Quantum thermodynamics summer school: preliminaries

22–27 August 2021

Here we list the basic preliminary knowledge to understand the language that will be used in the lectures of the summer school. No panic – the odds are that you have already seen most (if not all) of the topics listed. We are simply making sure that we are all standing on the same – in this case quantum mechanical – ground.

Parts of notes are borrowed from the lecture notes typed by Lorenzo Laneve and Giulia Carocari in the fall semester of 2020, which followed the course "quantum physics for non-physicists" given by Lídia del Rio.

"No fear, don't you turn like Orpheus."

Sara Bareilles, *Orpheus*

Contents

1 Hilbert spaces, measurement, evolution, and system composition **3**
- 1.1 Hilbert spaces and the braket notation . . . 3
- 1.2 The wave function . . . 4
- 1.3 Measurements . . . 4
- 1.4 Probability of outcomes in a measurement . . . 5
- 1.5 Observables . . . 6
- 1.6 Post-measurement state . . . 7
- 1.7 Reversible Evolution . . . 8
- 1.8 Evolution in qubits: quantum gates . . . 8
- 1.9 Unitary dynamics . . . 9
- 1.10 Deriving the Schrödinger equation . . . 9
- 1.11 Composing systems and tensor product . . . 10
- 1.12 Entanglement and measurement . . . 11

2 Position and Momentum **14**
- 2.1 The momentum basis . . . 14
- 2.2 The momentum operator . . . 15
- 2.3 The position-momentum commutator . . . 16
- 2.4 Transformation of observables in the commutator . . . 17
- 2.5 Hamiltonian of a particle in one dimension . . . 18
- 2.6 Evolution of an observable: Ehrenfest's theorem . . . 18
- 2.7 Schrödinger equation for a free particle . . . 19

3 Modeling Uncertainty **21**
3.1 Information in quantum mechanics . . . 21
3.2 The density matrix . . . 22
3.3 Properties of the density matrix . . . 24
3.4 Ignorance about local information and the partial trace . . . 25
3.5 Superposition vs probabilistic mixture . . . 27
3.6 Entanglement vs probabilistic mixture . . . 29
3.7 Uncertainty about evolution . . . 32
3.8 Kraus decomposition . . . 33
3.9 Stinespring dilation . . . 34
3.10 Example with CNOT gate . . . 35
3.11 Measurements as quantum channels . . . 37
3.12 Intuition: mixed states in the Bloch sphere . . . 39
3.13 Heisenberg's uncertainty principle . . . 40

4 Dynamics of Open Systems **42**
4.1 Schrödinger equation for mixed states . . . 42
4.2 Open systems . . . 42
4.3 Lindblad equation . . . 43

5 Physical implementation of measurements **47**
5.1 Indirect measurement and the modular momentum . . . 47
5.2 The Stern-Gerlach experiment . . . 51

Appendix **55**

A Recap on Analysis **56**
A.1 Exponential function . . . 56
A.2 Dirac delta function . . . 56
A.3 Complex numbers . . . 57
A.4 Fourier transform . . . 58
A.5 Solving differential equations . . . 60

B Recap on Linear Algebra **62**
B.1 Properties of the trace . . . 62
B.2 Inner product spaces . . . 62
B.3 Unitary matrices . . . 63
B.4 Eigenvalues and eigenvectors . . . 64
B.5 Hermiticity and the spectral theorem . . . 65
B.6 Positive semi-definiteness . . . 66

C Recap on Probability Theory **68**
C.1 Probability space and random variables . . . 68
C.2 Expectation . . . 69
C.3 Variance . . . 69
C.4 Gaussian distribution . . . 70

Chapter 1
Hilbert spaces, measurement, evolution, and system composition

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.

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:

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

1.7 Reversible Evolution

1.8 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:

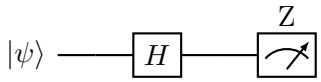
$$M = V\Lambda V^\dagger = \sum_i \lambda_i |v_i\rangle\langle v_i| \implies M' = U(V\Lambda V^\dagger)U^\dagger = \sum_i \lambda_i U|v_i\rangle\langle v_i|U^\dagger$$

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.

1.9 Unitary dynamics

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

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

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.

1.10 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:

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

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

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

Theorem 1.4. *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**:

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

1.11 Composing systems and 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:

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

1.12 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 1.5 (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:

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

implying that the induced probability space gives:

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

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:

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

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).

Chapter 2
Position and Momentum

2.1 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]:

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

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 2.1. *$\{|p\rangle\}_p$ 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 $\langle X \rangle$.

2.2 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:

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

2.3 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 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.

2.4 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 2.2. *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 2.2, 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 2.3. $[X^n, P] = i\hbar \cdot nX^{n-1}$

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:

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

2.5 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.

2.6 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 2.4 (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 2.5 (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.

2.7 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 2.2, 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.

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 3
Modeling Uncertainty

3.1 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.

3.2 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**[3]. 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:

$$\begin{aligned}\mathbf{P}(x)_{U(\mathcal{B})} &= \sum_i p_i \mathbf{P}(x)_{U(\psi_i)} \\ &= \sum_i p_i \operatorname{Tr}(\Pi_x|\psi_i'\rangle\langle\psi_i'|) \\ &= \sum_i p_i \operatorname{Tr}\left(\Pi_x U|\psi_i\rangle\langle\psi_i|U^\dagger\right)\end{aligned}$$

[3]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.

3.3 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 3.1. $\mathrm{Tr}(\rho) = 1$.

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 3.2. *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 3.3. *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 3.2 ensures that the eigenbasis of ρ is orthonormal, which is a property we always appreciate in quantum theory. Theorems 3.1 and 3.3 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 3.2, 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 $\rho = \frac{\mathbb{1}}{|\mathcal{H}|}$[4], 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 $\mathcal{H}$.

3.4 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 3.4 (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.

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:

$$\mathrm{Tr}\left(\Pi_x \rho_A\right) = \mathrm{Tr}\left((\Pi_x \otimes \mathbb{1}_B)\rho_{AB}\right)$$

[4] Although this is an extreme abuse of notation, you can imagine that this also works if $\mathcal{H}$ is infinite-dimensional.

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:

$$\begin{aligned}\mathrm{Tr}_B(\rho'_{AB}) &= \mathrm{Tr}_B\left((U_A \otimes V_B)\rho_{AB}(U_A^\dagger \otimes V_B^\dagger)\right) \\ &= \sum_{j} \left(\mathbb{1}_A \otimes \langle j|_B\right) (U_A \otimes V_B)\rho_{AB}(U_A^\dagger \otimes V_B^\dagger) \left(\mathbb{1}_A \otimes |j\rangle_B\right) \\ &= \sum_{j} \left(U_A \otimes \langle j|_B V_B\right) \rho_{AB} \left(U_A^\dagger \otimes V_B^\dagger |j\rangle_B\right)\end{aligned}$$

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:

$$\mathrm{Tr}_B(\rho'_{AB}) = \sum_{j} \left(U_A \otimes \langle j|_B\right) \rho_{AB} \left(U_A^\dagger \otimes |j\rangle_B\right)$$

$$= U \left(\sum_j \left(\mathbb{1}_A \otimes \langle j|_B \right) \rho_{AB} \left(\mathbb{1}_A \otimes |j\rangle_B \right) \right) U^\dagger$$
$$= U \rho_A U^\dagger$$

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 3.5 (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.

$$\begin{aligned}
\mathrm{Tr}_B(\alpha\rho_1 + \beta\rho_2) &= \sum_j \left(\mathbb{1}_A \otimes \langle j|_B \right) (\alpha\rho_1 + \beta\rho_2) \left(\mathbb{1}_A \otimes |j\rangle_B \right) \\
&= \sum_j \left(\mathbb{1}_A \otimes \langle j|_B \right) (\alpha\rho_1) \left(\mathbb{1}_A \otimes |j\rangle_B \right) + \sum_j \left(\mathbb{1}_A \otimes \langle j|_B \right) (\beta\rho_2) \left(\mathbb{1}_A \otimes |j\rangle_B \right) \\
&= \alpha \sum_j \left(\mathbb{1}_A \otimes \langle j|_B \right) \rho_1 \left(\mathbb{1}_A \otimes |j\rangle_B \right) + \beta \sum_j \left(\mathbb{1}_A \otimes \langle j|_B \right) \rho_2 \left(\mathbb{1}_A \otimes |j\rangle_B \right) \\
&= \alpha \,\mathrm{Tr}_B(\rho_1) + \beta \,\mathrm{Tr}_B(\rho_2)
\end{aligned}$$

□

Theorem 3.6. $\mathrm{Tr}_B(\rho_A \otimes \rho_B) = \rho_A$.

Proof.

$$\begin{aligned}
\mathrm{Tr}_B(\rho_A \otimes \rho_B) &= \sum_j \left(\mathbb{1}_A \otimes \langle j|_B \right) (\rho_A \otimes \rho_B) \left(\mathbb{1}_A \otimes |j\rangle_B \right) \\
&= \sum_j \left(\rho_A \otimes \langle j|_B \rho_B \right) \left(\mathbb{1}_A \otimes |j\rangle_B \right) \\
&= \sum_j \left(\rho_A \otimes \langle j|_B \rho_B \right) \left(\mathbb{1}_A \otimes |j\rangle_B \right) \\
&= \rho_A \sum_j \langle j|\rho_B|j\rangle \\
&= \rho_A \,\mathrm{Tr}(\rho_B) = \rho_A \qquad \text{by Theorem 3.1}
\end{aligned}$$

□

3.5 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 3.7. *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*.

3.6 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 3.6, 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.

3.7 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 3.1 and thus still be expressed as probabilistic mixture;

- **completely positive**: the new state $\mathcal{E}(\rho)$ must also preserve Theorem 3.3, 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.

3.8 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 3.8. *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 3.9. *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.

3.9 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**.

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 is still a reversible, unitary evolution we are already familiar with. We can define the mapping as follows:
$$\begin{aligned}\mathcal{E}_{A\mapsto B}(\rho_A) &= \operatorname{Tr}_{E'}\left(U_{AE}(\rho_A \otimes |0\rangle\langle 0|_E)U_{AE}^\dagger\right) \\ &= \operatorname{Tr}_{E'}\left(U_{AE}(\mathbb{1}_A \otimes |0\rangle_E)\rho_A(\mathbb{1}_A \otimes \langle 0|_E)U_{AE}^\dagger\right) \\ &=: \operatorname{Tr}_{E'}\left(V\rho_A V^\dagger\right)\end{aligned}$$
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 $V^\dagger V = \mathbb{1}$[5]:

$$\begin{aligned}
\mathbb{1}_A &= V^\dagger V \\
&= \sum_{k,\ell} (E_k^\dagger \otimes \langle k|)(E_\ell \otimes |\ell\rangle) \\
&= \sum_{k,\ell} E_k^\dagger E_\ell \langle k|\ell\rangle \\
&= \sum_k E_k^\dagger E_k
\end{aligned}$$

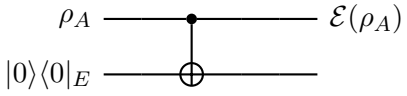
i.e. E_k are valid Kraus operators. Now let us take the partial trace:

$$\begin{aligned}
\mathcal{E}(\rho_A) &= \mathrm{Tr}_{E'}(V \rho_A V^\dagger) \\
&= \mathrm{Tr}_{E'} \left(\left(\sum_k E_k \otimes |k\rangle_E \right) \rho_A \left(\sum_\ell E_k^\dagger \otimes \langle \ell|_E \right) \right) \\
&= \mathrm{Tr}_{E'} \left(\sum_{k,\ell} E_k \rho_A E_\ell^\dagger \otimes |k\rangle\langle \ell|_E \right) \\
&= \sum_j (\mathbb{1}_A \otimes \langle j|) \left(\sum_{k,\ell} E_k \rho_A E_\ell^\dagger \otimes |k\rangle\langle \ell|_E \right) (\mathbb{1}_A \otimes |j\rangle) \\
&= \sum_{j,k,\ell} E_k \rho_A E_\ell^\dagger \langle j|k\rangle\langle \ell|j\rangle \\
&= \sum_k E_k \rho_A E_k^\dagger
\end{aligned}$$

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 5.1, when we showed how a global evolution could be written in a different way, exploiting the linearity of the tensor product.

3.10 Example with CNOT gate

We see here a simple example of quantum channel, with a global reversible operation we know: the CNOT.



[5]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.

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:

$$\begin{aligned} \rho_{AE} = \rho_A \otimes |0\rangle\langle 0|_E &= \sum_{i=0}^{1}\sum_{j=0}^{1} c_{ij}|i\rangle\langle j|_A \otimes |0\rangle\langle 0|_E \\ &= \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} c_{00} & 0 & c_{10} & 0 \\ 0 & 0 & 0 & 0 \\ c_{01} & 0 & c_{11} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

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:

$$\begin{aligned} \mathcal{E}(\rho_A) &= \mathrm{Tr}_{E'}\left(U(\rho_A \otimes |0\rangle\langle 0|)U^\dagger\right) \\ &= \mathrm{Tr}_{E'}\left(\sum_{i=0}^{1}\sum_{j=0}^{1} c_{ij}|i\rangle\langle j| \otimes |i\rangle\langle j|\right) \\ &= \sum_{i=0}^{1}\sum_{j=0}^{1} c_{ij}\,\mathrm{Tr}_{E'}\left(|i\rangle\langle j| \otimes |i\rangle\langle j|\right) && \text{linearity of partial trace} \\ &= \sum_{i=0}^{1}\sum_{j=0}^{1} c_{ij}\sum_{\ell}(\mathbb{1}_B \otimes \langle \ell|)(|i\rangle\langle j| \otimes |i\rangle\langle j|)(\mathbb{1}_B \otimes |\ell\rangle) \\ &= \sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{\ell} c_{ij}|i\rangle\langle j|\langle \ell|i\rangle\langle j|\ell\rangle \\ &= \sum_{i=0}^{1} c_{ii}|i\rangle\langle i| = \begin{pmatrix} c_{00} & 0 \\ 0 & c_{11} \end{pmatrix} \end{aligned}$$

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.

3.11 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 5.1 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 $|\psi\rangle = \sum_k c_k|k\rangle \otimes |\psi_0\rangle$:

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

$$=: \sum_k c_k |k\rangle \otimes |\psi_k\rangle$$

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:

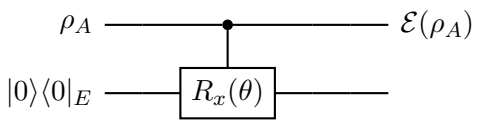
$$\begin{aligned}
\mathcal{E}(\rho_A) &= \mathrm{Tr}_x \left(\sum_{k,j} c_k c_j^* |k\rangle\langle j| \otimes |\psi_k\rangle\langle\psi_j| \right) \\
&= \sum_{k,j} c_k c_j^* \, \mathrm{Tr}_x \left(|k\rangle\langle j| \otimes |\psi_k\rangle\langle\psi_j| \right) \\
&= \sum_{k,j} c_k c_j^* \int_{\mathbb{R}} \left(\mathbb{1}_A \otimes \langle x|_x \right) \left(|k\rangle\langle j| \otimes |\psi_k\rangle\langle\psi_j| \right) \left(\mathbb{1}_A \otimes |x\rangle_x \right) dx \\
&= \sum_{k,j} c_k c_j^* |k\rangle\langle j| \int_{\mathbb{R}} \langle x|\psi_k\rangle\langle\psi_j|x\rangle dx \\
&= \sum_{k,j} c_k c_j^* |k\rangle\langle j| \int_{\mathbb{R}} \langle\psi_j|x\rangle\langle x|\psi_k\rangle dx \\
&= \sum_{k,j} c_k c_j^* |k\rangle\langle j| \langle\psi_j| \left(\int_{\mathbb{R}} |x\rangle\langle x| dx \right) |\psi_k\rangle \\
&= \sum_{k,j} c_k c_j^* |k\rangle\langle j| \langle\psi_j|\psi_k\rangle \\
&\simeq \sum_k |c_k|^2 |k\rangle\langle k| \qquad\qquad \text{overlap } \langle\psi_k|\psi_{k'}\rangle \simeq \delta(k - k')
\end{aligned}$$

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:

$$\sum_k \Pi_k^\dagger \Pi_k = \sum_k \Pi_k^2 = \sum_k \Pi_k = \mathbb{1}$$

3.12 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: a state $|\psi\rangle$ is identified by two spherical coordinates (θ, ϕ):

$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$$

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).

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, 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.

3.13 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 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 3.10. $[\Delta A, \Delta B] = [A, B]$.

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 3.11. *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 3.12. *For any two Hermitian operators A, B:*

1. *$[A, B]$ is anti-Hermitian;*
2. *$\{A, B\}$ is Hermitian;*
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 3.13 (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.
$$\begin{aligned}
\Delta A^2 \Delta B^2 &= \langle\psi|(\Delta A)^2|\psi\rangle\langle\psi|(\Delta B)^2|\psi\rangle \\
&= \langle\psi|(\Delta A)^\dagger(\Delta A)|\psi\rangle\langle\psi|(\Delta B)^\dagger(\Delta B)|\psi\rangle && \Delta A, \Delta B \text{ are Hermitian} \\
&\geq |\langle\psi|(\Delta A)(\Delta B)|\psi\rangle|^2 && \text{Cauchy-Schwartz inequality} \\
&= \left|\langle\psi|\left(\frac{1}{2}[\Delta A, \Delta B] + \frac{1}{2}\{\Delta A, \Delta B\}\right)|\psi\rangle\right|^2 && \text{by Lemma 3.11} \\
&= \frac{1}{4}\left|\langle\psi|[A,B]|\psi\rangle + \langle\psi|\{\Delta A, \Delta B\}|\psi\rangle\right|^2 && \text{by Lemma 3.10}
\end{aligned}$$

By Lemma 3.12 the first term is purely imaginary and, since $\{A,B\}$ is Hermitian, it has real labels and the expectation is real[6]. 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 3.14 (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.

Proof. It is sufficient to apply Theorem 3.13 where $[X,P] = i\hbar\mathbb{1}$ is the canonical commutation relation we derived in Section 2.3. □

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. In fact, for Gaussian wavepackets:
$$\Delta X = \sigma \Longrightarrow \Delta P = \frac{\hbar}{2\sigma}$$
which also proves that the bound is tight.

[6]The expectation of a real random variable is always real, convince yourself!

Chapter 4
Dynamics of Open Systems

4.1 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:

$$\frac{d\rho}{dt} = \frac{1}{i\hbar}[H,\rho] \tag{4.1}$$

and this is the time-independent Schrödinger equation for mixed states.

4.2 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.

4.3 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 3.13, 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 4.1, 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:

$$\frac{d\rho}{dt} = \frac{1}{i\hbar}[H_S, \rho] - \frac{1}{2}\left\{\sum_{k=1}^{\infty} L_k^\dagger L_k, \rho\right\} + \sum_{k=1}^{\infty} L_k \rho L_k^\dagger \tag{4.2}$$

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:

$$\rho(t) = \begin{pmatrix} 1 - c_{11}(0)e^{-\lambda^2 t} & c_{01}(0)e^{-\lambda^2 t/2} \\ c_{10}(0)e^{-\lambda^2 t/2} & c_{11}(0)e^{-\lambda^2 t} \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = |0\rangle\langle 0|.$$

Chapter 5
Physical implementation of measurements

5.1 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. 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:

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

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):

$$\begin{aligned}
U(t)|\tilde{\psi}_0\rangle &= \exp\left(-\frac{it}{\hbar}\left(M_A \otimes P_x\right)\right)\left(\sum_k c_k |k\rangle_A \otimes |\psi_0\rangle_x\right) \\
&= \exp\left(-\frac{it}{\hbar}\left(\sum_k a_k |k\rangle\langle k|_A \otimes P_x\right)\right)\left(\sum_k c_k |k\rangle_A \otimes |\psi_0\rangle_x\right) \\
&= \sum_{n=0}^{\infty} \frac{1}{n!}\left(-\frac{it}{\hbar}\right)^n \left(\sum_k a_k |k\rangle\langle k|_A \otimes P_x\right)^n \left(\sum_k c_k |k\rangle_A \otimes |\psi_0\rangle_x\right) \\
&= \sum_{n=0}^{\infty} \frac{1}{n!}\left(-\frac{it}{\hbar}\right)^n \left(\left(\sum_k a_k |k\rangle\langle k|_A\right)^n \otimes P_x^n\right)\left(\sum_k c_k |k\rangle_A \otimes |\psi_0\rangle_x\right) \\
&= \sum_{n=0}^{\infty} \frac{1}{n!}\left(-\frac{it}{\hbar}\right)^n \left(\sum_k a_k^n |k\rangle\langle k|_A \otimes P_x^n\right)\left(\sum_k c_k |k\rangle_A \otimes |\psi_0\rangle_x\right) \\
&= \sum_k \sum_{k'} \sum_{n=0}^{\infty} \frac{1}{n!}\left(-\frac{it}{\hbar}\right)^n \left(a_k^n |k\rangle\langle k|_A \otimes P_x^n\right)\left(c_{k'} |k'\rangle_A \otimes |\psi_0\rangle_x\right)
\end{aligned}$$

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

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:

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

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:

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

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 5.1, and generally acts on the position wave function as

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

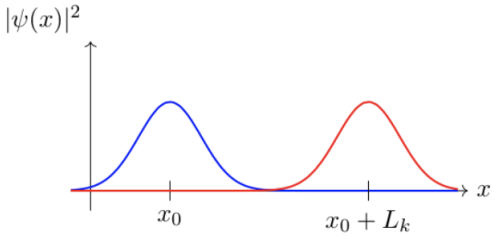


Figure 5.1: The modular momentum operator shifts the position wave function by L_k towards the positive direction of the axis.

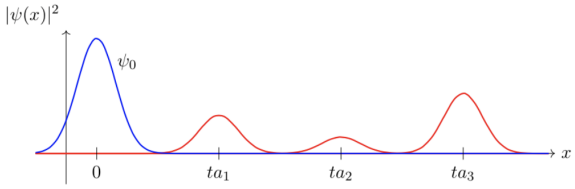


Figure 5.2: The initial Gaussian wave function (blue) evolves to become a superposition of Gaussians with different means (red).

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

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

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 5.2) 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 5.3). 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

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

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

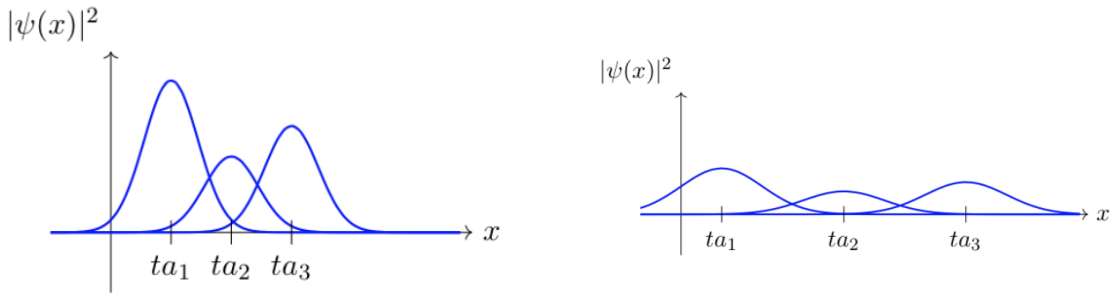


Figure 5.3: Examples of weak measurements due to small t (left) and high initial uncertainty in position (right).

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:

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

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

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

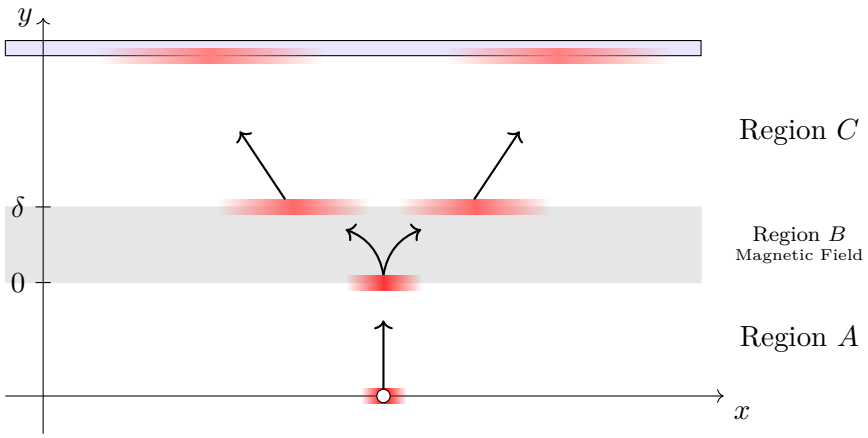


Figure 5.4: A scheme for the Stern-Gerlach experiment, which allows to indirectly determine the spin of an electron by looking at its position.

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.

5.2 The Stern-Gerlach experiment

Here we present a more concrete example of the Stern-Gerlach setup. 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 5.4. 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.

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

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, 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, $\langle P\rangle = (0, \hbar k)^T$).

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 5.5. The computation for a general initial Σ requires minor changes, which are not relevant here.

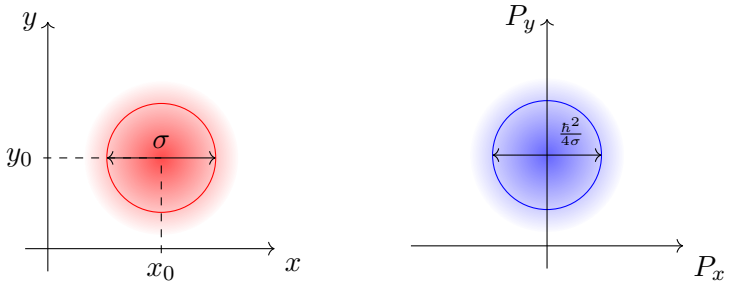


Figure 5.5: Gaussian waves in position (left) and momentum (right) basis of the initial state represented from above.

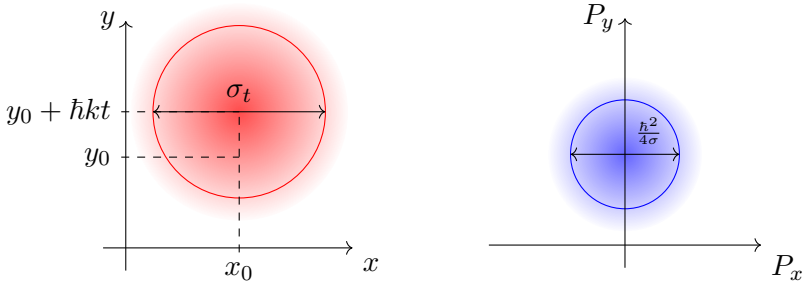


Figure 5.6: Gaussian waves in position and momentum basis for the state after the evolution in region A.

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:

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

And for Gaussian wave packets:

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

As before, we give some intuituion for what the waves packets look like in Figure 5.6.

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:

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

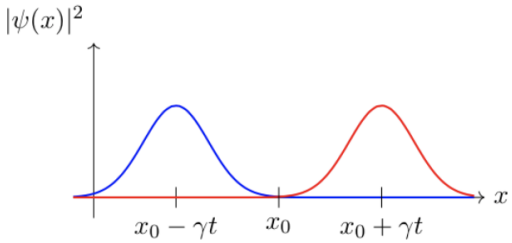


Figure 5.7: The wave function resulting from the evolution in region B.

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 5.1. 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 5.1 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 5.1, 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 5.7).

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:

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

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

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 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:

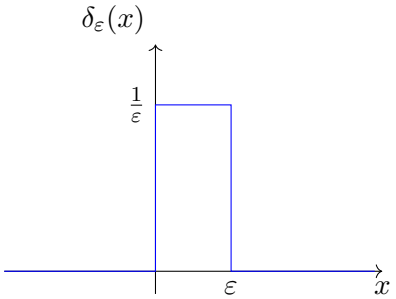
$$e^x = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

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 $\Lambda t = \operatorname{diag}(\lambda_1 t, \ldots, \lambda_n t)$. □

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 $e^{\Lambda} = \operatorname{diag}(e^{\lambda_1, \ldots, \lambda_n})$, but this immediately follows from the fact that $\Lambda^k = \operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k)$. □

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.