Quantum Mechanics 2

Dr Juan Rojo
VU Amsterdam and Nikhef Theory Group
http://www.juanrojo.com/, j.rojo@vu.nl
Current version: **January 28, 2021**

1 Chapter 3: The Formalism of Quantum Mechanics

Learning Goals

- To describe quantum states in terms of the Dirac notation.
- To apply the language of linear algebra to describe the basic operations of quantum mechanics.
- To determine which conditions a linear operator must satisfy to represent a physical observable and what are the implications of these conditions.
- To utilise the generalised statistical interpretation to predict the outcome of measurements in quantum systems.
- To generalise Heisenberg's uncertainty principle to an arbitrary pair of observables.
- To apply the formalism of quantum mechanics to simple finite-dimensional Hilbert spaces.

In this section of the lecture notes, we present the main concepts discussed in **Chapter 3 ("Formalism")** of the course textbook. The goal of these lecture notes is to provide a self-consistent study resource for the students, which is then complemented by the live lectures (and their recordings), the tutorial sessions, as well as their own study of the textbook. The relevant textbook sections are indicated below, material from other sections not listed there will not be required for the examination .

Textbook sections

- **3.1**: Hilbert space.
- **3.2**: Observables.
- **3.3**: Eigenfunctions of a Hermitian operator.
- **3.4**: Generalised statistical interpretation.
- **3.5**: The uncertainty principle.
- **3.6**: Vectors and operators.

In previous courses, in particular in *Quantumfysica 1* and *Quantum Concepten*, you have been introduced to various important concepts about quantum mechanics, such as the wave function, the Schroedinger equation, and Heisenberg's uncertainty principle. Now we would like to take a step back and present the formalism of quantum mechanics in a more abstract way. While this formulation might appear at first disconnected from your previous knowledge of the quantum world, we will show how the formalism of quantum mechanics is rather more powerful, and how once we have the right tools we will be able to readily reproduce and derive the results and equations that you are already familiar with.

1.1 Hilbert spaces and Dirac notation

Perhaps one of the most far-reaching concepts in quantum mechanics is that physical states are members of an abstract vector space called the *Hilbert space* and thus they obey the standard rules of linear algebra. This might not seem a world-shattering statement, but as we will see it has striking consequences, such as those leading to the famous Schroedinger cat which is nor-dead-nor-alive.

The formalism of quantum mechanics is built upon two fundamental concepts:

- The state of quantum system is characterised by its **state vector** $|\alpha\rangle$, an element of an abstract complex vector space known as the **Hilbert space** $\mathcal{H}$, $|\alpha\rangle \in \mathcal{H}$. *All physical information* about a given system is encapsulated in its state vector.
- These state vectors are modified by **linear operators** that act upon them and that determine physical properties, such as for example their time evolution.

Mathematically, quantum mechanics is most naturally described by the language of **linear algebra**.

You can think by analogy of the vector space in Cartesian coordinates living $\mathbb{R}^3$. There, we have a basis composed by three linearly independent vectors, say $\vec{v}_1 = (1,0,0)$, $\vec{v}_2 = (0,1,0)$, and $\vec{v}_3 = (0,0,1)$, and *any other element* of the same vector space admits an expression of the form $\vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + a_3\vec{v}_3$ with a_i being real coefficients. The same holds for quantum mechanics, with the difference that the now the basis elements correspond to specific quantum states.

Dirac notation. The formalism of quantum mechanics can be neatly formulated by means of the renowned *Dirac notation*. Assume that we have a quantum system characterised by an n-dimensional Hilbert space. This means that each possible state of the system can be represented by a state vector $|\alpha\rangle$ with n complex components, which we will write as a column vector

$$|\alpha\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} . \tag{1.1}$$

We denote the state vector in the Dirac notation $|\alpha\rangle$ as a *ket*, for reasons that will become obvious soon. Since the Hilbert space is a complex vector space, the components $\{a_i\}$ will be **complex numbers**.

Note that, as for any vector space, the values of these components depend on the specific choice of **basis**.

For example, a possible basis of this Hilbert space $\mathcal{H}$ could be

$$|\alpha_1\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad |\alpha_2\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad |\alpha_n\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \tag{1.2}$$

and in this basis one can express our original state vector $|\alpha\rangle$ as

$$|\alpha\rangle = a_1|\alpha_1\rangle + a_2|\alpha_2\rangle + \ldots + a_n|\alpha_n\rangle\,. \tag{1.3}$$

While the state vector $|\alpha\rangle$ itself is basis-independent, the values of its components $\{a_i\}$ will depend on the choice of basis. An appropriate choice of basis vectors for our Hilbert space can simplify many problems.

Formulating quantum mechanics as a linear algebra problem brings in a lot of perks. For example, we know that if we multiply a vector by a scalar, we end up with another *bona fide* vector. The same holds in Hilbert space: if the state vector $|\alpha\rangle$ belongs to the Hilbert space $\mathcal{H}$, then $|\beta\rangle = \gamma|\alpha\rangle$, with $\gamma \in \mathbb{C}$ an arbitrary complex number, is also an element of the same Hilbert space with components

$$|\beta\rangle = \gamma|\alpha\rangle = \begin{pmatrix} \gamma\, a_1 \\ \gamma\, a_2 \\ \vdots \\ \gamma\, a_n \end{pmatrix} \in \mathcal{H}\,. \tag{1.4}$$

While this property appears at first a tad too abstract, note that has physical consequences: it tells us that the state vector $|\beta\rangle$ represents another **physically allowed state** of our quantum system.

A two-dimensional Hilbert space: the electron spin

A particularly simple, yet extremely important, quantum system is that of the **spin of the electron**. An electron spin can adopt two configurations: it can point "up" (along some axis), or it can point "down". So this system has associated a **two-dimensional Hilbert space** ($n = 2$). If the two basis state vectors are chosen to correspond to the "up" and "down" directions,

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{1.5}$$

then a general element of this Hilbert space $\mathcal{H}$ can be expressed as

$$|\alpha\rangle = c_+|+\rangle + c_-|-\rangle = \begin{pmatrix} c_+ \\ c_- \end{pmatrix}, \qquad c_+, c_- \in \mathbb{C}\,. \tag{1.6}$$

You can then see how this very simple quantum system actually can host an **infinite number** of different states, corresponding to different choices of the coefficients c_i. This radical departure with classical physics is the underpinning of *e.g.* the greatly superior performance of quantum computers.

By refreshing your linear algebra, you can easily derive yourselves many of the relevant properties of state vectors in quantum mechanics. For example, if $|\alpha\rangle$ and $|\beta\rangle$ represent two possible states of our quantum system (that is, they are two elements of the Hilbert space $\mathcal{H}$ associated to the system), then their arbitrary

linear superposition will also belong to the same Hilbert space,

$$|\rho\rangle \equiv \gamma|\alpha\rangle + \lambda|\beta\rangle = \begin{pmatrix} \gamma\, a_1 + \lambda\, b_1 \\ \gamma\, a_2 + \lambda\, b_2 \\ \vdots \\ \gamma\, a_n + \lambda\, b_n \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_n \end{pmatrix} \in \mathcal{H}\,, \tag{1.7}$$

and thus can be associated to another possible physical state of the quantum system. In this course we will extensively exploit this connection between the space of quantum states and the properties of linear algebra.

Dual Hilbert space and internal product. As you have learned in your linear algebra course, another important property of vector spaces is the existence of an **internal product** between two elements of the space. For example, you are familiar with the *scalar product*, whereby I can multiply two vectors, say elements of $\mathbb{R}^n$, to obtain a real scalar:

$$\vec{a}, \vec{b} \in \mathbb{R}^n\,, \quad \vec{a}\cdot\vec{b} = \sum_{i=1}^{n} a_i b_i \in \mathbb{R}\,, \tag{1.8}$$

when the components of $\vec{a}$ and $\vec{b}$ correspond to the same (orthonormal) basis. Inner products are essential to define many of the properties of a vector space. For example, the *magnitude* of a vector is defined in terms of the scalar product with itself,

$$|\vec{a}| = \sqrt{\vec{a}\cdot\vec{a}} = \sqrt{\sum_{i=1}^{n} a_i^2}\,, \tag{1.9}$$

while the *distance* between two points in $\mathbb{R}^n$ is also evaluated in terms of the inner product

$$d(\vec{a},\vec{b}) = |\vec{a}-\vec{b}| = \sqrt{\left(\vec{a}-\vec{b}\right)\cdot\left(\vec{a}-\vec{b}\right)} = \sqrt{\sum_{i=1}^{n}\left(a_i-b_i\right)^2}\,. \tag{1.10}$$

Surely, if quantum mechanics is formulated in the language of linear algebra, there also needs to be some kind of inner product between state vectors?

In order to define an inner product for state vector, we need to introduce first the concept of *dual Hilbert space* $\mathcal{H}^*$. Putting aside mathematical subtleties, if a n-dimensional Hilbert space is the vector space spanned by (column) state vectors of the form

$$|\alpha\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}\,, \tag{1.11}$$

then its dual $\mathcal{H}^*$ is the vector space spanned by elements of the form

$$\langle\alpha| = \left(a_1^*, a_2^*, \cdots, a_n^*\right)\,, \tag{1.12}$$

where always following the Dirac notation we denote $\langle\alpha|$ by a *bra* vector. Note that to move from a state vector $|\alpha\rangle \in \mathcal{H}$ to its dual $\langle\alpha| \in \mathcal{H}^*$ is have transposed the column vector into a row vector and taken the complex conjugate of its components.

Armed with the notion of a dual space, we can now introduce the inner product in Hilbert space. If we have two state vectors living in the same Hilbert space, $|\alpha\rangle, |\beta\rangle \in \mathcal{H}$, then their inner product is defined as

$$\langle\alpha|\beta\rangle \equiv (a_1^*, a_2^*, \cdots, a_n^*) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \sum_{i=1}^{n} a_i^* b_i \,. \tag{1.13}$$

Note that the inner product is a (complex) **scalar**, and therefore the result of $\langle\alpha|\beta\rangle$ will be the same irrespectively of the choice of basis. For obvious reasons, the object $\langle\alpha|\beta\rangle$ is called a **braket**, explaining why we denote dual state vectors $\langle\alpha|$ as *bras* and state vectors $|\beta\rangle$ as *kets*. in the Dirac notation. The inner product between two elements $|\alpha\rangle$ and $|\beta\rangle$ of a vector space is a measure of their **overlap** in the Hilbert space. A pair of vectors such as $\langle\alpha|\beta\rangle = 0$ are said to be *orthogonal* (in the Hilbert space sense, not in the Cartesian vector space sense). In this course we will almost always work with *normalised state vectors*, which are those for which satisfy $\langle\alpha|\alpha\rangle = 1$ in analogy with Eq. (1.9) (unit norm vectors).

Let's go back to the quantum system describing the spin of an electron. The basis vectors $|+\rangle$ and $|-\rangle$ are obviously orthonormal. Now assume that we have two elements of this Hilbert space:

$$|\alpha\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle - i|-\rangle\right) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} ,$$

$$|\beta\rangle = \left(b_1|+\rangle + b_2|-\rangle\right) = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} .$$

We would like to determine b_1 and b_2 such that $|\beta\rangle$ is normalised and orthogonal to $|\alpha\rangle$. First of all imposing the orthogonality condition we find

$$\langle\alpha|\beta\rangle = \frac{1}{\sqrt{2}}(1, +i) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = 0 \qquad b_1 = -ib_2 \,, \tag{1.14}$$

and then imposing the normalisation condition

$$\langle\beta|\beta\rangle = (+ib_2^*, b_2^*) \begin{pmatrix} -ib_2 \\ b_2 \end{pmatrix} = 1 \qquad b_2 = \frac{1}{\sqrt{2}} \,. \tag{1.15}$$

In this course we will make extensive user of the inner product in Hilbert spaces in order to understand and quantify the properties of quantum systems.

Linear superposition and completeness Let us further exploit the fact that quantum systems are described by the formalism of linear algebra and vector spaces to derive a number of important properties. Assume that we have an n-dimensional Hilbert space $\mathcal{H}$ equipped with a basis of orthonormal vectors $\{|\psi_i\rangle\}$. This set of basis vectors is said to be **complete**, since I can express any element $|\Psi\rangle$ of the same Hilbert space as a linear superposition of the basis elements, that is

$$|\Psi\rangle = \sum_{i=1}^{n} c_i |\psi_i\rangle \,, \quad \forall\, |\Psi\rangle \in \mathcal{H} \,, \quad c_i \in \mathbb{C} \,. \tag{1.16}$$

The complex coefficients c_i can be evaluated by taking the inner product of $|\Psi\rangle$ with the basis vector $|\psi_j\rangle$,

$$\langle\psi_j|\Psi\rangle = \langle\psi_j|\left(\sum_{i=1}^{n} c_i|\psi_i\rangle\right) = \sum_{i=1}^{n} c_i\langle\psi_j|\psi_i\rangle = c_j\,, \tag{1.17}$$

where I have used that the basis vectors are orthonormal and hence $\langle\psi_j|\psi_i\rangle = \delta_{ij}$ with δ_{ij} the Kronecker delta. A particularly elegant question to express the concept of **completeness** is given by the fact that the basis vectors satisfy

$$\sum_{i=1}^{n} |\psi_i\rangle\langle\psi_i| = 1\,. \tag{1.18}$$

Note that the LHS of Eq. (1.18) is an *operator* acting on elements of the Hilbert space (more about operators very soon). To demonstrate this relation, we can act with the LHS of Eq. (1.18) onto the general element of the Hilbert space Eq. (1.16) and then we find

$$\left(\sum_{i=1}^{n} |\psi_i\rangle\langle\psi_i|\right)|\Psi\rangle = \left(\sum_{i=1}^{n} |\psi_i\rangle\langle\psi_i|\right)\left(\sum_{j=1}^{n} c_j|\psi_j\rangle\right) = \sum_{i,j=1}^{n} c_j|\psi_i\rangle\langle\psi_i|\psi_j\rangle = |\Psi\rangle\,, \tag{1.19}$$

which since $|\Psi\rangle$ is fully general confirms Eq. (1.18).

Infinite-dimensional Hilbert spaces. At this point, you might be convincing ourselves that this course is unrelated to the previous quantum courses, since what is more representative of quantum mechanics than the wave function in position space $\Psi(x)$, which so far has been nowhere to be seen? The reason is that the formalism of quantum mechanics is most succinctly described for Hilbert spaces of *finite dimension*, such as the (two-dimensional) electron spin. However, a particle moving on one dimension can in principle occupy any value of the position x, and thus the associated Hilbert space will be *infinite dimensional*. As we show now, the derivations and properties discussed below for the case of finite Hilbert spaces can be generalised to the case of infinitely-dimensional ones.

In the case of an infinite-dimensional Hilbert space, such as that composed by all the possible positions that a particle can occupy in one dimensions, the basis kets will be given by $|x\rangle$. In the same way as before we expressed a general quantum state as a linear superposition of its basis vectors, Eq. (1.16), now we can write this general quantum state as

$$|\Psi\rangle = \int dx\,\psi(x)|x\rangle\,, \tag{1.20}$$

where the sum has been replaced by an integral due to the infinite dimensions of this space, and the support of the integration extends to the physically allowed positions of the particle. We can thus observe how the coefficients of the basis kets $|x\rangle$ are represented by a continuous function $\psi(x)$, which is nothing but the **wave function** that you have seen in previous courses.

Furthermore, in analogy with Eq. (1.13) the inner product between two states of an infinite-dimensional Hilbert space is given by

$$\begin{aligned}\langle\Psi_1|\Psi_2\rangle &= \left(\int dx'\psi_1^*(x')\langle x'|\right)\left(\int dx\,\psi_2(x)|x\rangle\right) \\ &= \int dx'dx\,\psi_1^*(x')\psi_2(x)\langle x'|x\rangle = \int dx\,\psi_1^*(x)\psi_2(x)\,,\end{aligned} \tag{1.21}$$

where we have used the *orthonormality condition* between basis vectors

$$\langle x'|x\rangle = \delta(x-x')\,, \tag{1.22}$$

in terms of Dirac's delta function (rather than the Kronecker delta for finite Hilbert spaces). One can derive the relation Eq. (1.22) using for example the solutions of the free-particle Schroedinger equation. From here the usual normalisation condition of the wave functions follows, which now expressed in Dirac notation reads

$$|\Psi| = \left(\langle\Psi|\Psi\rangle\right)^{1/2} = \int dx\, \psi^*(x)\psi(x) = 1\,. \tag{1.23}$$

Property	Finite-dim $\mathcal{H}$	Infinite-dim $\mathcal{H}$
Completeness	$\|\Psi_1\rangle = \sum_{i=1}^n c_{1,i}\|\psi_i\rangle$ $\|\Psi_2\rangle = \sum_{i=1}^n c_{2,i}\|\psi_i\rangle$	$\|\Psi_1\rangle = \int dx\, \psi_1(x)\|x\rangle$ $\|\Psi_2\rangle = \int dx'\, \psi_2(x')\|x'\rangle$
Inner product	$\langle\Psi_1\|\Psi_2\rangle = \sum_{i=1}^n c_{1,i}^* c_{2,i}$	$\langle\Psi_1\|\Psi_2\rangle = \int dx\, \psi_1^*(x)\psi_2(x)$
Orthonormality	$\langle\psi_i\|\psi_i\rangle = \delta_{ij}$	$\langle x'\|x\rangle = \delta(x-x')$
Completeness v2	$\sum_{i=1}^n \|\psi_i\rangle\langle\psi_i\| = 1$	$\int dx\, \|x\rangle\langle x\| = 1$

Table 1: Summary of the correspondence between finite and infinite dimensional Hilbert spaces.

In this rest of the chapter, we will alternate between Hilbert spaces of finite and of infinite dimension, with the understanding that the relations summarised in Table 1 establish a correspondence between them.

1.2 Observables and operators

We now turn to present how this formalism of quantum mechanics makes possible systematizing the discussion concerning physical observables. We will see that physical observables are represented by a special class of (finite- or infinite-dimensional) **operators** that act upon the elements of the Hilbert space of the system. We will relate this discussion to how we can make predictions about the outcome of measurements in general quantum systems.

Operators in finite Hilbert spaces. In full generality, we define an operator $\widehat{\mathcal{O}}$ as a mathematical transformation between two elements of a given Hilbert vector space $\mathcal{H}$, that is

$$\widehat{\mathcal{O}}|\Psi_1\rangle = |\Psi_2\rangle\,, \quad |\Psi_1\rangle, |\Psi_2\rangle \in \mathcal{H}\,. \tag{1.24}$$

Note that this transformation is different from the *inner product*, which maps two vectors into a complex scalar. The explicit form of the operator $\widehat{\mathcal{O}}$ will be different depending on whether we deal with finite- or infinite-dimensional Hilbert spaces. In the former cases, operators can be represented as *n-dimensional matrices*. That is, if in a certain basis $\{|\psi_i\rangle\}$ we have that the coefficients of $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are

$$|\Psi_1\rangle = \sum_{i=1}^n a_i|\psi_i\rangle\,, \qquad |\Psi_2\rangle = \sum_{i=1}^n b_i|\psi_i\rangle\,, \tag{1.25}$$

then the components of the matrix $\widehat{\mathcal{O}}$ that implements the transformation in Eq. (1.24) are given by

$$\widehat{\mathcal{O}}|\Psi_1\rangle = |\Psi_2\rangle \quad \rightarrow \quad \begin{pmatrix} O_{1,1} & O_{1,2} & \dots & O_{1,n} \\ O_{2,1} & O_{2,2} & \dots & O_{2,n} \\ \vdots & \vdots & \dots & \vdots \\ O_{n,1} & O_{n,2} & \dots & O_{n,n} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} . \tag{1.26}$$

Note that in the same way as the coefficients of the linear expansion Eq. (1.25) are basis-dependent, also the matrix representation of operators will depend on the specific choice of basis (although the operator itself is basis independent).

Furthermore, you can convince yourselves that the matrix elements of a given operator are uniquely specified by the action of this operator on the basis vectors. To see this, note that by definition the action of $\widehat{\mathcal{O}}$ transforms an element of $\mathcal{H}$ into another element of the same Hilbert space, and thus with all generality one can write

$$\widehat{\mathcal{O}}|\psi_i\rangle = \sum_{j=1}^{n} c_{j,i}|\psi_j\rangle \, , \qquad i = 1, \dots, n \, , \tag{1.27}$$

where we have used that the RHS is true for any element of $\mathcal{H}$, given a suitable choice of coefficients $\{c_{j,i}\}$. We can now multiply from the left with the bra $\langle\psi_k|$ to get

$$\langle\psi_k| \left(\widehat{\mathcal{O}}|\psi_i\rangle\right) = \langle\psi_k| \left(\sum_{j=1}^{n} c_{j,i}|\psi_j\rangle\right) = \sum_{j=1}^{n} c_{j,i}\langle\psi_k|\psi_j\rangle = c_{k,i} \, , \qquad i,k = 1, \dots, n \, , \tag{1.28}$$

where again we have used orthonormality of the basis vectors. This tells us that indeed the matrix representation of this operator

$$\widehat{\mathcal{O}} = \begin{pmatrix} O_{1,1} & O_{1,2} & \dots & O_{1,n} \\ O_{2,1} & O_{2,2} & \dots & O_{2,n} \\ \vdots & \vdots & \dots & \vdots \\ O_{n,1} & O_{n,2} & \dots & O_{n,n} \end{pmatrix} \tag{1.29}$$

is fixed by its action on the basis vectors,

$$O_{i,j} = \langle\psi_i| \left(\widehat{\mathcal{O}}|\psi_j\rangle\right) \equiv \langle\psi_i|\widehat{\mathcal{O}}\psi_j\rangle \, , \qquad i,j = 1, \dots, n \, , \tag{1.30}$$

where note that we have used a generalised *bracket notation* to indicate that the operator acts on the ket to its right. Eq. (1.30) is often denoted as the **matrix element** of the operator $\widehat{\mathcal{O}}$ between the states $|\psi_i\rangle$ and $|\psi_j\rangle$. Note that in general these matrix elements are not symmetric, that is $O_{i,j} \neq O_{j,i}$.

The previous discussion of operators focuses on Hilbert spaces with finite dimension. From *Quantumfysica 1* you also know that for Hilbert spaces of infinite dimension we have operators that act on the wave function. In particular you are familiar with the position $\hat{x}$ and linear momentum $\hat{p}$ operators, whose action on a general wave function $\psi(x)$ is defined by

$$\hat{x}\psi(x) = x\psi(x) \, , \quad \hat{p}\psi(x) = -i\hbar\frac{d}{dx}\psi(x) \, . \tag{1.31}$$

Other operators can be constructed by combining these two, for example for the kinetic energy $\hat{T}$ we have

$$\hat{T} = \frac{1}{2}m\hat{v}^2 = \frac{\hat{p}^2}{2m} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\,. \tag{1.32}$$

Operators and the electron spin

Let us come back to our favorite $n = 2$ Hilbert space, the one spanned by the electron spin. Assume that we have an operator $\widehat{S}$ defined by its action on the basis vectors, given by

$$\hat{S}|+\rangle = i|-\rangle\,, \quad \hat{S}|-\rangle = -i|+\rangle\,.$$

We would like to determine the matrix representation of $\hat{S}$ and determine its action on the state $|\Psi\rangle = (1/\sqrt{2})\,(|+\rangle - i|-\rangle)$. First, we use the definition of the matrix elements Eq. (1.30) to find

$$S_{+,-} = \langle +|\left(\widehat{S}|-\rangle\right) = \langle +|\,(-i|+\rangle) = -i\,, \quad S_{-,+} = \langle -|\left(\widehat{S}|+\rangle\right) = \langle -|\,(i|-\rangle) = i\,,$$

and since $S_{+,+} = S_{-,-} = 0$, the resulting matrix representation and the action of $\hat{S}$ on $|\Psi\rangle$ are

$$\widehat{S} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\,, \quad \widehat{S}|\Psi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -i \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ i \end{pmatrix}\,.$$

Operators in infinite Hilbert spaces. In general, for a physical observable depending on the position and linear momentum, $\mathcal{O}(x,p)$, there will be an operator constructed from the appropriate combination of the position and momentum operators, $\widehat{\mathcal{O}}(\hat{x},\hat{p})$. As in the case of finite Hilbert spaces, we can compute the matrix element of a given operator $\mathcal{O}$ between two states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ with associated wave functions $\psi_1(x)$ and $\psi_2(x)$, which recalling the correspondence in Table 1 will be given by

$$\langle\Psi_1|\left(\widehat{\mathcal{O}}|\Psi_2\rangle\right) = \langle\Psi_1|\widehat{\mathcal{O}}\Psi_2\rangle = \int dx\,\psi_1^*(x)\,\widehat{\mathcal{O}}\,\psi_2(x)\,, \tag{1.33}$$

which is the analog expression of Eq. (1.30) in the case of continuous (infinite-dimensional) Hilbert spaces.

Observables in quantum mechanics. So up to here we have seen that in quantum mechanics physical observables are represented by operators, and that these operators admit a matrix representation in the case of finite Hilbert spaces. What else we can say about these operators? Is every possible operator eligible to represent a physical observable? Actually this is not the case, and there are stringent mathematical requirements that those operators must satisfy.

First of all, we know that the expectation value associated to a measurement of the physical observable $\widehat{\mathcal{O}}$ given a quantum state $|\Psi\rangle$ is given by the following matrix element:

$$\langle\mathcal{O}\rangle \equiv \langle\Psi|\widehat{\mathcal{O}}\Psi\rangle\,, \tag{1.34}$$

which for continuous Hilbert spaces is given by

$$\langle\mathcal{O}\rangle = \langle\Psi|\widehat{\mathcal{O}}\Psi\rangle = \int dx\,\psi^*(x)\,\widehat{\mathcal{O}}\,\psi(x)\,, \tag{1.35}$$

while for the discrete counterparts you can verify that

$$\langle \mathcal{O} \rangle \equiv \langle \Psi | \widehat{\mathcal{O}} \Psi \rangle = \sum_{i,j=1}^{n} c_j^* \mathcal{O}_{ji} c_i \,, \tag{1.36}$$

where $|\Psi\rangle = \sum_{i=1}^{n} c_i |\psi_i\rangle$, which is of course nothing but the usual matrix multiplication recipe once we express the state and the operator in the representations associated to the basis $\{|\psi_i\rangle\}$.

In both cases, the bottom line is that the expectation value $\langle \mathcal{O} \rangle$ is something that we can measure from the quantum system. Given that the outcome of any measurement is a real quantity, it follows that $\langle \mathcal{O} \rangle \in \mathbb{R}$ for *any operator* $\widehat{\mathcal{O}}$ and *any quantum state* $|\Psi\rangle$. This property is rather non-trivial, since both $\widehat{\mathcal{O}}$ and $|\Psi\rangle$ are in general complex quantities. Another way to state the same fact is to write that the expectation value is equal to its complex conjugate $\langle \mathcal{O} \rangle = \langle \mathcal{O} \rangle^*$. One can see how taking $\langle \mathcal{O} \rangle^*$ changes the expression for the inner product. For continuous spaces

$$\langle \mathcal{O} \rangle^* = \langle \Psi | \widehat{\mathcal{O}} \Psi \rangle^* = \left(\int dx \, \psi^*(x) \, \widehat{\mathcal{O}} \, \psi(x) \right)^* = \int dx \, \psi(x) \, \widehat{\mathcal{O}}^* \, \psi^*(x) = \int dx \, \left(\widehat{\mathcal{O}}^* \, \psi^*(x) \right) \, \psi(x) \,, \tag{1.37}$$

from which follows that $\langle \mathcal{O} \rangle^* = \langle \widehat{\mathcal{O}} \Psi | \Psi \rangle$, that is, taking the complex conjugate *reverses the order* of the inner product. The same property takes place for discrete Hilbert spaces, since

$$\langle \mathcal{O} \rangle^* = \langle \Psi | \widehat{\mathcal{O}} \Psi \rangle^* = \sum_{i,j=1}^{n} c_j \mathcal{O}_{ji}^* c_i^* = \sum_{i,j=1}^{n} c_i^* \mathcal{O}_{ji}^* c_j = \sum_{i,j=1}^{n} c_j^* \mathcal{O}_{ij}^* c_i = \langle \widehat{\mathcal{O}} \Psi | \Psi \rangle \,. \tag{1.38}$$

Therefore, we have derived a very important property of those operators in quantum mechanics that represent physical observables:

In quantum mechanics, operators $\widehat{\mathcal{O}}$ representing physical observables must satisfy

$$\langle \Psi | \widehat{\mathcal{O}} \Psi \rangle = \langle \widehat{\mathcal{O}} \Psi | \Psi \rangle \,, \quad \forall \, |\Psi\rangle \in \mathcal{H} \,, \tag{1.39}$$

which are called **Hermitian operators**. These are by definition equal to their *Hermitian conjugates*, $\widehat{\mathcal{O}}^\dagger = \widehat{\mathcal{O}}$, where the latter are defined by the relation $\langle \Psi_1 | \widehat{\mathcal{O}} \Psi_2 \rangle = \langle \widehat{\mathcal{O}}^\dagger \Psi_1 | \Psi_2 \rangle$ for all $|\Psi_1\rangle, |\Psi_2\rangle \in \mathcal{H}$.[a]

[a] Note that (see **Problem 3.3** in textbook) if $\langle \Psi | \widehat{\mathcal{O}} \Psi \rangle = \langle \widehat{\mathcal{O}} \Psi | \Psi \rangle$ then it follows that $\langle \Psi_1 | \widehat{\mathcal{O}} \Psi_2 \rangle = \langle \widehat{\mathcal{O}} \Psi_1 | \Psi_2 \rangle$.

This way, we find that now all possible operators can be associated to physical observables in quantum mechanics, but only Hermitian ones are accepted. Therefore if we are asked to assess whether or not a given operator $\widehat{\mathcal{O}}$ can be associated to a physical observable, what we need to determine is whether or not it is Hermitian. In the case of *discrete operators*, the Hermiticity condition is particularly simple to verify: imposing that Eq. (1.38) equals Eq. (1.36) we find

$$\sum_{i,j=1}^{n} c_j^* \mathcal{O}_{ji} c_i = \sum_{i,j=1}^{n} c_j^* \mathcal{O}_{ij}^* c_i \,, \quad \forall \, \{c_i\} \quad \rightarrow \quad \mathcal{O}_{ij}^* = \mathcal{O}_{ji} \,. \tag{1.40}$$

Therefore, for discrete operators taking the *Hermitian conjugates* (also known sometimes as *adjoint*) $\widehat{\mathcal{O}}^\dagger$ corresponds to *transposing the matrix* (exchange rows by columns) and then *taking the complex conjugate* of each element.

We can verify that the operators we are already familiar with are represented (as expected) with Hermitian

operators. For instance, let's demonstrate that the linear momentum operator $\hat{p}$ is indeed Hermitian.

$$\begin{aligned}\langle\Psi_1|\hat{p}\Psi_2\rangle &= \int_{-\infty}^{\infty} dx\,\psi_1^*(x)\,\hat{p}\,\psi_2(x) = \int_{-\infty}^{\infty} dx\,\psi_1^*(x)\left(-i\hbar\frac{d}{dx}\right)\psi_2(x) \\ &= -i\hbar\int_{-\infty}^{\infty} dx\,\psi_1^*(x)\frac{d\psi_2(x)}{dx} = -i\hbar\left(\psi_1^*(x)\psi_2^*(x)\right)\Big|_{-\infty}^{\infty} + i\hbar\int_{-\infty}^{\infty} dx\,\psi_2(x)\frac{d\psi_1^*(x)}{dx}\end{aligned} \tag{1.41}$$

where in the last equality we have used integration by parts.[1] Since the wave functions vanish at infinity (else they would not be normalisable and hence non-physical) the boundary term goes to zero, and we end up with

$$\langle\Psi_1|\hat{p}\Psi_2\rangle = \int_{-\infty}^{\infty} dx\left(-i\hbar\frac{d\psi_1(x)}{dx}\right)^*\psi_2(x) = \langle\hat{p}\Psi_1|\Psi_2\rangle\,, \tag{1.42}$$

which is the condition that an Hermitian operator must satisfy, as indicated in Eq. (1.39), and thus we confirm (reassuringly) that $\hat{p}$ can indeed be associated to a physical observable. Similar derivations can be applied to show that other commonly used operators, such as the ones representing the kinetic $\widehat{T}$ and the total energy $\widehat{H}$, are also Hermitian.

Consider a 3-dimensional Hilbert space. We are given the following two operators belonging to this Hilbert space. Can these operators represent *bona fide* physical observables?

$$\widehat{\mathcal{O}}_1 = \begin{pmatrix} 0 & -i & 1 \\ i & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}\,, \qquad \widehat{\mathcal{O}}_2 = \begin{pmatrix} 2 & 1 & 0 \\ i & 3 & i \\ 0 & -i & 0 \end{pmatrix}\,. \tag{1.43}$$

As indicated by Eq. (1.40), for finite Hilbert spaces the Hermiticity condition $\widehat{\mathcal{O}}^\dagger = \widehat{\mathcal{O}}$ translates into $\mathcal{O}_{ij}^* = \mathcal{O}_{ji}$ for the matrix elements, that is, if I *transpose* the matrix and then take the *complex conjugates* of its elements I should recover the original matrix. Verifying this condition:

$$\widehat{\mathcal{O}}_1^\dagger = \left(\widehat{\mathcal{O}}_1^T\right)^* = \begin{pmatrix} 0 & i & 1 \\ -i & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}^* = \begin{pmatrix} 0 & -i & 1 \\ i & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} = \widehat{\mathcal{O}}_1\,, \tag{1.44}$$

and thus $\widehat{\mathcal{O}}_1$ is Hermitian and could represent a physical observable. By using the same method you can show that instead $\widehat{\mathcal{O}}_2$ is not Hermitian and thus cannot be associated to a physical observable.

In the following the study the consequences of the fact that observables in quantum mechanics are represented by Hermitian operators, and what does this property tell us about the allowed states of a quantum system.

1.3 Eigenfunctions and eigenvalues of Hermitian operators

Many important properties of quantum systems can be derived from the fact that they arise from eigenvalue equations involving Hermitian operators. For this reason, here first we briefly review eigenvalue equations in linear algebra, and then we discuss how to extend the concept to quantum mechanics.

[1] Where for any two functions $f(x), g(x)$ one has that $\int_a^b dx f(x) d(g(x)) = (f(x)g(x))\Big|_a^b - \int_a^b dx g(x) d(f(x))$.

Eigenvalue equations in linear algebra. You are familiar with matrix equations of the form $A\vec{v} = \lambda\vec{v}$,

$$\begin{pmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,n} \\ A_{2,1} & A_{2,2} & \dots & A_{2,n} \\ \vdots & \vdots & \dots & \vdots \\ A_{n,1} & A_{n,2} & \dots & A_{n,n} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} , \tag{1.45}$$

where A represents a square matrix of dimensions $n \times n$ and $\vec{v}$ is a column vector with dimension n. These equations, where the outcome of acting with a matrix on a vector is proportional to the original vector, are known as *eigenvalue equations*. The scalar λ is known as the *eigenvalue* of the equation, while the vector $\vec{v}$ is known as the associated *eigenvector*. The key feature of Eq. (1.45) is that applying a matrix A to the vector $\vec{v}$ returns the original vector with an overall rescaling, $\lambda\vec{v}$. In order to compute the eigenvalues of a matrix, we need to evaluate the solutions of the *characteristic equation* of the matrix A, defined as

$$\det\left(A - \lambda \cdot I\right) = 0\,, \tag{1.46}$$

where I is the identity matrix of dimensions $n \times n$, and det is the determinant. Once we have evaluated the eigenvalues λ_i associated to a given matrix A, we can compute the corresponding eigenvectors $\vec{v}_i$, defined as the vectors that satisfy

$$A\mathbf{v}_i = \lambda_i \mathbf{v}_i\,, \quad i = 1, \dots, n\,, \tag{1.47}$$

which is an n–dimensional system of linear equations. Note that the *normalisation of the eigenvectors is arbitrary*: if $\mathbf{v}_i$ is an eigenvector of a given matrix A, then $b\mathbf{v}_i$, with b being an arbitrary complex number, will also be an eigenvector of the same matrix equation.

To illustrate how to compute eigenvalues and eigenvectors, one can consider the following matrix

$$A = \begin{pmatrix} 1 & -3 \\ 3 & -5 \end{pmatrix} , \tag{1.48}$$

which has associated the following characteristic equation

$$\det\left(A - \lambda \cdot I\right) = \begin{vmatrix} 1-\lambda & -3 \\ 3 & -5-\lambda \end{vmatrix} = \lambda^2 + 4\lambda + 4 = 0\,. \tag{1.49}$$

with (degenerate) solutions $\lambda_1 = \lambda_2 = -2$. We can determine the single eigenvector by solving

$$\begin{pmatrix} 1 & -3 \\ 3 & -5 \end{pmatrix} \begin{pmatrix} v_{1,1} \\ v_{1,2} \end{pmatrix} = \lambda_1 \begin{pmatrix} v_{1,1} \\ v_{1,2} \end{pmatrix} , \quad \begin{aligned} v_{1,1} - 3v_{1,2} &= -2v_{1,1} \\ 3v_{1,1} - 5v_{1,2} &= -2v_{1,1} \end{aligned} \tag{1.50}$$

which can be solved to get $\vec{v}_1 = (1, 1)$. Note that $b\vec{v}_1$, with $b \in \mathbb{C}$, is also a possible eigenvector.

Eigenvalue equations in quantum mechanics. In Eq. (1.24) we defined an operator $\widehat{\mathcal{O}}$ as a mathematical transformation between two elements of a given Hilbert vector space $\mathcal{H}$, that is

$$\widehat{\mathcal{O}}|\Psi_1\rangle = |\Psi_2\rangle\,, \quad |\Psi_1\rangle, |\Psi_2\rangle \in \mathcal{H}\,, \tag{1.51}$$

where $\mathcal{H}$ can be either finite- or infinite-dimensional. An eigenvalue equation in quantum mechanics, by analogy with the linear algebra case, is an operator equation where $|\Psi_2\rangle$ is proportional to $|\Psi_1\rangle$,

$$\widehat{\mathcal{O}}|\Psi_1\rangle = \lambda_1|\Psi_1\rangle\,, \quad |\Psi_1\rangle \in \mathcal{H}\,, \tag{1.52}$$

where now we say that λ_1 is an eigenvalue of the operator $\widehat{\mathcal{O}}$ with $|\Psi_1\rangle$ the associated eigenvector. Eigenvalue equations of the form of Eq. (1.52) are extremely important in quantum mechanics: for example, you can see how the ubiquitous Schroedinger equation is nothing but the eigenvalue equation associated to the Hamiltonian operator, with the particle energy being the corresponding eigenvector

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\psi(x) = E\psi(x)\,, \quad \rightarrow \quad \mathcal{H}|\Psi\rangle = E|\Psi\rangle\,. \tag{1.53}$$

Determinate states

In quantum theory, we denote a **determinate** state $|\Psi\rangle$ as one where a measurement of the physical observable $\mathcal{O}$ returns always the same value, say λ. One can show that these states are eigenvectors of the eigenvalue equation associated with $\mathcal{O}$, namely $\widehat{\mathcal{O}}|\Psi\rangle = \lambda|\Psi\rangle$. Therefore, we conclude that determinate states of the observable $\mathcal{O}$ are the eigenvectors of $\widehat{\mathcal{O}}$. This property implies for example, considering Eq. (1.53) that if a given quantum state is an eigenvector of $\mathcal{H}$, a measurement of its energy will always return the same value.

The eigenvector spectrum of an Hermitian operator can be either *discrete* or *continuous*, depending on whether the associated Hilbert space is finite- or infinite-dimensional. Let us demonstrate now three important properties associated to eigenvalue equations involving Hermitian operators.

(a) Hermitian operators have associated **real eigenvalues**. This can be shown starting from the relevant eigenvalue equation for $\widehat{\mathcal{O}}$

$$\widehat{\mathcal{O}}|\Psi_1\rangle = \lambda_1|\Psi_1\rangle\,, \tag{1.54}$$

and then imposing the Hermiticity property $\langle\Psi_1|\widehat{\mathcal{O}}\Psi_1\rangle = \langle\widehat{\mathcal{O}}\Psi_1|\Psi_1\rangle$ we see that

$$\langle\Psi_1|\widehat{\mathcal{O}}\Psi_1\rangle = \lambda_1\langle\Psi_1|\Psi_1\rangle\,, \quad \langle\widehat{\mathcal{O}}\Psi_1|\Psi_1\rangle = \lambda_1^*\langle\Psi_1|\Psi_1\rangle\,, \tag{1.55}$$

hence $\lambda_1 = \lambda_1^*$ which implies that $\lambda_1 \in \mathbb{R}$. This property confirms that if we measure the observable $\mathcal{O}$ on the determinate state $|\Psi_1\rangle$ we will obtain a real number, which is reassuring.

(b) The eigenvectors $|\Psi_1\rangle$ and $|\Psi_2\rangle$ associated to different eigenvalues, $\lambda_2 \neq \lambda_1$, are **orthogonal**. To show this again we start from the corresponding eigenvalue equations

$$\widehat{\mathcal{O}}|\Psi_1\rangle = \lambda_1|\Psi_1\rangle\,, \quad \widehat{\mathcal{O}}|\Psi_2\rangle = \lambda_2|\Psi_2\rangle\,, \tag{1.56}$$

and now we impose Hermiticity $\langle\Psi_2|\widehat{\mathcal{O}}\Psi_1\rangle = \langle\widehat{\mathcal{O}}\Psi_2|\Psi_1\rangle$ which implies that $\lambda_2\langle\Psi_2|\Psi_1\rangle = \lambda_1\langle\Psi_2|\Psi_1\rangle$, which given that $\lambda_2 \neq \lambda_1$ can only be satisfied if $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are orthogonal: $\langle\Psi_2|\Psi_1\rangle = 0$. Even in the presence of *degenerate spectra* with different eigenvectors sharing the same eigenvalue, it is possible to construct orthogonal eigenvectors within each degenerate subspace.

(c) The eigenvectors of an Hermitian operator **span the complete Hilbert space**, in other words (given

that they are already orthonormal) they represent **a complete basis** in this Hilbert space.[2] This means that if a given Hermitian operator $\widehat{\mathcal{O}}$ living in an n-dimensional Hilbert space has as eigenvalue spectrum

$$\widehat{\mathcal{O}}|\Psi_i\rangle = \lambda_i|\Psi_i\rangle\,, \quad i = 1,\ldots,n\,, \tag{1.57}$$

then I can always use these eigenvectors to represent any element of $\mathcal{H}$:

$$|\Psi\rangle = \sum_{i=1}^{n} c_i|\Psi_i\rangle\,, \quad c_i = \langle\Psi_i|\Psi\rangle \in \mathbb{C}\,, \quad \forall\,|\Psi\rangle \in \mathcal{H}\,. \tag{1.58}$$

Note however that $|\Psi\rangle$ itself is *not* an eigenvalue of $\widehat{\mathcal{O}}$, since

$$\widehat{\mathcal{O}}|\Psi\rangle = \sum_{i=1}^{n} c_i\widehat{\mathcal{O}}|\Psi_i\rangle = \sum_{i=1}^{n} c_i\lambda_i|\Psi_i\rangle \neq \lambda|\Psi_i\rangle\,. \tag{1.59}$$

These three properties of Hermitian operators are immensely important, and will be thoroughly exploited in the rest of this course.

1.4 Generalised statistical interpretation

In previous courses, you have learned some important properties about the statistical interpretation of the wave function $\psi(x)$ of a quantum system and what are the expected outcomes if we attempt specific measurements in this system. In particular, you saw that $P(x)dx = |\psi(x)|^2dx$ is the probability of finding a particle, upon a measurement of its position, in the region of space defined by $[x, x+dx]$. Now, armed with the powerful mathematical tools that constitute the formalism of quantum mechanics, we can state another of the core axioms that define the foundations of quantum mechanics:

Generalised statistical interpretation

Consider a general physical observable $\mathcal{O}$ represented by the Hermitian operator $\widehat{\mathcal{O}}$. If now we attempt to measure $\mathcal{O}$ on the state $|\Psi\rangle$, the **outcome of this measurement** will be **one of the eigenvalues** λ_i associated to the eigenvector equation $\widehat{\mathcal{O}}|\Psi_i\rangle = \lambda_i|\Psi_i\rangle$. For discrete spectra, recall Eq. (1.58), the *probability* of measuring λ_i in the state $|\Psi\rangle$ will be given by

$$P(\lambda_i) = |c_i|^2 \quad \text{where} \quad c_i = \langle\Psi_i|\Psi\rangle\,, \tag{1.60}$$

Furthermore, upon this measurement the quantum state **collapses** into the relevant eigenvector. For instance, always in the discrete case, if I measure observable $\mathcal{O}$ on the state $|\Psi\rangle$ and find λ_i, after the measurement the quantum state will collapse from $|\Psi\rangle$ to $|\Psi_i\rangle$:

$$\text{Before the measurement}: \qquad |\Psi\rangle = \sum_{i=1}^{n} c_i|\Psi_i\rangle\,, \quad c_i = \langle\Psi_i|\Psi\rangle\,,$$

$$\text{After the measurement (having measured } \mathcal{O} = \lambda_\text{i}): \qquad |\Psi\rangle = |\Psi_i\rangle\,.$$

[2] To be precise, the proof that the eigenvalues of an Hermitian operator span the full Hilbert space applies only for finite-dimensional spaces. In the continuous case, it is usually taken to be an **axiom** that only Hermitian operators whose eigenfunctions are *complete* can represent physical observables.

If you find this *collapse of the wave function* bewildering, you are not alone: it is still the subject of endless discussions about the interpretations of quantum mechanics. However, For the purposes of this course, we will put epistemology aside and take this collapse as part of the axioms of the theory.

This generalised statistical interpretation can also be applied to *continuous spectra*. In such case, the use of the Dirac notation is sometimes cumbersome, and we will interchange frequently the abstract quantum state $|\Psi(z)\rangle$, with z labeling the relevant continuous quantum number (such as the particle energy or momentum) with its wave function in the position-space representation, $\psi_z(x)$. Then, if we have an operator $\widehat{\mathcal{O}}$ with a continuous spectrum in the observable z,

$$\widehat{\mathcal{O}}\,\psi_z(x) = q(z)\psi_z(x)\,, \tag{1.61}$$

with $q(z)$ the real eigenvalues and $\psi_z(x)$ the orthonormal eigenfunctions (in the Dirac sense), the generalised statistical interpretation tells us that, given a general quantum state $\psi(x)$, the *probability of finding a result for* $\mathcal{O}$ *in the range* $[q(z), q(z+dz)]$ is

$$P(z)dz = |c(z)|^2 dz \quad \text{where} \quad c(z) = \langle \psi_z|\psi\rangle = \int_{-\infty}^{\infty} dx\; \psi_z^*(x)\psi(x)\,, \tag{1.62}$$

fully analogous to its counterpart for discrete spectra, Eq. (1.60).

For discrete (continuous) spectra, the value of the coefficient c_i ($c(z)$) quantifies the **overlap** between the full quantum state $|\Psi\rangle$ ($\psi(x)$) and the specific eigenvector $|\Psi_i\rangle$ ($\psi_z(x)$) associated to the eigenvalue λ_i ($q(z)$). Recall that measurements of the observable $\mathcal{O}$ always yield one of the eigenvalues λ_i ($q(z)$).

You can easily recover the usual statistical interpretation of the wave function that you learned in *Quantumfysica 1* by applying this discussion to the case of the position operator $\hat{x}$. In this case, one has that the corresponding eigenvalue equation is trivial

$$\hat{x}\,\psi_{x'}(x) = x'\psi_{x'}(x)\,, \tag{1.63}$$

where we have separated the value of the position x from the eigenfunction label x', which in principle are different quantities. The eigenvectors are Dirac delta functions of the form $\psi_{x'}(x) = \delta(x-x')$, and hence the probability of finding the particle in the interval $[x', x'+dx']$ is nothing but

$$P(x')dx' = |c(x')|^2 dx' = \left|\int_{-\infty}^{\infty} dx\; \psi_{x'}^*(x)\psi(x)\right|^2 dx = |\psi(x')|^2\, dx'\,, \tag{1.64}$$

namely the statistical representation of the square of position-space wave function as a probability density.

An straightforward consequence of this generalised statistical interpretation is that the total probability (understood as the sum over the probabilities associated to all possible outcomes of the measurement of the observable $\mathcal{O}$) must be unity:

$$\sum_{i=1}^{n} |c_i|^2 = 1 \quad \text{(discrete)}\,, \qquad \int_z dz\; |c(z)|^2 = 1 \quad \text{(continuous)}\,, \tag{1.65}$$

where in the continuous case the support of the integral over z is the physically allowed range for the observable $\mathcal{O}$. Again, the second relation is nothing but the requirement that the wave function is normalised,

if we take $\mathcal{O} = x$,

$$\int_z dz\ |c(z)|^2 = \int_\infty^\infty dz\, |\psi(z)|^2 = 1\,. \tag{1.66}$$

Another way to show that the sum over $|c_i|^2$ gives unity is exploit that we deal with normalised wave functions,

$$1 = \langle\Psi|\Psi\rangle = \left(\sum_{j=1}^n c_j^* \langle\Psi_j|\right)\left(\sum_{i=1}^n c_i|\Psi_i\rangle\right) = \sum_{i,j=1}^n c_j^* c_i \langle\Psi_j|\Psi_i\rangle = \sum_{i,j=1}^n |c_i|^2\,, \tag{1.67}$$

where in the last step I have used the orthonormality of the eigenvectors. Likewise, you can easily demonstrate that the expectation value of $\mathcal{O}$ in the quantum state $|\Psi\rangle$ is given by

$$\langle\mathcal{O}\rangle = \langle\Psi|\widehat{O}\Psi\rangle = \sum_{i=1}^n \lambda_i |c_i|^2 = \sum_{i=1}^n \lambda_i\, P(\lambda_i)\,, \tag{1.68}$$

as couldn't be otherwise: the expected value of $\mathcal{O}$ is the sum over all possible outcomes of its measurement λ_i each weighted by the corresponding probability $P(\lambda_i) = |c_i|^2$.

1.5 The generalised uncertainty principle

Again, you are already familiar with one of the key concepts of quantum mechanics, *Heisenberg's* uncertainty principle. This principle told us that if we measure the position x and momentum p with an error of σ_x and σ_p respectively, then the following relation should hold:

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}\,, \tag{1.69}$$

which tells us that there exists a *fundamental limit* on how accurately I can determine the position and the momentum of a particle. So far, the Heisenberg uncertainty principle has been given to you as if it were another of the theory axioms, but instead we will show there how this principle (rather, its generalised version) follows directly from the formalism of quantum mechanics assembled so far.

Assume that we have a physical observable, $\mathcal{O}$, represented by its corresponding Hermitian operator $\widehat{\mathcal{O}}$. For a general quantum state $|\Psi\rangle$, measurements of $\mathcal{O}$ will result in a range of different outcomes. If this range is very small, I will know quite accurately what to expect from measurements of $\mathcal{O}$ in $|\Psi\rangle$, and hence I'll say I can determine these observables with a *small uncertainty*. Conversely, if the range of outcomes is very broad, I will have limited knowledge about what to expect from measurements of $\mathcal{O}$, and in this case my uncertainties will be much larger. One possible measure to quantify how narrow or broad is the range of outcomes for measurements of an observable $\mathcal{O}$ (and thus of how small or large my uncertainty will be) is given by the **standard deviation**, which you might have seen in statistics courses and that is defined as

$$\sigma_{\mathcal{O}}^2 \equiv \left\langle \left(\mathcal{O} - \langle\mathcal{O}\rangle\right)^2 \right\rangle = \langle\Psi|\left(\widehat{\mathcal{O}} - \langle\mathcal{O}\rangle\right)^2\Psi\rangle = \langle\left(\widehat{\mathcal{O}} - \langle\mathcal{O}\rangle\right)\Psi|\left(\widehat{\mathcal{O}} - \langle\mathcal{O}\rangle\right)\Psi\rangle\,, \tag{1.70}$$

where in the last equality we have used the fact that the operator is Hermitian ($\widehat{\mathcal{O}}^\dagger = \widehat{\mathcal{O}}$).

Now let's say we have two different physical observables, denoted by A and B, represented by the Hermitian operators $\hat{A}$ and $\hat{B}$. They standard deviations can be computed in the same way as above

$$\sigma_A^2 \equiv \left\langle \left(A - \langle A\rangle\right)^2 \right\rangle = \langle\Psi|\left(\hat{A} - \langle A\rangle\right)^2\Psi\rangle = \langle\left(\hat{A} - \langle A\rangle\right)\Psi|\left(\hat{A} - \langle A\rangle\right)\Psi\rangle\,, \tag{1.71}$$

$$\sigma_B^2 \equiv \left\langle (B - \langle B \rangle)^2 \right\rangle = \langle \Psi | \left(\hat{B} - \langle B \rangle \right)^2 \Psi \rangle = \langle \left(\hat{B} - \langle B \rangle \right) \Psi | \left(\hat{B} - \langle B \rangle \right) \Psi \rangle \, , \tag{1.72}$$

Now we can use the Schwartz inequality[3] to express the product of the two standard deviations as follows:

$$\begin{aligned} \sigma_A^2 \sigma_B^2 &= \langle (\hat{A} - \langle A \rangle)\Psi | (\hat{A} - \langle A \rangle)\Psi \rangle \langle (\hat{B} - \langle B \rangle)\Psi | (\hat{B} - \langle B \rangle)\Psi \rangle \\ &\geq \left| \langle (\hat{A} - \langle A \rangle)\Psi | (\hat{B} - \langle B \rangle)\Psi \rangle \right|^2 . \\ &\geq \left[\frac{1}{2i} \left(\langle (\hat{A} - \langle A \rangle)\Psi | (\hat{B} - \langle B \rangle)\Psi \rangle - \langle (\hat{B} - \langle B \rangle)\Psi | (\hat{A} - \langle A \rangle)\Psi \rangle \right) \right]^2 , \end{aligned} \tag{1.73}$$

where in the last step I have used that $|z^2| \geq |(z - z^*)/(2i)|^2$ for any complex number z. Now making use of the fact that both $\hat{A}$ and $\hat{B}$ are Hermitian operators, some algebra tells us that

$$\begin{aligned} &\langle (\hat{A} - \langle A \rangle)\Psi | (\hat{B} - \langle B \rangle)\Psi \rangle = \langle \Psi | (\hat{A} - \langle A \rangle)(\hat{B} - \langle B \rangle)\Psi \rangle \\ &= \langle \Psi | (\hat{A}\hat{B} - \langle A \rangle \hat{B} - \hat{A} \langle B \rangle + \langle A \rangle \langle B \rangle)\Psi \rangle = \left\langle \hat{A}\hat{B} \right\rangle - \langle A \rangle \langle B \rangle \, . \end{aligned}$$

Recall that $\hat{A}$ and $\hat{B}$ are operators, and hence we should be careful with their order. On the other hand $\langle A \rangle$ and $\langle B \rangle$ are just numbers so they obey the usual commutative properties.

Putting everything together, we end up with the following relation:

$$\begin{aligned} \sigma_A^2 \sigma_B^2 &\geq \left[\frac{1}{2i} \left(\langle (\hat{A} - \langle A \rangle)\Psi | (\hat{B} - \langle B \rangle)\Psi \rangle - \langle (\hat{B} - \langle B \rangle)\Psi | (\hat{A} - \langle A \rangle)\Psi \rangle \right) \right]^2 \\ &= \left[\frac{1}{2i} \left(\left\langle \hat{A}\hat{B} \right\rangle - \langle A \rangle \langle B \rangle - \left\langle \hat{B}\hat{A} \right\rangle - \langle B \rangle \langle A \rangle \right) \right]^2 \\ &= \left[\frac{1}{2i} \left(\left\langle \hat{A}\hat{B} \right\rangle - \left\langle \hat{B}\hat{A} \right\rangle \right) \right]^2 \equiv \left(\frac{1}{2i} \left\langle \left[\hat{A}, \hat{B} \right] \right\rangle \right)^2 \end{aligned} \tag{1.74}$$

where we have defined the **commutator** between two operators as

$$\left[\hat{A}, \hat{B} \right] = \hat{A}\hat{B} - \hat{B}\hat{A} \, , \tag{1.75}$$

and hence we have nicely derived the **generalised uncertainty principle**:

$$\sigma_A \sigma_B \geq \left| \frac{1}{2i} \left\langle \left[\hat{A}, \hat{B} \right] \right\rangle \right| . \tag{1.76}$$

Note that the standard deviation of a physical observable is by construction a *positive-definite* quantity, and hence we take the absolute value. While this derivation might appear to be rather abstract, it is powerful and we will show next how the well-known Heisenberg's uncertainty relations can be derived from it.

It is worth emphasizing here that this generalised uncertainty principle is **not** one of the axioms of quantum theory, but rather a consequence of those, specifically of the axiom that posits that all physical observables in quantum theory are represented by Hermitian operators.

As usual in science, when we derive a general result the first thing that one needs to do is to verify that one is able to recover specific, known cases. So let's see if we can reproduce the vanilla Heisenberg

[3] In a nutshell, for elements $\vec{v}$ and $\vec{u}$ of a vector space equipped with an inner product, the Cauchy–Schwartz relation tell us that $|\vec{u} \cdot \vec{v}|^2 \leq (\vec{u} \cdot \vec{u}) \times (\vec{v} \cdot \vec{v})$ which is the same as stating that $|\vec{u} \cdot \vec{v}| \leq |\vec{u}||\vec{v}|$.

uncertainty principle. We start from Eq. (1.88) and now we apply it to the position x and linear momentum p observables. We need to evaluate their commutator

$$\begin{aligned}[\hat{x},\hat{p}]\,\psi(x) &= \left[x,-i\hbar\frac{d}{dx}\right]\psi(x) = -i\hbar\left(x\frac{d}{dx}\psi(x)-\frac{d}{dx}\left(x\psi(x)\right)\right)\\ &= -i\hbar\left(x\frac{d}{dx}\psi(x)-\psi(x)-x\frac{d}{dx}\psi(x)\right) = i\hbar\psi(x)\end{aligned}$$

and hence $[\hat{x},\hat{p}] = i\hbar$, which leads to $\sigma_x\sigma_p \geq \hbar/2$ as expected. We thus see that Heisenberg's principle is just a specific case of a much more general property of quantum mechanics: a similar relation will hold for any pair of observables A and B whose operators *do not commute*, $\left[\hat{A},\hat{B}\right] \neq 0$. In previous courses, you have also seen that the Heisenberg uncertainty principle also applies to energy and time, and hence one has a relation of the form $\Delta t\,\Delta E \geq \hbar/2$. However you might immediately object that I cannot derive this relation using Eq. (1.88) since *time is not a physical observable* in non-relativistic quantum mechanics, and hence there is no "time operator". Below we will show how one can formally derive the $\Delta t\,\Delta E \geq \hbar/2$ relation.

Compatible and incompatible observables

In quantum mechanics, we denote a pair of physical observables A and B for which the associated commutator vanishes, $\left[\hat{A},\hat{B}\right] = 0$, as **compatible observables**. In this case, I can measure simultaneously A and B with arbitrarily good precision (the uncertainty principle does not apply) and further $\hat{A}$ and $\hat{B}$ admit a *complete set of common eigenfunctions*. On the contrary, we denote as **incompatible observables** those for which $\left[\hat{A},\hat{B}\right] \neq 0$. In such case, A and B cannot be measured simultaneously since they have associated a restriction of the form Eq. (1.88) and they cannot have *a complete set of shared eigenfunctions*.

Note that while you have been mostly exposed to the Heisenberg uncertainty principle in the case of infinite-dimensional Hilbert spaces, Eq. (1.88) also applies to finite spaces. For example, in the case of the electron spin, we can measure the spin along the x, y, or z directions. We denote the physical observables in this case as S_x, S_y, and S_z, and as we will see these are represented by the following operators:

$$\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}\,,\quad \hat{S}_y = \frac{\hbar}{2}\begin{pmatrix}0&-i\\i&0\end{pmatrix}\,,\quad \hat{S}_z = \frac{\hbar}{2}\begin{pmatrix}1&0\\0&-1\end{pmatrix}\,, \tag{1.77}$$

which you can easily check are Hermitian. If you evaluate their commutators, you can verify that these operators do not commute and hence the corresponding physical observables *are incompatible* and will have associated a Heisenberg-like uncertainty relation. For example, we have that

$$\begin{aligned}\left[\hat{S}_x,\hat{S}_z\right] &= \frac{\hbar^2}{4}\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}1&0\\0&-1\end{pmatrix} - \frac{\hbar^2}{4}\begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}0&1\\1&0\end{pmatrix}\\ &= \frac{\hbar^2}{4}\begin{pmatrix}0&-1\\1&0\end{pmatrix} - \frac{\hbar^2}{4}\begin{pmatrix}0&1\\-1&0\end{pmatrix} = \frac{\hbar^2}{4}\begin{pmatrix}0&-2\\2&0\end{pmatrix} = -i\hbar\,\hat{S}_y \neq 0\,.\end{aligned}$$

Note that in general the commutator between two operators will be *another operator*, as shown in this specific case. Therefore we confirm that S_x and S_z are incompatible operators, and that their associated

Heisenberg-like uncertainty principle will be given by

$$\sigma_{S_x}\sigma_{S_z} \geq \left| \frac{1}{2i} \left\langle -i\hbar \hat{S}_y \right\rangle \right| = \frac{\hbar}{2} \left| \left\langle \hat{S}_y \right\rangle \right| . \tag{1.78}$$

In this case the lower bound on the product $\sigma_{S_x}\sigma_{S_z}$ depends on the specific quantum state $|\Psi\rangle$ under consideration. Interestingly, we note that for a quantum state such that $\left\langle \hat{S}_y \right\rangle = 0$, then S_x and S_z become *compatible observables* and can be simultaneously measured!

The energy-time uncertainty principle. We have mentioned below that the energy-time version of Heisenberg's uncertainty principle, $\Delta t\, \Delta E \geq \hbar/2$, cannot really be derived from Eq. (1.88) since time is not an observable in quantum mechanics: it is rather an **independent variable** (that is, an external parameter to the system) and hence it does not admit an operator representation. We will need to use a different approach to derive this relation using the quantum theory formalism spelled out in this chapter.

Assume that we have a time-dependent quantum system in one dimension, whose wave function is then the solution of the time-dependent Schroedinger equation

$$i\hbar \frac{\partial \Psi(x,t)}{\partial t} = \hat{H}\Psi(x,t) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \right) \Psi(x,t) , \tag{1.79}$$

and assume that we have a general observable in the system which is constructed out of the position and momentum and also depends on time, $\mathcal{O}(x,p,t)$. Let us evaluate the first derivative with time of the expectation value of this observable:

$$\frac{d}{dt}\left\langle \mathcal{O}(x,p,t) \right\rangle = \frac{d}{dt}\left\langle \Psi \middle| \widehat{\mathcal{O}}\Psi \right\rangle = \left\langle \frac{\partial \Psi}{\partial t} \middle| \widehat{\mathcal{O}}\Psi \right\rangle + \left\langle \Psi \middle| \widehat{\mathcal{O}}\frac{\partial \Psi}{\partial t} \right\rangle + \left\langle \Psi \middle| \frac{\partial \widehat{\mathcal{O}}}{\partial t}\Psi \right\rangle , \tag{1.80}$$

where we have used the chain rule to make sure we take into account all the possible time dependencies (in particular that of the observable $\mathcal{O}$). The time derivative of the wave function is given by the RHS of the time-dependent Schroedinger equation, Eq. (1.79), and hence we can write the above expression as

$$\frac{d}{dt}\left\langle \mathcal{O} \right\rangle = -\frac{1}{i\hbar}\left\langle \hat{H}\Psi \middle| \widehat{\mathcal{O}}\Psi \right\rangle + \frac{1}{i\hbar}\left\langle \Psi \middle| \widehat{\mathcal{O}}\hat{H}\Psi \right\rangle + \left\langle \frac{\partial \widehat{\mathcal{O}}}{\partial t} \right\rangle = \frac{i}{\hbar}\left\langle \left[\hat{H}, \widehat{\mathcal{O}}\right] \right\rangle + \left\langle \frac{\partial \widehat{\mathcal{O}}}{\partial t} \right\rangle , \tag{1.81}$$

where we have used the Hermiticity of $\hat{H}$. This result is often called the **generalised Ehrenfest theorem**, and determines the time variation of a given physical observable in terms of its commutator with the Hamiltonian. There are two interesting limiting cases of this theorem:

- For observables that *do not depend explicitely on time*, $\mathcal{O}(x,p)$, we have that

$$\frac{d}{dt}\left\langle \mathcal{O} \right\rangle = \frac{i}{\hbar}\left\langle \left[\hat{H}, \widehat{\mathcal{O}}\right] \right\rangle , \tag{1.82}$$

  which tells us that for these observables their time dependence is entirely determined by their commutator with the Hamiltonian. A particularly interesting example, that you might have encountered already, arises when $\mathcal{O}$ is either x or p, since then

$$\frac{d}{dt}\left\langle x \right\rangle = \frac{i}{\hbar}\left\langle \left[\hat{H}, \hat{x}\right] \right\rangle = \left\langle \frac{\hat{p}}{m} \right\rangle \quad \rightarrow \quad m\frac{d}{dt}\left\langle x \right\rangle = \left\langle p \right\rangle , \tag{1.83}$$

$$\frac{d}{dt}\langle p\rangle = \frac{i}{\hbar}\left\langle\left[\hat{H},\hat{p}\right]\right\rangle = -\left\langle\frac{dV(x)}{dx}\right\rangle\,, \tag{1.84}$$

which when combined yield the so-called **Ehrenfest theorem**,

$$m\frac{d^2}{dt^2}\langle x\rangle = -\left\langle\frac{dV(x)}{dx}\right\rangle\,, \tag{1.85}$$

which is the quantum-mechanical version of Newton's equations of motion,

$$m\frac{d^2}{dt^2}x = -\frac{dV(x)}{dx} = F\,, \tag{1.86}$$

for conservative potentials.

- Furthermore, for observables that in addition commute with the Hamiltonian, $\left[\hat{H},\widehat{\mathcal{O}}\right] = 0$ we have

$$\frac{d}{dt}\langle\mathcal{O}\rangle = 0\,, \tag{1.87}$$

indicating that the expectation value of $\mathcal{O}$ is *time-independent* (constant).

Why it was necessary for our purposes to derive Eq. (1.81)? Assume now that we apply the generalised uncertainty principle to a time-independent observable $\mathcal{O}$ and to the Hamiltonian H,

$$\sigma_H\sigma_{\mathcal{O}} \geq \left|\frac{1}{2i}\left\langle\left[\hat{H},\widehat{\mathcal{O}}\right]\right\rangle\right| = \frac{\hbar}{2}\left|\frac{d\langle\mathcal{O}\rangle}{dt}\right|\,, \tag{1.88}$$

which relates the uncertainty in H and in $\mathcal{O}$ to the time-derivative of the expected value of the latter. If we now define $\Delta E \equiv \sigma_H$ and $\Delta t \equiv \sigma_{\mathcal{O}}/|d\langle\mathcal{O}\rangle/dt|$, we reproduce the sought-for time-energy version of Heisenberg's uncertainty principle:

$$\Delta E\,\Delta t \geq \frac{\hbar}{2}\,. \tag{1.89}$$

However, thanks to this derivation we can be more precise about the intepretation of Δt. By writing

$$\sigma_{\mathcal{O}} = \left|\frac{d\langle\mathcal{O}\rangle}{dt}\right|\Delta t\ \,, \tag{1.90}$$

we see that in the interval of time Δt the expectation value $\langle\mathcal{O}\rangle$ will have changed by one standard deviation $\sigma_{\mathcal{O}}$. So Δt is not an absolute interval, but depends rather on the observable under consideration: it is a measure of how fast or slow the expectation value of an observable varies. For observables that vary with time very rapidly, Δt will be small and hence the uncertainty in the energy ΔE will be large.

Summary

We can now recapitulate what have we learned in this chapter concerning the basic formalism of quantum mechanics:

I/ The state of a quantum system is characterised by its state vector $|\Psi\rangle$, which is an element of an abstract complex vector space, known as the Hilbert space $\mathcal{H}$, equipped with an inner product. All the information about this quantum state is fully encoded in this state vector $|\Psi\rangle$.

II/ These state vectors are modified by a specific type of linear transformations known as Hermitian operators such that $\widehat{\mathcal{O}}|\Psi_1\rangle = |\Psi_2\rangle$ where $|\Psi_1\rangle, |\Psi_2\rangle \in \mathcal{H}$.

III/ Every physical observable in quantum theory is represented by an Hermitian operator $\widehat{\mathcal{O}} = \widehat{\mathcal{O}}^\dagger$. Their hermiticity implies that the corresponding eigenvalues will be real and that its eigenvectors will be orthogonal and provide a complete basis of the Hilbert space $\mathcal{H}$.

IV/ The outcome of measurements of $\mathcal{O}$ in the quantum system $|\Psi\rangle$ is predicted by the generalised statistical interpretation, which tells us that the possible outcomes are defined by the eigenvalues of $\widehat{\mathcal{O}}$ with well-defined probabilities.

V/ Any pair of physical observables A and B for which their commutator does not vanish, $\left[\hat{A}, \hat{B}\right] \neq 0$, will have associated a Heisenberg-like uncertainty relation that limits how precisely we can measure these two observables simultaneously.