prove that this term has no physical meaning.

Theorem 1.3. *Let $|\psi\rangle$ be a vector in a Hilbert space representing the state of a quantum system, and consider $\phi \in [0, 2\pi)$. The vectors $|\psi\rangle$ and $e^{i\phi}|\psi\rangle$ represent the same state.*

Proof. Consider an arbitrary basis of the Hilbert space $\{|x\rangle\}_x$. The measurement of $e^{i\phi}|\psi\rangle$ induces a probability space $(\Omega, \mathcal{F}, \mathbf{P}_\phi)$ such that:

$$\mathbf{P}_\phi(x) = |\langle x|e^{i\phi}|\psi\rangle|^2 = |e^{i\phi}\langle x|\psi\rangle|^2 = |\langle x|\psi\rangle|^2$$

Therefore the probability spaces induced by the two states are equal with respect to any measurement basis. □

On the other hand, we have a **relative phase** when different components of the state vector have different phases. In any case, we may represent a state with a vector where one of the components is real, i.e. normalize the phases. For example, in the case of a qubit:

$$|\psi\rangle = ae^{i\phi_A}|0\rangle + be^{i\phi_B}|1\rangle = e^{i\phi_A}\left(a|0\rangle + be^{i(\phi_B-\phi_A)}|1\rangle\right)$$

1.5 Observables

An observable is an operator representing a certain quantity of the system we want to observe.

$$A = \int_{\mathbb{R}} f(x)|x\rangle\langle x|dx$$

where $\{|x\rangle\}$ is a basis of the Hilbert space and $f(x) \in \mathbb{R}$ is the observed quantity, i.e. $f(x)$ is the quantity we would observe if the system were in state $|x\rangle$ upon measurement. Obviously A is Hermitian since:

$$A^\dagger = \left(\int_{\mathbb{R}} f(x)|x\rangle\langle x|dx\right)^\dagger = \int_{\mathbb{R}} f(x)\left(|x\rangle\langle x|\right)^\dagger dx = \int_{\mathbb{R}} f(x)|x\rangle\langle x|dx = A$$

In particular, notice that $f(x)$ is the eigenvalue associated with the eigenvector $|x\rangle$ of A. These eigenvalues are also called **labels** of the observable.

For example, in the case of the position of an electron, the observable of the position is:

$$X = \int_{\mathbb{R}} x|x\rangle\langle x|dx$$

If we want to compute the expected value of a quantity in a state $|\psi\rangle$ we can simply compute the inner product with the corresponding observable. In the case of the position of an electron we have:

$$\begin{aligned}\langle X\rangle &= \langle x|X|x\rangle \\ &= \langle\psi|\left(\int_{\mathbb{R}} x|x\rangle\langle x|dx\right)|\psi\rangle \\ &= \int_{\mathbb{R}} x\langle\psi|x\rangle\langle x|\psi\rangle dx \\ &= \int_{\mathbb{R}} x|\psi(x)|^2 dx\end{aligned}$$

In a probabilistic formalization, one can see an observable as a **random variable**: the observable A above assumes value $f(x)$ if the event $x \in \Omega$ occurs.

1.6 Post-measurement state

Suppose that we measure a state $|\psi\rangle$ with a set of projection operators $\{P_A, P_B, P_C\}$, and that this causes a collapse of the state in the subspace of P_A. How can we compute the post-measurement state? We know that the new state must be the projection of $|\psi\rangle$ onto the subspace of P_A, and we also need to impose that the new state is normalized. This leads to:

$$\begin{aligned}|\psi'\rangle = \frac{P_A|\psi\rangle}{|P_A|\psi\rangle|} &= \frac{P_A|\psi\rangle}{\sqrt{\langle\psi|P_A^2|\psi\rangle}} = \frac{P_A|\psi\rangle}{\sqrt{\langle\psi|P_A|\psi\rangle}} \\ &= \frac{1}{\sqrt{\langle\psi|P_A|\psi\rangle}}\int_A |x\rangle\langle x|dx\int_{\mathbb{R}}\psi(x')|x'\rangle dx' \\ &= \frac{1}{\sqrt{\langle\psi|P_A|\psi\rangle}}\int_A dx\int_{\mathbb{R}}\psi(x')|x\rangle\langle x|x'\rangle dx' \\ &= \frac{1}{\sqrt{\langle\psi|P_A|\psi\rangle}}\int_A\int_{\mathbb{R}}\psi(x')|x\rangle\delta(x-x')dxdx' \\ &= \int_A \frac{\psi(x)}{\sqrt{\langle\psi|P_A|\psi\rangle}}|x\rangle dx \stackrel{!}{=} \int_{\mathbb{R}}\psi'(x)|x\rangle dx\end{aligned}$$

Therefore, the collapsed wave function is:

$$\psi'(x) = \begin{cases}\frac{\psi(x)}{\sqrt{\langle\psi|P_A|\psi\rangle}} & x \in A \\ 0 & x \notin A\end{cases}$$

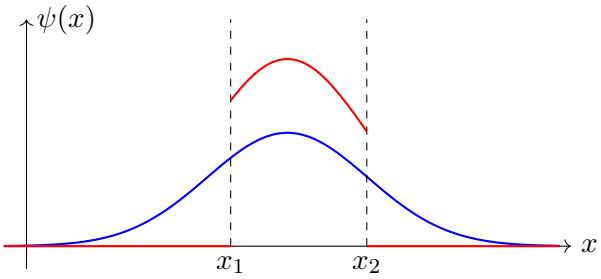


Figure 1.1: Collapse of the wave function after the particle has been measured in the region between x_1 and x_2.

Example. Assume we are measuring the position of a free particle travelling on a line. In Figure 1.1 we observe how, given that we have measured that the particle is located between points x_1 and x_2, the original wave function $\psi(x)$ (in blue) collapses to the wave function of the post measurement state $\psi'(x)$ (in red).

1.7 Qubits: the Bloch sphere

Consider a qubit in an arbitrary state $|\psi\rangle$:

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

where $\alpha, \beta \in \mathbb{C}$. We can rewrite these two numbers in complex exponential form:

$$\alpha = ae^{i\theta_a}, \beta = be^{i\theta_b}$$

Now, if we let $\phi = \theta_b - \theta_a$ and factorize $e^{i\theta_a}$ like we did in Section 1.4, we can remove the global phase and assume that the component of $|0\rangle$ is real:

$$|\psi\rangle = a|0\rangle + be^{i\phi}|1\rangle$$

Notice that a qubit has exactly 2 degrees of freedom: in general we have 2 absolute values and 2 phases, but irrelevance of the global phase and normalization, i.e. $|\alpha|^2 + |\beta|^2 = 1$, remove one degree of freedom each. In order to emphasize the fact that we have two degrees of freedom we parameterize a and b:

$$\begin{cases} a = \cos(\theta/2) \\ b = \sin(\theta/2) \end{cases} \implies |\psi\rangle = \cos(\theta/2)|0\rangle + \sin(\theta/2)e^{i\phi}|1\rangle$$

If we consider (θ, ϕ) as spherical coordinates, these span a point set in $\mathbb{R}^3$ called the **Bloch sphere**. In fact, one can see the xy-plane intersecting this sphere as the complex plane. With this in mind, $\theta \in [0, \pi]$ is the height angle, while $\phi \in [0, 2\pi]$ can be seen as the azimut. Let us look as some special cases (depicted in Figure 1.2):

- When $\theta = 0$, then $|\psi\rangle = |0\rangle$ regardless of ϕ;
- When $\theta = \pi$, we have $|\psi\rangle = e^{i\phi}|1\rangle$, and ϕ becomes a global phase, implying that the resulting state is equivalent to $|1\rangle$ for any choice of ϕ.

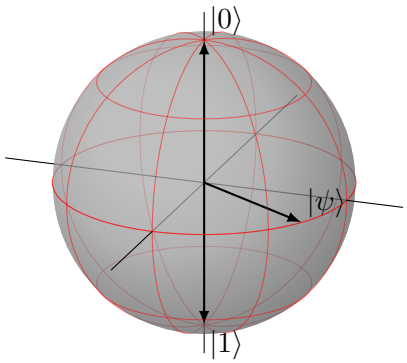


Figure 1.2: A visual representation of the Bloch Sphere and various states.

- When $\theta = \pi/2$, the vector is in the complex plane, and the state is in the form:

$$|\psi\rangle = \frac{|0\rangle + e^{i\phi}|1\rangle}{\sqrt{2}}$$

Another interesting property of the Bloch sphere is that two elements of a basis (i.e. two orthogonal elements) are always in opposite points of the sphere.

Chapter 2
Reversible Evolution

2.1 Evolution in qubits: quantum gates

An evolution of the state $|\psi\rangle$ of a quantum system can be expressed as a unitary transformation in its Hilbert space (the definition of unitary operator can be found in Appendix B). In the case of a qubit, we have 2×2 matrices that represent **single-qubit quantum gates**. The first quantum gate we discuss is the **Hadamard gate**:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

This matrix transforms the components of the computational basis as follows:

$$H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} = |+\rangle, \quad H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}} = |-\rangle$$

where $\{|+\rangle, |-\rangle\}$ forms a basis of the qubit space called **Hadamard basis**.

Other interesting quantum gates are the so called **Pauli matrices**:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

One can see that X acts exactly as a NOT gate with respect to the computational basis:

$$X|0\rangle = |1\rangle, \quad X|1\rangle = |0\rangle$$

while it acts as follows on the Hadamard basis:

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

that is, the Hadamard basis is the eigenbasis of X (with $+1$ and -1 being the eigenvalues associated to $|+\rangle$ and $|-\rangle$ respectively). On the other hand, one can see that Z acts in the exact opposite way as X: while it swaps $|+\rangle$ and $|-\rangle$, the computational basis is its eigenbasis.

Pauli matrices as observables. Since X, Y, Z are Hermitian, they can also be seen as observables. In particular notice that:

$$X = |+\rangle\langle+| - |-\rangle\langle-|$$
$$Z = |0\rangle\langle 0| - |1\rangle\langle 1|$$

Therefore, if we take for example X, a $|+\rangle$ is measured with a label $+1$, while a $|-\rangle$ is measured with a label -1.

One last thing to notice is that transforming a state $|\psi\rangle$ with an operator U before measuring with respect to an observable M is equivalent to a measurement with respect to the following observable:

$$M' = UMU^\dagger$$

In order to see this, consider the spectral decomposition of M:

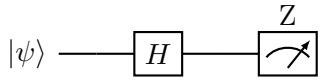
$$M = V\Lambda V^\dagger = \sum_i \lambda_i |v_i\rangle\langle v_i| \implies M' = U(V\Lambda V^\dagger)U^\dagger = \sum_i \lambda_i U|v_i\rangle\langle v_i|U^\dagger$$

implying that, while the labels do not change, the eigenvectors of the new observables are $\{U|v_i\rangle\}_i$. Keep in mind that this also works for infinite-dimensional spaces.

Example with qubits. Assume we want to perform a measurement in the X basis. This is equivalent to first evolving the system via the H gate and then measuring with respect to the Z basis, as:

$$\begin{aligned} HZH^\dagger &= H(|0\rangle\langle 0| - |1\rangle\langle 1|)H && \text{since } H \text{ is Hermitian} \\ &= H(|0\rangle\langle 0|)H - H(|1\rangle\langle 1|)H \\ &= |+\rangle\langle +| - |-\rangle\langle -| = X \end{aligned}$$

In general, we represent the evolution of qubit systems using quantum circuit diagrams. This transform-then-measure example would look as follows:



We read the line from left to right as an evolution over time. A gate is represented as rectangle enclosing an identifier for the type of gate, whereas the meter symbol is used to represent a measurement, specifically in the Z basis. We will go more into detail in Chapter 3, when analysing systems with multiple qubits.

2.2 Unitary dynamics

We already said that a reversible evolution is expressed with a unitary operator. Note that unitarity implies preservation of the inner products:

$$|\psi'\rangle = U|\psi\rangle, \quad |\phi'\rangle = U|\phi\rangle \quad \implies \quad \langle\phi'|\psi'\rangle = \langle\phi|U^\dagger U|\psi\rangle = \langle\phi|\psi\rangle$$

and, in particular, $\langle\psi|\psi\rangle$ remains 1 over time, for any state $|\psi\rangle$.

Moreover, we know that a unitary operator U admits an orthonormal eigenbasis:

$$U = \sum_k u_k |k\rangle\langle k|$$

$$1 = \langle k|k\rangle = \langle k|U^\dagger U|k\rangle = \langle k|u_k^* u_k|k\rangle = |u_k|^2\langle k|k\rangle = |u_k|^2$$

implying that every eigenvalue has unitary absolute value. Thus, we can directly express U as:

$$U = \sum_k e^{i\alpha_k}|k\rangle\langle k|$$

In this case, $\{|k\rangle\}_k$ are said to be the **eigenstates** of the evolution U.

2.3 Deriving the Schrödinger equation

We consider a state evolving over time:

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle$$

The temporal derivative of $|\psi(t)\rangle$ can easily be defined:

$$\begin{aligned}
\frac{\partial}{\partial t}|\psi(t)\rangle &= \lim_{\Delta t \to 0} \frac{|\psi(t+\Delta t)\rangle - |\psi(t)\rangle}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{U(t+\Delta t, t)|\psi(t)\rangle - |\psi(t)\rangle}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{U(t+\Delta t, t) - \mathbb{1}}{\Delta t}|\psi(t)\rangle \\
&= G|\psi(t)\rangle
\end{aligned}$$

We found that the temporal derivative of a state can be expressed as a linear operator.

Theorem 2.1. *G is anti-Hermitian.*

Proof. We know that $\langle\psi(t)|\psi(t)\rangle = 1$, thus it does not change over time:

$$\begin{aligned}
0 &= \frac{\partial}{\partial t}\left(\langle\psi(t)|\psi(t)\rangle\right) \\
&= \frac{\partial}{\partial t}\left(\langle\psi(t)|\right)|\psi(t)\rangle + \langle\psi(t)|\frac{\partial}{\partial t}\left(|\psi(t)\rangle\right) \\
&= \langle\psi(t)|G^\dagger|\psi(t)\rangle + \langle\psi(t)|G|\psi(t)\rangle \\
&= \langle\psi(t)|\left(G^\dagger + G\right)|\psi(t)\rangle
\end{aligned}$$

Thus $G + G^\dagger = 0$. □

We now define $H = i\hbar \cdot G$ as the **Hamiltonian** of the system, where $\hbar$ is Planck's constant, measured in Joule times second $[J \cdot s]$. Notice that the Hamiltonian is Hermitian, and it can be seen as an observable of the total energy of the system: in fact, G is expressed in inverse seconds $[s^{-1}]$ (since it is a temporal derivative operator) and, together with the Planck constant, we get that the eigenvalues of H are expressed in Joule $[J]$.

The definition of the Hamiltonian yields the **Schrödinger equation**:

$$H|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle.$$

Chapter 3
Composing Systems

3.1 Tensor product

Consider two Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$. Given $|x\rangle \in \mathcal{H}_A, |y\rangle \in \mathcal{H}_B$ we define a tensor product:

$$|x\rangle_A \otimes |y\rangle_B$$

where we usually explicit the subscripts A, B on the kets indicating which space the states belong to. For any $|x_1\rangle, |x_2\rangle \in \mathcal{H}_1, |y_1\rangle, |y_2\rangle \in \mathcal{H}_2$, we have that:

- The tensor product is **distributive** over addition

$$(|x_1\rangle + |x_2\rangle)_A \otimes |y_1\rangle_B = |x_1\rangle_A \otimes |y_1\rangle_B + |x_2\rangle_A \otimes |y_1\rangle_B$$
$$|x_1\rangle_A \otimes (|y_1\rangle + |y_2\rangle)_B = |x_1\rangle_A \otimes |y_1\rangle_B + |x_1\rangle_A \otimes |y_2\rangle_B$$

- Scalar constants $a \in \mathbb{C}$ can be taken out of the product

$$(a|x_1\rangle)_A \otimes |y_1\rangle_B = |x_1\rangle_A \otimes (a|y_1\rangle)_B = a(|x_1\rangle_A \otimes |y_1\rangle_B)$$

- The tensor product of operators is applied independently to each component:

$$(U_1 \otimes U_2)(|x_1\rangle_A \otimes |y_1\rangle_B) = (U_1|x_1\rangle_A) \otimes (U_2|y_1\rangle_B)$$

- Inner product acts linearly on the tensor product (i.e. the order of application of inner product and tensor product can be reversed):

$$(\langle x_1|_A \otimes \langle y_1|_B)(|x_2\rangle_A \otimes |y_2\rangle_B) = \langle x_1|x_2\rangle\langle y_1|y_2\rangle$$

From now on, we will write $|x\rangle \otimes |y\rangle$ when the spaces we are referring to are clear from the context. It is also possible to find $|x\rangle|y\rangle$ when it is clear that a tensor product is involved, or even $|xy\rangle$ when it is clear which state belongs to which space.

We extend the definition of the tensor product to Hilbert spaces:

$$\mathcal{H}_1 \otimes \mathcal{H}_2 = \text{span}\,\{|x\rangle \otimes |y\rangle \mid |x\rangle \in \mathcal{H}_1, |y\rangle \in \mathcal{H}_2\}$$

3.2 Qubits: multiple qubit gates

We will now see how to use the tensor product, starting with two qubits. Consider a product space $\mathcal{H}_A \otimes \mathcal{H}_B$. We know that, in the case of a qubit, the state space of each subspace is $\mathbb{C}^2$, and the elements of the computational basis can be expressed using the classical vector notation:

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

How does the tensor product work here? We can resort to its properties: for example, let us compute $|0\rangle|1\rangle$

$$|0\rangle|1\rangle = (1|0\rangle + 0|1\rangle)(0|0\rangle + 1|1\rangle) = 0|0\rangle|0\rangle + 1|0\rangle|1\rangle + 0|1\rangle|0\rangle + 0|1\rangle|1\rangle$$

If we conveniently choose $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ as the standard basis of our product space (in this order), we can write $|01\rangle$ as:

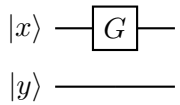
$$|0\rangle|1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

This is convenient because, this way, the element $|x\rangle$, where x is the binary representation of a natural number n, is exactly the vector where the n-th component (starting from 0) is set to 1, and all the others to 0. In fact, after some computation one can find:

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

An interesting property here is that n qubits yield a state space isomorphic to $\mathbb{C}^{2^n}$, i.e. the dimension of the state space of a system of n qubits grows exponentially fast with n. This is one of the crucial factors that allows quantum computation to outspeed any model of computation involving classical systems.

Single-qubit gates in multiple-qubit systems. Before talking about gates involving multiple inputs, we discuss how to represent gates acting only on one of the qubits at a time. Suppose to have a quantum circuit with two qubits, where we apply a gate G to the first qubit and leave the second qubit unaltered.



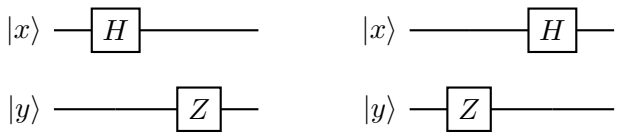
Leaving a qubit unaltered is conceptually equivalent to applying the identity $\mathbb{1}$ to it. Therefore, the above circuit can be represented with the following "product gate":

$$G \otimes \mathbb{1} \Longrightarrow (G \otimes \mathbb{1})|xy\rangle = G|x\rangle \otimes \mathbb{1}|y\rangle = G|x\rangle \otimes |y\rangle$$

Example. Suppose $G = H$ is the Hadamard gate and the input of the circuit is $|00\rangle$. The output of the circuit will be:

$$(H_A \otimes \mathbb{1}_B)|00\rangle_{AB} = H_A|0\rangle_A \otimes |0\rangle_B = |+\rangle_A \otimes |0\rangle_B = \frac{|00\rangle + |10\rangle}{\sqrt{2}}$$

As another example, imagine that H is applied to the first qubit and then Z is applied to the second, as follows:



This is equivalent to applying $(H \otimes \mathbb{1})$ and then $(\mathbb{1} \otimes Z)$ to both qubits, i.e.:

$$(H \otimes \mathbb{1})(\mathbb{1} \otimes Z)|xy\rangle = (H \cdot \mathbb{1}) \otimes (\mathbb{1} \cdot Z)|xy\rangle = (H \otimes Z)|xy\rangle$$

The calculation suggests that this is also equivalent to first applying Z and then H, which makes sense, since the qubits are not interacting with each other.

(a) Controlled NOT gate (b) Controlled Q gate

Figure 3.1: Quantum circuit diagram representation of controlled gates.

A first two-qubit gate: the CNOT gate. We now introduce a gate involving two qubits, which is the *CNOT* gate:

$$CNOT = \bigwedge X = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes X$$

As hinted by the definition, this gate has a clear meaning in the computational basis: if the first qubit is $|0\rangle$, the second qubit is left unchanged

$$\left(\bigwedge X\right)|0x\rangle = |0\rangle\langle 0|0\rangle \otimes \mathbb{1}|x\rangle + |1\rangle\langle 1|0\rangle \otimes X|x\rangle = |0x\rangle$$

otherwise, having $|1\rangle$ on the first qubit causes the second qubit to flip:

$$\left(\bigwedge X\right)|1x\rangle = |0\rangle\langle 0|1\rangle \otimes \mathbb{1}|x\rangle + |1\rangle\langle 1|1\rangle \otimes X|x\rangle = |1\rangle \otimes X|x\rangle = |1\overline{x}\rangle$$

where $\overline{x}$ denotes the negation of the bit x. Here we give the matrix representation with respect to the standard computational basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$:

$$\bigwedge X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Another way to see the CNOT gate is the fact that it implements a XOR in the computational basis (reading the result off the second qubit):

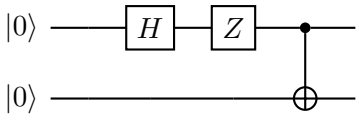
$$\left(\bigwedge X\right)|xy\rangle = |x\rangle|x \oplus y\rangle$$

The CNOT is a widely used gate, to the point that it has its own symbol in the graphical notation for quantum circuits. The concept can be generalized to an arbitrary gate Q:

$$\bigwedge Q = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes Q$$

i.e. the Q gate is applied to the second qubit if the first qubit is $|1\rangle$. The visual representation of these gates is depicted in Figure 3.1.

Example. Suppose to have the following circuit with two qubits:



If the input is $|00\rangle$, the output will be:

$$\left(\bigwedge X\right)(Z_A \otimes \mathbb{1}_B)(H_A \otimes \mathbb{1}_B)|00\rangle = (|0\rangle\langle 0|_A \otimes \mathbb{1}_B + |1\rangle\langle 1|_A \otimes X_B)(Z_A H_A \otimes \mathbb{1}_B)|00\rangle$$

$$
\begin{aligned}
&= (|0\rangle\langle 0|_A \otimes \mathbb{1}_B + |1\rangle\langle 1|_A \otimes X_B)(Z_A H_A |0\rangle_A \otimes \mathbb{1}_B |0\rangle_B) \\
&= (|0\rangle\langle 0|_A \otimes \mathbb{1}_B + |1\rangle\langle 1|_A \otimes X_B)(Z_A |+\rangle_A \otimes |0\rangle_B) \\
&= (|0\rangle\langle 0|_A \otimes \mathbb{1}_B + |1\rangle\langle 1|_A \otimes X_B)(|-\rangle_A \otimes |0\rangle_B) \\
&= |0\rangle\langle 0|-\rangle_A \otimes \mathbb{1}_B |0\rangle_B + |1\rangle\langle 1|-\rangle_A \otimes X_B |0\rangle_B \\
&= |0\rangle\langle 0|-\rangle_A \otimes |0\rangle_B + |1\rangle\langle 1|-\rangle_A \otimes |1\rangle_B \\
&= \frac{|00\rangle - |11\rangle}{\sqrt{2}}
\end{aligned}
$$

The last equality follows from the fact that $\langle 0|-\rangle = \frac{1}{\sqrt{2}}, \langle 1|-\rangle = -\frac{1}{\sqrt{2}}$.

3.3 Entanglement and measurement

At the end of the previous section the quantum circuit ended up with this output:

$$|\psi\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}$$

What is the state of the first qubit in this case? Actually there is no answer to this question, in the sense that there is not a well-defined state for the single qubits. This is because the states of the two qubits are somehow tied to each other, i.e. the two qubits are **entangled**.

Definition 3.1 (Quantum entanglement)**.** *Let $\mathcal{H}_1, \mathcal{H}_2$ be the state spaces of two sub-systems, and consider the product space $\mathcal{H}_1 \otimes \mathcal{H}_2$. A state $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ is said to be **unentangled** (or product state) if it can be written as a tensor product of the states of the single subsystems:*

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$$

*Otherwise, the state is said to be **entangled**.*

In order to give more intuition about entanglement, we talk about something more familiar: *independence of random variables in probability.* We can think of the two subsystems as two random variables: if the state of the whole system is unentangled, we can think of them as two independent random variables, where knowing something about one system does not tell us anything about the other.

In fact, this analogy is not a coincidence: as we said in Section 1.4 a state $|\psi\rangle$ with wave function $\psi(x)$ induces a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with probability $|\psi(x)|^2$ (with or without the differential dx, depending on the type of system we are considering, either discrete or continuous). This holds in this case as well: if $\{|x\rangle\}_x$ and $\{|y\rangle\}_y$ are bases of the two subsystems, then $\{|x\rangle \otimes |y\rangle\}_{x,y}$ is a basis of the whole system and:

$$|\psi\rangle = \iint_{\mathbb{R}^2} \psi(x,y)(|x\rangle \otimes |y\rangle) dx dy$$

implying that the induced probability space gives:

$$\mathbf{P}(x,y) = |\psi(x,y)|^2 dx dy$$

The interesting thing comes when $|\psi\rangle$ is **unentangled**. In this case, we can rewrite $|\psi\rangle$ as:

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle = \left(\int_{\mathbb{R}} \psi_A(x)|x\rangle dx\right) \otimes \left(\int_{\mathbb{R}} \psi_B(y)|y\rangle dy\right)$$

$$= \iint_{\mathbb{R}^2} \psi_A(x)\psi_B(y)|x\rangle \otimes |y\rangle dxdy$$

Thus, the induced probability space yields:

$$\mathbf{P}(x,y) = |\psi_A(x)\psi_B(y)|^2 dxdy = |\psi_A(x)|^2 dx \cdot |\psi_B(y)|^2 dy = \mathbf{P}_A(x) \cdot \mathbf{P}_B(y)$$

which is exactly the definition of independence in probability. These calculations answers our questions about what the probability of measuring an outcome is.

Our attention now goes to **what happens** when an outcome is actually measured. It will not be a surprise if we say that measuring only one of the subsystems when the global system is in an **unentangled** state, the other subsystem will not be affected. More formally, measuring only the first subsystem with an observable A is equivalent to measuring the **whole** system with the observable $A \otimes \mathbb{1}$.

Why is measuring with the identity operator equivalent to not measuring at all? Think about the post-measurement state we derived in Section 1.6. We only have one subspace, i.e. the whole space, where we end up with probability 1. In this case, the post-measurement state is:

$$\frac{\mathbb{1}|\psi\rangle}{|\mathbb{1}|\psi\rangle|} = |\psi\rangle$$

Thus, nothing changes with probability 1. Also, we retrieve no information out of such measurement, as all the eigenvalues of $\mathbb{1}$ are 1.

Measurement of a subsystem in unentangled state. If we suppose that, upon measurement of state $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ with the observable A, we end up in the subspace associated with the projection operator P, the post-measurement state of the whole system is:

$$\begin{aligned}
\frac{(P \otimes \mathbb{1})|\psi\rangle}{|(P \otimes \mathbb{1})|\psi\rangle|} &= \frac{(P \otimes \mathbb{1})|\psi\rangle}{\sqrt{\langle\psi|(P \otimes \mathbb{1})^2|\psi\rangle}} \\
&= \frac{P|\psi_A\rangle \otimes |\psi_B\rangle}{\sqrt{\langle\psi|(P \otimes \mathbb{1})^2|\psi\rangle}} \\
&= \frac{P|\psi_A\rangle \otimes |\psi_B\rangle}{\sqrt{\langle\psi|(P^2 \otimes \mathbb{1}^2)|\psi\rangle}} \\
&= \frac{P|\psi_A\rangle \otimes |\psi_B\rangle}{\sqrt{\langle\psi_A|P|\psi_A\rangle\langle\psi_B|\psi_B\rangle}} \\
&= \frac{P|\psi_A\rangle}{\sqrt{\langle\psi_A|P|\psi_A\rangle}} \otimes |\psi_B\rangle
\end{aligned}$$

Measurement of a subsystem in entangled state. For this case we will be a bit more concrete, and directly see an example. Let us take the example of the two qubits in the entangled state from Section 3.2:

$$|\psi\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}$$

We measure only the first qubit using Z (i.e. the whole system using $Z \otimes \mathbb{1}$):

$$Z = |0\rangle\langle 0| - |1\rangle\langle 1|$$

The probability of measuring $|0\rangle$ (or, more precisely, the label $+1$) is:

$$\mathbf{P}_A(0) = \langle\psi|(|0\rangle\langle 0| \otimes \mathbb{1})|\psi\rangle$$

$$
\begin{aligned}
&= \frac{\langle 00| - \langle 11|}{\sqrt{2}}(|0\rangle\langle 0| \otimes \mathbb{1})\frac{|00\rangle - |11\rangle}{\sqrt{2}} \\
&= \frac{\langle 00|(|0\rangle\langle 0| \otimes \mathbb{1}) - \langle 11|(|0\rangle\langle 0| \otimes \mathbb{1})}{\sqrt{2}} \cdot \frac{|00\rangle - |11\rangle}{\sqrt{2}} \\
&= \frac{\langle 0|0\rangle\langle 0| \otimes \langle 0| - \langle 1|0\rangle\langle 0| \otimes \langle 1|}{\sqrt{2}} \cdot \frac{|00\rangle - |11\rangle}{\sqrt{2}} \\
&= \frac{\langle 00|}{\sqrt{2}} \cdot \frac{|00\rangle - |11\rangle}{\sqrt{2}} = \frac{1}{2}
\end{aligned}
$$

by symmetry also the second qubit has probability $\frac{1}{2}$ of being measured as 0. Suppose that we measure the first qubit and it collapses to 0. The post-measurement state is:

$$
|\psi'\rangle = \frac{(|0\rangle\langle 0| \otimes \mathbb{1})|\psi\rangle}{\sqrt{\langle\psi|(|0\rangle\langle 0| \otimes \mathbb{1})|\psi\rangle}} = \sqrt{2} \cdot (|0\rangle\langle 0| \otimes \mathbb{1})\frac{|00\rangle - |11\rangle}{\sqrt{2}} = |00\rangle
$$

while the first qubit collapsed to state 0, also the second qubit will now surely be 0 when measured with respect to the computational basis. In some sense, $|\psi\rangle$ was telling us that the two qubits still act like $|-\rangle$ when seen singularly, but measuring one of them also causes a collapse in the state of the other one.

Observe that we could also perform a joint measurement of the two qubits via $Z \otimes Z$, the probabilities of the outcomes would not change (i.e. the post measurement states are $|00\rangle$ and $|11\rangle$ with probability $\frac{1}{2}$ and it thus can never be $|01\rangle$ or $|10\rangle$).

Computing the tensor product of two matrices. We end this section by looking at how the observable $Z \otimes \mathbb{1}$ actually looks like. First, notice that the projection operators are:

$$
\mathcal{M}_{AB} = \{|0\rangle\langle 0|_A \otimes \mathbb{1}_B, |1\rangle\langle 1|_A \otimes \mathbb{1}_B\}
$$

with labels $+1$ and -1, respectively. In order to see this, we again take advantage of the properties of the tensor product:

$$
\begin{aligned}
Z \otimes \mathbb{1} &= (|0\rangle\langle 0| - |1\rangle\langle 1|) \otimes (|0\rangle\langle 0| + |1\rangle\langle 1|) \\
&= |0\rangle\langle 0|_A \otimes |0\rangle\langle 0|_B + |0\rangle\langle 0|_A \otimes |1\rangle\langle 1|_B - |1\rangle\langle 1|_A \otimes |0\rangle\langle 0|_B - |1\rangle\langle 1|_A \otimes |1\rangle\langle 1|_B \\
&= |00\rangle\langle 00| + |01\rangle\langle 01| - |10\rangle\langle 10| - |11\rangle\langle 11| \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}
\end{aligned}
$$

and the matrix clearly shows its eigenspaces (since it is diagonal).

In general, by playing around with the properties of the tensor product one can derive the following general formula for the tensor product between two 2×2 matrices (which can then be generalized to higher dimensions):

$$
\begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes Q = \begin{pmatrix} a \cdot Q & b \cdot Q \\ c \cdot Q & d \cdot Q \end{pmatrix} \in \mathbb{C}^{4\times 4}
$$

3.4 The Stern-Gerlach experiment: part 1

We have an electron and we would like to measure its spin. Consider the following setup: the

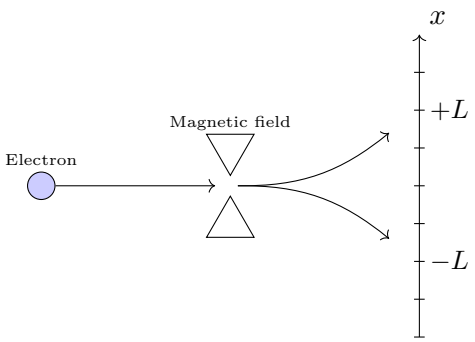


Figure 3.2: A simplified version of the Stern-Gerlach experiment to measure the spin of a particle.

spin can be schematized as a qubit, where $|0\rangle$ indicates a spin-up, and $|1\rangle$ indicates a spin-down. Moreover, we consider its position along the x-axis, as depicted in Figure 3.2. For simplicity we assume that the position is in an integral domain. We have $\mathcal{H}_A, \mathcal{H}_B$ as the Hilbert spaces of the qubit and the position of the electron respectively. Assuming that the state of the spin is $|\psi\rangle$, and the position is initially $|0\rangle$, we apply the following transformation (through e.g. a magnetic field):

$$|0\rangle\langle 0| \otimes \Delta^+ + |1\rangle\langle 1| \otimes \Delta^-$$

where $\Delta^+ = \sum_{n\in\mathbb{Z}} |n+L\rangle\langle n|$ and $\Delta^- = \sum_{n\in\mathbb{Z}} |n-L\rangle\langle n|$ shift the position of the electron respectively up and down by L. If $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, the new state of the electron is:

$$\begin{aligned}(|0\rangle\langle 0| \otimes \Delta^+ + |1\rangle\langle 1| \otimes \Delta^-)|\psi\rangle_A|0\rangle_B &= |0\rangle\langle 0|\psi\rangle \otimes \Delta^+|0\rangle_B + |1\rangle\langle 1|\psi\rangle \otimes \Delta^-|0\rangle_B \\ &= \alpha|0\rangle_A|{+}\mathrm{L}\rangle_\mathrm{B} + \beta|1\rangle_\mathrm{A}|{-}\mathrm{L}\rangle_\mathrm{B}\end{aligned}$$

In this way we created an entanglement between the spin of the electron and its position, and we can indirectly measure the spin by looking at the position (indeed, measuring the position will also change the spin, and one can formally derive this in a similar way we computed the post-measurement state in Section 3.3). We will give more physical details of this experiment later in the course.

3.5 Schrödinger's cat

In this section we analyze the famous thought experiment of Schrödinger's cat: suppose to have a cat in a closed, isolated box. In this box we also have an atom which may or may not decay. If the atom decays, a vial of poison is released on the cat, and the latter dies. Otherwise, the atom does not decay and the cat is still alive. We leave it, and after an hour we open the box to see if the cat is still alive. A scheme of the experiment (bonus: my cat Giuliano!) is depicted in Figure 3.3.

Formally, we can define the states of the atom, the vial of poison, and the cat with three qubits, where the one representing the decay of the atom is in state $|+\rangle$ ($|1\rangle$ indicates that the atom decayed). The other two qubits, respectively of the vial and the cat, are initially in state $|0\rangle$ (vial closed and cat alive). The quantum circuit for this system with three qubits looks as follows:

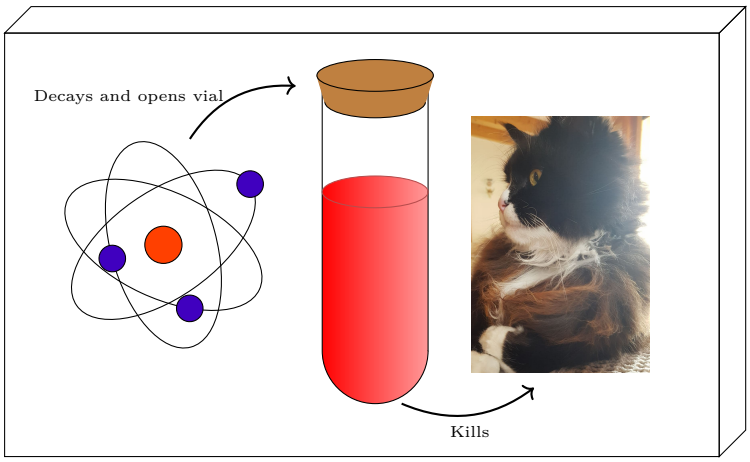
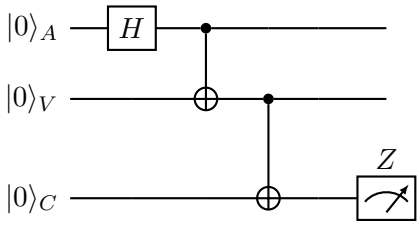


Figure 3.3: Schroedinger's thought experiment.



Thus, the state of the overall system is:

$$\begin{aligned}
\left(\mathbb{1}\otimes\bigwedge X\right)\left(\bigwedge X\otimes\mathbb{1}\right)|+\rangle|00\rangle &= \left(\mathbb{1}\otimes\bigwedge X\right)\left(\left(\bigwedge X\right)|+\rangle|0\rangle\otimes\mathbb{1}|0\rangle\right)\\
&= \left(\mathbb{1}\otimes\bigwedge X\right)\left(\left(\bigwedge X\right)\frac{|0\rangle+|1\rangle}{\sqrt{2}}|0\rangle\otimes|0\rangle\right)\\
&= \frac{1}{\sqrt{2}}\left(\mathbb{1}\otimes\bigwedge X\right)\left(\left(\left(\bigwedge X\right)|00\rangle+\left(\bigwedge X\right)|10\rangle\right)\otimes|0\rangle\right)\\
&= \frac{1}{\sqrt{2}}\left(\mathbb{1}\otimes\bigwedge X\right)\left(\left(|00\rangle+|11\rangle\right)\otimes|0\rangle\right)\\
&= \frac{1}{\sqrt{2}}\left(\mathbb{1}\otimes\bigwedge X\right)\left(|000\rangle+|110\rangle\right)\\
&= \frac{1}{\sqrt{2}}\left(\mathbb{1}\otimes\bigwedge X\right)|000\rangle+\frac{1}{\sqrt{2}}\left(\mathbb{1}\otimes\bigwedge X\right)|110\rangle\\
&= \frac{1}{\sqrt{2}}\left(\mathbb{1}|0\rangle\otimes\left(\bigwedge X\right)|00\rangle\right)+\frac{1}{\sqrt{2}}\left(\mathbb{1}|1\rangle\otimes\left(\bigwedge X\right)|10\rangle\right)\\
&= \frac{1}{\sqrt{2}}\left(|0\rangle\otimes|00\rangle\right)+\frac{1}{\sqrt{2}}\left(|1\rangle\otimes|11\rangle\right)\\
&= \frac{|000\rangle+|111\rangle}{\sqrt{2}}
\end{aligned}$$

Thus, at the end of the experiment, right before opening the box, the system will be in a superposition between two situations:

- Either the atom decayed, the poison was released, and the cat is dead or
- Nothing changed, the poison vial is still closed and the cat is still alive.

Opening the box corresponds to see whether the cat is alive or dead, i.e. using the observable

$$\mathbb{1} \otimes \mathbb{1} \otimes Z$$

which either yields $|000\rangle$ or $|111\rangle$ with equal probability. We presented this experiment not only to have another excuse to practice with qubits calculations, but also because it raises an important question in quantum theory: when does the system exactly *stop* being in a superposition, i.e. when does the measurement of the state actually occur? Trying to answer this question raises different interpretations of quantum theory, and all of them are perfectly acceptable (i.e. do not fall in contradiction with the theory).

Chapter 4
Uniform Dynamics

4.1 General setting

Consider a system with evolution Hamiltonian H. We talk about **uniform dynamics** when the Hamiltonian does not depend on time. If we consider a unitary evolution operator $U(t, t_0)$, as discussed in Section 2.2, we can find its form by solving the Schrödinger equation:

$$H|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle$$
$$H\ U(t,t_0)\ |\psi(t_0)\rangle = i\hbar\frac{\partial}{\partial t}\left(U(t,t_0)|\psi(t_0)\rangle\right).$$

This equation holds for every initial state $|\psi(t_0)\rangle$; in particular we could pick a set of initial states $\{|\psi_i(t_0)\rangle\}_i$ that form an orthonormal basis for the Hilbert space, obtaining the fact that the equation must hold at the level of the operators themselves:

$$H\ U(t,t_0) = i\hbar\frac{\partial}{\partial t}U(t,t_0).$$

This differential equation has the solution

$$U(t,t_0) = e^{-\frac{i}{\hbar}H(t-t_0)}$$

where the exponential of an operator is defined in Appendix A. This implies that the evolution along the time interval $[t_0, t]$ does not depend on the initial time t_0, i.e. it is **time-invariant**:

$$U(t,t_0) \equiv U(\Delta t)$$

Note that this does not imply that the actual evolution of a state does not depend on the *initial state*, but only that an initial state always causes the same evolution regardless of the instant the evolution takes place. The following properties can be derived from this assumption:

- $U(\Delta t)\cdot U(\Delta s) = U(\Delta s + \Delta t)$;
- $U^k(\Delta t) = U(k\Delta t)$;

4.2 Stationary states

Just to see another notation, this time we will assume a discrete Hilbert space. The above condition gives another important property for the evolution operator $U(t)$.

Theorem 4.1. *In uniform dynamics, $U(t)$ has time-invariant eigenvectors.*

Proof. Consider the Hamiltonian H, and write its spectral decomposition:

$$H = VEV^\dagger$$

Notice that V is unitary and E is real, by Hermiticity of the Hamiltonian. We rewrite the solution of the Schrödinger equation:

$$U(t) = e^{-\frac{i}{\hbar}Ht}$$

$$
\begin{aligned}
&= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar} Ht\right)^n && \text{Definition of exponential} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{it}{\hbar}\right)^n \left(VEV^\dagger\right)^n && \text{Eigendecomposition of } H \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{it}{\hbar}\right)^n VE^nV^\dagger && \text{Unitarity of } V \\
&= V \left(\sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{it}{\hbar}\right)^n E^n \right) V^\dagger \\
&= Ve^{-\frac{i}{\hbar}Et}V^\dagger
\end{aligned}
$$

Thus, the eigenvectors of H are also eigenvectors of $U(t)$, and they are independent of time. □

We also proved that, if spectral decomposition of H is:

$$H = \sum_{k\in\mathbb{Z}} E_k |k\rangle\langle k|$$

then the spectral decomposition of $U(t)$ is:

$$U = \sum_{k\in\mathbb{Z}} \exp\left(-\frac{iE_kt}{\hbar}\right) |k\rangle\langle k|$$

The eigenvectors (or eigenstates) of U and H are called **stationary states** of the evolution, in the sense that during the evolution they only change their phase, and they do so with frequency:

$$\omega_k = \frac{E_k}{\hbar}$$

In fact, consider for example a stationary $|\psi(0)\rangle = |k\rangle$ and how it evolves under these assumptions:

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle = \sum_{k'} e^{-i\omega_{k'}t}|k'\rangle\langle k'|k\rangle = e^{-i\omega_k t}|k\rangle$$

Therefore, if we take a period $t = \frac{2\pi}{\omega_k}$, we have that the system goes back to the initial state: $U(\frac{2\pi}{\omega_k})|k\rangle = |k\rangle$.

Hence, in order to know how a system evolves under uniform dynamics, it is sufficient to know the eigenbasis of the Hamiltonian $\{|k\rangle\}_k$ and its eigenvalues E_k (or equivalently, the frequencies ω_k) and then solve the **time-independent** Schrödinger equation:

$$H|k\rangle = E_k|k\rangle$$

4.3 Evolution of a time-independent observable

Consider an observable A with eigendecomposition:

$$A = \sum_{k\in\mathbb{Z}} a_k |k\rangle\langle k|$$

and assume it is time-independent, i.e. $A(t) \equiv A$. We work on the temporal derivative of the expectation of A in the state $|\psi\rangle$:

$$\frac{d}{dt}\langle A\rangle = \frac{d}{dt}\langle\psi(t)|A|\psi(t)\rangle$$

$$
\begin{aligned}
&= \langle\psi(t)|A\frac{d}{dt}\left(|\psi(t)\rangle\right) + \frac{d}{dt}\left(\langle\psi(t)|\right)A|\psi(t)\rangle \\
&= \langle\psi(t)|A\left(\frac{1}{i\hbar}H|\psi(t)\rangle\right) - \left(\langle\psi(t)|\frac{1}{i\hbar}H^\dagger\right)A|\psi(t)\rangle \qquad \text{by Schrödinger's equation} \\
&= \langle\psi(t)|A\left(\frac{1}{i\hbar}H|\psi(t)\rangle\right) - \left(\langle\psi(t)|\frac{1}{i\hbar}H\right)A|\psi(t)\rangle \qquad \text{by Hermiticity of } H \\
&= \frac{1}{i\hbar}\left(\langle\psi(t)|(AH - HA)|\psi(t)\rangle\right) \\
&=: \frac{1}{i\hbar}\left(\langle\psi(t)|[A,H]|\psi(t)\rangle\right)
\end{aligned}
$$

The minus in the third equality comes from $\left(\frac{1}{i\hbar}\right)^\dagger = -\frac{1}{i\hbar}$ (just check that this number is purely imaginary). Here we define $[H,A] = HA - AH$ as the **commutator** between H and A. Notice that if $[H,A] = 0$ then A observes a conserved quantity since:

$$
\frac{d}{dt}\langle A\rangle = \frac{1}{i\hbar}\left(\langle\psi(t)|[A,H]|\psi(t)\rangle\right) = 0
$$

and, in fact, $[H,H] = 0$ tells us that we have conservation of energy (the quantum systems we schematize with Hilbert spaces are closed by definition, we will see how to represent open systems later). In order to give more intuition about commutators we state and prove the following result, which will also be useful in the future.

Theorem 4.2. *Two operators A, B commute if and only if they can be decomposed with the same basis of eigenvectors.*

Proof. Suppose there is an eigenbasis U such that:

$$
A = U\Lambda_A U^{-1}, \quad B = U\Lambda_B U^{-1}
$$

Then we can rewrite the commutator $[A,B]$ as:

$$
\begin{aligned}
[A,B] &= U\Lambda_A U^{-1}U\Lambda_B U^{-1} - U\Lambda_B U^{-1}U\Lambda_A U^{-1} \\
&= U\Lambda_A\Lambda_B U^{-1} - U\Lambda_B\Lambda_A U^{-1} \\
&= U[\Lambda_A,\Lambda_B]U^{-1} = 0
\end{aligned}
$$

since two diagonal operators always commute.

To prove the other implication, we assume that each eigenvalue of A generate an eigenspace of dimension 1 (the general case needs some more effort, but the idea is the same). If $[A,B] = 0$, consider an eigenvector $|\psi_a\rangle$ of A associated with the eigenvalue a. We have that:

$$
A(B|\psi_a\rangle) = B(A|\psi_a\rangle) = aB|\psi_a\rangle
$$

i.e. $|\psi_a\rangle$ and $B|\psi_a\rangle$ are equal up to a constant factor (since they are both eigenvectors of the same eigenvalue), if we let this constant factor be b we say:

$$
B|\psi_a\rangle = b|\psi_a\rangle
$$

i.e. $|\psi_a\rangle$ is also an eigenvector of B. Hence, every eigenvector of A is also an eigenvector of B and vice versa (with a symmetric argument). □

Exponential of a matrix. Here we make a little note about the exponential of a matrix. Notice that the property of exponentials:

$$e^{A+B} = e^A e^B$$

holds if A and B commute, i.e. $[A, B] = 0$, but it is not true in general unlike the exponentials of scalars, where the multiplication is commutative.

Theorem 4.3. *If* $[A, B] = 0$ *then* $e^{A+B} = e^A e^B$.

Proof. Since $AB = BA$, we can use the binomial theorem on $(A + B)^n$:

$$\begin{aligned}
e^{A+B} &= \sum_{n=0}^{\infty} \frac{1}{n!}(A+B)^n && \text{definition of exponential} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{m=0}^{n} \binom{n}{m} A^m B^{n-m} && \text{by binomial theorem} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{m=0}^{n} \frac{n!}{m!(n-m)!} A^m B^{n-m} \\
&= \sum_{n=0}^{\infty} \sum_{m=0}^{n} \frac{A^m}{m!} \frac{B^{n-m}}{(n-m)!} \\
&= \left(\sum_{n=0}^{\infty} \frac{A^n}{n!} \right) \left(\sum_{m=0}^{\infty} \frac{B^m}{m!} \right) = e^A e^B && \text{by Cauchy formula}
\end{aligned}$$

□

The Heisenberg picture. One may ask what changes when the observable A is not time-independent, and what the physical meaning of an observable that changes over time is, when the whole information about the evolution is already enclosed in the state $|\psi\rangle$. The first question is easily addressed: we would get a new additive term of the form

$$\langle \psi(t)| \left(\frac{d}{dt} A \right) |\psi(t)\rangle$$

For the second question we can say that, in general, one may always assume that the observable is time-independent, and any change on observed quantity can be captured by the state. This is called the **Schrödinger picture** of an evolution. A different but equivalent view of the formalism is called the **Heisenberg picture**, in which we aggregate the evolution operator to the observable:

$$\langle \psi(t)|A|\psi(t)\rangle = (\langle \psi|U^\dagger(t))A(U(t)|\psi\rangle) = \langle \psi|(U^\dagger(t)AU(t))|\psi\rangle =: \langle \psi|A(t)|\psi\rangle$$

4.4 Qubits: implementing a rotation on the Bloch sphere

Consider a matrix $R_Z(\theta)$ representing a rotation on the Bloch sphere around the z-axis (the one containing $|0\rangle$ and $|1\rangle$ in Figure 4.1).

Such a rotation will have a "uniform" property (not with respect to time, but with respect to the angle θ):

$$R_Z(\theta_1 + \theta_2) = R_Z(\theta_1) R_Z(\theta_2)$$

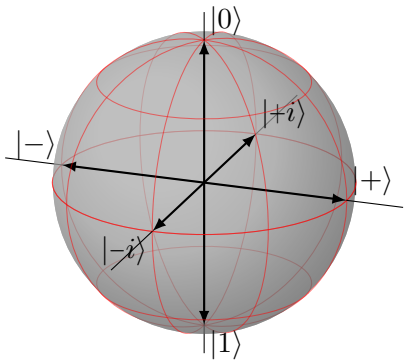


Figure 4.1: Rotations of the Hadamard basis around the Z axis on the Bloch sphere.

We would like to compute such a matrix (in order to, say, express it as a function of the Hamiltonian and implement it in some way). We start by looking for the **stationary states**. Notice that the only states which are not changed by R_Z are the ones on the rotation axis, i.e. $|0\rangle, |1\rangle$:

$$\begin{cases} R_Z(\theta)|0\rangle = e^{i\omega_0\theta}|0\rangle \\ R_Z(\theta)|1\rangle = e^{i\omega_1\theta}|1\rangle \end{cases}$$

In order to find ω_0, ω_1 we do the following calculation:

$$\begin{aligned} R_Z\left(\frac{\pi}{2}\right)|+\rangle &= R_Z\left(\frac{\pi}{2}\right) \cdot \frac{|0\rangle + |1\rangle}{\sqrt{2}} \\ &= \frac{1}{\sqrt{2}}\left(R_Z\left(\frac{\pi}{2}\right)|0\rangle + R_Z\left(\frac{\pi}{2}\right)|1\rangle\right) \\ &= \frac{1}{\sqrt{2}}\left(e^{i\omega_0\frac{\pi}{2}}|0\rangle + e^{i\omega_1\frac{\pi}{2}}|1\rangle\right) \\ &\stackrel{!}{=} e^{i\phi}\frac{|0\rangle + i|1\rangle}{\sqrt{2}} \end{aligned}$$

We would like the last equality to hold (since that would be the result after a rotation of $\pi/2$ around the z-axis, see the Figure 4.1). The term $e^{i\phi}$ is there to take into account a possible global phase (which we know is irrelevant). Possible values are $\omega_0 = -\frac{1}{2}, \omega_1 = \frac{1}{2}$, which yield a global phase $\phi = \frac{\pi}{4}$ (check it!). Since we now know the stationary states of $R_Z(\theta)$ and the relative eigenvalues, we can write a general expression for it:

$$R_Z(\theta) = e^{-i\theta/2}|0\rangle\langle 0| + e^{i\theta/2}|1\rangle\langle 1| = e^{-\frac{i\theta}{2}Z}$$

Note that this is not the only way to define $R_Z(\theta)$ (the two frequencies of the stationary states are not unique, and they may just yield a different global phase along the rotation), but choosing this representation makes Z show up, which is particularly neat. Moreover, one can check that:

$$R_Z(2\pi) = e^{-i\pi}|0\rangle\langle 0| + e^{i\pi}|1\rangle\langle 1| = -(|0\rangle\langle 0| + |1\rangle\langle 1|) = -\mathbb{1}$$

You can imagine that, if we wanted to implement this rotation, we would need $U(t) = R_Z(\omega t)$ for some frequency ω, and the Hamiltonian of our evolution should be of the form cZ for some constant c. Indeed:

$$\frac{d}{dt}|\psi(t)\rangle = \frac{d}{dt}(R_Z(\omega t))|\psi(0)\rangle$$

$$= \left(-\frac{i\omega}{2} e^{-i\omega t/2} |0\rangle\langle 0| + \frac{i\omega}{2} e^{i\omega t/2} |1\rangle\langle 1| \right) |\psi(0)\rangle$$

If we multiply both sides by $i\hbar$ we obtain:

$$\begin{aligned} i\hbar \frac{d}{dt} |\psi(t)\rangle &= i\hbar \left(-\frac{i\omega}{2} e^{-i\omega t/2} |0\rangle\langle 0| + \frac{i\omega}{2} e^{i\omega t/2} |1\rangle\langle 1| \right) |\psi(0)\rangle \\ &= \frac{\hbar\omega}{2} \left(e^{-i\omega t/2} |0\rangle\langle 0| - e^{i\omega t/2} |1\rangle\langle 1| \right) |\psi(0)\rangle \end{aligned}$$

and the derivative on left-hand side is clearly $H|\psi(t)\rangle = HR_Z(\omega t)|\psi(0)\rangle$, by Schrödinger equation. We guess that $H = \frac{\hbar\omega}{2} Z$ and, in fact, we obtain that:

$$\begin{aligned} HR_Z(\omega t) &= \frac{\hbar\omega}{2} \left(|0\rangle\langle 0| - |1\rangle\langle 1| \right) \left(e^{-i\omega t/2} |0\rangle\langle 0| + e^{i\omega t/2} |1\rangle\langle 1| \right) \\ &= \frac{\hbar\omega}{2} \left(e^{-i\omega t/2} |0\rangle\langle 0| - e^{i\omega t/2} |1\rangle\langle 1| \right) \end{aligned}$$

which is exactly the expression above. Notice that we may similarly construct $R_X(\omega t)$ for a rotation around the x-axis (the one containing $|+\rangle$ and $|-\rangle$).

4.5 Symmetries

We now talk about a notion that will become useful later, e.g. in Chapter 10 when introducing potential.

Definition 4.4. *A unitary matrix V is a symmetry if and only if*

$$VU(t)|\psi\rangle = U(t)V|\psi\rangle$$

for any state $|\psi\rangle$, and instant t.

This definition tells us that a transformation by a symmetry gives the same effect if applied before or after a certain evolution. In particular this means that V commutes with $U(t)$:

$$[V, U(t)] = 0 \implies [V, H] = 0$$

We can say that an observable A observes a conserved quantity if and only if e^{iA} is a symmetry.

Chapter 5
Position and Momentum

5.1 The momentum basis

Suppose to have a continuous Hilbert space $\mathcal{H}$ spanned by a basis $\{|x\rangle\}_x$, which you can imagine to be the position basis of a particle along an axis. A state $|\psi\rangle$ of a particle can be expressed as:

$$|\psi\rangle = \mathbb{1}|\psi\rangle = \left(\int_{\mathbb{R}} |x\rangle\langle x| dx\right)|\psi\rangle = \int_{\mathbb{R}} |x\rangle\langle x|\psi\rangle dx = \int_{\mathbb{R}} \psi(x)|x\rangle dx$$

Now assume to have another basis $\{|p\rangle\}_p$ of $\mathcal{H}$. Following the same argument, we can express $|\psi\rangle$ also as:

$$|\psi\rangle = \int_{\mathbb{R}} \bar{\psi}(p)|p\rangle dx$$

For some wave function $\bar{\psi}(p) = \langle p|\psi\rangle$, which will be different from $\psi(x)$ in general. We now choose a very particular basis here, i.e. the one that satisfies[2]:

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \cdot e^{ipx/\hbar} \implies |p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} e^{ipx/\hbar}|x\rangle dx$$

Notice also that $\langle p|x\rangle = (\langle x|p\rangle)^\dagger = \frac{1}{\sqrt{2\pi\hbar}} \cdot e^{-ipx/\hbar}$, and we can also write $|x\rangle$ in terms of the basis $|p\rangle$ in a similar way:

$$|x\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} e^{-ipx/\hbar}|p\rangle dp$$

Theorem 5.1. *$\{|p\rangle\}_p$ is an orthonormal basis for $\mathcal{H}$.*

Proof. We need to show that, given $\langle x|x'\rangle = \delta(x - x')$, we have $\langle p|p'\rangle = \delta(p - p')$. In fact:

$$\begin{aligned}
\langle p|p'\rangle &= \left(\frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} e^{ipx/\hbar}|x\rangle dx\right)^\dagger \left(\frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} e^{ip'x/\hbar}|x\rangle dx\right) \\
&= \frac{1}{2\pi\hbar} \left(\int_{\mathbb{R}} e^{-ipx/\hbar}\langle x| dx\right)\left(\int_{\mathbb{R}} e^{ip'x/\hbar}|x\rangle dx\right) \\
&= \frac{1}{2\pi\hbar} \int_{\mathbb{R}}\int_{\mathbb{R}} e^{-i(p'x'-px)/\hbar}\langle x|x'\rangle dx dx' \\
&= \frac{1}{2\pi\hbar} \int_{\mathbb{R}} e^{-ix(p'-p)/\hbar} dx = \delta(p - p')
\end{aligned}$$

where the last equality derives from the properties of the Dirac delta (Corollary A.9). □

If we define the basis $\{|p\rangle\}_p$ in this way, what happens to $\bar{\psi}(p)$? We find that:

$$\begin{aligned}
\bar{\psi}(p) &= \langle p|\psi\rangle \\
&= \langle p|\left(\int_{\mathbb{R}} |x\rangle\langle x| dx\right)|\psi\rangle \\
&= \int_{\mathbb{R}} \langle p|x\rangle\langle x|\psi\rangle dx
\end{aligned}$$

[2] It is also possible to find the definition of $\langle x|p\rangle$ with a minus sign in the exponential. The equation for $\langle p|x\rangle$ would then change accordingly. In the following we used a definition consistent with Schumacher and Westmoreland [1].

$$= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} e^{-ipx/\hbar} \psi(x) dx$$

which is exactly the **fourier transform** of $\psi(x)$. With the exact same argument we find that:

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} e^{ipx/\hbar} \bar{\psi}(p) dp$$

i.e. the **inverse fourier transform** of $\bar{\psi}(p)$. Later in this chapter we will see how in fact this definition of momentum influences the speed, i.e. the derivative of $\langle X \rangle$.

5.2 The momentum operator

We already seen the position operator, i.e. its observable:

$$X = \int_{\mathbb{R}} x|x\rangle\langle x| dx$$

In a similar way, we can define the momentum operator:

$$P = \int_{\mathbb{R}} p|p\rangle\langle p| dp$$

Let us see what happens to the wave function $\psi(x)$ when we apply the momentum operator:

$$\begin{aligned}
\langle x|P|\psi\rangle &= \langle x| \left(\int_{\mathbb{R}} p|p\rangle\langle p| dp \right) |\psi\rangle \\
&= \int_{\mathbb{R}} p\langle x|p\rangle\langle p|\psi\rangle dp \\
&= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} p \cdot e^{ipx/\hbar} \cdot \bar{\psi}(p) dp \\
&= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \left(\frac{\hbar}{i} \frac{\partial}{\partial x} e^{ipx/\hbar} \right) \cdot \bar{\psi}(p) dp \\
&= -i\hbar \frac{\partial}{\partial x} \left(\frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} e^{ipx/\hbar} \cdot \bar{\psi}(p) dp \right) \\
&= -i\hbar \frac{\partial}{\partial x} \psi(x)
\end{aligned}$$

Therefore, we usually write that P is an operator such that:

$$P : \psi(x) \to -i\hbar \frac{\partial}{\partial x} \psi(x)$$

Moreover, a way to write the expected momentum, which will be useful in certain cases, is:

$$\begin{aligned}
\langle P \rangle &= \langle \psi|P|\psi\rangle \\
&= \langle \psi| \left(\int_{\mathbb{R}} |x\rangle\langle x| dx \right) P|\psi\rangle \\
&= \langle \psi| \left(\int_{\mathbb{R}} |x\rangle\langle x| dx \right) P|\psi\rangle \\
&= \int_{\mathbb{R}} \langle \psi|x\rangle\langle x|P|\psi\rangle dx \\
&= -i\hbar \int_{\mathbb{R}} \psi^*(x) \frac{\partial}{\partial x} \psi(x) dx
\end{aligned}$$

5.3 The position-momentum commutator

In this section we will compute $[X, P]$ and discuss the implications of the result. Let us start by computing the wave function of the state $XP|\psi\rangle$:

$$\begin{aligned}\langle x|XP|\psi\rangle &= (\langle x|X)\, P|\psi\rangle \\ &= (x\langle x|)\, P|\psi\rangle && \text{since } |x\rangle \text{ is an eigenstate of } X \\ &= -i\hbar x\frac{\partial}{\partial x}\psi(x)\end{aligned}$$

The computation of the other part of the commutator $PX|\psi\rangle$ is a bit less straightforward:

$$\begin{aligned}\langle x|PX|\psi\rangle &= \langle x|P\,(X|\psi\rangle) \\ &= \langle x|P\left(\int_{\mathbb{R}} x|x\rangle\langle x|\psi\rangle dx\right) \\ &= \langle x|P\left(\int_{\mathbb{R}} x\cdot\psi(x)|x\rangle dx\right) \\ &= \langle x|P|\psi'\rangle\end{aligned}$$

here we are abusing the notation a little bit: notice that $|\psi'\rangle$ is **not** a valid state, since $x\psi(x)$ is not normalized in general. However, it is useful to think of it as a state because now we know immediately that, by the properties of P:

$$\begin{aligned}\langle x|PX|\psi\rangle &= \langle x|P|\psi'\rangle \\ &= -i\hbar\frac{\partial}{\partial x}\,(x\cdot\psi(x)) \\ &= -i\hbar x\frac{\partial}{\partial x}\psi(x) - i\hbar\psi(x)\end{aligned}$$

Putting the two results together we get:

$$\begin{aligned}\langle x|\,(XP-PX)\,|\psi\rangle &= -i\hbar x\frac{\partial}{\partial x}\psi(x) + i\hbar x\frac{\partial}{\partial x}\psi(x) + i\hbar\psi(x) \\ &= i\hbar\cdot\psi(x)\end{aligned}$$

We just found that $[X, P]$ acts as:

$$[X, P] : \psi(x) \to i\hbar\cdot\psi(x)$$

Therefore $[X, P]|\psi\rangle = i\hbar|\psi\rangle$ for every state $|\psi\rangle$, implying:

$$[X, P] = i\hbar\mathbb{1}$$

This is called **canonical commutation relation** between position and momentum. What does this mean? Recall from Section 4.3 that the derivative of the expected position $\langle X\rangle$ of a state changes according to the commutator $[X, H]$, where H is the Hamiltonian of the evolution. We will see later in this chapter that the position operator P and the Hamiltonian are closely related to each other, in a way that nicely resembles the total mechanical energy of a classical system.

5.4 Transformation of observables in the commutator

We proved that $[X, P] = i\hbar\mathbb{1}$, but what if we wanted to compute $[f(X), P]$, for some function f? We will take advantage of the following simple result:

Theorem 5.2. *The commutator $[A, B]$ between two operators A, B is bilinear.*

Proof. If we consider a linear combination on the left operator:

$$\begin{aligned}[a_1A_1 + a_2A_2, B] &= (a_1A_1 + a_2A_2)B - B(a_1A_1 + a_2A_2) \\ &= a_1A_1B + a_2A_2B - a_1BA_1 - a_2BA_2 \\ &= a_1(A_1B - BA_1) + a_2(A_2B - BA_2) \\ &= a_1[A_1, B] + a_2[A_2, B]\end{aligned}$$

The argument for a linear combination on the second operator is the same. □

Now take the Taylor series of f:

$$f(X) = \sum_{n=0}^{\infty} a_n X^n$$

By Theorem 5.2, we can rewrite $[f(X), P]$ as:

$$[f(X), P] = \left[\sum_{n=0}^{\infty} a_n X^n, P\right] = \sum_{n=0}^{\infty} a_n [X^n, P]$$

Thus all we're left to do is compute $[X^n, P]$ for every n:

Theorem 5.3. $[X^n, P] = i\hbar \cdot nX^{n-1}$

Proof. We prove this by induction: for $n = 1$ we already know that $[X, P] = i\hbar\mathbb{1} = i\hbar \cdot X^0$. If we assume the claim to be true for $n - 1$ we have:

$$\begin{aligned}[X^n, P] &= X^nP - PX^n \\ &= X^{n-1}(XP) - PX^n \\ &= X^{n-1}(PX + i\hbar\mathbb{1}) - PX^n && \text{since } XP = PX + [X, P] \\ &= i\hbar X^{n-1} + X^{n-1}PX - PX^n \\ &= i\hbar X^{n-1} + [X^{n-1}, P]X \\ &= i\hbar X^{n-1} + i\hbar \cdot (n-1)X^{n-2}X && \text{by induction} \\ &= i\hbar \cdot nX^{n-1}\end{aligned}$$

□

For a bit of intuition about this, you may notice that the expression of $[X^n, P]$ gives $i\hbar$ times the "derivative of X^n with respect to X", which aligns with the fact that P maps $\psi(x)$ to $i\hbar \cdot \frac{\partial}{\partial x}\psi(x)$.

Now we are ready to plug this expression into our Taylor series.

$$[f(X), P] = \sum_{n=0}^{\infty} a_n [X^n, P]$$

$$
\begin{aligned}
&= i\hbar \cdot \sum_{n=0}^{\infty} a_n n X^{n-1} \\
&= i\hbar \cdot \sum_{n=0}^{\infty} a_n \frac{\partial}{\partial X}(X^n) \\
&= i\hbar \cdot \frac{\partial}{\partial X}\left(\sum_{n=0}^{\infty} a_n X^n\right) \\
&= i\hbar \frac{\partial}{\partial X} f(X)
\end{aligned}
$$

With an argument symmetric to the one we presented, it is also possible to prove the following:

$$[X, f(P)] = i\hbar \frac{\partial}{\partial P} f(P)$$

5.5 Hamiltonian of a particle in one dimension

Consider a particle moving along the x-axis. The general Hamiltonian for a particle of (real and constant) mass μ is:

$$H = \frac{P^2}{2\mu} + V(X)$$

Recall that the Hamiltonian is the observable for the total energy of the system. Here we are decomposing the total energy in a classical way:

- The first term indicates the **kinetic energy**, where P is the momentum operator, i.e. the observable of the momentum of the particle. This term should not be a surprise: in classical physics, a body of mass m and speed v has a total kinetic energy of:

$$\mathcal{E}_K = \frac{1}{2}mv^2 = \frac{1}{2m}(mv)^2 = \frac{p^2}{2m}$$

- The second term is a symmetry called **potential energy**. As you may recall from classical physics courses, the potential energy is a quantity of energy due to the position of the body in a space containing a (conservative) **force field**. Classical examples are gravitational and electrical fields.

5.6 Evolution of an observable: Ehrenfest's theorem

For an observable A we already saw that, in order to analyze the derivative of $\langle A \rangle$, we need $[A, H]$, where H is the Hamiltonian of the observed system. Therefore:

$$[A, H] = \frac{1}{2\mu}\left[A, P^2\right] + [A, V(X)]$$

which is extremely neat, since usually the observables we work with either depend on X or P (and thus one of the two terms vanishes). Let us see some examples, starting from $[X, H]$:

$$[X, H] = \frac{1}{2\mu}\left[X, P^2\right] + [X, V(X)]$$

Notice that $[X, V(X)] = 0$ since $V(X)$ can be expressed as a Taylor series involving $[X, X^n]$, which are all zero (why?). Thus, we are only left with:

$$[X, H] = \frac{1}{2\mu}[X, P^2] = i\hbar\frac{1}{2\mu}2P = i\hbar\frac{P}{\mu}$$

This gives a relation between the momentum of the particle as we defined it, and its **speed**, finally aligning our definition from the classical meaning of linear momentum.

Theorem 5.4 (Ehrenfest's Theorem I)**.** *The derivative of the expected position of a particle is its expected speed, i.e. the expected momentum over its mass:*

$$\frac{\partial}{\partial t}\langle X\rangle = \frac{1}{\mu}\langle P\rangle$$

Proof.

$$\frac{\partial}{\partial t}\langle X\rangle = \frac{1}{i\hbar}\langle\psi|[X, H]|\psi\rangle = \frac{1}{i\hbar}\langle\psi|\left(i\hbar\cdot\frac{P}{\mu}\right)|\psi\rangle = \frac{1}{\mu}\langle\psi|P|\psi\rangle = \frac{1}{\mu}\langle P\rangle$$

□

Let us now work with $[P, H]$:

$$[P, H] = \frac{1}{2\mu}[P, P^2] + [P, V(X)] = [P, V(X)] = i\hbar\frac{\partial}{\partial X}V(X)$$

this leads to a result which closely resembles the second principle of classical dynamics:

Theorem 5.5 (Ehrenfest's Theorem II)**.** *The derivative of the expected momentum of a particle is the total external force acting on it, i.e. the spatial derivative of the potential.*

$$\frac{\partial}{\partial t}\langle P\rangle = \langle\frac{\partial}{\partial X}V(X)\rangle$$

Proof.

$$\frac{\partial}{\partial t}\langle P\rangle = \frac{1}{i\hbar}\langle\psi|[P, H]|\psi\rangle = \frac{1}{i\hbar}\langle\psi|\left(i\hbar\cdot\frac{\partial}{\partial X}V(X)\right)|\psi\rangle = \langle\frac{\partial}{\partial X}V(X)\rangle$$

□

These theorems more closely resemble their classical counterpart when we are in 2 or 3 dimensions and these actually are written in **vector** form. We will see how to generalize our formalization to more dimensions later.

5.7 Schrödinger equation for a free particle

We have a **free** particle when no external force is applied, i.e. $V(x) = 0$ and $H = \frac{P^2}{2\mu}$. With a similar argument as the one for P given in Section 5.2, P^2 acts on a state $|\psi\rangle$ in the following way:

$$P^2 : \psi(x) \to (-i\hbar)^2\frac{\partial^2}{\partial x^2}\psi(x) = -\hbar^2\frac{\partial^2}{\partial x^2}\psi(x)$$

Therefore we can multiply the Schrödinger equation we know with $\langle x|$ in order to derive the one for the wave function:

$$\frac{P^2}{2\mu}|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle$$

$$\frac{1}{2\mu}\langle x|P^2|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}\langle x|\psi(t)\rangle$$
$$-\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial x^2}\psi(x,t) = i\hbar\frac{\partial}{\partial t}\psi(x,t)$$

On the other hand, in order to solve the Schrödinger equation in this case, it may be useful to work in the momentum basis, since only the momentum operator appears. Hence let us try to multiply by $\langle p|$:

$$\frac{P^2}{2\mu}|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle$$
$$\frac{1}{2\mu}\langle p|P^2|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}\langle p|\psi(t)\rangle$$
$$\frac{p^2}{2\mu}\bar{\psi}(p) = i\hbar\frac{\partial}{\partial t}\bar{\psi}(p)$$

and this is a simple homogeneous differential equation with solution:

$$\bar{\psi}(p,t) = \bar{\psi}(p,0)\cdot e^{-\frac{ip^2t}{2\mu\hbar}}$$

Finally, in order to find $\psi(x,t)$ we can simply apply an inverse Fourier transform to this solution. In general, it is a good idea to:

- Start with the initial state $\psi(x,0)$;
- Transform $\psi(x,0) \to \bar{\psi}(p,0)$;
- Evolve the state in the momentum basis;
- Apply the inverse Fourier transform to find the evolution in the position basis.

We will see a practical example of this in Section 9.3 when discussing the double slit experiment.

Stationary states. Recall that a stationary state is an eigenstate of the Hamiltonian and only their global phase changes over time. We can derive the same conclusion for the wave function of stationary states:

$$|\psi(t)\rangle = e^{-iE_kt/\hbar}|\psi(0)\rangle$$
$$\langle x|\psi(t)\rangle = e^{-iE_kt/\hbar}\langle x|\psi(0)\rangle$$
$$\psi(x,t) = e^{-iE_kt/\hbar}\psi(x,0)$$

Finding the stationary states of a Hamiltonian is a very important problem in many branches of science, from chemistry to pharmacy, and one of the main promises of quantum computation is to be able to compute them in feasible time.

Chapter 6
Dynamics in 3D

6.1 Position and momentum in space

In this chapter we will see how to extend what we saw for continuous systems involving the position of a particle to the general setting of a three-dimensional space. Just like when we have to consider multiple qubits together, we can express the Hilbert space of a three dimensional system as **tensor product** of the spaces corresponding to the three axes of the reference frame we want to consider:

$$\mathcal{H} = \mathcal{H}_x \otimes \mathcal{H}_y \otimes \mathcal{H}_z = \text{span}\{|x\rangle \otimes |y\rangle \otimes |z\rangle, x, y, z \in \mathbb{R}\} = \text{span}\{|\mathbf{r}\rangle, \mathbf{r} \in \mathbb{R}^3\}$$

where $\mathbf{r}$ denotes the position vector. Thus, the state of a particle $|\psi\rangle$ can be expressed as:

$$\begin{aligned}
|\psi\rangle &= (\mathbb{1}_x \otimes \mathbb{1}_y \otimes \mathbb{1}_z)\,|\psi\rangle \\
&= \left(\left(\int_{\mathbb{R}} |x\rangle\langle x| dx\right) \otimes \left(\int_{\mathbb{R}} |y\rangle\langle y| dy\right) \otimes \left(\int_{\mathbb{R}} |z\rangle\langle z| dz\right)\right)|\psi\rangle \\
&= \left(\iiint_{\mathbb{R}^3} |x\rangle\langle x| \otimes |y\rangle\langle y| \otimes |z\rangle\langle z| dxdydz\right)|\psi\rangle \\
&= \left(\iiint_{\mathbb{R}^3} |x\rangle|y\rangle|z\rangle\langle x|\langle y|\langle z| dxdydz\right)|\psi\rangle \\
&= \iiint_{\mathbb{R}^3} |\mathbf{r}\rangle\langle\mathbf{r}|\psi\rangle dxdydz \\
&= \iiint_{\mathbb{R}^3} \psi(\mathbf{r})|\mathbf{r}\rangle dxdydz = \iiint_{\mathbb{R}^3} \psi(\mathbf{r})|\mathbf{r}\rangle d^3\mathbf{r}
\end{aligned}$$

The momentum basis is, by the properties of the tensor product, immediately derived using three independent Fourier transforms:

$$\langle x|p_x\rangle = \frac{1}{\sqrt{2\pi\hbar}} \cdot e^{ip_x x/\hbar}, \quad \langle y|p_y\rangle = \frac{1}{\sqrt{2\pi\hbar}} \cdot e^{ip_y y/\hbar}, \quad \langle z|p_z\rangle = \frac{1}{\sqrt{2\pi\hbar}} \cdot e^{ip_z z/\hbar}$$

Hence also the following is a basis for our space:

$$\mathcal{H} = \text{span}\{|p_x\rangle \otimes |p_y\rangle \otimes |p_z\rangle, p_x, p_y, p_z \in \mathbb{R}\} = \text{span}\{|\mathbf{p}\rangle, \mathbf{p} \in \mathbb{R}^3\}$$

We can directly express $|\mathbf{p}\rangle$ in terms of $|\mathbf{r}\rangle$ in a very elegant way:

$$\begin{aligned}
\langle\mathbf{r}|\mathbf{p}\rangle &= (\langle x|\langle y|\langle z|)(|p_x\rangle|p_y\rangle|p_z\rangle) \\
&= \langle x|p_x\rangle\langle y|p_y\rangle\langle z|p_z\rangle \\
&= \frac{1}{\sqrt{2\pi\hbar}} \cdot e^{ip_x x/\hbar} \cdot \frac{1}{\sqrt{2\pi\hbar}} \cdot e^{ip_y y/\hbar} \cdot \frac{1}{\sqrt{2\pi\hbar}} \cdot e^{ip_z z/\hbar} \\
&= \frac{1}{(2\pi\hbar)^{3/2}} \cdot e^{i(p_x x + p_y y + p_z z)/\hbar} \\
&= \frac{1}{(2\pi\hbar)^{3/2}} \cdot e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}
\end{aligned}$$

And this gives us a natural extension of what we already know:

$$\bar{\psi}(\mathbf{p}) = \frac{1}{(2\pi\hbar)^{3/2}} \iiint_{\mathbb{R}^3} \psi(\mathbf{r}) \cdot e^{i\mathbf{p}\cdot\mathbf{r}/\hbar} dxdydz$$

6.2 Momentum operator in 3D space

Using a simple tensor product, in a similar way we did when we tried to measure only one out of two qubits, we can define an observable to each component. For example, we can define the observables for the components of the momentum as follows:

$$P_x = \left(\int_{\mathbb{R}} p_x |p_x\rangle\langle p_x| dp_x \right) \otimes \mathbb{1}_y \otimes \mathbb{1}_z$$
$$P_y = \mathbb{1}_x \otimes \left(\int_{\mathbb{R}} p_y |p_y\rangle\langle p_y| dp_y \right) \otimes \mathbb{1}_z$$
$$P_z = \mathbb{1}_x \otimes \mathbb{1}_y \otimes \left(\int_{\mathbb{R}} p_z |p_z\rangle\langle p_z| dp_z \right)$$

They act on the wave function in the way we have already seen (the calculations to derive this do not change at all, we simply keep the other two variables fixed):

$$P_x : \psi(\mathbf{r}) \to -i\hbar \frac{\partial}{\partial x} \psi(\mathbf{r})$$
$$P_y : \psi(\mathbf{r}) \to -i\hbar \frac{\partial}{\partial y} \psi(\mathbf{r})$$
$$P_z : \psi(\mathbf{r}) \to -i\hbar \frac{\partial}{\partial z} \psi(\mathbf{r})$$

Since we will use it to observe the kinetic energy of a particle, here we also define the squared momentum operator P^2:

$$P^2 = P_x^2 + P_y^2 + P_z^2 = \iiint_{\mathbb{R}^3} \left(p_x^2 + p_y^2 + p_z^2 \right) |\mathbf{p}\rangle\langle\mathbf{p}| d^3\mathbf{p}$$

And, by exploiting linearity, P^2 acts on the wave function analogously as what we have already seen in the one dimensional case:

$$P^2 : \psi(\mathbf{r}) \to -\hbar^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \psi(\mathbf{r}) = -\hbar^2 \nabla^2 \psi(\mathbf{r})$$

where ∇^2 is the *Laplace operator*.

6.3 Schrödinger equation for a free particle in space

The Schrödinger equation remains unchanged in this extension to multiple dimensions:

$$H|\psi\rangle = i\hbar \frac{\partial}{\partial t} |\psi\rangle$$

and, for a free particle, we again know that the Hamiltonian H is:

$$H = \frac{P^2}{2\mu}$$

Thus, if ψ is the wave function with respect to the position basis:

$$\frac{P^2}{2\mu}|\psi\rangle = i\hbar \frac{\partial}{\partial t} |\psi\rangle$$
$$\langle x| \frac{P^2}{2\mu} |\psi\rangle = i\hbar \frac{\partial}{\partial t} \langle x|\psi\rangle$$
$$\frac{-\hbar^2}{2\mu} \nabla^2 \psi = i\hbar \frac{\partial}{\partial t} \psi$$

Chapter 7
Variance and Wave Packets

7.1 Motion of a free particle

Classical physics tells us that, whenever the net force acting on an object is zero, this object has constant velocity, i.e. it is either at rest or it moves at constant speed towards a constant direction. Ehrenfest's theorems tell us that this classical behaviour is followed (in expectation) also in quantum systems. In fact,

$$\frac{\partial}{\partial t}\langle P\rangle = \langle\frac{\partial}{\partial X}V(X)\rangle = 0 \Longrightarrow \langle P\rangle_t = \langle P\rangle_0$$

which means that the expected momentum is **conserved**. Moreover, as outlined in Figure 7.1,

$$\frac{\partial}{\partial t}\langle X\rangle = \frac{\langle P\rangle}{\mu} \Longrightarrow \langle X\rangle_t = \langle X\rangle_0 + \frac{\langle P\rangle}{\mu}t$$

and this extends to each component of a possible two- or three-dimensional space.

What we need now is to quantify **uncertainty** and, to do so, we use a well-known quantity: the **variance**. We denote the variance of an observable A as ΔA^2. From probability theory (a short recap is given in Appendix C) we know that the variance can be rewritten as:

$$\Delta A^2 = \langle A^2\rangle - \langle A\rangle^2$$

Example with the position operator. For a Gaussian wave, we can visualize the variance ΔX^2 as the "spread" of the probability density function (see Figure 7.2). Later in the chapter we will formally see how we can use these Gaussian wave functions to model the behaviour of the considered systems.

7.2 Variance of the momentum

In order to better understand the motion of a free particle, we take a look at the variance of its momentum over time:

$$\left(\Delta P^2\right)_t = \langle P^2\rangle_t - \langle P\rangle_t^2$$

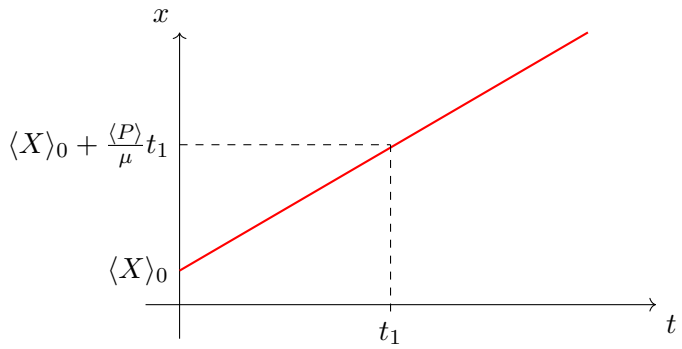


Figure 7.1: When momentum is conserved, the average position changes linearly over time.

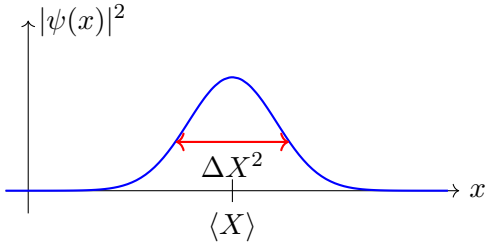


Figure 7.2: Approximate representation of the variance of a 1-dimensional Gaussian wave.

The second term is clearly constant, by Ehrenfest's Theorem II. What is interesting is the fact that also the first term is constant:

$$\frac{\partial}{\partial t}\langle P^2 \rangle = \frac{1}{i\hbar}\langle [P^2, H] \rangle = 0$$

since $[P^2, H] = 0$, as $H = P^2/2\mu$ for a free particle. We just found that the momentum of a free particle not only preserves its expectation, but also its variance, i.e. we do not accumulate uncertainty about the momentum of a free particle over time.

7.3 Variance of the position

Now we look at the variance ΔX^2 of the position:

$$\begin{aligned}
\left(\Delta X^2\right)_t &= \langle X^2 \rangle_t - \langle X \rangle_t^2 \\
&= \langle X^2 \rangle_t - \left(\langle X \rangle_0 + \frac{t}{\mu}\langle P \rangle\right)^2 && \text{by Ehrenfest's Theorem I} \\
&= \langle X^2 \rangle_t - \langle X \rangle_0^2 - \frac{t^2}{\mu^2}\langle P \rangle^2 + \frac{2t}{\mu}\langle X \rangle_0 \langle P \rangle
\end{aligned}$$

We take the derivative of $\langle X^2 \rangle$ in order to analyze it:

$$\frac{\partial}{\partial t}\langle X^2 \rangle = \frac{1}{i\hbar}\langle [X^2, H] \rangle = \frac{1}{2i\mu\hbar}\langle [X^2, P^2] \rangle$$

In order to go ahead we'll use the following trick:

Theorem 7.1. $[AB, C] = A[B, C] + [A, C]B$

Proof.

$$\begin{aligned}
A[B, C] + [A, C]B &= A(BC - CB) + (AC - CA)B \\
&= ABC - ACB + ACB - CAB \\
&= ABC - CAB = [AB, C]
\end{aligned}$$

□

Thus, $[X^2, P^2]$ can be rewritten as:

$$\begin{aligned}
[X^2, P^2] = [XX, P^2] &= X[X, P^2] + [X, P^2]X \\
&= X\left(i\hbar\frac{\partial}{\partial P}(P^2)\right) + \left(i\hbar\frac{\partial}{\partial P}(P^2)\right)X
\end{aligned}$$

$$= 2i\hbar(XP + PX)$$

Thus, if we go back to the derivative above:

$$\frac{\partial}{\partial t}\langle X^2\rangle = \frac{1}{2i\mu\hbar}\langle [X^2, P^2]\rangle = \frac{1}{\mu}\langle XP + PX\rangle$$

which still does not tell us much. We try to see if the second derivative give us some relevant information:

$$\frac{\partial^2}{\partial t^2}\langle X^2\rangle = \frac{1}{\mu}\frac{\partial}{\partial t}\langle XP + PX\rangle$$

And we (again) replace the temporal derivative of the expectation of this observable:

$$\begin{aligned}
\frac{\partial}{\partial t}\langle XP + PX\rangle &= \frac{1}{i\hbar}\langle [XP + PX, H]\rangle \\
&= \frac{1}{2i\mu\hbar}\langle [XP + PX, P^2]\rangle \\
&= \frac{1}{2i\mu\hbar}\langle [XP, P^2] + [PX, P^2]\rangle && \text{linearity of commutator} \\
&= \frac{1}{2i\mu\hbar}\Big(\langle [XP, P^2]\rangle + \langle [PX, P^2]\rangle\Big) && \text{linearity of expectation}
\end{aligned}$$

Let us treat the two commutators separately:

$$\begin{aligned}
[XP, P^2] &= X[P, P^2] + [X, P^2]P && \text{by Theorem 7.1} \\
&= [X, P^2]P \\
&= \left(i\hbar\frac{\partial}{\partial P}P^2\right)P \\
&= 2i\hbar P^2
\end{aligned}$$

and, by the same argument:

$$[PX, P^2] = P[X, P^2] + [P, P^2]X = P(2i\hbar P) = 2i\hbar P^2$$

Thus the second derivative of $\langle X^2\rangle$ becomes:

$$\begin{aligned}
\frac{\partial^2}{\partial t^2}\langle X^2\rangle &= \frac{1}{\mu}\frac{\partial}{\partial t}\langle XP + PX\rangle \\
&= \frac{1}{\mu}\frac{1}{2i\mu\hbar}\Big(\langle [XP, P^2]\rangle + \langle [PX, P^2]\rangle\Big) \\
&= \frac{1}{\mu}\frac{1}{2i\mu\hbar}\langle 4i\hbar P^2\rangle \\
&= \frac{2}{\mu^2}\langle P^2\rangle
\end{aligned}$$

This leaves us with the following second-order differential equation:

$$\frac{\partial^2}{\partial t^2}\langle X^2\rangle = \frac{2}{\mu^2}\langle P^2\rangle \Longrightarrow \langle X^2\rangle_t = \frac{1}{\mu^2}\langle P^2\rangle t^2 + Ct + D$$

In order to find C, D we impose the initial conditions:

$$\begin{cases} C = \left[\frac{\partial}{\partial t}\langle X^2\rangle\right]_{t=0} = \frac{1}{\mu}\langle XP + PX\rangle_0 \\ D = \langle X^2\rangle_0 \end{cases}$$

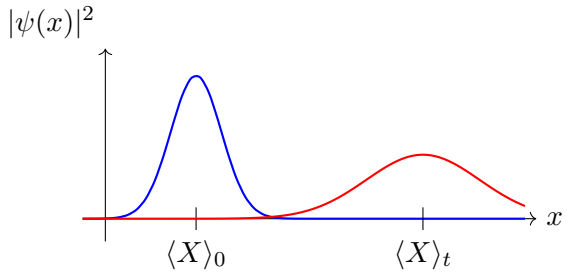


Figure 7.3: The Gaussian wave of the position at time t (in red) has spread with respect to the initial wave function (blue).

Putting it all together we finally obtain the expression for ΔX^2:

$$\begin{aligned}\left(\Delta X^2\right)_t &= \frac{t^2}{\mu^2}\langle P^2\rangle + \frac{t}{\mu}\langle XP+PX\rangle_0 + \langle X^2\rangle_0 - \langle X\rangle_0^2 - \frac{t^2}{\mu^2}\langle P\rangle^2 + \frac{2t}{\mu}\langle X\rangle_0\langle P\rangle \\ &= \frac{t^2}{\mu^2}\left(\Delta P^2\right) + \frac{t}{\mu}\left(\langle XP+PX\rangle_0 + 2\langle X\rangle_0\langle P\rangle\right) + \left(\Delta X^2\right)_0\end{aligned}$$

This is telling us that the variance (and thus our uncertainty about the position), grows quadratically over time. This concept is visualized in Figure 7.3.

7.4 Wave packets

A special form of wave functions are the so called **wave packets**.

Definition 7.2. *A wave packet is a wave function of the form:*

$$\psi(x) = \phi(x - x_0)\cdot e^{ikx}$$

*where $\phi(x)$ is a real function called the **envelope function** of the wave packet, while k is the **frequency**.*

In this case, the expectation $\langle X\rangle$ corresponds to:

$$\begin{aligned}\langle X\rangle &= \int_{\mathbb{R}} x|\psi(x)|^2dx \\ &= \int_{\mathbb{R}} x|\phi(x-x_0)|^2dx \\ &= \int_{\mathbb{R}} (x+x_0)|\phi(x)|^2dx && \text{substitution } x \leftarrow x + x_0 \\ &= x_0 + \int_{\mathbb{R}} x|\phi(x)|^2dx\end{aligned}$$

i.e. $\langle X\rangle$ depends exclusively on the envelope. From now on we will assume without loss of generality that $\phi(x)$ is zero-mean, i.e. the second term vanishes and the expectation of the wave packet is exactly the shift x_0: notice that we can always reduce to this case by redefining ϕ with a shift. Wave packets are particularly interesting also because of the neat form of the momentum wave function (which also explains the role of the term e^{ikx}):

Theorem 7.3. *Consider a wave packet $\psi(x)$ with frequency k and envelope ϕ, where the envelope has expectation x_0. The momentum wave function $\bar{\psi}(p)$ is:*

$$\bar{\psi}(p) = \bar{\phi}(p - \hbar k) \cdot e^{-ipx_0/\hbar}$$

where $\bar{\phi}$ denotes the Fourier transform of the envelope.

Proof.

$$\begin{aligned}
\bar{\psi}(p) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \psi(x) e^{-ipx/\hbar} dx \\
&= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \phi(x - x_0) e^{ikx} e^{-ipx/\hbar} dx \\
&= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \phi(x) e^{ik(x+x_0)} e^{-ip(x+x_0)/\hbar} dx \qquad \text{substitution } x \leftarrow x + x_0 \\
&= \frac{1}{\sqrt{2\pi\hbar}} e^{ikx_0} e^{-ipx_0/\hbar} \int_{\mathbb{R}} \phi(x) e^{ikx} e^{-ipx/\hbar} dx \\
&= \frac{1}{\sqrt{2\pi\hbar}} e^{-i(p-\hbar k)x_0/\hbar} \int_{\mathbb{R}} \phi(x) e^{-i(p-\hbar k)x/\hbar} dx \\
&= \bar{\phi}(p - \hbar k) \cdot e^{-i(p-\hbar k)x_0/\hbar}
\end{aligned}$$

where the $\hbar k$ at the exponent gives only a global phase. □

Notice the symmetry: while ψ has expectation x_0 and frequency k, the momentum wave function has expectation $p_0 = \hbar k$ and frequency $x_0/\hbar$ [this needs to be further justified; check proof of this statement in book, and turn it into a lemma.]. The symmetry would be even more evident if we redefined a wave packet as:

$$\psi(x) = \phi(x - x_0) \cdot e^{ip_0 x/\hbar} \longleftrightarrow \bar{\psi}(p) = \bar{\phi}(p - p_0) \cdot e^{-ipx_0/\hbar}$$

On the other hand, the elegance of this statement not only lies on the symmetry it shows. We expect that using $\phi(x)$ with mean x_0 instead of $\phi(x - x_0)$ with zero-mean does not change anything. Indeed, by property A.7 of the Fourier transform, a shift x_0 yields a factor $e^{-ipx_0/\hbar}$ on the total transform.

7.5 Gaussian wave packets

A wave packet is said to be *Gaussian* if $\phi(x)$ has the following form:

$$\phi(x - x_0) = \frac{1}{\sqrt[4]{2\pi\sigma^2}} \cdot e^{-\frac{(x-x_0)^2}{4\sigma^2}}$$

This name comes from the fact that $|\psi(x)|^2$ is exactly a Gaussian distribution (the term e^{ikx} always has magnitude 1):

$$|\psi(x)|^2 = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-x_0)^2}{2\sigma^2}}$$

and we know from probability theory (Theorems C.10 and C.11) that this yields $\langle X \rangle = x_0$ and $\Delta X^2 = \sigma^2$.

Gaussian wave packets nicely explain where the term "envelope" comes from. In fact, if we plot the real and imaginary parts of $\psi(x)$ (Figure 7.4), we can see that they are nicely enclosed between $\phi(x)$ and $-\phi(x)$.

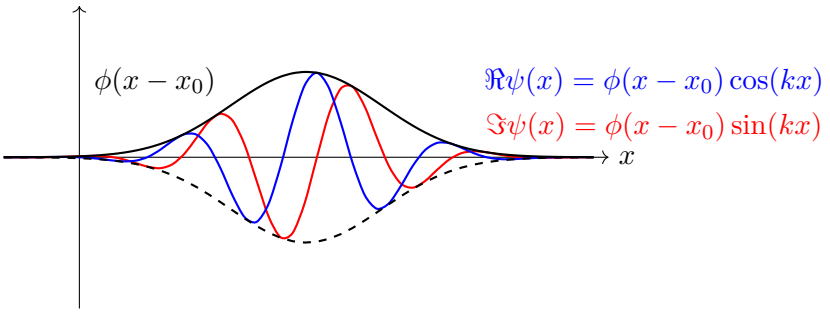


Figure 7.4: Plot of the real and imaginary part of a Gaussian wave packet.

The derivation of the momentum wave function is a bit more challenging: Gaussian wave packets are interesting because Gaussians are **eigenfunctions** of the Fourier transform (i.e. transforming a Gaussian still yields a Gaussian):

$$\phi(x) = \frac{1}{\sqrt[4]{2\pi\sigma^2}} \cdot e^{-\frac{x^2}{4\sigma}} \Longrightarrow \bar{\phi}(p) = \sqrt[4]{\frac{2\sigma^2}{\pi\hbar^2}} \cdot e^{-\frac{\sigma^2 p^2}{\hbar^2}}$$

Thus, we deduce (with the aid of Theorem 7.3) that the momentum wave function yields a Gaussian wave packet with expectation $\langle P \rangle = \hbar k$ and variance $\Delta P^2 = \frac{\hbar^2}{4\sigma^2}$.

Free evolution of a Gaussian wave packet. Let us see an example of free evolution for these types of waves: now that we know expectation and variance of position and momentum we can apply the expressions we derived in Sections 7.2 and 7.3:

$$\begin{aligned} \langle X \rangle_t &= \langle X \rangle_0 + \frac{\langle P \rangle}{\mu} t = x_0 + \frac{\hbar k}{\mu} t \\ \langle P \rangle_t &= \langle P \rangle = \hbar k \\ \left(\Delta P^2\right)_t &= \Delta P^2 = \frac{\hbar^2}{4\sigma^2} \end{aligned}$$

What is left to compute is the variance of the position. We will not show here that, in the case of a Gaussian wave packet:

$$\langle XP + PX \rangle_0 + 2\langle X \rangle_0 \langle P \rangle = 0$$

i.e. the linear term of the expression for $\left(\Delta X^2\right)_t$ vanishes. This leaves us with:

$$\Delta X^2 = \frac{\Delta P^2}{\mu^2} t^2 + \left(\Delta X^2\right)_0 = \frac{\hbar^2}{4\sigma^2\mu^2} t^2 + \sigma^2$$

7.6 Group and phase velocity

Consider a generic wave packet for a free particle evolving over time, with momentum wave function:

$$\bar{\psi}(p, t) = \bar{\phi}(p - p_0) \cdot e^{-ipx_0/\hbar} \cdot e^{-i\omega(p/\hbar)t}$$

The rightmost term is actually new, and it directly comes from the term $e^{-itH/\hbar}$ in the evolving state: $\omega(p/\hbar) = \omega(k)$ is a temporal frequency, and it tells us how the phase changes through

time (note that the term $e^{-ipx_0/\hbar}$ we already had gives us the evolution of the phase over space). The position wave function can be found by an inverse Fourier transform:

$$\begin{aligned}
\psi(x,t) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \bar{\phi}(p-p_0) \cdot e^{-ipx_0/\hbar} \cdot e^{-i\omega(p/\hbar)t} \cdot e^{ipx/\hbar} dp \\
&= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \bar{\phi}(p-p_0) \cdot e^{ip(x-x_0)/\hbar} \cdot e^{-i\omega(p/\hbar)t} dp \\
&= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \hbar\bar{\phi}(\hbar k-p_0) \cdot e^{-i\omega(k)t} \cdot e^{ik(x-x_0)} dk \qquad \text{replacing } p \leftarrow \hbar k
\end{aligned}$$

Now we will do a Taylor expansion of $\omega(k)$ around a particular point $k = k_0 = p_0/\hbar$ (which, in general, is the peak of the envelope), in particular we take k_0 as the peak of the envelope of $\bar{\phi}(\hbar k - p_0)$:

$$\omega(k) \simeq \omega(k_0) + \omega'(k_0)(k-k_0) =: \omega_0 + \omega_0'(k-k_0)$$

The transform can be rewritten as:

$$\begin{aligned}
\psi(x,t) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \bar{\phi}(\hbar k-p_0) \cdot e^{-i\omega(k)t} \cdot e^{ik(x-x_0)} dk \\
&\simeq \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \hbar\bar{\phi}(\hbar k-p_0) \cdot e^{-i(\omega_0+\omega_0'(k-k_0))t} \cdot e^{ik(x-x_0)} dk \\
&= \frac{1}{\sqrt{2\pi\hbar}} \cdot e^{i(k_0x-k_0x_0-\omega_0 t)} \int_{\mathbb{R}} \hbar\bar{\phi}(\hbar k-p_0) \cdot e^{i(k-k_0)(x-x_0-\omega_0' t)} dk \\
&= \frac{1}{\sqrt{2\pi\hbar}} \cdot e^{ik_0(x-x_0-\frac{\omega_0}{k_0}t)} \int_{\mathbb{R}} \bar{\phi}(p-p_0) \cdot e^{i(p-p_0)(x-x_0-\omega_0' t)/\hbar} dp \\
&= e^{ik_0(x-x_0-\frac{\omega_0}{k_0}t)} \cdot \phi(x-x_0-\omega_0' t)
\end{aligned}$$

We found that the position wave function can be decomposed into:

- An envelope term, whose peak moves with velocity $v_{gr} = \omega_0'$, called **group velocity**;
- A plane wave term, with a phase propagating with velocity $v_{ph} = \frac{\omega_0}{k_0}$ called **phase velocity**.

On this Wikipedia page there is a nice visualization of some evolving wave packets with different group and phase velocities.

Chapter 8
Dynamics of Composed Systems

8.1 Non-interacting systems

Consider two systems with state spaces $\mathcal{H}_A, \mathcal{H}_B$ that are **non-interacting**, i.e. the Hamiltonian H of the total system $\mathcal{H}_A \otimes \mathcal{H}_B$ can be expressed as

$$H = H_A \otimes \mathbb{1}_B + \mathbb{1}_A \otimes H_B$$

In this section we will see how the evolution of the two systems will remain separated (which makes sense as they are non-interacting, even if we are considering them as a single system).

Computing the evolution operator. We want to know what the evolution operator $U(t)$ of two non-interacting systems in tensor product looks like. We start by writing $U(t)$ in terms of the Hamiltonian:

$$U(t) = \exp\left(-\frac{it}{\hbar}H\right) = \exp\left(-\frac{it}{\hbar}(H_A \otimes \mathbb{1}_B + \mathbb{1}_A \otimes H_B)\right)$$

Now notice that the commutator $[H_A \otimes \mathbb{1}, \mathbb{1} \otimes H_B]$ is:

$$\begin{aligned}[H_A \otimes \mathbb{1}_B, \mathbb{1}_A \otimes H_B] &= (H_A \otimes \mathbb{1}_B)(\mathbb{1}_A \otimes H_B) - (\mathbb{1}_A \otimes H_B)(H_A \otimes \mathbb{1}_B) \\ &= H_A \otimes H_B - H_A \otimes H_B = 0\end{aligned}$$

i.e. operators acting on non-interacting systems always **commute**, which again makes sense: transformations acting on independent systems should not interfere which each other. Thus the property of exponentials (Theorem 4.3) holds:

$$\begin{aligned}U(t) &= \exp\left(-\frac{it}{\hbar}(H_A \otimes \mathbb{1}_B + \mathbb{1}_A \otimes H_B)\right) \\ &= \exp\left(-\frac{it}{\hbar}(H_A \otimes \mathbb{1}_B)\right) \cdot \exp\left(-\frac{it}{\hbar}(\mathbb{1}_A \otimes H_B)\right)\end{aligned}$$

We found that the evolution operator can be expressed as the product of the exponential of local operators (i.e. operators acting only on one subsystem). Moreover, notice that

$$(H_A \otimes \mathbb{1}_B)^n = (H_A^n \otimes \mathbb{1}_B^n) = (H_A^n \otimes \mathbb{1}_B)$$

as we can just distribute the operator multiplication with respect to the tensor product. With a Taylor series trick, we obtain:

$$\begin{aligned}\exp\left(-\frac{it}{\hbar}(H_A \otimes \mathbb{1}_B)\right) &= \sum_{n=0}^{\infty}\left(-\frac{it}{\hbar}\right)^n (H_A \otimes \mathbb{1}_B)^n \\ &= \sum_{n=0}^{\infty}\left(-\frac{it}{\hbar}\right)^n (H_A^n \otimes \mathbb{1}_B) \\ &= \left(\sum_{n=0}^{\infty}\left(-\frac{it}{\hbar}\right)^n H_A^n\right) \otimes \mathbb{1}_B \\ &= \exp\left(-\frac{it}{\hbar}H_A\right) \otimes \mathbb{1}_B\end{aligned}$$

A similar argument can be applied for the other term by bilinearity of the tensor product, and putting it all together we obtain:

$$\begin{aligned}
U(t) &= \exp\left(-\frac{it}{\hbar}(H_A \otimes \mathbb{1}_B)\right) \cdot \exp\left(-\frac{it}{\hbar}(\mathbb{1}_A \otimes H_B)\right) \\
&= \left(\exp\left(-\frac{it}{\hbar}H_A\right) \otimes \mathbb{1}_B\right) \cdot \left(\mathbb{1}_A \otimes \exp\left(-\frac{it}{\hbar}H_B\right)\right) \\
&= \left(\exp\left(-\frac{it}{\hbar}H_A\right) \cdot \mathbb{1}_A\right) \otimes \left(\mathbb{1}_B \cdot \exp\left(-\frac{it}{\hbar}H_B\right)\right) \\
&= \exp\left(-\frac{it}{\hbar}H_A\right) \otimes \exp\left(-\frac{it}{\hbar}H_B\right) \\
&= U_A(t) \otimes U_B(t)
\end{aligned}$$

In fact, this is the only result we could expect: in this way, any unentangled state will keep evolving the subsystems independently as

$$(U_A(t) \otimes U_B(t))(|\psi_A\rangle \otimes |\psi_B\rangle) = (U_A(t)|\psi_A\rangle) \otimes (U_B(t)|\psi_B\rangle)$$

8.2 Local observables

Suppose now that we have a **local** observable, i.e. that only observes one of the subsystems:

$$M = M_A \otimes \mathbb{1}_B$$

We would like to see what happens to its expectation over time. As we already know, we need to look at the commutator with the Hamiltonian:

$$\frac{\partial}{\partial t}\langle M\rangle = \frac{1}{i\hbar}\langle [M, H]\rangle$$

Let us see if we can simplify the commutator:

$$\begin{aligned}
[M, H] &= [M_A \otimes \mathbb{1}_B, (H_A \otimes \mathbb{1}_B + \mathbb{1}_A \otimes H_B)] \\
&= [M_A \otimes \mathbb{1}_B, H_A \otimes \mathbb{1}_B] + [M_A \otimes \mathbb{1}_B, \mathbb{1}_A \otimes H_B] \\
&= (M_A \otimes \mathbb{1}_B)(H_A \otimes \mathbb{1}_B) - (H_A \otimes \mathbb{1}_B)(M_A \otimes \mathbb{1}_B) + \\
&\quad + (M_A \otimes \mathbb{1}_B)(\mathbb{1}_A \otimes H_B) - (\mathbb{1}_A \otimes H_B)(M_A \otimes \mathbb{1}_B) \\
&= M_A H_A \otimes \mathbb{1}_B - H_A M_A \otimes \mathbb{1}_B + M_A \otimes H_B - M_A \otimes H_B \\
&= (M_A H_A - H_A M_A) \otimes \mathbb{1} \\
&= [M_A, H_A] \otimes \mathbb{1}
\end{aligned}$$

And this is enough to prove that only the commutator of the first subsystem influences the evolution of the observable as

$$\begin{aligned}
\frac{\partial}{\partial t}\langle M\rangle &= \frac{1}{i\hbar}\langle [M, H]\rangle \\
&= \frac{1}{i\hbar}\langle [M_A, H_A] \otimes \mathbb{1}_B\rangle \\
&= \frac{1}{i\hbar}\langle \psi(t)| \left([M_A, H_A] \otimes \mathbb{1}_B\right) |\psi(t)\rangle \\
&= \frac{1}{i\hbar}\langle \psi(0)|U^\dagger(t) \cdot ([M_A, H_A] \otimes \mathbb{1}_B) \cdot U(t)|\psi(0)\rangle
\end{aligned}$$

$$= \frac{1}{i\hbar}\langle\psi(0)|\left(U_A^\dagger(t)[M_A, H_A]U_A(t)\right) \otimes \left(U_B^\dagger(t)\cdot U_B(t)\right)|\psi(0)\rangle$$
$$= \frac{1}{i\hbar}\langle\psi(0)|\left(U_A^\dagger(t)[M_A, H_A]U_A(t)\right) \otimes (\mathbb{1}_B)|\psi(0)\rangle$$

The last two steps are quite insightful: they are basically saying that no matter the evolution of the second system, we will get the identity on its term, i.e. the local observable M cannot distinguish possible evolutions of the system B, which seems intuitive, given the definition of M. Indeed, if $|\psi\rangle$ is unentangled, i.e. $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$, the above expression simplifies to:

$$\frac{\partial}{\partial t}\langle M\rangle = \frac{1}{i\hbar}\langle\psi_A(0)|\left(U_A^\dagger(t)[M_A, H_A]U_A(t)\right)|\psi_A(0)\rangle\langle\psi_B(0)|\mathbb{1}_B|\psi_B(0)\rangle$$
$$= \frac{1}{i\hbar}\langle\psi_A(t)|[M_A, H_A]|\psi_A(t)\rangle$$

This is what we call **non-signaling principle** between systems.

8.3 Indirect measurement and the modular momentum

In this section we discuss a general method to implement observables that are not directly measurable, then we will see a concrete example with a deeper analysis of the Stern-Gerlach experiment, which we discussed briefly in Section 3.4. Consider two Hilbert spaces:

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

and an observable M on $\mathcal{H}_A$ that is not directly measurable, for example an internal degree of freedom (like spin) of a particle:

$$M_A = \sum_k a_k |k\rangle\langle k|_A$$

The second subsystem represents something easily measurable, such as the position of some particle. The idea is to evolve the system using a suitable interaction Hamiltonian, in such a way that, after the evolution, the subsystem we *want* to measure and the one we *can* measure are entangled in a convenient way. At this point it will be sufficient to carry out a measurement on the latter. For this purpose, we propose the following interaction Hamiltonian:

$$H = M_A \otimes P_x \Longrightarrow U(t) = \exp\left(-\frac{it}{\hbar}(M_A \otimes P_x)\right)$$

It is worth noticing that this is not the same as the Hamiltonian in non-interacting systems we discussed in Section 8.1.

Given an initial state

$$|\tilde{\psi}_0\rangle = \left(\sum_k c_k|k\rangle_A\right) \otimes |\psi_0\rangle_x = \sum_k c_k|k\rangle_A \otimes |\psi_0\rangle_x$$

its evolution can be computed by using the definition of exponential (as usual):

$$U(t)|\tilde{\psi}_0\rangle = \exp\left(-\frac{it}{\hbar}(M_A \otimes P_x)\right)\left(\sum_k c_k|k\rangle_A \otimes |\psi_0\rangle_x\right)$$
$$= \exp\left(-\frac{it}{\hbar}\left(\sum_k a_k|k\rangle\langle k|_A \otimes P_x\right)\right)\left(\sum_k c_k|k\rangle_A \otimes |\psi_0\rangle_x\right)$$

$$
\begin{aligned}
&= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{it}{\hbar}\right)^n \left(\sum_k a_k |k\rangle\langle k|_A \otimes P_x\right)^n \left(\sum_k c_k |k\rangle_A \otimes |\psi_0\rangle_x\right) \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{it}{\hbar}\right)^n \left(\left(\sum_k a_k |k\rangle\langle k|_A\right)^n \otimes P_x^n\right) \left(\sum_k c_k |k\rangle_A \otimes |\psi_0\rangle_x\right) \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{it}{\hbar}\right)^n \left(\sum_k a_k^n |k\rangle\langle k|_A \otimes P_x^n\right) \left(\sum_k c_k |k\rangle_A \otimes |\psi_0\rangle_x\right) \\
&= \sum_k \sum_{k'} \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{it}{\hbar}\right)^n \left(a_k^n |k\rangle\langle k|_A \otimes P_x^n\right) \left(c_{k'} |k'\rangle_A \otimes |\psi_0\rangle_x\right) \\
&= \sum_k \sum_{k'} \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{it}{\hbar}\right)^n \left(a_k^n c_{k'} |k\rangle\langle k|k'\rangle_A \otimes P_x^n |\psi_0\rangle_x\right) \\
&= \sum_k \sum_{k'} \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{it}{\hbar} a_k\right)^n \left(c_{k'} \delta_{k,k'} |k\rangle \otimes P_x^n |\psi_0\rangle_x\right) \\
&= \sum_k \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{it}{\hbar} a_k\right)^n \left(c_k |k\rangle_A \otimes P_x^n |\psi_0\rangle_x\right) \\
&= \sum_k c_k |k\rangle_A \otimes \left(\sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{it}{\hbar} a_k P_x\right)^n\right) |\psi_0\rangle_x \\
&= \sum_k c_k |k\rangle_A \otimes \exp\left(-\frac{it}{\hbar} a_k P_x\right) |\psi_0\rangle_x
\end{aligned}
$$

This tells us that this Hamiltonian perfectly entangles the two subsystems in such a way that the state of the second subsystem has a term a_k upon measurement of state $|k\rangle$ on the first subsystem. We let $L_k = t \cdot a_k$, and define a new operator called the **modular momentum** operator:

$$\exp\left(-\frac{i}{\hbar} L_k P\right)$$

We use the modular momentum operator to analyze the evolution of the position state, expressing it in the momentum basis:

$$
\begin{aligned}
e^{-iL_k P/\hbar} |\psi_0\rangle &= \int_{\mathbb{R}} e^{-iL_k P/\hbar} \bar{\psi}_0(p) |p\rangle dp \\
&= \int_{\mathbb{R}} \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar} L_k P\right)^n \bar{\psi}_0(p) |p\rangle dp \\
&= \int_{\mathbb{R}} \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar} L_k p\right)^n \bar{\psi}_0(p) |p\rangle dp \qquad\qquad \text{since } P^n |p\rangle = p^n |p\rangle \\
&= \int_{\mathbb{R}} \bar{\psi}_0(p)\ e^{-iL_k p/\hbar}\ |p\rangle dp
\end{aligned}
$$

Thus the momentum wave function of $|\psi\rangle = e^{-iL_k P} |\psi_0\rangle$ is:

$$\bar{\psi}(p) = \bar{\psi}_0(p)\ e^{-iL_k p/\hbar}$$

We just found that the modular momentum operator acts on the momentum wave function as:

$$\exp\left(-\frac{i}{\hbar} L_k P\right) : \bar{\psi}_0(p) \to \bar{\psi}_0(p)\ e^{-iL_k p/\hbar}$$

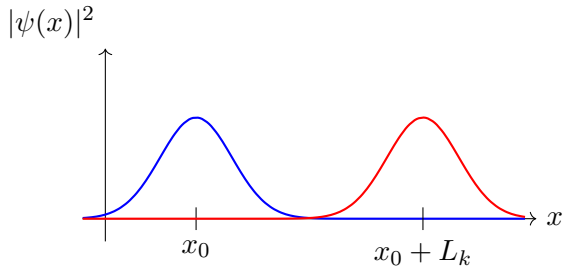


Figure 8.1: The modular momentum operator shifts the position wave function by L_k towards the positive direction of the axis.

Now we can find the position wave function with an inverse Fourier transform:

$$\begin{aligned}\psi(x) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \bar{\psi}_0(p)\ e^{-iL_k p/\hbar}\ e^{ipx/\hbar} dp \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \bar{\psi}_0(p)\ e^{ip(x-L_k)/\hbar} dp \\ &= \psi_0(x - L_k).\end{aligned}$$

We found that the modular momentum operator $\exp(-iLP/\hbar)$ shifts the **position** by L_k to the right, as depicted in Figure 8.1, and generally acts on the position wave function as

$$\exp\left(-\frac{i}{\hbar} L_k P\right) : \psi_0(x) \to \psi_0(x - L_k).$$

If we go back to our evolution and replace L_k we obtain:

$$\begin{aligned}U(t)|\tilde{\psi}_0\rangle &= \sum_k c_k |k\rangle_A \otimes \exp\left(-\frac{it}{\hbar} a_k P_x\right) |\psi_0\rangle_x \\ &= \sum_k c_k |k\rangle_A \otimes \int_{\mathbb{R}} \psi_0(x - t a_k) |x\rangle dx\end{aligned}$$

We actually associated the state of the first subsystem with a position shift in the second subsystem. In the case where $|\psi_0\rangle$ has a Gaussian wave function, we can also see that the wave stayed intact, but its peak is shifted. Thus, if we wait for long enough (i.e. t sufficiently large, Figure 8.2) we can make the points $\{t \cdot a_k\}_k$ sufficiently distant from each other, making the overlaps between waves negligible. At this point one can construct an observable on the position by partitioning the x-axis in regions centered on the peaks $\{t \cdot a_k\}_k$. Notice that the actual Gaussians in the induced probability mixture are also scaled by the factors c_i: peaks associated with more probable $|k\rangle$ must also be observed with higher probability.

We want to stress the fact that the time t we let the system evolve before measuring the position must be chosen in relation to the variance of the initial position state. If the value of t is too small, it will give rise to a problem of **weak measurements** (we obtain little information about the first subsystem through a measurement of the position, as the peaks are too close: see Figure 8.3). The same thing could happen if the initial variances of the position were too large. There is still active research on this topic.

Another choice of coupling between the system to be measured and the pointer is given by the Hamiltonian

$$H = M_A \otimes X_x \implies U(t) = \exp\left(-\frac{it}{\hbar}(M_A \otimes X_x)\right).$$

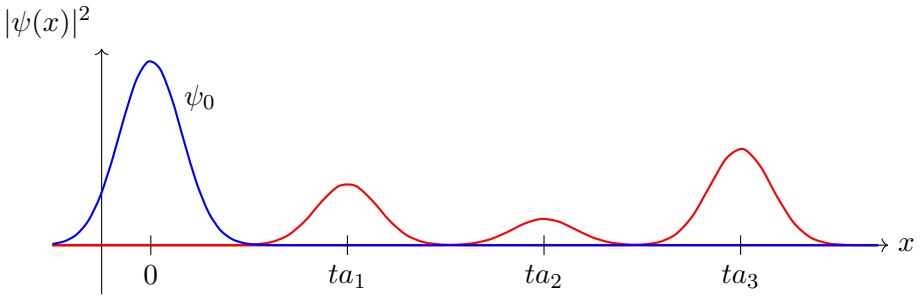


Figure 8.2: The initial Gaussian wave function (blue) evolves to become a superposition of Gaussians with different means (red).

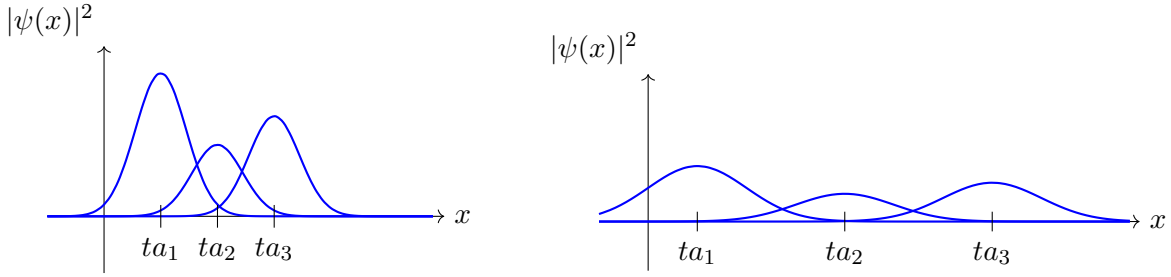


Figure 8.3: Examples of weak measurements due to small t (left) and high initial uncertainty in position (right).

The final effect will be analogous to the previous Hamiltonian: it will shift the momentum wave function and multiply the position wave function by an exponential. To see this, we can apply the unitary to the initial state:

$$\begin{aligned} U(t)|\tilde{\psi}_0\rangle &= \exp\left(-\frac{it}{\hbar}\left(M_A \otimes X_x\right)\right)\left(\sum_k c_k|k\rangle_A \otimes |\psi_0\rangle_x\right) \\ &= \sum_k c_k|k\rangle_A \otimes \exp\left(-\frac{it}{\hbar}a_k X_x\right)|\psi_0\rangle_x. \end{aligned}$$

This leads us to define the **modular position** operator,

$$\exp\left(-\frac{i}{\hbar}L_k X\right),$$

where again $L_k = a_k t$. We can see how this operator acts on the initial state by expanding the latter, this time in the position basis,

$$\begin{aligned} e^{-iL_k X/\hbar}|\psi_0\rangle &= \int_{\mathbb{R}} e^{-iL_k X/\hbar}\psi_0(x)|x\rangle dx \\ &= \int_{\mathbb{R}} \psi_0(p)\ e^{-iL_k x/\hbar}\ |x\rangle dx \end{aligned}$$

Thus the modular position acts on the position wave function as

$$\exp\left(-iL_k X/\hbar\right):\ \psi_0(x) \to \psi_0(p)\ e^{-iL_k x/\hbar}.$$

By applying the reverse Fourier transform we can see how it acts on the position wave function:

$$\bar{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}}\int_{\mathbb{R}} \psi_0(x)\ e^{-iL_k x/\hbar}\ e^{-ipx/\hbar}dp$$

$$= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \psi_0(x)\ e^{ix(p+L_k)/\hbar} dp$$
$$= \psi_0(p + L_k).$$

Overall, we have that our evolution results in the state transformation

$$U(t)|\tilde{\psi}_0\rangle = \sum_k c_k |k\rangle_A \otimes \exp\left(-\frac{it}{\hbar} a_k X_x\right) |\psi_0\rangle_x$$
$$= \sum_k c_k |k\rangle_A \otimes \int_{\mathbb{R}} \bar{\psi}_0(p + ta_k)|p\rangle dp.$$

How do we know which interaction Hamiltonian to choose — the one that couples to momentum (and shifts position) or the one that couples to position (and shifts momentum)? This depends on the experimental setup, how we are planning to observe our pointer system, and also what physics is available for us to implement one or the other. In the following section we will see an example through the Stern-Gerlach experiment. There we will end up choosing to couple the internal spin of a particle to its position, leading to a selective shift in momentum. This is because we can then let this particle evolve under a free Hamiltonian, so that a small kick in momentum can result in a large difference in position after some time.

8.4 The Stern-Gerlach experiment: part 2

Here we present a more concrete example, which was first introduced in Section 3.4. An electron moving on a plane has one **internal** degree of freedom (its spin), which can be modeled as qubit:

$$\mathcal{H}_S = \text{span}\{|0\rangle, |1\rangle\}$$

and two **orbital** degrees of freedom, i.e. its position in the plane:

$$\mathcal{H}_x = \text{span}\{|x\rangle, x \in \mathbb{R}\}, \quad \mathcal{H}_y = \text{span}\{|y\rangle, y \in \mathbb{R}\}$$

The total state space of the electron can thus be modeled as

$$\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_x \otimes \mathcal{H}_y$$

We already saw what the Stern-Gerlach experiment is about: measuring the spin of the electron indirectly by looking at its position. We identify three regions in the x-y plane as depicted in Figure 8.4. In particular, red shaded areas provide some intuition about the uncertainty in the position of the particle with respect to the x axis, whereas arrows indicate how the momentum changes in the different regions of space:

- A region A, where the electron starts in a state $|\tilde{\psi}_0\rangle$ with a two-dimensional Gaussian wave such that the expected momentum $\langle \mathbf{P} \rangle$ points towards the positive y-axis;
- A region B of depth δ influenced by a magnetic field;
- A region C with a screen where the electron will land, which is used for measurement.

In order to analyze the motion of the particle we decompose the Hamiltonian into a free particle component H^0 and an **interaction** Hamiltonian H^{int}, which should take into account the force due to the magnetic field in region B.

$$H = H^0 + H^{int}$$

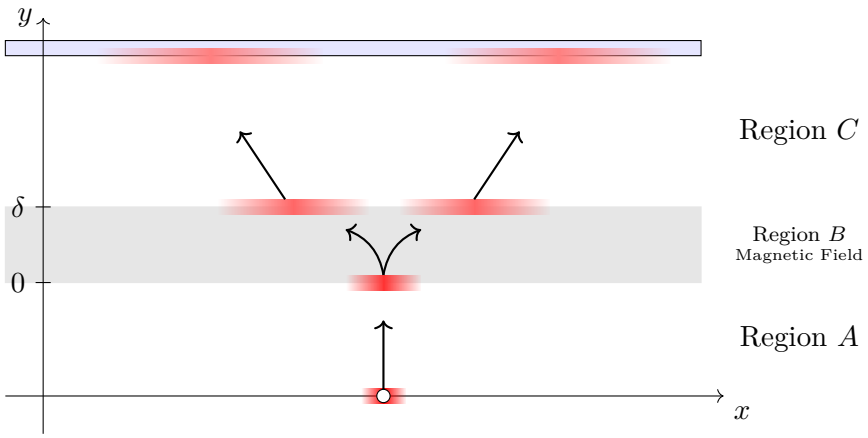


Figure 8.4: A scheme for the Stern-Gerlach experiment, which allows to indirectly determine the spin of an electron by looking at its position.

H^0 has no secrets for us:

$$H^0 = \frac{P^2}{2\mu} = \frac{1}{2\mu}\left(\mathbb{1}_S \otimes P_x^2 \otimes \mathbb{1}_y + \mathbb{1}_S \otimes \mathbb{1}_x \otimes P_y^2\right)$$

The form of the interaction Hamiltonian is a bit more interesting:

$$H^{int} = H_{sx}^{int} \otimes \Pi_y^B = H_{sx}^{int} \otimes \left(\int_0^\delta |y\rangle\langle y| dy\right)$$

This way of expressing a tensor product of an observable and a projection operator (which we have already seen in Section 3.2, when we discussed the CNOT gate), is powerful but straightforward: the interaction Hamiltonian is nonzero only when we are in region B and, in fact $\Pi_y^B|y\rangle = 0$ for every $(x, y) \notin B$. The definition of H_{sx}^{int} is something we will derive when we discuss the motion in region B.

The initial state. Let us discuss the form of $|\tilde{\psi}_0\rangle$: we consider a particle with spin and position unentangled:

$$|\tilde{\psi}_0\rangle = |\phi\rangle \otimes |\psi_0\rangle$$

where $|\phi\rangle$ is a qubit of the form:

$$|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$$

while the position of the particle in the plane is expressed as a two-dimensional Gaussian wave packet:

$$|\psi_0\rangle = \iint_{\mathbb{R}^2} e^{iky} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{4}(\mathbf{r}-\mu)^T\Sigma^{-1}(\mathbf{r}-\mu)}|\mathbf{r}\rangle d^2\mathbf{r}$$

You can check that the square of the absolute value of the wave function is exactly the probability density function of a gaussian distribution with mean $\mu = (x_0, y_0)^T$ and covariance matrix Σ. Note also that the term of the form e^{ikx} is missing, and this is because we said that the expected momentum points towards the positive y-axis (in particular, $\langle P\rangle = (0, \hbar k)^T$).

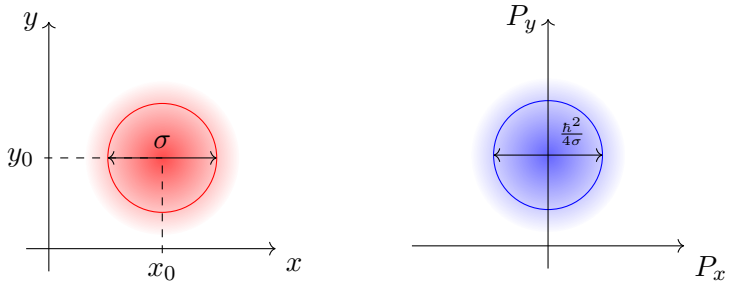


Figure 8.5: Gaussian waves in position (left) and momentum (right) basis of the initial state represented from above.

For simplicity we also assume here that $\Sigma = \sigma^2 \mathbb{1}$, so that x and y are unentangled and with the same initial variance:

$$\left(\Delta X^2\right)_0 = \left(\Delta Y^2\right)_0 = \sigma^2, \ \left(\Delta P_x^2\right)_0 = \left(\Delta P_y^2\right)_0 = \frac{\hbar^2}{4\sigma^2}$$

We can imagine to look at these Gaussian waves from above, representing higher measurement probabilities with a darker color shade as in Figure 8.5. The computation for a general initial Σ requires minor changes, which are not relevant here.

Motion in region A. Since spin and positions are unentangled here, we know that the evolution operator here can be written as tensor product:

$$\begin{aligned} U(t) &= \mathbb{1}_S \otimes \exp\left(-\frac{it}{\hbar} H_x\right) \otimes \exp\left(-\frac{it}{\hbar} H_y\right) \\ &= \mathbb{1}_S \otimes \exp\left(-\frac{it}{\hbar} \frac{P_x^2}{2\mu}\right) \otimes \exp\left(-\frac{it}{\hbar} \frac{P_y^2}{2\mu}\right) \end{aligned}$$

Thus the qubit remains unchanged:

$$\langle Z_S \rangle_t = \langle Z_S \rangle_0 = |\alpha|^2 - |\beta|^2$$

Ehrenfest's theorems give us the expression for the expectations of position and momentum:

$$\begin{cases} \langle X \rangle_t = \langle X \rangle_0 = x_0 \\ \langle Y \rangle_t = \langle Y \rangle_0 + \hbar k t = y_0 + \hbar k t \end{cases} \wedge \begin{cases} \langle P_x \rangle_t = \langle P_x \rangle_0 = 0 \\ \langle P_y \rangle_t = \langle P_y \rangle_0 = \hbar k \end{cases}$$

And we also computed the variances for Gaussian wave packets in Section 7.5:

$$\begin{cases} \left(\Delta X^2\right)_t = \frac{\hbar^2}{4\sigma^2\mu^2} t^2 + \sigma^2 = \sigma_t^2 \\ \left(\Delta Y^2\right)_t = \frac{\hbar^2}{4\sigma^2\mu^2} t^2 + \sigma^2 = \sigma_t^2 \end{cases} \wedge \begin{cases} \left(\Delta P_x^2\right)_t = \left(\Delta P_x^2\right)_0 = \frac{\hbar^2}{4\sigma^2} \\ \left(\Delta P_y^2\right)_t = \left(\Delta P_y^2\right)_0 = \frac{\hbar^2}{4\sigma^2} \end{cases}$$

As before, we give some intuituion for what the waves packets look like in Figure 8.6.

Motion in region B. Here we would like to have a precise effect: the magnetic field should push the electron to the left or to the right according to its spin. According to the evolution we had in region A, the electron must be entering region B with a state of the form:

$$|\tilde{\psi}_0\rangle = (\alpha|0\rangle_S + \beta|1\rangle_S) \otimes |\psi_0\rangle_{xy}$$

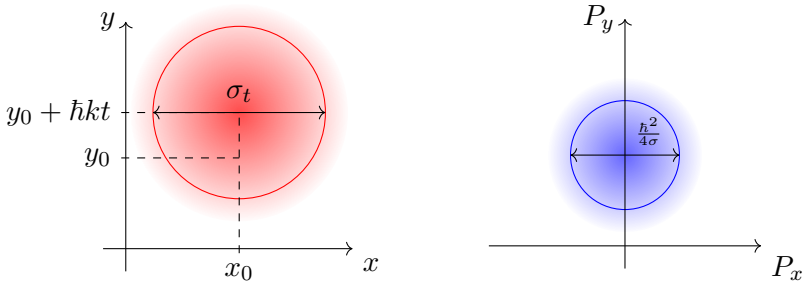


Figure 8.6: Gaussian waves in position and momentum basis for the state after the evolution in region A.

where $|\psi_0\rangle_{xy}$ has a Gaussian wave with $\langle P_y\rangle = \hbar k$ and $\langle P_x\rangle = 0$. Ideally, the final state (i.e. when the electron reaches the end of region B) should be of the form:

$$|\tilde{\psi}\rangle = \alpha|0\rangle_S \otimes |\psi^L\rangle_{xy} + \beta|1\rangle_S \otimes |\psi^R\rangle_{xy}$$

where the states $|\psi^L\rangle, |\psi^R\rangle$ enclose Gaussian waves with $\langle P_y\rangle = \hbar k$ and $\langle P_x\rangle$ respectively $-\hbar a$ and $+\hbar a$ for some $a > 0$. We anticipate that we will take a sufficiently small $\delta \simeq 0$ such that the region B has negligible thickness, and we can immediately return to the assumption of free motion. Thus, now we need to find a suitable choice of H^{int}_{sx} which gives the desired result. We will try the two coupling Hamiltonians from Section 8.3. First we try the one that couples the observable to momentum:

$$H^{int}_{sx} = M_S \otimes P_x,$$

where M is the observable on the qubit we cannot directly measure. We let $M = \gamma Z$: the γ factor will be useful to separate the waves in superposition and overcome the problem of weak measurements discussed at the end of Section 8.3 more easily. Measuring with γZ instead of Z, however, makes no difference to us. Another problem is that H^{int}_{sx}, unlike in the discussion of Section 8.3, is not the only Hamiltonian we have here, as we also have the free particle Hamiltonian H^0. Sweeping this problem under the rug for now (think of an approximate Hamiltonian here), we compute the evolution due to the interaction:

$$\begin{aligned}
U(t)|\tilde{\psi}_0\rangle &= \exp\left(-\frac{it}{\hbar}H^{int}_{sx}\right)|\tilde{\psi}_0\rangle \\
&= \exp\left(-\frac{it}{\hbar}\gamma Z \otimes P_x\right)|\tilde{\psi}_0\rangle \\
&= \alpha|0\rangle \otimes \exp\left(-\frac{it}{\hbar}\gamma P_x\right)|\psi_0\rangle + \beta|1\rangle \otimes \exp\left(+\frac{it}{\hbar}\gamma P_x\right)|\psi_0\rangle \\
&= \alpha|0\rangle \otimes \left(\int_{\mathbb{R}} \psi_0(x - t\gamma)|x\rangle dx\right) + \beta|1\rangle \otimes \left(\int_{\mathbb{R}} \psi_0(x + t\gamma)|x\rangle dx\right)
\end{aligned}$$

We obtain the decomposition of the original wave function into two wave functions, where the area under each curve is roughly $|\alpha|^2$ for the left one and $|\beta|^2$ for the right (Figure 8.7).

But at the start we said we would shift the momentum, not the position, and this would also avoid having to deal with the free particle Hamiltonian in region B (this, again, because $\delta \simeq 0$); finally this would create a sufficiently large gap in position with the free motion in region C. If we want to implement this shift in momentum we can replace the momentum operator with the position operator. We then obtain the **modular position** operator and with the exact same computation we get:

$$U(t)|\tilde{\psi}_0\rangle = \exp\left(-\frac{it}{\hbar}H^{int}_{sx}\right)$$

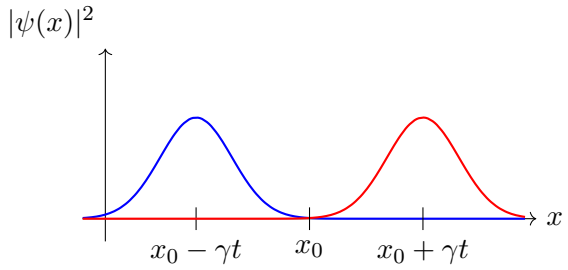


Figure 8.7: The wave function resulting from the evolution in region B.

$$
\begin{aligned}
&= \exp\left(-\frac{it}{\hbar}\gamma Z \otimes X\right)|\tilde{\psi_0}\rangle \\
&= \alpha|0\rangle \otimes \exp\left(-\frac{it}{\hbar}\gamma X\right)|\psi_0\rangle + \beta|1\rangle \otimes \exp\left(+\frac{it}{\hbar}\gamma X\right)|\psi_0\rangle \\
&= \alpha|0\rangle \otimes \left(\int_{\mathbb{R}} \tilde{\psi}_0(p + t\gamma)|p\rangle dp\right) + \beta|1\rangle \otimes \left(\int_{\mathbb{R}} \tilde{\psi}_0(p - t\gamma)|p\rangle dp\right)
\end{aligned}
$$

We presented the modular momentum first because a shift in position is more intuitive and straightforward to visualize; it will also be useful later.

Motion in region C. Since we are back to the free particle assumption we have the same evolution as in region A:

$$
U(t) = \mathbb{1}_S \otimes \exp\left(-\frac{it}{\hbar}\frac{P_x^2}{2\mu}\right) \otimes \exp\left(-\frac{it}{\hbar}\frac{P_y^2}{2\mu}\right)
$$

But this time, $\langle P_x \rangle$ is $\pm\hbar\gamma$ (the sign depends on the component of the superposition), and the particle will deviate either to the left or to the right.

State tomography. What if we wanted to estimate α and β? Assuming we are able to prepare several identical initial states, we can simply repeat the experiment and do an estimation of $|\alpha|^2$ by looking at the fraction of qubits that are measured as $|0\rangle$, ($|\beta|^2$ will follow by the normalization constraint). We know, however, that a qubit has another degree of freedom, i.e. its relative phase. We can find a second constraint, linearly independent from the first, by doing the same estimation with respect to another basis: for example we replace the Z in the interaction Hamiltonian for region B with the Pauli matrix X. This procedure however requires us to have many copies of the same state $|\psi\rangle$ we want to estimate (remember that we **cannot** clone a qubit).

Chapter 9
The Double Slit Experiment

9.1 Interference

Consider a continuous Hilbert space $\mathcal{H} = \text{span}\{|x\rangle\}_{x\in\mathbb{R}}$ (e.g. the position of a particle) and consider a state of the form:

$$|\psi\rangle = \alpha|\psi_1\rangle + \beta|\psi_2\rangle = \alpha\left(\int_{\mathbb{R}} \psi_1(x)|x\rangle dx\right) + \beta\left(\int_{\mathbb{R}} \psi_2(x)|x\rangle dx\right)$$

i.e. a superposition of two states $|\psi_1\rangle$ and $|\psi_2\rangle$. We can use linearity of integrals:

$$|\psi\rangle = \int_{\mathbb{R}} \left(\alpha\psi_1(x) + \beta\psi_2(x)\right)|x\rangle dx$$

obtaining a new wave function we will call (for obvious reasons) $\psi(x)$. The probability that the particle is measured in position $x \in [x_0, x_0 + \Delta]$ is:

$$\mathbf{P}\left([x_0, x_0+\Delta]\right) = \langle\psi|\left(\int_{x_0}^{x_0+\Delta} |x\rangle\langle x|dx\right)|\psi\rangle = \int_{x_0}^{x_0+\Delta} |\psi(x)|^2 dx$$

We can have two cases here:

- $|\psi(x)| \gg |\psi_1(x)|, |\psi_2(x)|$, i.e. although the two individual states have a low probability to be in $|x\rangle$, the resulting superposition has a high probability there. This is what we call **constructive interference**.
- $|\psi(x)| \ll |\psi_1(x)|, |\psi_2(x)|$, i.e. the resulting superposition has a much smaller probability than the single composing states, and here we talk about **destructive interference**.

This is somehow counter-intuitive: how can a state have a high probability of being at some point, say x, when it consists of a superposition of two states where the probability of being at x is low? This yields a fundamental difference between superpositions in quantum theory and probabilistic mixtures (which is also what the experiment we are going to discuss here wants to show). The mathematical reason is that, while probabilistic mixtures are expressed as real terms, wave functions in quantum mechanics are complex in general. In fact, the absolute value of the sum of two complex numbers strongly depends on the **phases** of the adding terms: if we consider a complex number of the form

$$e^{i\theta_1} + e^{i\theta_2}$$

This number can have modulo as high as 2 (e.g. when $\theta_1 = \theta_2$) or even zero (e.g. when $\theta_1 = \theta_2 + \pi$). This can also be seen graphically with sine waves: suppose that

$$\Re\{\psi_1(x)\} = \sin(x), \quad \Re\{\psi_2(x)\} = \sin(x+\phi)$$

where $\Re$ denotes the real part of a complex number (we can also analyze the imaginary part $\Im$ in a similar way, then the complex number is simply a linear combination of $\Re$ and $\Im$). Figure 9.1 shows a plot of the two sine waves (in blue and red respectively) and their sum (in black).

Moreover, as we already said, in the case where the sine waves have the same phase ($\phi = 0$), the two sine waves always have the same sign, whereas the sum of two sines with relative phase $\phi = \pi$ would be identically zero, as the two waves are pointwise opposite.

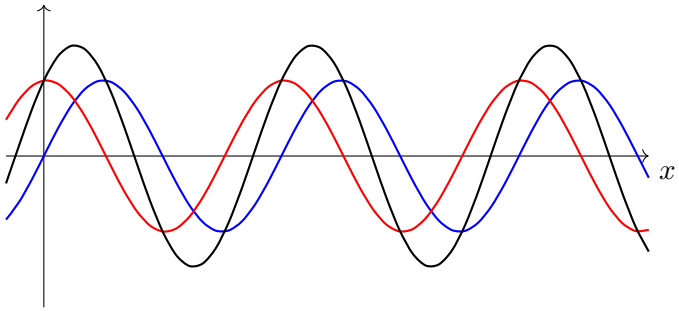


Figure 9.1: Summing sine waves.

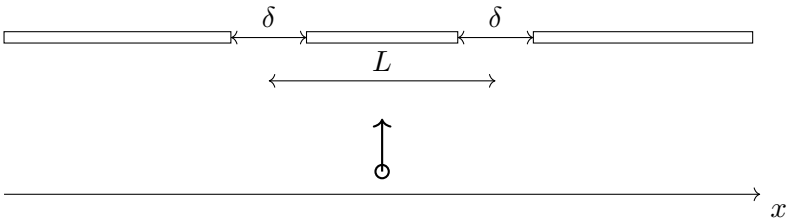


Figure 9.2: The setup for the double slit experiment.

9.2 The double slit

Suppose to have a wall of negligible thickness with two slits of width δ at distance L from each other ($x = 0$ is exactly in the middle), and a particle moving towards the wall, as depicted in Figure 9.2.

We already know that the particle propagates in the plane as a wave. If the particle hits the wall, it stops, otherwise it may pass through one of the slits[3]. This is equivalent to a **measurement**: with a certain probability the particle is in front of the slits at the moment of impact with the wall, and this corresponds to the following projection operator:

$$\Pi_{slits} = \Pi_L + \Pi_R = \int_{-\frac{L}{2}-\frac{\delta}{2}}^{-\frac{L}{2}+\frac{\delta}{2}} |x\rangle\langle x| dx + \int_{\frac{L}{2}-\frac{\delta}{2}}^{\frac{L}{2}+\frac{\delta}{2}} |x\rangle\langle x| dx$$

Thus, the state of a particle passing through one of the slits corresponds to the post-measurement state:

$$\begin{aligned}
|\psi^+\rangle &= \frac{\Pi_{slits}|\psi\rangle}{\sqrt{\langle\psi|\Pi_{slits}|\psi\rangle}} \\
&= \frac{1}{\sqrt{\langle\psi|\Pi_{slits}|\psi\rangle}} \int_{-\frac{L}{2}-\frac{\delta}{2}}^{-\frac{L}{2}+\frac{\delta}{2}} |x\rangle\langle x|\psi\rangle dx + \frac{1}{\sqrt{\langle\psi|\Pi_{slits}|\psi\rangle}} \int_{\frac{L}{2}-\frac{\delta}{2}}^{\frac{L}{2}+\frac{\delta}{2}} |x\rangle\langle x|\psi\rangle dx \\
&= \frac{1}{\sqrt{\langle\psi|\Pi_{slits}|\psi\rangle}} \int_{-\frac{L}{2}-\frac{\delta}{2}}^{-\frac{L}{2}+\frac{\delta}{2}} \psi(x)|x\rangle dx + \frac{1}{\sqrt{\langle\psi|\Pi_{slits}|\psi\rangle}} \int_{\frac{L}{2}-\frac{\delta}{2}}^{\frac{L}{2}+\frac{\delta}{2}} \psi(x)|x\rangle dx
\end{aligned}$$

Here we define two new states $|\psi_L\rangle, |\psi_R\rangle$ with wave functions:

$$\psi_L(x) = \begin{cases} \frac{1}{\alpha\sqrt{\langle\psi|\Pi_{slits}|\psi\rangle}}\psi(x) & x \in \left(-\frac{L}{2} \pm \frac{\delta}{2}\right) \\ 0 & x \notin \left(-\frac{L}{2} \pm \frac{\delta}{2}\right) \end{cases}, \quad \psi_R(x) = \begin{cases} \frac{1}{\beta\sqrt{\langle\psi|\Pi_{slits}|\psi\rangle}}\psi(x) & x \in \left(+\frac{L}{2} \pm \frac{\delta}{2}\right) \\ 0 & x \notin \left(+\frac{L}{2} \pm \frac{\delta}{2}\right) \end{cases}$$

[3] In order to get an idea of how the propagation of the wave will look like, you may want to take a look at this link.

where $\alpha, \beta \in \mathbb{C}$ are chosen in such a way that $|\psi_L\rangle, |\psi_R\rangle$ are valid states, i.e.

$$\int_{\mathbb{R}} |\psi_L(x)|^2 dx = \int_{\mathbb{R}} |\psi_R(x)|^2 dx = 1$$

Hence, we can rewrite the post-measurement state as:

$$|\psi^+\rangle = \alpha|\psi_L\rangle + \beta|\psi_R\rangle$$

From now on we will assume that the wave is symmetric with respect to the hyperplane $x = 0$ implying $|\alpha|^2 = |\beta|^2 = \frac{1}{2}$ by symmetry. Hence the only degrees of freedom that remain are given by the phases:

$$|\psi^+\rangle = \frac{e^{i\theta_a}}{\sqrt{2}}|\psi_L\rangle + \frac{e^{i\theta_b}}{\sqrt{2}}|\psi_R\rangle = e^{i\theta_a}\frac{|\psi_L\rangle + e^{i\theta_b - i\theta_a}|\psi_R\rangle}{\sqrt{2}}$$

Thus, if we let $\theta = \theta_b - \theta_a$ and remove the global phase, we obtain the equivalent state:

$$|\psi_\theta^+\rangle = \frac{|\psi_L\rangle + e^{i\theta}|\psi_R\rangle}{\sqrt{2}}$$

By the symmetry assumption, we also know that the two collapsed waves must be equal in front of the two slits, i.e.

$$\psi_L(x + L) = \psi_R(x)$$

Moreover, the two waves do not overlap in any point: one of them is zero for any x and, in particular we evince that

$$\psi_L^*(x)\psi_R(x) = 0 \quad \forall x \in \mathbb{R}$$

Before starting with the evolution, we make another important observation. If $\psi(x)$ was a wave packet, then also $\psi_L(x)$ and $\psi_R(x)$ are wave packets:

$$\psi_L(x) = \begin{cases} \frac{1}{\alpha\sqrt{\langle\psi|\Pi_{slits}|\psi\rangle}}\psi(x) & x \in \left(-\frac{L}{2} \pm \frac{\delta}{2}\right) \\ 0 & x \notin \left(-\frac{L}{2} \pm \frac{\delta}{2}\right) \end{cases} = \begin{cases} \frac{1}{\alpha\sqrt{\langle\psi|\Pi_{slits}|\psi\rangle}}\phi(x)e^{ikx} & x \in \left(-\frac{L}{2} \pm \frac{\delta}{2}\right) \\ 0 & x \notin \left(-\frac{L}{2} \pm \frac{\delta}{2}\right) \end{cases}$$

(the same applies to ψ_R). Also the envelope function is the same (up to truncations and multiplicative constants), therefore if we assumed to have a Gaussian wave packet, then we will still have a (truncated) Gaussian wave.

Now we are ready to let $|\psi_\theta^+\rangle$ evolve:

$$|\psi_\theta^+(t)\rangle = U(t)|\psi_\theta^+\rangle = \frac{1}{\sqrt{2}}\left(U(t)|\psi_L\rangle + e^{i\theta}U(t)|\psi_R\rangle\right)$$

which clearly shows that the relative phase θ has an important influence on the interference.

9.3 Evolution beyond the slits: stationary wave approximation

Consider a generic wave packet $\psi(x, t)$ over time, where:

$$\psi(x, 0) = \phi(x - x_0) \cdot e^{ip_0 x/\hbar}$$

such that $\phi(x)$ has zero mean. We will compute the evolution of ψ in the momentum basis, as anticipated in Section 5.7. By Theorem 7.3,

$$\bar{\psi}(p, 0) = \bar{\phi}(p - p_0) \cdot e^{-ipx_0/\hbar}$$

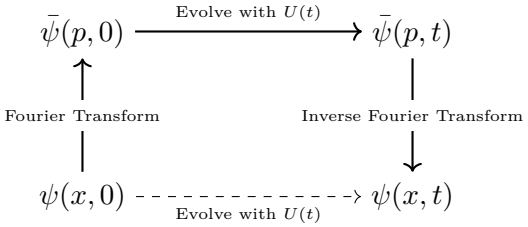


Figure 9.3: Equivalent ways of evolving a state.

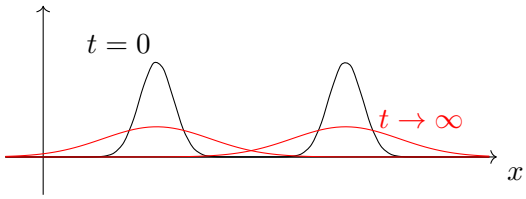


Figure 9.4: The wave function of the state as $t \to \infty$

where $\bar{\phi}(p)$ is the Fourier transform of $\phi(x)$. Beyond the slits, we have again a free particle evolution, thus:

$$H = \frac{P^2}{2\mu} \Longrightarrow U(t) = \exp\left(-\frac{iP^2}{2\mu\hbar}t\right)$$

Applying the definition of exponential, this gives us, at time t:

$$\bar{\psi}(p,t) = e^{-\frac{ip^2}{2\mu\hbar}t}\bar{\psi}(p,0)$$

which we will anti-transform:

$$\begin{aligned}\psi(x,t) &= \frac{1}{\sqrt{2\pi\hbar}}\int_{\mathbb{R}} \bar{\psi}(p,t)\cdot e^{ipx/\hbar}dp \\ &= \frac{1}{\sqrt{2\pi\hbar}}\int_{\mathbb{R}} \bar{\psi}(p,0)\cdot e^{-i\frac{p^2t}{2\mu\hbar}}e^{ipx/\hbar}dp \\ &= \frac{1}{\sqrt{2\pi\hbar}}\int_{\mathbb{R}} \bar{\psi}(p,0)\cdot \exp\left(\frac{i}{\hbar}\left(xp - \frac{p^2}{2\mu}t\right)\right)dp\end{aligned}$$

Figure 9.3 summarizes with bold lines the steps we have performed above. Now we "only" need to evaluate the integral.

Stationary wave approximation. In order to simplify our computation here, we will use an approximation: $t \to \infty$, i.e. we wait for a sufficiently large amount of time and the particle has moved far away ($x = \nu t$ for some constant ν). The evolution of the wave function in this case is depicted in Figure 9.4.

Now we use this assumption to approximate the integral above, working with the term ip^2 at the exponential. Take the Taylor expansion of the following expression:

$$f(p) = xp - \frac{p^2}{2\mu}t \simeq f(\mu\nu) + f'(\mu\nu)(p-\mu\nu) + \frac{1}{2}f''(\mu\nu)(p-\mu\nu)^2$$

where $\mu\nu = \mu x/t$ is a critical point of f (check it!). This is because, by the above assumption:

$$\frac{\partial}{\partial p}\left(xp - \frac{p^2}{2\mu}t\right) = t\left(\nu - \frac{p}{\mu}\right) \to 0 \quad \text{for } t \to \infty$$

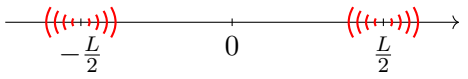


Figure 9.5: Evolution of the wave function along the x axis immediately after the slits.

which means that $p \to \mu\nu$: the wave tends to be stationary as time passes. Thus, the Taylor expansion evaluates to:

$$\begin{aligned} f(p) &\simeq f(\mu\nu) + f'(\mu\nu)(p-\mu\nu) + \frac{1}{2}f''(\mu\nu)(p-\mu\nu)^2 \\ &= x\mu\nu - \frac{\mu\nu^2}{2}t + (x-\nu t)(p-\mu\nu) + \frac{1}{2}\left(-\frac{t}{\mu}\right)(p-\mu\nu)^2 \\ &= \frac{\mu\nu^2}{2}t - \frac{t}{2\mu}(p-\mu\nu)^2 \end{aligned}$$

On the other hand, since $p \to \mu\nu$, we approximate the rest of the integrand by replacing this value in p.

$$\begin{aligned} \psi(\nu t, t) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \bar{\psi}(p,0) \cdot \exp\left(\frac{i}{\hbar}\left(xp - \frac{p^2}{2\mu}t\right)\right) dp \\ &\simeq \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \bar{\psi}(\mu\nu,0) \cdot \exp\left(\frac{i}{\hbar}\left(\frac{\mu\nu^2}{2}t - \frac{t}{2\mu}(p-\mu\nu)^2\right)\right) dp \\ &= \frac{1}{\sqrt{2\pi\hbar}} \cdot \bar{\psi}(\mu\nu,0) \cdot e^{i\mu\nu^2 t/2\hbar} \int_{\mathbb{R}} e^{-it(p-\mu\nu)^2/2\mu\hbar} dp \\ &= \frac{1}{\sqrt{2\pi\hbar}} \cdot \bar{\psi}(\mu\nu,0) \cdot e^{i\mu\nu^2 t/2\hbar} \sqrt{\frac{2\hbar\mu}{it}} \int_{\mathbb{R}} e^{-p^2} dp \\ &= \sqrt{-i}\sqrt{\frac{\mu}{t}} \cdot \bar{\psi}(\mu\nu,0) \cdot e^{i\mu\nu^2 t/2\hbar} \end{aligned}$$

where in the second-to-last step we did a substitution $p \leftarrow \sqrt{\frac{2\hbar\mu}{it}}(p+\mu\nu)$, and the remaining Gaussian integral evaluates to $\sqrt{\pi}$. What we found is that the position wave function ψ will tend to have the form of the momentum wave function in $p = \mu\nu = \mu x/t$ during the propagation in space. From now on we will exclude the global phase $\sqrt{-i}$.

Let us see now how we can apply this result to the double slit experiment: for simplicity we will only look at the x-axis as in Figure 9.5 (the evolution also propagates along the y-axis, but it is sufficient to analyze the interference immediately after the slits).

Assume to have the two wave packets on the x-axis, with zero expected momentum[4]:

$$\psi_L(x,0) = \phi\left(x + \frac{L}{2}\right), \quad \psi_R(x,0) = \phi\left(x - \frac{L}{2}\right)$$

Moreover, we consider the two wave packets to be Gaussian. They are enough to fix the idea, and at the same time we can use the neat properties we derived in Section 7.5:

$$\begin{aligned} \phi(x - x_0) &= \frac{1}{\sqrt[4]{2\pi\sigma^2}} \cdot e^{-\frac{(x-x_0)^2}{4\sigma^2}} \\ \bar{\psi}(p,0) = \bar{\phi}(p) \cdot e^{-ipx_0/\hbar} &= \sqrt[4]{\frac{2\sigma^2}{\pi\hbar^2}} \cdot e^{-\frac{\sigma^2 p^2}{\hbar^2}} \cdot e^{-ipx_0/\hbar} \end{aligned}$$

[4]This assumption on the expected momentum is not disturbing at all: in a symmetric double slit experiment it is natural to think that the momentum along the x-axis is zero in expectation (recall that the particle's initial expected momentum along the x-axis was 0).

The only thing that differs from the derivation of Section 7.5 is that here we directly transformed $\phi(x - x_0)$ and not $\phi(x)$, i.e. the envelope is not zero-mean, and the Fourier transform yields another factor $e^{-ipx_0/\hbar}$ for the shift (which will contribute to the frequency of the momentum wave packet, as we have already seen). Now we have everything we need to compute $\psi_L(x,t), \psi_R(x,t)$:

$$\begin{aligned}\psi_L(\nu t, t) &\simeq \sqrt{\frac{\mu}{t}} \cdot \bar{\phi}(\mu\nu) \cdot e^{i\mu\nu L/2\hbar} \cdot \exp\left(\frac{i\mu\nu}{2\hbar} \cdot \nu t\right) \\ &= \sqrt{\frac{\mu}{t}} \cdot \bar{\phi}(\mu\nu) \cdot \exp\left(\frac{i\mu\nu}{2\hbar} \cdot (\nu t + L)\right)\end{aligned}$$

$$\begin{aligned}\psi_R(\nu t, t) &\simeq \sqrt{\frac{\mu}{t}} \cdot \bar{\phi}(\mu\nu) \cdot e^{-i\mu\nu L/2\hbar} \cdot \exp\left(\frac{i\mu\nu}{2\hbar} \cdot \nu t\right) \\ &= \sqrt{\frac{\mu}{t}} \cdot \bar{\phi}(\mu\nu) \cdot \exp\left(\frac{i\mu\nu}{2\hbar} \cdot (\nu t - L)\right)\end{aligned}$$

And we can now see what happens to the wave function $\psi_\theta(x,t)$ of the state $|\psi_\theta^+\rangle$:

$$\begin{aligned}\psi_\theta(\nu t, t) &\simeq \frac{1}{\sqrt{2}}\left(\psi_L(\nu t, t) + e^{i\theta}\psi_R(\nu t, t)\right) \\ &= \frac{1}{\sqrt{2}} \cdot \sqrt{\frac{\mu}{t}} \cdot \bar{\phi}(\mu\nu)\left(\exp\left(\frac{i\mu\nu}{2\hbar} \cdot (\nu t + L)\right) + e^{i\theta} \cdot \exp\left(\frac{i\mu\nu}{2\hbar} \cdot (\nu t - L)\right)\right)\end{aligned}$$

Just to simplify the calculation, let B be everything out of the parentheses:

$$\begin{aligned}\psi_\theta(\nu t, t) &\simeq B \cdot \left(\exp\left(\frac{i\mu\nu}{2\hbar} \cdot (\nu t + L)\right) + e^{i\theta} \cdot \exp\left(\frac{i\mu\nu}{2\hbar} \cdot (\nu t - L)\right)\right) \\ &= Be^{i\theta/2} \cdot \left(e^{-i\theta/2} \exp\left(\frac{i\mu\nu}{2\hbar} \cdot (\nu t + L)\right) + e^{i\theta/2} \cdot \exp\left(\frac{i\mu\nu}{2\hbar} \cdot (\nu t - L)\right)\right) \\ &= Be^{i\theta/2} \exp\left(\frac{i\mu\nu^2 t}{2\hbar}\right) \cdot \left(e^{-i\theta/2} \exp\left(\frac{i\mu\nu L}{2\hbar}\right) + e^{i\theta/2} \cdot \exp\left(-\frac{i\mu\nu L}{2\hbar}\right)\right)\end{aligned}$$

We again simplify and call D everything out of the parentheses:

$$\begin{aligned}\psi_\theta(\nu t, t) &\simeq D \cdot \left(e^{-i\theta/2} \exp\left(\frac{i\mu\nu L}{2\hbar}\right) + e^{i\theta/2} \cdot \exp\left(-\frac{i\mu\nu L}{2\hbar}\right)\right) \\ &= D \cdot \left(\exp\left(i\left(\frac{\mu\nu L}{2\hbar} - \frac{\theta}{2}\right)\right) + \exp\left(-i\left(\frac{\mu\nu L}{2\hbar} - \frac{\theta}{2}\right)\right)\right) \\ &= 2D\cos\left(\frac{\mu\nu L}{2\hbar} - \frac{\theta}{2}\right)\end{aligned}$$

Let us analyze what we obtained, starting from the terms in $2D$:

$$2D = \sqrt[4]{\frac{8\sigma^2}{\pi^3\hbar^2}} \cdot \sqrt{\frac{\mu}{t}} \cdot \exp\left(-\frac{\sigma^2\mu^2\nu^2}{\pi^2}\right) \cdot \exp\left(i\left(\frac{\theta}{2} + \mu\frac{\nu^2 t}{2\hbar}\right)\right)$$

In order from left to right we have:

- A normalization term;
- A **dampening term**, which will become a factor $\frac{\mu}{t}$ in the probability distribution $|\phi_\theta|^2$;
- A **Gaussian term**;

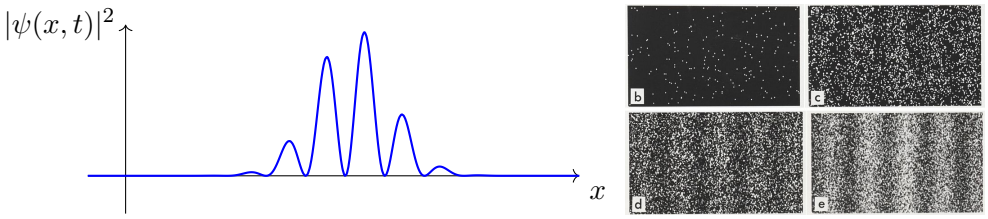


Figure 9.6: Plot of the squared modulus of the wave function approximation we computed (left) and experimental results (right, source).

- A **global phase**;

Computing the squared modulo of this we obtain:

$$|2D|^2 = \sqrt{\frac{8\sigma^2}{\pi^3\hbar^2}} \cdot \frac{\mu}{t} \cdot \exp\left(-\frac{2\sigma^2\mu^2\nu^2}{\hbar^2}\right)$$

Thus the total wave function becomes:

$$|\psi(\nu t,t)|^2 \simeq \sqrt{\frac{8\sigma^2}{\pi^3\hbar^2}} \cdot \frac{\mu}{t} \cdot \exp\left(-\frac{2\sigma^2\mu^2\nu^2}{\hbar^2}\right) \cdot \cos^2\left(\frac{\mu\nu L}{2\hbar} - \frac{\theta}{2}\right)$$

$$|\psi(x,t)|^2 \simeq \sqrt{\frac{8\sigma^2}{\pi^3\hbar^2}} \cdot \frac{\mu}{t} \cdot \exp\left(-\frac{2\sigma^2\mu^2}{\hbar^2 t^2}x^2\right) \cdot \cos^2\left(\frac{\mu L x}{2\hbar t} - \frac{\theta}{2}\right) \qquad \text{replacing back } \nu \leftarrow \frac{x}{t}$$

In Figure 9.6 we can clearly see the effect of double slit: when the cosine is low (and this fact heavily depends on θ, as one can see on the expression above), ψ_L, ψ_R add up with opposite phases, and thus we obtain a **destructive interference**.

9.4 Observing the relative phase

In this section we look for an observable that we can use to conveniently measure the relative phase θ between the slits, which would allow us to describe what the interference fringes look like. Let us see what happens with observables X^n and P^n for every n.

$$\begin{aligned}\langle X\rangle &= \langle\psi_\theta^+|X|\psi_\theta^+\rangle \\ &\left(\int_{\mathbb{R}} \psi_\theta^*(x)\langle x|dx\right)\left(\int_{\mathbb{R}} x|x\rangle\langle x|dx\right)\left(\int_{\mathbb{R}} \psi_\theta(x)|x\rangle dx\right) \\ &\int_{\mathbb{R}} x\psi_\theta^*(x)\psi_\theta(x)dx\end{aligned}$$

By definition of $|\psi_\theta^+\rangle$:

$$\psi_\theta(x) = \frac{\psi_L(x) + e^{i\theta}\psi_R(x)}{\sqrt{2}}$$

And we can replace it in the integral:

$$\begin{aligned}\langle X\rangle &= \int_{\mathbb{R}} x\psi_\theta^*(x)\psi_\theta(x)dx \\ &= \frac{1}{2}\int_{\mathbb{R}} x\left(\psi_L^*(x) + e^{-i\theta}\psi_R^*(x)\right)\left(\psi_L(x) + e^{i\theta}\psi_R(x)\right)dx\end{aligned}$$

$$= \frac{1}{2}\int_{\mathbb{R}} x\left(\psi_L^*(x)\psi_L(x) + \psi_R^*(x)\psi_R(x) + e^{-i\theta}\psi_R^*(x)\psi_L(x) + e^{i\theta}\psi_L^*(x)\psi_R(x)\right)dx$$
$$= \frac{1}{2}\int_{\mathbb{R}} x\left(|\psi_L(x)|^2 + |\psi_R(x)|^2\right)dx$$

where the last step follows from the fact that $\psi_L^*\psi_R = \psi_R^*\psi_L \equiv 0$. We found that the position of a particle passing through the slits does not depend on the relative phase θ between $|\psi_L\rangle, |\psi_R\rangle$, i.e. no measurement of the position can distinguish different values of θ. This argument immediately extends to X^n as it only changes the label of the measurement.

Let us see if observing the momentum can help us: remembering the transformation of the position wave function with P we derived in Section 5.2 we can write

$$\begin{aligned}\langle P\rangle &= \langle\psi_\theta^+|P|\psi_\theta^+\rangle \\ &= \left(\int_{\mathbb{R}}\psi_\theta^*(x)\langle x|dx\right)\left(\int_{\mathbb{R}} -i\hbar\frac{\partial}{\partial x}\psi_\theta(x)|x\rangle dx\right) \\ &= -i\hbar\int_{\mathbb{R}}\psi_\theta^*(x)\frac{\partial}{\partial x}\psi_\theta(x)dx \\ &= -\frac{i\hbar}{2}\int_{\mathbb{R}}\left(\psi_L^*(x) + e^{-i\theta}\psi_R^*(x)\right)\frac{\partial}{\partial x}\left(\psi_L(x) + e^{i\theta}\psi_R(x)\right)dx \\ &= -\frac{i\hbar}{2}\int_{\mathbb{R}}\left(\psi_L^*(x)\frac{\partial}{\partial x}\psi_L(x) + \psi_R^*(x)\frac{\partial}{\partial x}\psi_R(x)\right)dx\end{aligned}$$

Notice that the other two terms that show up in the product on the last equality here vanish, since:

$$\psi_L(x) \neq 0 \Longrightarrow \psi_R(x) = 0 \Longrightarrow \frac{\partial}{\partial x}\psi_R(x) = 0$$

i.e. $\psi_R(x)$ is constant in the region when $\psi_L(x) \neq 0$ and vice versa.

The above integral can be rewritten as (check it!):

$$\langle P\rangle = \frac{1}{2}\langle\psi_L|P|\psi_L\rangle + \frac{1}{2}\langle\psi_R|P|\psi_R\rangle$$

therefore, for the exact same reason as X, also P (and thus P^n for every n) cannot distinguish different values of θ.

Our next try involves the **modular momentum** operator. This seems to be a good idea, since we know what this operator does to the wave function from Section 8.3:

$$e^{iLP/\hbar} : \psi(x) \to \psi(x + L)$$

Thus its expectation will be:

$$\begin{aligned}\langle e^{iLP/\hbar}\rangle &= \int_{\mathbb{R}}\psi^*(x)e^{iLP/\hbar}\psi(x)dx \\ &= \int_{\mathbb{R}}\psi^*(x)\psi(x+L)dx \\ &= \frac{1}{2}\int_{\mathbb{R}}(\psi_L^*(x) + e^{-i\theta}\psi_R^*(x))(\psi_L(x+L) + e^{i\theta}\psi_R(x+L))dx\end{aligned}$$

Now, from the fact that $\psi_R(x + L) = 0$ (there is nothing at the right of the right slit) and $\psi_L(x + L) = \psi_R(x)$ (by symmetry of the slits), we obtain:

$$\langle e^{iLP/\hbar}\rangle = \frac{1}{2}\int_{\mathbb{R}}(\psi_L^*(x) + e^{-i\theta}\psi_R^*(x))(\psi_L(x+L) + e^{i\theta}\psi_R(x+L))dx$$

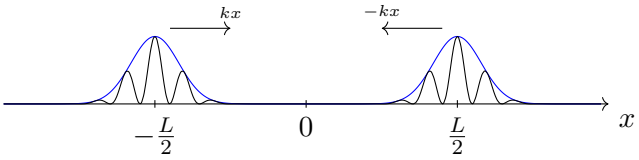


Figure 9.7: Two one-dimensional waves moving towards each other with the same momentum.

$$
\begin{aligned}
&= \frac{1}{2}\int_{\mathbb{R}} (\psi_L^*(x) + e^{-i\theta}\psi_R^*(x))\psi_R(x)dx \\
&= \frac{1}{2}\int_{\mathbb{R}} e^{-i\theta}|\psi_R(x)|^2 dx \\
&= \frac{1}{2}e^{-i\theta}
\end{aligned}
$$

Actually, we will not use this observable directly, but the following will be more useful:

$$\cos\left(\frac{pL}{\hbar}\right) = \frac{e^{iPL/\hbar} + e^{-iPL/\hbar}}{2}$$

and one can check that, by linearity of expectation we can still derive information about the phase:

$$\langle \cos\left(\frac{pL}{\hbar}\right)\rangle = \frac{e^{-i\theta} + e^{i\theta}}{4} = \frac{\cos(\theta)}{2}$$

A more in-depth discussion of this method to analyse interference is given in Sandu Popescu's 2019 lecture available on YouTube.

9.5 Evolution beyond the slits: constant variance approximation

We will now look at a one-dimensional approximation of the experiment, in which the wave packets, instead of moving on the two-dimensional plane, only move towards each other.

Let us go back at the instant where the particle passes through the slits, and suppose to have two wave packets centered on the two slits, moving towards the center with the same momentum, as in Figure 9.7:

$$
\begin{aligned}
\psi_L(x,0) &= \phi\left(x + \frac{L}{2}\right)\cdot e^{ikx} \\
\psi_R(x,0) &= \phi\left(x - \frac{L}{2}\right)\cdot e^{-ikx}
\end{aligned}
$$

The total state $|\psi_\theta^+\rangle$ has, as before, wave function:

$$\psi_\theta(x,t) = \frac{1}{\sqrt{2}}\Big(\psi_L(x,t) + e^{i\theta}\psi_R(x,t)\Big)$$

Here, we will do another approximation: the idea is based on the fact that, assuming $t \to 0$ and the mass μ very large, we can approximate the variance of the position.

$$\Big(\Delta X^2\Big)_t = \frac{\Delta P^2}{\mu^2}t^2 + \Big(\Delta X^2\Big)_0 \simeq \Big(\Delta X^2\Big)_0$$

i.e. the variance of the position will not change much during the evolution around $t \simeq 0$.

Our calculation begins with a simple but powerful trick on the Hamiltonian, where $p_0 = \hbar k$:

$$H = \frac{P^2}{2\mu} = \frac{(p_0\mathbb{1} + P - p_0 identity)^2}{2\mu}$$

We call $P' = P - p_0\mathbb{1}$. Let us plug this replacement in the Hamiltonian:

$$\begin{aligned} H &= \frac{P^2}{2\mu} \\ &= \frac{(p_0\mathbb{1} + P')^2}{2\mu} \\ &= \frac{p_0^2\mathbb{1}^2}{2\mu} + \frac{p_0 P'\mathbb{1}}{2\mu} + \frac{p_0\mathbb{1}P'}{2\mu} + \frac{(P')^2}{2\mu} \\ &= \frac{p_0^2\mathbb{1}}{2\mu} + \frac{p_0 P'}{\mu} + \frac{(P')^2}{2\mu} \end{aligned}$$

Since $t \simeq 0$, P is sharply concentrated around p_0, thus $P' \simeq 0$, and we can neglect the last term.

$$\begin{aligned} H &\simeq \frac{p_0^2\mathbb{1}}{2\mu} + \frac{p_0 P'}{\mu} \\ &= \frac{p_0^2\mathbb{1}}{2\mu} + \frac{p_0(P - p_0\mathbb{1})}{\mu} \\ &= \frac{p_0 P}{\mu} - \frac{p_0^2\mathbb{1}}{2\mu} \end{aligned}$$

We use $v_0 = \frac{p_0}{\mu}$ to denote the **group velocity** (as introduced in Section 7.6). The second term, on the other hand, is just a constant, and it can only contribute to the global phase since:

$$\begin{aligned} U(t) &= \exp\left(-\frac{it}{\hbar} \cdot \frac{p_0}{\mu} P\right) \cdot \exp\left(\frac{it}{\hbar} \cdot \frac{p_0^2}{2\mu}\mathbb{1}\right) \\ &= \exp\left(-\frac{it}{\hbar} \cdot \frac{p_0}{\mu} P\right) \cdot \exp\left(\frac{it}{\hbar} \cdot \frac{p_0^2}{2\mu}\right) \end{aligned}$$

Notice that we could use $e^{A+B} = e^A e^B$ because $\mathbb{1}$ and P commute. Thus, this constant term is irrelevant for our purposes, and we can work with the approximate Hamiltonian:

$$H' = v_0 P$$

Something particularly convenient for us happens on the evolution of a wave function $\psi(x,t)$:

$$\psi(x,t) = \exp\left(-\frac{it}{\hbar} v_0 P\right) \cdot \psi(x,0)$$

i.e. we found a modular momentum operator in the evolution. This means that

$$\psi(x,t) = \psi(x - v_0 t, 0)$$

which is perfectly coherent with our approximation: we neglected the change in the variance ΔX^2 over time, while the shift can also be seen as a direct consequence of Ehrenfest's theorem I, as the shift acts on the expectation. We apply this evolution to ψ_L, ψ_R now: recall that $v_0^L = \frac{\hbar k}{\mu} =: v_0$ and $v_0^R = -\frac{\hbar k}{\mu} = -v_0$, hence

$$\psi_L(x,t) = \psi_L(x - v_0 t, 0) = \phi\left(x - t v_0 + \frac{L}{2}\right) \cdot e^{ik(x - v_0 t)}$$

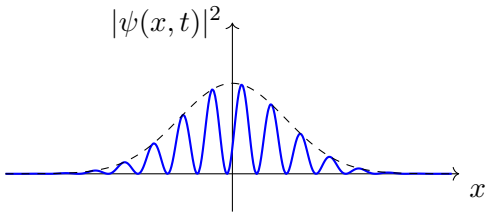


Figure 9.8: Plot of the probability density function of the position of the particle when using constant variance approximation.

$$\psi_R(x,t) = \psi_R(x+v_0t,0) = \phi\left(x+tv_0-\frac{L}{2}\right)\cdot e^{-ik(x+v_0t)}$$

Putting it all together in $\psi_\theta(x,t)$ we obtain:

$$\begin{aligned}
\psi_\theta(x,t) &= \frac{1}{\sqrt{2}}\left(\psi_L(x,t)+e^{i\theta}\psi_R(x,t)\right)\\
&= \frac{1}{\sqrt{2}}\left(\phi\left(x-tv_0+\frac{L}{2}\right)\cdot e^{ik(x-v_0t)}+\phi\left(x+tv_0-\frac{L}{2}\right)\cdot e^{i\theta}\cdot e^{-ik(x+v_0t)}\right)\\
&= \frac{1}{\sqrt{2}}e^{i\theta/2}\left(\phi\left(x-tv_0+\frac{L}{2}\right)\cdot e^{-i\theta/2}\cdot e^{ik(x-v_0t)}+\phi\left(x+tv_0-\frac{L}{2}\right)\cdot e^{i\theta/2}\cdot e^{-ik(x+v_0t)}\right)
\end{aligned}$$

Now we would like to take the instant t where the two waves ψ_L, ψ_R have the same expected value, which means (since $\phi(x-x_0)$ has expectation x_0)

$$tv_0-\frac{L}{2} = -tv_0+\frac{L}{2} \iff tv_0 = \frac{L}{2} \iff t = \frac{L}{2v_0}$$

By plugging this t into ψ_θ, the envelopes are equal and can be factored out of the superposition:

$$\begin{aligned}
\psi_\theta\left(x,\frac{L}{2v_0}\right) &= \frac{1}{\sqrt{2}}e^{i\theta/2}\phi(x)\cdot\left(e^{-i\theta/2}\cdot e^{ik(x-\frac{L}{2})}+e^{i\theta/2}\cdot e^{-ik(x+\frac{L}{2})}\right)\\
&= \frac{1}{\sqrt{2}}e^{i\theta/2}\phi(x)\cdot\left(\exp\left(-i\left(kx-k\frac{L}{2}-\frac{\theta}{2}\right)\right)+\exp\left(i\left(kx-k\frac{L}{2}-\frac{\theta}{2}\right)\right)\right)\\
&= \sqrt{2}\cdot e^{i\theta/2}\phi(x)\cdot\cos\left(kx-k\frac{L}{2}-\frac{\theta}{2}\right)
\end{aligned}$$

And the probability distribution at this point follows:

$$\left|\psi_\theta\left(x,\frac{L}{2v_0}\right)\right|^2 = 2|\phi(x)|^2\cdot\cos^2\left(kx-k\frac{L}{2}-\frac{\theta}{2}\right)$$

Again, depending on the cosine we either have destructive or constructive interference, as we observe in Figure 9.8. Observe that in this case we can also directly acquire information about the phase θ by looking at the distance between the y-axis and the first peak, which is $\frac{\theta}{2k}+\frac{L}{2}$.

9.6 Measuring the path of the particle

In Figure 9.9 we consider again the two slits and only dimension x: what if we wanted to know which slit a particle passes through? An idea is to place a **pointer particle** near the right slit that **reacts** in some way, e.g. a particle which gets deflected due to the electric field of a particle passing through the right slit.

In this section we will consider a toy example: this particle on the right slit can be represented as a qubit. Thus, we have a product space:

$$\mathcal{H} = \mathcal{H}_x \otimes \mathcal{H}_S$$

and we consider the following model for interaction:

$$U = \Pi_L \otimes \mathbb{1}_S + \Pi_R \otimes X_S = \left(\int_{-\infty}^{0} |x\rangle\langle x| dx\right) \otimes \mathbb{1}_S + \left(\int_{0}^{+\infty} |x\rangle\langle x| dx\right) \otimes X_S$$

i.e. if the particle passes through the left slit, nothing happens to the qubit, while passing through the right slit causes a flip of the qubit with the Pauli matrix $X = |0\rangle\langle 1| + |1\rangle\langle 0|$. Indeed, what we created here is nothing more than a CNOT gate (which we have seen in Section 3.2), if you think of the two slits as $|0\rangle$ and $|1\rangle$. We are implementing the trick of using projector operators in the tensor product as "selectors" for the subspace on which we want to apply the operator on the qubit:

$$\begin{aligned}
\Pi_L|\psi_L\rangle &= |\psi_L\rangle \\
\Pi_R|\psi_R\rangle &= |\psi_R\rangle \\
\Pi_L|\psi_R\rangle &= \Pi_R|\psi_L\rangle = 0
\end{aligned}$$

Suppose our qubit starts in state $|0\rangle_S$, and our particle passes through the slits in the usual superposition, yielding a total state:

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|\psi_L\rangle_x + e^{i\theta}|\psi_R\rangle\right) \otimes |0\rangle_S = \frac{1}{\sqrt{2}}\left(|\psi_L\rangle_x|0\rangle_S + e^{i\theta}|\psi_R\rangle|0\rangle_S\right)$$

Then, applying U to this state we obtain:

$$\begin{aligned}
|\psi'\rangle &= U|\psi\rangle \\
&= U\frac{1}{\sqrt{2}}\left(|\psi_L\rangle_x|0\rangle_S + e^{i\theta}|\psi_R\rangle|0\rangle_S\right) \\
&= \frac{1}{\sqrt{2}}\left(U|\psi_L\rangle_x|0\rangle_S + e^{i\theta}U|\psi_R\rangle_x|0\rangle_S\right) \\
&= \frac{1}{\sqrt{2}}\left(|\psi_L\rangle_x \otimes (\mathbb{1}_S|0\rangle_S) + e^{i\theta}|\psi_R\rangle \otimes (X_S|0\rangle_S)\right) \\
&= \frac{1}{\sqrt{2}}\left(|\psi_L\rangle_x|0\rangle_S + e^{i\theta}|\psi_R\rangle_x|1\rangle_S\right)
\end{aligned}$$

Suppose now we measure the qubit with Z (i.e. with respect to the computational basis). Looking at the induced probability space:

$$\begin{aligned}
\mathbf{P}(1) &= \langle\psi'|\left(\mathbb{1}_x \otimes |1\rangle\langle 1|\right)|\psi'\rangle \\
&= \frac{1}{2}e^{-i\theta}e^{i\theta}\langle\psi_R|\mathbb{1}_x|\psi_R\rangle\langle 1|(|1\rangle\langle 1|)|1\rangle = \frac{1}{2}
\end{aligned}$$

Figure 9.9: A pointer particle is placed on one of the two slits.

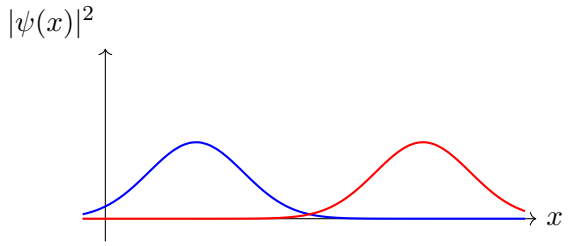


Figure 9.10: The blue wave function is associated with the measurement of outcome $|0\rangle$ on the qubit, whereas the red one corresponds to state $|1\rangle$ on the quibit.

and the post-measurement state is:

$$\frac{(\Pi_R \otimes |1\rangle\langle 1|)|\psi'\rangle}{\sqrt{\langle\psi'|\left(\mathbb{1}_x \otimes |1\rangle\langle 1|\right)|\psi'\rangle}} = e^{i\theta}|\psi_R\rangle_x \otimes |1\rangle_S$$

from which we can remove the global phase. If we evolve it now, we will only see the right wave. In general, upon measurement of the qubit we will only see the wave of one of the two slits. This will thus also remove all the observed interference, as pointed out in Figure 9.10.

9.7 Application of interference

We conclude this chapter by providing some pointers to interesting applications of what we have discussed so far.

The fields of application of interference include:

- The construction of **interferometers to detect gravitational waves**. For example, you can check out this Wikipedia page;
- **Crystallography**: when a crystal is exposed to a laser beam, the interference pattern is analysed to find the molecule structure of said crystal (Wikipedia); There are groups at ETH Zurich researching this matter.
- Analysis of the structure of uneven surfaces.

If you are interested in the topic, you can check also check out the Quantum Mechanics 2 course offered by the D-PHYS or the courses offered by the Crystallography Group of the D-MATL

Chapter 10
Stationary States in Potential

These lectures, in particular starting from Section 10.3, closely follow Chapter 15 (*Stationary states in 1-D*) of Schumacher and Westmoreland [1]. For the *Particle in a box setting* (Section 10.2) we refer you to Chapter 11, Section 5, of the same book.

10.1 Particle in a potential

Recall from Section 5.5 the definition of general Hamiltonian:

$$H = \frac{P^2}{2\mu} + V(X)$$

Up to now, we always considered a free particle, i.e. $V(x) \equiv 0$. In this chapter we will see how this term can influence the wave function of the state of a particle. We will again reduce to the one-dimensional case for simplicity, but the extension to multi-dimensional cases are immediate.

The first example of non-zero potential we want to explore are **potential wells**:

$$V(x) = \begin{cases} 0 & 0 \leq x \leq L \\ V & \text{otherwise} \end{cases}$$

We could also have multiple wells next to each other, and various configurations of wave functions for the particle inside and outside the well, as depicted in Figure 10.1.

Quantum engineering can actually build these types of wells, using **quantum dots** or **cavities**. Double wells in particular can be useful in quantum memory implementations (i.e. storing a qubit), where we can see the two wells as states $|0\rangle$ and $|1\rangle$, according to the position of the particle in one of the two wells.

You may recall from classical physics and/or calculus classes, that the potential of a conservative field is not unique. If $U(\mathbf{r})$ is a potential for a vector field $\mathbf{F}(\mathbf{r})$, then also $U + C$ is a valid potential, for any constant C. However, the choice of the constant in the potential causes an additive constant factor in the Hamiltonian, which in turn only adds an irrelevant global phase term $e^{itC/\hbar}$ on the evolution operator.

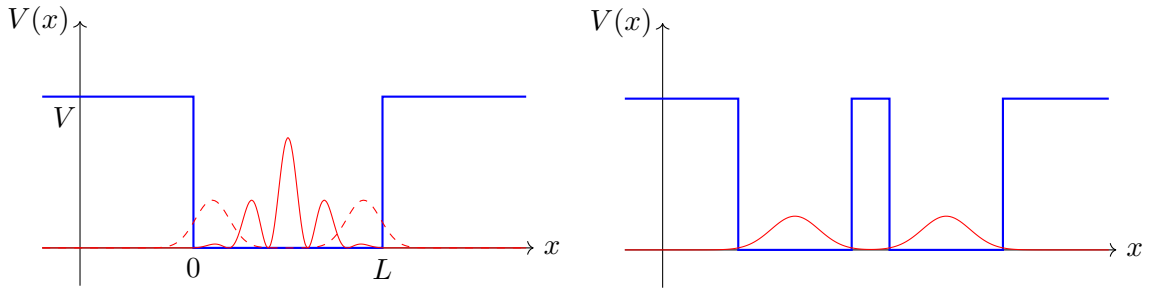


Figure 10.1: Two possible configurations for potential wells (single and double).

10.2 Warm up: particle in a box

Consider a well of the form:

$$V(x) = \begin{cases} 0 & 0 \leq x \leq L \\ \infty & \text{otherwise} \end{cases}$$

i.e. going from $[0, L]$ to outside the "box" requires infinite energy. Another way of saying this, used also in classical physics, is that there are two **potential barriers** located at $x = 0, L$ for particles within these two points, because the energy of the particle is not enough to pass those points. This concept can be extended to the case in which the potential outside of the box is finite, but in this case the box endpoints are not potential barriers for a particle with sufficiently high energy. For this reason it is convenient to use the infinite case as a warm up. Keep in mind, however, that an infinite potential is un-physical, and the setting we are considering is only ideal.

With this in mind, we can do the following observation for the state $|\psi\rangle$ of a particle:

$$V(x) \to \infty \implies \psi(x) \to 0$$

as the particle needs a high amount of energy in order to stay in zones with higher potential (remember that force fields always tend to push from high to low potential). In particular, $\psi(0) = \psi(L) = 0$ by **boundary conditions** (wave functions must be continuous). How do we compute the wave function *within* the box?

Looking for stationary states. Recall that stationary states are eigenstates of the Hamiltonian:

$$H|\psi_E\rangle = E|\psi_E\rangle$$

where $E = \hbar\omega_E$ is the energy eigenvalue of $|\psi_E\rangle$. Moreover, the evolution of an eigenstate is as follows:

$$|\psi_E(t)\rangle = e^{-iHt/\hbar}|\psi_E\rangle = e^{-i\omega_E t}|\psi_E\rangle$$

We go ahead with the following educated guess for the form of an eigenstate $|\psi\rangle$:

$$\psi(x) = A\sin(kx)$$

Indeed, the Schrödinger equation will give a second order differential equation with solutions that are either real or complex exponentials (we will see this in detail later), and the only solutions that can have $\psi(0) = \psi(L) = 0$ are of this form. Given our guess, $\psi(0) = 0$ holds "for free". Applying the second boundary condition we obtain:

$$\psi(L) = 0 \quad \implies \quad A\sin(kL) = 0 \quad \implies \quad k_n = \frac{n\pi}{L},\ n \in \mathbb{N}$$

Hence, we can define ψ_n as the wave with n peaks:

$$\psi_n(x) = A\sin(k_n x)$$

On the other hand, since the potential is 0 within the box, we can use the free particle Hamiltonian:

$$\langle x|\frac{P^2}{2\mu}|\psi\rangle = \frac{-\hbar^2}{2\mu}\frac{\partial^2}{\partial x^2}\psi_n(x) \qquad \text{Recall from Section 5.7}$$

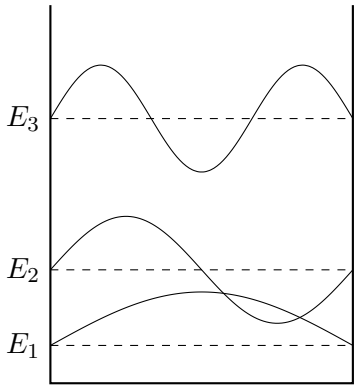


Figure 10.2: Wave functions of stationary states of a particle in a box for the first three energy levels. Observe how the value of the n-th energy level grows quadratically in n.

$$\begin{aligned}
&= \frac{-\hbar^2}{2\mu}\frac{\partial^2}{\partial x^2}(A\sin(k_n x)) \\
&= \frac{\hbar^2 k_n^2}{2\mu} A\sin(k_n x) \\
&= \frac{\hbar^2 k_n^2}{2\mu}\psi_n(x)
\end{aligned}$$

Thus, our guess was correct, and $\{|\psi_n\rangle\}_n$ are stationary states with eigenvalues

$$E_n = \frac{k_n^2\hbar^2}{2\mu} = \frac{\hbar^2}{2\mu}\frac{\pi^2}{L^2}n^2 = E_1 n^2$$

If we evolve the stationary states over time (as we have first seen in Section 4.2):

$$|\psi_n(t)\rangle = e^{-itH/\hbar}|\psi_n\rangle = e^{-i\omega_1 n^2 t}|\psi_n\rangle$$

and we obtain $\omega_n = \omega_1 n^2 = \frac{E_1}{\hbar}n^2$. This gives us the energy level of every stationary state. It follows that the wave functions of the stationary states have the shape depicted in Figure 10.2.

The only degree of freedom we are left with is A, and one can imagine that this is constrained by normalization:

$$\begin{aligned}
1 &= \int_{\mathbb{R}} |\psi_n(x)|^2 dx \\
&= |A|^2 \int_0^L \sin^2(k_n x)dx \\
&= |A|^2 \int_0^L \left(\frac{1-\cos(2k_n x)}{2}\right) dx && \text{by prosthaphaeresis} \\
&= \frac{|A|^2}{2}\left[x - \frac{\sin(2k_n x)}{2k_n}\right]_0^L \\
&= \frac{|A|^2 L}{2}
\end{aligned}$$

Implying $A = \sqrt{\frac{2}{L}}$ (any phase would be global, so let us assume it to be real). Putting all together we have that the stationary states of the Hamiltonian in the box are:

$$\psi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{\pi n x}{L}\right)$$

The energy levels of the stationary states in the well are not continuous but discrete, i.e. divided into quanta, and so is their emission spectrum.

Time evolution of non-stationary states. Recall that, since the Hamiltonian operator is Hermitian, the stationary states form an orthonormal basis of the state space, i.e. any state can be represented as linear combination of stationary states. For example, if we have $|\psi(0)\rangle = \alpha|\psi_\ell\rangle + \beta|\psi_m\rangle$, where $|\psi_m\rangle, |\psi_\ell\rangle$ are stationary states, an evolution over time yields:

$$|\psi(t)\rangle = \alpha e^{-i\omega_1 t\ell^2}|\psi_\ell\rangle + \beta e^{-i\omega_1 t m^2}|\psi_m\rangle$$

In the special case $\alpha = \beta = \frac{1}{\sqrt{2}}$, and assuming $m^2 - \ell^2 = \Delta \geq 0$, we obtain

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}} e^{-i\omega_1 t\ell^2}\left(|\psi_\ell\rangle + e^{-i\omega_1 \Delta\cdot t}|\psi_m\rangle\right)$$

We can plug in the wave function of the stationary states to derive the wave function of $|\psi(t)\rangle$:

$$\begin{aligned}
\psi(x,t) &= \frac{1}{\sqrt{L}} e^{-i\omega_1 \ell^2 t}\left(\sin\left(\frac{\pi\ell x}{L}\right) + e^{-i\omega_1\Delta t}\sin\left(\frac{\pi m x}{L}\right)\right) \\
|\psi(x,t)|^2 &= \frac{1}{L}\left(\sin^2\left(\frac{\pi\ell x}{L}\right) + \sin^2\left(\frac{\pi m x}{L}\right) + \left(e^{-i\omega_1\Delta t} + e^{i\omega_1\Delta t}\right)\sin\left(\frac{\pi m x}{L}\right)\cdot\sin\left(\frac{\pi\ell x}{L}\right)\right) \\
&= \frac{1}{L}\left(\sin^2\left(\frac{\pi\ell x}{L}\right) + \sin^2\left(\frac{\pi m x}{L}\right) + 2\cos\left(\omega_1\Delta t\right)\cdot\sin\left(\frac{\pi m x}{L}\right)\cdot\sin\left(\frac{\pi\ell x}{L}\right)\right)
\end{aligned}$$

One can notice that this state is a sort of mean between $|\psi_\ell(x)|^2$ and $|\psi_m(x)|^2$, and the cosine (which is the only part dependent on time) makes the values oscillate between these two peaks. This resembles the behaviour in classical physics: imagine a ball moving without friction and bouncing with elastic collision on the walls of the box. Figure 10.3 shows how the wave function evolves in the case of $l = 1$ and $m = 1$.

10.3 Interaction zone and boundary conditions

Let us now move to a slightly more general case. We take the general Hamiltonian:

$$H = \frac{P^2}{2\mu} + V(X)$$

Again, we want to find the stationary states, since we know we will always be able to reconstruct the evolution of an arbitrary state as a superposition of stationary states. Therefore, let us write down the time-independent Schrödinger equation, where E is the energy eigenvalue of $|\psi_E\rangle$:

$$\left(\frac{P^2}{2\mu} + V(X)\right)|\psi_E\rangle = E|\psi_E\rangle$$

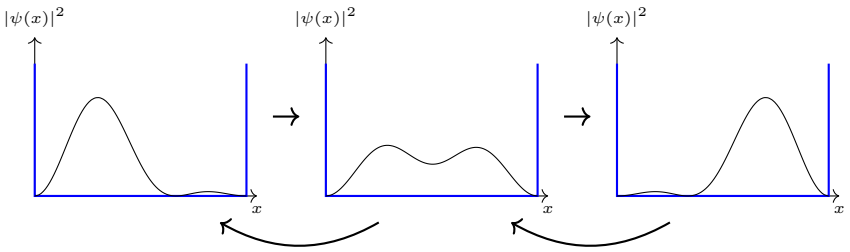


Figure 10.3: Evolution of a non-stationary wave with $l = m = 1$.

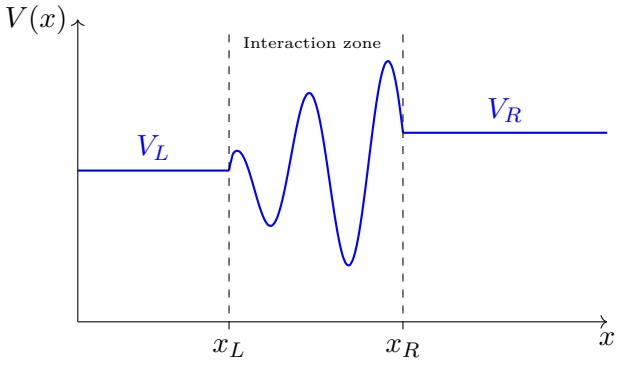


Figure 10.4: We divide the axis into three zones based on the value of the potential.

$$\begin{aligned}
\frac{P^2}{2\mu}|\psi_E\rangle &= (E\mathbb{1} - V(X))|\psi_E\rangle \\
\langle x|\frac{P^2}{2\mu}|\psi_E\rangle &= \langle x|(E\mathbb{1} - V(X))|\psi_E\rangle \\
-\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial x^2}\psi_E(x) &= (E - V(x))\psi_E(x) \\
\frac{\partial^2}{\partial x^2}\psi_E(x) &= \frac{2\mu}{\hbar^2}(V(x) - E)\psi_E(x)
\end{aligned}$$

In order to continue the analysis (i.e. solve the differential equation above) we will divide the x-axis into "easy" zones: two tail zones, where we will assume the potential to be constant (there is no force after a certain point), and a middle interval called **interaction zone**, where the potential may behave in an arbitrary way. The idea is visualized in Figure 10.4.

We will solve the Schrödinger equation separately for each zone and then impose continuity conditions:

- $\psi_E(x)$ must be continuous: this is important because the induced probability space denotes an absolutely continuous probability distribution. If $|\psi_E(x)|^2$ has discontinuities, the distribution is ill-formed.
- $\psi'_E(x) = \frac{\partial}{\partial x}\psi_E(x)$ is continuous when $V(x)$ is finite. By integrating the Schrödinger equation once we see that the first derivative is:

  $$\psi'_E(x) = \psi'_E(x_0) + \frac{2\mu}{\hbar^2}\int_{x_0}^{x}(V(x) - E)\psi_E(x)dx$$

  which is continuous by construction. We already saw that this condition may fail in the case of the particle in a box.
- $\psi''_E(x)$ is continuous when $V(x)$ is continuous; this can be seen directly from the Schrödinger equation.

Other conditions imposed by physics are:

- $\psi_E(x)$ must be normalized. In particular, this rules out **divergent** solutions.
- $\langle H\rangle \geq V_{min} =: \min_x V(x)$, where $\langle H\rangle$ is the expected total energy. This because

  $$\langle H\rangle = \frac{1}{2\mu}\langle P^2\rangle + \langle V(X)\rangle \qquad \text{by linearity of expectation}$$

$$
\begin{aligned}
&\geq \langle V(X) \rangle && \text{since } P^2 \text{ always yields a non-negative value} \\
&\geq \min_x V(x) && \text{since expectation} \geq \text{minimum}
\end{aligned}
$$

Solving the equation in zones of constant potential. Consider a zone with constant potential V_0. We can have two cases: either $E > V_0$ or $E < V_0$.

- If $E > V_0$, then the Schrödinger equation becomes, for $d = E - V_0 > 0$,

  $$\frac{\partial^2}{\partial x^2}\psi_E(x) = -\frac{2\mu}{\hbar^2} d \cdot \psi_E(x)$$

  which induces a **complex exponential** on the wave function:

  $$\psi_E(x) = Ae^{-ikx} + Be^{ikx}$$

  for constants A, B and $k = \sqrt{\frac{2\mu}{\hbar^2}(E - V_0)}$.

- If $E < V_0$, then the energy eigenvalue of the state is not enough for the potential, and this results in a potential barrier, as we have already seen. In fact, we have, for $d = V_0 - E > 0$

  $$\frac{\partial^2}{\partial x^2}\psi_E(x) = \frac{2\mu}{\hbar^2} d \cdot \psi_E(x)$$

  whose solution is a **real exponential**:

  $$\psi_E(x) = Ce^{-bx} + De^{bx}$$

  again, for constants C, D and $b = \sqrt{\frac{2\mu}{\hbar^2}(V_0 - E)}$. Indeed, we will see in a bit that only negative exponentials (i.e. rapidly dropping to zero) can actually satisfy the normalization constraint.

10.4 Reflecting, scattering and bound states

We now move on to a more concrete example. Suppose to have a bounded interaction zone:

$$V(x) = \begin{cases} V_L & x < x_L \\ v(x) & x_L \leq x \leq x_R \\ V_R & x > x_R \end{cases}$$

We consider a stationary state $|\psi_E\rangle$ with energy eigenvalue E in the following cases. First of all, notice that E is exactly the expected total energy of $|\psi_E\rangle$:

$$\langle H \rangle = \langle \psi_E | H | \psi_E \rangle = E \langle \psi_E | \psi_E \rangle = E$$

In Figure 10.5 we expand the diagram introduced in Figure 10.4 with what we know about the shapes of the wave function in the constant potential zones.

Observe now that we can have various combinations for the values of V_L and V_R with respect to the energy level of the stationary state E:

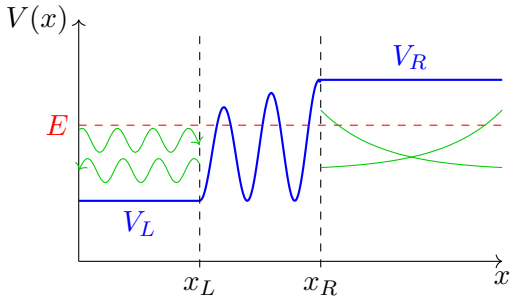


Figure 10.5: We divide the x-axis int zones with different wave function shapes according to the energy value of the state.

- Suppose $V_L < E$ but $V_R > E$:

  By what we have seen in the previous section, we will have two complex exponential terms on the left tail, and two real exponentials on the right tail:

  $$\psi_E(x) = \begin{cases} Ae^{-ikx} + Be^{ikx} & x < x_L \\ Ce^{-bx} + De^{bx} & x > x_R \end{cases}$$

  If $D \neq 0$, then the positive exponential will continue to grow all the way to infinity, and the integral

  $$\int_{x_R}^{+\infty} |\psi_E(x)|^2 dx$$

  diverges. This case will surely not satisfy the normalization constraint. So we deduce $D = 0$:

  $$\psi_E(x) = \begin{cases} Ae^{-ikx} + Be^{ikx} & x < x_L \\ Ce^{-bx} & x > x_R \end{cases}$$

  What happens in the interaction zone depends on the behaviour of the potential and is not of our interest here. However, once we compute the wave function in this interval, we can "glue" the three parts together using boundary conditions. A state with $V_L < E < V_R$ (or the other way around, for which the analysis is analogous) is called **reflecting state**. This name comes from the fact that, in e^{ikx}, e^{-ikx}, $\pm\hbar k$ is the expected momentum of the particle (recall what we saw for wave packets in Chapter 7.4), and in this case any wave going to the right gets killed by a real negative exponential, thus e^{-ikx} can be seen as the reflected wave propagating towards $-\infty$.

  One could argue that also complex exponentials propagating to infinity may give problems with the normalization constraint. However, these terms represent **plane waves** and, as we saw in Section 7.6, these simply represent propagation of possible wave packets (imagine a wave packet $\phi(x)e^{ikx}$ with a very wide envelope function). In general, when we have functions with these types of plane waves, we cannot (and will not) normalize them. Instead, we will just look at the relationships between the constants (in this case A, B, C) to analyze "how much" of a wave is reflected or transmitted.

- Now consider a state with high energy, i.e. $E > V_L, V_R$. In this case we have complex exponential on both tails:

  $$\psi_E(x) = \begin{cases} Ae^{-ik_L x} + Be^{ik_L x} & x < x_L \\ Ce^{-ik_R x} + De^{ik_R x} & x > x_R \end{cases}$$

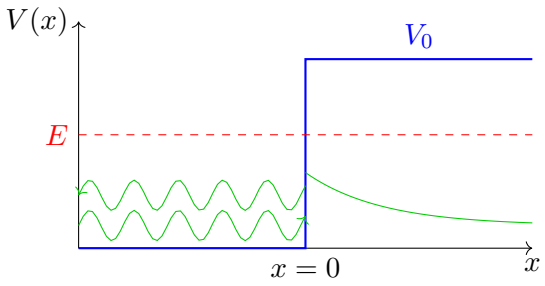


Figure 10.6: Diagram of a potential step.

For this case, it is convenient for us to see it as a superposition of two different wave functions $\psi_E^+(x), \psi_E^-(x)$

$$\psi_E^+(x) = \begin{cases} Ae^{-ik_L x} + Be^{ik_L x} & x < x_L \\ De^{ik_R x} & x > x_R \end{cases}, \quad \psi_E^-(x) = \begin{cases} Ae^{-ik_L x} & x < x_L \\ Ce^{-ik_R x} + De^{ik_R x} & x > x_R \end{cases},$$

In particular, $\psi_E^+(x)$ presents a case where, given a particle coming from $-\infty$, part of it will be transmitted, and the rest will be reflected. Also $\psi_E^-(x)$ can be seen analogously, for a particle coming from $+\infty$. These wave functions represent states we call **scattering**.

- The last case is when the energy of the state is low, i.e. $E < V_L, V_R$. By now one can imagine what will happen: on both tails we will have real exponentials, and they need to converge to 0 at $\pm\infty$, i.e.

$$\psi_E(x) = \begin{cases} Be^{b_L x} & x < x_L \\ Ce^{-b_R x} & x > x_R \end{cases}$$

An important observation to do here is that, although tails are at a higher energy level than the eigenvalue of $|\psi_E\rangle$, the potential must drop below E somewhere within the interaction zone, otherwise we would violate the constraint $\langle H \rangle \geq V_{min}$. These wave functions represent states we call **bound**.

10.5 State of a particle with a potential step

Here we derive a full wave function for a first, very simply case of interaction zone: a **potential step** (Figure 10.6):

$$V(x) = \begin{cases} 0 & x < 0 \\ V_0 & x > 0 \end{cases}$$

Reflecting states Consider a particle in stationary state $|\psi_E\rangle$ with energy eigenvalue $0 < E < V_0$. We already know $|\psi_E\rangle$ will be a reflecting state:

$$\psi_E(x) = \begin{cases} Ae^{ikx} + Be^{-ikx} & x < 0 \\ Ce^{-bx} & x > 0 \end{cases}$$

with $k = \sqrt{\frac{2\mu}{\hbar^2}E}$ and $b = \sqrt{\frac{2\mu}{\hbar^2}(V_0 - E)}$. As anticipated, in a case like this we are interested in transmission and reflection, namely $\frac{|B|^2}{|A|^2}$ and $\frac{|C|^2}{|A|^2}$, thus for simplicity we set $A = 1$. All we are left to do is to constrain B, C using continuity conditions on $x = 0$:

$$e^{ik0} + Be^{-ik0} = Ce^{-b0} \iff 1 + B = C$$

Also the first derivative must be continuous, i.e.

$$ike^{ik0} - iBke^{-ik0} = bCe^{-b0} \iff ik(1 - B) = bC$$

Solving this system of linear equations we get the values for B, C:

$$B = \frac{ik + b}{ik - b} = \frac{i\sqrt{E} + \sqrt{V_0 - E}}{i\sqrt{E} - \sqrt{V_0 - E}}$$
$$C = \frac{2ik}{ik - b} = \frac{2i\sqrt{E}}{i\sqrt{E} - \sqrt{V_0 - E}}$$

This result shows how the relationship between V_0 and E influences transmission and reflection: if $V_0 \gg E$ then the energy of the state is too low and $x = 0$ is a potential barrier. This yields a very low transmission, and in fact we would have $|C|^2 \ll 1$. In any case, the induced probability of finding the particle in a position x_R far beyond the barrier is:

$$\mathbf{P}\left([x_R, +\infty)\right) = \int_{x_R}^{+\infty} |\psi(x)|^2 dx = |C|^2 \int_{x_R}^{+\infty} e^{-2bx} dx = \frac{|C|^2}{2b}\left[-e^{-2bx}\right]_{x_R}^{+\infty} = \frac{|C|^2}{2b} e^{-2bx_R}$$

which decays exponentially in x_R for $x_R > 0$.

Scattering states Now we assume that $E > V_0$. This means that we have a scattering state with a wave function of the form:

$$\psi_E(x) = \begin{cases} Ae^{-ik_L x} + Be^{ik_L x} & x < 0 \\ Ce^{ik_R x} & x > 0 \end{cases},$$

with $k_L = \sqrt{\frac{2\mu}{\hbar^2}E}, k_R = \sqrt{\frac{2\mu}{\hbar^2}(E - V_0)}$. Analogously to the case of reflecting state, we set $A = 1$, and continuity conditions give:

$$\begin{cases} 1 + B = C \\ ik_L(1 - B) = ik_R C \end{cases} \implies \begin{cases} B = \frac{k_L - k_R}{k_L + k_R} = \frac{\sqrt{E} - \sqrt{E - V_0}}{\sqrt{E} + \sqrt{E - V_0}} \\ C = \frac{2k_L}{k_L + k_R} = \frac{2\sqrt{E}}{\sqrt{E} + \sqrt{E - V_0}} \end{cases}$$

10.6 Quantum tunneling and resonant scattering

Consider again a state $|\psi_E\rangle$ of a particle with energy eigenvalue E, this time along the x-axis with a potential function of the form:

$$V(x) = \begin{cases} 0 & x < 0 \\ V_0 & 0 < x < L \\ 0 & x > L \end{cases}$$

i.e. we have a barrier in the middle, as in Figure 10.7.

Quantum tunneling. If $V_0 > E$, by what we saw we should obtain a wave function of the form:

$$\psi_E(x) = \begin{cases} Ae^{ikx} + Be^{-ikx} & x < 0 \\ Ce^{bx} + De^{-bx} & 0 < x < L \\ Fe^{ikx} & x > L \end{cases}$$

with $k = \sqrt{\frac{2\mu}{\hbar^2}E}, b = \sqrt{\frac{2\mu}{\hbar^2}(V_0 - E)}$ as usual. Again, we just want to find out $\frac{|B|^2}{|A|^2}$ and $\frac{|F|^2}{|A|^2}$ so we set $A = 1$.

This time we will have to impose continuity conditions on two points:

- $x = 0$:

$$e^{ik0} + Be^{-ik0} = Ce^{b0} + De^{-b0} \iff 1 + B = C + D$$
$$ike^{ik0} - ikBe^{-ik0} = bCe^{b0} - bDe^{-b0} \iff ik(1 - B) = b(C - D)$$

- $x = L$:

$$Ce^{bL} + De^{-bL} = Fe^{ikL}$$
$$bCe^{bL} - bDe^{-bL} = ikFe^{ikL}$$

These give a system of 4 linearly independent equations, and we can easily solve it. We will get B, C, D, F as functions of L, V_0, E, and μ. For example, F will be:

$$F = \frac{4ikbe^{-ikL}}{(b + ik)^2 e^{-bL} - (b - k)^2 e^{bL}}$$

Once again, we are interested to know "how much" of the mass of the probability distribution is **transmitted**, i.e. will go towards $+\infty$. This can be quantified by the **transmission rate**:

$$T = \frac{|F|^2}{|A|^2}$$

where A is the coefficient in front of e^{ikx} in the left tail, while F is its counterpart on the right tail. In the case of the square barrier, we computed F, and set $A = 1$. After some computation we obtain:

$$\frac{1}{T} = \frac{1}{|F|^2} = 1 + \frac{V_0^2}{4E(V_0 - E)} \sinh^2(bL)$$

Let us analyze some cases here:

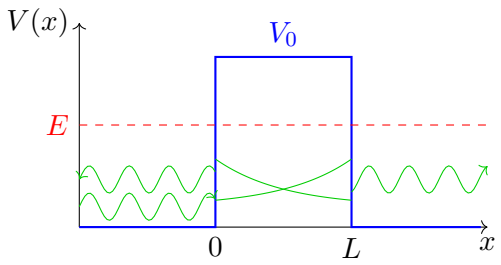


Figure 10.7: Diagram of a potential barrier.

- If $bL \ll 1$, we have a **narrow barrier**: the sine will tend to zero an thus $\frac{1}{T}$ will tend to 1 (along with T), having (nearly) full transmission;
- If $bL \gg 1$, the barrier to pass is **wide**. In this case the sine will tend to infinity and, in particular, the additive 1 above is negligible:

$$\begin{aligned} T &\simeq \frac{4E(V_0 - E)}{V_0^2 \sinh^2(bL)} \\ &= \frac{4E(V_0 - E)}{V_0^2} \frac{4}{(e^{bL} - e^{-bL})^2} \\ &\simeq \frac{4E(V_0 - E)}{V_0^2} \frac{4}{e^{2bL}} \\ &= \frac{16E(V_0 - E)}{V_0^2} e^{-2bL} \end{aligned}$$

  which decreases exponentially fast with the thickness L of the barrier. This phenomenon, called **quantum tunneling**, is somehow counter-intuitive: we never saw a ball passing through a wall[5]. Let us analyze the exponential term we obtained:

$$bL = \frac{\sqrt{2}}{\hbar} L \sqrt{\mu} \sqrt{V_0 - E}$$

  Since $\hbar \sim 10^{-34}$, the first term alone is telling us that this phenomenon is **highly unlikely** (since $\frac{1}{\hbar}$ on an exponential decreases quite fast). The term $\sqrt{\mu}$ tells us that the mass of the particle should be low enough for the quantum tunneling to happen. This is why we never encounter this phenomenon in classical physics, where we work with bodies with such a high mass that T becomes completely negligible. Also the difference $V_0 - E$ is important (as we might have expected).

  In a typical example with an electron, $\mu \sim 10^{-30}, V_0 - E \sim 10^{-19}$, thus the exponential (although still decreasing) does not make T negligible so easily.

Resonant scattering. Consider now a particle with state $|\psi_E\rangle$ which, this time, has an energy eigenvalue $E > V_0$. The picture remains essentially the same as in the previous paragraph, except we need to change the form of the wave function inside the potential barrier. In fact, this gives us a scattering state. Thus, let us consider only one of the elements of the superposition we discussed in Section 10.4 (the argument for the other wave is symmetric):

$$\psi(x) = \begin{cases} e^{ikx} + Be^{-ikx} & x < 0 \\ Ce^{ik'x} + De^{-ik'x} & 0 < x < L \\ Fe^{ikx} & x > L \end{cases}$$

with $k = \sqrt{\frac{2\mu}{\hbar^2} E}, k' = \sqrt{\frac{2\mu}{\hbar^2}(E - V_0)}$. We apply the continuity conditions once again:

- $x = 0$:

$$\begin{aligned} e^{ik0} + Be^{-ik0} &= Ce^{ik'0} + De^{-ik'0} \\ ike^{ik0} - ikBe^{-ik0} &= ik'Ce^{ik'0} - ik'De^{-ik'0} \end{aligned}$$

[5]This example is not out of the blue: in classical physics, a tall wall requires a ball to be high enough from the ground to pass it, which means having a high gravitational potential energy. This means that the shape of the wall in this case is actually the profile of what we called $V(x)$.

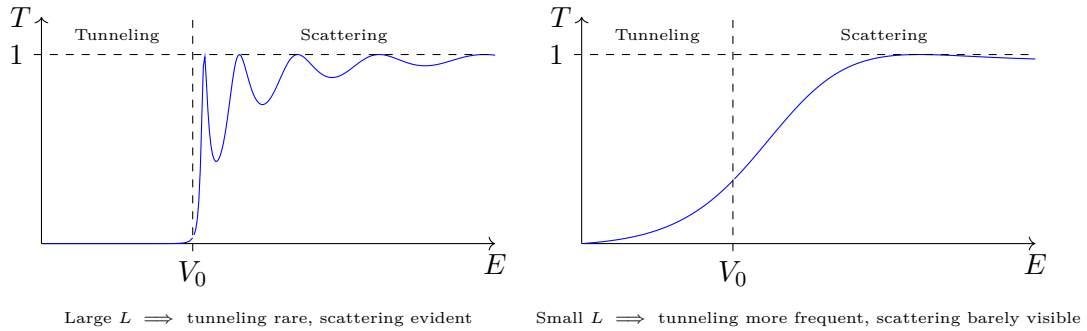


Figure 10.8: Transmission rate as a function of the energy of the state for two different barrier widths.

- $x = L$:

$$Ce^{ik'L} + De^{-ik'L} = Fe^{ikL}$$
$$ik'Ce^{ik'L} - ik'De^{-ik'L} = ikFe^{ikL}$$

After solving the linear system, the transmission rate becomes:

$$\frac{1}{T} = \frac{1}{|F|^2} = 1 + \frac{V_0^2}{4E(E - V_0)} \sin^2(k'L)$$

Let us analyze the cases as we did in the previous paragraph:

- If $k'L \ll 1$ (i.e. **narrow barrier**), then the sine becomes negligible and we obtain $T = 1$;
- The case of wide barrier becomes quite interesting: remember that $\hbar k'$ is the expected momentum of a particle inside the barrier.

  If we get $k'L = n\pi$ for some integer n, the sine above becomes 0, and we get full transmission again. Moreover, we obtain stationary waves similar to the ones we found in Section 10.2 for the particle in the box:

$$n\pi = \frac{\sqrt{2}}{\hbar} L\sqrt{\mu}\sqrt{E - V_0}$$
$$E - V_0 = \left(\frac{\hbar^2\pi^2}{2\mu L^2}\right) n^2 =: E_1 n^2$$

We can now put the two cases together (tunnelling and scattering) and look at what the transmission rate looks like as a function of the energy of the system. We observe in Figure 10.8 how there is a trade-off between the tunneling and scattering effects based on the thickness of the potential wall L.

Some applications of this analysis are used in **metrology**, where we would like to measure L (which is, say, the thickness of some material) by shooting some electrons at various energy levels (assuming we can control the energy eigenvalue E) and looking at how many of them get transmitted to the other side, estimating T. From this we can revert the equations above to find L. Conversely, if we know the mass μ but not the energy E, we can control L or V_0 to measure it in a analogous way.

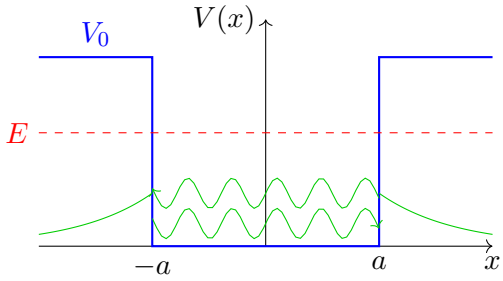


Figure 10.9: Diagram of a finite well.

10.7 Particle in a finite well

We conclude this chapter by looking at the finite counterpart of the particle in a box. In Figure 10.9, we consider a particle with stationary state $|\psi_E\rangle$ in a potential of the form:

$$V(x) = \begin{cases} V_0 & |x| > a \\ 0 & |x| < a \end{cases}$$

Of course we could choose any point for the center of the well, but choosing the origin makes the computations we are going to carry out more straightforward.

Bound state. If $E < V_0$, then we have a bound state:

$$\psi(x) = \begin{cases} Ae^{bx} & x < -a \\ Be^{ikx} + Ce^{-ikx} & |x| < a \\ De^{-bx} & x > a \end{cases}$$

and, with no surprise, $k = \sqrt{\frac{2\mu}{\hbar^2}E}, b = \sqrt{\frac{2\mu(V_0-E)}{\hbar^2}}$. Here comes the trick: the Hamiltonian now is symmetric with respect to $x = 0$. This gives us something useful that will restrict our solution space significantly.

Theorem 10.1. *If $V(x)$ is even, then the eigenstates of the Hamiltonian must have even or odd wave functions.*

Proof. We are going to use something called the **parity operator**:

$$\Gamma = \int_{\mathbb{R}} |-x\rangle\langle x| dx, \quad \Gamma|x\rangle = |-x\rangle$$

Notice that Γ is Hermitian and unitary as:

$$\langle x|\Gamma^\dagger|x'\rangle = \langle -x|x'\rangle = \delta(x + x') = \langle x|-x'\rangle = \langle x|\Gamma|x'\rangle$$
$$\langle x|\Gamma^\dagger\Gamma|x'\rangle = \langle -x|-x'\rangle = \delta(x - x') = \langle x|x'\rangle = \langle x|\mathbb{1}|x'\rangle$$

Thus the only possible eigenvalues are ± 1. For a generic state $|\psi\rangle$ with wave function $\psi(x)$, the parity operator does the following transformation:

$$\Gamma|\psi\rangle = \Gamma \int_{\mathbb{R}} \psi(x)|x\rangle dx = \int_{\mathbb{R}} \psi(x)\Gamma|x\rangle dx = \int_{\mathbb{R}} \psi(x)|-x\rangle dx = \int_{\mathbb{R}} \psi(-x)|x\rangle dx$$

i.e. the parity operator flips the sign of the argument in the wave function. Thus, if $|\psi\rangle$ is an eigenstate of Γ:

$$\Gamma|\psi\rangle = |\psi\rangle \iff \psi(-x) = \psi(x)$$
$$\Gamma|\psi\rangle = -|\psi\rangle \iff \psi(-x) = -\psi(x)$$

This means that the eigenstates associated with $+1$ and -1 have, respectively, even and odd wave functions. Now it is sufficient to prove that $[H, \Gamma] = 0$: at this point Theorem 4.2 would tell us that Γ and H have the same set of eigenstates, which means that any stationary state has either an even or an odd solution.

We start with $[\Gamma, P^2]$. For an arbitrary state $|\psi\rangle$ we have:

$$\begin{aligned}\langle x|\Gamma P|\psi\rangle &= \langle x|\Gamma \int_{\mathbb{R}} \left(-i\hbar\frac{\partial}{\partial x}\psi(x)\right)|x\rangle dx \\ &= -i\hbar\frac{\partial}{\partial x}\psi(-x)\end{aligned}$$

$$\begin{aligned}\langle x|P\Gamma|\psi\rangle &= \langle x|P \int_{\mathbb{R}} \psi(-x)|x\rangle dx \\ &= -i\hbar\frac{\partial}{\partial x}\psi(-x)\end{aligned}$$

i.e. both ΓP and $P\Gamma$ act in the same way on an arbitrary wave function, implying $[\Gamma, P] = 0$, and thus also $[\Gamma, P^2] = 0$. On the other hand:

$$\begin{aligned}V(X)\Gamma &= \int_{\mathbb{R}} V(x)|x\rangle\langle x|dx \int_{\mathbb{R}} |x\rangle\langle -x|dx \\ &= \int_{\mathbb{R}} V(x)|x\rangle\langle -x|dx \\ &= \int_{\mathbb{R}} V(-x)|-x\rangle\langle x|dx \\ &= \int_{\mathbb{R}} V(x)|-x\rangle\langle x|dx && \text{since } V(x) \text{ is even} \\ \Gamma V(X) &= \Gamma^\dagger V(X) && \text{since } \Gamma \text{ is Hermitian} \\ &= \int_{\mathbb{R}} |-x\rangle\langle x|dx \int_{\mathbb{R}} V(x)|x\rangle\langle x|dx \\ &= \int_{\mathbb{R}} V(x)|-x\rangle\langle x|dx\end{aligned}$$

so also $[\Gamma, V(X)] = 0$. Putting it all together:

$$[\Gamma, H] = \frac{1}{2\mu}[\Gamma, P^2] + [\Gamma, V(X)] = 0$$

as desired. □

From this reasoning, we can split into two cases for the wave function $\psi(x)$:

- $\psi(x)$ is even: we must have $B = C$, and $A = D$

$$\psi(x) = \begin{cases} Ae^{bx} & x < a \\ B\cos(kx) & |x| < a \\ Ae^{-bx} & x > a \end{cases}$$

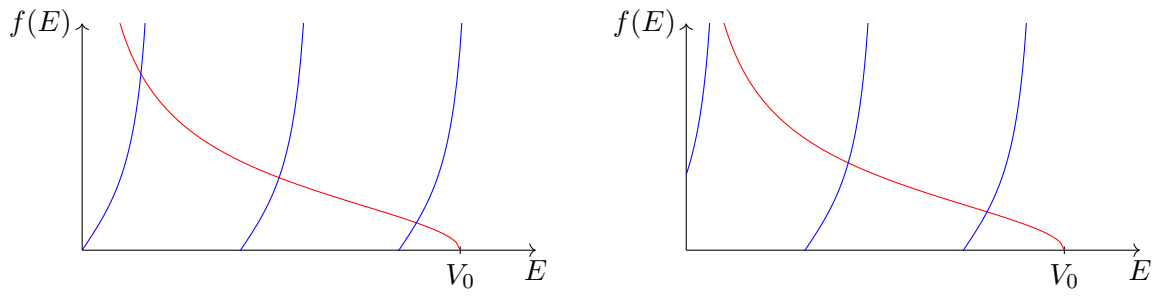


Figure 10.10: Solutions to the systems of equations for even (left) and odd (right) wave functions.

Symmetry is also nice because we will only need to apply continuity conditions at $x = a$, and the will get the other case for free.

$$\begin{aligned} Ae^{-ba} &= B\cos(ka) \\ -bAe^{-ba} &= -kB\sin(ka) \end{aligned}$$

If we divide the second equation by the first we obtain:

$$\frac{b}{k} = \tan(ka)$$

The choices of k and b for which this equation holds yield a stationary state. In order to get a flavour of what these solutions look like, we plot the two functions in Figure 10.10.

- $\psi(x)$ is odd: then $B = -C$ and $A = -D$

$$\psi(x) = \begin{cases} -Ae^{bx} & x < -a \\ B\sin(kx) & |x| < a \\ Ae^{-bx} & x > a \end{cases}$$

Here continuity conditions at $x = a$ give us

$$\begin{aligned} Ae^{-ba} &= B\sin(ka) \\ -bAe^{-ba} &= kB\cos(ka) \end{aligned}$$

implying that choices of b, k for stationary states this time must satisfy:

$$\frac{b}{k} = -\cot(ka)$$

To sum up, the allowed values for the energy eigenvalue E must satisfy one of the two equations we derived: the ratio b/k must be either $\tan(ka)$ (obtaining an even wave function) or $\cot(ka)$ (obtaining an odd wave function).

Chapter 11
Modeling Uncertainty

11.1 Information in quantum mechanics

Suppose that, for some reason, we do not know exactly which quantum state we have (in a closed box, say), but we nonetheless want to have a mathematical object that helps us describe the (partial) information we have about this state.

For this purpose, we model this uncertainty with a black box $\mathcal{B}$, which gives us some state $|\psi_i\rangle$ with probability p_i. If $(\Omega, \mathcal{F}, \mathbf{P}_{\psi_\mathbf{i}})$ is the probability space induced by $|\psi_i\rangle$, we would like to have a probability space for $\mathcal{B}$ in such a way that:

- this representation gives us the true distribution of an outcome x which is, by the law of total probability:

$$\mathbf{P}(x)_{\mathcal{B}} = \sum_i p_i \mathbf{P}(x)_i$$

- after a unitary evolution U, this distribution remains consistent with what happened:

$$\mathbf{P}(x)_{U(\mathcal{B})} = \sum_i p_i \mathbf{P}(x)_{U(\psi_i)}$$

  where $(\Omega, \mathcal{F}, \mathbf{P}_{U(\psi_i)})$ is the probability space induced by $U|\psi_i\rangle$.

Let us also talk about measurements: if we have an observable $\mathcal{O}$ of the form

$$\mathcal{O} = \sum_x \lambda_x \Pi_x$$

then the probability of measuring x is:

$$\mathbf{P}(x)_\psi = \langle\psi|\Pi_x|\psi\rangle$$

as we already know.

Let us now introduce an operator we know from linear algebra: the **trace**. The trace of a matrix A can be simply seen as the sum of the elements in the diagonal of A. Another way to define it is:

$$\mathrm{Tr}(A) = \sum_k \langle k|A|k\rangle$$

where $\{|k\rangle\}_k$ is an orthonormal basis of the Hilbert space in which A is an endomorphism. The trace operator can also be extended to continuous operators:

$$\mathrm{Tr}(A) = \int_{\mathbb{R}} \langle x|A|x\rangle dx$$

Now we can rewrite the outcome probability in an interesting way:

$$\begin{aligned}\mathbf{P}(x)_\psi &= \langle\psi|\Pi_x|\psi\rangle \\ &= \mathrm{Tr}\left(\langle\psi|\Pi_x|\psi\rangle\right) && \text{every scalar is the trace of itself}\end{aligned}$$

$$= \operatorname{Tr}(\Pi_x|\psi\rangle\langle\psi|) \qquad \text{by cyclic property (Theorem B.2)}$$

Another way to prove that the probability of an outcome is given by the above trace is the following: let us choose an orthonormal basis $\{|\psi_j\rangle\}_j$ to express the operator A (and its trace), where $|\psi_1\rangle = |\psi\rangle$ is our state.

$$\begin{aligned}\operatorname{Tr}(\Pi_x|\psi\rangle\langle\psi|) &= \sum_j \langle\psi_j|\Pi_x|\psi\rangle\langle\psi|\psi_j\rangle \\ &= \langle\psi|\Pi_x|\psi\rangle = \mathbf{P}(x)_\psi\end{aligned}$$

Other useful properties of the trace can be found in Section B.1.

11.2 The density matrix

We introduced a fancy use of the trace operator, but we did not solve our problem yet: how can we conveniently describe $\mathcal{B}$? Let us look at the total probability again now:

$$\begin{aligned}\mathbf{P}(x)_{\mathcal{B}} &= \sum_i p_i \mathbf{P}(x)_i \\ &= \sum_i p_i \langle\psi_i|\Pi_x|\psi_i\rangle \\ &= \sum_i p_i \operatorname{Tr}(\Pi_x|\psi_i\rangle\langle\psi_i|) \\ &= \operatorname{Tr}\left(\sum_i p_i \Pi_x|\psi_i\rangle\langle\psi_i|\right) && \text{linearity of trace} \\ &= \operatorname{Tr}\left(\Pi_x \sum_i p_i|\psi_i\rangle\langle\psi_i|\right) \\ &=: \operatorname{Tr}(\Pi_x\rho)\end{aligned}$$

We found that the total probability can be written as the trace of a product between two matrices: the projector operator relative to the outcome of the measurement Π_x, and a new matrix ρ, which we call the **density matrix**[6]. Notice that by "density" here we mean probability density. Also keep in mind that we tacitly assumed that the probability distribution of the states returned by the black box $\mathcal{B}$ is discrete, but nothing prevents us to define the same black box for continuous distributions, in which the density matrix will be defined with an integral sum. Moreover, this matrix does not necessarily have to be unitary (it does not even have to be invertible), since the possible states $|\psi_i\rangle$ are not orthonormal in general.

Evolving a distribution. Suppose we apply a unitary evolution U to whichever state $\mathcal{B}$ will give us. At the end we will have a set of states $\{|\psi_i'\rangle\}_i$, where $|\psi_i'\rangle = U|\psi_i\rangle$. The total probability will become:

$$\begin{aligned}\mathbf{P}(x)_{U(\mathcal{B})} &= \sum_i p_i \mathbf{P}(x)_{U(\psi_i)} \\ &= \sum_i p_i \operatorname{Tr}(\Pi_x|\psi_i'\rangle\langle\psi_i'|) \\ &= \sum_i p_i \operatorname{Tr}\left(\Pi_x U|\psi_i\rangle\langle\psi_i|U^\dagger\right)\end{aligned}$$

[6] This is also called density operator, this again depends on the system we model. Since in this chapter we will mainly talk about qubits, we will use the term matrix.

$$
\begin{aligned}
&= \operatorname{Tr}\left(\sum_i p_i \Pi_x U|\psi_i\rangle\langle\psi_i|U^\dagger\right) && \text{linearity of trace}\\
&= \operatorname{Tr}\left(\Pi_x U\left(\sum_i p_i|\psi_i\rangle\langle\psi_i|\right)U^\dagger\right)\\
&= \operatorname{Tr}\left(\Pi_x U\rho U^\dagger\right)
\end{aligned}
$$

We found that a mixture of states with density matrix ρ, after a unitary evolution U, becomes a mixture of states with density matrix $U\rho U^\dagger$.

Post-measurement states. The argument for the measurements does not work only with unitary operators, but with Hermitian operators in general: this gives us an expression for the post-measurement state for free. If we observe the subspace of a projector Π_k upon measurement, the projector will transform $\rho \mapsto \Pi_k \rho \Pi_k^\dagger$ and then, analogously to what we do with normal states, we will need to normalize ($\operatorname{Tr}\rho' = 1$, see next section):

$$
\begin{aligned}
\rho \mapsto & \frac{\Pi_k \rho \Pi_k^\dagger}{\operatorname{Tr}(\Pi_k \rho \Pi_k)}\\
&= \frac{\Pi_k \rho \Pi_k}{\operatorname{Tr}(\Pi_k \rho \Pi_k)} && \text{projector is Hermitian}\\
&= \frac{\Pi_k \rho \Pi_k}{\operatorname{Tr}(\Pi_k^2 \rho)} && \text{cyclic property of trace}\\
&= \frac{\Pi_k \rho \Pi_k}{\operatorname{Tr}(\Pi_k \rho)} && \text{projector is idempotent}\\
&= \frac{\Pi_k \rho \Pi_k}{\mathbf{P}(k)_\rho}
\end{aligned}
$$

Expectation. Given an observable $A = \sum_k a_k |k\rangle\langle k|$, with $\{|k\rangle\}_k$ orthonormal basis, we can also find a neat expression for the expectation of an observable A under a given state ρ:

$$
\begin{aligned}
\langle A\rangle^\rho &= \sum_k a_k \mathbf{P}(k)_\rho\\
&= \sum_k a_k \operatorname{Tr}(|k\rangle\langle k|\rho)\\
&= \operatorname{Tr}\left(\sum_k a_k |k\rangle\langle k|\rho\right) && \text{linearity of trace}\\
&= \operatorname{Tr}(A\rho)
\end{aligned}
$$

Examples with qubits. Suppose that the black box $\mathcal{B}_1$ returns a qubit with states $|0\rangle$ or $|1\rangle$ uniformly at random, i.e. with probability $\frac{1}{2}$ each. The density matrix ρ_1 of this distribution is:

$$\rho_1 = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}\mathbb{1}$$

Now consider a black box $\mathcal{B}_2$ returns a qubit with states $|+\rangle$ or $|-\rangle$ uniformly at random. The density matrix ρ_2 is:

$$\rho = \frac{1}{2}\left(|+\rangle\langle +| + |-\rangle\langle -|\right) = \frac{1}{2}\mathbb{1}$$

We obtained the same density matrix. This means that in practice we cannot distinguish $\mathcal{B}_1$ and $\mathcal{B}_2$, even after an arbitrary evolution U since $U\mathbb{1}U^\dagger = \mathbb{1}$, i.e. the density matrix does not change upon evolution. This property of quantum information is quite unique, and sets an important difference from classical information: clearly $|+\rangle, |-\rangle$ are different from $|0\rangle, |1\rangle$, yet it is impossible to tell the two cases apart.

11.3 Properties of the density matrix

Let us analyze ρ, and derive some properties. First of all, we notice that ρ can be seen an endomorphism of the Hilbert space $\mathcal{H}$ containing the states $|\psi_i\rangle$ in the mixture.

Theorem 11.1. $\mathrm{Tr}(\rho) = 1$.

Proof. Let $\{|\psi_i\rangle\}_i$ be the set of possible states with mixing probabilities p_i.

$$\begin{aligned}
\mathrm{Tr}(\rho) &= \mathrm{Tr}\left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right) \\
&= \sum_i \mathrm{Tr}\left(p_i |\psi_i\rangle\langle\psi_i|\right) && \text{linearity of trace} \\
&= \sum_i \mathrm{Tr}\left(p_i \langle\psi_i|\psi_i\rangle\right) && \text{cyclic property of trace} \\
&= \sum_i p_i = 1
\end{aligned}$$

□

Theorem 11.2. *ρ is Hermitian.*

Proof.

$$\rho^\dagger = \left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right)^\dagger = \sum_i p_i (|\psi_i\rangle\langle\psi_i|)^\dagger = \sum_i p_i |\psi_i\rangle\langle\psi_i| = \rho$$

□

Theorem 11.3. *ρ is positive semi-definite.*

Proof. For any $|\phi\rangle$ in $\mathcal{H}$:

$$\begin{aligned}
\langle\phi|\rho|\phi\rangle &= \sum_i p_i \langle\phi|\psi_i\rangle\langle\psi_i|\phi\rangle \\
&= \sum_i p_i |\langle\psi_i|\phi\rangle|^2 \\
&\geq 0
\end{aligned}$$

□

These properties tell us something important about the spectral decomposition of ρ: Theorem 11.2 ensures that the eigenbasis of ρ is orthonormal, which is a property we always appreciate in quantum theory. Theorems 11.1 and 11.3 tell us something about its eigenvalues: positive semi-definiteness implies that the eigenvalues are all non-negative while, on the other hand,

the sum of the eigenvalues of a matrix always equals the trace, which in our case is 1. The eigendecomposition seems to yield a probability mixture of new states:

$$\rho = UDU^\dagger = \sum_x \mathbf{P}(x)_x |\psi_x\rangle\langle\psi_x|$$

where D is diagonal, $\{|\psi_x\rangle\}$ forms a orthonormal basis of $\mathcal{H}$, and $\mathbf{P}_x$ is the probability induced by $|\psi_x\rangle$. Just like in the example with qubits we have introduced in Section 11.2, if we construct a black box $\mathcal{B}'$ using the eigenbasis of ρ as probability mixture, we would obtain a total probability that is indistinguishable from the original (ρ did not change after all).

We present here two important special cases:

- if $\rho = |\phi\rangle\langle\phi|$, meaning that the mixture yields $|\phi\rangle$ with probability 1, we call such mixture **pure state**;
- if $\rho = \frac{\mathbb{1}}{|\mathcal{H}|}$[7], the distribution will be uniform among the states in $\mathcal{H}$, and this corresponds to what we call **fully mixed state**.

From now on we will extend the term "state" also to refer to such distributions. Moreover, we will denote with $\mathcal{S}(\mathcal{H}) \subseteq \mathrm{End}(\mathcal{H})$ the space of density matrices in the Hilbert space $\mathcal{H}$.

11.4 Ignorance about local information and the partial trace

Consider an example where we have a state $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ shared by two players, Alice and Bob. Alice has only access to $\mathcal{H}_A$, and Bob only to $\mathcal{H}_B$ (you can imagine two qubits A, B). We would like to represent the knowledge that only one of the players has about the global system.

Definition 11.4 (Partial trace)**.** *Given a composite system $\mathcal{H}_A \otimes \mathcal{H}_B$, the partial trace with respect to $\mathcal{H}_A$ is a function*

$$\begin{aligned} \mathrm{Tr}_B : \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B) &\mapsto \mathcal{S}(\mathcal{H}_A) \\ \rho_{AB} &\mapsto \rho_A \end{aligned}$$

i.e. it yields the density matrix of the subsystem A, given the density matrix of the global system.

Let us derive a general expression that we can use. We know that the partial trace above must satisfy the following conditions:

- A local measurement on $\mathcal{H}_A$ (i.e. an observable of the form $M_A \otimes \mathbb{1}_B$) must behave with the correct outcome probability distribution;
- A local evolution on $\mathcal{H}_B$ should not change the value of the partial trace.

Notice that these are the exact same properties we found in Sections 8.1 and 8.2, when we talked about local observables and non-interacting systems.

[7] Although this is an extreme abuse of notation, you can imagine that this also works if $\mathcal{H}$ is infinite-dimensional.

Local observables. Suppose to have an observable $\mathcal{O} = M_A \otimes \mathbb{1}_B$, as anticipated, where $M = \sum_x a_x \Pi_x$, with $\Pi_x = |x\rangle\langle x|$. The property we described above should translate to the following constraint in the induced probability spaces:

$$\mathbf{P}(x)_{\rho_A} = \mathbf{P}(x)_{\rho_{AB}}$$

which, rewritten in terms of traces becomes:

$$\mathrm{Tr}\left(\Pi_x \rho_A\right) = \mathrm{Tr}\left((\Pi_x \otimes \mathbb{1}_B)\rho_{AB}\right)$$

Let us now choose an orthonormal basis $\{|i\rangle_A |j\rangle_B\}_{i,j}$ of $\mathcal{H}_A \otimes \mathcal{H}_B$ which we use to express the trace:

$$\begin{aligned}\mathrm{Tr}\left((\Pi_x \otimes \mathbb{1}_B)\rho_{AB}\right) &= \sum_{i,j} \langle i|\langle j| \left((\Pi_x \otimes \mathbb{1}_B)\rho_{AB}\right) |i\rangle|j\rangle \\ &= \sum_{i,j} \left(\langle i|\Pi_x\right) \otimes \left(\langle j|\mathbb{1}_B\right) \rho_{AB} |i\rangle|j\rangle\end{aligned}$$

We are going to choose $\{|i\rangle\} = \{|x\rangle\}$, since we could choose any orthonormal basis to define the trace.

$$\begin{aligned}\mathrm{Tr}\left((\Pi_x \otimes \mathbb{1}_B)\rho_{AB}\right) &= \sum_{i,j} \left(\langle i|x\rangle\langle x|\right) \otimes \left(\langle j|\mathbb{1}_B\right) \rho_{AB} |i\rangle|j\rangle \\ &= \sum_{j} \left(\langle x|_A \langle j|_B\right) \rho_{AB} \left(|x\rangle_A |j\rangle_B\right) \\ &= \langle x| \left(\sum_j \left(\mathbb{1}_A \otimes \langle j|_B\right) \rho_{AB} \left(\mathbb{1}_A \otimes |j\rangle_B\right) \right) |x\rangle\end{aligned}$$

Notice that the matrix in the sum is in $\mathcal{S}(\mathcal{H}_A)$ and in fact, under the assumption that Π_x projects onto a single basis element $|x\rangle$, what we have within the tuples is exactly ρ_A since:

$$\mathrm{Tr}(|x\rangle\langle x|\rho_A) = \mathrm{Tr}(\langle x|\rho_A|x\rangle) = \langle x|\rho_A|x\rangle$$

as in the above expression.

Therefore, a good candidate for the definition of the partial trace can be:

$$\mathrm{Tr}_B(\rho_{AB}) := \sum_j \left(\mathbb{1}_A \otimes \langle j|_B\right) \rho_{AB} \left(\mathbb{1}_A \otimes |j\rangle_B\right)$$

Since the trace is a linear operator we can use the following notation:

$$\mathrm{Tr}_B(\rho_{AB}) := \left(\mathbb{1}_A \otimes \mathrm{Tr}_B\right) \rho_{AB}$$

Local evolutions. We still need to prove that the definition we found is independent from possible evolutions in $\mathcal{H}_B$. Suppose we evolve the two systems independently with an operator $(U_A \otimes V_B)$. We already proved that the density matrix becomes:

$$\rho_{AB} \mapsto \rho'_{AB} = (U_A \otimes V_B)\rho_{AB}(U_A^\dagger \otimes V_B^\dagger)$$

Let us see what happens to the partial trace with our definition:

$$\mathrm{Tr}_B(\rho'_{AB}) = \mathrm{Tr}_B\left((U_A \otimes V_B)\rho_{AB}(U_A^\dagger \otimes V_B^\dagger)\right)$$

$$= \sum_j (\mathbb{1}_A \otimes \langle j|_B)\,(U_A \otimes V_B)\rho_{AB}(U_A^\dagger \otimes V_B^\dagger)\,(\mathbb{1}_A \otimes |j\rangle_B)$$
$$= \sum_j (U_A \otimes \langle j|_B V_B)\,\rho_{AB}\left(U_A^\dagger \otimes V_B^\dagger |j\rangle_B\right)$$

Now we simply do a change of basis $|j\rangle \leftarrow V|j\rangle$ (note that it is still an orthonormal basis since V is unitary). We know that the trace is independent of the basis and:

$$\begin{aligned}\mathrm{Tr}_B(\rho'_{AB}) &= \sum_j (U_A \otimes \langle j|_B)\,\rho_{AB}\left(U_A^\dagger \otimes |j\rangle_B\right)\\ &= U\left(\sum_j (\mathbb{1}_A \otimes \langle j|_B)\,\rho_{AB}\,(\mathbb{1}_A \otimes |j\rangle_B)\right)U^\dagger\\ &= U\rho_A U^\dagger\end{aligned}$$

which means that a local evolution in B does not change anything on the partial trace, exactly what we wanted. This confirms that the definition we found was exactly what we were looking for. It is worth noticing that in order to find $\rho'_A = U\rho_A U^\dagger$, one can either evolve the partial trace $\mathrm{Tr}_B(\rho_{AB})$, or compute the partial trace on the evolved state $\mathrm{Tr}_B(\rho'_{AB})$, i.e.

$$\mathrm{Tr}_B\left((U_A \otimes V_B)\rho_{AB}(U_A^\dagger \otimes V_B^\dagger)\right) = U\,\mathrm{Tr}_B(\rho_{AB})U^\dagger$$

We close this section by showing two properties of the partial trace, which will be useful later:

Theorem 11.5 (Linearity of partial trace)**.** $\mathrm{Tr}_B(\alpha\rho_1 + \beta\rho_2) = \alpha\,\mathrm{Tr}_B(\rho_1) + \beta\,\mathrm{Tr}_B(\rho_2)$.

Proof.

$$\begin{aligned}\mathrm{Tr}_B(\alpha\rho_1 + \beta\rho_2) &= \sum_j (\mathbb{1}_A \otimes \langle j|_B)\,(\alpha\rho_1 + \beta\rho_2)\,(\mathbb{1}_A \otimes |j\rangle_B)\\ &= \sum_j (\mathbb{1}_A \otimes \langle j|_B)\,(\alpha\rho_1)\,(\mathbb{1}_A \otimes |j\rangle_B) + \sum_j (\mathbb{1}_A \otimes \langle j|_B)\,(\beta\rho_2)\,(\mathbb{1}_A \otimes |j\rangle_B)\\ &= \alpha\sum_j (\mathbb{1}_A \otimes \langle j|_B)\,\rho_1\,(\mathbb{1}_A \otimes |j\rangle_B) + \beta\sum_j (\mathbb{1}_A \otimes \langle j|_B)\,\rho_2\,(\mathbb{1}_A \otimes |j\rangle_B)\\ &= \alpha\,\mathrm{Tr}_B(\rho_1) + \beta\,\mathrm{Tr}_B(\rho_2)\end{aligned}$$

□

Theorem 11.6. $\mathrm{Tr}_B(\rho_A \otimes \rho_B) = \rho_A$.

Proof.

$$\begin{aligned}\mathrm{Tr}_B(\rho_A \otimes \rho_B) &= \sum_j (\mathbb{1}_A \otimes \langle j|_B)\,(\rho_A \otimes \rho_B)\,(\mathbb{1}_A \otimes |j\rangle_B)\\ &= \sum_j (\rho_A \otimes \langle j|_B\rho_B)\,(\mathbb{1}_A \otimes |j\rangle_B)\\ &= \sum_j (\rho_A \otimes \langle j|_B\rho_B)\,(\mathbb{1}_A \otimes |j\rangle_B)\\ &= \rho_A \sum_j \langle j|\rho_B|j\rangle\\ &= \rho_A\,\mathrm{Tr}(\rho_B) = \rho_A \qquad \text{by Theorem 11.1}\end{aligned}$$

□

11.5 Superposition vs probabilistic mixture

In this section we want to evidence an important difference between a qubit in a superposition and a bit chosen uniformly at random. Consider two different states:

- ρ_1 yields $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$ with probability 1;
- ρ_2 gives one of $|0\rangle, |1\rangle$ uniformly at random.

Analysis of the superposition. Let us start by analyzing ρ_1: its density matrix can be computed directly using Dirac notation:

$$\begin{aligned}\rho_1 &= |+\rangle\langle+| \\ &= \frac{1}{2}(|0\rangle + |1\rangle)(\langle 0| + \langle 1|) \\ &= \frac{1}{2}(|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|) \\ &= \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\end{aligned}$$

Notice that ρ_1, like any density matrix of a pure state, is idempotent:

$$\rho_1^2 = (|+\rangle\langle+|)(|+\rangle\langle+|) = |+\rangle\langle+|+\rangle\langle+| = |+\rangle\langle+| = \rho_1$$

This is a good way to check whether a density matrix (expressed in vector notation, say) corresponds to a pure state or not.

Now let us compute the outcome probabilities when we measure ρ_1:

- Using the Pauli matrix Z:

$$\begin{aligned}\mathbf{P}(0)_{\rho_1} &= \mathrm{Tr}(|0\rangle\langle 0|\rho_1) \\ &= \mathrm{Tr}(|0\rangle\langle 0|+\rangle\langle+|) \\ &= \mathrm{Tr}\left(\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\right) \\ &= \frac{1}{2}\mathrm{Tr}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2} = \mathbf{P}(1)_{\rho_1}\end{aligned}$$

- Using the Pauli matrix X:

$$\begin{aligned}\mathbf{P}(+)_{\rho_1} &= \mathrm{Tr}(|+\rangle\langle+|\rho_1) \\ &= \mathrm{Tr}(|+\rangle\langle+|+\rangle\langle+|) \\ &= \frac{1}{2}\mathrm{Tr}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 1\end{aligned}$$

It is worth mentioning that nobody forces us to compute the trace using matrix notation. Sometimes it may be even faster to use Dirac notation and take advantage of the linearity of trace.

Now let us consider an example of evolution of ρ_1: we make the qubit pass through a Hadamard gate H. The state after the transformation is:

$$\rho_1 \mapsto H\rho_1 H^\dagger = |0\rangle\langle 0|$$

i.e. the state will be $|0\rangle$ with probability 1, which totally makes sense, since H always transform $|+\rangle$ to $|0\rangle$.

Analysis of the probabilistic mixture. We immediately see the first difference by computing the density matrix:

$$\rho_2 = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}\mathbb{1}$$

And it is clearly not a pure state:

$$\rho_2^2 = \frac{1}{4}\mathbb{1} \neq \rho_2$$

Indeed, let us see what happens with a measurement here:

- with the Pauli matrix Z:

$$\begin{aligned}\mathbf{P}(0)_{\rho_2} &= \mathrm{Tr}\left(|0\rangle\langle 0|\rho_2\right) \\ &= \frac{1}{2}\mathrm{Tr}\left(|0\rangle\langle 0|\mathbb{1}\right) \\ &= \frac{1}{2}\mathrm{Tr}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}\end{aligned}$$

- Using the Pauli matrix X:

$$\begin{aligned}\mathbf{P}(+)_{\rho_2} &= \mathrm{Tr}\left(|+\rangle\langle +|\rho_2\right) \\ &= \frac{1}{2}\mathrm{Tr}\left(|+\rangle\langle +|\mathbb{1}\right) \\ &= \frac{1}{2}\mathrm{Tr}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \frac{1}{2}\end{aligned}$$

Moreover, any evolution of ρ_2 does not change any information we have about a fully mixed state:

$$\rho_2 \mapsto U\rho_2 U^\dagger = \frac{1}{2}U\mathbb{1}U^\dagger = \frac{1}{2}\mathbb{1} = \rho_2$$

From these we can infer a different characterization of a fully mixed state.

Theorem 11.7. *A state $\rho \in \mathcal{S}(\mathcal{H})$ is fully mixed if and only if the probability distribution of outcomes is **uniform** for any chosen measurement basis.*

On the other hand, we have already seen in Section 1.4 that the notion of superposition is relative to a particular measurement basis, and thus there exists a basis (indeed, infinitely many) in which the outcome is deterministic. This is why we called this type of density matrices *pure state*.

11.6 Entanglement vs probabilistic mixture

We will see now how entanglement behaves with this new formalization, and we will see three example cases.

Pure entangled state. Let us now consider the density matrix $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ of two qubits, and suppose it gives a pure state:

$$\rho_{AB} = |\psi\rangle\langle\psi|_{AB}$$

where $|\psi\rangle$ is the entangled state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. We rewrite ρ_{AB}:

$$\begin{aligned}\rho_{AB} &= \frac{1}{2}\left(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|\right) \\ &= \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}\end{aligned}$$

What happens now if we take the partial trace?

$$\begin{aligned}\rho_A &= \mathrm{Tr}_B(\rho_{AB}) \\ &= \frac{1}{2}\left(\mathrm{Tr}_B(|00\rangle\langle 00|) + \mathrm{Tr}_B(|00\rangle\langle 11|) + \mathrm{Tr}_B(|11\rangle\langle 00|) + \mathrm{Tr}_B(|11\rangle\langle 11|)\right) \\ &= \frac{1}{2}\sum_j \left(\mathbb{1}_A \otimes \langle j|_B\right)\rho_{AB}\left(\mathbb{1}_A \otimes |j\rangle_B\right)\end{aligned}$$

Now notice that $\mathrm{Tr}_B(|00\rangle\langle 11|) = \mathrm{Tr}_B(|11\rangle\langle 00|) = 0$, since in the terms of the sum $\langle j|0\rangle\langle 1|j\rangle$ or $\langle j|1\rangle\langle 0|j\rangle$ will appear; if we choose a basis $\{|j\rangle\}$ containing $|0\rangle, |1\rangle$, all these terms cancel out since at most one of $|0\rangle, |1\rangle$ can be equal to $|j\rangle$. Thus we are left with:

$$\begin{aligned}\rho_A &= \frac{1}{2}\mathrm{Tr}\left(|00\rangle\langle 00| + |11\rangle\langle 11|\right) \\ &= \frac{1}{2}\left(\sum_j \left(\mathbb{1}_A \otimes \langle j|_B\right)|00\rangle\langle 00|\left(\mathbb{1}_A \otimes |j\rangle_B\right) + \sum_j \left(\mathbb{1}_A \otimes \langle j|_B\right)|11\rangle\langle 11|\left(\mathbb{1}_A \otimes |j\rangle_B\right)\right) \\ &= \frac{1}{2}\left(\sum_j \left(|0\rangle_A \otimes \langle j|0\rangle_B\right)\left(\langle 0|_A \otimes \langle 0|j\rangle_B\right) + \sum_j \left(|1\rangle_A \otimes \langle j|1\rangle_B\right)\left(\langle 1|_A \otimes \langle 1|j\rangle_B\right)\right) \\ &= \frac{1}{2}\left(|0\rangle\langle 0|_A + |1\rangle\langle 1|_A\right) = \frac{1}{2}\mathbb{1}_A\end{aligned}$$

i.e. if we have an entangled pure state globally, then locally we get a mixed state. While this is somewhat counter-intuitive (a global superposition is giving a local mixture after all), keep in mind that having a locally pure state would mean that the global state could be written as tensor product of local states, which falls in contradiction with the fact that $|\psi\rangle$ is entangled.

Classical correlation. Now suppose that ρ_{AB} gives us one of $|00\rangle, |11\rangle$ uniformly at random, i.e.

$$\begin{aligned}\rho_{AB} &= \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11| \\ &= \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\end{aligned}$$

This is what we call **classical correlation**, because the two qubits are correlated in the sense of probability theory. The partial trace $\mathrm{Tr}_B(\rho_{AB})$ is the same as in the previous case, because we only lack the terms $|00\rangle\langle 11|, |11\rangle\langle 00|$, which canceled out anyway in the computations above:

$$\begin{aligned}\rho_A &= \mathrm{Tr}_B(\rho_{AB}) \\ &= \frac{1}{2}\left(|0\rangle\langle 0|_A + |1\rangle\langle 1|_A\right) = \frac{1}{2}\mathbb{1}_A\end{aligned}$$

i.e. only looking at the first qubit gives us a random bit. We found that the two cases are **locally indistinguishable**: we cannot tell if two qubits are entangled or only classically correlated if we only have access to one of them, but the situation can get worse.

Local mixtures. In this case, we have two completely unentangled qubits in fully mixed state:

$$\rho_{AB} = \rho_A \otimes \rho_B = \frac{1}{2}\mathbb{1}_A \otimes \frac{1}{2}\mathbb{1}_B = \frac{1}{4}\mathbb{1}_{AB} = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

By Theorem 11.6, also in this case the partial trace becomes:

$$\mathrm{Tr}_B\left(\rho_A \otimes \rho_B\right) = \rho_A = \frac{1}{2}\mathbb{1}_A$$

Let us see which post-measurement states we get by measuring in the following three different bases, namely,

- with $Z_A \otimes Z_B$, i.e. the computational basis:

$$\{|0\rangle, |1\rangle\} \otimes \{|0\rangle, |1\rangle\} = \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$$

- with $X_A \otimes X_B$, i.e. the Hadamard basis:

$$\{|+\rangle, |-\rangle\} \otimes \{|+\rangle, |-\rangle\} = \{|+\rangle|+\rangle, |+\rangle|-\rangle, |-\rangle|+\rangle, |-\rangle|-\rangle\}$$

- with B, i.e. the Bell basis:

$$\left\{|\Phi^+\rangle, |\Psi^+\rangle, |\Phi^-\rangle, |\Psi^-\rangle\right\} = \left\{\frac{|00\rangle + |11\rangle}{\sqrt{2}}, \frac{|01\rangle + |10\rangle}{\sqrt{2}}, \frac{|00\rangle - |11\rangle}{\sqrt{2}}, \frac{|01\rangle - |10\rangle}{\sqrt{2}}\right\}$$

	$Z_A \otimes Z_B$	$X_A \otimes X_B$	B
Pure entangled	$\|00\rangle, \|11\rangle$ u.a.r.	$\|+\rangle\|+\rangle, \|-\rangle\|-\rangle$ u.a.r.	$\|\Phi^+\rangle$
Classical correlation	$\|00\rangle, \|11\rangle$ u.a.r.	any basis state u.a.r.	$\|\Phi^+\rangle, \|\Phi^-\rangle$ u.a.r.
Local mixtures	any basis state u.a.r.	any basis state u.a.r.	any basis state u.a.r.

Notice that these measurements can actually distinguish the three cases (the distribution is different, a statistical test is sufficient), but they are **not** local measurements.

11.7 Uncertainty about evolution

In the previous sections we formalized the concept of density matrix, in order to describe a system for which we do not completely know the state. Now we want to address the case in which we are uncertain about the **evolution**. We model this in an analogous way: we have a black box $\mathcal{B}$ just as before, returning a state ρ. Moreover, we have another black box $\mathcal{E}$ which takes ρ as input, and returns the evolved state. Thus, if we were certain about the evolution, i.e. we know it is a unitary operator U, we would already know what the evolved state would look like:

$$\mathcal{E} : \rho \mapsto U\rho U^\dagger$$

On the other hand, if we suppose that $\mathcal{E}$ evolves the input state using operator U_i with probability p_i, then we can use the law of total probability:

$$\mathcal{E} : \rho \mapsto \mathcal{E}(\rho) = \sum_i p_i \cdot U_i \rho U_i^\dagger$$

A different way to express uncertainty about evolution is to introduce a state σ_E, which in some sense represents the state of the environment, and then evolve the state $\rho \otimes \sigma_E$ with a known evolution operator U:

$$\mathcal{E} : \rho \mapsto \mathcal{E}(\rho) = \left(U_{AE}(\rho \otimes \sigma_E)U_{AE}^\dagger\right)$$

In this second representation, the uncertainty about the evolution lies in what we do not know about the environment, while the evolution itself is well-known. It is important to note that the transformations defined above are, in general, non-reversible, i.e. they are not invertible.

How can we define a model for a general case that takes into account both the definitions above? We would like such a map to be:

- **linear**: this is because we want in particular that:

  $$\mathcal{E}\left(\sum_i p_i \rho_i\right) = \sum_i p_i \mathcal{E}(\rho_i)$$

  since we can always express a state as mixture of other states, and this would keep the probabilities consistent with the evolution;

- **trace-preserving**: $\mathrm{Tr}(\rho) = \mathrm{Tr}(\mathcal{E}(\rho))$, in order for the new state $\mathcal{E}(\rho)$ to preserve Theorem 11.1 and thus still be expressed as probabilistic mixture;

- **completely positive**: the new state $\mathcal{E}(\rho)$ must also preserve Theorem 11.3, i.e must remain positive semi-definite. This must be true also if we apply $\mathcal{E}$ to a subsystem:

  $$\rho_{AB} \succcurlyeq 0 \Longrightarrow (\mathcal{E}_A \otimes \mathbb{1}_B)(\rho_{AB}) \succcurlyeq 0$$

Thus, we formalize the concept of uncertain evolution with a mapping:

$$\mathcal{E}_{A \mapsto B} : \mathcal{S}(\mathcal{H}_A) \mapsto \mathcal{S}(\mathcal{H}_B)$$

satisfying the three properties we mentioned. Notice that the Hilbert space changes because we also use it to model transformations from a system to another. These maps are called (non-ironically) **trace-preserving completely positive maps** (or **TPCPM**). **Quantum channels** is another very popular term we will use.

11.8 Kraus decomposition

We present here the first representation of TPCPMs, generalizing the idea of total probability presented in the previous section. It is called **operator-sum** (or **Kraus**) decomposition:

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger, \quad \text{where } \sum_k E_k^\dagger E_k = \mathbb{1}_A$$

and $\{E_k\}_k$ are called **Kraus operators**. Linearity comes natural, so let us see if it satisfies the other conditions.

Theorem 11.8. *The Kraus decomposition preserves the trace.*

Proof.

$$\begin{aligned}
\operatorname{Tr}\left(\sum_k E_k \rho E_k^\dagger\right) &= \sum_k \operatorname{Tr}\left(E_k \rho E_k^\dagger\right) && \text{linearity of trace} \\
&= \sum_k \operatorname{Tr}\left(\rho E_k^\dagger E_k\right) && \text{cyclic property of trace} \\
&= \operatorname{Tr}\left(\rho \sum_k E_k^\dagger E_k\right) \\
&= \operatorname{Tr}(\rho) && \text{since } \sum_k E_k^\dagger E_k = \mathbb{1}_A
\end{aligned}$$

□

Theorem 11.9. *The Kraus decomposition is completely positive.*

Proof. Considering $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \succcurlyeq 0$:

$$\begin{aligned}
\sum_k E_k \rho E_k^\dagger &= \sum_k E_k \left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right) E_k^\dagger \\
&= \sum_i p_i \sum_k E_k |\psi_i\rangle\langle\psi_i| E_k^\dagger
\end{aligned}$$

One can see that $\sum_k |\psi_{i,k}\rangle\langle\psi_{i,k}|$ is a sum of outer products $E_k|\psi\rangle(E_k|\psi\rangle)^\dagger$, which is positive semi-definite by Theorems B.23 and B.24. Therefore, for any vector $|\phi\rangle$, and since $p_i \geq 0$:

$$\langle\phi| \left(\sum_i p_i \sum_k E_k |\psi_i\rangle\langle\psi_i| E_k^\dagger\right) |\phi\rangle = \sum_i p_i \langle\phi| \left(\sum_k E_k |\psi_i\rangle\langle\psi_i| E_k^\dagger\right) |\phi\rangle \geq 0$$

Moreover, if we suppose to have a state in a composite system $\rho_{AE} = \sum_i p_i |\psi_i\rangle\langle\psi_i|$:

$$\begin{aligned}
(\mathcal{E}_{A\mapsto B} \otimes \mathbb{1}_E)(\rho_{AE}) &= \sum_k (E_k \otimes \mathbb{1}_E) \rho_{AE} (E_k^\dagger \otimes \mathbb{1}_E) \\
&= \sum_i p_i \left(\sum_k (E_k \otimes \mathbb{1}_E) |\psi_i\rangle\langle\psi_i| (E_k \otimes \mathbb{1}_E)^\dagger\right)
\end{aligned}$$

just like in the single-system case, we have a sum of outer products, and the result remains positive semi-definite with the same argument. □

Let us see some examples with the Kraus decomposition:

- A **unitary map** is indeed a special case of Kraus decomposition

$$\mathcal{E}(\rho) = U\rho U^\dagger, \quad U^\dagger U = \mathbb{1}$$

- **Erasure channel**: maps every state to a fixed state $\rho_A \mapsto |\psi\rangle_B$

$$\mathcal{E}(\rho) = \sum_k |\psi\rangle\langle k|\rho|k\rangle\langle\psi|$$

for an orthonormal basis $\{|k\rangle\}_k$ of $\mathcal{H}_A$. One can see that this indeed gives us a pure state:

$$\begin{aligned}\mathcal{E}(\rho) &= \sum_k |\psi\rangle\langle k|\rho|k\rangle\langle\psi| \\ &= \sum_k \langle k|\rho|k\rangle|\psi\rangle\langle\psi| \\ &= |\psi\rangle\langle\psi| \sum_k \langle k|\rho|k\rangle \\ &= |\psi\rangle\langle\psi| \operatorname{Tr}(\rho) = |\psi\rangle\langle\psi|\end{aligned}$$

moreover, $\sum_k E_k^\dagger E_k = \sum_k |k\rangle\langle\psi|\psi\rangle\langle k| = \sum_k |k\rangle\langle k| = \mathbb{1}$, since $|\psi\rangle$ is normalized and $\{|k\rangle\}$ is an orthonormal basis.

- **White noise (depolarizing channel)**: this channel introduces noise with some probability p

$$\mathcal{E}(\rho) = (1-p)\rho + p\frac{\mathbb{1}}{2}$$

With "introducing noise" we mean that the state becomes a fully mixed state (the 2 is because we are assuming that this is a qubit, but it can be generalized to arbitrary state spaces), losing the information ρ we had. Extracting the Kraus operators requires some more effort, but in the end one can find:

$$E_1 = \sqrt{1 - \frac{3p}{4}}\mathbb{1}, \; E_2 = \sqrt{\frac{p}{4}}X, \; E_3 = \sqrt{\frac{p}{4}}Y, \; E_4 = \sqrt{\frac{p}{4}}Z$$

We conclude the section by mentioning that the Kraus decomposition, just like the decomposition of a density matrix, is not always unique.

11.9 Stinespring dilation

Here we present a different and more intuitive way to represent an uncertain evolution, based on the environment representation we anticipated.

$$\mathcal{E}(\rho_A) \otimes |x\rangle_{E'} = U_{AE}(\rho_A \otimes |0\rangle\langle 0|_E)U_{AE}^\dagger$$

here $|0\rangle_E$ represents the initial state of the environment, called **ancillary state**. The circuit for this evolution is represented in Figure 11.1.

We have an initial state space $\mathcal{H}_A \otimes \mathcal{H}_E$ and a final state space $\mathcal{H}_B \otimes \mathcal{H}_{E'}$, such that they are **isomorphic**. If we talk about qubits, one can imagine this constraint as having the same number of qubits both in input and output to the circuit. Note that the evolution written above

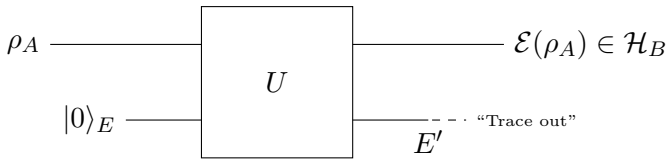


Figure 11.1: Circuit for the evolution of a density matrix using an ancillary state.

is still a reversible, unitary evolution we are already familiar with. We can define the mapping as follows:

$$\begin{aligned}\mathcal{E}_{A\mapsto B}(\rho_A) &= \mathrm{Tr}_{E'}\left(U_{AE}(\rho_A \otimes |0\rangle\langle 0|_E)U_{AE}^\dagger\right)\\ &= \mathrm{Tr}_{E'}\left(U_{AE}(\mathbb{1}_A \otimes |0\rangle_E)\rho_A(\mathbb{1}_A \otimes \langle 0|_E)U_{AE}^\dagger\right)\\ &=: \mathrm{Tr}_{E'}\left(V\rho_A V^\dagger\right)\end{aligned}$$

where $V = U_{AE}(\mathbb{1}_A \otimes |0\rangle_E)$ is an **isometry** from $\mathcal{H}_A$ to $\mathcal{H}_B \otimes \mathcal{H}_{E'}$.

From Stinespring dilation to Kraus decomposition. Let us discuss the equivalence between the two representations we obtained. We rewrite the isometry V as follows:

$$V = \sum_k E_k \otimes |k\rangle_E$$

where $\{|k\rangle\}_k$ is an orthonormal basis for $\mathcal{H}_E$. With this expression, we obtain from the fact that $V^\dagger V = \mathbb{1}$[8]:

$$\begin{aligned}\mathbb{1}_A &= V^\dagger V\\ &= \sum_{k,\ell}(E_k^\dagger \otimes \langle k|)(E_\ell \otimes |\ell\rangle)\\ &= \sum_{k,\ell} E_k^\dagger E_\ell \langle k|\ell\rangle\\ &= \sum_k E_k^\dagger E_k\end{aligned}$$

i.e. E_k are valid Kraus operators. Now let us take the partial trace:

$$\begin{aligned}\mathcal{E}(\rho_A) &= \mathrm{Tr}_{E'}(V\rho_A V^\dagger)\\ &= \mathrm{Tr}_{E'}\left(\left(\sum_k E_k \otimes |k\rangle_E\right)\rho_A\left(\sum_\ell E_k^\dagger \otimes \langle \ell|_E\right)\right)\\ &= \mathrm{Tr}_{E'}\left(\sum_{k,\ell} E_k\rho_A E_\ell^\dagger \otimes |k\rangle\langle\ell|_E\right)\\ &= \sum_j(\mathbb{1}_A \otimes \langle j|)\left(\sum_{k,\ell} E_k\rho_A E_\ell^\dagger \otimes |k\rangle\langle\ell|_E\right)(\mathbb{1}_A \otimes |j\rangle)\\ &= \sum_{j,k,\ell} E_k\rho_A E_\ell^\dagger \langle j|k\rangle\langle\ell|j\rangle\end{aligned}$$

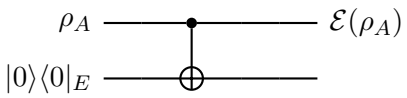
[8]Remember that, for an isometry, it holds that $||Vx||^2 = ||x||^2$, implying $V^\dagger V = \mathbb{1}$. If the isometry is unitary, then also $VV^\dagger = \mathbb{1}$ holds.

$$= \sum_k E_k \rho_A E_k^\dagger$$

which is exactly a Kraus decomposition. This derivation also gives a reason why the Kraus decomposition is not unique: we know that the partial trace here gives us only information about local states, hence different global evolutions can lead to the same local effect, while yielding possibly different Kraus representations. This is something we saw indirectly in Section 8.3, when we showed how a global evolution could be written in a different way, exploiting the linearity of the tensor product.

11.10 Example with CNOT gate

We see here a simple example of quantum channel, with a global reversible operation we know: the CNOT.



Notice that $\mathcal{H}_A \equiv \mathcal{H}_B, \mathcal{H}_E \equiv \mathcal{H}_{E'}$. The evolution is given by the CNOT gate:

$$U = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

If the first qubit is in state:

$$\rho_A = \sum_{i=0}^{1} \sum_{j=0}^{1} c_{ij} |i\rangle\langle j| = \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix}$$

The total input of the circuit is:

$$\rho_{AE} = \rho_A \otimes |0\rangle\langle 0|_E = \sum_{i=0}^{1} \sum_{j=0}^{1} c_{ij} |i\rangle\langle j|_A \otimes |0\rangle\langle 0|_E$$

$$= \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} c_{00} & 0 & c_{10} & 0 \\ 0 & 0 & 0 & 0 \\ c_{01} & 0 & c_{11} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

After some calculations we get that the evolved state is:

$$U \rho_{AE} U^\dagger = \begin{pmatrix} c_{00} & 0 & 0 & c_{10} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ c_{01} & 0 & 0 & c_{11} \end{pmatrix} = \sum_{i=0}^{1} \sum_{j=0}^{1} c_{ij} |ii\rangle\langle jj| = \sum_{i=0}^{1} \sum_{j=0}^{1} c_{ij} |i\rangle\langle j| \otimes |i\rangle\langle j|$$

and finally, we take the partial trace:

$$\mathcal{E}(\rho_A) = \mathrm{Tr}_{E'} \left(U(\rho_A \otimes |0\rangle\langle 0|) U^\dagger \right)$$

$$
\begin{aligned}
&= \mathrm{Tr}_{E'} \left(\sum_{i=0}^{1} \sum_{j=0}^{1} c_{ij} |i\rangle\langle j| \otimes |i\rangle\langle j| \right) \\
&= \sum_{i=0}^{1} \sum_{j=0}^{1} c_{ij} \, \mathrm{Tr}_{E'} \left(|i\rangle\langle j| \otimes |i\rangle\langle j| \right) && \text{linearity of partial trace} \\
&= \sum_{i=0}^{1} \sum_{j=0}^{1} c_{ij} \sum_{\ell} (\mathbb{1}_B \otimes \langle \ell|)(|i\rangle\langle j| \otimes |i\rangle\langle j|)(\mathbb{1}_B \otimes |\ell\rangle) \\
&= \sum_{i=0}^{1} \sum_{j=0}^{1} \sum_{\ell} c_{ij} |i\rangle\langle j|\langle \ell|i\rangle\langle j|\ell\rangle \\
&= \sum_{i=0}^{1} c_{ii} |i\rangle\langle i| = \begin{pmatrix} c_{00} & 0 \\ 0 & c_{11} \end{pmatrix}
\end{aligned}
$$

Thus we found that the CNOT gate does what we call a **pinch** of the matrix, which means zeroing out all the entries that are not in the diagonal:

$$
\mathcal{E} \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix} = \begin{pmatrix} c_{00} & 0 \\ 0 & c_{11} \end{pmatrix}
$$

and the Kraus representation of this is easily found (check it!):

$$
E_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}
$$

In general one can notice that, in vector notation, the partial trace can be easily computed:

$$
\mathrm{Tr}_B \begin{pmatrix} a_1 & & b_1 & \\ & a_2 & & b_2 \\ c_1 & & d_1 & \\ & c_2 & & d_2 \end{pmatrix} = \begin{pmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 \end{pmatrix}
$$

i.e. the trace is computed in blocks, in a similar way to how we distribute the tensor product.

Let us choose, to make the example more concrete, a pure state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, i.e.

$$
\rho_A = |\psi\rangle\langle\psi| = |\alpha|^2 |0\rangle\langle 0| + \alpha^* \beta |1\rangle\langle 0| + \alpha\beta^* |0\rangle\langle 1| + |\beta|^2 |1\rangle\langle 1| = \begin{pmatrix} |\alpha|^2 & \alpha^*\beta \\ \alpha\beta^* & |\beta|^2 \end{pmatrix}
$$

Pinching the matrix gives us:

$$
\begin{aligned}
\mathcal{E} \begin{pmatrix} |\alpha|^2 & \alpha^*\beta \\ \alpha\beta^* & |\beta|^2 \end{pmatrix} &= \begin{pmatrix} |\alpha^2| & 0 \\ 0 & |\beta|^2 \end{pmatrix} \\
&= |\alpha|^2 |0\rangle\langle 0| + |\beta|^2 |1\rangle\langle 1| \\
&= \mathbf{P}(0)_\rho \, |0\rangle\langle 0| + \mathbf{P}(1)_\rho \, |1\rangle\langle 1|
\end{aligned}
$$

which is not a superposition, but a probability mixture of **post-measurement states**. In particular, we can think of this as the view of the post measurement state for an observer that does only know a measurement has been performed, but not its outcome. Thus, we showed here how to schematize a projective measurement as a quantum channel, and this can be generalized (with some effort) to an arbitrary observable and an arbitrary Hilbert space.

11.11 Measurements as quantum channels

Suppose to have a Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_x$, where the latter is continuous, and consider a (discrete) observable M_A on $\mathcal{H}_A$:

$$M_A = \sum_k a_k |k\rangle\langle k|$$

We may have more complex projectors for the same label, but we keep single elements (outer products) $|k\rangle\langle k|$, for simplicity. We already saw in Section 8.3 that a possible interaction Hamiltonian to implement M_A is:

$$H = M_A \otimes P_x \Longrightarrow U(t) = \exp\left(-\frac{it}{\hbar} M_A \otimes P_x\right)$$

We also saw that, with an initial state $|\psi\rangle = \sum_k c_k |k\rangle \otimes |\psi_0\rangle$:

$$\begin{aligned} U(t)|\psi\rangle &= \sum_k c_k |k\rangle \otimes \exp\left(-\frac{it}{\hbar} M_A \otimes P_x\right) |\psi_0\rangle \\ &= \sum_k c_k |k\rangle \otimes \int_{\mathbb{R}} \psi_0(x - t a_k) |x\rangle dx \\ &=: \sum_k c_k |k\rangle \otimes |\psi_k\rangle \end{aligned}$$

We fix some time $t > 0$, and we define the pure state ρ after the evolution as:

$$\rho = |\psi(t)\rangle\langle\psi(t)| = \sum_{k,j} c_k c_j^* |k\rangle\langle j| \otimes |\psi_k\rangle\langle\psi_j|$$

Let us assume we do not have a weak measurement, which means t and a_k are chosen such that the overlap between different $|\psi_k\rangle$ is negligible. In order to see the local effect, we apply the partial trace, as before:

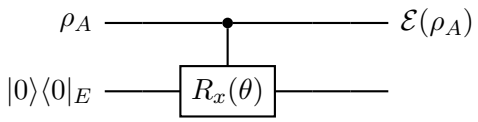
$$\begin{aligned} \mathcal{E}(\rho_A) &= \mathrm{Tr}_x \left(\sum_{k,j} c_k c_j^* |k\rangle\langle j| \otimes |\psi_k\rangle\langle\psi_j| \right) \\ &= \sum_{k,j} c_k c_j^* \, \mathrm{Tr}_x \left(|k\rangle\langle j| \otimes |\psi_k\rangle\langle\psi_j| \right) \\ &= \sum_{k,j} c_k c_j^* \int_{\mathbb{R}} \left(\mathbb{1}_A \otimes \langle x|_x \right) \left(|k\rangle\langle j| \otimes |\psi_k\rangle\langle\psi_j| \right) \left(\mathbb{1}_A \otimes |x\rangle_x \right) dx \\ &= \sum_{k,j} c_k c_j^* |k\rangle\langle j| \int_{\mathbb{R}} \langle x|\psi_k\rangle\langle\psi_j|x\rangle dx \\ &= \sum_{k,j} c_k c_j^* |k\rangle\langle j| \int_{\mathbb{R}} \langle\psi_j|x\rangle\langle x|\psi_k\rangle dx \\ &= \sum_{k,j} c_k c_j^* |k\rangle\langle j| \langle\psi_j| \left(\int_{\mathbb{R}} |x\rangle\langle x| dx \right) |\psi_k\rangle \\ &= \sum_{k,j} c_k c_j^* |k\rangle\langle j| \langle\psi_j|\psi_k\rangle \\ &\simeq \sum_k |c_k|^2 |k\rangle\langle k| \qquad\qquad \text{overlap } \langle\psi_k|\psi_{k'}\rangle \simeq \delta(k - k') \end{aligned}$$

i.e. locally we see exactly a mixture of the post-measurement states, with probabilities $|c_k|^2$. The problem here is that the assumption that the overlap is negligible is not always reasonable: a method to better model the measurement is to take into account the overlaps using the structure of the density matrix. Let us review the last step of our computation:

$$\mathcal{E}(\rho) = \sum_{k,j} c_k c_j^* |k\rangle\langle j| \langle\psi_j|\psi_k\rangle$$

If $j = k$, then the term is 1 anyway, but if $i \neq j$, then the coefficient $\langle\psi_j|\psi_k\rangle$ will be somewhere in $[0, 1]$ (in absolute value). In the diagonal of $\mathcal{E}(\rho)$ we will still see $|c_k|^2$, but the pinch of the matrix will not be perfect, i.e. the matrix can still have non-zero values outside of the diagonal. This is called **non-projective measurement**: although we cannot implement it, it is a good ideal model to work with.

Example in the discrete case. We can modify the example with the CNOT gate from the previous section, by replacing the X gate with a rotation by an arbitrary angle around the x axis of the Bloch sphere:



Depending on the rotation angle θ, we obtain different unitary evolution operators. In particular, if $\theta = 0$, nothing happens on the system and $U(t) = \mathbb{1}$. On the other hand, for $\theta = \pi$ we get exactly the $CNOT$ gate. Finally, other rotations will yield non-projective measurements with non-zero values off the diagonals.

Kraus decomposition of an observable. We apply the law of total probability: with probability $\mathbf{P}(k)_\rho$ we end up with the corresponding post-measurement state (we avoid terms with probability 0):

$$\begin{aligned}\mathcal{E}(\rho) &= \sum_k \mathbf{P}(k)_\rho \frac{\Pi_k \rho \Pi_k^\dagger}{\mathbf{P}(k)_\rho} \\ &= \sum_k \Pi_k \rho \Pi_k^\dagger\end{aligned}$$

hence, $E_k = \Pi_k$ are valid Kraus operators since:

$$\sum_k \Pi_k^\dagger \Pi_k = \sum_k \Pi_k^2 = \sum_k \Pi_k = \mathbb{1}$$

11.12 Intuition: mixed states in the Bloch sphere

In this section we show an elegant geometric property of the Bloch sphere, which is helpful to better understand what uncertainty is: we already saw in Section 1.7 that a state $|\psi\rangle$ is identified by two spherical coordinates (θ, ϕ):

$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$$

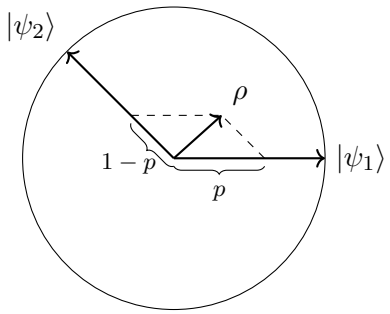


Figure 11.2: A mixed state ρ on the plane on which its composing states lie.

where $\theta \in [0, \pi), \phi \in [0, 2\pi)$. Following what we saw in this chapter, this holds for any pure state $\rho = |\psi\rangle\langle\psi|$. Now let us add a third coordinate $r \in [0, 1]$, ending up with a total coordinate system (r, θ, ϕ): this means that we completed the Bloch sphere by adding its interior points. If we have a state:

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

this state is a convex combination of pure states, i.e. it can be identified by a point within the sphere (which is exactly the convex hull of the space we are considering).

Consider for example a mixed state $\rho = p|\psi_1\rangle\langle\psi_1| + (1-p)|\psi_2\rangle\langle\psi_2|$. If for simplicity we only look at the plane on which the two pure states lie on the Bloch sphere, we get the diagram for the mixed state in Figure 11.2.

At this point, the fully mixed state (one may have imagined at this point that this is actually **unique**), corresponds exactly to the center of the sphere, i.e. when $r = 0$, while pure states are identified by (θ, ϕ) as usual when $r = 1$.

What about measurement probabilities now? We already know that a measurement basis corresponds to two opposite points of the sphere, thus a segment connecting these two points passes exactly through the origin. For example, suppose we measure with respect to Z, i.e. the computational basis $\{|0\rangle, |1\rangle\}$: the segment is exactly a vertical segment

$$\left\{(x, y, z) \in \mathbb{R}^3 \;\middle|\; x = y = 0, z \in \{-1, 1\}\right\}$$

The measurement probabilities $\mathbf{P}(0)_\rho, \mathbf{P}(1)_\rho$ induced by a state ρ can be interpreted geometrically as follows: take the point in the Bloch sphere corresponding to ρ and **project it** onto the segment. This gives a point in the Bloch sphere $\mathbf{x}$ that can be expressed as:

$$\mathbf{x} = (1-t)\mathbf{x}_0 + (t)\mathbf{x}_1$$

where $\mathbf{x}_0, \mathbf{x}_1$ are the points of $|0\rangle, |1\rangle$ on the sphere. At this point, the probability of measuring $|0\rangle$ is exactly t, and the other is $1-t$. Roughly speaking, as you can see in Figure 11.3, the closer the projected point is to one of the elements of the measurement basis, the more likely is that element to be observed.

This not only can be extended to an arbitrary measurement basis (i.e. an arbitrary diameter of the sphere), but also shows that the fully mixed state is the only one that will be exactly in the middle of any diameter, i.e. chosen **any** measurement basis, yielding a uniform distribution.

11.13 Heisenberg's uncertainty principle

We close the chapter by showing an important and famous theoretical limit to the degree of certainty we have about the state of a particle in space. For an observable A and a state ρ we define the deviation observable:

$$\Delta A = A - \langle A \rangle^\rho \mathbb{1}$$

It is easy to see that the variance of A as defined in Section 7.1 (and as introduced for the first time in any probability course) is exactly the expectation of $(\Delta A)^2$. Whenever it is not ambiguous we will denote by ΔA also the square root of ΔA^2, i.e. the standard deviation of the observable A.

We first prove some results that we will use:

Lemma 11.10. $[\Delta A, \Delta B] = [A, B]$.

Proof. For any state ρ the following holds:

$$\begin{aligned}[\Delta A, \Delta B] &= [A - \langle A \rangle^\rho \mathbb{1}, B - \langle B \rangle^\rho \mathbb{1}] \\ &= [A, B] - \langle A \rangle^\rho [\mathbb{1}, B] - \langle B \rangle^\rho [A, \mathbb{1}] + \langle A \rangle^\rho \langle B \rangle^\rho [\mathbb{1}, \mathbb{1}] \\ &= [A, B]\end{aligned}$$

since commutator is bilinear and everything commutes with the identity. □

Lemma 11.11. *For any two operators A, B, $AB = \frac{1}{2}[A, B] + \frac{1}{2}\{A, B\}$, where*

$$\begin{aligned}[A, B] &:= AB - BA \\ \{A, B\} &:= AB + BA\end{aligned}$$

are respectively the commutator and the anti-commutator of A, B.

Proof. $[A, B] + \{A, B\} = AB - BA + AB + BA = 2AB$ □

Lemma 11.12. *For any two Hermitian operators A, B:*

1. *$[A, B]$ is anti-Hermitian;*
2. *$\{A, B\}$ is Hermitian;*

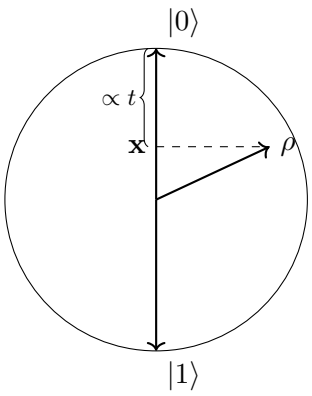


Figure 11.3: Measuring a mixed state: interpretation on the Bloch sphere.

3. $\langle\psi|[A,B]|\psi\rangle$ *is purely imaginary.*

Proof. We prove the three claims separately:

1. $[A,B]^\dagger = (AB-BA)^\dagger = (AB)^\dagger - (BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = BA - AB = -[A,B]$
2. $\{A,B\}^\dagger = (AB+BA)^\dagger = (AB)^\dagger + (BA)^\dagger = B^\dagger A^\dagger + A^\dagger B^\dagger = BA + AB = \{A,B\}$
3. We proved that $[A,B]^\dagger = -[A,B]$, hence
$$(\langle\psi|[A,B]|\psi\rangle)^* = \langle\psi|[A,B]^\dagger|\psi\rangle = -\langle\psi|[A,B]|\psi\rangle$$
and any z with $z^* = -z$ is purely imaginary.

□

We are ready to prove our main result:

Theorem 11.13 (Robertson, Schrödinger)**.** *Let A, B be two observables. For any state $|\psi\rangle$ the following bound holds:*

$$\Delta A\ \Delta B \geq \frac{1}{2}\left|\langle[A,B]\rangle^\psi\right|$$

Proof.

$$\begin{aligned}
\Delta A^2 \Delta B^2 &= \langle\psi|(\Delta A)^2|\psi\rangle\langle\psi|(\Delta B)^2|\psi\rangle \\
&= \langle\psi|(\Delta A)^\dagger(\Delta A)|\psi\rangle\langle\psi|(\Delta B)^\dagger(\Delta B)|\psi\rangle && \Delta A, \Delta B \text{ are Hermitian} \\
&\geq |\langle\psi|(\Delta A)(\Delta B)|\psi\rangle|^2 && \text{Cauchy-Schwartz inequality} \\
&= \left|\langle\psi|\left(\frac{1}{2}[\Delta A,\Delta B] + \frac{1}{2}\{\Delta A,\Delta B\}\right)|\psi\rangle\right|^2 && \text{by Lemma 11.11} \\
&= \frac{1}{4}\left|\langle\psi|[A,B]|\psi\rangle + \langle\psi|\{\Delta A,\Delta B\}|\psi\rangle\right|^2 && \text{by Lemma 11.10}
\end{aligned}$$

By Lemma 11.12 the first term is purely imaginary and, since $\{A,B\}$ is Hermitian, it has real labels and the expectation is real[9]. This means that the two terms are orthogonal in $\mathbb{C}$ and we can separate them:

$$|a+ib|^2 = |a|^2 + |b|^2$$

Thus, we found that:

$$\begin{aligned}
\Delta A^2 \Delta B^2 &\geq \frac{1}{4}\left|\langle\psi|[A,B]|\psi\rangle\right|^2 + \frac{1}{4}\left|\langle\psi|\{\Delta A,\Delta B\}|\psi\rangle\right|^2 \\
&\geq \frac{1}{4}\left|\langle\psi|[A,B]|\psi\rangle\right|^2
\end{aligned}$$

Taking the square root completes the proof. □

As a corollary we obtain the **Heisenberg uncertainty principle**.

Principle 11.14 (Heisenberg)**.** *For a particle in space with position operator X and momentum operator P:*

$$\Delta X\ \Delta P \geq \hbar/2$$

where $\hbar$ is Planck's constant.

[9] The expectation of a real random variable is always real, convince yourself!

Proof. It is sufficient to apply Theorem 11.13 where $[X, P] = i\hbar\mathbb{1}$ is the canonical commutation relation we derived in Section 5.3. □

This result tells us that we cannot be very sure about **both** position and momentum at the same time: if the variance of the position is very low, then the variance of the momentum will be inevitably high and vice versa. We also saw this effect implicitly when we talked about Gaussian wave packets in Section 7.5. In fact, we have seen that:

$$\Delta X = \sigma \Longrightarrow \Delta P = \frac{\hbar}{2\sigma}$$

which also proves that the bound is tight.

Chapter 12
Dynamics of Open Systems

12.1 Schrödinger equation for mixed states

In this section we generalize the Schrödinger equation to mixed states in uniform dynamics. We already know that the evolution of a state $\rho(t)$ with Hamiltonian H, over a tiny period of time δt is determined by:

$$\rho(t+\delta t) = U(\delta t)\ \rho(t)\ U^\dagger(\delta t)$$

and this is because of uniform dynamics. Now let $\delta t \to 0$ and take the Taylor expansion of $U(\delta t)$:

$$U(\delta t) = \exp\left(-\frac{i\ \delta t}{\hbar}H\right) = \mathbb{1} - \frac{i\ \delta t}{\hbar}H + \mathcal{O}(\delta t^2)$$

Using this expansion, let us rewrite the expression above:

$$\begin{aligned}
\rho(t+\delta t) &= \left(\mathbb{1} - \frac{i\ \delta t}{\hbar}H + \mathcal{O}(\delta t^2)\right)\rho(t)\left(\mathbb{1} + \frac{i\ \delta t}{\hbar}H + \mathcal{O}(\delta t^2)\right) \\
&= \rho(t) + \frac{i\ \delta t}{\hbar}\mathbb{1}\rho(t)H - \frac{i\ \delta t}{\hbar}H\rho(t)\mathbb{1} + \mathcal{O}(\delta t^2) \\
&= \rho(t) - \frac{i\ \delta t}{\hbar}H\rho(t) + \frac{i\ \delta t}{\hbar}\rho(t)H + \mathcal{O}(\delta t^2) \\
&= \rho(t) - \frac{i\ \delta t}{\hbar}[H,\rho(t)] + \mathcal{O}(\delta t^2)
\end{aligned}$$

By rearranging the terms we obtain:

$$\frac{\rho(t+\delta t) - \rho(t)}{\delta t} = -\frac{i}{\hbar}[H,\rho(t)] = \frac{1}{i\hbar}[H,\rho(t)]$$

and, by taking the limit for $\delta t \to 0$ we obtain exactly the derivative on the left-hand side:

$$\frac{d\rho}{dt} = \frac{1}{i\hbar}[H,\rho] \tag{12.1}$$

and this is the time-independent Schrödinger equation for mixed states.

12.2 Open systems

Now we consider the notion of **open system**: the concept of "openness" in general physics implies that the system can interact with the external environment. In order to schematize this in quantum theory, we consider a Hilbert space of the form:

$$\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$$

where S is our open system and E represents the environment. Generally, $\dim \mathcal{H}_S \ll \dim \mathcal{H}_E$, i.e. the environment is typically a much larger system. If we try to analyze the evolution of the

system S, we need to take into account the influence of the environment. More concretely, this means solving the Schrödinger equation on the whole system:

$$H_{SE} \mapsto U_{SE}(t) = \exp\left(-\frac{it}{\hbar} H_{SE}\right)$$

Now here it comes the problem: H_{SE} may be dramatically large, and we may not even know all the details of the interaction enough to approximate the Hamiltonian. On the other hand, we do not need to keep track of the evolution of the state ρ_{SE} of the whole system, in particular it suffices for us to know enough of the state of the system ρ_S, i.e. the partial trace

$$\rho_S(t) = \mathrm{Tr}_E\left(U_{SE}(t)\ \rho_{SE}(0)\ U_{SE}^\dagger(t)\right)$$

However, again, computing this partial trace may be infeasible, for the reasons we already mentioned. The ideal model is to have a TPCPM $\mathcal{E}$ dependent on time which **approximates** what happens in the environment:

$$\rho_S(t) = \mathcal{E}_S(t, \rho_S(0))$$

In order for this approximation to be good enough, we will assume that the dynamics of the environment is much faster than the interaction, which means that previous correlations between the system and the environment become negligible. A concrete example is the cup of coffee: when heat is transferred from the coffee to the surrounding air, there are particles of air that receive energy. Since the particles of a gas move really fast, we can assume that the particles in the coffee interact with "new pieces of air" every time, and that it is unlikely to interact with the same particle twice.

12.3 Lindblad equation

In this section we present the Lindblad equation, which is an extension of the Schrödinger equation to the setting of the open systems. We take the idea we presented in the previous section:

$$\rho(t + \delta t) = \mathcal{E}(\delta t, \rho_S(t))$$

where, again, $\mathcal{E}(\delta t, \cdot)$ is a TPCPM for every δt. We consider a tiny variation $\delta\rho$:

$$\mathcal{E}(\delta t, \rho_S(t)) \simeq \rho(t) + \delta\rho$$

where $|\delta\rho| \ll |\rho|$. Since this is a quantum channel, we can write down its Kraus decomposition:

$$\mathcal{E}(\delta t, \rho) = \rho + \delta\rho = \sum_k A_k \rho A_k^\dagger$$

and we choose the following Kraus operators:

$$\begin{cases} A_0 = \mathbb{1} + \delta t(L_0 - iK) \\ A_k = L_k\sqrt{\delta t} & k > 0 \end{cases}$$

where K and L_k for every k are bounded operators.

Let us analyze the terms of the sum in the Kraus decomposition:

$$A_0 \rho A_0^\dagger = (\mathbb{1} + \delta t\ (L_0 - iK))\ \rho\ (\mathbb{1} + \delta t\ (L_0 + iK))$$

$$= \rho + \delta t L_0 \rho - i\delta t K\rho + \delta t \rho L_0 + i\delta t \rho K + \mathcal{O}(\delta t^2)$$
$$= \rho + \delta t\{L_0, \rho\} + i\delta t[\rho, K] + \mathcal{O}(\delta t^2)$$

where $\{A, B\} := AB + BA$ is the **anti-commutator** we first introduced in Section 11.13, which has the same bilinearity properties of the commutator (check it!). The terms with $k > 0$, on the other hand, become simply:

$$A_k \rho A_k^\dagger = L_k \rho L_k^\dagger \, \delta t$$

If we put everything together we obtain:

$$\rho(t) + \delta\rho = \sum_{k=0}^{\infty} A_k \rho A_k^\dagger$$
$$= \rho + \delta t\{L_0, \rho\} + i\delta t[\rho, K] + \delta t \sum_{k=1}^{\infty} L_k \rho L_k^\dagger + \mathcal{O}(\delta t^2)$$

Hence, remembering that $\rho + \delta\rho = \rho(t + \delta t)$ and letting $\delta t \to 0$, we are left with:

$$\rho(t + \delta t) = \rho(t) + \delta t \left(\{L_0, \rho\} + i[\rho, K] + \sum_{k=1}^{\infty} L_k \rho L_k^\dagger \right)$$
$$\frac{d\rho}{dt} = \{L_0, \rho\} - i[K, \rho] + \sum_{k=1}^{\infty} L_k \rho L_k^\dagger$$

and this last result we obtain is the **Lindblad equation**. While the anti-commutator and the Kraus decomposition are new to us, the term in the middle closely resembles what we obtained in the Schrödinger equation of Section 12.1, thus let us pick $K = H_S/\hbar$, where H_S is the Hamiltonian of our system. The Lindblad equation now becomes:

$$\frac{d\rho}{dt} = \frac{1}{i\hbar}[H_S, \rho] + \{L_0, \rho\} + \sum_{k=1}^{\infty} L_k \rho L_k^\dagger$$

This clearly shows that this is an extension of the Schrödinger equation. Indeed, let us consider an isolated system, i.e. the system S and the environment E do not interact:

$$H_{SE} = H_S \otimes \mathbb{1}_E + \mathbb{1}_S \otimes H_E \implies U_{SE}(t) = U_S(t) \otimes U_E(t)$$

implying that the evolution of our state is:

$$U_S(t) \rho_S(0) U_S^\dagger(t)$$

which means that this state satisfies the Schrödinger equation and, in turn, the Lindblad equation with $L_k \equiv 0$ for every k.

Now we would like to understand a bit more what these operators L_k look like, and in order to infer this we will take advantage of some constraints that must hold. In particular, we check that the trace is **preserved**:

$$\operatorname{Tr}\rho \equiv 1 \implies \frac{d}{dt}\operatorname{Tr}\rho = \operatorname{Tr}\left(\frac{d\rho}{dt}\right) = 0$$

where the last equality follows from linearity of trace. Now we replace the Lindblad equation here, and apply all the properties of trace we know:

$$0 = \operatorname{Tr}\left(\frac{d\rho}{dt}\right) = \frac{1}{i\hbar}\operatorname{Tr}[H_S, \rho] + \operatorname{Tr}\{L_0, \rho\} + \sum_{k=1}^{\infty} \operatorname{Tr}\left(L_k \rho L_k^\dagger\right)$$

The trace of a commutator is always zero (why?), and the cyclic property of trace also ensures us that

$$\mathrm{Tr}(L_0\rho + \rho L_0) = \mathrm{Tr}(L_0\rho) + \mathrm{Tr}(\rho L_0) = 2\,\mathrm{Tr}(L_0\rho)$$

The sum in the third term, on the other hand, yields

$$\sum_{k=1}^{\infty} \mathrm{Tr}\left(L_k \rho L_k^\dagger\right) = \sum_{k=1}^{\infty} \mathrm{Tr}\left(L_k^\dagger L_k \rho\right)$$

Putting all together we obtain that:

$$\begin{aligned}
0 &= 2\,\mathrm{Tr}(L_0\rho) + \sum_{k=1}^{\infty} \mathrm{Tr}\left(L_k^\dagger L_k \rho\right) \\
0 &= \mathrm{Tr}(L_0\rho) + \frac{1}{2}\sum_{k=1}^{\infty} \mathrm{Tr}\left(L_k^\dagger L_k \rho\right) \\
0 &= \mathrm{Tr}\left(L_0\rho + \frac{1}{2}\sum_{k=1}^{\infty} L_k^\dagger L_k \rho\right) \\
0 &= \mathrm{Tr}\left(\left(L_0 + \frac{1}{2}\sum_{k=1}^{\infty} L_k^\dagger L_k\right)\rho\right)
\end{aligned}$$

Since this must hold for any state ρ, we evince that the expression in the tuples must be the null operator, i.e.

$$L_0 + \frac{1}{2}\sum_{k=1}^{\infty} L_k^\dagger L_k = 0 \iff L_0 = -\frac{1}{2}\sum_{k=1}^{\infty} L_k^\dagger L_k$$

Hence we can replace L_0 in the Lindblad equation:

$$\frac{d\rho}{dt} = \frac{1}{i\hbar}[H_S, \rho] - \frac{1}{2}\left\{\sum_{k=1}^{\infty} L_k^\dagger L_k, \rho\right\} + \sum_{k=1}^{\infty} L_k \rho L_k^\dagger \tag{12.2}$$

and L_k for $k > 0$ are called **Lindblad operators**. Our goal now is to understand what these L_k mean and, more concretely, how to compute them for a given physical system we want to analyze.

Example with qubits. We consider our system to be a qubit:

$$\mathcal{H}_S = \mathrm{span}\{|0\rangle, |1\rangle\}$$

with degenerate Hamiltonian, i.e. $H_S = 0$ and the two states are at the same energy level. This qubit is in a generic state ρ:

$$\rho_S(0) = \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix}$$

We would like ρ to converge to $|0\rangle\langle 0|$ for any initial state (which means that the qubit is erased):

$$\lim_{t\to\infty} \rho_S(t) = |0\rangle\langle 0|$$

The elegance of Lindbladian analysis here is that we only need a single Lindblad operator to get the desired effect:

$$L_1 = \lambda|0\rangle\langle 1|$$

which means that $|1\rangle$ is always mapped to $|0\rangle$, and λ influences how fast the evolution will be. Since $L_1^\dagger L_1 = \lambda^2|1\rangle\langle 1|$ and $H_S = 0$, the Lindblad equation in this case becomes:

$$\begin{aligned}
\frac{d\rho}{dt} &= -\frac{\lambda^2}{2}\left\{|1\rangle\langle 1|, \rho\right\} + \lambda^2|0\rangle\langle 1|\rho|1\rangle\langle 0| \\
&= \lambda^2\left(-\frac{1}{2}|1\rangle\langle 1|\rho - \frac{1}{2}\rho|1\rangle\langle 1| + \langle 1|\rho|1\rangle|0\rangle\langle 0|\right) \\
&= \lambda^2\left(-\frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix} - \frac{1}{2}\begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + c_{11}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right) \\
&= -\lambda^2\begin{pmatrix} 0 & 0 \\ c_{10}/2 & c_{11}/2 \end{pmatrix} - \lambda^2\begin{pmatrix} 0 & c_{01}/2 \\ 0 & c_{11}/2 \end{pmatrix} + \lambda^2\begin{pmatrix} c_{11} & 0 \\ 0 & 0 \end{pmatrix}
\end{aligned}$$

We found that the evolved state satisfies the following linear system of differential equations:

$$\frac{d}{dt}\begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix} = \lambda^2\begin{pmatrix} c_{11} & -c_{01}/2 \\ -c_{10}/2 & -c_{11} \end{pmatrix}$$

We immediately see that $c_{10}(t), c_{01}(t), c_{11}(t)$ have an equation of the form $y' = -dy$, thus they die exponentially fast ($y(t) = y(0)e^{-ct}$). The only term surviving will be c_{00}:

$$\begin{aligned}
\frac{dc_{00}}{dt} &= \lambda^2 c_{11}(t) = \lambda^2 c_{11}(0)e^{-\lambda^2 t} \\
c_{00}(t) - c_{00}(0) &= \lambda^2\int_0^t c_{11}(0)e^{-\lambda^2 t}dt = \lambda^2\left[-c_{11}(0)\frac{e^{-\lambda^2 t}}{\lambda^2}\right]_0^t = c_{11}(0) - c_{11}(0)e^{-\lambda^2 t}
\end{aligned}$$

Hence we found that $c_{00}(t) = c_{00}(0) + c_{11}(0) - c_{11}(0)e^{-\lambda^2 t} = 1 - c_{11}(0)e^{-\lambda^2 t}$, and this is because $c_{00} + c_{11} = \operatorname{Tr}\rho = 1$. Putting all together, we found the expression for the evolution of ρ:

$$\rho(t) = \begin{pmatrix} 1 - c_{11}(0)e^{-\lambda^2 t} & c_{01}(0)e^{-\lambda^2 t/2} \\ c_{10}(0)e^{-\lambda^2 t/2} & c_{11}(0)e^{-\lambda^2 t} \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = |0\rangle\langle 0|.$$

We will see a more practical example in Section 13.7, where a qubit interacting with an environment at temperature T will thermalize through Linbladian evolution.

Chapter 13
Quantum Thermodynamics

13.1 Work

Unlike other branches of physics such as astrophysics, which were created for purely theoretical interest or a better understanding of the universe, thermodynamics was born for **engineering purposes**. In fact, in thermodynamics we talk about energy, how we can characterize incoming and outgoing flows of energy by analyzing properties of the system at a macroscopic level. For example, the average velocity (or kinetic energy) of the particles in a gas or fluid is tightly correlated to the concept of **temperature**, which is the first quantity that comes into mind when we talk about thermodynamics.

Before starting, we present (or recall) the classical definition of work, for completeness. In classical mechanics, suppose to have a particle at some point in space and, when we apply a force $\mathbf{F}$ to it, we move it by a certain distance $d\mathbf{r}$. The amount of work done by $\mathbf{F}$ is $F \cdot d\mathbf{r}$. If the particle moves along a curve $\gamma : [a, b] \to \mathbb{R}^3$ in space, we sum up all the terms with a line integral:

$$W = \int_\gamma \mathbf{F} \cdot d\mathbf{r} = \int_a^b \mathbf{F}(\mathbf{r}_\gamma(t)) \cdot \mathbf{r}'_\gamma(t)\ dt$$

Let us give an example of this: we have a hanging mass m at height h from the ground, connected to a toy train on a shelf by a pulling system, as in Figure 13.1.

One can see that the force pulling the train towards the positive x axis is the force pulling the hanging mass towards the ground, i.e. the gravitational force $m\mathbf{g}$ ($|\mathbf{g}| = g \simeq 9.81 \frac{m}{s^2}$ is the approximated gravitational acceleration). Thus the total work done by gravity to move the train is:

$$\begin{aligned} W &= \int_\gamma m\mathbf{g} \cdot d\mathbf{r} \\ &= \int_h^0 -mg\ dy && \mathbf{g} \text{ points towards negative } y \text{ axis} \\ &= mgh, \end{aligned}$$

which is exactly the loss of potential energy of the mass (work is a form of energy and has energy units). Indeed, in the initial state, the energy of the system was exactly mgh: nothing was moving, so the kinetic energy is 0, while we assume the ground $y = 0$ to be at zero potential (remember that the potential is defined up to an additive constant!).

The term **ground state** in quantum physics, which is a stationary state of the Hamiltonian with the minimum energy eigenvalue, comes from this example: when the ball reaches the ground, it is at its lowest level of potential energy, and we need to apply some external energy to it (i.e. do some work) in order to reuse it (say, by picking the ball up). In thermodynamics, we use the term work to denote some amount of well-ordered energy that we can use/redirect.

13.2 Energy of batteries

We return to quantum physics now, and we will build a system that can store energy analogously to how we stored potential energy by holding a ball at a certain height. We will call this system

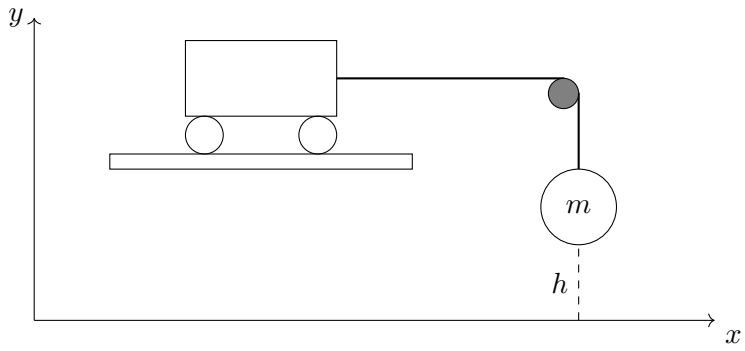


Figure 13.1: A train being pulled by a ball through gravitational force.

a **battery**, and the energy stored there **work**, if it fulfils some properties that we will see later. Consider a Hilbert space:

$$\mathcal{H}_b = \text{span}\{|k\rangle\}_{k\in\mathbb{Z}},$$

where the elements of the basis $\{|k\rangle\}_k$ correspond the different energy levels of the system, i.e. they are the eigenbasis of the Hamiltonian, which has equally spaced energy levels:

$$H_b = \sum_{k\in\mathbb{Z}} k\Delta \ |k\rangle\langle k|.$$

There, Δ is the difference of energy between levels. Note that this Hamiltonian is not physically possible to have, because it does not have a ground state (imagine the ball on the example above falling without a ground, it would lose an indefinite amount of potential energy). Actual batteries will be truncated at the bottom, but for simplicity we will for now study this case.

Also, we **discretized** the Hamiltonian here: for reasons that we will see ahead, ideally it would be convenient to have $\Delta \to 0$ and pass to a continuous basis. However in practice it's easier to have better experimental control over discrete quantum systems than continuous ones, in the same way that we use discrete bits to store and process information in computers, and not analogous continuous systems.

In the ideal case where we have a pure state $\rho_b = |\psi\rangle\langle\psi|_b = |k\rangle\langle k|_b$, we would have:

$$\langle H_b\rangle = \text{Tr}(H_b \ \rho_b) = \text{Tr}(H_b|k\rangle\langle k|_b) = k\Delta$$

and variance $(\Delta H)^2 = 0$. In a real situation, we usually do not have a pure state, but a mixed state giving us a range of possible energy levels for the battery.

How can we use the energy stored in our battery? We will spend it to implement quantum operations in other systems, much like there are schemes to convert potential gravitational energy (for example of the water in a dam) into electricity or other useful forms of energy. To see this in a simple example, take a composite system:

$$\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_b$$

where $\mathcal{H}_S$ is a qubit, and $\mathcal{H}_b$ is the battery we discussed above. In particular, the Hamiltonian of the qubit will be:

$$H_S = 0|0\rangle\langle 0| + \varepsilon|1\rangle\langle 1| = \varepsilon|1\rangle\langle 1|$$

which means that the state $|0\rangle$ has the lowest energy 0, while $|1\rangle$ reaches an amount ε of energy. Thus, the Hamiltonian of the composite system is given by the sum of the two:

$$H = H_S \otimes \mathbb{1}_b + \mathbb{1}_S \otimes H_b = \varepsilon|1\rangle\langle 1| \otimes \mathbb{1}_b + \mathbb{1}_S \otimes \sum_{k\in\mathbb{Z}} k\Delta \ |k\rangle\langle k|.$$

Now we want to understand how the battery and the qubit exchange energy. First we will impose a constraint: that the system is isolated, that is, we don't allow for energy exchanges with the environment. In practice this means that the energy inside the system is conserved, so that we can keep track of the internal energy flows between the battery and the qubit. For an analogy, if you want to keep track of the money flows between two Monopoly players, you can demand that they can only exchange money with each other, and not take from the bank.

What does "conservation of energy" mean in mathematical terms? We must demand that any evolution U_{Sb} on the joint system commute with the total Hamiltonian,

$$[U, H] = 0.$$

This implies conservation of energy since:

$$\begin{aligned}
\langle H\rangle_{U\rho U^\dagger} &= \mathrm{Tr}(HU\rho U^\dagger) \\
&= \mathrm{Tr}(UH\rho U^\dagger) && U, H \text{ commute} \\
&= \mathrm{Tr}(H\rho U^\dagger U) && \text{cyclic property of trace} \\
&= \mathrm{Tr}(H\rho) = \langle H\rangle_\rho
\end{aligned}$$

Let us see an example state $|\psi\rangle = |0\rangle \otimes |k\rangle$, i.e. S is initially in its ground state, and the battery has energy $k\Delta$. If we want to implement the Pauli bit-flip gate X_S, which switches the qubit from $|0\rangle$ to $|1\rangle$, we must transfer a quantity ε of energy from the battery to S, i.e. the battery will lose

$$q = \frac{\varepsilon}{\Delta}$$

levels of energy. Vice versa, if we want to switch a qubit from $|1\rangle$ to $|0\rangle$, the energy ε returns to the battery, which regains q levels of energy. The evolution U we need to implement is:

$$U = |0\rangle\langle 1| \otimes \sum_{k\in\mathbb{Z}} |k+q\rangle\langle k| + |1\rangle\langle 0| \otimes \sum_{k\in\mathbb{Z}} |k-q\rangle\langle k|$$

This is a way to implement the gate X that (ideally) preserves the energy:

$$\begin{aligned}
U(|0\rangle \otimes |k\rangle) &= |1\rangle \otimes |k-q\rangle \\
U(|1\rangle \otimes |k\rangle) &= |0\rangle \otimes |k+q\rangle
\end{aligned}$$

The initial energy of the qubit can be computed using the partial trace:

$$\begin{aligned}
\langle H_S\rangle^{|\psi\rangle} &= \mathrm{Tr}(H_S \otimes \mathbb{1}_b \ |\psi\rangle\langle\psi|) \\
&= \mathrm{Tr}(H_S \ \mathrm{Tr}_b(|\psi\rangle\langle\psi|)) \\
&= \mathrm{Tr}(H_S|0\rangle\langle 0|) \\
&= \mathrm{Tr}(\varepsilon|1\rangle\langle 1|0\rangle\langle 0|) = 0
\end{aligned}$$

while $\langle H_b\rangle^{|\psi\rangle} = k\Delta$, as we already derived above. The total initial energy, is:

$$\langle H\rangle^{|\psi\rangle} = \langle H_S \otimes \mathbb{1}_b + \mathbb{1}_S \otimes H_b\rangle$$

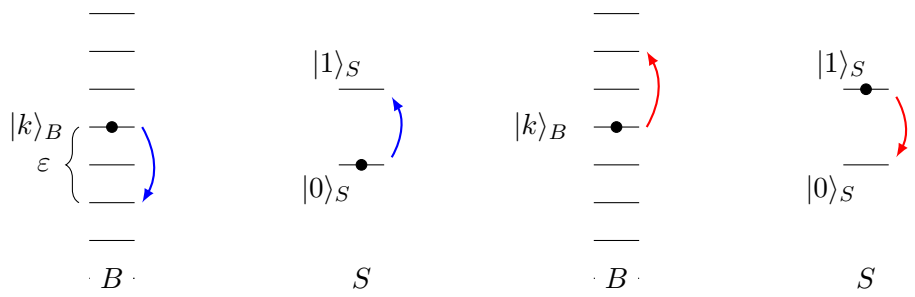


Figure 13.2: Implementing the X gate on a qubit using a battery. The inputs are $|0\rangle_S \otimes |k\rangle_B$ and $|1\rangle_S \otimes |k\rangle_B$ respectively.

$$\begin{aligned}
&= \langle H_S \otimes \mathbb{1}_b\rangle + \langle \mathbb{1}_S \otimes H_b\rangle && \text{linearity of expectation} \\
&= \mathrm{Tr}(H_S \otimes \mathbb{1}_b |\psi\rangle\langle\psi|) + \mathrm{Tr}(\mathbb{1}_S \otimes H_b |\psi\rangle\langle\psi|) \\
&= \langle H_S\rangle^{|\psi\rangle} + \langle H_b\rangle^{|\psi\rangle} \\
&= 0 + k\Delta = k\Delta
\end{aligned}$$

On the other hand, if $\rho' = U|\psi\rangle\langle\psi|U^\dagger = |1\rangle \otimes |k-q\rangle$ is the state after the evolution U:

$$\begin{aligned}
\langle H_S\rangle^{\rho'} &= \mathrm{Tr}(H_S \rho'_S) = \mathrm{Tr}(\varepsilon |1\rangle\langle 1|1\rangle\langle 1|) = \varepsilon \\
\langle H_b\rangle^{\rho'} &= \mathrm{Tr}(H_b \rho'_b) = \mathrm{Tr}(H_b |k\rangle\langle k|) = \Delta(k-q) = k\Delta - \varepsilon
\end{aligned}$$

and, by the same derivation as above, the total energy in this state is still given by the sum $\langle H_S\rangle^{\rho'} + \langle H_b\rangle^{\rho'} = \varepsilon + k\Delta - \varepsilon = k\Delta$, which is coherent with the fact that the energy is conserved. These evolutions are depicted in Figure 13.2.

Now, if we wanted the battery to be compatible with many different systems (like more qubits S_1, S_2, S_3, etc), we would need Δ to divide all the energy gaps in the Hamiltonians of those systems. This would mean that for each transition that we wanted to implement in the systems, there would be an integer number of steps in the battery that we could move up or down. The simplest way to achieve this (if we don't have much control over the Hamiltonians of those systems) is just to make the energy gap of the battery very small, $\Delta \to 0$, which is why we said above that it would be convenient to go to the continuous limit of the Hamiltonian of the battery. In practice, for a given implementation of a quantum computer, it's more realistic to fix a small but finite Δ for the battery and demand that all the other subsystems have Hamiltonians compatible with that (much like we can fix the shape of the USB port and just demand that all devices be compatible with it, instead of building adaptable chargers).

13.3 Thermodynamics and computation

We now turn our attention to more sophisticated questions, for example:

- Must computers always heat up (i.e. dissipate energy)?

Maybe a more accurate question is: how much energy does a computer need to dissipate in order to carry out some sort of computation? For this consider a quantum computer, in particular a qubit: what is its Hamiltonian? In general it depends on the implementation, but consider the simplest case, which is $H = 0$. This means that the states $|0\rangle, |1\rangle$ have the same energy level (and as a consequence, so does any superposition/mixture of them). This situation is said to have **degenerate** energy levels, which means that the two eigenstates of H have the same

eigenvalue or, for linear algebra fans, H is congruent to the identity $\mathbb{1}$. Such a situation is still possible (recall the example of the **quantum cavities** we mentioned in Section 10.1), although it requires non-trivial physical evolutions to pass from one state to another.

In this case, notice that any evolution $U_{Sb} = U_S \otimes \mathbb{1}_b$ (in the case of qubit and battery) commutes with the Hamiltonian H_S of the qubit (and thus, with the total Hamiltonian):

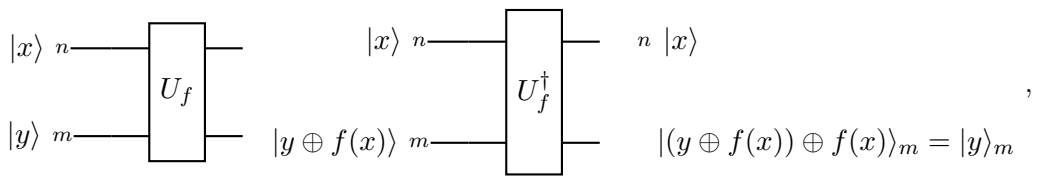
$$\begin{aligned}
[U_{Sb}, H] &= [U_S \otimes \mathbb{1}_b, H_S \otimes \mathbb{1}_b] + [U_S \otimes \mathbb{1}_b, \mathbb{1}_S \otimes H_b] && \text{linearity of commutator} \\
&= [U_S \otimes \mathbb{1}_b, H_S \otimes \mathbb{1}_b] && \text{non-interacting operators commute} \\
&= 0 && \text{since } H_S \equiv 0
\end{aligned}$$

This means that any U_S can be implemented "for free", meaning that, using this degenerate Hamiltonian with the qubit-battery construction we discussed in Section 13.2, the battery will not lose energy by switching the state of the qubit (i.e. $\varepsilon = 0$). Of course, this does not mean that computation in the practical sense will be completely conservative with respect to energy: we may have energy dissipations for clocks and control systems of the gates. But, excluding these technicalities, we can say that the term **reversible computation** used in quantum computing, not only comes from the fact that transformations applied by quantum gates are invertible, but it also refers to the thermodynamical definition of reversibility, i.e. conservation of energy. We will later see that these two notions are the same.

What about irreversible computation? What we mean this time is just computation of a non-invertible function, such as logical AND and OR[10]. It turns out that **any** integer function on n bits,

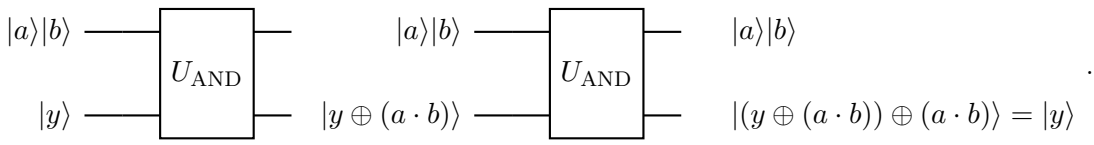
$$\begin{aligned} f: \quad & \{0,1\}^n \to \{0,1\}^m \\ & x \mapsto f(x) \end{aligned}$$

can be implemented with a reversible quantum circuit, which will have the form:



where $\oplus$ is the bit-wise addition. The unitary U_f acts on the full input (an n-qubit register) and an auxiliary m-qubit system, linearly as $U_f(|x\rangle_n \otimes |y\rangle_m) = |x\rangle_n \otimes |y \oplus f(x)\rangle_m$. Often we have $y = 0$, so that the second register has directly $|f(x)\rangle_m$. You can check through a quick calculation that $U_f^\dagger = U_f$.

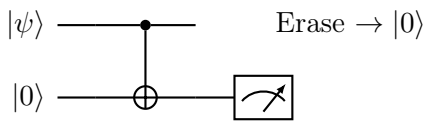
For example, in order to make a reversible AND gate, we need three qubits in total (two for the input $x = (a, b)$ and one for the output $f(x) = a \cdot b$. The gate U_{AND} that achieves this is called the **Toffoli gate**, as is essentially a CCNOT.



[10] We talk about boolean functions with respect to a fixed a basis of the Hilbert space: for example, an AND expressed in the computational basis may not necessarily act like an AND with respect to the Hadamard basis.

After we apply U_f once we have the outputs $|x\rangle_n, |y \oplus f(x)\rangle$ as above we might want to store the latter, while the former can be discarded, using a procedure of **erasure**, which disentangles the qubit from our output and safely brings it back to, say, state $|0\rangle$. This is analogous to wanting to reformat a hard-drive when you no longer need what is stored there.

Other examples of irreversible computation include measurements, and also here we can expand it as a CNOT gate as shown in the figure below, in order to move the state to the second qubit, while the first qubit can be erased like in the implementation of the AND above.



In general, we can schematize an irreversible operation as a TPCPM $\mathcal{E}$, consisting of a reversible operator acting on the input and an ancillary state (in the same way we did in Section 11.9), followed by an erasure operation $\rho \mapsto |0\rangle$.

$$\mathrm{Tr}_C \left(U_{AC}(\rho_A \otimes |0\rangle\langle 0|_C) U_{AC}^\dagger \right)$$

Erasure operations. Consider a classical example: we have a bit string s of 10 bits, and we want to reset it to all zeroes

$$\texttt{011?0100??} \mapsto \texttt{0000000000}$$

The problem is "easy" within certain limits: we flip whatever bit is set to 1 and this is still a reversible evolution, but what about bits marked as `?`? They are unknown bits, and if we set them to 0, we incur a cost C, i.e. by erasing a string s we incur a total cost of:

$$C \cdot H(s)$$

where $H(s)$ is the number of unknown bits in s, also known as its **entropy**.[11] In other words, $H(s)$ is the "amount of randomness" contained in s. We will elaborate on this "cost" and the definition of entropy later. Let us look at a quantum counterpart now:

$$\rho_s = |0\rangle\langle 0| \otimes |1\rangle\langle 1| \otimes \rho_3 \otimes |1\rangle\langle 1|$$

In this case, the reversible evolution is given by $\mathbb{1} \otimes X \otimes \mathbb{1} \otimes X$, where X is the Pauli matrix. All we are left to do is to apply some erasure protocol in order to transform $\rho_3 \mapsto |0\rangle\langle 0|$.

Consider the simplest example where ρ is the fully mixed state $\mathbb{1}/2$:

$$\rho_{Sb} = \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} \otimes |k\rangle\langle k|$$

This state cannot be transformed into $|0\rangle$ with a reversible evolution U: this would imply that both $U|0\rangle = |0\rangle$ and $U|1\rangle = |0\rangle$, i.e. the transformation is not even injective.

13.4 Landauer's principle

Before trying to find a way to physically erase information, however, we need to clarify what the "cost" we defined in the previous section is: the following gives us a clear definition of what we lose after an irreversible computation, and also creates an important bridge between thermodynamics and information theory.

[11] See Section 13.12 for a proper definition of entropy. The concept of entropy is treated more thoroughly in the QIT and AT QIT courses.

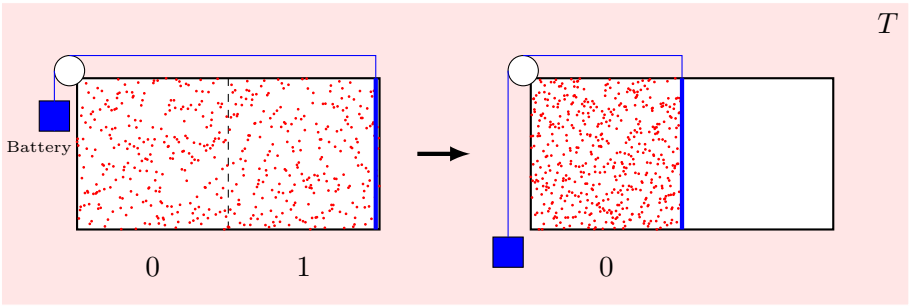


Figure 13.3: Encoding of a bit as perfect gas in a box.

Principle 13.1 (Landauer)**.** *Erasing one bit of information in an environment of temperature T releases an amount of heat:*

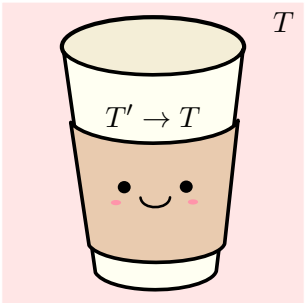
$$\Delta E = k_B T \ln 2$$

where $k_B \simeq 10^{-23} \frac{J}{K}$ is Boltzmann's constant.

To give an intuition about Landauer's principle, consider a bit encoded as depicted in Figure 13.3. We have two cells representing the states 0 and 1 respectively, and these cells may be filled with (perfect) gas: if the gas is all in the left cell, then we say that the bit is in state 0, and on the opposite case we say that the bit is in state 1. If the bit has an unknown value, then we can see this situation as the gas being distributed in the two cells. In order to reset the bit to, say, 0, we need to **compress** the gas, and this takes us some work due to the pressure of the gas. By the law of perfect gases, $pV = nRT$, one can see that the pressure p applied by the gas on the walls of the cell increases linearly with its temperature T: intuitively, pushing some gas with a piston becomes harder and harder, as the volume V of the gas decreases.

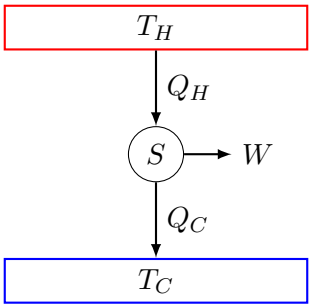
13.5 Heat baths and temperature

Consider a classical example: a cup of coffee at temperature T' in a room filled with gas at temperature T. Traditional thermodynamics tells us that if we leave the cup in the room for long enough (and the volume of the coffee is negligible with respect to the volume of the room), T' will eventually tend to T, i.e. the coffee will reach the room's temperature. In other words, the cup of coffee will be **thermalized**, and this type of system is what we call a **heat bath**. We observe that heat baths alone are useless: the evolution of the system is irreversible and we cannot extract work (i.e. energy we can use) from it.



Classical thermodynamics usually models a complex system like a gas consisting of many particles, which may have different velocities and energies, with macroscopic quantities involving some statistics, like temperature and pressure. Before continuing we give an **operational**

definition of temperature: we look at a **heat engine**, which is a system composed of two heat baths, one at a high temperature T_H, and the other one at a lower temperature T_C:



Thermodynamics tells us that, due to the difference of temperature, we have a flow of heat between the two baths (since the two temperatures will tend to become equal): the heat engine can extract some work W from the heat Q_H coming from the hot bath, while the rest of the heat Q_C reaches the cold bath. A bit of notation: we use Q to denote heat, and W to define work, but keep in mind that they are both energies and are measured in Joules $[J]$. A practical example of heat engine is the steam locomotive of a train: the hot heat bath is created by burning some coal, while the cold bath is the external air. The work extracted from the engine is transmitted to the wheels.

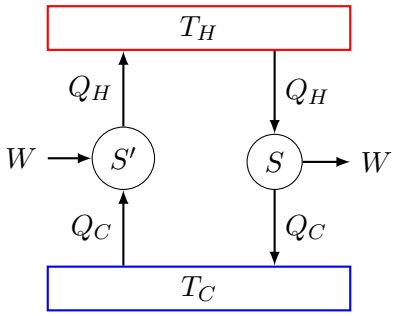
By conservation of energy we have that:

$$Q_H = W + Q_C \implies W = Q_H - Q_C$$

and then we can define the **efficiency** of the engine as follows:

$$\eta = \frac{W}{Q_H} \leq 1$$

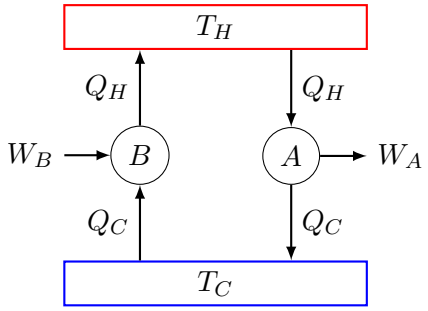
We call an engine S **reversible** if we can construct an engine S' that reverts the process, giving the heat back to the hot bath, if we give it the same work we extracted from S:



One nice property of reversible engines is the following:

Theorem 13.2. *Reversible engines between two given heat baths all have the same efficiency.*

Proof. Suppose for a contradiction that we have two reversible engines A, B with efficiencies $\eta_A > \eta_B$. We use these two engines as follows:

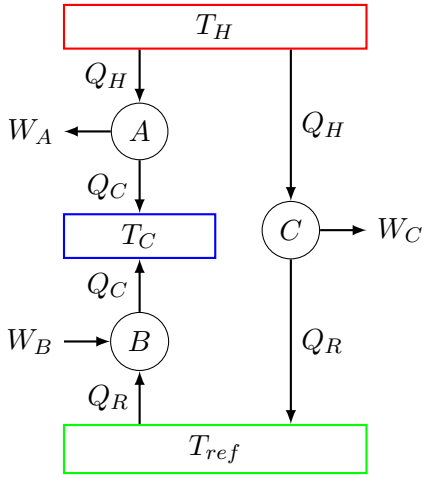


This scheme tells us that we introduce some work W_B and extract some other work W_A. The difference between these amounts of work will be:

$$\begin{aligned}\Delta W &= W_A - W_B \\ &= \eta_A Q_H - \eta_B Q_H \qquad \text{definition of efficiency} \\ &= Q_H(\eta_A - \eta_B) > 0\end{aligned}$$

which means that we constructed an engine which gives us free energy, i.e. a perpetual motion machine. Contradiction. □

Now consider three baths, at temperatures $T_H > T_C > T_{ref}$.



We add three heat engines: one from T_H to T_C, giving us some work W_A; one from T_{ref} to T_C, absorbing some work W_B; one from T_H to T_{ref}, giving W_C. For simplicity we set $T_{ref} = 1$ $[K]$. By conservation of energy we have:

$$W_A = W_C - W_B$$

One can check that this also holds from the definitions of heat engines:

$$\begin{cases} W_A = Q_H - Q_C \\ W_B = Q_C - Q_R \\ W_C = Q_H - Q_R \end{cases} \implies Q_H - Q_C = (Q_H - Q_R) - (Q_C - Q_R)$$

We define the amount of heat Q as a function of the temperature:

$$Q = Q_R \cdot f(T) =: Q_R \cdot T$$

Notice that we can define $f(T) := T$ for convenience because we have not defined what the temperature exactly is yet. Thus, from the definitions

$$\begin{aligned} W_C &= Q_H - Q_R = Q_R(T_H - 1) \\ W_B &= Q_C - Q_R = Q_R(T_C - 1) \\ W_A &= W_C - W_B = Q_R(T_H - 1) - Q_R(T_C - 1) = Q_R(T_H - T_C) \end{aligned}$$

At this point, we have an expression for the efficiency of engines A, C:

$$\begin{aligned} \eta_A &= \frac{W_A}{Q_H} = \frac{T_H - T_C}{T_H} \\ \eta_C &= \frac{W_C}{Q_C} = \frac{T_H - 1}{T_H} \end{aligned}$$

Hence we found a definition for T_H:

$$T_H = \frac{1}{1 - \eta_C} \equiv \frac{1}{1 - \eta}$$

and η is uniquely defined by Theorem 13.2 as long as we take a reversible engine for C.

We derive the following conclusions:

- For every reversible heat engine between two temperatures T_H, T_C:

$$\eta = \frac{W}{Q_H} = \frac{T_H - T_C}{T_H}$$

- The temperature of a bath can be defined by setting $T_C = 1\ [K]$ in a reversible heat engine, at this point we have:

$$T = T_H = \frac{1}{1 - \eta}$$

- We cannot extract energy from a single bath. A single bath can be seen as two baths of the same temperature T, and the efficiency of any heat engine acting on these two baths would be:

$$\eta = \frac{T - T}{T} = 0$$

  i.e. no work is extracted (this is called **passivity**).

13.6 Quantum thermal states

We need to translate the ideas we presented in the previous sections into quantum theory. For this purpose, consider a Hilbert space $\mathcal{H}_A$, with Hamiltonian H_A and a state $\tau_A \equiv \tau_A(H_A, T_A)$. We would like to build a (reversible) heat engine between two such systems $\mathcal{H}_A, \mathcal{H}_B$ in states $\tau_A(H_A, T_A), \tau_B(H_B, T_B)$ with efficiency:

$$\eta = \frac{T_B - T_A}{T_B}$$

in particular, the efficiency should not depend on the form of the Hamiltonians, and it should present the same behaviour as classical heat baths, which means that, for $T_A = T_B$, no energy can be extracted.

It turns out that all the properties we listed are satisfied by the following state:

$$\tau_A(H_A, T_A) = \exp\left(-\frac{1}{k_B T_A} H_A\right) / \operatorname{Tr}\left[\exp\left(-\frac{1}{k_B T_A} H_A\right)\right]$$

where k_B is Boltzmann's constant and the denominator normalizes the density operator, preserving Theorem 11.1. This state gives a **macroscopic** view of the system, in a similar way as we model systems in classical thermodynamics. Indeed, we never said that $\mathcal{H}_A$ was a single particle: it can be an arbitrarily complex system.

If we write the Hamiltonian as:

$$H_A = \sum_k E_k |k\rangle\langle k|$$

then the density operator can be rewritten as:

$$\tau_A(H_A, T) = \sum_k \exp\left(-\frac{1}{k_B T_A} E_k\right) |k\rangle\langle k| / \operatorname{Tr} \exp\left(-\frac{1}{k_B T_A} H_A\right)$$

This form tells us that τ_A is probabilistic mixture of energy levels: we have no superpositions here. The term $\langle k|\tau_A|k\rangle$, which is the probability to be at energy level k, is called **population** or **occupation** of the level.

Let us introduce two elements to simplify the notation:

- The **inverse temperature** $\beta = \frac{1}{k_B T}$
- The normalization factor, or **partition function**:

$$Z \equiv Z(H_A, T_A) = \operatorname{Tr} \exp\left(-\frac{1}{k_B T_A} H_A\right) = \sum_k e^{-E_k / k_B T_A}$$

With these two terms the expression for τ_A can be simplified:

$$\tau_A = \frac{1}{Z}\, e^{-\beta H_A}$$

We do some other important observations, which may be useful later:

- If $T \to 0$, i.e. we tend to the absolute zero:

  $$\tau_A(H_A, T) \to |0\rangle\langle 0|$$

  which means that the state tends to the lowest energy level. This can be seen by looking at the decomposition above, since

  $$e^{-\beta E_k} = o\left(e^{-\beta E_0}\right)$$

  for all $k > 0$.

- If $T \to \infty$, then all the exponentials tend to 1, and in particular the state tends to be fully mixed:

  $$\tau_A(H_A, T_A) \to \frac{1}{Z} \sum_k |k\rangle\langle k| = \mathbb{1}/Z$$

- States with $T < 0$, although they are not physically realizable, are said to have **population inversion** and we may want to engineer them, as they have high energy; for example we will see ahead that in heat engines there is a "virtual qubit" with $T < 0$, and this "negative temperature" characterizes the performance of the engine.

Some applications of these thermal states include:

- **Quantum chemistry**, where one has to find low energy eigenstates, which can be used to approximate τ at low temperatures, allowing to model the behaviour of the molecules of materials. This is important in **nitrogen fixation**, i.e. building fertilizers for soil, and in building **pharmaceuticals**.
- **Adiabatic quantum computing**, where we encode a problem with a Hamiltonian H that can be decomposed into local Hamiltonians. For example, consider $3SAT$, a well-known NP-complete problem in computer science: we have a formula in propositional logic in conjunctive normal form with exactly 3 variables per clause

  $$(x_1 \vee x_2 \vee \bar{x}_3) \wedge (\bar{x}_2 \vee \bar{x}_4 \vee x_5) \wedge \cdots$$

  and we want to know whether there exists an assignment of the variables $x_1, \ldots, x_n$ that satisfies the formula. We can encode this clause with a Hamiltonian and decompose it into a sum of operators regarding each of the clauses. At this point we need to evolve the system be tweaking the Hamiltonian over time in some way and look at the energy eigenvalue $\langle H \rangle^\tau$, which tells us whether a satisfying assignment exists or not[12].

Properties of a thermal state. Here we show some interesting properties of thermal states.

Theorem 13.3 (Composability)**.** *If a system of non-interacting subsystems is in thermal state, then each of the subsystems is in thermal state.*

Proof. Let $\mathcal{H}$ be the Hilbert space of the whole system:

$$\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$$

and suppose we have a thermal state:

$$\tau(T) = \frac{1}{Z} \exp(-\beta H)$$

Since the subsystems are non-interacting, we can decompose the Hamiltonian H:

$$H = \sum_i \mathbb{1}_1 \otimes \cdots \otimes \mathbb{1}_{i-1} \otimes H_i \otimes \mathbb{1}_{i+1} \otimes \cdots \otimes \mathbb{1}_n$$

These terms pairwise commute, thus we can use the property of the exponential $e^{A+B} = e^A e^B$:

$$\begin{aligned} \tau(T) &= \frac{1}{Z} \exp\left(-\beta \sum_i \mathbb{1}_1 \otimes \cdots \otimes \mathbb{1}_{i-1} \otimes H_i \otimes \mathbb{1}_{i+1} \otimes \cdots \otimes \mathbb{1}_n\right) \\ &= \frac{1}{Z} \prod_i \exp(-\beta \mathbb{1}_1 \otimes \cdots \otimes \mathbb{1}_{i-1} \otimes H_i \otimes \mathbb{1}_{i+1} \otimes \cdots \otimes \mathbb{1}_n) \end{aligned}$$

[12] If you are wondering about this: no, this does not give us a model of computation that solves $3SAT$ in polynomial time, as H can become exponentially large.

$$
\begin{aligned}
&= \frac{1}{Z}\prod_i \exp\left(\mathbb{1}_1 \otimes \cdots \otimes \mathbb{1}_{i-1} \otimes (-\beta H_i) \otimes \mathbb{1}_{i+1} \otimes \cdots \otimes \mathbb{1}_n\right) \\
&= \frac{1}{Z}\prod_i \sum_{k=0}^{\infty} \frac{1}{k!}\left(\mathbb{1}_1 \otimes \cdots \otimes \mathbb{1}_{i-1} \otimes (-\beta H_i) \otimes \mathbb{1}_{i+1} \otimes \cdots \otimes \mathbb{1}_n\right)^k \\
&= \frac{1}{Z}\prod_i \left(\mathbb{1}_1 \otimes \cdots \otimes \mathbb{1}_{i-1} \otimes \left(\sum_{k=0}^{\infty} \frac{1}{k!}(-\beta H_i)^k\right) \otimes \mathbb{1}_{i+1} \otimes \cdots \otimes \mathbb{1}_n\right) \\
&= \frac{1}{Z}\prod_i \left(\mathbb{1}_1 \otimes \cdots \otimes \mathbb{1}_{i-1} \otimes \exp(-\beta H_i) \otimes \mathbb{1}_{i+1} \otimes \cdots \otimes \mathbb{1}_n\right) \\
&= \frac{1}{Z}\exp(-\beta H_1) \otimes \cdots \otimes \exp(-\beta H_n)
\end{aligned}
$$

We found that the state is decomposed into single thermal states. □

This argument can be extended with an approximation for weakly interacting subsystems, but we redirect to the course *Advanced Topics in Quantum Information Theory* for this result.

Another important property of thermal states is typicality.

Definition 13.4 (Typicality)**.** *A Hamiltonian is typical if any pure state tends to a thermal state upon evolution after a sufficiently large time, i.e. if the evolution is:*

$$|\psi(t)\rangle = \exp\left(-\frac{it}{\hbar}H\right)|\psi(0)\rangle$$

Then we have that most subsystems appear thermalized:

$$\rho_S(t) = \mathrm{Tr}_E(|\psi(t)\rangle\langle\psi(t)|) \to \tau(H_S, T)$$

for $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$, $\dim \mathcal{H}_S \ll \dim \mathcal{H}_E$.

The last assumption about the dimensions of the Hilbert space means that we are looking only at a small portion of the system (for example, a few particles of the gas), and we see it tending to thermalize. This property is very common in "natural Hamiltonians" such as gases, liquid, metals, and in general systems with many local interactions. We refer to Section 2 of Brunner et al. [2] for some discussions about typicality.

The third property is **complete passivity**.

Definition 13.5 (Complete passivity)**.** *No work can be extracted from a single heat bath, i.e. a state of the form:*

$$\tau_A(H_A, T) \otimes \tau_B(H_B, T) \otimes \cdots$$

This property closely resembles the idea we mentioned earlier in classical physics, when we said that there is no flow of heat between two heat baths at the same temperature. A deeper analysis of this notion is treated in *Advanced Topics in Quantum Information Theory.*

13.7 Thermalization of a qubit

How can we model the thermalization process of a small system in contact with a large environment? If we know the global dynamics and global initial state we can represent the full evolution as $\rho_S(t) = \mathrm{Tr}_E(U_{SE}(t)\rho_{SE}(0)U_S E^\dagger(t))$. In practice, this is usually hard to compute, and we

can instead use what we've learned about open system dynamics do design a Lindbladian that models the local effect of the global evolution.

Consider a qubit with Hamiltonian $H_S = E|1\rangle\langle 1|$, i.e. the state $|1\rangle$ is at energy level E, in a generic state ρ:

$$\rho_S(0) = \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix}$$

We will set up two Lindblad operators:

- $L_1 = \alpha\lambda|0\rangle\langle 1|$, which promotes the transition from $|1\rangle$ to $|0\rangle$,
- $L_2 = \beta\lambda|1\rangle\langle 0|$, which promotes the transition from $|0\rangle$ to $|1\rangle$,

where λ again influences how fast the evolutions will be and α, β are parameters we will set later. Notice that having only one of them would give us a situation similar to what we found in the example of Section 12.3. Hence, we can rewrite the Lindblad equation:

$$\frac{d\rho}{dt} = \frac{1}{i\hbar}[H_S, \rho] - \frac{1}{2}\left\{L_1^\dagger L_1 + L_2^\dagger L_2, \rho\right\} + L_1^\dagger \rho L_1 + L_2^\dagger \rho L_2$$

Let us consider the three terms separately, starting from the commutator of ρ and the Hamiltonian:

$$\begin{aligned}
\frac{1}{i\hbar}[H_S, \rho] &= \frac{E}{i\hbar}[|1\rangle\langle 1|, \rho] \\
&= \frac{E}{i\hbar}\left(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix} - \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) \\
&= \frac{E}{i\hbar}\left(\begin{pmatrix} 0 & 0 \\ c_{10} & c_{11} \end{pmatrix} - \begin{pmatrix} 0 & c_{01} \\ 0 & c_{11} \end{pmatrix}\right) = \frac{E}{i\hbar}\begin{pmatrix} 0 & -c_{01} \\ c_{10} & 0 \end{pmatrix}
\end{aligned}$$

The anti-commutator in the second term gives us:

$$\begin{aligned}
\frac{1}{2}\left\{L_1^\dagger L_1 + L_2^\dagger L_2, \rho\right\} &= \frac{1}{2}\left\{L_1^\dagger L_1, \rho\right\} + \left\{L_2^\dagger L_2, \rho\right\} \\
&= \frac{1}{2}\left\{|\alpha|^2\lambda^2|1\rangle\langle 0|0\rangle\langle 1| + |\beta|^2\lambda^2|0\rangle\langle 1|1\rangle\langle 0|, \rho\right\} \\
&= \frac{1}{2}\left\{|\alpha|^2\lambda^2|1\rangle\langle 1| + |\beta|^2\lambda^2|0\rangle\langle 0|, \rho\right\} \\
&= \frac{1}{2}\left\{\begin{pmatrix} |\beta|^2\lambda^2 & 0 \\ 0 & |\alpha|^2\lambda^2 \end{pmatrix}, \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix}\right\} \\
&= \frac{1}{2}\begin{pmatrix} |\beta|^2\lambda^2 & 0 \\ 0 & |\alpha|^2\lambda^2 \end{pmatrix}\begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix}\begin{pmatrix} |\beta|^2\lambda^2 & 0 \\ 0 & |\alpha|^2\lambda^2 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} c_{00}|\beta|^2\lambda^2 & c_{01}|\beta|^2\lambda^2 \\ c_{10}|\alpha|^2\lambda^2 & c_{11}|\alpha|^2\lambda^2 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} |\beta|^2\lambda^2 c_{00} & |\alpha|^2\lambda^2 c_{01} \\ |\beta|^2\lambda^2 c_{10} & |\alpha|^2\lambda^2 c_{11} \end{pmatrix} \\
&= \begin{pmatrix} |\beta|^2\lambda^2 c_{00} & \frac{|\alpha|^2+|\beta|^2}{2}\lambda^2 c_{01} \\ \frac{|\alpha|^2+|\beta|^2}{2}\lambda^2 c_{10} & |\alpha|^2\lambda^2 c_{11} \end{pmatrix}
\end{aligned}$$

Finally, the last term simplifies to:

$$L_1^\dagger \rho L_1 + L_2^\dagger \rho L_2 = |\alpha|^2\lambda^2|0\rangle\langle 1|\rho|1\rangle\langle 0| + |\beta|^2\lambda^2|1\rangle\langle 0|\rho|0\rangle\langle 1|$$

$$= |\alpha|^2\lambda^2 c_{11}|0\rangle\langle 0| + |\beta|^2\lambda^2 c_{00}|1\rangle\langle 1|$$
$$= \begin{pmatrix} |\alpha|^2\lambda^2 c_{11} & 0 \\ 0 & |\beta|^2\lambda^2 c_{00} \end{pmatrix}$$

Putting everything together, the Lindblad equation becomes:

$$\frac{d\rho}{dt} = \frac{E}{i\hbar}\begin{pmatrix} 0 & -c_{01} \\ c_{10} & 0 \end{pmatrix} - \begin{pmatrix} |\beta|^2\lambda^2 c_{00} & \frac{|\alpha|^2+|\beta|^2}{2}\lambda^2 c_{01} \\ \frac{|\alpha|^2+|\beta|^2}{2}\lambda^2 c_{10} & |\alpha|^2\lambda^2 c_{11} \end{pmatrix} + \begin{pmatrix} |\alpha|^2\lambda^2 c_{11} & 0 \\ 0 & |\beta|^2\lambda^2 c_{00} \end{pmatrix}$$
$$= \begin{pmatrix} |\alpha|^2\lambda^2 c_{11} - |\beta|^2\lambda^2 c_{00} & \left(-\frac{|\alpha|^2+|\beta|^2}{2}\lambda^2 - \frac{E}{i\hbar}\right) c_{01} \\ \left(-\frac{|\alpha|^2+|\beta|^2}{2}\lambda^2 + \frac{E}{i\hbar}\right) c_{10} & |\beta|^2\lambda^2 c_{00} - |\alpha|^2\lambda^2 c_{11} \end{pmatrix}$$

This gives us a linear system of differential equations:

$$\begin{cases} \frac{dc_{00}}{dt} = |\alpha|^2\lambda^2 c_{11} - |\beta|^2\lambda^2 c_{00} \\ \frac{dc_{01}}{dt} = \left(-\frac{|\alpha|^2+|\beta|^2}{2}\lambda^2 - \frac{E}{i\hbar}\right) c_{01} \\ \frac{dc_{10}}{dt} = \left(-\frac{|\alpha|^2+|\beta|^2}{2}\lambda^2 + \frac{E}{i\hbar}\right) c_{10} \\ \frac{dc_{11}}{dt} = |\beta|^2\lambda^2 c_{00} - |\alpha|^2\lambda^2 c_{11} \end{cases}$$

Notice that the equations for c_{01}, c_{10} are self-contained, implying they yield exponentials:

$$c_{01}(t) = c_{01}(0)\exp\left(-\frac{|\alpha|^2+|\beta|^2}{2}\lambda^2 t - \frac{Et}{i\hbar}\right) = c_{01}(0)\ e^{-\lambda^2(|\alpha|^2+|\beta|^2)t/2}e^{iEt/\hbar}$$
$$c_{10}(t) = c_{10}(0)\exp\left(-\frac{|\alpha|^2+|\beta|^2}{2}\lambda^2 t + \frac{Et}{i\hbar}\right) = c_{10}(0)\ e^{-\lambda^2(|\alpha|^2+|\beta|^2)t/2}e^{-iEt/\hbar}$$

and one can see they both tend to zero as $t \to \infty$, since they are exponentials with negative real part. Now we need to solve the equations for c_{00}, c_{11}:

$$\frac{d}{dt}\begin{pmatrix} c_{00} \\ c_{11} \end{pmatrix} = \lambda^2 \begin{pmatrix} -|\beta|^2 & |\alpha|^2 \\ |\beta|^2 & -|\alpha|^2 \end{pmatrix}\begin{pmatrix} c_{00} \\ c_{11} \end{pmatrix} =: \lambda^2 A \begin{pmatrix} c_{00} \\ c_{11} \end{pmatrix}$$

The solution of a homogeneous linear system of differential equations is the exponential of A:

$$\begin{pmatrix} c_{00}(t) \\ c_{11}(t) \end{pmatrix} = e^{\lambda^2 At}\begin{pmatrix} c_{00}(0) \\ c_{11}(0) \end{pmatrix}$$

We can simplify the solution above by computing the eigendecomposition of A (we obtain the decomposition of $A\lambda^2 t$ by multiplying the eigenvalues of A by $\lambda^2 t$, why?):

$$A = \begin{pmatrix} 1 & |\alpha|^2 \\ -1 & |\beta|^2 \end{pmatrix}\begin{pmatrix} -|\alpha|^2 - |\beta|^2 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & |\alpha|^2 \\ -1 & |\beta|^2 \end{pmatrix}^{-1} =: UDU^{-1}$$

One can check that this is indeed the eigendecomposition of A. The general integral of the differential system hence becomes:

$$\begin{pmatrix} c_{00}(t) \\ c_{11}(t) \end{pmatrix} = \begin{pmatrix} C_0 & C_1 \end{pmatrix} U e^{\lambda^2 Dt} U^{-1}$$
$$= C_0 e^{0\lambda^2 t}\begin{pmatrix} |\alpha|^2 \\ |\beta|^2 \end{pmatrix} + C_1 e^{-\lambda^2(|\alpha|^2+|\beta|^2)t}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$= \begin{pmatrix} C_0|\alpha|^2 + C_1 e^{-\lambda^2(|\alpha|^2+|\beta|^2)t} \\ C_0|\beta|^2 - C_1 e^{-\lambda^2(|\alpha|^2+|\beta|^2)t} \end{pmatrix}$$

Let us put everything together and let $t \to \infty$:

$$\rho(t) = \begin{pmatrix} C_0|\alpha|^2 + C_1 e^{-\lambda^2(|\alpha|^2+|\beta|^2)t} & c_{01}(0)\ e^{-\lambda^2(|\alpha|^2+|\beta|^2)t/2} e^{iEt/\hbar} \\ c_{10}(0)\ e^{-\lambda^2(|\alpha|^2+|\beta|^2)t/2} e^{-iEt/\hbar} & C_0|\beta|^2 - C_1 e^{-\lambda^2(|\alpha|^2+|\beta|^2)t} \end{pmatrix}$$
$$\to \begin{pmatrix} C_0|\alpha|^2 & 0 \\ 0 & C_0|\beta|^2 \end{pmatrix}$$

If we impose the state to converge to a thermal state, i.e.:

$$\begin{pmatrix} C_0|\alpha|^2 & 0 \\ 0 & C_0|\beta|^2 \end{pmatrix} \stackrel{!}{=} \tau(H_S, T) = \frac{1}{Z} \begin{pmatrix} 1 & 0 \\ 0 & e^{-E/k_B T} \end{pmatrix}$$

Then we set $|\alpha|^2 = 1, |\beta|^2 = e^{-E/k_B T}$, while the constant C_0 of the general integral is set to:

$$C_0 = \frac{1}{Z} = \frac{1}{1 + e^{-E/k_B T}} = \frac{1}{|\alpha|^2 + |\beta|^2}$$

Keep in mind that these constants can be changed (α, β are the coefficients of the Lindblad operators, C_0 is the constant of the general integral) while obtaining the same result at $t \to \infty$. The other constant C_1 can then be determined by plugging $t = 0$:

$$\begin{cases} c_{00}(0) = \frac{1}{Z} + C_1 \\ c_{11}(0) = \frac{e^{-E/k_B T}}{Z} - C_1 \end{cases} \implies \begin{cases} C_1 = c_{00}(0) - \frac{1}{Z} \\ C_1 = \frac{e^{-E/k_B T}}{Z} - c_{11}(0) \end{cases}$$

One can see that these two conditions on C_1 are linearly dependent, as $c_{00}(0) + c_{11}(0) = \operatorname{Tr} \rho = 1$. Hence, the final solution is:

$$\rho(t) = \begin{pmatrix} \frac{1}{Z} + \left(c_{00}(0) - \frac{1}{Z}\right) e^{-\lambda^2 Z t} & c_{01}(0)\ e^{-\lambda^2 Z t/2} e^{iEt/\hbar} \\ c_{10}(0)\ e^{-\lambda^2 Z t/2} e^{-iEt/\hbar} & \frac{e^{-E/k_B T}}{Z} + \left(c_{11}(0) - \frac{e^{-E/k_B T}}{Z}\right) e^{-\lambda^2 Z t} \end{pmatrix}$$

when choosing the following Lindblad operators for the evolution:

$$L_0 = |0\rangle\langle 1|,$$
$$L_1 = e^{-E/2k_B T} |1\rangle\langle 0|.$$

13.8 Heat flow between two baths

Let us look at how we can formalize the concept of heat engine in quantum thermodynamics. First we will start with a toy example, where we just have a cold qubit connected to a hot qubit, and we observe the heat flowing between them. Further ahead we will add other ingredients to build heat engines and quantum fridges.

Consider a 2-qubit system $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, with total Hamiltonian:

$$H = H_0 + H_{int}$$
$$H_0 = H_1 \otimes \mathbb{1} + \mathbb{1} \otimes H_2 = E|1\rangle\langle 1| \otimes \mathbb{1} + \mathbb{1} \otimes E|1\rangle\langle 1|$$
$$H_{int} = \gamma(|10\rangle\langle 01| + |01\rangle\langle 10|)$$

Let us analyze this Hamiltonian: H_0 gives a non-interactive contribution, when each qubit has energy E in state $|1\rangle$ and zero energy in state $|0\rangle$. The form of the interaction Hamiltonian, on the other hand, hints us that, when the first qubit switches from $|0\rangle$ to $|1\rangle$, the second will go from $|1\rangle$ to $|0\rangle$ and vice versa. The term γ basically determines how fast this interaction evolves (indeed it will appear at the exponent in the evolution operator).

Now suppose that the two qubits are at temperature T_H, T_C respectively, with $T_H > T_C$. This means that, since:

$$\mathbf{P}(1) \propto e^{-E/k_B T}$$

the first qubit will be in state $|1\rangle$ with higher probability. Thus, if we let the system evolve for some time, the second transition (the one identified by $|01\rangle\langle 10|$) will be more likely. In other words, we will have a flow of heat Q going from the first qubit to the second, which makes sense since the two qubits model a hot and a cold bath respectively. Let us do some computation to have a more concrete idea of the evolution. We rewrite H_0:

$$\begin{aligned} H_0 &= E(|1\rangle\langle 1| \otimes \mathbb{1} + \mathbb{1} \otimes |1\rangle\langle 1|) \\ &= E|1\rangle\langle 1| \otimes (|0\rangle\langle 0| + |1\rangle\langle 1|) + E(|0\rangle\langle 0| + |1\rangle\langle 1|) \otimes |1\rangle\langle 1| \\ &= E|10\rangle\langle 10| + E|11\rangle\langle 11| + E|01\rangle\langle 01| + E|11\rangle\langle 11| \\ &= 0|00\rangle\langle 00| + E|01\rangle\langle 01| + E|10\rangle\langle 10| + 2E|11\rangle\langle 11| \end{aligned}$$

which is something we could expect, since each qubit in state $|1\rangle$ adds E to the total energy of the system. Our initial state $\rho(0)$ is a product of two thermal states:

$$\begin{aligned} \rho(0) &= \tau(H_1, T_H) \otimes \tau(H_2, T_C) \\ &= \frac{1}{Z_H} \exp\left(-\frac{1}{k_B T_H} H_1\right) \otimes \frac{1}{Z_C} \exp\left(-\frac{1}{k_B T_C} H_2\right) \\ &= \frac{1}{Z_H Z_C} \left(|0\rangle\langle 0| + e^{-E/k_B T_H}|1\rangle\langle 1|\right) \otimes \left(|0\rangle\langle 0| + e^{-E/k_B T_C}|1\rangle\langle 1|\right) \\ &= \frac{1}{Z_H Z_C} \left(|00\rangle\langle 00| + e^{-E/k_B T_H}|10\rangle\langle 10| + e^{-E/k_B T_C}|01\rangle\langle 01| + e^{-E\left(\frac{1}{k_B T_H} + \frac{1}{k_B T_C}\right)}|11\rangle\langle 11|\right) \end{aligned}$$

This density matrix has the same eigenbasis as H_0, thus they commute, $[\rho(0), H_0] = 0$. Moreover, $[H_{int}, H_0] = 0$.

In order to see what happens in the first moments of evolution, when we disconnect the two qubits from the respective baths and let them evolve under the joint Hamiltonian, we apply the Schrödinger equation for mixed states of the previous chapter, Eq. 12.1,

$$\frac{d\rho(t)}{dt} = -\frac{i}{\hbar}[H, \rho(t)].$$

We will see that as time passes, the commutator will change, making this differential equation cumbersome to handle analytically — and is in fact more sensible to solve it numerically. Here we can solve it "numerically by hand" by evolving the density matrix through two small time steps. This is not necessarily the most elegant to do it, but it suffices to give us an intuition. After this we briefly discuss the long-term behaviour of the system.

For a very small time interval $\delta \to 0$, we can approximate the derivative to first order in time with

$$\frac{\rho(\delta) - \rho(0)}{\delta} \approx \frac{d\rho(0)}{dt}$$

$$
\begin{aligned}
&= -\frac{i}{\hbar}[H, \rho(0)] \\
&= -\frac{i}{\hbar}[H_0 + H_{int}, \rho(0)] \\
&= -\frac{i}{\hbar}\left(\underbrace{[H_0, \rho(0)]}_{=0} + [H_{int}, \rho(0)]\right).
\end{aligned}
$$

We are left to compute the commutator $[H_{int}, \rho(0)]$. This can be done directly in either Dirac or matrix notation,

$$
\begin{aligned}
[H_{int}, \rho(0)] &= \frac{\gamma}{Z_H Z_X}\left[\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{-E/k_B T_H} & 0 & 0 \\ 0 & 0 & e^{-E/k_B T_C} & 0 \\ 0 & 0 & 0 & e^{-E/k_B(1/T_H+1/T_C)} \end{pmatrix}\right] \\
&= \frac{1}{Z_H Z_X}\underbrace{\gamma(e^{-E/k_B T_H} - e^{-E/k_B T_C})}_{:=\eta}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.
\end{aligned}
$$

Plugging this into the previous expression, we obtain

$$
\begin{aligned}
\rho(\delta) &\approx \rho(0) + \frac{i\delta\eta}{\hbar Z_H Z_C}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
&= \frac{1}{Z_H Z_C}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{-E/k_B T_H} & \frac{i\delta\eta}{\hbar} & 0 \\ 0 & -\frac{i\delta\eta}{\hbar} & e^{-E/k_B T_C} & 0 \\ 0 & 0 & 0 & e^{-E/k_B(1/T_H+1/T_C)} \end{pmatrix}.
\end{aligned}
$$

After this small time step, the reduced states of the individual qubits have not changed yet,[13] but there are already changes in the global state — hidden as correlations between the two qubits, and found in the off-diagonal terms of the global density matrix.

Evolving this state by another small time step δ will already show changes in the reduced state of each qubit. We start again as

$$
\frac{\rho(2\delta) - \rho(\delta)}{\delta} \approx -\frac{i}{\hbar}[H, \rho(\delta)].
$$

This computation is slightly more involved, as the density matrix is not diagonal, and better solved by computer. We have uploaded a Mathematica notebook that does precisely this on Moodle. It takes two time steps to see that some energy transfer from the hotter to the colder qubit (in standard thermodynamics this is called a heat flow). This is computed by evolving the global state, taking the partial state to find the reduced state of the first qubit, and then computing its average energy via $\mathrm{Tr}(H_1\ \rho_1(2\delta))$, or alternatively by evaluating the local energy directly on the global state, $\mathrm{Tr}(H_1 \otimes \mathbb{1}_2\ \rho(2\delta))$, which as we saw, yields the same result.

What happens after a long time? Intuitively you may think that if you bring a colder body in contact with a hotter one, they will eventually equilibrate at the same temperature. However, in the case of very small quantum systems like two qubits, the dynamics will never equilibrate,

[13] You can check that $\rho_1(\delta) = \mathrm{Tr}_2(\rho_{12}(\delta)) = \rho_1(0) = \tau(H_1, T_H)$ through a direct calculation.

and instead keep cycling back to the original state periodically. The evolution of the local energy of the two qubits is depicted in Fig. 13.4.

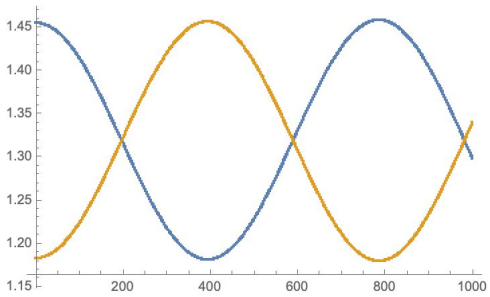


Figure 13.4: Evolution of the local energy of the first (blue) and second (orange) qubits under Schrödinger evolution, without dissipation. Here we can observe a cyclical behaviour called recurrence or ergodicity — as we have unitary evolution in a very small system, there is no long-term equilibration. This plot can be reproduced in the Mathematica file found on Moodle. For simplicity, we gave arbitrary, non-physical values to all the constants like $\hbar$ and k_B.

The intuition of why this happens is that once the hot qubit gave most of its energy to the cold one, the latter is effectively hotter, so the heat starts flowing in the opposite direction. To make this more precise, we can check the populations of the cross levels for the initial state:

$$\mathbf{P}\,(10)_{\rho(0)} = \langle 10|\rho(0)|10\rangle = \frac{1}{Z_H Z_C}\; e^{-E/k_B T_H}$$
$$\mathbf{P}\,(01)_{\rho(0)} = \langle 01|\rho(0)|01\rangle = \frac{1}{Z_H Z_C}\; e^{-E/k_B T_C}$$

For the transition $|10\rangle \to |01\rangle$ to be favoured (that is, energy flowing from the first to the second qubit), we need:

$$\mathbf{P}\,(10)_{\rho(0)} > \mathbf{P}\,(01)_{\rho(0)}$$

which is true if and only if $T_H > T_C$, as expected (in other words, if if the first qubit is actually hotter). Once these populations are reversed, the opposite transition is favoured.

This phenomenon of an evolving system returning periodically to its initial state is called *ergodicity*. The rough idea is that for a simple enough Hamiltonian there will be a time $\tilde{t}$ such that $e^{i\frac{H}{\hbar}\tilde{t}} \approx \mathbb{1}$, and so cyclically, $e^{i\frac{H}{\hbar}t} \approx e^{i\frac{H}{\hbar}(t-n\tilde{t})}$, $n \in \mathbb{N}$. For a computer science analogy, you can think of cellular automata: if your state space and the update rules are complex enough, a system can achieve non-trivial evolution (like in Conway's game of life, which is even Turing-complete), but if you create a very simple 1D automaton you can easily reach an oscillating structure whose behaviour loops on itself (more generally, think of Wolfram's Class 2 automata).

So how could we recover the classical behaviour of larger systems that equilibrate? For this we need to recouple each qubit to the respective heat bath, at temperature T_H and T_C. We model this through a Lindbladian, so that the evolution depends on both the internal Hamiltonian and the coupling to the baths. Depending on the relative weight of the Lindblad operators relative to H we may see more or less of the equilibration at the midpoint temperature.

Following what we've seen in the previous subsection, we will need two Lindblad operators to thermalize the hot qubit, and two for the cold qubit:

$$L_{0H} = \lambda_H\; |0\rangle\langle 1|_H \otimes \mathbb{1}_C,$$
$$L_{1H} = \lambda_H\; e^{-E/2k_B T_H} |1\rangle\langle 0|_H \otimes \mathbb{1}_C,$$

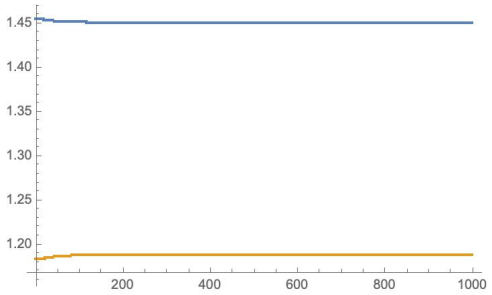


Figure 13.5: Evolution of the local energy of the first (blue) and second (orange) qubits under dissipative Lindblad evolution, where each qubit is coupled to a different heat bath. After a first short equilibration period, the local energy values stabilize. This plot can be reproduced in the Mathematica file found on Moodle. For simplicity, we gave arbitrary, non-physical values to all the constants like $\hbar$ and k_B.

$$L_{0C} = \lambda_C \ \mathbb{1}_H \otimes |0\rangle\langle 1|_C,$$
$$L_{1C} = \lambda_C \ \mathbb{1}_H \otimes e^{-E/2k_BT_C}|1\rangle\langle 0|_C,$$

where the coefficients λ_H and λ_C determine how quickly the hot and cold qubits thermalize respectively. For now we have no reason to wish one will be faster than the other, so we can set $\lambda_H = \lambda_C = \lambda$.

The Lindbladian equation (Eq. 12.2) is now

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H_0, \rho] - \frac{i}{\hbar}[H_{int}, \rho] - \frac{1}{2}\left\{ \sum_{\ell=H,C} \sum_{k=0}^{1} L_{k\ell}^\dagger L_{k\ell}, \rho \right\} + \sum_{\ell=H,C} \sum_{k=0}^{1} L_{k\ell} \rho L_{k\ell}^\dagger,$$

which is slightly more complicated than the one of the previous section where we thermalized a single qubit. However, the different terms can be easily computed in Mathematica (or equivalent software) to find the long-term behaviour of the two-qubit system analogously to what we've done before. The script can be found on Moodle, and the result in Fig. 13.5. There, we can see that there is a short transient period followed by equilibration — but not at the same energy, as each qubit is coupled to its own bath. Try playing with the coupling paremeter λ on the Mathematica script and see what happens to the final state.

As a side note, to achieve reversibility we need the energy to flow from one qubit to the other, and not stay stuck in correlations or be dissipated to the heat baths; that is, we will want that the total energy equals the sum of local terms. In particular this implies that whatever energy is lost by the first qubit can be found in the second. In turn this implies $T_H \simeq T_C$, thus in order to have little energy dissipation we would like T_H to be "just a little larger" than T_C. To learn more about this topic, see the paper by Brunner et al. [3].

13.9 Quantum fridges

Here we extend the ideas and computations above from two to three qubits, in order to create a **fridge**. In Figure 13.6 we have three qubits, with a total Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$, with energies in state $|1\rangle$ respectively E_1, E_2, E_3 and temperatures T_1, T_2, T_3, such that:

$$E_2 = E_1 + E_3, \quad T_3 > T_1, T_2$$

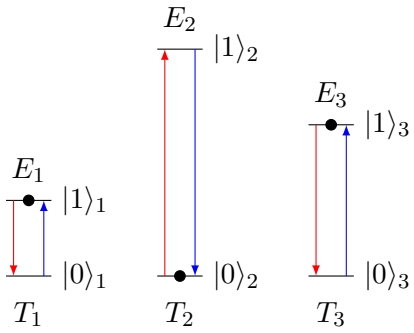


Figure 13.6: A quantum fridge with three qubits. The red evolution cools the qubits 1 and 3 while heating the second one. The blue evolution reverses the previous one.

The first condition is needed to have conservation of energy, and we would like to cool down the first qubit. The interaction Hamiltonian will be:

$$H_{int} = |101\rangle\langle 010| + |010\rangle\langle 101|$$

which means that when the second qubit goes from $|1\rangle$ to $|0\rangle$, the other two qubits must do the opposite and vice versa (this is why the first condition guarantees conservation of energy). The probabilities for the two states, by computations similar to the ones in the previous section, are:

$$\mathbf{P}\,(010)_{\rho(0)} = \frac{1}{Z_1 Z_2 Z_3}\; e^{-E_2/k_B T_2}$$
$$\mathbf{P}\,(101)_{\rho(0)} = \frac{1}{Z_1 Z_2 Z_3}\; e^{-E_1/k_B T_1}\; e^{-E_3/k_B T_3}$$

In order for the the transition we want (cooling first qubit) to be favoured we need:

$$e^{-E_1/k_B T_1}\; e^{-E_3/k_B T_3} > e^{-E_2/k_B T_2} \Longrightarrow \frac{E_1}{T_1} + \frac{E_3}{T_3} < \frac{E_2}{T_2} = \frac{E_1 + E_3}{T_2}$$

and, again, for the process to be reversible, this last constraint should hold with approximate equality. The actual evolution of the global density matrix can be computed through the Lindbladian equation, analogously to the previous section. An example can be found on the Mathematica file on Moodle. As we can see in Fig. 13.7 the target qubit (the first one) cools down after some time, stabilizing at a fixed temperature, which will depend on all the parameters of the setting. You can also try to find the final state analytically by computing the fixed point of the Lindbladian equation (that is a state that is stable under the differential equation).

13.10 Heat engines with battery

Now we would like to heat up (or cool down) a qubit as before, but this time we would like to use a battery to store the energy, as depicted in Figure 13.8. Consider a Hilbert space $\mathcal{H} = \mathcal{H}_b \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$, where $\mathcal{H}_b$ models a battery with Hamiltonian:

$$H_b = \sum_k E_1 k\; |k\rangle\langle k|$$

i.e. the gap between energy levels is E_1. The other two spaces identify two qubits with temperatures $T_2 = T_H$, $T_3 = T_C$ and energy gaps E_2, E_3 respectively. Again, we consider $E_2 = E_1 + E_3$

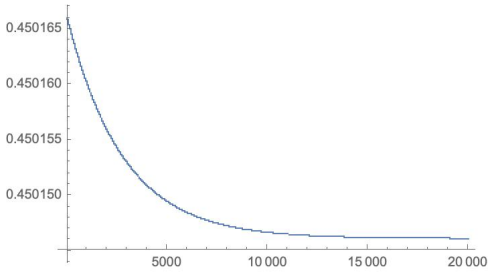


Figure 13.7: Evolution of the local energy of the target qubit under dissipative Lindblad evolution, in a quantum fridge. The qubit cools significantly, equilibrating to a fixed temperature after a while. This plot can be reproduced in the Mathematica file found on Moodle. For simplicity, we gave arbitrary, non-physical values to all the constants like $\hbar$ and k_B.

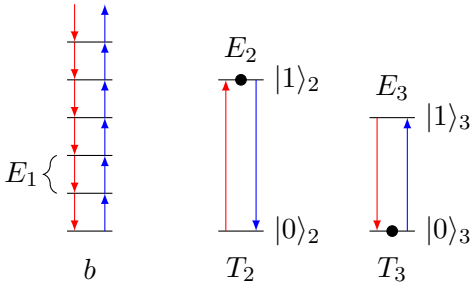


Figure 13.8: A heat engine with two qubits and a battery.

and $T_H > T_C$. The goal of this engine is to charge the battery. The interaction Hamiltonian is:

$$H_{int} = \sum_k |k+1\rangle|01\rangle\langle k|\langle 10| + \sum_k |k-1\rangle|10\rangle\langle k|\langle 01|$$

which is telling us that the energy either goes from the middle qubit to the other two subsystems (E_2 is split into E_1 for the battery and the E_3 for the rightmost qubit), or vice versa. The probabilities of the initial state are, again:

$$\mathbf{P}\,(10)_{\rho(0)} = \frac{1}{Z_2 Z_3}\, e^{-E_2/k_B T_2}$$
$$\mathbf{P}\,(01)_{\rho(0)} = \frac{1}{Z_2 Z_3}\, e^{-E_3/k_B T_3}$$

and, if we want the heat up transition to be favoured we need:

$$e^{E_2/k_B T_2} > e^{E_3/k_B T_3} \Longrightarrow \frac{T_2}{E_2} > \frac{T_3}{E_3}$$

Using this setting, we can also say something about **efficiency**: we know that, for the process to be reversible, we need $\frac{T_2}{E_2} \simeq \frac{T_3}{E_3}$, and we schematize the energy in the battery as work W (since we can actually use it). Thus, the efficiency is defined, as in the classical case:

$$\eta = \frac{W}{Q_H} = \frac{Q_H - Q_C}{Q_H} = \frac{E_2 - E_3}{E_2} \simeq \frac{T_2 - T_3}{T_2}$$

The first equality holds from conservation of energy $W = Q_H - Q_C$, and the second comes from the fact that $\frac{Q_H}{Q_C} = \frac{E_2}{E_3}$ (as E_2 is the energy coming from the hot bath, and E_3 is the energy going to the cold bath). The approximation on the right is implied by reversibility, and it also gives us an **optimal** efficiency, i.e. an upper bound on the efficiency we can obtain.

13.11 Erasing a qubit

We now have all the elements to implement our erasure operator. Suppose to have a system composed by a qubit S in fully mixed state with the two states at the same energy level, a battery with energy E, and a heat bath in thermal state $\tau(T)$. The total state is:

$$\rho = \frac{\mathbb{1}}{2} \otimes |E\rangle\langle E| \otimes \tau(T)$$

Analogously to the classical case, we would like to obtain a state, following Landauer's principle:

$$\rho' = |0\rangle\langle 0| \otimes |E - k_B T \ln 2\rangle\langle E - k_B T \ln 2| \otimes \tau(T)$$

i.e. we lose energy $k_B T \ln 2$ due to information erasure. In reality, we will never obtain this unentangled state, as real engines may leave some correlations, but the result will approximately be the same.

In order to continue, we need to better describe the heat bath and its state $\tau(T)$: we decompose the Hilbert space of the heat bath in a product space of $N+1$ qubits.

$$\mathcal{H}_H = \bigotimes_{\ell=0}^{N} \mathcal{H}_{H,i} = \bigotimes_{\ell=0}^{N} \text{span}\{|0\rangle, |1\rangle\}$$

and the Hamiltonian of the system is the sum of local Hamiltonians, i.e. qubits are non-interacting:

$$H_H = \sum_{\ell=0}^{N} \mathbb{1}_{H,0} \otimes \cdots \otimes \mathbb{1}_{H,\ell-1} \otimes H_{H,\ell} \otimes \mathbb{1}_{H,\ell+1} \otimes \cdots \otimes \mathbb{1}_{H,N} =: \sum_{\ell=0}^{N} H_{H,\ell} \otimes \mathbb{1}_{H,-\ell}$$

where each local Hamiltonian is $H_\ell = \ell \cdot \Delta \ |1\rangle\langle 1|$, which means that each qubit of the sequence has a higher energy gap than the previous one. This allows us to schematize the heat bath in a convenient way, so that, for any amount of energy the battery loses or gains we can adjust the energy of the heat bath by looking at only one of the qubits. And in fact Δ is the energy gap between energy levels of the battery:

$$H_b = \sum_{k \in \mathbb{Z}} k \cdot \Delta \ |E_k\rangle\langle E_k|$$

This is a simplifying assumption since, as we already mentioned earlier, a real battery cannot have an arbitrarily low energy (it must have a ground state), hence we are neglecting the lower bound on the energy levels.

Now we need a valid interaction Hamiltonian: we will consider a setting similar to the one of the fridge, where we have the two transitions between the states as in Figure :

$$|E_k\rangle_b |1\rangle_S |0\rangle_{H,\ell} \longleftrightarrow |E_{k-\ell}\rangle_b |0\rangle_S |1\rangle_{H,\ell}$$

where we want the transition from left to right to be more likely (i.e. we want to push the state of the qubit S towards $|0\rangle$, transferring the energy of the battery to the heat bath). These transitions can be summarized with the following unitary evolution:

$$\begin{aligned} U^{(\ell)} = &\sum_{k \in \mathbb{Z}} \left(|E_{k+\ell}\rangle_b |1\rangle_S |0\rangle_{H,\ell} \langle E_k|_b \langle 0|_S \langle 1|_{H,\ell} + |E_k\rangle_b |0\rangle_S |1\rangle_{H,\ell} \langle E_{k+\ell}|_b \langle 1|_S \langle 0|_{H,\ell} \right) \otimes \mathbb{1}_{H,-\ell} \\ &+ \mathbb{1}_b \otimes \left(|0\rangle_S |0\rangle_{H,\ell} \langle 0|_S \langle 0|_{H,\ell} + |1\rangle_S |1\rangle_{H,\ell} \langle 1|_S \langle 1|_{H,\ell} \right) \otimes \mathbb{1}_{H,-\ell} \end{aligned}$$

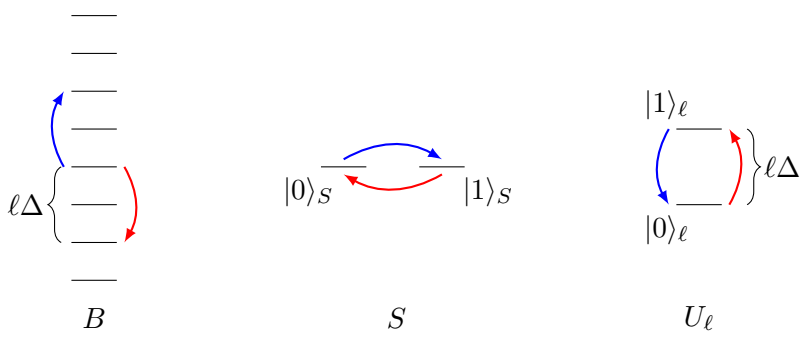


Figure 13.9: Qubit erasure using a battery and a heat bath: the red transition discharges the battery, erases the qubit in S and excites the qubit in the heat bath; the blue one, cools the heat bath, sets the qubit of interest to $|1\rangle$ and charges the battery.

One can see that $[U^{(\ell)}, H] = 0$ (convince yourself!) and indeed, since qubit S has the same energy level for the two states, this transition conserves the energy.

Let us analyze the first interaction $U^{(1)}$. We have an initial state:

$$\begin{aligned}\rho &= |E_0\rangle\langle E_0|_b \otimes \frac{\mathbb{1}}{2} \otimes \tau(H_1, T) \\ &= |E_0\rangle\langle E_0|_b \otimes \frac{\mathbb{1}_S}{2} \otimes \frac{1}{Z(H_1,T)} \left(|0\rangle\langle 0| + e^{-\Delta/k_BT}|1\rangle\langle 1|\right) \\ &= \frac{1}{2Z(H_1,T)}|E_0\rangle\langle E_0|_b \otimes \left(|00\rangle\langle 00| + e^{-\Delta/k_BT}|11\rangle\langle 11| + |01\rangle\langle 01| + e^{-\Delta/k_BT}|10\rangle\langle 10|\right)\end{aligned}$$

Observe that in the line above $e^{-\Delta/k_BT} < 1$, thus the transition that heats the bath and erases the qubit ($|01\rangle\langle 01|$) is favored over the one that goes in the other direction ($|10\rangle\langle 10|$), as desired.

The evolved state can be computed as usual (note that $U = U^\dagger$). Moreover, only terms with $|E_0\rangle$ will remain, the others will cancel out.

$$\begin{aligned}\rho' &= U\rho U^\dagger \\ &= \frac{1}{2Z}|E_0\rangle\langle E_0| \otimes \left(|00\rangle\langle 00| + |11\rangle\langle 11|e^{-\Delta/k_BT}\right) \\ &+ \frac{1}{2Z}|E_1\rangle\langle E_1| \otimes |10\rangle\langle 10|e^{-\Delta/k_BT} + \frac{1}{2Z}|E_{-1}\rangle\langle E_{-1}| \otimes |01\rangle\langle 01|\end{aligned}$$

And, if we compute the partial trace:

$$\begin{aligned}\rho'_b &= \mathrm{Tr}_{SH}(\rho') \\ &= \frac{1}{2Z}|E_0\rangle\langle E_0|\,\mathrm{Tr}\left(|00\rangle\langle 00| + e^{-\Delta/k_BT}|11\rangle\langle 11|\right) \\ &+ \frac{1}{2Z}|E_1\rangle\langle E_1|\,\mathrm{Tr}\left(e^{-\Delta/k_BT}|10\rangle\langle 10|\right) \\ &+ \frac{1}{2Z}|E_{-1}\rangle\langle E_{-1}|\,\mathrm{Tr}\left(|01\rangle\langle 01|\right) \\ &= \frac{1}{2Z}|E_0\rangle\langle E_0|\left(1 + e^{-\Delta/k_BT}\right) + \frac{1}{2Z}|E_1\rangle\langle E_1|\,e^{-\Delta/k_BT} + \frac{1}{2Z}|E_{-1}\rangle\langle E_{-1}|\end{aligned}$$

We found that there is a certain probability that the energy of the battery either goes down or up by one level. Let us see what happens to the qubit S:

$$\rho'_S = \mathrm{Tr}_{bH}(\rho')$$

$$
\begin{aligned}
&= \frac{1}{2Z}|0\rangle\langle 0|_S \operatorname{Tr}\left(|E_0\rangle|0\rangle\langle E_0|\langle 0| + |E_{-1}\rangle|1\rangle\langle E_{-1}|\langle 1|\right) \\
&+ \frac{1}{2Z}|1\rangle\langle 1|_S \operatorname{Tr}\left(e^{-\Delta/k_BT}|E_0\rangle|1\rangle\langle E_0|\langle 1| + e^{-\Delta/k_BT}|E_1\rangle|0\rangle\langle E_1|\langle 0|\right) \\
&= \frac{1}{Z}|0\rangle\langle 0| + \frac{1}{Z}e^{-\Delta/k_BT}|1\rangle\langle 1|
\end{aligned}
$$

Notice that this state is exactly the state of the qubit of the heat bath had before the transformation. This means that, during this process, the qubit of the heat bath and the qubit S **swapped** their states, and one can appreciate the fact that ρ'_S is not fully mixed anymore, as it is tending to the pure state $|0\rangle\langle 0|$. If we repeat these calculations with the other qubits in the heat bath (i.e. $U^{(\ell)}$), which have higher and higher energy gaps, we make S closer and closer to $|0\rangle\langle 0|$.

13.12 Entropy

We close this chapter by giving the definition of a notion that is heavily used in thermodynamics and information theory: the **entropy**.

Definition 13.6 (Information-theoretical entropy)**.** *Let X be a random variable defined over a probability space $(\Omega, \mathcal{F}, \boldsymbol{P})$. The entropy $\mathbb{H}[X]$ of X is defined as:*

$$\mathbb{H}[X] = -\mathbb{E}\left[\log_2 p_X(X)\right]$$

where $p_X(x)$ is the probability density function of X.

The entropy is measured in **bits** (because of the base 2 of the logarithm), and in some sense it measures "how much" the random variable is uniform in the space. Here are some examples:

- If X is constant, i.e. it has a value c with probability 1, the entropy is:

  $$\mathbb{H}[X] = -\mathbf{P}\left(X = c\right)\log_2 \mathbf{P}\left(X = c\right) = -1\log_2 1 = 0$$

  in other words, X carries out zero bits of information.

- If X is taken uniformly at random in $\{1, \ldots, n\}$, then its entropy will be:

  $$\mathbb{H}[X] = -\sum_{i=1}^{n} \mathbf{P}\left(X = i\right)\log_2 \mathbf{P}\left(X = i\right) = \log_2 n$$

  or, in other words, X has maximum entropy, giving $\log_2 n$ bits of information. Indeed, if $n = 2^k$, the random variable can be seen as k bits chosen uniformly at random.

This definition can be extended to the density matrices we gave in Chapter 11, which are nothing more than distributions of states in a Hilbert space.

Definition 13.7. *Let ρ be a density matrix representing a mixed state in a Hilbert space $\mathcal{H}$ with eigenvalues $p_1, \ldots, p_n$. The entropy carried by ρ is defined as:*

$$\mathbb{H}[\rho] = -\sum_i p_i \log_2 p_i$$

One may ask: why is it necessary to use the eigenbasis? A density matrix can be decomposed in different convex combinations of states in general. What is so special about the convex combination given by the spectral decomposition is that the states of the combinations are the eigenstates of ρ, which are pairwise orthogonal (ρ is Hermitian!). This means that, taken any orthonormal basis for a measurement, the distribution defined by $p_1, \ldots, p_n$ does not have any "overlap" between states. Let us see some examples:

- If $\rho = |\psi\rangle\langle\psi|$ is a pure state, then ρ has only one eigenvalue at 1 while the others are 0, and the entropy turns out to be 0 like in the example of the constant random variable presented above.

  This also shows how possible superpositions of states **do not carry any information**.

- If $\rho = \frac{\mathbb{1}}{|\mathcal{H}|}$ is a fully mixed state, then we have that all the eigenvalues are $\frac{1}{|\mathcal{H}|}$, and the entropy is exactly $\log_2 |\mathcal{H}|$.

Below we derive other natural properties of entropy for mixed states:

Theorem 13.8. $\mathbb{H}[\rho_A \otimes \rho_B] = \mathbb{H}[\rho_A] + \mathbb{H}[\rho_B]$.

Proof. Let $\rho_A = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ and $\rho_B = \sum_j q_j |\phi_j\rangle\langle\phi_j|$ be the eigenvalues of ρ_A and ρ_B, respectively. Then we have $\rho_A \otimes \rho_B = \sum_{i,j} p_i q_j |\psi_i\rangle\langle\psi_i| \otimes |\phi_j\rangle\langle\phi_j|$, i.e. $\{p_i q_j\}_{i,j}$ are the eigenvalues of the total state:

$$\begin{aligned}
\mathbb{H}[\rho_A \otimes \rho_B] &= \sum_{i,j} p_i q_j \log_2(p_i q_j) \\
&= \sum_{i,j} p_i q_j (\log_2 p_i + \log_2 q_j) \\
&= \sum_{i,j} p_i q_j \log_2 p_i + \sum_{i,j} p_i q_j \log_2 q_j \\
&= \sum_i p_i \log_2 p_i \sum_j q_j + \sum_j q_j \log_2 q_j \sum_i p_i \\
&= \sum_i p_i \log_2 p_i + \sum_j q_j \log_2 q_j \\
&= \mathbb{H}[\rho_A] + \mathbb{H}[\rho_B]
\end{aligned}$$

where the last step follows from $\sum_i p_i = \sum_j q_j = 1$. □

Theorem 13.9. $\mathbb{H}[U\rho U^\dagger] = \mathbb{H}[\rho]$. *In other words, reversible evolution preserves entropy.*

Proof. It is sufficient to notice that a unitary only changes the eigenstates, but preserves the eigenvalues of ρ. □

From these considerations, one can extend Landauer's principle to the concept of entropy.

Principle 13.10 (Landauer)**.** *Erasing h bits of entropy at temperature T releases an amount of heat:*

$$\Delta E = h \cdot k_B T \ln 2$$

where k_B is Boltzmann's constant.

If we give a thermodynamical definition of the **entropy of a system** as the ratio between the total energy of the system and its temperature:

$$S = \frac{E}{T}$$

Landauer's principle gives the more elegant statement:

Principle 13.11 (Landauer)**.** *Erasing h bits of entropy at temperature T releases an amount of entropy:*

$$\Delta S = h \cdot k_B \ln 2$$

where k_B is Boltzmann's constant.

In other words, reducing the entropy of a bit string by a quantity h, irrevocably produces an increase of the entropy of the system by

$$\Delta S = \frac{\Delta E}{T} = h k_B \ln 2 = k_B \ln 2^h$$

and this inevitably happens when we transform a string s using an irreversible computation. In other words, the uncertainty about a state does not disappear, but it simply passes to the surrounding system.

Chapter 14
Spin and Rotation

14.1 Pauli matrices and spin observables

We have already encountered the Pauli matrices for a qubit:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

What we did not mention in the previous chapters is that X, Y, Z are the names of the axes of the Bloch sphere: one can see that each matrix performs a rotation of the Bloch sphere of π around the corresponding axis. In fact, it is possible to derive the following properties:

$$X^2 = Y^2 = Z^2 = \mathbb{1}$$

Moreover:

$$[X, Y] = 2i\ Z, \quad [Y, Z] = 2i\ X, \quad [Z, X] = 2i\ Y.$$

We have also seen that Pauli matrices can be used as observables. In particular, when we measure the spin of an electron, say in the Stern-Gerlach experiment (Section 8.4), the actual observables we implement are:

$$S_x = \frac{\hbar}{2}\ X, \quad S_y = \frac{\hbar}{2}\ Y, \quad S_z = \frac{\hbar}{2}\ Z.$$

By linearity of the commutator, we have

$$[S_x, S_y] = i\hbar\ S_z, \quad [S_y, S_z] = i\hbar\ S_x, \quad [S_z, S_x] = i\hbar\ S_y. \tag{14.1}$$

Note: there may be some inconsistency with other chapters of these lecture notes (and the handwritten notes, and the exam) where we defined $S_x = \hbar X$ and so on. The factor of $\frac{1}{2}$ makes the commutation relations easier to handle. In physical terms, the constant factor doesn't matter, so long as we keep track of it and adjust the eigenvalues accordingly: for example, for a continuous system we could define the observable $\hat{X}$ to be "position in meters" or an equally valid observable $\hat{X}' = 100\,\hat{X}$ to be "position in centimeters." However, having the factor $\frac{1}{2}$ helps connects to the conclusion that the spin of an electron can be $+\frac{\hbar}{2}$ or $-\frac{\hbar}{2}$, which will be derived ahead and is the most popular convention in physics. In future editions of these lecture notes, the notation will be more uniform.

Examples of spin systems. We have often mentioned the spin of an election throughout our discussion. Indeed, electrons orbiting around the nucleus of an atom can have different energy levels which depend, among other factors, on their orbit. However, for electrons at the same energy level, the spin is a further degree of freedom which can lead to small differences in the overall energy of the particle (depending on whether it is up or down). Another example worth mentioning is **ferromagnetism**: consider a material in which the atoms are arranged on a grid, so that the only components that are free to move are its electrons. According to the Ising model, we can decompose the Hamiltonian of the system into an interaction component and a component that depends on the individual spins:

$$H = -\sum_{i,j} J_{ij} S_i^Z \otimes S_j^Z \otimes \mathbb{1}_{-ij} - \sum_i h_i S_i^{X/2} \otimes \mathbb{1}_{-i}$$

If we place the metal into a magnetic field, since the system will thermalize (and at low temperature, this means going closer to the ground state, which has the lowest energy), we have that the spins of the electrons will tend to align with the magnetic field. Moreover, we see how,x by the same principle, the interaction component tells us that the spins tend to agree (notice the minus sign in front).

14.2 Generalized spin

We try to generalize the concept of spin to more complex systems, i.e. to Hilbert spaces with a dimension higher than 2. Let us take a Hilbert space $\mathcal{H}$ of dimension n. Suppose that in this Hilbert space there are three observables which we name S_x, S_y, S_z satisfying the commutation relations of Eq. 14.1. We will now try to find more about the structure of this Hilbert space, including the eigenstates and eigenvalues of these observables. We will see that just imposing these commutation relations is sufficient to learn almost everything about the Hilbert space.

First of all, in order to keep a structure similar to the one we use for qubits, we would like to define a "computational basis", and we take the eigenbasis of S_z for this, in analogy with the qubit. More precisely, we will have n vectors $|\psi_{m,\alpha}\rangle$ such that:

$$S_z|\psi_{m,\alpha}\rangle = m\hbar\ |\psi_{m,\alpha}\rangle$$

where $m\hbar \in \mathbb{R}$ are the eigenvalues of S_z (remember that as an observable S_z is Hermitian!), and α is needed to distinguish eigenstates associated with the same eigenvalue. Here, m will represent the spin of the state (in a qubit we would have two values for spin, i.e. the two elements of the computational basis). At this stage, we don't yet know what values m can take, and it could be any real number — as we will see, it will turn out to be always an integer or half-integer. We also don't know anything about the state $\{|\psi_{m,\alpha}\rangle\}_{m,\alpha}$ yet, just that they are eigenstates of S_z.

Now we define the **total spin** operator S^2:

$$S^2 = S_x^2 + S_y^2 + S_z^2$$

Since also S^2 is Hermitian, we can use it to decompose $\mathcal{H}$ into the eigenspaces of S^2:

$$\mathcal{H} = \bigoplus_s \mathcal{H}_s =: \bigoplus_s \operatorname{span}\left\{|\psi\rangle \in \mathcal{H} \;\middle|\; S^2|\psi\rangle = \hbar^2 f(s)|\psi\rangle\right\},$$

where we index the eigenvalues of S^2, $\hbar f(s)$ by a number s. Here $f(s)$ is a function of s to be determined later. The subspace $\mathcal{H}_s$ is said to have **quantum number** s and is thus called a **spin-s system**. From now on we will focus on one of these subspaces, and we make use of the following lemma.

Lemma 14.1. $[S^2, S_x] = [S^2, S_y] = [S^2, S_z] = 0$

Proof. We prove $[S^2, S_z] = 0$ (the arguments for the other two commutators are symmetric):

$$\begin{aligned}
[S^2, S_z] &= [S_x^2, S_z] + [S_y^2, S_z] + [S_z^2, S_z] && \text{linearity of commutator} \\
&= [S_x^2, S_z] + [S_y^2, S_z] \\
&= S_x[S_x, S_z] + [S_x, S_z]S_x + S_y[S_y, S_z] + [S_y, S_z]S_y && \text{by Theorem 7.1} \\
&= -S_x[S_z, S_x] - [S_z, S_x]S_x + S_y[S_y, S_z] + [S_y, S_z]S_y \\
&= -i\hbar S_x S_y - i\hbar S_y S_x + i\hbar S_y S_x + i\hbar S_x S_y = 0
\end{aligned}$$

□

This lemma is telling us that S^2 and S_z share a basis of eigenstates by Theorem 4.2. In particular, acting with eigenstates of S^2 with any of the operators S_x, S_y, S_z leaves the state in the same subspace $\mathcal{H}_s$.

Lemma 14.2. *Let $|\psi\rangle \in \mathcal{H}_s$ (that is, $|\psi\rangle$ is an eigenstate of S^2 with eigenvalue $\hbar^2 f(s)$). Then the state*

$$\left(\sum_{j=x,y,z} \alpha_j S_j\right)|\psi\rangle$$

is also in $\mathcal{H}_s$, for any $\alpha_j \in \mathbb{C}$.

Proof. To see this, we just need to check that the new state has the same S^2 eigenvalue as the original state. We will use the fact that S^2 commutes with the spin operators (Lemma 14.1),

$$\begin{aligned} S^2\left(\left(\sum_{j=x,y,z} \alpha_j S_j\right)|\psi\rangle\right) &= \sum_{j=x,y,z} \alpha_j S^2 S_j|\psi\rangle \\ &= \sum_{j=x,y,z} \alpha_j S_j S^2|\psi\rangle \\ &= \sum_{j=x,y,z} \alpha_j S_j\ \hbar f(s)\ |\psi\rangle \\ &= \hbar f(s)\ \left(\sum_{j=x,y,z} \alpha_j S_j\right)|\psi\rangle. \end{aligned}$$

□

However, for each subspace $\mathcal{H}_s$ we could still have several values of m. To narrow down our search for structure, we will decompose $\mathcal{H}_s$ into a direct sum of eigenspaces of S_z:

$$\mathcal{H}_s = \bigoplus_m \mathcal{H}_{s,m} =: \bigoplus_m \text{span}\left\{|\psi\rangle \in \mathcal{H} \;\middle|\; S_z|\psi\rangle = \hbar m|\psi\rangle \text{ and } S^2|\psi\rangle = \hbar^2 f(s)|\psi\rangle\right\}.$$

So by now we decomposed the whole Hilbert space into subspaces of "total spin" s, and each of those into subspaces of "z-direction spin" m,

$$\mathcal{H} = \bigoplus_s \bigoplus_m \mathcal{H}_{s,m}.$$

We still don't know what values m can take for each s, the range of s, or the structure of each subspace $\mathcal{H}_{s,m}$. To study it, we need to introduce two **ladder operators**

$$\begin{aligned} S_+ &= S_x + iS_y, \\ S_- &= S_x - iS_y. \end{aligned}$$

By Lemma 14.2, we know that acting with $S_\pm$ (read "S_+ or S_-") leaves us inside the original $\mathcal{H}_s$ subspace. We will see that (roughly) these operators move us up and down through the S_z subspaces, from $\mathcal{H}_{s,m}$ to $\mathcal{H}_{s,m\pm1}$, hence the name "ladder operators." To show this, we start by computing the commutator $[S_z, S_+]$:

$$\begin{aligned} [S_z, S_+] &= [S_z, S_x] + i[S_z, S_y] \\ &= [S_z, S_x] - i[S_y, S_z] \\ &= i\hbar S_y + \hbar S_x = \hbar S_+. \end{aligned}$$

This allows us to derive the following:

$$\begin{aligned} S_z(S_+|\psi_{m,\alpha}\rangle) &= (S_+S_z + \hbar S_+)|\psi_{m,\alpha}\rangle \\ &= S_+(S_z + \hbar\mathbb{1})|\psi_{m,\alpha}\rangle \\ &= S_+(m\hbar + \hbar)|\psi_{m,\alpha}\rangle \\ &= (m+1)\hbar(S_+|\psi_{m,\alpha}\rangle), \end{aligned}$$

which means that applying S_+ to the eigenstate associated with the value $m\hbar$ gives us an eigenstate associated to the value $(m+1)\hbar$, i.e. the spin level immediately "above". With the exact same argument we find that applying S_- gives us the opposite effect, i.e. the eigenvalue will be $(m-1)\hbar$. To sum up, we obtained that:

$$\begin{aligned} S_+|\psi_{m,\alpha}\rangle &= K|\psi_{m+1,\alpha'}\rangle \\ S_-|\psi_{m,\alpha}\rangle &= K'|\psi_{m-1,\alpha''}\rangle, \end{aligned}$$

for some constants K, K' that we will find later (and which can depend on m, s). Actually, this is not very accurate: if K or K' can turn out to be 0, and we will see later this means that m represents the **maximal/minimal** spin level in $\mathcal{H}_s$. Notice also that $\alpha', \alpha'' \neq \alpha$ in general (recall that α is just a value to distinguish linearly independent eigenstates associated with the same eigenvalue, so associating equal values of α over different values of m would not even make sense in the first place). Here we are only stating that S_+ and S_- make the state "jump" from $\mathcal{H}_{s,m}$ to $\mathcal{H}_{s,m+1}$ and $\mathcal{H}_{s,m-1}$, respectively (we may have a state not entirely contained in one eigenspace, but that can be just seen as a superposition of states at different spin levels). Now it is clear why S_+, S_- are called **ladder** operators, in particular they are also called **raising** and **lowering** operators.

Bounding the eigenvalues. We would like to find an upper bound on the possible values for m. We will show in the following that $|m| \leq s$. We start from the positive semi-definiteness of the inner product:

$$\begin{aligned} 0 &\leq \langle\psi_{m,\alpha}|(S_+)^\dagger S_+|\psi_{m,\alpha}\rangle \\ &= \langle\psi_{m,\alpha}|S_-S_+|\psi_{m,\alpha}\rangle \\ &= \langle\psi_{m,\alpha}|(S_x - iS_y)(S_x + iS_y)|\psi_{m,\alpha}\rangle \\ &= \langle\psi_{m,\alpha}|(S_x^2 + S_y^2 + i[S_x, S_y])|\psi_{m,\alpha}\rangle \\ &= \langle\psi_{m,\alpha}|(S_x^2 + S_y^2 + i[S_x, S_y])|\psi_{m,\alpha}\rangle \\ &= \langle\psi_{m,\alpha}|(S^2 - S_z^2 + i(i\hbar S_z))|\psi_{m,\alpha}\rangle \\ &= \langle\psi_{m,\alpha}|(S^2 - S_z(S_z + \hbar\mathbb{1}))|\psi_{m,\alpha}\rangle \\ &= \langle\psi_{m,\alpha}|(\hbar^2 f(s) - \hbar^2 m(m+1))|\psi_{m,\alpha}\rangle \\ &= (\hbar^2 f(s) - \hbar^2 m(m+1))\underbrace{\langle\psi_{m,\alpha}|\psi_{m,\alpha}\rangle}_{1}, \end{aligned}$$

which gives us $m(m+1) \leq f(s)$. With the exact same calculations (replacing S_+ with S_-) we find $-m(m+1) \leq f(s)$, and we can say that

$$|m(m+1)| \leq f(s).$$

If we choose $f(s) = s(s+1)$ as the spectrum of S^2 (it is an educated guess), we can deduce $|m| \leq s$. Here comes a paradox: how can we accept both the fact that m is bounded by $\pm s$ but

$S_\pm$ shift the value of m without falling in a contradiction? We must infer that for $m = s, -s$ we have

$$S_+|\psi_{s,\alpha}\rangle = 0, \quad S_-|\psi_{-s,\alpha}\rangle = 0,$$

i.e. going out of these bounds would give us the null vector. Indeed, if this was not true, S_+ would make m increase indefinitely, contradicting $m \leq s$ (and the same for S_-). Since the extremal values for m are exactly $\pm s$, and since $S_\pm$ shift m by ± 1, we also infer that m and s must always differ by exactly an integer value and, in particular:

$$m \in \{-s, -s+1, \ldots, s-1, s\}$$

Moreover, this also tells us that $2s \in \mathbb{N}$, i.e. $s = n/2$ for some natural n (hence the name **quantum** number). For example, if $s = \frac{1}{2}$, then $m \in \{-\frac{1}{2}, \frac{1}{2}\}$, if $s = 1$, then $m \in \{-1, 0, 1\}$ and so on.[14] In particular, the first example is typical of **electrons**, where $\frac{1}{2}$ denotes a "spin up", and $-\frac{1}{2}$ gives a "spin down". Particles with a integer total spin ($s \in \mathbb{N}$) are called **bosons**, while particles with $s + \frac{1}{2} \in \mathbb{N}$ are called **fermions**. Typical examples of fermions are the elements of the atom (electrons, protons, neutrons...) which constitute the matter, while in the category of bosons we can find particles like photons, important for electromagnetic interactions.[15]

Bounding the eigenspaces. Now we would like to bound the number of orthogonal eigenstates in each eigenspace or, more concretely, the number of values α for each possible eigenvalue m. We start by noticing that, since S_z is Hermitian, there exists an orthonormal basis of $\mathcal{H}$ and, in particular, for $\alpha \neq \beta$:

$$\langle\psi_{m,\alpha}|\psi_{m,\beta}\rangle = 0$$

for a suitable choice of the basis $\{|\psi_{m,\alpha}\rangle\}_\alpha$ of the eigenspace associated with m. Now, with the same calculation as above (notice that the two vectors are associated with the same eigenvalue, although they are different):

$$\begin{aligned}\langle\psi_{m,\alpha}|(S_+)^\dagger S_+|\psi_{m,\beta}\rangle &= \langle\psi_{m,\alpha}|(\hbar^2 f(s) - \hbar^2 m(m+1))|\psi_{m,\beta}\rangle \\ &= (\hbar^2 f(s) - \hbar^2 m(m+1))\langle\psi_{m,\alpha}|\psi_{m,\beta}\rangle = 0\end{aligned}$$

Meaning that S_+ preserves pairwise orthogonality from $\mathcal{H}_{s,m}$ to $\mathcal{H}_{s,m+1}$. This in particular tells us that any orthonormal basis of $\mathcal{H}_{s,m}$ is contained in an orthonormal basis of $\mathcal{H}_{s,m+1}$ and

$$\dim \mathcal{H}_{s,m} \leq \dim \mathcal{H}_{s,m+1}$$

Doing the same reasoning for S_- yields the opposite bound, concluding that $\mathcal{H}_{s,m}$ have the same dimension for all m. The fact that the number of values for α are equal for every eigenspace allows us to **decompose** the Hilbert space in a clever way:

$$\mathcal{H}_s = \mathcal{H}_{Z_s} \otimes \mathcal{H}_{*_s}$$

where $\mathcal{H}_{Z_s}$ has dimension $2s+1$ and $\mathcal{H}_{*_s}$ has dimension $d = \dim \mathcal{H}_{s,m}$ for every m. A state $|\psi_{m,\alpha}\rangle$ can then be decomposed into a tensor product as well:

$$|\psi_{m,\alpha}\rangle = |\psi_m\rangle_{Z_s} \otimes |\psi_\alpha\rangle_{*_s}$$

[14] Someone asked why the bound $|m| \leq s$ must be saturated; why can't we have the values of m even more constrained, for example between $-s+2$ and $s-3$? This is a very good question! In this case we'd need to impose $S_+|\psi_{s-3,\alpha}\rangle = 0$ and derive the rest of the calculation in the following pages accordingly. When I have distilled the answer I will add it here, but here's the intuition: remember that $S^2 = S_x^2 + S_y^2 + S_z^2$, and if you over-constrain the possible eigenvalues of S_x, S_y, S_z then their squares will be too small for S^2 to have the large eigenvalue $s(s+1)\hbar^2$. You can check this for specific implementations of spin to physical systems.

[15] Incidentally, fermions and bosons have wildly different properties (studied in QM2), which is why we bother naming this split between integer and half-integer spin-s systems.

Keep in mind that we still do not know which $|\psi_{\alpha'}\rangle$ is obtained by applying S_+ or S_- to the above product state. At this point the total Hilbert space is decomposed into the sum

$$\mathcal{H} = \bigoplus_s (\mathcal{H}_{Z_s} \otimes \mathcal{H}_*)$$

14.3 Simple spin systems

In general, $\mathcal{H}_{*_s}$ is another independent Hilbert space. More concretely, if $d = \dim \mathcal{H}_{*_s}$, the dimension of a spin-s system is $\dim \mathcal{H}_s = d \cdot (2s+1)$ (recall that m takes $2s+1$ different values since $|m| \leq s$). In the following we will focus on so-called **simple spin systems**, for which $\dim \mathcal{H}_{*_s} = 1$; later on we will see what happens if $\mathcal{H}_{*_s}$ is another spin system.

The idea is to **fix** the lowest state $|\psi_{-s}\rangle$ and then recursively define:

$$\begin{aligned} |\psi_{m+1}\rangle &= \frac{S_+|\psi_m\rangle}{\sqrt{\langle\psi_m|S_+^\dagger S_+|\psi_m\rangle}} \\ &= \frac{S_+|\psi_m\rangle}{\hbar\sqrt{s(s+1) - m(m+1)}} \end{aligned}$$

or, equivalently:

$$S_+|\psi_m\rangle = \hbar\sqrt{s(s+1) - m(m+1)}|\psi_{m+1}\rangle$$

Hence we can finally give a definition for S_+ and, with a similar argument, also one for S_-:

$$\begin{aligned} S_+ &= \hbar \sum_m \sqrt{s(s+1) - m(m+1)}|\psi_{m+1}\rangle\langle\psi_m| \\ S_- &= \hbar \sum_m \sqrt{s(s+1) - m(m-1)}|\psi_{m-1}\rangle\langle\psi_m|. \end{aligned}$$

14.4 Composing spin systems

Let us take two simple spin systems:

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$$

with spin numbers s_1, s_2 respectively. The question here is: can we consider $\mathcal{H}$ as a spin system? The answer is yes: we can define global operators S_x, S_y, S_z satisfying the necessary commutation relations, and find subspaces with total spin s that depends on s_1 and s_2.

We start by defining S_z as:

$$S_z = S_{z,1} \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes S_{z,2},$$

and S_x, S_y analogously. Indeed, using linearity of commutation:

$$\begin{aligned} [S_x, S_y] &= [S_{x,1} \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes S_{x,2},\ S_{y,1} \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes S_{y,2}] \\ &= [S_{x,1} \otimes \mathbb{1}_2,\ S_{y,1} \otimes \mathbb{1}_2] + [\mathbb{1}_1 \otimes S_{x,2},\ S_{y,1} \otimes \mathbb{1}_2] \\ &+ [S_{x,1} \otimes \mathbb{1}_2,\ \mathbb{1}_1 \otimes S_{y,2}] + [\mathbb{1}_1 \otimes S_{x,2},\ \mathbb{1}_1 \otimes S_{y,2}] \\ &= [S_{x,1} \otimes \mathbb{1}_2,\ S_{y,1} \otimes \mathbb{1}_2] + [\mathbb{1}_1 \otimes S_{x,2},\ \mathbb{1}_1 \otimes S_{y,2}] \\ &= [S_{x,1}, S_{y,1}] \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes [S_{x,2},\ S_{y,2}] \end{aligned}$$

$$= i\hbar S_{z,1} \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes i\hbar S_{z,2} = i\hbar S_z$$

and, again, the derivation for the other two commutators is identical. Thus, we found that the commutation properties for these constructed spin operators immediately follow from the properties of the spin operators of the subsystems, and the results of the previous sections are extended to the global system $\mathcal{H}$. One can see that, if $|\psi_{m_1}\rangle_1 \otimes |\psi_{m_2}\rangle_2$ is a global state where the two subsystems have spin m_1, m_2 respectively, then:

$$\begin{aligned} S_z(|\psi_{m_1}\rangle_1 \otimes |\psi_{m_2}\rangle_2) &= (S_{z,1} \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes S_{z,2})(|\psi_{m_1}\rangle_1 \otimes |\psi_{m_2}\rangle_2) \\ &= S_{z,1}|\psi_{m_1}\rangle_1 \otimes |\psi_{m_2}\rangle_2 + |\psi_{m_1}\rangle_1 \otimes S_{z,2}|\psi_{m_2}\rangle_2 \\ &= \hbar m_1 |\psi_{m_1}\rangle_1 \otimes |\psi_{m_2}\rangle_2 + |\psi_{m_1}\rangle_1 \otimes \hbar m_2 |\psi_{m_2}\rangle_2 \\ &= \hbar(m_1 + m_2)|\psi_{m_1}\rangle_1 \otimes |\psi_{m_2}\rangle_2 \end{aligned}$$

i.e. the global state is a state of spin $m_1 + m_2$ for $\mathcal{H}$, and this also tells us that, since

$$\begin{aligned} m_1 &\in \{-s_1, \ldots, +s_1\} \\ m_2 &\in \{-s_2, \ldots, +s_2\}, \end{aligned}$$

then the sum belongs to the set

$$m_1 + m_2 \in \{-s_1 - s_2, \ldots, s_1 + s_2\},$$

which implies that $s = s_1 + s_2$ for product states. In the following example we will see also what happens for entangled states.

Example with two electrons. Consider now a system of (the internal spin of) two electrons:

$$\mathcal{H} = \mathcal{H}_{1/2} \otimes \mathcal{H}_{1/2}.$$

We represent the basis for the spins of the individual electrons with a spin up state $|\uparrow\rangle$ and a spin down state $|\downarrow\rangle$ (which are respectively $|0\rangle, |1\rangle$ when we schematize this electron as a qubit). Thus, the basis for the global system is:

$$\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$$

and, if we compute the eigenvalues as above:

$$\begin{aligned} S_z|\uparrow\uparrow\rangle &= \left(\frac{\hbar}{2} + \frac{\hbar}{2}\right)|\uparrow\uparrow\rangle = \hbar|\uparrow\uparrow\rangle \\ S_z|\uparrow\downarrow\rangle &= \left(\frac{\hbar}{2} - \frac{\hbar}{2}\right)|\uparrow\downarrow\rangle = 0 \\ S_z|\downarrow\uparrow\rangle &= \left(-\frac{\hbar}{2} + \frac{\hbar}{2}\right)|\downarrow\uparrow\rangle = 0 \\ S_z|\downarrow\downarrow\rangle &= \left(-\frac{\hbar}{2} - \frac{\hbar}{2}\right)|\downarrow\downarrow\rangle = -\hbar|\downarrow\downarrow\rangle \end{aligned}$$

There is a problem here: we said that each eigenspace $\mathcal{H}_{s,m}$ should have equal dimension, but here we found an eigenspace with $m = 0$ with two orthogonal eigenstate, while the eigenspaces with $m = \pm 1$ have one eigenstate each. In fact, the space is decomposed into two subspaces, one with $s = 1$ (i.e. 3 possible spin states $-1, 0, +1$) and one with $s = 0$ (with only the trivial

spin 0). In order to see this, we start with the down/down state $|\downarrow\downarrow\rangle$ and we apply the raising operator using the recursive definition derived above.

$$S_+|\downarrow\downarrow\rangle = S_+|\psi_{s,-1}\rangle = \hbar\sqrt{s(s+1)-(-1)(-1+1)}|\psi_{s,0}\rangle = \hbar\sqrt{s(s+1)}|\psi_{s,0}\rangle$$

We want to find the explicit representation of $|\psi_{s,0}\rangle$, in terms of spin up and down of the two electrons, and in order to do this we rewrite the expression:

$$\begin{aligned} S_+|\downarrow\downarrow\rangle &= (S_+\otimes\mathbb{1}+\mathbb{1}\otimes S_+)\,|\downarrow\downarrow\rangle \\ &= (S_+|\downarrow\rangle)\otimes|\downarrow\rangle + |\downarrow\rangle\otimes(S_+|\downarrow\rangle) \\ &= \hbar\sqrt{s_1(s_1+1)-m_1(m_1+1)}|\uparrow\downarrow\rangle + \hbar\sqrt{s_2(s_2+1)-m_2(m_2+1)}|\downarrow\uparrow\rangle \\ &= \hbar(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \end{aligned}$$

where in the last step we plugged $s_1 = s_2 = \frac{1}{2}$, which is the quantum number of both systems, and $m_1 = m_2 = -\frac{1}{2}$, since they are both at spin level $-\frac{1}{2}$ (spin down). Thus the actual state $|\psi_{s,0}\rangle$ is:

$$|\psi_{s,0}\rangle = \frac{1}{\sqrt{s(s+1)}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$$

the last equality, which follows from normalization, also implies that $s = 1$ is the quantum number of this subspace. Therefore, the three elements of the new subspace $\mathcal{H}_1$ (also called **triplet**) are:

$$\begin{aligned} |\psi_{1,-1}\rangle &= |\downarrow\downarrow\rangle \\ |\psi_{1,0}\rangle &= \frac{1}{\sqrt{2}}(|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle) \\ |\psi_{1,+1}\rangle &= |\uparrow\uparrow\rangle \end{aligned}$$

while the missing orthogonal state

$$|\psi_{0,0}\rangle = \frac{1}{\sqrt{2}}(|\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle)$$

belongs to the other subspace with quantum number 0 (also called **singlet**). It is immediate to see that this last state is orthogonal to the other three. Moreover, if we try to apply the ladder operators:

$$\begin{aligned} S_+|\psi_{0,0}\rangle &= \frac{1}{\sqrt{2}}(S_+\otimes\mathbb{1}+\mathbb{1}\otimes S_+)(|\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle) \\ &= \frac{1}{\sqrt{2}}((S_+|\downarrow\rangle)\otimes|\uparrow\rangle + |\downarrow\rangle\otimes(S_+|\uparrow\rangle) - (S_+|\uparrow\rangle)\otimes|\downarrow\rangle - |\uparrow\rangle\otimes(S_+|\downarrow\rangle)) \end{aligned}$$

Two terms cancel out since $S_+|\uparrow\rangle = 0$ and the other two terms turn out to be equal and opposite. The same calculation can be done for S_-, and this confirms that this state is alone in a subspace with $s = 0$. Hence, we derived that the tensor product of two spin-$\frac{1}{2}$ particles can be rewritten as direct sum of a triplet and a singlet:

$$\mathcal{H}_{\frac{1}{2}}\otimes\mathcal{H}_{\frac{1}{2}} = \mathcal{H}_0\oplus\mathcal{H}_1.$$

We want to generalize this derivation to a tensor product of spaces with arbitrary quantum numbers.

$$\mathcal{H} = \bigotimes_j \mathcal{H}_{s_j}$$

In other words, given the eigenbasis $\{\bigotimes_j |\psi_{s_j,m_{j,k}}\rangle\}_k$ of S^2 and S_z of each subsystem, we want to find an eigendecomposition for the operators S^2, S_z of the global system. In order to do this we rewrite the eigenstates $|\psi_{s,m}\rangle$ of the global system in terms of the eigenbases of the subsystems:

$$\begin{aligned}|\psi_{s,m}\rangle &= \sum_k \left(\bigotimes_j |\psi_{s_j,m_{j,k}}\rangle\right)\left(\bigotimes_j \langle\psi_{s_j,m_{j,k}}|\right)|\psi_{s,m}\rangle \\ &=: \sum_k c_k \left(\bigotimes_j |\psi_{s_j,m_{j,k}}\rangle\right)\end{aligned}$$

The constants $c_k \in \mathbb{C}$ are called the **Clebsch-Gordon coefficients**. These coefficients only depend on the quantum numbers s_j of the subsystems. Widely used combinations of s_j are already computed and can be found in lookup tables, but it is possible to compute them by hand (we did it in the example above). In general we can use the following "recipe":

- The total dimension of $\mathcal{H}$ is given by the product of the subspaces:

$$\dim \mathcal{H} = \prod_j \dim \mathcal{H}_{s_j} = \prod_j (2s_j + 1).$$

- We can also deduce that $c_k = 0$ unless $m = \sum_j m_{j,k}$. In fact, one can see that the tensor product above is an eigenstate of S_z:

$$\begin{aligned}S_z\left(\bigotimes_j |\psi_{s_j,m_{j,k}}\rangle\right) &= \left(\sum_j S_{z,j} \otimes \mathbb{1}_{-j}\right)\left(\bigotimes_j |\psi_{s_j,m_{j,k}}\rangle\right) \\ &= \left(\sum_j m_{j,k}\right)\left(\bigotimes_j |\psi_{s_j,m_{j,k}}\rangle\right),\end{aligned}$$

  which means that this state is an eigenstate associated to the eigenvalue $\sum_j m_{j,k}$ and, since S_z is Hermitian, any $|\psi_{s,m}\rangle$ with $m \neq \sum_j m_{j,k}$ has to be orthogonal, i.e.

$$c_k = \left(\bigotimes_j \langle\psi_{s_j,m_{j,k}}|\right)|\psi_{s,m}\rangle = 0.$$

14.5 Angular momentum

We conclude the chapter by looking at a particular spin system, which is **angular momentum**. Unlike the examples we saw before, angular momentum is defined on continuous systems, but most properties of spin carry out to this case. Let us first see the classical definition:

Definition 14.3. *The angular momentum $\mathbf{L}_O$ of a particle C with respect to a center O is defined as:*

$$\mathbf{L}_O = \mathbf{r} \times \mathbf{p}$$

where $\mathbf{r} = \mathbf{r}_C - \mathbf{r}_O$ is a vector starting from the point O pointing to the position $\mathbf{r}_C$ of the particle and $\times$ denotes the cross product of two vectors.

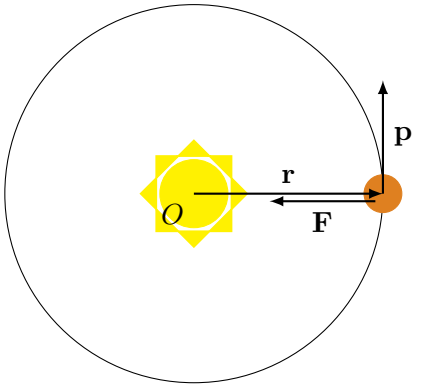


Figure 14.1: Conservation of angular momentum of a planet rotating around a star.

This object is extremely useful in certain analyses because, for a suitable choice of the center O, it is possible to obtain **conservation of angular momentum**. For example, imagine a planet with linear momentum $\mathbf{p} = m\mathbf{v}$ rotating around a star, as in Figure 14.1. The star will apply a gravitational force $\mathbf{F}$ to the planet. If we choose O as the center of the star, the angular momentum $\mathbf{L}_O$ of the planet will be conserved since:

$$\frac{d\mathbf{L}_O}{dt} = \frac{d\mathbf{r}}{dt} \times \mathbf{p} + \mathbf{r} \times \frac{d\mathbf{p}}{dt} = \mathbf{v} \times \mathbf{p} + \mathbf{r} \times \mathbf{F} = 0$$

The first term always vanishes because $\mathbf{p}$ and $\mathbf{v}$ are always parallel, by the classical definition of momentum. The clever choice of our center O also makes the second term vanish, because the gravitational force $\mathbf{F}$ is always pointing from the planet towards the star, parallel (and in opposite direction) to the vector $\mathbf{r}$. A typical use of the conservation of angular momentum in this context is to compute the velocity of the planet at any point of its orbit around the star.

Let us now find an analogous definition in quantum physics: the cross product $\mathbf{r} \times \mathbf{p}$ can be rewritten explicitly as:

$$\mathbf{r} \times \mathbf{p} = \begin{pmatrix} r_y p_z - r_z p_y \\ r_z p_x - r_x p_z \\ r_x p_y - r_y p_x \end{pmatrix}$$

If we define the Hilbert space for the position of a particle as usual

$$\mathcal{H} = \mathcal{H}_x \otimes \mathcal{H}_y \otimes \mathcal{H}_z$$

We can define a vector of **angular momentum operators** as follows:

$$\begin{aligned} L_x &= \mathbb{1}_x \otimes Y \otimes P_z - \mathbb{1}_x \otimes P_x \otimes Z \\ L_y &= P_x \otimes \mathbb{1}_y \otimes Z - X \otimes \mathbb{1}_y \otimes P_z \\ L_z &= X \otimes P_y \otimes \mathbb{1}_z - P_x \otimes Y \otimes \mathbb{1}_z \end{aligned}$$

To complete the definition we should also be able to express the center O: assuming X_O, Y_O, Z_O are the operators of the center, one can replace the position operators in the definition above with $X - X_O, Y - Y_O, Z - Z_O$, but to keep things simple we will assume that O is the origin of our space in the quantum case.

An important property here can be observed by looking at the three commutators $[L_x, L_y]$, $[L_y, L_z]$, and $[L_z, L_x]$:

$$[L_x, L_y] = L_x L_y - L_y L_x$$

$$
\begin{aligned}
&= (P_x \otimes Y \otimes P_z Z - X \otimes Y \otimes P_z^2 - P_x \otimes P_y \otimes Z^2 + X \otimes P_y \otimes Z P_z) \\
&- (P_x \otimes Y \otimes Z P_z - P_x \otimes P_y \otimes Z^2 - X \otimes Y \otimes P_z^2 + X \otimes P_y \otimes P_z Z) \\
&= P_x \otimes Y \otimes P_z Z + X \otimes P_y \otimes Z P_z - P_x \otimes Y \otimes Z P_z - X \otimes P_y \otimes P_z Z \\
&= P_x \otimes Y \otimes [P_z, Z] + X \otimes P_y \otimes [Z, P_z] \\
&= X \otimes P_y \otimes [Z, P_z] - P_x \otimes Y \otimes [Z, P_z] \\
&= i\hbar (X \otimes P_y \otimes \mathbb{1}_z - P_x \otimes Y \otimes \mathbb{1}_z) = i\hbar L_z
\end{aligned}
$$

where in the last step we used the canonical commutation relation $[Z, P_z] = i\hbar\mathbb{1}$. With similar calculations we also get $[L_y, L_z] = i\hbar L_x, [L_z, L_x] = i\hbar L_y$, thus we found that the angular momentum operators satisfy the properties of the spin operators, and the results we derived in the previous sections still hold. We will have a **total angular momentum operator** and the two ladder operators:

$$
\begin{aligned}
L^2 &= L_x^2 + L_y^2 + L_z^2 \\
L_+ &= L_x + iL_y \\
L_- &= L_x - iL_y
\end{aligned}
$$

whereas the eigenstates of L^2 (which are also the eigenstates of L_x, L_y, L_z) satisfy:

$$
\begin{aligned}
L^2 |\psi_{\ell,m}\rangle &= \ell(\ell+1)\hbar^2 |\psi_{\ell,m}\rangle \\
L_z |\psi_{\ell,m}\rangle &= m\hbar |\psi_{\ell,m}\rangle
\end{aligned}
$$

We would like to find the wave function $\psi_{\ell,m}(\mathbf{r})$:

$$
|\psi_{\ell,m}\rangle = \iiint_{\mathbb{R}^3} \psi_{\ell,m}(\mathbf{r}) \; |\mathbf{r}\rangle \; d^3\mathbf{r}
$$

In order to find this function we find the explicit form of L^2 and L_z:

$$
\begin{aligned}
\langle \mathbf{r}| L_z |\psi\rangle &= \langle \mathbf{r}| (X \otimes P_y \otimes \mathbb{1}_z - P_x \otimes Y \otimes \mathbb{1}_z) |\psi\rangle \\
&= x \langle \mathbf{r}| (\mathbb{1}_x \otimes P_y \otimes \mathbb{1}_z) |\psi\rangle - y \langle \mathbf{r}| (P_x \otimes \mathbb{1}_y \otimes \mathbb{1}_z) |\psi\rangle \\
&= i\hbar x \frac{\partial}{\partial y} \psi(\mathbf{r}) - i\hbar y \frac{\partial}{\partial x} \psi(\mathbf{r})
\end{aligned}
$$

By doing the same calculation for L_x, L_y, we find that the vector operator $\mathbf{L} = (L_x, L_y, L_z)^T$ transforms the wave function as follows:

$$
\mathbf{L} : \psi(\mathbf{r}) \mapsto i\hbar \; (\nabla_{x,y,z} \times \mathbf{r}\psi(\mathbf{r})) = i\hbar \; \mathrm{curl}\, (\mathbf{r}\psi(\mathbf{r}))
$$

where $\nabla_{x,y,z} \times \mathbf{r}\psi(\mathbf{r})$ is the **curl** of the vector $\mathbf{r}\psi(\mathbf{r})$. The rest of the procedure is not different from what we have already seen in the previous section: we find the eigenvalues of L^2, we restrict our attention to an eigenspace with quantum number ℓ and then, starting from the lowest eigenstate $|\psi_{\ell,-\ell}\rangle$, we can retrieve all the other eigenstates spanning this subspace by iteratively applying L_+.

14.6 Polar and spherical coordinates in space

In general, when we deal with angular momentum and spin on the plane or in space, Cartesian coordinates are not the best representation we can use to carry out the analysis. Like in classical

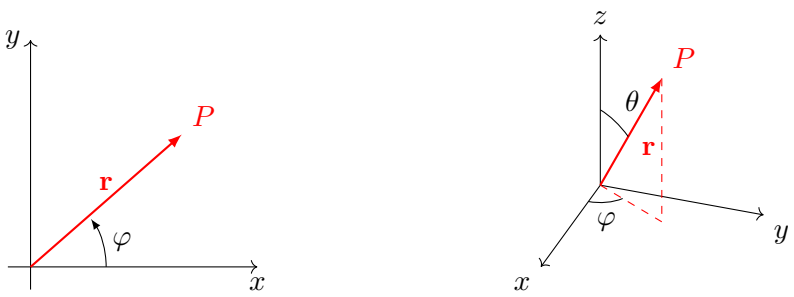


Figure 14.2: Relationship between polar and Cartesian coordinates on the plane and in a 3-dimensional space respectively.

physics, it would be a good idea in such cases to **switch to polar coordinates**. In the plane we have the following change of coordinates:

$$(x, y) \mapsto (r, \varphi) = \left(\sqrt{x^2 + y^2}, \arctan\left(\frac{y}{x}\right)\right)$$
$$(r, \varphi) \mapsto (x, y) = (r\cos\varphi, r\sin\varphi)$$

which means that r is the distance from the center of the system (in this case the origin, but we can generalize it by doing a translation), while θ is an angle that determines the point on the circle of radius r. It is clear that, in the case of planar rotation, r is constant and only θ changes over time.

We extend these coordinates to the three dimensional case by adding a third angle ϕ, which is tightly connected to the z-coordinate:

$$(x, y, z) \mapsto (r, \varphi, \theta) = \left(\sqrt{x^2 + y^2 + z^2}, \arctan\left(\frac{y}{x}\right), \arccos\left(\frac{z}{r}\right)\right)$$
$$(r, \varphi, \theta) \mapsto (x, y, z) = (r\cos\varphi\sin\theta, r\sin\varphi\cos\theta, r\cos\theta)$$

Here, φ is called **azimuthal angle**, while θ is called **zenith angle**. A similar coordinate system is the one we defined for the Bloch sphere, in Section 1.7. The relationship between polar and Cartesian coordinates is visualized in Figure 14.2.

Now, we would like to do a change in coordinates for an arbitrary state $|\psi\rangle$ and see what the wave function looks like. We define a basis $|r\rangle|\varphi\rangle|\theta\rangle$ such that:

$$(\langle r|\langle\varphi|\langle\theta|)(|x\rangle|y\rangle|z\rangle) = \delta(x - r\cos\varphi\sin\theta)\ \delta(y - r\sin\varphi\cos\theta)\ \delta(z - r\cos\theta)$$

and, as direct consequence:

$$\begin{aligned}
|\psi\rangle &= \iiint_{\mathbb{R}^3} \psi(x, y, z)\ |x\rangle|y\rangle|z\rangle dxdydz \\
&= \int_{\mathbb{R}^+}\int_0^{2\pi}\int_0^{\pi} |r\rangle|\varphi\rangle|\theta\rangle\langle r|\langle\varphi|\langle\theta| dr d\varphi d\theta \iiint_{\mathbb{R}^3} \psi(x, y, z)\ |x\rangle|y\rangle|z\rangle dxdydz \\
&= \int_{\mathbb{R}^+}\int_0^{2\pi}\int_0^{\pi} \iiint_{\mathbb{R}^3} \psi(x, y, z) \\
&\qquad \delta(x - r\cos\varphi\sin\theta)\ \delta(y - r\sin\varphi\cos\theta)\ \delta(z - r\cos\theta)|r\rangle|\varphi\rangle|\theta\rangle dr d\varphi d\theta \\
&= \int_{\mathbb{R}^+}\int_0^{2\pi}\int_0^{\pi} \psi(r\cos\varphi\sin\theta, r\sin\varphi\cos\theta, r\cos\theta)|r\rangle|\varphi\rangle|\theta\rangle dr d\varphi d\theta \\
&=: \int_{\mathbb{R}^+}\int_0^{2\pi}\int_0^{\pi} \tilde{\psi}(r, \varphi, \theta)|r\rangle|\varphi\rangle|\theta\rangle dr d\varphi d\theta
\end{aligned}$$

Thus, $\tilde{\psi}(r, \varphi, \theta) = \psi(r\cos\varphi\sin\theta, r\sin\varphi\cos\theta, r\cos\theta)$.

Angular momentum in spherical coordinates. Let us see how the angular momentum operators, in particular L^2 and L_z, act on the spherical wave function $\tilde{\psi}$. Considering the vector operator $\mathbf{L}$, we start from what we derived in the previous section:

$$L_z : \psi(\mathbf{r}) \mapsto i\hbar \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} \right) \psi(\mathbf{r})$$

We use the chain rule to replace the partial derivative operators:

$$\begin{aligned}\frac{\partial}{\partial x} &= \frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial \varphi}{\partial x}\frac{\partial}{\partial \varphi} + \frac{\partial \theta}{\partial x}\frac{\partial}{\partial \theta} \\ &= \sin\theta\cos\varphi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\cos\varphi\frac{\partial}{\partial \theta} - \frac{1}{r}\frac{\sin\varphi}{\sin\theta}\frac{\partial}{\partial \varphi}\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial y} &= \frac{\partial r}{\partial y}\frac{\partial}{\partial r} + \frac{\partial \varphi}{\partial y}\frac{\partial}{\partial \varphi} + \frac{\partial \theta}{\partial y}\frac{\partial}{\partial \theta} \\ &= \sin\theta\sin\varphi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\sin\varphi\frac{\partial}{\partial \theta} + \frac{1}{r}\frac{\cos\varphi}{\sin\theta}\frac{\partial}{\partial \varphi}\end{aligned}$$

On the other hand, we already know that $x = r\cos\varphi\sin\theta, y = r\sin\varphi\cos\theta, z = r\cos\theta$, hence we can replace everything to obtain:

$$\begin{aligned}x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} &= r\sin^2\theta\sin\varphi\cos\varphi\frac{\partial}{\partial r} - r\sin\theta^2\cos\varphi\sin\varphi\frac{\partial}{\partial r} \\ &\quad - \sin\theta\cos\theta\sin\varphi\cos\varphi\frac{\partial}{\partial \theta} + \sin\theta\cos\theta\sin\varphi\cos\varphi\frac{\partial}{\partial \theta} \\ &\quad + \cos^2\varphi\frac{\partial}{\partial \varphi} + \sin^2\varphi\frac{\partial}{\partial \varphi} \\ &= \frac{\partial}{\partial \varphi}\end{aligned}$$

And thus we conclude

$$L_z : \tilde{\psi}(r,\varphi,\theta) \mapsto -i\hbar\frac{\partial}{\partial \varphi}\tilde{\psi}(r,\varphi,\theta)$$

One can see a similarity here, between L_z and the linear momentum operators P_x, P_y, P_z: in some sense, the angular momentum is related with the angle φ in the exact same way P_x, P_y, P_z are related with x, y, z. Using similar arguments we saw in Chapter 5 one can also show that this quantity is tightly related to **angular velocity**.

Extending the calculations we find the following results:

$$\begin{aligned}L_x &: \tilde{\psi}(r,\varphi,\theta) \mapsto i\hbar\left(\sin\varphi\frac{\partial}{\partial\theta} + \cot\theta\cos\varphi\frac{\partial}{\partial\varphi}\right)\tilde{\psi}(r,\varphi,\theta) \\ L_y &: \tilde{\psi}(r,\varphi,\theta) \mapsto i\hbar\left(\cos\varphi\frac{\partial}{\partial\theta} + \cot\theta\sin\varphi\frac{\partial}{\partial\varphi}\right)\tilde{\psi}(r,\varphi,\theta) \\ L_z &: \tilde{\psi}(r,\varphi,\theta) \mapsto -i\hbar\frac{\partial}{\partial\varphi}\tilde{\psi}(r,\varphi,\theta)\end{aligned}$$

and, by composing the above, also the transformation of ladder operators and total angular momentum can be found:

$$L_+ : \tilde{\psi}(r,\varphi,\theta) \mapsto \hbar e^{i\varphi}\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\varphi}\right)\tilde{\psi}(r,\varphi,\theta)$$

$$L_- : \tilde{\psi}(r,\varphi,\theta) \mapsto \hbar e^{-i\varphi}\left(\frac{\partial}{\partial\theta} - i\cot\theta\frac{\partial}{\partial\varphi}\right)\tilde{\psi}(r,\varphi,\theta)$$
$$L^2 : \tilde{\psi}(r,\varphi,\theta) \mapsto -\hbar^2\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\tilde{\psi}(r,\varphi,\theta)}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\tilde{\psi}(r,\varphi,\theta)}{\partial\varphi^2}\right)$$

Finding the eigenstates of L_z and L^2. Using the ideas we presented at the beginning of the chapter, we want to find the eigenstates of L^2. The idea behind this is that, if the potential in the Hamiltonian is central, i.e.

$$H = \frac{P^2}{2\mu} + V(r)$$

then the total angular momentum is conserved, in a similar way we showed in the classical example of Section 14.5. This means that $[H, L^2] = 0$ and the eigenstates of L^2 are also the energy eigenstates of our system. For such a state $|\psi_{\ell,m}\rangle$ within a spin-ℓ system we have:

$$\begin{aligned}
L_z|\psi_{\ell,m}\rangle &= \hbar m|\psi_{\ell,m}\rangle \\
\langle r|\langle\varphi|\langle\theta|L_z|\psi_{\ell,m}\rangle &= \hbar m\langle r|\langle\varphi|\langle\theta|\psi_{\ell,m}\rangle \\
-i\hbar\frac{\partial}{\partial\varphi}\tilde{\psi}_{\ell,m}(r,\varphi,\theta) &= \hbar m\tilde{\psi}_{\ell,m}(r,\varphi,\theta) \\
\frac{\partial}{\partial\varphi}\tilde{\psi}_{\ell,m}(r,\varphi,\theta) &= im\tilde{\psi}_{\ell,m}(r,\varphi,\theta)
\end{aligned}$$

which yields the solution space:

$$\tilde{\psi}_{\ell,m}(r,\varphi,\theta) = C(r,\theta)e^{im\varphi}.$$

In order to find $C(r,\theta)$ we impose the condition at the bottom of the ladder (note that from now on we will take for granted that ℓ only takes integer values[16]):

$$L_-|\psi_{\ell,-\ell}\rangle \stackrel{!}{=} 0$$

$$\begin{aligned}
0 &= \hbar e^{-i\varphi}\left(\frac{\partial}{\partial\theta} - i\cot\theta\frac{\partial}{\partial\varphi}\right)\tilde{\psi}_{\ell,-\ell}(r,\varphi,\theta) \\
0 &= \left(\frac{\partial}{\partial\theta} - i\cot\theta\frac{\partial}{\partial\varphi}\right)C(r,\theta)e^{-i\ell\varphi} \\
&= e^{-i\ell\varphi}\frac{\partial}{\partial\theta}C(r,\theta) - iC(r,\theta)\cot\theta\frac{\partial}{\partial\varphi}e^{-i\ell\varphi} \\
&= e^{-i\ell\varphi}\left(\frac{\partial}{\partial\theta}C(r,\theta) - \ell C(r,\theta)\cot\theta\right)
\end{aligned}$$

giving us the following differential equation:

$$\frac{\partial}{\partial\theta}C(r,\theta) = \ell\cdot C(r,\theta)\cot\theta \Longrightarrow C(r,\theta) = C(r)\sin^\ell\theta$$

i.e. the lowest spin level has wave function of the form:

$$\psi_{\ell,-\ell}(r,\varphi,\theta) = C(r)\ e^{-i\ell\varphi}\sin^\ell\theta$$

[16]The intuition for this is that we want $\ell = 0$ to be a solution for systems that are not moving.

and from this we can apply L_+ for $m-\ell$ times to obtain $\psi_{\ell,m}(r,\varphi,\theta)$:

$$|\psi_{\ell,m+1}\rangle = \frac{1}{\sqrt{\ell(\ell+1)-m(m+1)}} L_+ |\psi_{\ell,m}\rangle$$
$$\psi_{\ell,m+1}(r,\varphi,\theta) = \frac{1}{\sqrt{\ell(\ell+1)-m(m+1)}} \hbar e^{i\varphi} \left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\varphi}\right)\psi_{\ell,m}(r,\varphi,\theta)$$

The terms of the wave function dependent from θ and ϕ (with normalization) are called **spherical harmonics**. To sum up, the transformation of L_+ yields the following recurrence for the spherical harmonics:

$$\begin{cases} P_{\ell,-\ell}(\theta,\varphi) = k_\ell \sin^\ell\theta \\ P_{\ell,m+1}(\theta,\varphi) = \frac{1}{\sqrt{\ell(\ell+1)-m(m+1)}}\left(\frac{\partial P_{\ell,m}}{\partial\theta} - mP_{\ell,m}(\theta,\varphi)\cot\theta\right) \end{cases}$$

where k_ℓ is a normalization constant. Some examples of well-known harmonics are:

- $P_{0,0}(\theta,\varphi) = \frac{1}{\sqrt{4\pi}}$;
- $P_{1,0}(\theta,\varphi) = \sqrt{\frac{3}{4\pi}}\cos\theta$;
- $P_{1,\pm1}(\theta,\varphi) = \mp\sqrt{\frac{3}{8\pi}}\sin\theta\ e^{\pm i\varphi}$.

Finally, to find the full expression for the wave function, we need to compute $C(r)$. In order to do so, we will exploit the fact that the states $|\psi_{l,m}\rangle$ are eigenstates of both L^2 and the Hamiltonian H:

$$H|\psi_{l,m}\rangle = E|\psi_{l,m}\rangle \tag{14.2}$$

We start by considering, as derived in Chapter 6, how P^2 acts on the wave function in Cartesian coordinates:

$$\begin{aligned} P^2 : \psi(x,y,z) \mapsto &-\hbar^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\psi(x,y,z) \\ &= -\hbar^2\nabla^2\psi(x,y,z) \end{aligned}$$

As usual, we can translate everything to spherical coordinates to get:

$$\begin{aligned} P^2 : \psi(r,\theta,\phi) \mapsto &-\hbar^2\left(\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\tilde{\psi}}{\partial r}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\tilde{\psi}}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\tilde{\psi}}{\partial\varphi^2}\right) \\ &= -\hbar^2\left(\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\tilde{\psi}}{\partial r}\right) - \frac{1}{\hbar^2 r^2}L^2(\tilde{\psi})\right) \end{aligned}$$

Where we simplified the expression using what we learned earlier in this section about the effect of L^2 on $\tilde{\psi}$. Thus, putting this together with the Hamiltonian for central potentials:[17]

$$H : \tilde{\psi}(r,\theta,\varphi) \mapsto -\frac{\hbar^2}{2\mu r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\tilde{\psi}}{\partial r}\right) + \frac{1}{2\mu r^2}L^2(\tilde{\psi}) + V(r)(\tilde{\psi}) \tag{14.3}$$

We can now combine Equation 14.3 and Equation 14.2 by considering a specific wave function $\tilde{\psi}_{\ell,m}$. Recall that $L^2|\tilde{\psi}_{\ell,m}\rangle = \hbar^2\ell(\ell+1))|\tilde{\psi}_{\ell,m}\rangle$; we obtain:

$$-\frac{\hbar^2}{2\mu r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\tilde{\psi}_{\ell,m}}{\partial r}\right) + \frac{\hbar^2\ell(\ell+1)}{2\mu r^2}(\tilde{\psi}_{\ell,m}) + V(r)(\tilde{\psi}_{\ell,m}) = E\tilde{\psi}_{\ell,m}$$

[17] There is a sign difference in the handwritten notes (a minus sign in the second term). For all practical purposes please follow the current version (with the + sign).

But we have already computed an explicit form for this wave function. We substitute it here and observe that we can simplify the terms $P_{\ell,m}(\theta,\varphi)$, as they do not depend on r:

$$-\frac{\hbar^2}{2\mu r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial C(r)}{\partial r}\right)+\frac{\hbar^2\ell(\ell+1)}{2\mu r^2}C(r)+V(r)C(r)=E\cdot C(r)$$

Let us rearrange the terms:

$$-\frac{\hbar^2}{2\mu r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial C(r)}{\partial r}\right)=\left(-\frac{\hbar^2\ell(\ell+1)}{2\mu r^2}+E-V(r)\right)C(r) \tag{14.4}$$

We observe here that $C(r)$ only depends on ℓ, not on m, whereas the energy might depend on some other quantity n, which we call **principal number**, we will learn about later, i.e. $E := E_{n,\ell}$; since C also depends on E, we write $C := C_{n,\ell}(r)$.

To solve the differential equation we apply a few tricks:

- Rewrite the potential as an **effective potential**:

$$V_\ell^{\text{eff}}(r)=V(r)+\frac{\hbar^2\ell(\ell+1)}{2\mu r^2}$$

- Simplify the term on the left by defining a new radial function $u_{n,\ell}(r)=r\cdot C(r)$. We obtain:

$$\begin{aligned}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial C(r)}{\partial r}\right)&=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\left(\frac{u_{n,l}(r)}{r}\right)\right)\\&=\frac{1}{r^2}\frac{\partial}{\partial r}\left(-u_{n,l}(r)+r\frac{\partial}{\partial r}u_{n,l}(r)\right)\\&=\frac{1}{r}\frac{\partial^2}{\partial r^2}u_{n,l}(r)\end{aligned}$$

In light of these two definitions, we can rewrite Equation 14.4 as

$$-\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial r^2}u_{n,l}(r)=\left(E-V_\ell^{\text{eff}}(r)\right)u_{n,l}(r) \tag{14.5}$$

If you look at it closely, you can notice that it resembles the one we derived for the stationary states of the Hamiltonian in a one-dimensional system in Section 10.3. Indeed, it also has the same solutions, which in this case depend on $V_\ell^{\text{eff}}(r)$. We will look at the solution in a concrete example in the next section.

14.7 Application: central potential in the Hydrogen atom

Let us now turn to a concrete example for the shape of $V_\ell^{\text{eff}}(r)$, namely the hydrogen atom, in which we have one proton in the nucleus and an electron orbiting around it, as depicted in Figure 14.3. The particles have mass $\mu_{p^+}=1.67\cdot10^{-27}\text{kg}$ and $\mu_{e^-}=9.11\cdot10^{-31}\text{kg}$ respectively. The electron also has a charge of $q_{e^-}=-1.60\cdot10^{-19}C:=-1e$ (the charge is measured in Coulomb $[C]$, but for simplicity we define this quantity as **one electron charge**); the proton has the same charge with opposite sign: $q_{p^+}=1e$.

Observe how the mass of the electron is several orders of magnitude smaller than the one of the proton. This leads to the following reasonable approximation: we neglect the effect of e^- on

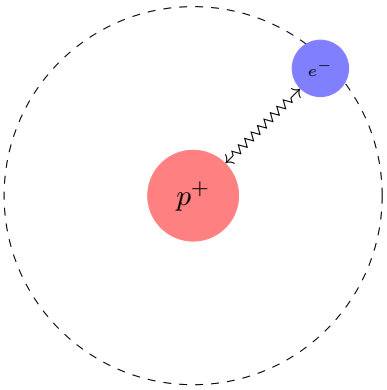


Figure 14.3: Diagram of a hydrogen atom.

p^+ and we assume the proton to be fixed at the center of our system of reference. The electron, on the other hand, is influenced by an electrical potential due to the presence of p^+.

Classically, the potential felt by a charge q_2 due to the presence of a charge q_1 at distance r is given by:

$$V(r) = k\frac{q_1 q_2}{r}$$

where $k \simeq 8.99 \cdot 10^9 Nm^2C^2$ is the **electric force constant**. In the case of the hydrogen atom it becomes:

$$V(r) = -k\frac{e^2}{r}$$

Given these premises, we will try to answer some questions that naturally arise, such as:

- What is the average value of r, i.e. what is the size of an atom?
- What does the energy spectrum $\{E_{n,l}\}_{n,l}$ look like?
- What shape does the orbit of the electron have, i.e. what can we learn about $|\psi_{n,l,m}|^2$?

To answer these, we try to find an explicit form for $u_{n,\ell}(r)$, by plugging in the specific form of the potential for this case into Equation 14.5:

$$-\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial r^2}u_{n,l}(r) = \left(E + k\frac{e^2}{r} - \frac{\hbar^2\ell(\ell+1)}{2\mu r^2}\right)u_{n,l}(r)$$

It turns out that this differential equation is, in fact, very hard to solve, even without taking into account the constraints associated with it, such as the positivity of the coefficients and the physical limits we want to impose to $u_{n,l}$ (e.g. when r goes to zero or to infinity). We refer you to Section 13.6 of Schumacher and Westmoreland [1] for an explicit step-by-step solution. Here, we just summarize a few key results from the computation:

- $n > \ell \geq |m| \geq 0$: this limits the different tuples for these values as summarized in Figure 14.4.
- The energy values are $E_{n,\ell} = E_n = -\frac{\mu k^2 e^4}{2\hbar^2} \cdot \frac{1}{n^2}$, where we call E_1 the **ground state**. From this we define the **Bohr radius** $a := \frac{\hbar^2}{\mu k e^2}$. This is the result we were looking for about the energy spectrum of the electron.

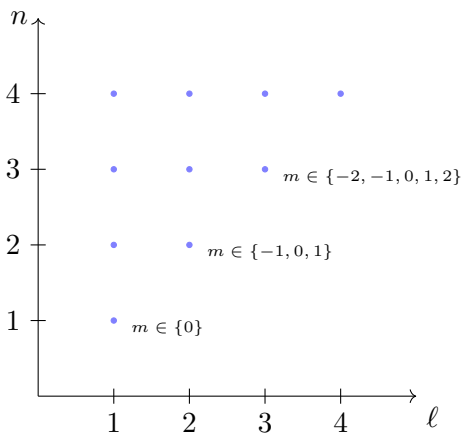


Figure 14.4: Possible values for n and ℓ in the hydrogen atom. To each value of ℓ corresponds the same set of possible values for m.

- The radial component of the wave function has the following form:

$$R_{n,\ell}(r) = A_{n,\ell}\left(\frac{r}{a}\right)^{\ell} \exp\left(-\frac{r}{a}\frac{1}{n}\right)\text{Poly}\left[\frac{r}{a}\right]$$

  where $A_{n,l}$ is a normalization factor and the last term is a **Laguerre polynomial**.

Examples of eigenstates. In atomic physics, each eigenstate defines an **orbital**. Here we list the radial components for the first few values of n and ℓ; we have computed the corresponding spherical harmonics in Section 14.6:

- $n = 1, \ell = 0, m = 0$: this is the ground state (orbital $1s$), in which $R_{1,0} = 2a^{-3/2}e^{-r/a}$;
- $n = 2, \ell = 0, m = 0$: orbital $2s$, in this case $R_{2,0} = \frac{a^{-3/2}}{\sqrt{2}}e^{-r/2a}(1 - \frac{r}{2a})$;
- $n = 2, \ell = 1$: these are called the p orbitals. The radial component is $R_{2,1} = \frac{a^{-3/2}}{\sqrt{24}}e^{-r/2a}(1-\frac{r}{a})$. Recall that we now have three different values for $m \in \{-1, 0, 1\}$, and thus three distinct spherical harmonics.

These answer one of our questions about the hydrogen atom, namely what shape do the orbitals have? You can find the shapes visualized on the corresponding Wikipedia page.

Computing the radius of an atom. The following theorem can often be used as a helpful trick to compute the average values of observables.

Theorem 14.4 (Feynman-Hellmann)**.** *Given a system with Hamiltonian H, let $|\psi_E\rangle$ be the eigenstate of H associated with eigenvalue E, i.e. $H|\psi_E\rangle = E|\psi_E\rangle$. If H, E and $|\psi_E\rangle$ all depend smoothly on some parameter λ, then we have that:*

$$\frac{\partial E}{\partial \lambda} = \langle\psi_E|\left(\frac{\partial H}{\partial \lambda}\right)|\psi_E\rangle$$

Proof. From $H|\psi_E\rangle = E|\psi_E\rangle$, since the state $|\psi_E\rangle$ is normalized, we evince that:

$$E = \langle\psi_E|H|\psi_E\rangle$$

Taking the derivative with respect to λ on both sides:

$$\begin{aligned}
\frac{\partial E}{\partial \lambda} &= \frac{\partial}{\partial \lambda}\langle\psi_E|H|\psi_E\rangle \\
&= \left(\frac{\partial}{\partial \lambda}\langle\psi_E|\right) H|\psi_E\rangle + \langle\psi_E|\left(\frac{\partial}{\partial \lambda}H\right)|\psi_E\rangle + \langle\psi_E|H\left(\frac{\partial}{\partial \lambda}|\psi_E\rangle\right) \\
&= E\left(\frac{\partial}{\partial \lambda}\langle\psi_E|\right)|\psi_E\rangle + \langle\psi_E|\left(\frac{\partial}{\partial \lambda}H\right)|\psi_E\rangle + E\langle\psi_E|\left(\frac{\partial}{\partial \lambda}|\psi_E\rangle\right) \\
&= E\left(\left(\frac{\partial}{\partial \lambda}\langle\psi_E|\right)|\psi_E\rangle + \langle\psi_E|\left(\frac{\partial}{\partial \lambda}|\psi_E\rangle\right)\right) + \langle\psi_E|\left(\frac{\partial}{\partial \lambda}H\right)|\psi_E\rangle \\
&= E\frac{\partial}{\partial \lambda}(\langle\psi_E|\psi_E\rangle) + \langle\psi_E|\left(\frac{\partial}{\partial \lambda}H\right)|\psi_E\rangle \\
&= \langle\psi_E|\left(\frac{\partial}{\partial \lambda}H\right)|\psi_E\rangle
\end{aligned}$$

□

We will now use the theorem to compute the average radius of the atom as a function of the energy level n. In particular, if we set $\lambda = k$, we obtain:

$$\begin{aligned}
\frac{\partial E_n}{\partial k} &= \langle\psi_{n,\ell,m}|\frac{\partial H}{\partial k}|\psi_{n,\ell,m}\rangle \\
\frac{\partial}{\partial k}\left(-\frac{\mu k^2 e^4}{2\hbar^2 n^2}\right) &= \langle\psi_{n,\ell,m}|\frac{\partial}{\partial k}\left(\frac{P^2}{2\mu} + k\frac{e^2}{R}\right)|\psi_{n,\ell,m}\rangle \\
-\frac{\mu k e^2}{\hbar^2 n^2} &= \langle\psi_{n,\ell,m}|\left(\frac{1}{R}\right)|\psi_{n,\ell,m}\rangle \\
\frac{1}{an^2} &= \langle\frac{1}{R}\rangle^{\psi_{n,l,m}}
\end{aligned}$$

where a is again the Bohr radius. Thus we found that $\langle R\rangle^{\psi_{n,l,m}} \simeq an^2$ (we approximate the inverse of the expectation with the expectation of the inverse using Jensen's inequality with proper additional assumptions on R). Indeed, a is the average radius of the hydrogen atom in the ground state.

Quantum theory and experiments for the hydrogen atom. The quantum model of the hydrogen atom is widely supported by experimental results. For example, when exciting the electron in some way, it will tend to relax back to the ground state. In doing so, by conservation of energy, it will release the excess energy in the form of a proton, whose energy is proportional to its frequency ν: $E_{\text{photon}} = \hbar\nu$. In turn, the frequency of the photon is related to its color and can thus be measured: the energy of the photon is indeed the difference in energy between the energy levels of the electron.

Fine tuning the model. Finally, one may ask how we can loosen the assumptions about the interaction we made in the beginning of the discussion, and even how to generalize the model to other atoms. In this latter case, one must also take into account the interaction between the electrons when computing the Hamiltonian; this increase in complexity means that the resulting differential equations need to be solved numerically: this is competence of computational physics.

There are however some refinements to the model that we can make with the tools we have acquired during the course. For example, we know that the electron has an internal degree of freedom given, by its spin. The finer-structured Hilbert space should also take this into account:

$$\mathcal{H}_{e^-} = \mathcal{H}_{3D} \otimes \mathcal{H}_{\text{spin}} = \mathcal{H}_R \otimes \mathcal{H}_\theta \otimes \mathcal{H}_\varphi \otimes \mathcal{H}_{\text{spin}}$$

The state of an electron now would be

$$|\psi\rangle_{e^-} = \sum_{n=1}^{\infty} \sum_{\ell=0}^{n-1} \sum_{m=-\ell}^{\ell} \sum_{m_s=-\frac{1}{2}}^{\frac{1}{2}} \alpha_{n,\ell,m,m_s} |\psi_{n,\ell,m}\rangle |\psi_{m_s}\rangle$$

and this is analogous to the "adding spins" scenario we have analysed in Section 14.4.

Another aspect one might want to take into account is the fact that the electrons travel around the nucleus at a fairly high velocity, and thus incorporate relativistic analysis into the model. Finally, going into the hyper-fine structure, one might also consider the fact that the nucleus has a spin itself. All these observations can be combined to create a more complex and complete vision of the atoms from a quantum perspective. As we refine the model — in this case the Hamiltonian — we find that the energy degeneracy of the each level is lifted, as there are small energy corrections depending on the values of ℓ, m, s and so on. These can be detected with sufficiently accurate measurement devices.

Chapter A
Recap on Analysis

Here we briefly list all the elements of analysis needed to understand the lectures.

A.1 Exponential function

The exponential function $f(x) = e^x$ is defined as:

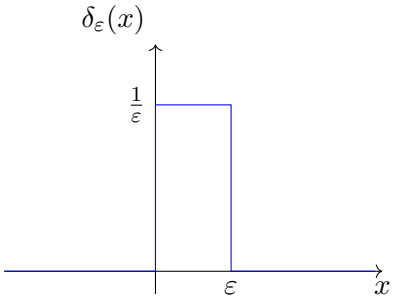
$$e^x = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Using these expressions, the definition of exponential can be extended to arbitrary objects with an algebraic structure providing an addition operator $+$ and a multiplication operator $\cdot$.

A.2 Dirac delta function

We define $\delta_\varepsilon(x)$ as the following function (also called **nascent delta**):

$$\delta_\varepsilon(x) = \begin{cases} \frac{1}{\varepsilon} & 0 < x < \varepsilon \\ 0 & \text{otherwise} \end{cases}$$



The **Dirac delta function** $\delta(x)$ can be defined as:

$$\delta(x) = \lim_{\varepsilon\to 0} \delta_\varepsilon(x)$$

This function has important properties that are used extensively in calculus applied to quantum physics. Here we show some of them.

Theorem A.1. $\int_{\mathbb{R}} \delta(x)dx = 1$.

Proof.

$$\int_{\mathbb{R}} \delta(x)dx = \int_{\mathbb{R}} \lim_{\varepsilon\to 0} \delta_\varepsilon(x)dx = \lim_{\varepsilon\to 0} \int_{\mathbb{R}} \delta_\varepsilon(x)dx = \lim_{\varepsilon\to 0} \int_0^\varepsilon \frac{1}{\varepsilon}\, dx = \lim_{\varepsilon\to 0} 1 = 1$$

□

Theorem A.2. *If f is continuous at $x = 0$, then* $\int_{\mathbb{R}} f(x)\delta(x)dx = f(0)$.

Proof.

$$\begin{aligned}\int_{\mathbb{R}} f(x)\delta(x)dx &= \lim_{\varepsilon\to 0}\int_0^{\varepsilon}\frac{1}{\varepsilon}f(x)dx \\ &= \lim_{\varepsilon\to 0}\int_0^{\varepsilon}\frac{1}{\varepsilon}f(0)dx && \text{by continuity of } f \\ &= f(0)\lim_{\varepsilon\to 0}\int_0^{\varepsilon}\frac{1}{\varepsilon}\,dx = f(0)\end{aligned}$$

□

This result can be generalized to a generic center:

$$\int_{\mathbb{R}} f(x)\delta(x-x_0)dx = f(x_0)$$

for any function f continuous in $x = x_0$.

We conclude this section by mentioning the analogous of the Dirac delta for the discrete case: the **Kronecker Delta**, which is defined as follows

$$\delta_{ij} = \begin{cases}1 & i = j \\ 0 & \text{otherwise}\end{cases}$$

A.3 Complex numbers

The set of complex numbers $\mathbb{C}$ can be defined as:

$$\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}\}$$

where $i := \sqrt{-1}$ is the imaginary unit. Any complex number $z \in \mathbb{C}$ can be expressed as above, where $\Re z := a, \Im z := b$ are respectively the real and imaginary parts of z.

Theorem A.3 (Euler's identity)**.** $e^{i\theta} = \cos(\theta) + i\sin(\theta)$.

Proof. We can rewrite the Taylor series of the exponential:

$$\begin{aligned}e^{i\theta} &= \sum_{n=0}^{\infty}\frac{(i\theta)^n}{n!} \\ &= \sum_{n=0}^{\infty}\frac{(i\theta)^{2n}}{(2n)!} + \sum_{n=0}^{\infty}\frac{(i\theta)^{2n+1}}{(2n+1)!} && \text{splitting the sum} \\ &= \sum_{n=0}^{\infty} i^{2n}\frac{\theta^{2n}}{(2n)!} + \sum_{n=0}^{\infty} i^{2n+1}\frac{\theta^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty}(-1)^n\frac{\theta^{2n}}{(2n)!} + i\sum_{n=0}^{\infty}(-1)^n\frac{\theta^{2n+1}}{(2n+1)!} && \text{since } i^2 = -1 \\ &= \cos(\theta) + i\sin(\theta) && \text{Taylor series of sin and cos}\end{aligned}$$

□

Using this identity we can always express $z \in \mathbb{C}$ in **complex exponential form**:

$$z = \rho e^{i\theta}$$

where $\rho \in \mathbb{R}_0^+$ is called **absolute value** (also written as $|z|$) and $\theta \in [0, 2\pi]$ is the **phase**. One can pass from one representation to another in the following way:

$$a + ib \mapsto \sqrt{a^2 + b^2} e^{i \arctan(b/a)}$$
$$\rho e^{i\theta} \mapsto \rho \cos(\theta) + i\rho \sin(\theta)$$

Definition A.4 (Complex conjugate)**.** *Let $z = a + ib \in \mathbb{C}$. We define the operator $z^* = a - ib$ as the **complex conjugate** of z.*

Some extremely important observations about complex conjugates:

- $(z_1 + z_2)^* = z_1^* + z_2^*$;
- $(z_1 z_2)^* = z_1^* z_2^*$;
- $z^* z = |z|^2$, since $(a + ib)(a - ib) = a^2 - (ib)^2 = a^2 + b^2 = |z|^2$;
- If $a = 0$ (i.e. z is purely imaginary), then $z^* = -z$;
- If $\rho = 1$ (i.e. z is unitary), then $z^* = \frac{1}{z}$.

The first two properties imply **linearity of conjugation**, while from the last two properties we also evince that $-i = i^* = \frac{1}{i}$, which we will use extensively.

A.4 Fourier transform

Let $f : \mathbb{R} \to \mathbb{R}$ be an integrable function. The **Fourier transform** f is a function F defined as:

$$F(t) = \mathcal{F}[f] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)\ e^{itx}\ dx$$

The inverse of the Fourier transform is another Fourier transform:

$$f(x) = \mathcal{F}[f] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} F(t)\ e^{-itx}\ dt$$

When we need to use the Fourier transform in quantum physics as relation between position p and momentum x, we add the Planck constant $\hbar$ for historical reasons, but also because $\hbar$ removes the units of measurements of the term px in the exponential:

$$F(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} f(x)\ e^{ipx/\hbar}\ dx$$
$$f(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} F(p)\ e^{-ipx/\hbar}\ dp$$

Here we briefly list some properties of the Fourier transform. We will directly refer to the position-momentum transform, but analogous results hold for a general Fourier transform:

Theorem A.5 (Linearity of the transform)**.** *$\mathcal{F}[af + bg] = a\mathcal{F}[f] + b\mathcal{F}[g]$ for $a, b \in \mathbb{C}$.*

Proof. Directly follows from linearity of integral. □

Theorem A.6 (Transform of derivative)**.** $\mathcal{F}[f'] = \frac{ip}{\hbar}\mathcal{F}[f]$

Proof.

$$\begin{aligned}
\mathcal{F}[f'] &= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} f'(x)\ e^{ipx/\hbar}\ dx \\
&= \frac{1}{\sqrt{2\pi\hbar}} \left[f(x)\ e^{ipx/\hbar} \right]_{\mathbb{R}} - \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} f(x)\ \left(-\frac{ip}{\hbar} \right) e^{ipx/\hbar}\ dx \qquad \text{integration by parts} \\
&= \frac{ip}{\hbar} \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} f(x)\ e^{ipx/\hbar}\ dx \\
&= \frac{ip}{\hbar} \mathcal{F}[f]
\end{aligned}$$

The first term of the integration by parts tends to 0 at $\pm\infty$, for the assumption that the integral $f(x)\ e^{ipx/\hbar}$ converges (i.e. the Fourier transform of $f(x)$ is finite). □

Theorem A.7 (Transform of the shift)**.** $\mathcal{F}[f(x - x_0)] = e^{ipx_0} \mathcal{F}[f]$

Proof.

$$\begin{aligned}
\mathcal{F}[f(x - x_0)] &= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} f(x - x_0)\ e^{ipx/\hbar}\ dx \\
&= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} f(x)\ e^{ip(x+x_0)/\hbar}\ dx \qquad \text{substitution } x \leftarrow x + x_0 \\
&= e^{ipx_0/\hbar} \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} f(x)\ e^{ipx/\hbar}\ dx \\
&= e^{ipx_0/\hbar} \mathcal{F}[f]
\end{aligned}$$

□

Theorem A.8 (Transform of the Dirac delta)**.** $\mathcal{F}[\delta] \equiv \frac{1}{\sqrt{2\pi\hbar}}$.

Proof. We use Theorem A.2:

$$\mathcal{F}[\delta] = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \delta(x)\ e^{ipx/\hbar}\ dx = \frac{1}{\sqrt{2\pi\hbar}}\ e^{ip0/\hbar} = \frac{1}{\sqrt{2\pi\hbar}}$$

□

Corollary A.9. *The Dirac delta function can be defined as:*

$$\delta(x) = \frac{1}{2\pi\hbar} \int_{\mathbb{R}} e^{-ipx/\hbar} dp$$

Proof. By Theorem A.8, $\delta(x)$ is the anti-transform of $\frac{1}{\sqrt{2\pi\hbar}}$, hence:

$$\begin{aligned}
\delta(x) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\hbar}} e^{-ipx/\hbar} dp \\
&= \frac{1}{2\pi\hbar} \int_{\mathbb{R}} e^{-ipx/\hbar} dp
\end{aligned}$$

□

A.5 Solving differential equations

Definition A.10 (Differential equation)**.** *A (ordinary) differential equation is an equation of the form:*

$$f_n(t, y, y', \dots, y^{(n)}) = 0$$

*where $y(t)$ is an unknown function of t, and $y^{(k)}$ denotes the k-th derivative of y. In this case, n is said to be the **order** of the equation.*

Definition A.11. *An ordinary differential equation is said to be **linear** if:*

$$f_n(t, y, y', \dots, y^{(n)}) = g(t) + g_0(t)y + g_1(t)y' + \cdots + g_n(t)y^{(n)} = g(t) + \sum_{k=0}^{n} g_k(t)y^{(k)}$$

*Moreover, if $g(t) = 0$, the equation is said to be **homogeneous**.*

Theorem A.12. *The solution space to a linear homogeneous ordinary differential equation yields a vector space.*

Proof. If y_1, y_2 are solutions, then also $\alpha y_1 + \beta y_2$ is a solution since:

$$\sum_{k=0}^{n} g_k(t)(\alpha y_1 + \beta y_2)^{(k)} = \alpha \sum_{k=0}^{n} g_k(t)y_1^{(k)} + \beta \sum_{k=0}^{n} g_k(t)y_2^{(k)} = 0$$

□

Theorem A.13. *An homogeneous, first order linear differential equation yields the following solution space:*

$$y' = ky \Longrightarrow y(t) = y(t_0)e^{k(t-t_0)}$$

for a fixed $t_0 \in \mathbb{R}$.

Proof. We integrate the equation once:

$$\begin{aligned}
y' = ky &\Leftrightarrow \frac{y'}{y} = k \\
&\Leftrightarrow \int_{t_0}^{t} \frac{y'}{y}\, dt = \int_{t_0}^{t} k\, dt \\
&\Leftrightarrow \int_{y(t_0)}^{y(t)} \frac{dy}{y} = k\int_{t_0}^{t} dt \qquad \text{substitution } y = y(t), dy = y'dt \\
&\Leftrightarrow \ln y(t) - \ln y(t_0) = k(t - t_0) \\
&\Leftrightarrow \ln y(t) = \ln y(t_0) + k(t - t_0) \\
&\Leftrightarrow y(t) = y(t_0)e^{k(t-t_0)}
\end{aligned}$$

□

Definition A.14. *A linear system of differential equations is of the form:*

$$y' = Ay + b$$

where $y, y' \in \mathbb{C}^n$, $A \in \mathbb{C}^{n\times n}$ and $b \in \mathbb{C}^n$. n is said to be the order of the system. If $b = 0$ the system is said to be homogeneous.

Theorem A.15. *An homogeneous linear system of differential equations of the form:*

$$y' = Ay$$

yields the following solution space:

$$y(t) = e^{A(t-t_0)}y(t_0)$$

Moreover, if $v_1, \ldots, v_n$ are eigenvectors of A associated to the eigenvalues $\lambda_1, \ldots, \lambda_n$ then the solution can be rewritten as:

$$y(t) = \sum_{k=1}^{n} e^{\lambda_k(t-t_0)} v_k v_k^{\dagger} y(t_0)$$

For a recap on eigenvalues and eigenvectors, see Chapter B.

Chapter B
Recap on Linear Algebra

Here we list the main notions of linear algebra. Keep in mind that we talk about **complex** field. Some definitions and results are not identical from the linear algebra over real field you may be more familiar with, but they are natural extensions.

B.1 Properties of the trace

Theorem B.1 (Linearity of trace)**.** $\mathrm{Tr}(\alpha A + \beta B) = \alpha \,\mathrm{Tr}(A) + \beta \,\mathrm{Tr}(B)$.

Theorem B.2 (Cyclic property of the trace)**.** *Let $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times m}$. The following holds:*

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$$

Proof.

$$\mathrm{Tr}(AB) = \sum_i [AB]_{ii} = \sum_i \sum_j a_{ji} b_{ij} = \sum_j \sum_i a_{ij} b_{ji} = \sum_j [BA]_{jj} = \mathrm{Tr}(BA)$$

□

Theorem B.3. *The trace of a matrix $\mathrm{Tr}(A)$ is independent of the basis chosen to represent A.*

Proof. In other words, $\mathrm{Tr}(UAU^{-1}) = \mathrm{Tr}(A)$ for any full-rank matrix U, but we get this for free from the cyclic property:

$$\mathrm{Tr}(UAU^{-1}) = \mathrm{Tr}(AU^{-1}U) = \mathrm{Tr}(A)$$

□

B.2 Inner product spaces

Definition B.4. *Let $\mathcal{X}$ be a vector space over complex field, and let $\langle\cdot,\cdot\rangle : \mathcal{X}\times\mathcal{X} \to \mathbb{R}$ be a function. A tuple $(\mathcal{X}, \langle\cdot,\cdot\rangle)$ is said to form a **inner product space** or **pre-Hilbert space** if the following holds for $\langle\cdot,\cdot\rangle$:*

- ***Linearity**: $\langle a+b, c\rangle = \langle a, c\rangle + \langle b, c\rangle$ and $\langle \alpha a, c\rangle = \alpha\langle a, c\rangle$ for $a, b, c \in \mathcal{X}, \alpha \in \mathbb{C}$;*
- ***Hermitian symmetry**: $\langle a, b\rangle = \langle b, a\rangle^*$ where $\cdot^*$ denotes the complex conjugate;*
- ***Positive definiteness**: $\langle a, a\rangle > 0$ for $a \neq 0$.*

A inner product space can be defined also for reals and, in that case, we obtain the definition of **Euclidean space**.

Definition B.5. *Let $\mathcal{X}$ be a vector space and $d : \mathcal{X}\times\mathcal{X} \to \mathbb{R}_0^+$. The tuple $(\mathcal{X}, d)$ is said to form a **metric space** if the following hold for d:*

- ***Identity of discernibles**: $d(x, y) = 0 \iff x = y$;*
- ***Symmetry**: $d(x, y) = d(y, x)$;*

- ***Triangular inequality:*** $d(x,z) \leq d(x,y) + d(y,z)$.

A inner product space $(\mathcal{X}, \langle\cdot,\cdot\rangle)$ naturally induces a metric space $(\mathcal{X}, d)$ where:

$$d(x,y) = \langle x-y, x-y\rangle$$

The **standard** (dot) product between complex vectors in $\mathbb{C}^n$ is redefined as:

$$\langle x,y\rangle = x^\dagger y = x_1^* y_2 + \cdots + x_n^* y_n$$

And two vectors $x, y \in \mathbb{C}^n$ are said to be **orthogonal** if and only if $x^\dagger y = 0$.

Theorem B.6 (Cauchy-Schwarz)**.** *For any two vectors x, y of an inner product space $(\mathcal{X}, \langle\cdot,\cdot\rangle)$ we have $\langle x,x\rangle\langle y,y\rangle \geq |\langle x,y\rangle|^2$.*

Proof. If $\langle y,y\rangle = 0$, the claim is trivial. Therefore, assume $\langle y,y\rangle \neq 0$. Define the following vector:

$$z = x - \frac{\langle x,y\rangle}{\langle y,y\rangle} y$$

One can notice that, by linearity of the inner product:

$$\langle z,y\rangle = \langle x - \frac{\langle x,y\rangle}{\langle y,y\rangle} y, y\rangle = \langle x,y\rangle - \frac{\langle x,y\rangle}{\langle y,y\rangle}\langle y,y\rangle = 0$$

i.e. z and y are orthogonal, and x can be represented as:

$$x = \frac{\langle x,y\rangle}{\langle y,y\rangle} y + z$$

since this is a sum of orthogonal vectors, we can use the Pythagorean theorem:

$$||x||^2 = \left|\frac{\langle x,y\rangle}{\langle y,y\rangle}\right|^2 ||y||^2 + ||z||^2 \geq \frac{|\langle x,y\rangle|^2}{||y||^2}$$

implying, $||x||^2||y||^2 \geq |\langle x,y\rangle|^2$, as claimed. □

B.3 Unitary matrices

Definition B.7. *The transpose conjugate $A^\dagger$ of a matrix A is defined as:*

$$A^\dagger = (A^*)^T \equiv (A^T)^*$$

The complex conjugate applies to every entry of a matrix or vector and is interchangeable with any linear operator, by linearity of conjugation.

Definition B.8. *A square matrix $U \in \mathbb{C}^{n\times n}$ is said to be unitary if $U^{-1} = U^\dagger$.*

A real unitary matrix yields exactly the definition of **orthogonal matrix**. The following properties hold for a unitary matrix U:

- U has orthogonal columns;
- $|\det(U)| = 1$;
- All the eigenvalues of U are unitary, i.e. $|\lambda| = 1$ for every eigenvalue λ of U (see next section for the definition of eigenvalues).

B.4 Eigenvalues and eigenvectors

Definition B.9 (Eigenvalues)**.** *Let $A \in \mathbb{C}^{n\times n}$. An eigenvalue λ of A is such that, there exists $v \in \mathbb{C}^n$ such that:*

$$Av = \lambda v$$

*v is said to be an **eigenvector** of A for λ.*

Computing the eigendecomposition of a matrix. Here we show how to compute the eigendecomposition of a matrix $A \in \mathbb{C}^{n\times n}$, namely find all the vectors $v \in \mathbb{C}^n$ such that:

$$Av = \lambda v$$

along with the corresponding eigenvalues λ. First of all, we rewrite the above constraint as:

$$Av \stackrel{!}{=} \lambda \mathbb{1} v \iff (A - \lambda \mathbb{1})v \stackrel{!}{=} 0$$

This is a linear system of equations, and we would like to find all the values of λ such that the system is non-trivial, i.e. the solution is not only the zero vector. This means that we would like to find λ such that the rank of the matrix $A - \lambda\mathbb{1}$ is not maximum, and we can use the determinant for this:

$$\mathrm{rk}(A - \lambda\mathbb{1}) < n \iff \det(A - \lambda\mathbb{1}) = 0$$

The determinant on the right is called **characteristic polynomial** of A. Since this is a polynomial, by the fundamental theorem of algebra we know that it has exactly n roots, which will be A's eigenvalues (the number of times a root is counted in the characteristic polynomial is called **algebraic multiplicity** of the eigenvalue).

Now that we found the eigenvalues, in order to find an eigenvector associated to an eigenvalue λ_i, it is sufficient to find a non-trivial solution to the linear system:

$$(A - \lambda_i \mathbb{1})v = 0$$

The solution space, which is $\ker(A - \lambda_i\mathbb{1})$, is a linear space called **eigenspace** of A associated to λ_i. The dimension of this subspace is called **geometric multiplicity** of λ_i.

If the geometric multiplicity and the algebraic multiplicity of each eigenvalue of A coincide, then the direct sum of the eigenspaces of A span the whole space $\mathbb{C}^n$, i.e. the eigenvectors of A form a basis of $\mathbb{C}^n$ called the **eigenbasis** of A.

Observation B.10. *For $A \in \mathbb{C}^{n\times n}$, $\ker(A)$ is exactly the eigenspace of A for the eigenvalue 0.*

Theorem B.11. *Let $A \in \mathbb{C}^{n\times n}$. Any set $v_1, \dots, v_k$ of non-null eigenvectors for pairwise distinct eigenvalues $\lambda_1, \dots, \lambda_k$ are linearly independent.*

Proof. Consider the first two vectors v_1, v_2, and a linear combination $v = a_1v_1 + a_2v_2 = 0$. Consider Av:

$$\begin{aligned} Av &= A(a_1v_1 + a_2v_2) \\ &= a_1\lambda_1v_1 + a_2\lambda_2v_2 = 0 \end{aligned}$$

Since $a_2v_2 = -a_1v_1$, the equation above becomes:

$$a_1(\lambda_1 - \lambda_2)v_1 = 0$$

This implies $a_1 = 0$ since $\lambda_1 \neq \lambda_2$, and thus also $a_2 = 0$ as $v_2 \neq 0$. Suppose by induction $v_1, \dots, v_{k-1}$ are linearly independent. We can apply the same reasoning by plugging $a_kv_k = -\sum_{i=1}^{k-1} a_iv_i$. □

B.5 Hermiticity and the spectral theorem

Definition B.12. *A matrix A is said to be Hermitian if $A = A^\dagger$. Notice that a real Hermitian matrix is also symmetric.*

Theorem B.13. *Any hermitian matrix A has real eigenvalues.*

Proof. Consider an eigenvalue λ with an eigenvector v and its complex conjugate λ^*.

$$\begin{aligned} \lambda v^\dagger v &= v^\dagger (Av) \\ &= v^\dagger A^\dagger v && \text{by Hermiticity} \\ &= (Av)^\dagger v \\ &= (\lambda v)^\dagger v \\ &= \lambda^* v^\dagger v \end{aligned}$$

Hence $\lambda = \lambda^*$. □

Theorem B.14 (Spectral theorem)**.** *If a matrix A is Hermitian, there exists an orthogonal basis of eigenvectors of A, i.e. it is unitarily diagonalizable.*

Proof. We prove this by induction on the size n. If $n = 1$, the claim is trivial as any unitary vector is an orthonormal eigenbasis of A. If $n > 1$, then by the fundamental theorem of algebra we must have n roots of $\det(A - \lambda \mathbb{1})$. Take one, and call it λ_1, along with an eigenvector v_1. Let $v_2, \ldots, v_n$ be an orthonormal basis for the subspace orthogonal to the one spanned by v_1. A can be rewritten as:

$$V^\dagger A V = \{v_i^\dagger A v_j\}_{ij} \begin{pmatrix} \lambda_1 & 0 \\ 0 & A' \end{pmatrix}$$

where V is a unitary matrix formed by $v_1, \ldots, v_n$, and A' is a $n-1 \times n-1$ matrix. By induction A' is unitarily diagonalizable by a matrix U with columns $u_2, \ldots, u_n$. Thus given the following:

$$V' = V \begin{pmatrix} 1 & 0 \\ 0 & U \end{pmatrix} \Longrightarrow (V')^\dagger A V' = \begin{pmatrix} \lambda_1 & 0 \\ 0 & U^\dagger A' U \end{pmatrix}$$

which is diagonal. □

Definition B.15 (Spectral decomposition)**.** *If U_λ is the matrix with columns formed by the eigenvectors associated with the eigenvalue λ, then $P_\lambda = U_\lambda U_\lambda^\dagger$ is the orthogonal projection matrix onto the eigenspace of λ. Any Hermitian matrix A with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ can be decomposed as follows:*

$$A = \lambda_1 P_{\lambda_1} + \cdots + \lambda_k P_{\lambda_k}$$

Definition B.16 (Eigendecomposition)**.** *A diagonalizable matrix can be written as:*

$$A = U \Lambda U^{-1}$$

where U is a unitary matrix containing all the eigenvectors of A as columns, and Λ is a diagonal matrix containing, in order, the eigenvalues.

Observation B.17. *For any matrix A, $A^\dagger A$ and $A A^\dagger$ are Hermitian.*

Theorem B.18. *If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, $\lambda_1 t, \ldots, \lambda_n t$ are the eigenvalues of At.*

Proof. If $A = U\Lambda U^{-1}$ is an eigendecomposition of A then

$$At = U\Lambda U^{-1}t = U(\Lambda t)U^{-1}$$

where $\Lambda t = \text{diag}(\lambda_1 t, \ldots, \lambda_n t)$. □

Theorem B.19. *If $A = U\Lambda U^{-1}$ is an eigendecomposition of A, then $A^k = U\Lambda^k U^{-1}$.*

Proof. We prove this by induction on k. If $k = 1$ the claim is trivial. If $k > 1$ we have:

$$\begin{aligned} A^k = (U\Lambda U^{-1})^k &= U\Lambda U^{-1}A^{k-1} \\ &= U\Lambda U^{-1}U\Lambda^{k-1}U^{-1} && \text{by induction} \\ &= U\Lambda^k U^{-1} && \text{since } U^{-1}U = \mathbb{1} \end{aligned}$$

□

Theorem B.20. *Let A be an $n \times n$ matrix with eigenvectors $v_1, \ldots, v_n$ associated with eigenvalues $\lambda_1, \ldots, \lambda_n$. The exponential e^A has eigenvectors $v_1, \ldots, v_n$ associated with the eigenvalues $e^{\lambda_1}, \ldots, e^{\lambda_n}$.*

Proof. Let $A = U\Lambda U^{-1}$ be an eigendecomposition of A. By definition of exponential:

$$\begin{aligned} e^A &= \sum_{k=0}^{\infty} \frac{1}{k!} A^k \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} U\Lambda^k U^{-1} && \text{by Theorem B.19} \\ &= U \left(\sum_{k=0}^{\infty} \frac{1}{k!} \Lambda^k \right) U^{-1} \\ &= U e^{\Lambda} U^{-1} \end{aligned}$$

We conclude the proof by showing that $e^{\Lambda} = \text{diag}(e^{\lambda_1, \ldots, \lambda_n})$, but this immediately follows from the fact that $\Lambda^k = \text{diag}(\lambda_1^k, \ldots, \lambda_n^k)$. □

B.6 Positive semi-definiteness

Definition B.21 (Positive semi-definiteness)**.** *An Hermitian matrix $A \in \mathbb{C}^{n \times n}$ is positive semi-definite ($A \succcurlyeq 0$) if and only if, for any $v \in \mathbb{C}^n$:*

$$v^\dagger A v \geq 0$$

Theorem B.22. *An Hermitian matrix A is positive semi-definite if and only if every eigenvalue of A is non-negative.*

Proof. Consider the spectral decomposition of A:

$$v^\dagger A v = \lambda_1 v_1^\dagger v_1 + \cdots + \lambda_n v_n^\dagger v_n$$

where v_i is the orthogonal projection onto the eigenspace of λ_i. Notice that:

- The inner product is necessarily real and positive as $z^* z = |z|^2$.

- Every eigenvalue is real since A is Hermitian.

□

Theorem B.23. *If A, B are two Hermitian positive (semi-)definite matrices, $A+B$ is Hermitian and positive (semi-)definite.*

Proof.

$$v^\dagger(A+B)v = v^\dagger Av + v^\dagger Bv \geq 0$$

□

Theorem B.24. *For any matrix A, $A^\dagger A$ is positive semi-definite.*

Proof. Let $z = Av$:

$$v^\dagger A^\dagger Av = (Av)^\dagger Av = z^\dagger z \geq 0$$

□

Chapter C
Recap on Probability Theory

In quantum theory we extensively use these concepts, although the notation may slightly differ.

C.1 Probability space and random variables

Definition C.1 (Probability space)**.** *A probability space is a tuple* $(\Omega, \mathcal{F}, \mathbf{P})$ *where:*

- Ω *is a non-empty set of elementary events;*
- $\mathcal{F} \subseteq \mathbf{2}^{\Omega}$ *is the* σ*-algebra of events;*
- $\mathbf{P} : \Omega \to [0, 1]$ *assigns a probability to each elementary event such that*

$$\sum_{\omega \in \Omega} \boldsymbol{P}(\omega) = 1$$

The following must hold:

- *The* σ*-algebra* $\mathcal{F}$ *must contain both* $\emptyset$ *and* Ω*, and it must be closed under any countably infinite intersection of events* $\{A_i\}_{i \in \mathbb{N}}$*:*

$$\forall i \; A_i \in \mathcal{F} \Longrightarrow \bigcap_i A_i \in \mathcal{F}$$

- *The definition of* $\mathbf{P}$ *is extended to* $\mathcal{F}$ *as follows:*

$$\boldsymbol{P}(A) = \sum_{\omega \in A} \boldsymbol{P}(\omega) \;\; \forall A \in \mathcal{F}$$

- $\boldsymbol{P}(A) = 0$ *if and only if* $A = \emptyset$*;*

Definition C.2 (Conditional probability)**.** *Let* $A, B \in \mathcal{F}$ *be two events in a probability space* $(\Omega, \mathcal{F}, \mathbf{P})$ *such that* $\boldsymbol{P}(B) \neq 0$*. The conditional probability is defined as:*

$$\boldsymbol{P}(A \mid B) = \frac{\boldsymbol{P}(A \cap B)}{\boldsymbol{P}(B)}$$

Theorem C.3 (Law of total probability)**.** *Let* $A_1, \ldots, A_n \in \mathcal{F}$ *be a partition of* Ω *in a probability space* $(\Omega, \mathcal{F}, \mathbf{P})$*. Then, for any event* $B \in \mathcal{F}$ *the following holds:*

$$\boldsymbol{P}(B) = \sum_{i=1}^{n} \boldsymbol{P}(B \mid A_i) \, \boldsymbol{P}(A_i)$$

Proof.

$$\mathbf{P}(B) = \mathbf{P}\left(\bigcup_{i=1}^{n} (B \cap A_i)\right) = \sum_{i=1}^{n} \mathbf{P}(B \cap A_i) = \sum_{i=1}^{n} \mathbf{P}(B \mid A_i) \, \mathbf{P}(A_i)$$

□

Definition C.4 (Random variable)**.** *A random variable in a probability space* $(\Omega, \mathcal{F}, \mathbf{P})$ *is a function* $X : \Omega \to \mathbb{R}$*.*

C.2 Expectation

Definition C.5. *Let X be a random variable defined under a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The expectation of X is defined as follows:*

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} \boldsymbol{P}(\omega) X(\omega)$$

or, equivalently (by regrouping events):

$$\mathbb{E}[X] = \sum_{x} x \boldsymbol{P}(X = x)$$

The sum can become an integral sum in case X is absolutely continuous (in this case $\boldsymbol{P}(X = x)$ is of the form $f(x)dx$, where $f(x)$ is called probability density function).

Theorem C.6. *Let X, Y be random variables defined under a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The following holds:*

1. $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$ *(linearity of expectation);*
2. $\inf_\omega X \leq \mathbb{E}[X] \leq \sup_\omega X$*;*

Proof. We prove the two statements separately:

1. Directly follows from linearity of the sum.
2. $$\inf_\omega X = \sum_{\omega \in \Omega} \mathbf{P}(\omega) \inf_\omega X \leq \sum_{\omega \in \Omega} \mathbf{P}(\omega) X(\omega) \leq \sum_{\omega \in \Omega} \mathbf{P}(\omega) \sup_\omega X = \sup_\omega X$$

□

C.3 Variance

Definition C.7. *Let X be a random variable defined under a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The variance of X can be defined as:*

$$Var[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$$

Theorem C.8. $Var[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.

Proof.

$$\begin{aligned}
\mathbb{E}\left[(X - \mathbb{E}[X])^2\right] &= \mathbb{E}\left[X^2 + \mathbb{E}[X]^2 + 2X\mathbb{E}[X]\right] \\
&= \mathbb{E}\left[X^2\right] + \mathbb{E}\left[\mathbb{E}[X]^2\right] - 2\mathbb{E}[X]^2 && \text{linearity of expectation} \\
&= \mathbb{E}\left[X^2\right] + \mathbb{E}[X]^2 - 2\mathbb{E}[X]^2 \\
&= \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 && \text{by Theorem C.6}
\end{aligned}$$

□

Theorem C.9. $Var[X] = Var[X + c]$ *for any* $c \in \mathbb{R}$.

Proof. The claim follows immediately by seeing that $(X + c) - \mathbb{E}[X + c] = X + c - \mathbb{E}[X] - c = X - \mathbb{E}[X]$. □

C.4 Gaussian distribution

An absolutely continuous random variable X defined under a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ is said to follow a Gaussian distribution (i.e. $X \sim \mathcal{N}(\mu, \sigma^2)$) if:

$$\mathbf{P}(X = x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Theorem C.10. $\mathbb{E}[X] = \mu$.

Proof.

$$\begin{aligned}
\mathbb{E}[X] &= \int_{\mathbb{R}} x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \int_{\mathbb{R}} (\sqrt{2}\sigma x + \mu) \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2} \sqrt{2}\sigma dx && \text{replacing } x \leftarrow \sqrt{2}\sigma x + \mu \\
&= \int_{\mathbb{R}} \sqrt{2}\sigma x \frac{1}{\sqrt{\pi}} e^{-x^2} dx + \mu \int_{\mathbb{R}} \frac{1}{\sqrt{\pi}} e^{-x^2} dx \\
&= \mu \int_{\mathbb{R}} \frac{1}{\sqrt{\pi}} e^{-x^2} dx && \text{since } xe^{-x} \text{ is odd} \\
&= \mu && \text{since } \int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}
\end{aligned}$$

□

Theorem C.11. $Var[X] = \sigma^2$.

Proof. To simplify the proof, we compute the variance of $X - \mu \sim \mathcal{N}(0, \sigma^2)$ (we know the variance does not change under translations by Theorem C.9).

$$\begin{aligned}
\mathrm{Var}[X] &= \mathrm{Var}[X - \mu] \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} x^2 e^{-x^2/2\sigma^2} dx - \mathbb{E}[(X-\mu)]^2 \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} x^2 e^{-x^2/2\sigma^2} dx && \text{by Theorem C.10} \\
&= \frac{2\sigma^2}{\sqrt{\pi}} \int_{\mathbb{R}} x^2 e^{-x^2} dx && \text{substitution } x \leftarrow \sqrt{2}\sigma x \\
&= \sigma^2 && \text{since } \int_{\mathbb{R}} x^2 e^{-x^2} dx = \frac{\sqrt{\pi}}{2}
\end{aligned}$$

□

Bibliography

[1] Benjamin Schumacher and Michael Westmoreland. *Quantum Processes Systems, and Information*. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511814006.

[2] John Goold, Marcus Huber, Arnau Riera, Lídia del Rio, and Paul Skrzypczyk. The role of quantum information in thermodynamics—a topical review. *Journal of Physics A: Mathematical and Theoretical*, 49(14):143001, feb 2016. DOI: 10.1088/1751-8113/49/14/143001. URL https://doi.org/10.1088/1751-8113/49/14/143001.

[3] Nicolas Brunner, Noah Linden, Sandu Popescu, and Paul Skrzypczyk. Virtual qubits, virtual temperatures, and the foundations of thermodynamics. *Phys. Rev. E*, 85:051117, May 2012. DOI: 10.1103/PhysRevE.85.051117. URL https://link.aps.org/doi/10.1103/PhysRevE.85.051117.