Quantum Mechanics 2

Dr Juan Rojo
VU Amsterdam and Nikhef Theory Group
http://www.juanrojo.com/
j.rojo@vu.nl

Lecture notes, current version: March 2, 2021

Contents

Page

Quantum Mechanics 2

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

Current version: **March 2, 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

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

where $\mathbb{1}$ is the **identity operator** of this Hilbert space. 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	$\vert\Psi_1\rangle = \sum_{i=1}^n c_{1,i}\vert\psi_i\rangle$ $\vert\Psi_2\rangle = \sum_{i=1}^n c_{2,i}\vert\psi_i\rangle$	$\vert\Psi_1\rangle = \int dx\,\psi_1(x)\vert x\rangle$ $\vert\Psi_2\rangle = \int dx'\,\psi_2(x')\vert x'\rangle$
Inner product	$\langle\Psi_1\vert\Psi_2\rangle = \sum_{i=1}^n c_{1,i}^* c_{2,i}$	$\langle\Psi_1\vert\Psi_2\rangle = \int dx\,\psi_1^*(x)\psi_2(x)$
Orthonormality	$\langle\psi_i\vert\psi_i\rangle = \delta_{ij}$	$\langle x'\vert x\rangle = \delta(x-x')$
Completeness v2	$\sum_{i=1}^n \vert\psi_i\rangle\langle\psi_i\vert = 1$	$\int dx\,\vert x\rangle\langle x\vert = 1$

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.

The dimensions of a Hilbert space

The easiest way to evaluate the dimensionality of a Hilbert space is by counting **how many basis vectors** are required to span the whole space. That is, what is the *minimal number of independent vectors whose linear combinations can construct any arbitrary element* of this Hilbert space?
Note also that you should not confuse the dimensions of $\mathcal{H}$ with the *physical dimensions* of the quantum system. For example, a free particle moves in one spatial dimension, but its quantum states are elements of an infinite-dimensional Hilbert space.

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

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

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

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

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

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

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

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

is fixed by its action on the basis vectors,

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

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

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

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 $|\Psi\rangle = (1/\sqrt{2})\left(|+\rangle - i|-\rangle\right)$. First, we use the definition of the matrix elements Eq. (1.30) to find

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

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

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

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.

---

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

Consider a 3-dimensional Hilbert space. We are given the following two operators belonging to this Hilbert space. Can these operators represent appropriate 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.

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

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

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

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

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

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

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

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

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 $\mathcal{H}$:

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

---

[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.

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

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

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 $P(x)dx = |\psi(x)|^2 dx$ 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 $|\Psi_i\rangle$:

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

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

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

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

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

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

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

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

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

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

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

where in the last step I have used that $|z^2| \geq |(z - z^*)/(2i)|^2$ for any complex number z. Now making use

---

[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 $|\vec{u} \cdot \vec{v}| \leq |\vec{u}||\vec{v}|$.

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:

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

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

$$\left[\hat{A},\hat{B}\right] = \hat{A}\hat{B}-\hat{B}\hat{A}\,, \qquad (1.75)$$

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

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

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 uncertainty principle. We start from Eq. (1.76) and now we apply it to the position x and linear momentum p observables. We need to evaluate their commutator

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

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. (1.76) 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. (1.76) and they cannot have *a complete set of shared eigenfunctions*.

The uncertainty principle in finite spaces. Note that while you have been mostly exposed to the Heisenberg uncertainty principle in the case of infinite-dimensional Hilbert spaces, Eq. (1.76) also applies to finite Hilbert 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 as 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 do not exhibit a Heisenberg-like restriction and in this case $\sigma_{S_x}\sigma_{S_z}\geq 0$. In other words, there exist states $|\Psi\rangle$ for which the bound saturates and hence $\sigma_{S_x}\sigma_{S_z}=0$. We can construct an example of such configuration. Any quantum state of the form

$$|\Psi\rangle=c_+|+_y\rangle+c_i|-_y\rangle\,, \tag{1.79}$$

with $|c_+|^2 = |c_-|^2$ and $|+_y\rangle, |-_y\rangle$ being the eigenvectors of $\hat{S}_y$ will satisfy this condition since then

$$\left\langle \hat{S}_y \right\rangle = P\left(S_y = \frac{\hbar}{2}\right)\frac{\hbar}{2} + P\left(S_y = -\frac{\hbar}{2}\right)\frac{-\hbar}{2} = \frac{\hbar}{2}\left(|c_+|^2 - |c_-|^2\right) = 0\,, \tag{1.80}$$

where I have used the prescription from the generalised statistical interpretation. If now I express the eigenstates of $\hat{S}_y$ in terms of those of $\hat{S}_z$ I obtain

$$|+_y\rangle = \frac{1}{\sqrt{2}}\left(|+_z\rangle + |-_z\rangle\right)\,, \quad |-_y\rangle = \frac{1}{\sqrt{2}}\left(|+_z\rangle - |-_z\rangle\right)\,, \tag{1.81}$$

and we find that a possible quantum state for which $\left\langle \hat{S}_y \right\rangle = 0$ is given by

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|+_y\rangle + |-_y\rangle\right) = |+_z\rangle\,, \tag{1.82}$$

for which indeed the bound saturates, $\sigma_{S_x}\sigma_{S_z} = 0$: if I measure S_z first, I will have a 100% certainty that $S_z = \hbar/2$ and thus $\sigma_{S_z} = 0$. Note that however this bound will not be saturated if the order of the measurements is different, in particular if S_x is measured first, highlighting that at the end of the day S_x and S_z remain **incompatible observables**.

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. (1.76) 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

$$i\hbar\frac{\partial\Psi(x,t)}{\partial t} = \hat{H}\Psi(x,t) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\Psi(x,t)\,, \tag{1.83}$$

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:

$$\frac{d}{dt}\left\langle \mathcal{O}(x,p,t)\right\rangle = \frac{d}{dt}\left\langle \Psi \middle| \widehat{\mathcal{O}}\Psi\right\rangle = \left\langle \frac{\partial\Psi}{\partial t}\middle|\widehat{\mathcal{O}}\Psi\right\rangle + \left\langle \Psi\middle|\widehat{\mathcal{O}}\frac{\partial\Psi}{\partial t}\right\rangle + \left\langle \Psi\middle|\frac{\partial\widehat{\mathcal{O}}}{\partial t}\Psi\right\rangle\,, \tag{1.84}$$

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. (1.83), and hence we can write the above expression as

$$\frac{d}{dt}\left\langle \mathcal{O}\right\rangle = -\frac{1}{i\hbar}\left\langle \hat{H}\Psi\middle|\widehat{\mathcal{O}}\Psi\right\rangle + \frac{1}{i\hbar}\left\langle \Psi\middle|\widehat{\mathcal{O}}\hat{H}\Psi\right\rangle + \left\langle\frac{\partial\widehat{\mathcal{O}}}{\partial t}\right\rangle = \frac{i}{\hbar}\left\langle\left[\hat{H},\widehat{\mathcal{O}}\right]\right\rangle + \left\langle\frac{\partial\widehat{\mathcal{O}}}{\partial t}\right\rangle\,, \tag{1.85}$$

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*, $\mathcal{O}\left(x,p\right)$, we have that

  $$\frac{d}{dt}\left\langle \mathcal{O}\right\rangle = \frac{i}{\hbar}\left\langle \left[\hat{H},\widehat{\mathcal{O}}\right]\right\rangle \,, \tag{1.86}$$

  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

  $$\frac{d}{dt}\left\langle x\right\rangle = \frac{i}{\hbar}\left\langle \left[\hat{H},\hat{x}\right]\right\rangle = \left\langle \frac{\hat{p}}{m}\right\rangle \quad \rightarrow \quad m\frac{d}{dt}\left\langle x\right\rangle = \left\langle p\right\rangle \,, \tag{1.87}$$

  $$\frac{d}{dt}\left\langle p\right\rangle = \frac{i}{\hbar}\left\langle \left[\hat{H},\hat{p}\right]\right\rangle = -\left\langle \frac{dV(x)}{dx}\right\rangle \,, \tag{1.88}$$

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

  $$m\frac{d^2}{dt^2}\left\langle x\right\rangle = -\left\langle \frac{dV(x)}{dx}\right\rangle \,, \tag{1.89}$$

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

  $$m\frac{d^2}{dt^2}x = -\frac{dV(x)}{dx} = F \,, \tag{1.90}$$

  for conservative potentials.

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

  $$\frac{d}{dt}\left\langle \mathcal{O}\right\rangle = 0 \,, \tag{1.91}$$

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

Why it was necessary for our purposes to derive Eq. (1.85)? Assume now that we apply the generalised uncertainty principle to a time-independent observable $\mathcal{O}$ and to the Hamiltonian H,

$$\sigma_H\sigma_{\mathcal{O}} \ge \left|\frac{1}{2i}\left\langle \left[\hat{H},\widehat{\mathcal{O}}\right]\right\rangle\right| = \frac{\hbar}{2}\left|\frac{d\langle\mathcal{O}\rangle}{dt}\right| , \tag{1.92}$$

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:

$$\Delta E\,\Delta t \ge \frac{\hbar}{2} \,. \tag{1.93}$$

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

$$\sigma_{\mathcal{O}} = \left|\frac{d\langle\mathcal{O}\rangle}{dt}\right|\Delta t \,, \tag{1.94}$$

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.

Quantum Mechanics 2

Dr Juan Rojo
VU Amsterdam and Nikhef Theory Group
http://www.juanrojo.com/ , **j.rojo@vu.nl**
Current version: **March 2, 2021**

2 Chapter 4: Quantum Mechanics in Three Dimensions

Learning Goals

- To formulate and solve the Schroedinger equation in three dimensions.
- To identify and represent graphically the main features of the radial and angular components of the wave function of the hydrogen atom.
- To determine the conditions that lead to the quantisation of the energies and the wave functions in the three dimensional case.
- To apply the formalism of quantum mechanics to the cases of the extrinsic and intrinsic angular momentum.
- To understand the concept of the spin of elementary particles and its description in quantum theory, in particular for spin-1/2 particles.

In this section of the lecture notes we present the main concepts discussed in **Chapter 4 ("Quantum Mechanics in Three Dimensions")** 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

- **4.1**: The Schroedinger Equation.
- **4.2**: The Hydrogen Atom.
- **4.3**: Angular Momentum.
- **4.4**: Spin

Up to now, we have considered quantum systems composed by particles that move in **one spatial dimension**, such as the free particle, the quantum harmonic oscillator, and the particle in a well (remember that the dimensions of the Hilbert space are unrelated to the physical spatial dimensions of the system). While these systems are extremely useful to illustrate the basic concepts of quantum theory, to describe many realistic situations we need to apply the same quantum formalism to the cases of particles that move in **three spatial dimensions**, such as the electron in the hydrogen atom. In doing so, we will have to deal with the quantum version of *angular momentum* (which is obviously irrelevant for the one-dimensional case), and discuss new quantum phenomena that do not have a classical counterpart, such as the *spin* (intrinsic angular momentum) of elementary particles.

2.1 The Schroedinger equation in three dimensions

The starting point of our discussion will be the Schroedinger equation in three dimensions. Given the operator representation of the Hamiltonian, $\hat{H}$, one can write

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \hat{H}\Psi(\mathbf{r},t) = \left(-\frac{\hbar^2}{2m}(\hat{p}_x^2+\hat{p}_y^2+\hat{p}_z^2)+V(\mathbf{r})\right)\Psi(\mathbf{r},t)\,, \tag{2.1}$$

where[4] $\mathbf{r}=(x,y,z)$. By using the operator presentation of the linear momentum associated to the spatial coordinate r_i, namely $\hat{p}_{r_i}=-i\hbar(\partial/\partial r_i)$, we find that the Schroedinger equation reads

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \left(-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{r})\right)\Psi(\mathbf{r},t)\,, \tag{2.2}$$

in terms of the **Laplacian** expressed in Cartesian coordinates:

$$\nabla^2 = \frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\,. \tag{2.3}$$

The main difference as compared to the usual one-dimensional systems is that now the wave function (and the potential energy) depends on $\mathbf{r}$, so we need three coordinates to specify a point in space. In the following, we will focus on **stationary** solutions that satisfy the time-independent Schroedinger equation,

$$\left(-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{r})\right)\psi(\mathbf{r}) = E\psi(\mathbf{r})\,. \tag{2.4}$$

By deploying the formalism derived in the previous chapter, you can extend some of its key results to the case of three spatial dimensions, for example the commutation relations between coordinates and linear momenta:

$$[\hat{r}_i,\hat{p}_j]=i\hbar\delta_{ij}\,,\quad [\hat{r}_i,\hat{r}_j]=[\hat{p}_i,\hat{p}_j]=0\,,\quad i=1,2,3\,. \tag{2.5}$$

In this chapter, we will deal with quantum systems with a very important characteristic: **spherically symmetric** (also known as **central**) potentials. These are potential that depend *exclusively* on the magnitude of $\mathbf{r}$ but not on its direction, that is

$$V(\mathbf{r})=V(x,y,z)=V\left(\sqrt{x^2+y^2+z^2}\right)=V(r)\,,\quad r=\sqrt{x^2+y^2+z^2}=|\mathbf{r}|\,. \tag{2.6}$$

For this reason, it is convenient to switch from Cartesian coordinates to *spherical coordinates*, since these

---

[4] In the following, we will denote two- and three-dimensional vectors in bold font.

make explicit the symmetries that characterise the potential of the problem, in particular the fact that $V(r)$ depends only on the radial coordinate but not on the angular ones.

Examples of important *central potentials* are the Coulomb potential between two electric charges,

$$V(r) \propto \frac{Q_1 Q_2}{r}\,, \tag{2.7}$$

the Yukawa potential that describes interactions mediated by a massive (rather than a massless, as is the case of electromagnetism) particle,

$$V(r) \propto \frac{Q_1 Q_2 \exp\left(-Br\right)}{r}\,, \tag{2.8}$$

the spherical harmonic oscillator

$$V(r) \propto kr^2\,, \tag{2.9}$$

and the *infinite spherical well*, which is defined by $V(r) = 0$ inside a sphere of radius a $(r < a)$ and then $V(r) = \infty$ outside the sphere (for $r > a$), which is the three-dimensional analog of the infinite square well.

Spherical coordinates. Spherical coordinates (r, θ, ϕ) (where θ and ϕ are called the *polar* and *azimuthal* angles respectively) are related to Cartesian coordinates by the following transformation:

$$\begin{aligned} x &= r\sin\theta\cos\phi\,, \\ y &= r\sin\theta\sin\phi\,, \\ z &= r\cos\theta\,, \end{aligned} \tag{2.10}$$

where $r^2 = x^2 + y^2 + z^2$. Spherical coordinates span the following ranges:

$$0 \le r \le \infty\,, \qquad 0 \le \theta \le \pi\,, \qquad 0 \le \phi \le 2\pi\,. \tag{2.11}$$

The inverse transformation, from Cartesian coordinates to spherical coordinates, takes the form

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2}\,, \\ \theta &= \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) = \arccos\left(\frac{z}{r}\right)\,, \\ \phi &= \arctan\left(\frac{y}{x}\right)\,. \end{aligned} \tag{2.12}$$

Separation of variables. We can now express the time-independent Schroedinger equation in three-dimensions, Eq. (2.4) in spherical coordinates using the known expression of the Laplacian in this coordinate system, and we find

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\left(\frac{\partial^2\psi}{\partial\phi^2}\right)\right] + V(r)\psi = E\psi\,. \tag{2.13}$$

This equation appears rather more complex than those you have seen up to know, so we need to do some preparatory work. The most efficient method to solve this equation is to use the method of **separation of variables**. The starting point is an *ansatz* where we post that the dependence of the wave function on the radial coordinate factorises with respect to the angular one, that is,

$$\psi(r, \theta, \phi) = R(r)Y(\theta, \phi)\,. \tag{2.14}$$

Note that this is a hypothesis that might or might not work, and one needs to verify explicitely if this is indeed the case. Inserting the ansatz Eq. (2.14) and multiplying the time-independent Schroedinger equation, Eq. (2.13), by $-2mr^2/(YR\hbar^2)$, we end up with the following expression:

$$\left[\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR(r)}{dr}\right)-\frac{2mr^2}{\hbar^2}\left(V(r)-E\right)\right]=-\frac{1}{Y}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y(\theta,\phi)}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\left(\frac{\partial^2 Y(\theta,\phi)}{\partial\phi^2}\right)\right] \tag{2.15}$$

where we have made explicit that we assume a central potential $V(r)$. Our separation of variables strategy appears to have been successful, since we see that the LHS of the equation depends *only* on the radial coordinate r while the RHS *only* on the angular coordinates θ and ϕ.

The next step is based on the observation that the only way that two functions of different variables are equal to each other is if they are both *equal to some numerical constant.* If we denote this constant as $\ell(\ell+1)$, for reasons that will become clear soon, we end up with two separate second order differential equations

$$\left[\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR(r)}{dr}\right)-\frac{2mr^2}{\hbar^2}\left(V(r)-E\right)\right]=\ell(\ell+1)\,, \tag{2.16}$$

$$\frac{1}{Y}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y(\theta,\phi)}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\left(\frac{\partial^2 Y(\theta,\phi)}{\partial\phi^2}\right)\right]=-\ell(\ell+1)\,, \tag{2.17}$$

which we can solve sequentially. These equations are known as the *radial equation* and the *angular equation* respectively.

Note that separation of variables has lead to a very significant reduction of the complexity of the differential equations to solve, and that in particular the angular equation is **independent of the potential** and thus its solution will apply to any quantum system characterised by a central potential $V(r)$. While here we are mostly interested in applying the quantum formalism to the Coulomb potential that describes the hydrogen atom, the angular part of the solution will be the same for any other central potential, and therefore we will not need to reinvent the wheel in those cases.

2.1.1 The angular equation

Let us start by solving the angular equation, Eq. (2.17). Again, we will try the method of separation of variables, so we will posit an ansatz of the form $Y(\theta,\phi)=\Theta(\theta)\Phi(\phi)$. Inserting this ansatz and multiplying the equation by $\sin^2\theta$ we find that

$$\left[\frac{1}{\Theta}\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta(\theta)}{d\theta}\right)+\ell(\ell+1)\sin^2\theta\right]=-\frac{1}{\Phi}\frac{d^2\Phi(\phi)}{d\phi^2}\,, \tag{2.18}$$

where again we see that the separation of variables strategy has worked successfully, since the LHS depends only in θ while the RHS only on ϕ, and hence they must be equal to some constant, which we will call it without loss of generality m^2 (which for the time being denotes an arbitrary complex number). So at this point the radial equation involves two complex numbers, ℓ and m are complex numbers. As we will show next, there are boundary conditions that the wave function must satisfy that restrict in a significant manner the possible values that ℓ and m are allowed to take.

The differential equation in ϕ is particularly easy, and it is a good strategy in this life to start with the easy things and then move from there on to the most challenging tasks. The most general solution of the second order homogeneous differential equation

$$\frac{d^2\Phi(\phi)}{d\phi^2}=-m^2\Phi(\phi)\,, \tag{2.19}$$

will be given by

$$\Phi(\phi) = Ae^{im\phi} + Be^{-im\phi}\,, \tag{2.20}$$

with A and B being some integration constants.

Now we should use the fact that, from the definition of the spherical coordinates, it follows that the azimuthal angle ϕ is restricted to lie in the range Eq. (2.11). This property implies that ϕ and $\phi + 2\pi$ correspond to the same physical angle, and for example Eq. (2.20) must be the same wave function as

$$\Phi(\phi + 2\pi) = Ae^{im(\phi+2\pi)} + Be^{-im(\phi+2\pi)}\,. \tag{2.21}$$

(An attentive reader might have noticed a loophole here: wave functions are not observables, and what needs to be invariant under the $\phi \to \phi + 2\pi$ transformation is instead the *square* of the wave function, which is the physical observable given its probabilistic interpretation. Now, the square of Eq. (2.20) reads

$$|\Phi(\phi)|^2 = |A|^2 + |B|^2 + A^*Be^{-2im\phi} + AB^*e^{2im\phi}\,, \tag{2.22}$$

so the same argument as before applies, now in a more robust manner.)

The requirement that the wave function (and its square) must be invariant under the $\phi \to \phi + 2\pi$ transformation is what is known as a **periodic boundary conditions**, that is, a boundary condition applied to a periodic function (which is typical for wavefunctions that involve angles). This periodic boundary condition enforces the requirement

$$\Phi(\phi + 2\pi) = \Phi(\phi) \quad \text{implying} \quad e^{\pm i2\pi m} = 1 \quad \text{thus} \quad m = 0, \pm 1, \pm 2, \ldots \tag{2.23}$$

hence m must be an *integer number*. In other words, the solutions of Eq. (2.19) are **quantized** (discrete) with m being an *integer quantum number* that labels them. In summary, we can write the solution of Eq. (2.19) as

$$\Phi(\phi) = Ae^{im(\phi)}\,, \quad m = 0, \pm 1, \pm 2, \ldots\,, \tag{2.24}$$

which A some overall normalisation constant which will be fixed afterwards.

Quantisation from periodic boundary conditions

In previous courses, you saw how the quantisation of the wave function and of the allowed energies was often a consequence of the **boundary conditions** required on the wave function (and its derivative). For example, in the *particle in an infinite well* system, the quantisation of the energies followed from imposing that $\psi(x = 0) = \psi(x = L) = 0$, since the wave function must be continuous at the endpoints of the well (and the wavefunction vanished outside the well). Now we see that quantisation can also arise due to **periodic boundary conditions** in the case of angular variables.

In both cases, quantisation follows because the configuration of the quantum system imposes specific restrictions on the values that the wave function is allowed to take at certain points, and in turn by virtue of the Schroedinger equation this imposes restrictions on the allowed values for the energy E.

After having found the solutions for the azimuthal angle ϕ, we need to determine the dependence on the polar angle. The differential equation associated to the polar angle θ is a bit more involved:

$$\sin\theta \frac{d}{d\theta}\left(\sin\theta \frac{d\Theta(\theta)}{d\theta}\right) + \left(\ell(\ell+1)\sin^2\theta - m^2\right)\Theta = 0\,. \tag{2.25}$$

Fortunately, there is no need to reinvent the wheel, since this differential equation is well known and has been extensively studied. In the case of the constant ℓ being a *positive integer*, this differential equation corresponds to the **associated Legendre equation** whose solutions are given by

$$\Theta(\theta) = B P_l^m(\cos\theta)\,, \quad l = 0, 1, 2, 3, \ldots\,, \quad m = 0, \pm 1, \pm 2\, \ldots, \pm l\,, \tag{2.26}$$

with $P_l^m(x)$ being a special type of functions called the *associated Legendre functions* and B some overall normalisation constant that will be fixed in a while. The mathematical aficionados can learn more about these special functions online or in the textbook: for the purpose of this course, we will consider them as given and provide their explicit expressions whenever they are required. Rather than on mathematical derivations, our focus here will be on the properties of these solutions and on their physical interpretation. Suffice here to say that the associated Legendre functions are *real functions* (rather than complex) of its argument.

Note that since Eq. (2.26) depends on the azimuthal angle θ via the cosinus, it automatically satisfies the relevant **periodic boundary conditions**, given that

$$\Theta(\theta + 2\pi) = B P_l^m(\cos(\theta + 2\pi)) = B P_l^m(\cos(\theta)) = \Theta(\theta)\,.$$

Furthermore, the solution of the associated Legendre equation imposes that m, which previously could take any integer value, should satisfy the restriction that $|m| < l$ (else there is no solution).

The integration constants B in Eq. (2.26) and A in Eq. (2.24) can be fixed at the same time by the requirement that the wave function must be normalised. For practical reasons, it is convenient to impose that the radial $R(r)$ and angular $Y(\theta, \phi)$ components of the three-dimensional wave function are *separately normalised*, although in principle only its product needs to be normalised. In spherical coordinates, the normalisation conditions boils down to

$$\int_0^\infty |R(r)|^2 r^2\, dr = 1\,, \tag{2.27}$$

$$\int_0^\pi \int_0^{2\pi} |Y(\theta, \phi)|^2 \sin\theta\, d\theta\, d\phi = 1\,, \tag{2.28}$$

where the extra factors of r^2 (in Eq. (2.27)) and $\sin\theta$ (in Eq. (2.28)) as compared to the usual normalisation conditions that one applies in Cartesian coordinates arise from the Jacobian of the change of variables.

Taking into account this normalisation requirement, and labeling the angular wave function by the quantum numbers l and m which can only take the values listed in Eq. (2.26), we end up with the following expression of the angular component of the wave function that solves the Schroedinger equation in three-dimensions:

$$Y_\ell^m(\theta, \phi) = \sqrt{\frac{(2\ell + 1)}{4\pi} \frac{(\ell - m)!}{(\ell + m)!}} e^{im\phi} P_\ell^m(\cos\theta)\,, \quad \ell = 0, 1, 2, 3, \ldots \quad m = 0, \pm 1, \pm 2, \ldots, \pm \ell\,, \tag{2.29}$$

which are known as the **spherical harmonics**. Their explicit expressions are not particularly informative, and you can find them in the textbook. However, what will be rather useful in the subsequent discussion of atomic orbitals will be the *symmetry properties* of the spherical harmonics. Furthermore, the spherical

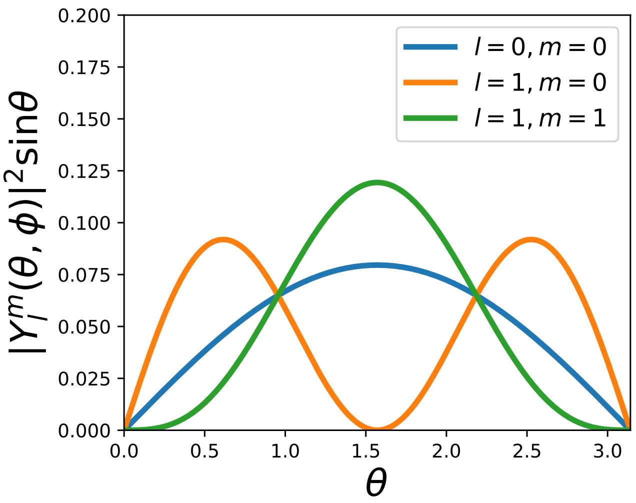

Figure 2.1: The probability densities in the polar angle θ associated to the first three spherical harmonics given in Eq. (2.32). The area under each curve integrates to one due to the normalisation condition Eq. (2.28).

harmonics defined this way are automatically orthonormal,

$$\int_0^\pi \int_0^{2\pi} \left[Y_{\ell'}^{m'}(\theta,\phi)\right]^* \left[Y_\ell^m(\theta,\phi)\right] \sin\theta \, d\theta \, d\phi = \delta_{\ell\ell'}\delta_{mm'} \, , \tag{2.30}$$

a property which will be important when assembling the wave functions of the hydrogen atom and related systems.

It is also interesting to work out what is the probability distribution associated to the angles θ and ϕ that one gets from the definition of the spherical harmonics. Taking into account the $P_l^m(x)$ are real functions, we can express the probability distribution in the angular variables as

$$P(\theta,\phi) = |Y_\ell^m(\theta,\phi)|^2 \sin\theta = \left(\frac{(2\ell+1)}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}\right)\left(P_\ell^m(\cos\theta)\right)^2 \sin\theta \tag{2.31}$$

The first thing we notice is that $P(\theta,\phi)$ does not depend on the azimuthal angle ϕ: this implies that all values of ϕ are *equally likely*. Then to determine the dependence on θ we need to indicate specific quantum numbers l and m. Below we provide some explicit examples:

$$\begin{aligned}
Y_0^0 &= \left(\frac{1}{4\pi}\right)^{1/2} & \rightarrow \qquad & P(\theta,\phi) = \frac{1}{4\pi}\sin\theta \, . \\
Y_1^0 &= \left(\frac{3}{4\pi}\right)^{1/2} \cos\theta & \rightarrow \qquad & P(\theta,\phi) = \left(\frac{3}{4\pi}\right)\cos^2\theta\sin\theta \, , \\
Y_1^{\pm 1} &= \mp\left(\frac{3}{8\pi}\right)^{1/2} \sin\theta e^{\pm i\phi} & \rightarrow \qquad & P(\theta,\phi) = \left(\frac{3}{8\pi}\right)\sin^3\theta \, .
\end{aligned} \tag{2.32}$$

These probability densities are shown in Fig. 2.1, where the area under each curve integrates to one due to the normalisation condition Eq. (2.28).

Before moving now to the radial equation, let us emphasize again that Eq. (2.29) will be the solution of any Schroedinger equation in three-dimensions with a radial potential. While for obvious reasons our main

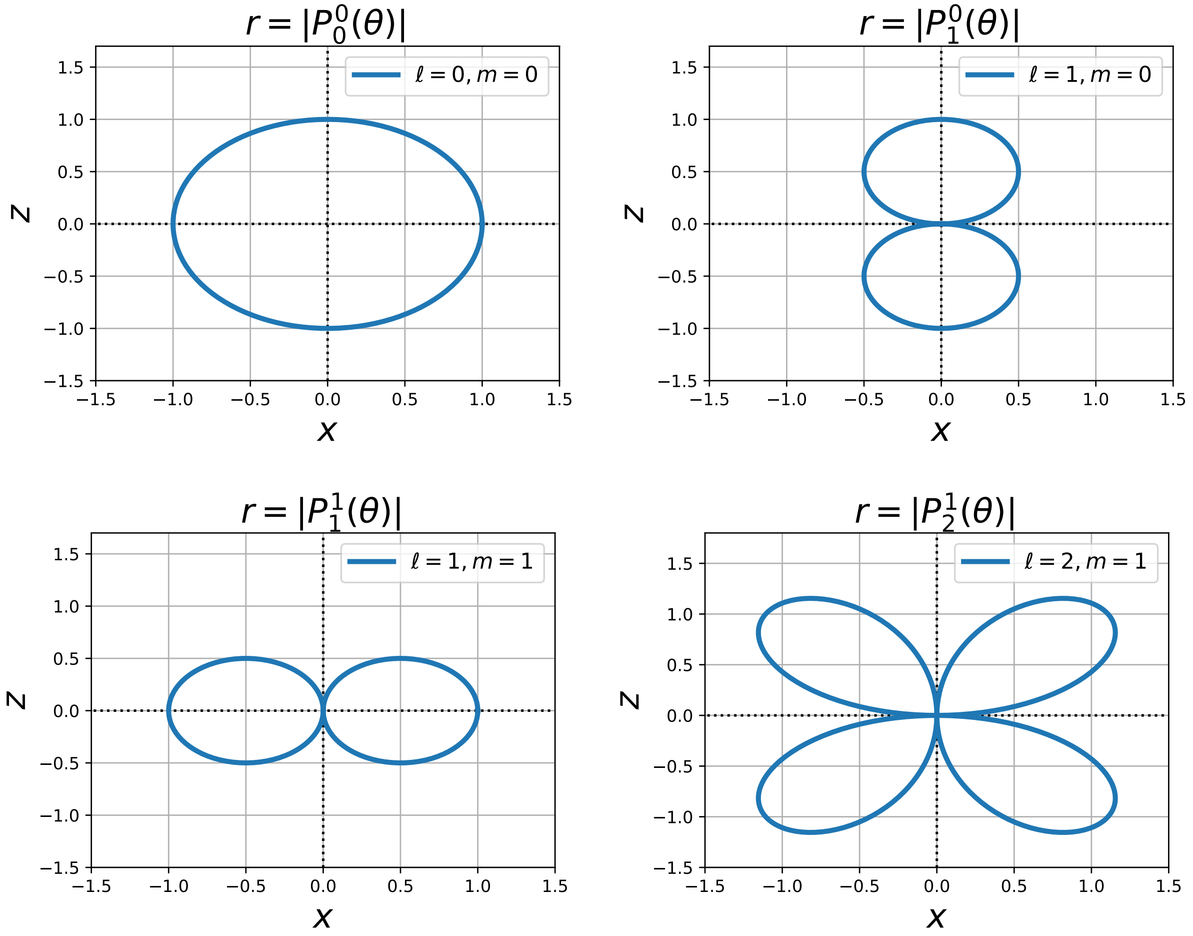

Figure 2.2: Graphical representation of the first few associated Legendre functions. The curves show the value of $r = |P_l^m(\theta)|$ as θ is varied from 0 to 2π.

interest is in the Coulomb potential, $V(r) \propto r^{-1}$, when we encounter other central potentials we can safely consider the angular part of the wave function as a done deal.

As we will demonstrate in the next section, the properties of the spherical harmonics will determine the **shape of the electronic orbitals** of the hydrogen atom. Fig. 2.2 displays a graphical representation of the first few associated Legendre functions:

$$\begin{aligned} P_0^0(\theta) &= 1\,, \\ P_1^0(\theta) &= \cos\theta\,, \\ P_1^1(\theta) &= -\sin\theta\,, \\ P_2^1(\theta) &= 3\sin\theta\cos\theta\,, \end{aligned}$$

where the curves indicate the value of $r = |P_l^m(\theta)|$ as θ is varied from 0 to 2π (there is rotational invariance around the z axis). Clearly, different values of ℓ and m will lead to different shapes of the electronic orbitals: for example $(\ell, m) = (0, 0)$ is spherically symmetric, while $(\ell, m) = (2, 1)$ corresponds to orbitals which are tilted $\pi/4$ with respect to the z axis.

2.1.2 The radial equation

We can now go back to Eq. (2.16) and attempt to solve the differential equation associated to the radial component of the total wave function

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR(r)}{dr}\right)-\frac{2mr^2}{\hbar^2}\left(V(r)-E\right)=\ell(\ell+1)\,, \tag{2.33}$$

with the added knowledge that ℓ is no longer an arbitrary complex number but that it must be a positive integer. This differential equation can be simplified quite a bit with a clever change of variables, $u(r)=rR(r)$, which after some algebra allows us to write Eq. (2.16) in a rather elegant way:

$$-\frac{\hbar^2}{2m}\frac{d^2u(r)}{dr^2}+\left[V(r)+\frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}\right]u(r)=Eu(r)\,, \tag{2.34}$$

which looks extremely similar to the standard one-dimensional Schroedinger equation for a wavefunction $u(r)$ if we define the **effective potential** as

$$V_{\rm eff}(r)\equiv V(r)+\frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}\,, \tag{2.35}$$

and hence we have that the radial component of the Schroedinger equation in three dimensions can be obtained by solving the following differential equation:

$$-\frac{\hbar^2}{2m}\frac{d^2u(r)}{dr^2}+V_{\rm eff}(r)u(r)=Eu(r)\,. \tag{2.36}$$

If you think a bit, this is a quite interesting result: it means that the radial component of the Schroedinger equation in *three-dimensions* is the solution of a differential equation that is nothing else but the Schroedinger equation in *one-dimensions* corresponding to the potential Eq. (2.35). Furthermore, recall that the normalisation condition Eq. (2.27) implies that

$$\int_0^\infty |u(r)|^2 dr = 1\,. \tag{2.37}$$

Without specifying Eq. (2.43) we cannot make any further progress. In the following we will solve the radial equation for two different potentials, first the *infinite spherical well* and then the *Coulomb potential* relevant for the description of the energy levels of the hydrogen atom.

The effective potential and angular momentum. Introducing the effective potential

$$V_{\rm eff}(r)\equiv V(r)+\frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}\,, \tag{2.38}$$

makes possible writing the differential equation for the radial component of the wave function as if it were that of a one-dimensional system. The second term in $V_{\rm eff}(r)$ is *positive definite*, that is, it tends to push the particle away from $r=0$, and the larger the value of ℓ the more intense this contribution will be.

Of course this effect is not unfamiliar: it is qualitatively the same as the *centrifugal force* that one experiences in a classical rotating body. As we will see below, this correspondence can be made more accurate once we introduce the angular momentum representation in quantum mechanics, and there we will see that the piece proportional to $\ell(\ell+1)$ arises indeed from the total angular momentum of the particle.

In other words, Eq. (2.43) might appear to describe a one-dimensional system, but it retains some *memory* that actually the particle is moving in three dimensions by means of the contribution to $V_{\rm eff}$ coming from the total angular momentum (the centrifugal force).

2.2 The infinite spherical well

One of the systems that you discussed in your earlier study of quantum mechanics in one dimension was the **infinite square well**, whose potential was defined as

$$\begin{aligned} V(x) &= 0 \quad \text{for } -a \le x \le a \\ V(x) &= \infty \quad \text{for } |x| > a \end{aligned}$$

with $2a$ being the width of the well. This potential corresponds to a configuration where a particle is restricted to move in the region $|x| < a$ where it obeys the free-particle Schroedinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) = E\psi(x)\,. \tag{2.39}$$

The solution to this equation is the usual free particle solution

$$\psi(x) = Ae^{ikx} + Be^{-ikx}\,, \quad k = \sqrt{2mE/\hbar^2} \tag{2.40}$$

subject to the continuity boundary conditions $\psi(x = -a) = \psi(x = a) = 0$. These boundary conditions implied that we can only have either *even* or *odd* solutions,

$$\psi(x) = A\cos(kx)\,, \quad \text{or} \quad \psi(x) = A\sin(kx)\,, \tag{2.41}$$

with the integration constant A now fixed by the normalisation condition of the wave function. Furthermore, for the *even solutions* the continuity boundary condition requires $\cos(ka) = 0$ so that $k = (\pi/2a)n$ with k being an odd integer ($n = 1, 3, 5, \ldots$). Hence we see that for the infinite quantum well, the **boundary conditions** on the wave function are the reason for the resulting **energy quantisation**:

$$E_n = \frac{\hbar^2\pi^2 n^2}{2m(2a)^2}\,, \quad n = 1, 3, 5, \ldots\,. \tag{2.42}$$

(the energy levels for $n = 2, 4, 6, \ldots$ are then obtained from imposing the continuity boundary conditions on the odd solutions).

We are now going to generalise this problem to the three dimensional case, which is known as the **infinite spherical well**, where now the corresponding potential is given by

$$\begin{aligned} V(r) &= 0 \quad \text{for } r \le a\,, \\ V(r) &= \infty \quad \text{for } r > a\,. \end{aligned}$$

In the same manner as for its one-dimensional counterpart, the wave function will vanish where the potential is infinite, $\psi(r, \theta, \phi) = 0$ for $r \ge a$, and hence we will also here have a *continuity boundary condition* demanding that $\psi(r = a, \theta, \phi) = 0$ for all values of the angular variables. This potential is clearly a *central potential*, since it depends only on r but not on the angular variables. Therefore, we can happily deploy all the machinery that we have developed in the previous sections concerning the solutions of the Schroedinger equation in three dimensions.

For the case of the infinite spherical well, the effective potential Eq. (2.35) is reduced to the term proportional to $\ell(\ell+1)$ (the centrifugal contribution) and hence the radial equation to be solved is:

$$-\frac{\hbar^2}{2m}\frac{d^2u(r)}{dr^2}+\frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}u(r)=Eu(r)\,. \tag{2.43}$$

which if we define (by analogy to the one-dimensional case) $k=\sqrt{2mE/\hbar^2}$ can be rewritten as

$$\frac{d^2u(r)}{dr^2}=\left[\frac{\ell(\ell+1)}{r^2}-k^2\right]u(r)\,. \tag{2.44}$$

Note that for the specific case of $\ell=0$ this differential equation for $u(r)$ has exactly the same form as that of its associated one-dimensional counterpart Eq. (2.39), and hence it will share its same solution. Note however that this does not mean that the physical wave function will be the same, since recall that we have used the transformation $u(r)=rR(r)$ (and also this is only the radial component of the total wave function).

Therefore, in the case $\ell=0$, the most general solution for the radial wave function in the infinite spherical well potential will be given by

$$R(r)=\frac{u(r)}{r}=A\frac{\cos(kr)}{r}+B\frac{\sin(kr)}{r}\,. \tag{2.45}$$

However, we note that not all possible values of A and B will lead to **physical solutions** for the wave function. In particular, since the wave function must be **finite everywhere**, we must impose that $A=0$, else the term proportional to $\cos(kr)/r$ would diverge in the $r\to 0$ limit. Furthermore, the *continuity of the wave function* requires that $R(r=a)=0$ since the wave function vanishes outside the barrier, as for the one-dimensional case. This requirement implies

$$B\frac{\sin(ka)}{a}=0\,,\quad\to\quad E_{N,0}=\frac{N^2\pi^2\hbar^2}{2ma^2}\,,\quad N=1,2,3,\ldots\,, \tag{2.46}$$

which are the same solutions for the one-dimensional well, once one accounts for both odd and even solutions. Finally we can determine the normalisation constant B by means of the normalisation condition Eq. (2.37) and we end up with

$$u_{N,0}(r)=\sqrt{\frac{2}{a}}\sin\left(\frac{N\pi r}{a}\right)\,,\quad r\le a\,,\quad N=1,2,3,\ldots\,, \tag{2.47}$$

and vanishes for $r>a$. So for the specific case of $\ell=0$ the wave functions $u(r)$ as the same as in the case of the infinite square well. We also note that the solutions Eq. (2.47) are labelled by an integer quantum number N.

A new type of boundary condition

We have seen from this example that another option to fix the integration constants that arise in the solutions of the Schroedinger equation is to demand that **only physical solutions are kept** and the unphysical ones are discarded. Specifically, we know that for a physical quantum state its **wave function must be finite everywhere**, else probabilities would be bigger than one which does not make sense. This requirement is specially handy in three-dimensional problems.

The radial equation in the infinite spherical well system, Eq. (2.44), can also be solved for an arbitrary positive integer ℓ in terms of two families of special functions, the *spherical Bessel functions* of order ℓ, $j_\ell(x)$

and the *spherical Neumann function* of order ℓ, $n_\ell(x)$:

$$u(r) = Arj_\ell(kr) + Brn_\ell(kr)\,, \tag{2.48}$$

where again the properties of these special functions have been thoroughly studied. Since $n_\ell(kr)$, diverges for $r \to 0$ we need to impose $B = 0$, and hence our candidate solution looks like

$$R(r) = Aj_\ell(kr)\,. \tag{2.49}$$

As in the case $\ell = 0$, we now need to impose the continuity boundary condition of the wave function at $r = a$. This boundary condition, $j_l(ka) = 0$, can be solved in terms of the zeros of the spherical Bessel functions and allows us to compute the quantised energy values $E_{N,\ell}$ for $\ell > 0$ in analogy with Eq. (2.46). If we note by $\beta_{N\ell}$ the N-th zero of the ℓ-th spherical Bessel function we find that the allowed energies are given by

$$E_{N,\ell} = \frac{\hbar^2}{2ma^2}\beta^2_{N\ell}\,. \tag{2.50}$$

Finally, we can combine the solution to the radial equation with the angular components of the wavefunction, which we get for free since the spherical harmonics are the same for any central potential, and we get

$$\psi_{nlm}(r,\theta,\phi) = A_{n\ell}j_\ell\left(\beta_{N\ell}\frac{r}{a}\right)Y_l^m(\theta,\phi)\,, \tag{2.51}$$

where $A_{n\ell}$ is determined from the normalisation condition and we have introduced the **principal quantum number** n which orders the allowed energies starting with $n = 1$ for the ground state of the system. Since

$$j_0(x) = \frac{\sin(x)}{x} \tag{2.52}$$

we see that $\beta_{N0} = N\pi$ and thus we reproduce the results that we have obtained for $\ell = 0$.

2.3 The Coulomb potential and the hydrogen atom

After this warm up with the infinite spherical well, we can move to the main goal of this chapter which is solving the Schroedinger equation in three-dimensions for a physical system defined by a proton and an electron interacting by means of the Coulomb potential:

$$V(r) = -\frac{e^2}{4\pi\epsilon_0}\frac{1}{r}\,, \tag{2.53}$$

which is required to construct the wavefunctions of the **hydrogen atom** and determine the spectrum of allowed energies. Therefore we now need to solve the radial equation for $u(r)$, Eq. (2.43), for the following effective potential

$$V_{\rm eff}(r) \equiv -\frac{e^2}{4\pi\epsilon_0}\frac{1}{r} + \frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}\,, \tag{2.54}$$

which is graphically represented in Fig. 2.3. for different values of the angular quantum number ℓ. Note how for ℓ there is no *centrifugal barrier* preventing the electron to become too close to $r = 0$, and indeed in this case as we will see $\psi(r = 0, \theta, \phi) \neq 0$. The larger the value of ℓ, the stronger the effect of this centrifugal barrier pushing the electron away from the origin.

We note that the Schroedinger equation for the Coulomb potential admits two types of solutions:

- The **bound states** of the hydrogen atom with negative energies, $E < 0$. We can see from Fig. 2.3

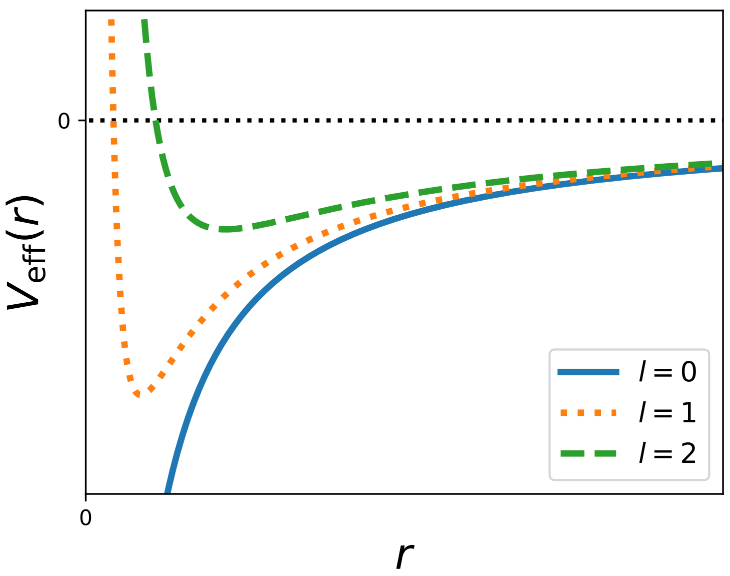

Figure 2.3: The effective potential $V_{\text{eff}}(r)$ associated to the Coulomb potential for different values of the angular quantum number ℓ. Note how for ℓ there is no *centrifugal barrier* preventing the electron to become too close to $r = 0$, and indeed in this case as we will see $\psi(r = 0, \theta, \phi) \neq 0$.

that if $E < 0$ the wave function eventually becomes exponentially suppressed for $r \to \infty$ (since there $E < V_{\text{eff}}(r)$, the classically forbidden region), which is the definition of a bound state (a wave function that vanishes asymptotically).

- The **scattering states** between an electron with positive energy, $E > 0$, and a proton. For $E > 0$, the wave function will displays for $r \to \infty$ the oscillatory behaviour characteristic of scattering states.

In the following we will focus on the bound states of the Coulomb potential, since these are the ones relevant for the description of the hydrogen atom. However, the eigenfunctions of the Schroedinger equation in three-dimensions will represent a complete basis only if one accounts both for the scattering and for the bound states.

The solution of the radial equation associated to the bound states of the Coulomb potential

$$-\frac{\hbar^2}{2m}\frac{d^2u(r)}{dr^2} + \left(-\frac{e^2}{4\pi\epsilon_0}\frac{1}{r} + \frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}\right)u(r) = Eu(r)\,, \quad E < 0\,, \tag{2.55}$$

is quite elegant from the mathematical point of view but provides limited insight from the physics point of view, so we will assume it as a given and point the interested student to the derivation in the textbook.[5] Here we only mention that in this derivation, much as in the case of the infinite spherical well of the previous section, we need to make sure to **eliminate unphysical solutions**, in particular those that lead to a blow-up of the wave function at either $r = 0$ or $r \to \infty$. One can also show that demanding the vanishing of the radial wave function $R(r)$ in the $r \to \infty$ limits leads to a **quantisation of the energy** and hence to the associated radial wave functions.

---

[5] In principle we should define r as the distance between the electron and the center of mass of the electron-proton system, and replace the mass of the electron by the *reduced mass* μ. However this is numerically a very small effect so effectively in what follows we are taking the proton to be infinitely heavy and thus motionless.

The end result of the calculation is the radial wave function for the Coulomb potential:

$$R_{n\ell}(r) = \sqrt{\left(\frac{2}{na}\right)^3 \frac{(n-\ell-1)!}{2n(n+\ell)!}} e^{-r/na} \left(\frac{2r}{na}\right)^l \left[L_{n-\ell-1}^{2\ell+1}(2r/na)\right] \,, \quad n=0,1,2,\ldots \quad \ell = 0,1,\ldots,n \tag{2.56}$$

where the overal prefactor ensures its appropriate normalisation:

$$\int_0^\infty r^2 R_{nl}^2(r)\, dr = 1 \,. \tag{2.57}$$

A number of observations about this important result, Eq. (2.56), are required at this point:

- The radial wave function $R_{n\ell}(r)$ depends on two quantum numbers: the **principal quantum number** n and the **angular quantum number** l. The latter was expected since ℓ entered the definition of the effective potential Eq. (2.54), which *sensu strictu* should have been labelled as $V_{\rm eff}^{(\ell)}(r)$. The former is a consequence of the requirement imposed by the boundary condition that $u(r)\to 0$ for $r\to\infty$ to ensure physical, normalisable wave functions (Since $R = u/r$).

  However, as we discuss below, the energy E_n depends only on the principal quantum number n but not on the angular one ℓ.

- The solution for $R_{n\ell}(r)$ depends on a special family of functions $L_q^p(x)$ known as the *associate Laguerre polynomials*. You can easily find their tabulated expressions online. Here the only property that we need to know about them is that they are finite when $x = 0$ and grow at most as $\propto x^q$ at large x.

- The dependence on the radial coordinate r appears always divided by a parameter a which has length dimensions and that is known as the **Bohr radius**:

  $$a \equiv \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 0.53\times 10^{-10}\,{\rm m}\,, \tag{2.58}$$

  which in some loose sense can be interpreted as the radius of the orbit that an electron has in the ground state of the hydrogen atom.[6]

- The explicit expressions for the radial wave function $R_{nl}(r)$ with low values of the quantum numbers n and l can be found in the textbook, and the reader is encouraged to take a closer look and represent some of them with the graphical software of their choice.

  Here we only point out that for $r \gtrsim na$ these radial wave functions are exponentially suppressed as $e^{-r/na}$, while for $r \lesssim na$ they exhibit a polynomial behaviour in r/a with the highest order being $(r/a)^{n-1}$ from the properties of the associate Laguerre polynomials.

- Of particular interest in the present discussion are the solutions where $\ell = 0$, that is, where the effective potential is purely Coulomb and there is no centrifugal contribution. In this case, you can verify that the associate Laguerre polynomials $L_{n-1}^1(2r/na)$ goes to a finite constant at $r = 0$, and hence the radial wave function exhibits the remarkable property that

  $$R_{n0}(r\to 0) \propto e^{-r/na} \,, \quad R_{n0}(r=0) = {\rm constant}\,, \tag{2.59}$$

  and hence the radial wave function *does not vanish at the origin* $r = 0$.

---

[6] Needless to day, the classical concept of orbits is meaningless in quantum mechanics, so below we will specify rather more precisely the physical interpretation of the Bohr radius.

> This might seem an odd result, since $r = 0$ corresponds of course to the spatial location of the proton (the nucleus of the hydrogen atom) and surely the proton and the electron cannot occupy exactly the same position? Note that this is a clearly unphysical configuration, since the Coulomb potential is singular for $r = 0$. The way out of this conundrum is the observation that while $R_{n0}(r = 0) \neq 0$, the corresponding **radial probability density** $P_{n\ell}(r) \equiv r^2|R_{n\ell}(r)|^2$ does indeed vanish at $r = 0$.
>
> Therefore indeed we confirm that the electron has a vanishing probability to be found at $r = 0$, as could be expected from general physics arguments.

Putting everything together, we conclude that the complete wave function for the **bound states of the hydrogen atom** is given by the following expression:

$$\psi_{n\ell m}(r,\theta,\phi) = R_{n\ell}(r)Y_{\ell m}(\theta,\phi)\,, \tag{2.60}$$

in terms of the radial component Eq. (2.56) and the spherical harmonics Eq. (2.29). These wave functions are also known as the **atomic orbitals**. The allowed ranges for the three quantum numbers n, ℓ, and m which label these solutions are given by

$$n = 1,2,3,\ldots\,, \quad \ell = 0,1,2,\ldots,n-1\,, \quad m = -\ell,-\ell+1,\ldots,\ell-1,\ell\,. \tag{2.61}$$

The radial and angular components of the wave function are separately normalised. While the wave functions depend on *three quantum numbers*, the energies themselves are **quantised** but depend only on the principal quantum number n:

$$E_n = -\left[\frac{m_e}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2\right]\frac{1}{n^2} = \frac{E_1}{n^2} = -\frac{13.6\,\mathrm{eV}}{n^2}\,, \tag{2.62}$$

with $E_1 = -13.6\,\mathrm{eV}$ representing the energy of the *ground state* ($n = 1$) of the hydrogen atom, also known as its **binding energy**: this is the energy that one needs to provide to remove an electron that occupies the ground state of the system.

The radius of the hydrogen atom. In Eq. (2.58) we introduced a numerical constant known as the *Bohr radius*, which we said had some relation with the size of a hydrogen atom. Let us make this connection more quantitative. We start with the radial wave function $R_{10}(r)$ corresponding to the *ground state of the hydrogen atom*:

$$R_{10}(r) = \frac{2}{a^{3/2}}\exp\left(-r/a\right)\,. \tag{2.63}$$

In quantum mechanics it makes no sense to ask a question such as *what is the radius of the electron orbit?*. Instead, questions that are meaningful to ask are:

- *What is the most likely value of radius of the electron orbit?*
- *What is the average value of the radius of the electron orbit?*

Note that in general the answer to these two questions, which superficially might appear to be the same, is different, so one has to be careful in checking what is being asked.

Let us answer these two questions in turn. We will need first to evaluate the corresponding radial probability density associated to this orbital:

$$P_{10}(r) = r^2|R_{10}(r)|^2 = \frac{4r^2}{a^3}\exp\left(-2r/a\right)\,. \tag{2.64}$$

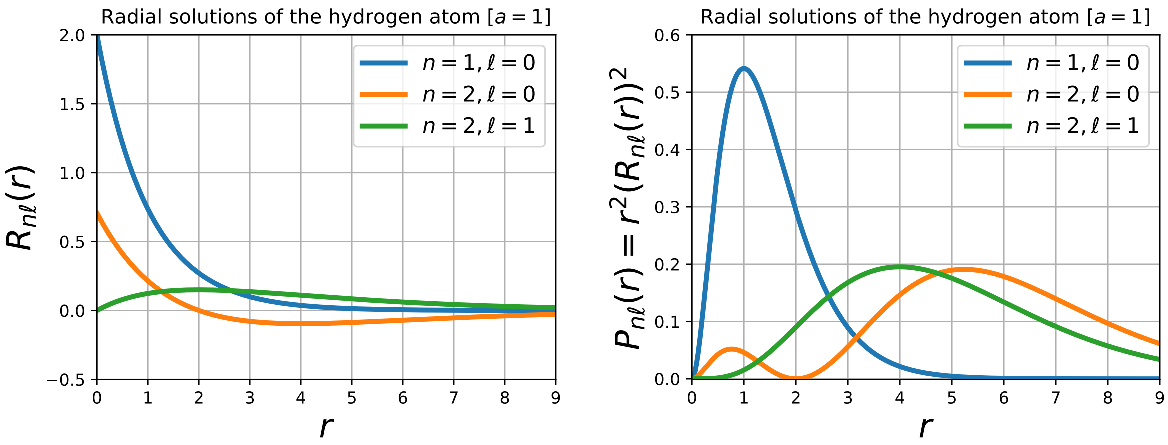

Figure 2.4: The radial wave functions $R_{n\ell}(r)$ and the associated probability distributions $P_{n\ell}(r)$ for the lowest three orbitals of the hydrogen atom (in units where $a = 1$). Note that while $R_{n0}(r = 0) \neq 0$, the associated probability densities vanish at the origin. Note also how the maximum of the probability distribution moves to the right (farther from the origin) as n increases; for $n = 1$ this maximum value is at $r^* = 1$, as we have derived in the text.

To evaluate the most likely value of the electron radius, r^*, the need to determine the local maximum of $P_{10}(r)$ by imposing that the first derivative vanishes:

$$\frac{d}{dr} P_{10}(r)\bigg|_{r=r^*} = 0 \quad \to \quad r^* = a \,, \tag{2.65}$$

and hence we see that the Bohr radius corresponds to the **most likely value of the radius** of the electron orbit in the ground state of hydrogen.

Next, we can evaluate the expectation value of the radius of the electron orbit:

$$\langle r \rangle = \int_0^\infty dr P_{10}(r) r = \frac{4}{a^3} \int_0^\infty dr\, r^3 \exp\left(-2r/a\right) = \frac{3}{2} a \,, \tag{2.66}$$

which is 50% higher than the Bohr radius, indicating that on average the electron orbits brings it farther away from the proton than the value indicated from the Bohr radius.

It is an interesting calculation to evaluate what is the dependence of both the most likely radius r^* and of its average value $\langle r \rangle$ of the electron's orbit as the quantum numbers n and ℓ are increased: you will evaluate some of them in the exercises of the tutorial session. Fig. 2.4 displays the radial wave functions $R_{n\ell}(r)$ and the associated probability distributions $P_{n\ell}(r)$ for the lowest three orbitals of the hydrogen atom (in units where the Bohr radius is $a = 1$). Note that while $R_{n0}(r = 0) \neq 0$, the associated probability densities vanish at the origin. Note also how the maximum of the probability distribution moves to the right (farther from the origin) as n increases; for $n = 1$ this maximum value is at $r^* = 1$, as we have derived above.

Visualisation of electronic orbitals of the hydrogen atom

In order to develop some intuition about the shape of the electronic orbitals of the hydrogen atom, it is useful to use onlive 3D visualisation tools such as this one:

`http://www.falstad.com/qmatom/`

Electronic transitions. An electron occupying an atomic orbital $\psi_{n\ell m}$ finds itself in an *stationary state*. However, external perturbations can induce **electronic transitions** between orbitals with different energy. For example, an electron occupying an orbital with $n_i > n_f$ can decay into the orbital with principal quantum number n_f. Since $E_{n_i} > E_{n_f}$, this process can only take place while ensuring **energy conservation** if energy is emitted, typically in the form of a photon. Using energy conservation, we get can evaluate the energy of the emitted photon

$$E_i = E_f + E_\gamma \quad \to \quad E_\gamma = E_i - E_f = (-13.6\,\mathrm{eV}) \times \left(\frac{1}{n_i^2} - \frac{1}{n_f^2} \right) , \tag{2.67}$$

where I have used the expression for the orbital energy in terms of the ionisation energy of hydrogen, Eq. (2.62). Likewise, an electron in an orbital with principal quantum number $n_i < n_f$ can transition to a higher orbital n_f via the *absorption of a photon*, whose energy is now

$$E_i + E_\gamma = E_f \quad \to \quad E_\gamma = E_f - E_i = (-13.6\,\mathrm{eV}) \times \left(\frac{1}{n_f^2} - \frac{1}{n_i^2} \right) . \tag{2.68}$$

Photons with energies corresponding to Eqns. (2.67) and (2.68) define what is known as the **spectrum of hydrogen**; you can relate these energies to wavelengths or frequencies using the Planck formula for the photons,

$$E_\gamma = \frac{hc}{\lambda} = h\nu \, . \tag{2.69}$$

The prediction of the spectra lines of hydrogen was one of the first historical triumphs of quantum theory.

Hydrogen-like atoms. The previous results have been derived assuming a Coulomb potential representing the attraction between the proton and the electron that compose a hydrogen atom,

$$V(r) = -\frac{e^2}{4\pi\epsilon_0}\frac{1}{r} \, . \tag{2.70}$$

However, it should be clear that the same discussion applies verbatim to the case of *hydrogen-like atoms*, systems defined as a single electron orbiting a nucleus composed by Z protons and $A - Z$ neutrons, with A being the atomic mass number (the neutrons are of course irrelevant in this context).

In this case, the Coulomb potential of this system is given by

$$V(r) = -\frac{Ze^2}{4\pi\epsilon_0}\frac{1}{r} \, , \tag{2.71}$$

so one can reuse all results obtained so far by doing the substitution $e^2 \to Ze^2$ in all formulae. For example, the bound state energy of a hydrogen-like atom will be given by

$$E_n(Z) = -\left[\frac{m_e}{2\hbar^2} \left(\frac{Ze^2}{4\pi\epsilon_0} \right)^2 \right] \frac{1}{n^2} = \frac{Z^2 E_1^H}{n^2} = -\frac{(13.6\,\mathrm{eV}) \times Z^2}{n^2} \, , \tag{2.72}$$

which are the same energies as for a hydrogen atom but now rescaled by a factor Z^2. In particular, the energy of the bound state of a hydrogen-like atom compares to that of the hydrogen atom by

$$E_1(Z) = E_1^H \times Z^2 \, , \tag{2.73}$$

which illustrates how the increase in the positive electric charge of the atomic nucleus leads to a stronger electronic bond in comparison with that of the hydrogen atom.

2.4 Angular momentum in quantum mechanics

Our main result in this Chapter so far has been Eq. (2.60), the complete wave function for the electron orbitals in the hydrogen atom, $\psi_{nlm}(r,\theta,\phi)$, which is labelled by three quantum numbers. The principal quantum number n determines the energy of the orbital via Eq. (2.62), while we know that l and m are somehow related to the *orbital angular momentum* (since they arise when considering the motion of the particle in the angular coordinates).

In order to formalize this connection in a more quantitative manner, our goal in this section is to present a quantum-mechanical theory of the **angular momentum**. Clearly, so far (in previous courses and in Chapter 3) we have limited ourselves to *one-dimensional problems* where angular momentum does not exist, so it is not surprising that it is only now, once we move to the three-dimensional configuration, that we need to worry about angular momentum in quantum mechanics.[7]

As you have learned in your study of classical mechanics, the angular momentum in three dimensions is defined as

$$\mathbf{L} = (L_x, L_y, L_x) = (yp_z - zp_y, zp_x - xp_z, xp_y - yp_x) = \mathbf{r} \times \mathbf{p}\,, \tag{2.74}$$

that is, the cross product between the position $\mathbf{r}$ of a particle (with respect to some reference point) and its linear momentum $\mathbf{p}$. The angular momentum is only non zero when the directions of $\mathbf{r}$ and $\mathbf{p}$ do not coincide (else the cross product vanishes).

What is the corresponding operator representation of the angular momentum $\widehat{\mathbf{L}}$ that one must deploy in the quantum-mechanical world? We just need to make the usual conversion between the position and linear momentum observables with their corresponding operators

$$\begin{gathered} x \to \hat{x} = x\,, \quad y \to \hat{y} = y\,, \quad z \to \hat{z} = z\,, \\ p_x \to \hat{p}_x = -i\hbar\frac{\partial}{\partial x}\,, \quad p_y \to \hat{p}_y = -i\hbar\frac{\partial}{\partial y}\,, \quad p_z \to \hat{p}_z = -i\hbar\frac{\partial}{\partial z}\,, \end{gathered} \tag{2.75}$$

and hence we find that the operator representation of the individual components of the angular momentum is given by

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\,, \tag{2.76}$$

$$\hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial x}\right)\,, \tag{2.77}$$

$$\hat{L}_x = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\,. \tag{2.78}$$

In the following, we will determine what are eigenvalues and the eigenvectors of $\widehat{\mathbf{L}}$ and demonstrate that these have a deep connection with the wave functions of the hydrogen atom that we have just derived.

First of all, one should derive the commutation relations for the angular momentum operator. In order to derive them, the only property that we need to use here is the commutation relations between the spatial coordinates and the linear momenta in the three-dimensional case

$$[\hat{r}_i, \hat{p}_j] = i\hbar\delta_{ij}\,, \quad [\hat{r}_i, \hat{r}_j] = [\hat{p}_i, \hat{p}_j] = 0\,, \tag{2.79}$$

---

[7] Actually, this statement is not quite true, since even for an electron moving in one dimension one needs to consider its spin, and thus a quantum theory of angular momentum is already necessary in that case.

and using these properties we find

$$\left[\hat{L}_x, \hat{L}_y\right] = [yp_z - zp_y, zp_x - zp_z] = [yp_z, zp_x] - [yp_z, zp_z] - [zp_y, zp_x] + [zp_y, zp_z] \, . \tag{2.80}$$

Next we need to apply the commutation relations for r_i and p_j, for example $[zp_y, zp_x] = 0$ and

$$[yp_z, zp_x] = \hat{y}\hat{p}_z\hat{z}\hat{p}_x - \hat{z}\hat{p}_x\hat{y}\hat{p}_z = \hat{y}\hat{p}_x\hat{p}_z\hat{z} - \hat{y}\hat{p}_x\hat{z}\hat{p}_z = \hat{y}\hat{p}_x\left[\hat{p}_z, \hat{z}\right] = -i\hbar\hat{y}\hat{p}_x \, .$$

Using these relations and after some algebra we find

$$[L_x, L_y] = i\hbar\left(xp_y - yp_x\right) = i\hbar L_z \, , \tag{2.81}$$

and likewise for the other commutators. In summary we find that the commutation relations between the individual components of the angular momentum operator are given by:

$$[L_x, L_y] = i\hbar L_z \, , \quad [L_y, L_z] = i\hbar L_x \, , \quad [L_z, L_x] = i\hbar L_y \, . \tag{2.82}$$

That is, the commutator between two components of the angular momentum is proportional to the other component. Following the discussion about the formalism of quantum mechanics that we had in the previous chapter, we can derive some some important consequences from this result:

Commutation relations for angular momentum

Since the three components of the angular momentum do not commute, they will have associated a Heisenberg-type **uncertainty relation**. Furthermore, it will not the possible to find a **complete set of functions** which are simultaneously eigenfunctions of L_i and L_j for $i \neq j$. Using the terminology that we derived in the previous section, we say that L_x, L_y, and L_z are **incompatible observables** that cannot be measured simultaneously, at least in a general quantum state.

The fact that L_i and L_j for $i \neq j$ do not commute implies therefore an uncertainty relation of the form

$$\sigma_{L_i}\sigma_{L_j} \geq \frac{\hbar}{2}\Big|\left\langle L_k\right\rangle\Big| \, , \quad i \neq j \neq k \, . \tag{2.83}$$

Note however how in the specific case where we have a quantum state such that $\langle L_k\rangle = 0$ then L_i and L_j *do commute* and thus can be measured simultaneously. But in general, it will be impossible to find a set of functions which are simultaneously eigenfunctions of σ_{L_i} and σ_{L_j} for $i \neq j$.

Furthermore, although at first sight it might seem unintuitive, the **square of the angular momentum** L^2 does commute with its individual components. Indeed, if one defines

$$L^2 \equiv L_x^2 + L_y^2 + L_z^2 \, , \tag{2.84}$$

it can be shown L^2 commutes with L_i. For example, let us evaluate the following commutator:

$$\left[L^2, L_x\right] = \left[L_x^2 + L_y^2 + L_z^2, L_x\right] = \left[L_y^2 + L_z^2, L_x\right] \tag{2.85}$$

since of course L_x commutes with itself. To simplify this expression, we can use a useful relation involving the commutators of three operators $\hat{A}, \hat{B}$ and $\hat{C}$,

$$\left[\hat{A}\hat{B}, \hat{C}\right] = \hat{A}\left[\hat{B}, \hat{C}\right] + \left[\hat{A}, \hat{C}\right]\hat{B}\,, \tag{2.86}$$

which as you can check follows from the definition of the commutator. Therefore, we have that

$$\begin{aligned}
\left[L^2, L_x\right] &= L_y\left[L_y, L_x\right] + \left[L_y, L_x\right]L_y + L_z\left[L_z, L_x\right] + \left[L_z, L_x\right]L_z \\
&= L_y(-i\hbar L_z) + (-i\hbar L_z)L_y + L_z(i\hbar L_y) + (i\hbar L_y)L_z \\
&= i\hbar\left(-L_yL_z - L_zL_y + L_zL_y + L_yL_z\right) = 0\,.
\end{aligned} \tag{2.87}$$

Since the same derivation applies to the other components, we have demonstrated that the total angular momentum commutes with its individual components:

$$\left[L^2, L_x\right] = \left[L^2, L_y\right] = \left[L^2, L_z\right] = 0\,. \tag{2.88}$$

Note that this result has the important consequence that:

Angular momentum measurements

In a general quantum system, it will be possible to *simultaneously determine* its **total angular momentum** and **its component in one specific direction** (but not for more than one direction!). Hence, each individual component of $\mathbf{L}$ is **compatible** with L^2 and one can find a set of eigenfunctions that simultaneously diagonalises the two operators. Specifically, there exists a complete set of functions that are at the same time eigenfunctions of L^2 and of L_z. As we discuss below, it happens that these eigenfunctions will also be eigenfunctions of the Hamiltonian $\hat{H}$.

The above property implies that it should be possible to find functions $f(\mathbf{x})$ that satisfy the corresponding eigenvalue equations:

$$L^2 f(\mathbf{x}) = \lambda f(\mathbf{x})\,, \quad L_z f(\mathbf{x}) = \mu f(\mathbf{x})\,, \tag{2.89}$$

or the equivalent relations with L_x or L_z. The allowed eigenvalues can be determined using a *ladder strategy*, similar as what you did with the harmonic oscillator. The derivation is mathematically elegant but not particularly enlightening from the physics point of view, so we refer the interested student to the textbook and here we just quote the final result:

$$L^2 f_\ell^m(\mathbf{x}) = \hbar^2\ell(\ell+1) f_\ell^m(\mathbf{x})\,, \quad L_z f_\ell^m(\mathbf{x}) = \hbar m f_\ell^m(\mathbf{x})\,, \tag{2.90}$$

with the quantised eigenfunctions being labelled by two quantum numbers ℓ and m which can take the following values:

$$\ell = 0, 1/2, 1, 3/2, \ldots\,; \quad m = -\ell, -\ell+1, \ldots, \ell-1, \ell\,. \tag{2.91}$$

This result should be familiar to you: it looks suspiciously close to the spherical harmonics that we defined in Eq. (2.29). Of course, this is no coincidence, and as we will discuss next the spherical harmonics are the sought-for **simultaneous eigenfunctions of** L^2 **and** L_z. The only difference appears to be that Eq. (2.91) allows for half-integer values of ℓ: these are *not* spurious solutions, but rather related to a new type of angular momentum which is unique to quantum mechanics: the **spin** of elementary particles, which will be discussed in Sect. 2.5.

On the relation between L^2 and L_z

From our discussion of the formalism of quantum mechanics, we know that the outcomes of a possible measurement of L^2 will be given by its eigenvectors, $\hbar^2\ell(\ell+1)$, and likewise the outcome of a given measurement of L_z will return $\hbar m$ subject to the constraint of Eq. (2.91). Interestingly, one notes that in all cases the outcome of the measurement of $\sqrt{L^2}$ will be **larger** that that of L_z since $\sqrt{\ell(\ell+1)} > \ell$ and $m_{\rm max} = \ell$. The reason for this result is that if a measurement of $\sqrt{L^2}$ returned the same outcome than that of L_z it would mean that $L_x = L_y = 0$. This would imply **absolute certainty** on the knowledge of all components of $\mathbf{L}$, which is incompatible with Heisenberg's uncertainty Eq. (2.83).

In order to determine the simultaneous eigenfunctions of L^2 and L_z that are defined by Eq. (2.90) it is convenient to switch to spherical coordinates, the reason being that the dependence on the radial coordinate drops out and the differential equations to be solved will depend only on the angular coordinates θ and ϕ. If you write the angular momentum $\mathbf{L} = -i\hbar\left(\mathbf{r}\times\nabla\right)$ and express the gradient in spherical coordinates after some algebra you find that the operator expression for the angular momentum in the z direction is given by

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi}\,, \tag{2.92}$$

whose eigenfunctions are the same ones as those of Eq. (2.19). Indeed, if we solve the associated eigenvalue equations in Cartesian coordinates

$$\hat{L}_z f_\ell^m(r,\theta,\phi) = -i\hbar\frac{\partial}{\partial\phi}f_\ell^m(r,\theta,\phi) = \hbar m f_\ell^m(r,\theta,\phi)\,, \tag{2.93}$$

we easily see that the eigenfunctions are

$$f_\ell^m(r,\theta,\phi) = g(r,\theta)e^{im\phi}\,, \tag{2.94}$$

with the periodic boundary conditions in ϕ being ensured by the constraints in Eq. (2.91) (if we discard the half-integer solutions, which in this case are *not* compatible with the periodic boundary conditions.)

Next, we note that the squared angular momentum, given by

$$L^2 = \mathbf{L}\cdot\mathbf{L} = -\hbar^2\left(\mathbf{r}\times\nabla\right)\cdot\left(\mathbf{r}\times\nabla\right) \tag{2.95}$$

can also be expressed in spherical coordinates and is given by

$$\hat{L}^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]\,. \tag{2.96}$$

You can derive this result either by expressing $\left(\mathbf{r}\times\nabla\right)\cdot\left(\mathbf{r}\times\nabla\right)$ in terms of the Laplacian and then transforming to spherical coordinates, or by using the raising and lowering operator method. Now, we note that we have already encountered this equation! Indeed, when solving the angular equation for θ, Eq. (2.25), we had that the θ-dependent part of the angular wave function obeyed

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta(\theta)}{d\theta}\right) + \left(\ell(\ell+1) - \frac{m^2}{\sin^2\theta}\right)\Theta = 0\,. \tag{2.97}$$

Therefore, the solution of the eigenvalue equation for $\hat{L}^2$

$$\hat{L}^2 f_\ell^m(r,\theta,\phi) = -\hbar^2 \left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right] f_\ell^m(r,\theta,\phi) = \hbar^2\ell(\ell+1) f_\ell^m(r,\theta,\phi)\,, \tag{2.98}$$

is nothing but the spherical harmonics $Y_l^m(\theta,\phi)$, given that we have demonstrated that:

$$f_\ell^m(r,\theta,\phi) = h(r)\Theta_l^m(\theta)e^{im\phi}\,, \tag{2.99}$$

Hence we can conclude that there spherical harmonics are the sought-for **simultaneous eigenfunctions of L^2 and L_z**, which makes possible achieving a deeper understanding of the properties of the electronic orbitals of the hydrogen atom.

Eigenvalues of the hydrogen atom orbitals

We have found that the solutions of the Schroedinger equation for the electronic orbitals of the hydrogen atom are simultaneous eigenfunctions of three operators: the Hamiltonian $\hat{H}$, the total angular momentum L^2 and its component in the z direction, L_z, with eigenvalues being given by

$$\begin{aligned}
\hat{H}\psi_{nlm}(r,\theta,\phi) &= E\psi_{nlm}(r,\theta,\phi)\,,\\
L^2\psi_{nlm}(r,\theta,\phi) &= \hbar^2\ell(\ell+1)\psi_{nlm}(r,\theta,\phi)\,,\\
L_z\psi_{nlm}(r,\theta,\phi) &= \hbar m\psi_{nlm}(r,\theta,\phi)\,.
\end{aligned}$$

This result implies that the can **simultaneously measure** E, L^2 and L_z of a given electronic orbital.

The property that $\hat{H}$, L^2, and L_z all commute between themselves (and hence a simultaneous eigenfunction basis exists) can be also noted by expressing the three-dimensional Schroedinger equation as

$$\hat{H}\psi_{nlm} = \frac{1}{2mr^2}\left[-\hbar^2\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \hat{L}^2\right]\psi_{nlm} + V(r)\psi_{nlm} = E\psi_{nlm} \tag{2.100}$$

and noting that $\hat{L}^2$ commutes with the Hamiltonian, $\left[\hat{H},\hat{L}^2\right]$, since it only acts on the angular variables while the non-$\hat{L}^2$ part of $\hat{H}$ acts only on the radial coordinate. Hence we can write

$$\left[\hat{H},\hat{L}^2\right] = \left[\hat{H},\hat{L}_z\right] = \left[\hat{L}_z,\hat{L}^2\right] = 0\,, \tag{2.101}$$

which is consistent with the existence of the simultaneous eigenvector basis that we have just derived.

2.5 Intrinsic angular momentum: Spin

When we determined the eigenvalues of the total angular momentum operator $\hat{L}^2$ in Eq. (2.90), we found that the angular quantum number ℓ could take both *integer* and *half-integer* values: $\ell = 0, 1/2, 1, 3/2, 2$, and so on. However, for the discussion of the solutions of the hydrogen atom, only the solutions with $\ell = 0, 1, 2, \ldots$ were relevant for the spherical harmonics, else the periodic boundary conditions would not be satisfied. For example, if $m = \pm 1/2$ then the spherical harmonics are not invariant under the $\phi \to \phi + 2\pi$ transformation. What about the half-integer solutions? Are they simply unphysical solutions to some differential equation and we can happily throw then away? As we will see, nothing would be more mistaken: these half-integer solutions play a crucial role in quantum mechanics and can be associated to one phenomenon that does not have a classical counterpart: the **spin** (or intrinsic angular momentum) of elementary particles.

Let us recall that in classical mechanics an object admits two types of angular momentum:

- **Orbital** or **extrinsic angular momentum**, associated to the motion of the center of mass of the object with respect to some external reference point: $\mathbf{L} = \mathbf{r} \times \mathbf{p}$.
- **Intrinsic angular momentum** or **spin**, associated with the rotation of the object with respect to some axis that passes through its center of mass, $\mathbf{L} = I\boldsymbol{\omega}$ where $\boldsymbol{\omega}$ represents the angular velocity and I the moment of inertia of the body with respect to the aforementioned axis.

  Note that this intrinsic angular momentum is not something conceptually different from the orbital one: it is only the sum of orbital angular momentum from all the constituents of the rigid body.

Hence, in classical mechanics the orbital (extrinsic) and the intrinsic (spin) angular momentum are essentially the same physical quantity, just evaluated in a different manner. However, in quantum mechanics it is essential to separate the extrinsic angular momentum $\mathbf{L}$ from the internal one, which we call spin and denote by the symbol $\mathbf{S}$.

Spin in quantum mechanics

The quantum mechanical spin $\mathbf{S}$ is a **fundamental property of elementary particles**, such as electron or the proton, with no counterpart in classical physics. In this respect, spin is on a similar footing to the charge or the mass of a particle: it is an intrinsic property that cannot be changed **without modifying the identity of the particle**. Also, do not push too far the analogy of imaging a tiny electron spinning around its axis, since one quickly finds unphysical results this way.

From the mathematical point of view, the distinction between orbital and intrinsic angular momentum remains as immaterial as in the classical case, and thus the **quantum theory of spin** follows readily from our previous discussion of angular momentum (it is essentially the same physical quantity, at the end of the day). In particular, given that we still live in three dimensions, we can define the components S_x, S_y, S_z of the spin of a particle (recall, its intrinsic angular momentum) whose associated operators $\hat{S}_x, \hat{S}_y, \hat{S}_z$ satisfy the usual commutation relations for angular momentum, recall Eq. (2.82),

$$\left[\hat{S}_x, \hat{S}_y\right] = i\hbar\hat{S}_z\,, \quad \left[\hat{S}_y, \hat{S}_z\right] = i\hbar\hat{S}_x \quad \left[\hat{S}_z, \hat{S}_x\right] = i\hbar\hat{S}_y\,. \tag{2.102}$$

Likewise, if we denote by s the quantum number that labels the total spin and m that of its component of in the z direction[8], and we denote by $|s,m\rangle$ an eigenstate of both $\hat{S}^2$ and $\hat{S}_z$, we know from our previous discussion of angular momentum that their eigenvalues will be

$$\hat{S}^2|s,m\rangle = \hbar^2 s(s+1)|s,m\rangle\,, \quad \hat{S}_z|s,m\rangle = \hbar m|s,m\rangle\,. \tag{2.103}$$

Furthermore, we can define the *raising* and *lowering* spin operators whose action on those eigenstates is

$$\hat{S}_\pm|s,m\rangle = \hbar\sqrt{(s(s+1) - m(m\pm1))}|s,m\pm1\rangle\,, \quad \hat{S}_\pm \equiv \hat{S}_x \pm i\hat{S}_y\,, \tag{2.104}$$

a relation that can be verified by using the commutation relations of Eq. (2.102), in particular

$$\left[\hat{S}_z, \hat{S}_\pm\right] = \pm\hbar S_\pm\,, \quad \left[\hat{S}^2, \hat{S}_\pm\right] = 0\,. \tag{2.105}$$

---

[8] In some textbooks one uses instead m_s, to separate it from m_l which is associated to the z component of the orbital spin. This distinction is not crucial here, but will become important in future chapters.

These operators are denoted as raising and lowering operators because of the following property:

$$\hat{S}_z\left(\hat{S}_\pm|s,m\rangle\right) = \hbar(m\pm 1)\left(\hat{S}_\pm|s,m\rangle\right)\,, \tag{2.106}$$

and hence $\hat{S}_\pm|s,m\rangle$ is also an eigenfunction of $\hat{S}_z$ but now with eigenvalue $\hbar(m\pm 1)$ instead of $\hbar m$ as in the original state $|s,m\rangle$, hence justifying the property of Eq. (2.104).

A crucial difference in the discussion of spin as compared to that of the orbital angular momentum is that now the eigenfunctions *do not need to be* the spherical harmonics (actually, there is not even a reference to the angles θ and ϕ when discussing spin!). That is, the spherical harmonics $Y_l^m(\theta,\phi)$ are certainly eigenvalues of $\hat{L}^2$ and $\hat{L}_z$ but they are not the only possible solutions of the angular momentum (recall that spin is just angular momentum) operator eigenvalue equations. Therefore, we should use the full set of values of s and m without excluding the half-integer values

$$s = 0, \frac{1}{2}, 1, \frac{3}{2}\, 2, \ldots\,, \qquad m = -s, -s+1, \ldots, s-1, s\,. \tag{2.107}$$

For each value of the spin s of a given particle, we can construct explicit matrix representations of the spin operators $\hat{S}_z$ and $\hat{S}^2$.

Now, you might ask, what determines the value of the total spin quantum number s of a given particle, and can s change as a consequence of some interactions? Actually it turns out that no, and that as mentioned above the spin s of a particle is one of its defining properties much in the same way as its mass m and its electric charge q:

The spin of elementary particles

According to their spin s, elementary particles can be divided into two groups: **fermions** with half-integer values of the spin, and **bosons** for integer spin values. For examples, electrons and quarks (and protons and neutrons) are fermions with spin $s = 1/2$, while photons are bosons with $s = 1$. The recently discovered Higgs boson is a spin-less particle with zero spin, $s = 0$. This classification is essential in that that the spin of a particle has a very significant impact on its properties and their interactions, as we will verify in the next Chapter.

The theory of spin-1/2. The above discussion holds for any value of the spin s associated to elementary particles. In this course, we will focus mostly on $s = 1/2$, which is the spin quantum number associated to electrons and protons, and hence of direct relevance to study quantum mechanics for electronic, atomic, and molecular physics. We will also consider spin $s = 1$, which is interesting both from the conceptual point of view as well as relevant to discuss the interactions of electromagnetic light (composed by photons, $s = 1$ particles) with matter.

With this motivation, let us now present the quantum theory of $s = 1/2$ particles. The eigenvalue equations from Eq. (2.103) in this case read

$$S^2\Big|\frac{1}{2},m\Big\rangle = \frac{3\hbar^2}{4}\Big|\frac{1}{2},m\Big\rangle\,, \quad S_z\Big|\frac{1}{2},m\Big\rangle = \hbar m\Big|\frac{1}{2},m\Big\rangle\,, \tag{2.108}$$

with $m = +1/2$ or $m = -1/2$ being the only two allowed values of the quantum number associated to S_z. Since s is fixed and $m = \pm 1/2$, spin-1/2 particles live in a **two-dimensional Hilbert space** (for the spin-related part of the wave functions). In the literature one can find a large variety of options in the notation to refer to these two eigenstates of $\hat{S}_z$, including:

- **spin up**: $\left|\frac{1}{2},\frac{1}{2}\right\rangle$, $|+_z\rangle$, $|\uparrow_z\rangle$, χ_+, ...
- **spin down**: $\left|\frac{1}{2},-\frac{1}{2}\right\rangle$, $|-_z\rangle$, $|\downarrow_z\rangle$, χ_-, ...

In most cases, it is easy to deduce by context which notation is being used to denote these eigenstates of $\hat{S}_z$, but please be careful specially when other components of $\hat{\mathbf{S}}$ are also involved.

Given that the eigenvectors of $\hat{S}_z$ form a complete basis ($\hat{S}_z$ being an Hermitian operator), we can express the **most general spin state** of a spin-1/2 particle as an electron as follows

$$|\Psi\rangle = c_+|+\rangle + c_-|-\rangle = \begin{pmatrix} c_+ \\ c_- \end{pmatrix}\,, \quad |+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\,, \quad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.109}$$

in terms of the eigenstates $|+\rangle$ and $|-\rangle$ expressed in their conventional (though by no means unique) column vector representation. We can work out the matrix representation of the operators $\hat{S}^2$ and $\hat{S}_z$ in this specific basis. Let us start with the total spin operator $\hat{S}^2$. The equations that its components must satisfy are

$$\hat{S}^2|+\rangle = \frac{3\hbar^2}{4}|+\rangle \quad \to \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{3\hbar^2}{4}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.110}$$

$$\hat{S}^2|-\rangle = \frac{3\hbar^2}{4}|-\rangle \quad \to \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{3\hbar^2}{4}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.111}$$

from where we see that

$$\hat{S}^2 = \frac{3\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\,. \tag{2.112}$$

Likewise, you can determine that the matrix representation of $\hat{S}_z$ in this basis is given by

$$\hat{S}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,. \tag{2.113}$$

You can verify that the eigenvalue equation for $\hat{S}_z$ has eigenvalues $\pm\hbar/2$ with eigenvectors $|+\rangle$ and $|-\rangle$, as expected from Eq. (2.108).

You can also work out the matrix representation for $\hat{S}_x$ and $\hat{S}_y$, and verify that you can express them as $\mathbf{S} = (\hbar/2)\boldsymbol{\sigma}$ in terms of the so-called Pauli matrices

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,, \quad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\,, \quad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,. \tag{2.114}$$

As they should, the three components of $\mathbf{S}$ are Hermitian and satisfy the defining commutation relations Eq. (2.102). You can check that $\hat{S}_x$ and $\hat{S}_y$ also have $\pm\hbar/2$ as eigenvalues.

From the discussions in Sect. 1.4, we know that, if our (normalised) quantum state is expressed in terms of the eigenvectors of $\hat{S}_z$,

$$|\Psi\rangle = c_+|+\rangle + c_-|-\rangle\,, \tag{2.115}$$

then the probability of measuring $S_z = +\hbar/2$ will be $|c_+|^2$ and that of $S_z = -\hbar/2$ will be $|c_-|^2$, with of course $|c_+|^2 + |c_-|^2 = 1$ given that there are the only two possible outcomes of the measurement. Furthermore, we can also evaluate the possible outcomes of a measurement of S_x or S_y in the same quantum state: to answer this question we need to solve the eigenvalue equations for $\hat{S}_x$ or $\hat{S}_y$. As mentioned above, in **any direction** the only two possible outcomes of a measurement of the spin are $\pm\hbar/2$ (as it could not be otherwise, given that the choice of z axis is ultimately arbitrary!).

For example, in the case of $\hat{S}_x$ we find that the normalised eigenvectors are

$$|+_x\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}\left(|+_z\rangle + |-_z\rangle\right) , \tag{2.116}$$

$$|-_x\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}}\left(|+_z\rangle - |-_z\rangle\right) , \tag{2.117}$$

in terms of the eigenstates of $\hat{S}_z$. Note that we have added a subindex to indicate explicitely to which direction the various spin eigenstates refer to. By inverting this linear system we get

$$|+_z\rangle = \frac{1}{\sqrt{2}}\left(|+_x\rangle + |-_x\rangle\right) , \quad |-_z\rangle = \frac{1}{\sqrt{2}}\left(|+_x\rangle - |-_x\rangle\right) , \tag{2.118}$$

or also using the frequently used arrow notation

$$|\uparrow_z\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_x\rangle + |\downarrow_x\rangle\right) , \quad |\downarrow_z\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_x\rangle - |\downarrow_x\rangle\right) . \tag{2.119}$$

We can use these relations to express our original state vector $|\Psi\rangle$ in terms of the eigenstates of $\hat{S}_x$:

$$|\Psi\rangle = \left(c_+|+_z\rangle + c_-|-_z\rangle\right) = \left(\frac{c_+ + c_-}{\sqrt{2}}\right)|+_x\rangle + \left(\frac{c_+ - c_-}{\sqrt{2}}\right)|-_x\rangle . \tag{2.120}$$

and hence the probabilities of measuring S_x in the positive or negative directions will be, following the prescription of the generalised probability interpretation, given by

$$P(S_x = +\hbar/2) = \left|\left(\frac{c_+ + c_-}{\sqrt{2}}\right)\right|^2 \qquad P(S_x = -\hbar/2) = \left|\left(\frac{c_+ - c_-}{\sqrt{2}}\right)\right|^2 \tag{2.121}$$

which taking into account the normalisation of the state vector $|\Psi\rangle$ can be shown to satisfy $P(S_x = +\hbar/2) + P(S_x = -\hbar/2)/ = 1$, as it could not be otherwise.

The uncertainty principle for spin-1/2 particles

You may have noticed from the previous derivation an interesting particular case: if $|\Psi\rangle$ is a **determinate state** of S_z (meaning that either c_+ or c_i vanish) then $P(S_x = +\hbar/2) = P(S_x = -\hbar/2) = 1$: we have the same likelihood for the two possible outcomes of measuring S_x. In other words, our uncertainty about the outcome of a measurement of S_x in this state is maximal. Since S_x and S_z are incompatible observables, good knowledge of one of them implies a worsened knowledge about the other, recall that the generalised uncertainty principle applied to spin-1/2 gives

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

The addition of angular momenta. Of crucial relevance for many applications of quantum mechanics is the situation where we have a quantum system composed by **more than one particle** (as we will study in the next Chapter) and we need to evaluate the **total angular momentum** of the system. We must be careful here, since angular momentum is a *vector* and not a *scalar*, and a naive addition of angular momenta will in general yield the wrong result.

For the sake of the argument, let us consider here a quantum system composed by two particles with well-defined spin states $|s_1, m_1\rangle$ and $|s_2, m_2\rangle$. We denote the composite state[9] of the system by $|s_1\ s_2, m_1\ m_2\rangle$, and the action of the corresponding spin operators on it is given by

$$\begin{aligned}
\hat{S}_1^2|s_1\ s_2, m_1\ m_2\rangle &= s_1(s_1+1)\hbar^2|s_1\ s_2, m_1\ m_2\rangle\,,\\
\hat{S}_2^2|s_1\ s_2, m_1\ m_2\rangle &= s_2(s_2+1)\hbar^2|s_1\ s_2, m_1\ m_2\rangle\,,\\
\hat{S}_{1,z}|s_1\ s_2, m_1\ m_2\rangle &= m_1\hbar\,|s_1\ s_2, m_1\ m_2\rangle\,,\\
\hat{S}_{2,z}|s_1\ s_2, m_1\ m_2\rangle &= m_1\hbar\,|s_1\ s_2, m_1\ m_2\rangle\,.
\end{aligned} \tag{2.122}$$

We would like to determine what is the total angular momentum of this state, defined as

$$\mathbf{S} \equiv \mathbf{S}_1 + \mathbf{S}_2\,. \tag{2.123}$$

Concerning the z component of the total angular momentum, one can see that

$$\hat{S}_z|s_1\ s_2, m_1\ m_2\rangle = \left(\hat{S}_{1,z} + \hat{S}_{2,z}\right)|s_1\ s_2, m_1\ m_2\rangle = (m_1\hbar + m_2\hbar)\,|s_1\ s_2, m_1\ m_2\rangle\,, \tag{2.124}$$

and hence $m = m_1 + m_2$ is the value of the total angular momentum in the z direction, S_z, associated to the quantum state $|s_1\ s_2, m_1\ m_2\rangle$.

Determining the values that S^2 can take in this composite system is rather more complicated, and here we will work out the explicit derivation in the case of $s_1 = s_2 = 1/2$, that is, for a system composed by two spin-1/2 particles. Since each particle's z-component of the spin can be positive or negative ($m_1 = \pm 1/2$ and $m_2 = \pm - 1/2$) we will have four possible combinations, each with a **well-defined value of** $m = m_1 + m_2$, constructed as

$$\begin{aligned}
|\uparrow\uparrow\rangle &\equiv \left|\frac{1}{2}\,\frac{1}{2},\ \frac{1}{2}\,\frac{1}{2}\right\rangle \quad (m=1)\\
|\uparrow\downarrow\rangle &\equiv \left|\frac{1}{2}\,\frac{1}{2},\ \frac{1}{2}\,\frac{-1}{2}\right\rangle \quad (m=0)\\
|\downarrow\uparrow\rangle &\equiv \left|\frac{1}{2}\,\frac{1}{2},\ \frac{-1}{2}\,\frac{1}{2}\right\rangle \quad (m=0)\\
|\downarrow\downarrow\rangle &\equiv \left|\frac{1}{2}\,\frac{1}{2},\ \frac{-1}{2}\,\frac{-1}{2}\right\rangle \quad (m=-1)\,.
\end{aligned} \tag{2.125}$$

where now we adopt a notation $|\uparrow\uparrow\rangle$ where the first entry in the ket indicates m_1 and the second m_2. Inspection of Eq. (2.125) reveals something peculiar. We see that m varies from $m = +1$ to $m = -1$, and this suggest that the composite states have total spin $s = 1$. But there are two different combinations that lead to $m = 0$, which appears to be inconsistent (since a quantum state with well-defined values of $\hat{S}^2$ and

---

[9] Sensu strictu we should introduce here the **tensor product** to describe quantum systems consisting of multiple subsystems, each of which associated to a different Hilbert space. But we can come back to this subtlety on the subsequent chapters.

$\hat{S}_z$ is unique up to an overall normalisation). How to solve this problem?

The solution to this conundrum is subtle: only *three of the four combinations* defined in Eq. (2.125) actually correspond to $s = 1$ states, while the other combination instead has associated a total spin of $s = 0$, explaining why $m = 0$ was repeated (since for $s = 0$ the only option is $m = 0$). To identify which is the combination that has associated $s = 0$, we can use the spin lowering operator that we defined in Eq. (2.104). Recall that this operator exhibited the useful property that

$$\hat{S}_-|s,m\rangle = \hbar\sqrt{(s(s+1)-m(m-1))}\;|s,m-1\rangle\,, \tag{2.126}$$

so its action on a state with $s = m = 1$ will turn it into a $s = 1$ and $m = 0$ state:

$$\hat{S}_-|s=1,m=1\rangle = \sqrt{2}\hbar\;|s=1,m=0\rangle\,. \tag{2.127}$$

Certainly $\big|\uparrow\uparrow\big\rangle$ is a $s = m = 1$ state, so let me act on it with the lowering operator:

$$\begin{aligned}\hat{S}_-\big|\uparrow\uparrow\big\rangle &= \hat{S}_{1,-}\left|\frac{1}{2}\,\frac{1}{2},\;\frac{1}{2}\,\frac{1}{2}\right\rangle + \hat{S}_{2,-}\left|\frac{1}{2}\,\frac{1}{2},\;\frac{1}{2}\,\frac{1}{2}\right\rangle \\ &= \left|\frac{1}{2}\,\frac{1}{2},\;\frac{-1}{2}\,\frac{1}{2}\right\rangle + \left|\frac{1}{2}\,\frac{1}{2},\;\frac{1}{2}\,\frac{-1}{2}\right\rangle = \big|\downarrow\uparrow\big\rangle + \big|\uparrow\downarrow\big\rangle\,.\end{aligned} \tag{2.128}$$

where we have used Eq. (2.126) for the $s_1 = m_1 = 1/2$ and $s_2 = m_2 = 1/2$ states. Hence we conclude that $\hat{S}_-\big|\uparrow\uparrow\big\rangle = \big|\downarrow\uparrow\big\rangle + \big|\uparrow\downarrow\big\rangle$ is the combination with $s = 1$ and $m = 0$, and then the linearly independent combination $\big|\downarrow\uparrow\big\rangle - \big|\uparrow\downarrow\big\rangle$ has instead $s = 0$ and hence $m = 0$.

Putting all this information together, we see that we have three states with total spin $s = 1$ (and $m = \pm1, 0$, as required by the properties of angular momentum) which are known as the **triplet states**

$$\begin{aligned}|1\;1\rangle &\equiv \big|\uparrow\uparrow\big\rangle \quad (s=1,m=1)\,,\\ |1\;0\rangle &\equiv \frac{1}{\sqrt{2}}\left(\big|\uparrow\downarrow\big\rangle + \big|\downarrow\uparrow\big\rangle\right) \quad (s=1,m=0)\,,\\ |1{-1}\rangle &\equiv \big|\downarrow\downarrow\big\rangle \quad (s=1,m=-1)\,.\end{aligned} \tag{2.129}$$

Note that we have normalised the combination with $m = 0$. Then, the combination associated to a total spin of $s = 0$ (which hence can only have $m = 0$) is denoted as the **singlet state** and is given by:

$$|0\;0\rangle \equiv \frac{1}{\sqrt{2}}\left(\big|\uparrow\downarrow\big\rangle - \big|\downarrow\uparrow\big\rangle\right)\,, \tag{2.130}$$

which again has been normalised. This derivation illustrates that a system composed by two spin-1/2 particles can carry either a total spin of $s = 1$ (for the triplet states) or of $s = 0$ (for the singlet state), depending on the configuration of the individual spins. As a further piece of evidence, we can demonstrate that the triplet states are eigenvectors of $\hat{S}^2$ with eigenvalues of $2\hbar^2$ while the application of $\hat{S}^2$ destroys the singlet state (since the eigenvalue is zero). That is, if we have that

$$S^2 = (\mathbf{S}_1 + \mathbf{S}_2)\cdot(\mathbf{S}_1 + \mathbf{S}_2) = S_1^2 + S_2^2 + 2\mathbf{S}_1\cdot\mathbf{S}_2\,, \tag{2.131}$$

then we will find

$$\hat{S}^2|1\;0\rangle = 2\hbar^2|1\;0\rangle\,,\quad \hat{S}^2|0\;0\rangle = 0\,. \tag{2.132}$$

where note that $\hat{\mathbf{S}}_1$ and $\hat{\mathbf{S}}_2$ commute between then since they act on *different Hilbert spaces* (more about this in the next chapters).

Towards the spin-statistics theorem

You might notice an interesting feature of the triplet combinations defined in Eq. (2.129): they are *invariant (symmetric) upon the* **exchange of the two particles**. However, the singlet state is instead *antisymmetric* upon the exchange of the two particles, since

$$|0\ 0\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right) \quad \to \quad \frac{1}{\sqrt{2}}\left(|\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle\right) = -|0\ 0\rangle\,. \tag{2.133}$$

This observation will become crucial in the next Chapter, when we discuss how the spin of a particle determines which energy levels can be occupied and which ones cannot.

Generalisation to other spin combinations. The previous example represents the simplest possible example of spin combination, that involving two spin-1/2 particles. The general problem of combining a spin s_1 particle with a spin s_2 particle is rather more involved. The general answer is that the allowed combinations will have s between $s_1 + s_2$ and $|s_1 - s_2|$ in integer steps. We can give some examples:

- The case we have just discussed above corresponded to $s_1 = 1/2$ and $s_2 = 1/2$. In this case the general answer is that the allowed combinations range between $s = s_1 + s_2 = 1$ and $s = |s_1 - s_2| = 0$ in integer steps, hence indeed $s = 0, 1$ as we just found.

- Assume you want to combine the spin of a photon ($s_1 = 1$) with that of a proton ($s_2 = 1/2$). The allowed combinations range between $s = s_1 + s_2 = 3/2$ and $s = |s_2 - s_1| = 1/2$ with integer steps, hence $s = 3/2$ and $s = 1/2$.

- Likewise, if you were to combine the spin of a gravitino (a hypothetical spin-3/2 particle) with that of a graviton (the quantum of the gravitational force, a spin-2 particle), then the allowed combinations range between $s = s_1 + s_2 = 7/2$ and $s = |s_2 - s_1| = 1/2$ with integer steps, hence $s = 7/2, 5/2, 3/2, 1/2$ will be possible values of the total spin of the system.

In general, it can be shown that the allowed state vectors representing two-particle systems with spins s_1 and s_2 will be given by:

$$|s\ m\rangle = \sum_{m_1+m_2=m} C^{s_1 s_2 m}_{m_1 m_2 m} |s_1\ s_2, m_1\ m_2\rangle\,, \tag{2.134}$$

with $C^{s_1 s_2 m}_{m_1 m_2 m}$ being known as the Clebsch-Gordan coefficients. The derivation of the Clebsch-Gordan coefficients is mathematically very interesting but not particularly relevant for the physics of this course, so we will not cover it further. The interested student is encouraged to read a bit about **group theory**, the mathematical theory of symmetries.

Finally, as repeatedly emphasized, from the mathematical point of view spin follows exactly the same rules as orbital angular momentum. Hence, the same rules that we have discussed for the combination of the spin states of two different particles will also apply to the **combination of spin with orbital angular momentum**. For instance, for an electron of the hydrogen atom with quantum numbers n, ℓ, m its total angular momentum (spin plus orbital) will be either $\ell - 1/2$ or $\ell + 1/2$, following the general prescription.

Summary

We can now recapitulate what have we learned in this chapter concerning the the quantum mechanics of three-dimensional systems.

I/ We can solve the Schroedinger equation in three dimensions by the method of separation of variables; in the case of central potentials the dependence on the angular coordinates decouples from that of the radial coordinate.

II/ The solutions of the Schroedinger equation in three dimensions for the Coulomb potential are simultaneous eigenfunctions of $\hat{H}$, $\hat{L}^2$, and $\hat{L}_z$, with quantum states labelled by the eigenvalues of these operators.

III/ The properties of the radial solution $R(r)$ and the spherical harmonics $Y_l^m(\theta, \phi)$ determine the properties of the electronic orbitals of the hydrogen atom, such as for example what are the most likely values of the radial coordinate r or the angular shape of the different orbitals.

IV/ The quantum theory of angular momentum can be applied to spin, the intrinsic angular momentum of elementary particles, and we can also determine the total spin content of two-particle systems.

Quantum Mechanics 2

Dr Juan Rojo
VU Amsterdam and Nikhef Theory Group
http://www.juanrojo.com/ , **j.rojo@vu.nl**
Current version: **March 2, 2021**

3 Chapter 5: Identical Particles in Quantum Mechanics

Learning Goals

- To formulate and solve the Schroedinger equation in the case of **two-particle systems**.
- To determine the implications that in quantum mechanics we deal with systems of **genuinely indistinguishable particles**.
- To formulate the **Pauli exclusion principle** and to derive its consequences for the structure of multi-electron atoms.
- To explain fundamental properties about the general **structure of solids** by means of basic quantum-mechanical properties.
- To formulate the concept of **band structure** in solids and use it to classify materials in terms of their electrical conductivity properties.

In this section of the lecture notes we present the main concepts discussed in **Chapter 5 ("Identical Particles")** 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

- **5.1**: Two-Particle Systems.
- **5.2**: Atoms.
- **5.3**: Solids.

Up to now, in your study of quantum mechanics you have only considered **one-particle systems** (including the hydrogen atom, which can be expressed essentially as a one-particle problem). As you know from classical mechanics, in general the description of many-body systems ramps up quickly in difficulty: the two-body problem in classical mechanics can be solved exactly but already the three-body problem does not admit a closed solution. However, at the classical level, systems of **non-interacting particles** are trivial, in that their solution is the combination of independent solutions of the one-particle problem.

However, the situation is completely different in the case of quantum mechanics, and the state vectors for systems composed by more than one particle will in general be different from a trivial combination of one-particle states. The reason for this marked difference is that in quantum mechanics we often deal with systems of **indistinguishable particles**: an electron is identical to any other electron anywhere in the Universe.

This appears at first sight to be a rather minor point, but careful consideration of the implications that dealing with identical particles has in quantum mechanics has deep consequences, from explaining what makes a solid well, "solid", to the allowed options in which spins can be combined in many-particle states (which has impact in for instance in quantum technologies). And these implications are already there even for **non-interacting particles**: as we will see we can obtain many interesting results even neglecting the interactions between the particles in our system.

3.1 Two-particle systems

As mentioned above, up to now we have considered quantum systems composed by **a single particle**. For example, we have solved the Schroedinger equation in three dimensions for the hydrogen atom and found the wave functions of the electronic orbitals $\psi_{nlm}(r,\theta,\phi)$. Note that although in this case the system was composed by a proton and an electron, the proton did not play any role other than generating the electric potential thus was not part of the wave function of the system (technically, it did not play any **dynamical role** in the system).

We will consider now a quantum state composed by **two dynamical particles**. Its wave function will now depend on the position of the two particles, $\Psi(\mathbf{r}_1,\mathbf{r}_2,t)$ and its time evolution will be determined by the solution of the time-dependent Schroedinger equation

$$i\hbar\frac{\partial\Psi(\mathbf{r}_1,\mathbf{r}_2,t)}{\partial t} = \hat{H}\Psi(\mathbf{r}_1,\mathbf{r}_2,t) = \left(-\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V(\mathbf{r}_1,\mathbf{r}_2,t)\right)\Psi(\mathbf{r}_1,\mathbf{r}_2,t)\,, \tag{3.1}$$

where in general the two particles will have different masses m_1 and m_2, and the potential energy now depends on the position vectors of the two particles as well as on the time, $V(\mathbf{r}_1,\mathbf{r}_2,t)$. The generalised statistical interpretation tells us that the square of the wave function, $|\Psi(\mathbf{r}_1,\mathbf{r}_2,t)|^2$, will be normalised to one when integrating over all possible values of $\mathbf{r}_1$ and $\mathbf{r}_2$.

Here we will consider *time-independent potentials*, such that the time evolution of the eigenfunctions of the Hamiltonian is given as usual by

$$\Psi(\mathbf{r}_1,\mathbf{r}_2,t) = \psi(\mathbf{r}_1,\mathbf{r}_2)\exp\left(-\frac{iEt}{\hbar}\right)\,, \tag{3.2}$$

where the spatial component of the wave function satisfies the time-independent Schroedinger equation

$$\left(\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V(\mathbf{r}_1,\mathbf{r}_2)\right)\psi(\mathbf{r}_1,\mathbf{r}_2) = E\psi(\mathbf{r}_1,\mathbf{r}_2)\,, \tag{3.3}$$

in terms of the kinetic terms for the two particles and of the potential energy. Solving a two-body Schroedinger equation is in general much more difficult that its one-body counterparts, but in some cases we actually can reduce the solution of Eq. (3.3) to the solutions of separate one-body problems.

The simplest scenario for Eq. (3.3) is that of **non-interacting particles**, where each particle moves in some external potential but they do not interact among them. Note that here the meaning of "non-interacting" is *not* that we have two free particles, but that the two particles in the system interact only with an external potential but not among them. In this case, the potential energy simplifies to

$$V(\mathbf{r}_1, \mathbf{r}_2) = V_1(\mathbf{r}_1) + V_2(\mathbf{r}_2)\,, \tag{3.4}$$

and it is easy to see that Eq. (3.3) admits a solution where the total wave function **factorises** into two one-particle wave functions

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)\,, \tag{3.5}$$

where ψ_a and ψ_b indicate two different one-particle wavefunctions, each of which satisfies their a separate one-particle Schroedinger equation:

$$\begin{aligned} \left(-\frac{\hbar^2}{2m_1}\nabla_1^2 + V_1(\mathbf{r}_1)\right)\psi_a(\mathbf{r}_1) &= E_a\psi_a(\mathbf{r}_1)\,, \\ \left(-\frac{\hbar^2}{2m_2}\nabla_2^2 + V_2(\mathbf{r}_2)\right)\psi_b(\mathbf{r}_2) &= E_b\psi_b(\mathbf{r}_2)\,, \end{aligned} \tag{3.6}$$

with the total energy of the system being given by $E = E_a + E_b$. Note that the situation where we have **two free particles** is a particular case of this, and the equations that we need to solve are:

$$\begin{aligned} -\frac{\hbar^2}{2m_1}\nabla_1^2\psi_a(\mathbf{r}_1) &= E_1\psi_a(\mathbf{r}_1)\,, \\ -\frac{\hbar^2}{2m_2}\nabla_2^2\psi_b(\mathbf{r}_2) &= E_2\psi_b(\mathbf{r}_2)\,. \end{aligned} \tag{3.7}$$

Note, however. that if we know the solutions of the one-particle equations, Eq. (3.6), then the case of two particles interacting in some external potential falls on the same complexity class as that of a system composed by two free particles, and it is *not* more difficult to solve.

Entanglement in two-particle systems

For non-interacting particles, the two-body wave function factorises into two one-body wave functions of the form

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)\,,$$

which are eigenvalues of the total Hamiltonian $\hat{H}$ with energy $E = E_a + E_b$. However, this does not imply that all valid quantum states of this two particle system can be described by such factorisable state vectors. For example, if ψ_c and ψ_d are two other one-particle solutions, surely I can have

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \sqrt{\frac{1}{3}}\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + \sqrt{\frac{2}{3}}\psi_c(\mathbf{r}_1)\psi_d(\mathbf{r}_2) \tag{3.8}$$

which **cannot be expressed as a product of one-particle wavefunctions**. These states are called **entangled states** and are of huge importance in quantum mechanics.

Another situation that admits a relatively compact analytic treatment is that of **central potentials**, such as the Coulomb potential that we discussed in the case of the hydrogen atom. In this case, the interaction between the two particles depends only on their **relative distance** but not on their relative orientation

$$V(\mathbf{r}_1, \mathbf{r}_2) = V(|\mathbf{r}_1 - \mathbf{r}_2|)\,. \tag{3.9}$$

Then the two-body Schroedinger equation

$$\left(\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V(|\mathbf{r}_1 - \mathbf{r}_2|) \right) \psi(\mathbf{r}_1, \mathbf{r}_2) = E\psi(\mathbf{r}_1, \mathbf{r}_2)\,, \tag{3.10}$$

can be factorised into **two separate one-body equations**, one for the motion around the center of mass of the system and the other for the free motion of the center of mass.

Thus we have found that there are two situations where solving a two-particle Schroedinger equation is relatively feasible:

- When the two particles **do not interact with each other**, but only with some external potential.
- When the two particles **interact with each other via a central potential**, but then do not interact with any external potential.

In both cases, we can reduce the two-body system to two one-particle problems that can be solved separately. However, in general the two particles will experience **both** interactions with some external force and mutual interactions between them. An important example in this context is the helium atom, a system composed by an atomic nucleus with two protons $Z = 2$ and then two electrons orbiting around it. The potential experienced by the electrons receives contributions both from the electrical attraction with the positively charged nucleus and from the negative repulsion between the two electrons:

$$V(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{4\pi\epsilon_0}\left(-\frac{2e^2}{|\mathbf{r}_1|} - \frac{2e^2}{|\mathbf{r}_2} + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right)\,. \tag{3.11}$$

The Schroedinger equation for the Helium atom cannot be solved analytically in closed form. We will need to introduce some approximations such as perturbation theory and the variational principle, which we will introduce later in the course.

Now, as we discuss next, even for the cases where we can solve exactly the Schroedinger equation for the two body problem (say for non-interacting particles), there remain some important issues to be addressed in the case that the two particles that compose the system are **indistinguishable**. To illustrate this point, when we wrote the factorised solution for the total wave function of a system composed by two non-interacting particles we had

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)\,, \tag{3.12}$$

we implicitely assumed that we could **distinguish the two particles**, since we assume that the first particle occupies the one-particle state ψ_a and the second ψ_b. But this is not possible if I can tell apart the two particles! We will see next how to quantum mechanics solves this conundrum.

Indistinguishable particles: fermions and bosons Let us assume that we have now two noninteracting particles, the first one occupying the one-particle state $\psi_a(\mathbf{r})$ and the second one in another one-particle state $\psi_b(\mathbf{r})$.[10] As demonstrated above, the total wave function of the two-particle system will be the product of

---

[10] While here I work with position-representation wavefunctions, the concept is fully general and can also be formulated in terms of the Dirac notation.

the two one-particle states:

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) \, . \tag{3.13}$$

Now, by writing the total wave function in this manner I am assuming that somehow we can **tell apart** the "first particle" from "the second particle" (that is, that we can stick labels on them). But in quantum mechanics **particles are indistinguishable**, not because of some technical limitation of our measurement apparatus, but because **every electron is identical to every other electron**, and every hydrogen atom is identical (and will always be) identical to any other hydrogen atom. In other words, in quantum mechanics it is not possible to say "this electron" or "that electron", but only "**a** electron".

To take into account that quantum mechanics deals with particles that are indistinguishable, we need to modify the two-body wavefunction for **identical particles** in a way that it does not implicitely assume that we can tell particles apart. This can de done in two ways. First, we can write

$$\psi_+(\mathbf{r}_1, \mathbf{r}_2) = A\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + \psi_b(\mathbf{r}_1)\psi_a(\mathbf{r}_2)\right] \, . \tag{3.14}$$

where the only thing that we assume is that there is "a particle" in orbital ψ_a and another of the same particles in orbital ψ_b, and where A is an overall normalisation constant. Another possible option that achieves the same goal is

$$\psi_-(\mathbf{r}_1, \mathbf{r}_2) = A\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) - \psi_b(\mathbf{r}_1)\psi_a(\mathbf{r}_2)\right] \, . \tag{3.15}$$

with the minus sign between the two terms.[11]

Actually, given a specific type of particle, only one of these two relations will be valid. If we assume normalised one-particle states, then we can have only one of the following two situations:

- We denote as **bosons** as those elementary particles for which their two-body wave function is

  $$\frac{1}{\sqrt{2}}\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + \psi_b(\mathbf{r}_1)\psi_a(\mathbf{r}_2)\right] \, .$$

  Bosons have the defining property that their total wave function is **symmetric** under the exchange of the two particles: $\psi_+(\mathbf{r}_1, \mathbf{r}_2) = \psi_+(\mathbf{r}_2, \mathbf{r}_1)$. All particles with **integer spin** are bosons (for example the photon, with $s = 1$ or the Higgs boson, with $s = 0$).

- We denote as **fermions** those elementary particles for which their two-body wave function is

  $$\frac{1}{\sqrt{2}}\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + -\psi_b(\mathbf{r}_1)\psi_a(\mathbf{r}_2)\right] \, .$$

  Fermions have the defining property that their total wave function is **antisymmetric** under the exchange of the two particles : $\psi_-(\mathbf{r}_1, \mathbf{r}_2) = -\psi_-(\mathbf{r}_2, \mathbf{r}_1)$. All particles with **half-integer spin** are fermions (for example the electron and the proton, both with $s = 1/2$).

Here we take this property as an axiom of **quantum mechanics**. The formal connection between the spin of a particle and the symmetry properties of the total wave function under the exchange of two particles can be demonstrated in quantum field theory and is known as the **spin-statistics theorem**. Here the situation is similar as for the Heisenberg uncertainty principle, which despite its name was not a fundamental principle but rather a consequence of the more fundamental axioms of quantum mechanics.

---

[11] Actually, there are cases where one can have also a *complex phase* between the two terms, leading to exotic quasi-particles such as anyons, but we will not discuss this case further in this course.

The Pauli exclusion principle

A direct consequence of the of the **spin-statistics axiom** of quantum mechanics is that **two fermions cannot occupy the same quantum state**. To see this, consider two non-interacting fermions that occupy the same one-particle quantum state ψ_a. The combined two particle wave function would then be:

$$\psi_-(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}\left[\psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2) - \psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2)\right] = 0\,. \tag{3.16}$$

hence only when the two one-particle states are different from each other the two-particle wave function of a fermion system will be non-zero. This central result of quantum theory is known as the **Pauli exclusion principle** and as we will see later in this chapter has crucial implications for the structure of matter.

The infinite square well with two particles. In order to illustrate how differences between fermions and bosons can have very significative consequences for the properties of quantum systems, let us consider the following system. Assume that we have **two noninteracting particles** with the same mass confined in the infinite square well potential in one dimension. For simplicity, let's work in units where the width of the well is $L = 1$ and that the particle mass is $m = 1$. The potential energy for this system is then

$$\begin{aligned} V(x) &= 0 \quad 0 < x < 1\,, \\ V(x) &= \infty \quad x \le 0\,, x \ge 1\,. \end{aligned}$$

The one-particle states and energies of this potential are are quantised in terms of a single principle quantum number n and are given by

$$\psi_n(x) = \sqrt{2}\sin(n\pi x)\,, \qquad E_n = \frac{n^2\pi^2\hbar^2}{2}\,, \qquad n = 1, 2, 3, \ldots\,. \tag{3.17}$$

Let us consider three different scenarios for this system, and as we will see whether or not particles are distinguishable or indistinguishable, and in the latter case whether they are fermions or bosons, has important consequences for the resulting energy spectrum of this system.

- The two particles are **distinguishable**, that is, they are **not identical**.

  If we call n_1 the energy level occupied by the first particle and n_2 that by the second particle, the total wave function is just the regular product of one-particle wave functions

  $$\psi_{n_1 n_2}(x) = \psi_{n_1}(x_1)\psi_{n_2}(x_2)\,, \qquad E_{n_1 n_2} = E_{n_1} + E_{n_2} = \frac{(n_1^2 + n_2^2)\pi^2\hbar^2}{2} \equiv (n_1^2 + n_2^2)K\,. \tag{3.18}$$

  For example, the ground state is $\psi_{11}(x)$ with energy $E_{11} = 2K$ and the first excited state is double-degenerate, with two possible two-particle states having associated the same total energy

  $$\begin{aligned} \psi_{21}(x_1, x_2) &= 2\sin(2\pi x_1)\sin(\pi x_2)\,, \qquad E_{21} = 5K\,, \\ \psi_{12}(x_1, x_2) &= 2\sin(\pi x_1)\sin(2\pi x_2)\,, \qquad E_{12} = 5K\,. \end{aligned} \tag{3.19}$$

  By varying n_1 and n_2, we can then construct all the possible two-particle wave functions allowed for this system.

Note that each two-particle state is **uniquely specified** by the choice of the two one-particle quantum numbers n_1 and n_2. In this respect, the two-particle states are trivial combinations of one-particle states, much in the same way as what would happen in classical physics for non-interacting particles.

- The two particles are **indistinguishable**, that is, they are **identical**, and specifically they are **bosons**.[12]

  In this case, the ground state will be unchanged, since clearly the two-particle wave function

  $$\psi_{11}(x_1,x_2) = 2\sin(\pi x_1)\sin(\pi x_2)\,, \tag{3.20}$$

  is **symmetric** under the exchange of the two particles:

  $$\psi_{11}(x_2,x_1) = 2\sin(\pi x_2)\sin(\pi x_1) = \psi_{11}(x_1,x_2)\,, \tag{3.21}$$

  as should be the case for the wave function of a system composed by bosons.

  However, you also notice that the (degenerate) first excited states in Eq. (3.19) are *not symmetric* under the exchange of the two particles. For example, if I take ψ_{21} and I exchange the particles

  $$\psi_{21}(x_2,x_1) = 2\sin(2\pi x_2)\sin(\pi x_1) = \psi_{12}(x_1,x_2) \neq \psi_{21}(x_1,x_2)\,, \tag{3.22}$$

  and hence $\psi_{21}(x_2,x_1)$ and $\psi_{12}(x_2,x_1)$ constructed as in Eq. (3.19) are not acceptable for a system composed by two identical bosons.

  Instead, we must symmetrise the two-particle wave function corresponding to the first excited state of the system by using the appropriate prescription for bosons. In this way we find that the wave function of the first excited state is given by

  $$\psi(x_1,x_2) = \sqrt{2}\left[\sin(2\pi x_1)\sin(\pi x_2) + \sin(\pi x_1)\sin(2\pi x_2)\right]\,, \qquad E = 5K\,, \tag{3.23}$$

  which is obviously invariant under the exchange of the two particles, $\psi(x_2,x_1) = \psi(x_1,x_2)$.

  This result illustrates a first important consequence of the quantum statistics: the first excited state, which is degenerate in the case of distinguishable particles, becomes **non-degenerate** in the case of indistinguishable bosons. That is, quantum statistics breaks the degeneracy which is present in the case of distinguishable particles.

- The two particles are **indistinguishable**, that is, they are **identical**, and specifically they are **fermions**.

  In this case the two-particle state which corresponding to the ground state of the previous two cases is not allowed due to the **Pauli exclusion principle**: two fermions cannot occupy the same one-particle quantum state. In this case the wave function of the ground state (the state with the lowest energy) is

  $$\psi(x_1,x_2) = \sqrt{2}\left[\sin(2\pi x_1)\sin(\pi x_2) - \sin(\pi x_1)\sin(2\pi x_2)\right]\,, \qquad E = 5K\,, \tag{3.24}$$

  which as requested for fermion systems is **antisymmetric** under the exchange of any two particles, given that $\psi(x_1,x_2) = -\psi(x_2,x_1)$.

  Here we note another very powerful consequence of quantum statistics: as compared to the case with distinguishable particles, or to the case with indistinguishable bosons, in the case of fermions the

---

[12] Note that the following discussion requires the two particles to be identical bosons, not *only* bosons. For example, a system composed by two *distinct* spin-1 particles can still be treated as in the previous case, since here we do not deal with genuinely indistinguishable particles.

energy of the ground state is increased. In other words, keeping everything the same (mass, potential etc) the **zero-point energy** of this quantum system is increases by a factor 2.5 just from replacing bosons by fermions, which is quite a remarkable result.

Hence, to summarize we have illustrated with this relatively simple example some important consequences of accounting for the fact that quantum mechanics deals with indistinguishable particles: as compared to system composed by distinguishable particles, we can

- Change the energy and eigenfunctions spectrum of the system by **removing degeneracies**.
- Change the energy and eigenfunctions spectrum of the system by **increasing the zero-point energy of the ground state**.

In the rest of this chapter, we will study other implications that systems of indistinguishable particles have in quantum mechanics.

Exchange forces. The previous example demonstrates that the quantum statistics that apply in systems composed by identical particles can lead to significance consequences in the properties of quantum states. Let us investigate a bit more what is the underlying reason for this behaviour.

Consider now a one-dimensional system composed by two particles, and consider two one-particle quantum states $\psi_a(x)$ and $\psi_b(x)$ (orthonormal between them). Following the previous discussion, the two-particle wave function will be given in the various relevant cases by:

- **Distinguishable particles**: $\psi(x_1, x_2) = \psi_a(x_1)\psi_b(x_2)$.
- **Indistinguishable particles (bosons)**: $\psi_+(x_1, x_2) = \frac{1}{\sqrt{2}}\left(\psi_a(x_1)\psi_b(x_2) + \psi_b(x_1)\psi_a(x_2)\right)$.
- **Indistinguishable particles (fermions)**: $\psi_-(x_1, x_2) = \frac{1}{\sqrt{2}}\left(\psi_a(x_1)\psi_b(x_2) - \psi_b(x_1)\psi_a(x_2)\right)$.

Let us now evaluate the expectation value of the **square of the separation between the two particles**:

$$\left\langle (x_1 - x_2)^2 \right\rangle = \left\langle x_1^2 \right\rangle + \left\langle x_2^2 \right\rangle - 2\left\langle x_1 x_2 \right\rangle \,, \tag{3.25}$$

in the three cases under consideration, and compare their results. The reason to choose this quantity is that, as we will demonstrate, quantum statistics leads to an effect on the average distance between the two particles which is akin to either a repulsive or attractive force, known as **exchange forces**. However, this term is a bit of a misnomer since the particles are **non-interacting**: it will *look like* there is a force acting upon them, but it is just a geometrical consequence of how wave functions of identical particles must be constructed in quantum mechanics.

- **Distinguishable particles**: in this case the relevant expectation values are

  $$\left\langle x_1^2 \right\rangle = \int dx_1 \int dx_2\, x_1 |\psi(x_1, x_2)|^2 = \int dx_1 x_1 |\psi_a(x_1)|^2 \int dx_2 |\psi_b(x_2)|^2 = \left\langle x^2 \right\rangle_a \,, \tag{3.26}$$

  where we have used that the one-particle wave functions are appropriately normalised. The result is then the expectation value of the position squared in the one-particle state $\psi_a(x)$. Using a similar strategy, we can easily evaluate the other matrix elements to find:

  $$\left\langle x_2^2 \right\rangle = \left\langle x^2 \right\rangle_b \,, \qquad \left\langle x_1 x_2 \right\rangle = \left\langle x \right\rangle_a \left\langle x \right\rangle_b \,, \tag{3.27}$$

and hence the sought-for expectation value for the square of the distance between the two particles in the system is given by

$$\left\langle (x_1-x_2)^2\right\rangle = \left\langle x^2\right\rangle_a + \left\langle x^2\right\rangle_b - 2\left\langle x\right\rangle_a \left\langle x\right\rangle_b \, . \tag{3.28}$$

Note how this result is consistent with the particles being distinguishable: $\left\langle (x_1-x_2)^2\right\rangle$ is determined by combining matrix elements that only involve *one-particle states*: the two particles **don't talk to each other at all** (as should be the case, since these are non-interacting particles).

Let's compare this result with the other two cases, corresponding now to indistinguishable particles.

- **Indistinguishable particles**: now the same calculation yields

$$\begin{aligned}
\left\langle x_1^2\right\rangle &= \int dx_1 \int dx_2\, x_1^2 |\psi_\pm(x_1,x_2)|^2 = \frac{1}{2}\int dx_1 \int dx_2\, x_1^2 |\left(\psi_a(x_1)\psi_b(x_2) \pm \psi_b(x_1)\psi_a(x_2)\right)|^2 \\
&= \frac{1}{2}\int dx_1 \int dx_2\, x_1^2 \Bigg(|\psi_a(x_1)\psi_b(x_2)|^2 + |\psi_b(x_1)\psi_a(x_2)|^2 \\
&\pm \psi_a^*(x_1)\psi_b^*(x_2)\psi_b(x_1)\psi_a(x_2) \pm \psi_b^*(x_1)\psi_a^*(x_2)\psi_a(x_1)\psi_b(x_2)\Bigg) \\
&= \frac{1}{2}\int dx_1 x_1^2 |\psi_a(x_1)|^2 \int dx_2 |\psi_b(x_2)|^2 + \frac{1}{2}\int dx_1 x_1^2 |\psi_b(x_1)|^2 \int dx_2 |\psi_a(x_2)|^2 \\
&\pm \frac{1}{2}\int dx_1 x_1^2 \psi_a^*(x_1)\psi_b(x_1) \int dx_2 \psi_b^*(x_2)\psi_a(x_2) \\
&\pm \frac{1}{2}\int dx_1 x_1^2 \psi_b^*(x_1)\psi_a(x_1) \int dx_2 \psi_a^*(x_2)\psi_b(x_2) = \frac{1}{2}\left(\left\langle x^2\right\rangle_a + \left\langle x^2\right\rangle_b\right) \, ,
\end{aligned} \tag{3.29}$$

where we have used that the one-particle states are normalised and *orthogonal* among them. Clearly, we will get the same result for $\left\langle x_2^2\right\rangle$, since the two particles are indistinguishable. Now for the third term in the calculation, it is when things become interesting:

$$\begin{aligned}
\left\langle x_1 x_2\right\rangle &= \int dx_1 \int dx_2\, x_1 x_2 |\psi_\pm(x_1,x_2)|^2 \\
&= \frac{1}{2}\int dx_1 \int dx_2\, x_1 x_2 |\left(\psi_a(x_1)\psi_b(x_2) \pm \psi_b(x_1)\psi_a(x_2)\right)|^2 \\
&= \frac{1}{2}\int dx_1 \int dx_2\, x_1 x_2 \Bigg(|\psi_a(x_1)\psi_b(x_2)|^2 + |\psi_b(x_1)\psi_a(x_2)|^2 \\
&\pm \psi_a^*(x_1)\psi_b^*(x_2)\psi_b(x_1)\psi_a(x_2) \pm \psi_b^*(x_1)\psi_a^*(x_2)\psi_a(x_1)\psi_b(x_2)\Bigg)
\end{aligned}$$

which by evaluating each of the individual integrals gives

$$\begin{aligned}
\left\langle x_1 x_2\right\rangle &= \frac{1}{2}\int dx_1\, x_1 |\psi_a(x_1)|^2 \int dx_2\, x_2 |\psi_b(x_2)|^2 \\
&+ \frac{1}{2}\int dx_1\, x_1 |\psi_b(x_1)|^2 \int dx_2\, x_2 |\psi_a(x_2)|^2 \\
&\pm \frac{1}{2}\int dx_1\, x_1 \psi_a^*(x_1)\psi_b(x_1) \int dx_2\; x_2 \psi_b^*(x_2)\psi_a(x_2) \\
&\pm \frac{1}{2}\int dx_1\, x_1 \psi_b^*(x_1)\psi_a(x_1) \int dx_2\, x_2 \psi_a^*(x_2)\psi_b(x_2) \\
&= \frac{1}{2}\left(2\left\langle x\right\rangle_a \left\langle x\right\rangle_b \pm \left\langle x\right\rangle_{ab}\left\langle x\right\rangle_{ba} \pm \left\langle x\right\rangle_{ba}\left\langle x\right\rangle_{ab}\right) = \left\langle x\right\rangle_a \left\langle x\right\rangle_b \pm |\left\langle x\right\rangle_{ab}|^2 \, ,
\end{aligned} \tag{3.30}$$

in terms of the **overlap integral** between the two one-particle states defined as

$$\langle x\rangle_{ab} \equiv \int dx\; x\; \psi_a^*(x)\psi_b(x)\,. \tag{3.31}$$

Note that this overlap integral arises only when we have indistinguishable particles: it it a consequence of the cross-terms that we get when we square the (anti-)symmetric two-particle wave function $\psi_\pm(x_1,x_2)$ (so it is a direct consequence of the quantum statistics requirements).

Putting together the three terms, we obtain the counterpart of Eq. (3.28) for indistinguishable particles:

$$\left\langle (x_1-x_2)^2\right\rangle = \left\langle x^2\right\rangle_a + \left\langle x^2\right\rangle_b - 2\left\langle x\right\rangle_a\left\langle x\right\rangle_b \mp 2|\left\langle x\right\rangle_{ab}|^2\,, \tag{3.32}$$

which is different from the case of the system composed by two distinguishable particles, and furthermore the result varies if we consider a system of bosons and a system of fermions.

Combining these results together, we therefore obtain that the distinguishability or not of the particles that compose a two-particle quantum system has a direct physical consequence in the expectation value of the square of their average spatial distance:

$$\left\langle (\Delta x)^2\right\rangle_\pm = \left\langle (\Delta x)^2\right\rangle_d \mp |\left\langle x\right\rangle_{ab}|^2\,, \tag{3.33}$$

which is different for bosons than for fermions:

- For *identical bosons*, on average the two particles will tend to be **closer together** as compared to the case of two distinguishable particles.
- For *identical fermions*, on average the two particles will tend to be **farther apart** as compared to the case of two distinguishable particles.

The exchange force for identical particles

The result of Eq. (3.33) indicates that a system of two identical bosons behaves as if there was an **attractive force** between then, while a system of two identical fermions behaves on the contrary as if there was a **repulsive force** between then. This putative force is often called **exchange force**, which is not really accurate since there is no physical interactions between the two particles (the system is non-interacting!), it is a consequence of the symmetry requirements on the two-particle state vectors of indistinguishable particles without classical counterpart.

We note that there is an interesting limit of the above result which corresponds to those configurations where the overlap integral vanishes (or becomes infinitesimally small):

$$\langle x\rangle_{ab} \equiv \int dx\; x\; \psi_a^*(x)\psi_b(x) \to 0\,. \tag{3.34}$$

For instance, this would be the case in situations where the one-particle states $\psi_a(x)$ and $\psi_b(x)$ are only non-zero in different regions of x, like when two electrons are really far from each other. This implies that systems of two identical particles for which the overlap integral Eq. (3.31) can be treated **as if they were distinguishable**: it makes full sense to tell two electrons apart if they are separated by kilometers (but not if the separation is of a few nanometers!).

The electron spin reloaded. In Sect. 2.5 we discussed how to combine the spin of two particles and to determine the possible options for the total spin S and its z-component S_z of the system. There we saw that we can combine the spin of two spin-1/2 fermions (such as electrons) into four states, three of them characterised by a total spin $s = 1$,

$$\begin{aligned} |1\ 1\rangle &= \left|\uparrow\uparrow\right\rangle \quad (s=1, m=1)\,, \\ |1\ 0\rangle &= \frac{1}{\sqrt{2}}\left(\left|\uparrow\downarrow\right\rangle + \left|\downarrow\uparrow\right\rangle\right) \quad (s=1, m=0), \\ |1{-1}\rangle &= \left|\downarrow\downarrow\right\rangle \quad (s=1, m=-1)\,. \end{aligned} \tag{3.35}$$

which are known as the **triplet states**, and another with vanishing total spin $s = 0$, the **singlet state**:

$$|0\ 0\rangle = \frac{1}{\sqrt{2}}\left(\left|\uparrow\downarrow\right\rangle - \left|\downarrow\uparrow\right\rangle\right)\,.$$

As we discussed ,the triplet states are *symmetric* under the exchange of two particles, while the singlet one is instead **antisymmetric**: if you exchange the two electrons you get

$$\frac{1}{\sqrt{2}}\left(\left|\downarrow\uparrow\right\rangle - \left|\uparrow\downarrow\right\rangle\right) = -\frac{1}{\sqrt{2}}\left(\left|\uparrow\downarrow\right\rangle - \left|\downarrow\uparrow\right\rangle\right) = -|0\ 0\rangle\,. \tag{3.36}$$

What is the connection between the spin combinations of a two-particle system and the general requirement that the wave function of a system of two fermions must be antisymmetric upon their exchange? Well, for a two electron system, the **total wave function** is the product of the spatial and the spin parts of the wave function. For example, we could have

$$\psi(\mathbf{r}_1, \mathbf{r}_2)|s\ m\rangle\,, \tag{3.37}$$

with the spin part $|s\ m\rangle$ being either one of the three triplet states or the singlet state defined above.[13] In such a case, the antisymmetry requirement of the total wave function of the two-electron system then implies:

- If the two electrons occupy the *same position state*, one has that $\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi(\mathbf{r}_2, \mathbf{r}_1)$, so then the spin state must the antisymmetric *singlet state* to ensure that the total wave function is antisymmetric upon the exchange of the two particles.
- If the two electrons have associated a *symmetric singlet state*, then the spatial part of the wave function will be non-zero only if it is antisymmetric $\psi(\mathbf{r}_1, \mathbf{r}_2) = -\psi(\mathbf{r}_2, \mathbf{r}_1)$ to ensure that the total wave function is antisymmetric upon the exchange of the two particles.

The Pauli exclusion principle for electrons

We have thus demonstrated that **two electrons can occupy the same position state** (say, the same electronic orbital ψ_{nlm} in the hydrogen atom) only if their spin state is the **singlet state**, that is, only if the z-component of their spins points in opposite directions. This is why in a given non-degenerate quantum state we can only accomodate two fermions: provided their spins are in opposite direction, this is the only configuration that ensures an antisymmetric total wave function.

---

[13] At least, if we want tout state to have well defined s and m values: as you know, any linear superposition of valid spin states represents also a *bona fide* state of the system.

The generalised symmetrisation principle. In the previous discussion we have restricted ourselves to systems of two non-interacting particles. However, the **generalised symmetrisation principle** of quantum mechanics has associated a much stronger requirement: a system of identical particles must be either **symmetric** (for bosons) or **antisymmetric** (for fermions) upon the exchange of any of the particles that compose it.

Let us denote the state vector of a general two-particle system as $|(1,2)\rangle$. Then we define the **exchange operator** $\hat{P}$ as follows

$$\hat{P}|(1,2)\rangle = |(2,1)\rangle \, . \tag{3.38}$$

The square of this operator is the identity operator $\mathbb{1}$, given that

$$\hat{P}^2|(1,2)\rangle = \hat{P}\left(\hat{P}|(1,2)\rangle\right) = \hat{P}|(2,1)\rangle = |(1,2)\rangle \, , \tag{3.39}$$

from where it follows that the eigenvalues of $\hat{P}$ are ±1. If two particles are identical, clearly the Hamiltonian will be invariant under the exchange $\mathbf{r}_1 \to \mathbf{r}_2$ and therefore $\hat{P}$ and $\hat{H}$ commute and represent compatible observables:

$$\left[\hat{H}, \hat{P}\right] = 0 \, . \tag{3.40}$$

Now, recall the **generalised Ehrenfest theorem** that determined the time evolution of an observable in terms of its commutation relations with the Hamiltonian operator:

$$\frac{d}{dt}\langle \mathcal{O}\rangle = \frac{i}{\hbar}\left\langle \left[\hat{H}, \widehat{\mathcal{O}}\right]\right\rangle + \left\langle \frac{\partial \widehat{\mathcal{O}}}{\partial t}\right\rangle \, , \tag{3.41}$$

This relation applied to the case of the exchange operator yields

$$\frac{d}{dt}\langle P\rangle = 0 \tag{3.42}$$

implying that if a quantum system starts in one of the eigenstates of $\hat{P}$ (that is, it is either symmetric $\langle P\rangle = 1$ or antisymmetric $\langle P\rangle = -1$ upon the exchange of the two particles), it will remain this way forever. In other words, the **symmetry properties** of a quantum state upon the exchange of two particles are **time-invariant**.

The **generalised symmetrisation axiom** of quantum mechanics tells us that these properties are the same for n-particle system composed by indistinguishable fermions or bosons. That is, only states which satisfy

$$\hat{P}_{ij}|(1,2,\ldots,i,\ldots,j,\ldots,n)\rangle = |(1,2,\ldots,j,\ldots,i,\ldots,n)\rangle = \pm|(1,2,\ldots,i,\ldots,j,\ldots,n)\rangle \, , \tag{3.43}$$

are **allowed** in quantum mechanics, where $\hat{P}_{ij}$ is the exchange operator applied to particles i and j of the system, and the positive sign holds for bosons and the negative one for fermions.

As mentioned above, actually this generalised symmetrisation requirement is not an axiom but a consequences of the principles of relativistic quantum mechanics (the **spin-statistics theorem**), though this difference will not have any practical implication for the discussions that we will have in the rest of the course.

3.2 Considerations for multi-electron atoms

Equipped with a better understanding of the quantum mechanical consequences and requirements of systems composed by identical (indistinguishable) particles, we can extend our discussion of the hydrogen atoms to **atoms composed by more than one electron**. The starting point is the Hamiltonian operator $\hat{H}$ for a neutral atom with Z protons, $A-Z$ neutrons, and Z electrons, which is given by

$$\hat{H}=\sum_{j=1}^{Z}\left[\frac{\hbar^2}{2m}\nabla_j^2-\left(\frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r_j}\right)\right]+\frac{1}{2}\sum_{j\neq k}^{Z}\frac{e^2}{|\mathbf{r}_j-\mathbf{r}_k|}\,, \tag{3.44}$$

which is composed by the kinetic energy term for each electron, the attractive potential between each electron and the positive charge $+Ze$ of the atomic nucleus, and the **electric pair-wise repulsion** between electrons. Note that the factor $1/2$ in the repulsive term is introduced to avoid double counting. Except for the hydrogen case ($Z=1$), this Hamiltonian is not solvable unless we introduce some approximations such as perturbation theory, which we will do later in the course.

The simplest possible multi-electron atom is helium ($Z=2$), for which the above Hamiltonian reduces to the following operator:

$$\hat{H}=\sum_{j=1}^{2}\left[\frac{\hbar^2}{2m}\nabla_j^2-\left(\frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r_j}\right)\right]+\frac{e^2}{|\mathbf{r}_1-\mathbf{r}_2|}\,, \tag{3.45}$$

The most drastic approximation which allows us to solve the corresponding Schroedinger equation is to completely **neglect the electron repulsion**. In this case, the system reduces to two non-interacting particles (since the electrons now experience only the attractive potential from the atomic nucleus) and as we know from the previous discussion the total two-particle wave function can now be computed exactly in terms of the one-particle wave functions.

With this approximation, the Hamiltonian of the helium atom simplifies to

$$\hat{H}=\sum_{j=1}^{2}\left[\frac{\hbar^2}{2m}\nabla_j^2-\left(\frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r_j}\right)\right]\,, \tag{3.46}$$

which has the same form of the Hamiltonian for a system of non-interacting particles. Therefore, we know that in this crude approximation the Helium wave function can be **factorised** into the product of the one-particle wavefunctions of a **hydrogen-like atom**, such that

$$\psi(\mathbf{r}_1,\mathbf{r}_2)=\psi_{nlm}(\mathbf{r}_1)\psi_{n'l'm'}(\mathbf{r}_2)\,, \tag{3.47}$$

where note that in general the quantum numbers of the two one-particle wave functions will be different.

Entangled states

However, this is certainly not the only option for the possible electronic wave functions of the Helium atom: we can have also linear combinations of the one-body eigenstates of the form

$$\psi(\mathbf{r}_1,\mathbf{r}_2)=\frac{3}{5}\psi_{nlm}(\mathbf{r}_1)\psi_{n'l'm'}(\mathbf{r}_2)+\frac{4}{5}\psi_{\tilde{n}\tilde{l}\tilde{m}}(\mathbf{r}_1)\psi_{\tilde{n}'\tilde{l}'\tilde{m}'}(\mathbf{r}_2)\,. \tag{3.48}$$

Note that these will be **entangled states** in that they cannot be written as a combination of specific one-particle states.

In these entangled states, the two particles are intrinsically, well, entangled: if I measure the energy (or the angular momentum) of particle 1, I will know for sure what will be the energy (or the angular momentum) of particle 2 **without measuring it**, even if the two particles are separated by a very large spatial distance.

Since the energies of hydrogen-like atoms with Z protons scale as $Z^2 E_n^H$ in terms of the energy levels of the hydrogen atom, recall Eq. (2.72), we know that in this approximation the total energy of the Helium atom system will be specified by the two principal quantum numbers n and n' and given by

$$E_{n,n'} = 4\left(E_n + E_{n'}\right) , \tag{3.49}$$

in terms of the energies E_n of the electronic orbitals of the hydrogen atom. The ground state in this approximation for the Helium atom will now be given by

$$\psi_0(\mathbf{r}_1, \mathbf{r}_2) = \psi_{100}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2) = \frac{8}{\pi a^3} e^{-2(r_1+r_2)/a} . \tag{3.50}$$

To derive this expression, we have used that the ground state wave function of the hydrogen atom is

$$\psi_{100}(\mathbf{r}) = R_{10}(r)Y_{00}(\theta,\phi) = \frac{2}{a^{3/2}} e^{-r/a}\sqrt{\frac{1}{4\pi}} = \frac{1}{\pi a^{3/2}} e^{-r/a} , \tag{3.51}$$

and then used that from the definition of the Bohr radius, Eq. (2.58), I need to replace a by a/Z if instead of a single proton I have Z protons in the atomic nucleus, as is the case for hydrogen-like atoms.

Since the spatial part of this Helium ground state wave function, Eq. (3.50), is symmetric, $\psi_0(\mathbf{r}_1, \mathbf{r}_2) = \psi_0(\mathbf{r}_2, \mathbf{r}_1)$, then the fact that electrons are fermions implies that the spin part of the wave function must be **antisymmetric**, and thus can only be the **singlet configuration**, with the z component of their spins pointing in opposite directions. Hence we conclude that the total wave function of the helium atom will be given by

$$\psi_0(\mathbf{r}_1, \mathbf{r}_2)|0\ 0\rangle = \frac{1}{\sqrt{2}}\psi_{100}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2)\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right) \tag{3.52}$$

in terms of the tensor product of the spatial and spin states of the wavefunctions, in the approximation where the neglect the electrostatic repulsion between the two electrons.

Neglecting the repulsive interactions between the two electrons is of course not a good approximation in any sense (note that the strength of the mutual electron repulsion is only half of the effect of the positive attraction from the nucleus). Below we provide some strategies in which to improve our predictions for the energies and the wave functions of the Helium atom.

Shielding and the effective charge. In the very crude approximation where we completely neglect the repulsive interaction between the two electrons, the ground state of Helium has an energy of

$$E_0 = 4 \times 2 \times (-13.6 \text{ eV}) = -109\,\text{eV} , \tag{3.53}$$

to be compared with the experimentally determined value of $E_0 = -78.975$ eV, which is off by almost 30%. Of course, this is not unexpected: the electron repulsive interactions is by no means a subleading effect, and hence it is expected that our crude approximation is not very close to the actual result.

The basic assumption that underlies the orbital approximation that we have used is that the motion of the two electrons in the system is *uncorrelated* between them: it is not affected by its mutual repulsive interaction. However, we know that this is not the case: the mutual repulsion between the two electrons will modify their motion around the nucleus to some extent. While we cannot compute fully this effect, we can account for it in an approximate manner:

The effective electric charge

In multi-electron atoms, one electron feels the **average presence** of the other electron. For example, for some configurations the two electrons will have little overlap, while for other configurations they will be on average closer. So effectively, the total positive electric charge felt by the electron will be smaller than the value Z from the protons in the nucleus, since the presence of the other electron will be reducing it. One then defines the **effective charge of the atomic nucleus** as the one that one electron experiences due to the *smearing* induced by the presence of the other electron.

In the case of the helium atom, we will have that this smearing will lead to an effective charge $Z_{\rm eff} < Z = 2$ smaller than the number of protons in the nucleus. This effect is also known as **shielding**: the positive charge of the nucleus felt by a given electron is partially shielded by the other electrons that are also orbiting around it. In this case, this effective electric charge of the nucleus turns out to be $Z_{\rm eff} = 1.62$. Note that in principle the value of $Z_{\rm eff}$ can be different depending on the specific orbital, but for the He atom with only two electrons it is the same for all electronic orbitals.

Hence we find that once we account for the effects of shielding via the effective electric charge, the spatial wave function of the ground state of the Helium atom will be

$$\psi_0(\mathbf{r}_1, \mathbf{r}_2) = \psi_{100}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2) = \frac{Z_{\rm eff}^3}{\pi a^3} e^{-Z_{\rm eff}(r_1+r_2)/a} \,, \tag{3.54}$$

with $1 \le Z_{\rm eff} \le 2$ quantifying the strength of the shielding effect. Therefore, we have shown how, within the **orbital approximation**, one can construct the wave functions for the electronic orbitals of the helium atom in terms of the wave functions that we derived in the previous chapter for the hydrogen-like atoms.

Furthermore, we know already that s-type orbitals have a higher likelihood to be found close to the nucleus at $r = 0$ and therefore the effective charge that they experience will be larger than that of the p-type orbitals, which on average are further away from the protons in the nucleus. This effect can be observed from the table below, which compares the value of the effective electric charge $Z_{\rm eff}$ of Helium with that associated to the different orbitals of Carbon (with $Z = 6$):

Element	Z	Orbital	$Z_{\rm eff}$
He	2	$1s$	1.6875
C	6	$1s$	5.6727
		$2s$	3.2166
		$2p$	3.1358

For similar reasons, the value of the shielding constant for d-type orbitals will also be different than for the s- and d-type orbitals: an electron in a d-type orbital is on average more likely be found farther from the nucleus than that of an p-type orbital, and thus one expects a more intense shielding. We can also see that for a given type of orbital, say a s-type orbital, the shielding constant will be larger for more excited states with larger n, which on average are found farther from the atomic nucleus.

Therefore, if we denote by $\sigma_{n,l}$ the shielding constant associated to an orbital with quantum numbers (n, l), we will have the following rules of thumb:

- For orbitals with a fixed principal quantum number n, the larger the value of l the higher the value of the shielding constant, for example we have that

$$\sigma_{3,s} \leq \sigma_{3,p} \leq \sigma_{3,l} \,. \tag{3.55}$$

- For orbitals with a fixed angular quantum number l, the larger the value of the principal quantum number n the higher the value of the shielding constant, for example we have that

$$\sigma_{1,s} \leq \sigma_{2,s} \leq \sigma_{3,s} \,. \tag{3.56}$$

Note also that this is not a fundamental rule and that there are exceptions to the above general principle.

Therefore, we see that in a multi-electron atom the amount of shielding experienced by the electrons will **depend on the orbital that they occupy**. For this reason, in order to evaluate for example the wavelength of photons involved in the transitions between the orbitals of a multi-electron atom, we need to use the following expression for the electron energy:

$$E_{n,\ell} = -\left[\frac{m_e}{2\hbar^2}\left(\frac{\left(Z_{\rm eff}^{(n,\ell)}\right)e^2}{4\pi\epsilon_0}\right)^2\right]\frac{1}{n^2} = \frac{\left(Z_{\rm eff}^{(n,\ell)}\right)^2 E_1^H}{n^2} = -\frac{(13.6\,{\rm eV})\times\left(Z_{\rm eff}^{(n,\ell)}\right)^2}{n^2} \,, \tag{3.57}$$

Note that now the energy levels depend **both** on the quantum numbers n and ℓ, rather than only an n as is the case of the hydrogen-like atoms. The reason of this difference is that now the electric charge $Z_{\rm eff}^{(n,\ell)}$ will vary with the orbital, and different values of ℓ will have associated different effective charges.

To summarise, the effective charge approximation allows us to improve our estimates of the energies and wave functions for multi-electron atoms, as compared to the scenario where we completely neglect the electrostatic repulsion between their electrons.

3.3 Implications for solid-state structure

The quantum mechanical theory of identical particles in general, and the associated Pauli exclusion principle in particular, have huge implications for our understanding of the structure of **solid-state matter**, such as crystalline solids. Here we will introduce two quantum-mechanical models for the solid state and highlight the role of quantum mechanical statistics in determining its properties.

3.3.1 The free-electron gas

We can understand a crystalline solid as a three-dimensional system of atoms (or molecules) which occupy some geometric, periodic lattice structure. The electrons in these atoms can be classified into **core electrons**, that to not participate in chemical bonding, and **valence electrons**, which corresponding to those electrons occupying the outermost orbitals and that are responsible to establish chemical bonds. These valence electrons are known to be **delocalized**: they are not bound to their original atoms, but can rather move freely across the whole solid, jumping their way across the different atoms in the crystalline lattice. Furthermore, these valence electrons are known to form a **electronic band**, which is a region of the phase space where the electrons can occupy any of a continuum of energy values (as opposed to the discrete energy spectrum that characterise the electronic orbitals of individual atoms).

Here, first of all we will study the consequences of the fact that the electrons are delocalised in the solid by means of the **free electron gas model**, and afterwards we will quantify the implications of the periodicity of the atoms in a crystal to explain the appearance of the electronic energy bands using a periodic potential composed by Dirac delta functions.

To very first approximation, we can therefore consider a solid as an **infinite square potential well** in three dimensions: an electron is free to move everywhere within the solid but cannot jump outside it. In other words, the potential that the electron will experience in this system is given by

$$\begin{aligned} V(x,y,z) &= 0 \qquad 0<x<l_x\,, 0<y<l_y\,, 0<z<l_z \\ V(x,y,z) &= \infty \qquad \text{otherwise} \end{aligned} \tag{3.58}$$

where l_x, l_y, l_z are the dimensions of the solid (the potential well) across the x, y, and z directions. For this potential, the Schroedinger equation is nothing but the free-particle Schroedinger equation in three dimensions. Its solutions will be subject to the usual **boundary condition** that the wave function must vanish at the surface of the solid, in the same way as in the infinite quantum well in one dimension:

$$\psi(x=0,y,z)=\psi(x=l_x,y,z)=\psi(x,y=0,z)=\psi(x,y=l_y,z)=\psi(x,y,z=0)=\psi(x,y,z=l_z)=0\,.$$

Since we have a free-particle time of equation, It is then to convince yourselves that its wave function factorises into a product of the one-particle solutions of the infinite quantum well,

$$\psi(x,y,z)=X(x)Y(y)Z(z)\,, \tag{3.59}$$

where we have defined the individual components of the wave function as:

$$X(x)=A_x\cos(k_xx)+B_x\sin(k_xx)\,,\quad k_x=\sqrt{\frac{2mE_x}{\hbar^2}}\,,$$

$$Y(y)-=A_y\cos(k_yy)+B_y\sin(k_xy)\,,\quad k_y=\sqrt{\frac{2mE_y}{\hbar^2}}\,,$$

$$Z(z)=A_z\cos(k_zz)+B_z\sin(k_xz)\,,\quad k_z=\sqrt{\frac{2mE_z}{\hbar^2}}\,,$$

which satisfy the boundary conditions if $A_x=A_y=A_z=0$ and $k_xl_x=n_x\pi$, $k_yl_y=n_y\pi$ and $k_zl_z=n_z\pi$ where as usual n_i are positive integers and B_i are integration constants fixed by the normalisation of the wave function.

Therefore we have demonstrated that the wave function for an electron confined into a three-dimensional box (our model for a solid) is given by

$$\psi_{n_xn_yn_z}(x,y,z)=\sqrt{\frac{8}{l_xl_yl_z}}\sin\left(\frac{n_x\pi}{l_x}x\right)\sin\left(\frac{n_y\pi}{l_y}y\right)\sin\left(\frac{n_z\pi}{l_z}z\right)\,, \tag{3.60}$$

while the corresponding electron energies will be

$$E_{n_1n_2n_3}=\frac{\hbar^2\pi^2}{2m}\left(\frac{n_x^2}{l_x^2}+\frac{n_y^2}{l_y^2}+\frac{n_z^2}{l_z^2}\right)=\frac{\hbar^2k^2}{2m}\,, \tag{3.61}$$

in terms of the magnitude of the **wave vector** $\mathbf{k}=(k_x,k_y,k_z)$. Every possible value of the wave vector corresponds to a different choice of quantum numbers:

$$(k_x, k_y, k_z) = \left(\frac{\pi n_x}{l_x}, \frac{\pi n_y}{l_y}, \frac{\pi n_z}{l_z}\right), \quad n_x, n_y, n_z = 1, 2, 3, \dots, \tag{3.62}$$

and therefore to a different electronic wavefunction. Note that the wave vector has units of 1/length, and it is often referred as to spanning the **conjugate space** of position space.

In order to understand the properties of this free-electron gas system, it is useful to think in terms of the three-dimensional vector space spanned by the **wave vector** $\mathbf{k} = (k_x, k_y, k_z)$. If you imagine a grid at integer values of k_i, then every node in the grid represents a possible **one-particle state** of the system. Since each grid point is separated by π/l_x, π/l_y, and π/l_z in the x, y, and z directions respectively, the volume of each block in the grid is

$$V_{\rm k-block} = \frac{\pi^3}{l_x l_y l_z} = \frac{\pi^3}{V}. \tag{3.63}$$

In these discussions one must be careful, since $V_{\rm k-block}$ refers to a **volume in the k-space** (the conjugate space), rather than V which is volume in position space.

If we had a single electron in this system, we would be done and we could end the discussion here. But of course, a solid is not composed by a single electron but by **a very large number of them**. For example, for a solid composed by N atoms, each atom will contribute with handful d of valence electrons, and hence the total number of electrons that will compose our **free electron gas** (free, except for not being able to jump outside the boundaries of the material) will be

$$N_{\rm electrons} = Nd. \tag{3.64}$$

Since electrons are fermions, they must obey the Pauli exclusion principle, and hence for a given value of wave vector $\mathbf{k} = (k_x, k_y, k_z)$ (hence, for a given volume $V_{\rm k-block} = \pi^3/V$) we can have at most **two electrons**. These electrons must form a **singlet state** with spins pointing in opposite directions.

Therefore, the total volume in k-space occupied by the electrons in this system will be

$$\frac{1}{2} \times V_{\rm k-block} \times N_{\rm electrons} = \frac{1}{2} \times \frac{\pi^3}{V} \times Nd, \tag{3.65}$$

where the 1/2 factor comes since two electrons occupy one $V_{\rm k-block}$. Since (k_x, k_y, k_z) is an array of positive integer numbers, electrons will occupy first the states with lowest k_i and once these are occupied they will occupy higher wave vectors. The volume of the resulting octant of a sphere in k-space, once all electrons have filled the available states, will be given by

$$\frac{1}{8}\left(\frac{4}{3}\pi k_F^3\right) = \frac{Nd}{2}\left(\frac{\pi^3}{V}\right), \tag{3.66}$$

where the radius of this k-space sphere is known as the Fermi momentum:

$$k_F = \left(\frac{3Nd\pi^2}{V}\right)^{1/3} = (3\rho\pi^2)^{1/3} \tag{3.67}$$

where we have defined the **free electron density** as $\rho \equiv Nd/V$, namely the number of free (valence) electrons in the free electron gas system per (physical) unit volume. The surface of this sphere, populated by the electrons with the highest values of the wave vector $\mathbf{k}$, is known as the **Fermi surface**.

The electrons with wave vector $\mathbf{k}$ spanning the surface of the Fermi sphere will have the highest energies in the whole material. By recalling the dispersion relation for free particles, $E = \hbar^2 k^2/2m$, we can determine the energy of those electrons, which is denoted as the **Fermi energy**:

$$E_F = \frac{\hbar^2 k_F^2}{2m} = \frac{\hbar^2}{2m}(3\rho\pi^2)^{2/3} \tag{3.68}$$

The position of the Fermi energy has a crucial impact of the properties of a solid, for example in determining its electrical conductivity properties. From this result we observe how the denser the material, the higher the value of its Fermi energy.

Position versus moment (wave number) space

You should be careful with the distinction between **position** and **moment** space, also denoted as x-space and k-space. The electrons occupy only a sphere (octant) in moment (that is, in energy space), but they remain **delocalised** in position space. Each electron, irrespectively of its wave vector $\mathbf{k}$, has some finite probability of being found everywhere within the solid, as indicated by the wave function Eq. (3.60). Hence the Fermi sphere is a geometrical construct in moment space, which does not have a counterpart in position space.

We can now evaluate the total energy of my system of Nd electrons in this free-electron gas model. Within the Fermi sphere in k-space, a shell with (infinitesimal) thickness dk contains a volume

$$dV = \frac{1}{8}\left(4\pi k^2\right) dk\,. \tag{3.69}$$

Now, the **density of electrons in k-space** is $2/V_{\text{k}-\text{block}} = 2V/\pi^3$, since each k-block within the Fermi sphere contains two electrons. Hence, we have that the number of electrons in this infinitesimal shell is

$$dN_{\text{el}} = \frac{1}{8}\left(4\pi k^2\right) dk \times \frac{2V}{\pi^3} = \frac{V}{\pi^2} k^2 dk\,. \tag{3.70}$$

Furthermore, a state represented by the wave vector $\mathbf{k}$ has associated an energy of $E = \hbar^2 k^2/2m$, and therefore the energy of the shell will be

$$dE = dN_{\text{el}} \times \frac{\hbar^2 k^2}{2m} = \left(\frac{V}{\pi^2} k^2 dk\right) \times \left(\frac{\hbar^2 k^2}{2m}\right) = \frac{V\hbar^2 k^4}{2\pi^2 m} dk\,. \tag{3.71}$$

At this point, in order to evaluate the total energy of the free-electron gas system, I only need to integrate over all available wave vectors, from $k \simeq 0$ for the ground state up to the Fermi energy k_F. We obtain the following result:

$$E_{\text{tot}} = \int_0^{k_F} \frac{V\hbar^2 k^4}{2\pi^2 m} dk = \frac{\hbar^2 k_F^5 V}{10\pi^2 m}\,. \tag{3.72}$$

It is convenient to simplify this important result by exploiting the expression of the Fermi energy Eq. (3.68), which remember is the energy of the electrons with the highest wave numbers $\boldsymbol{k}$ of the system. By doing this we find that the total energy of the free electron gas is:

$$E_{\text{tot}} = \frac{\hbar^2 (3\pi^2 Nd)^{5/3}}{10\pi^2 m} V^{-2/3}\,. \tag{3.73}$$

The total energy grows with the number of free electrons as $N_{\text{el}}^{5/3}$, and decreases with the volume of the solid. The latter result can be understood by the fact that in the $V \to \infty$ limit and fixed number of electrons, the volume of the k-block vanishes and electrons occupy states with very low energies.

While we are referring of this system as a free-electron gas, this *gas* is very different from the usual gases in that the electrons themselves are **non-interacting**, and thus their energy Eq. (3.73) is **purely quantum mechanical** in origin (as opposed to a regular gas, whose energy is ultimately thermal in origin). Despite this different origin, this quantum mechanical energy exerts a very measurable pressure on the walls of a solid, in the same way as the pressure of a gas arises from the collisions of its constituent molecules with the walls of the contained.

We can determine the **quantum mechanical pressure** exerted by our free-electron gas as follows. If the solid (container of the electron gas) expands by an amount dV, you can see that its total energy decreases:

$$dE_{\rm tot} = -\frac{2}{3}\frac{\hbar^2(3\pi^2 Nd)^{5/3}}{10\pi^2 m}V^{-5/3}dV = -\frac{2}{3}E_{\rm tot}\frac{dV}{V}\,. \tag{3.74}$$

If you recall from your thermal physics course the relation between **work, pressure and volume**,

$$dW = PdV\,, \tag{3.75}$$

we find that the quantum mechanical pressure exerted by the electrons on the walls of the solid is given by:

$$P = \frac{2}{3}\frac{E_{\rm tot}}{V} = \frac{(3\pi^2)^{2/3}\hbar^2}{5m}\rho^{5/3}\,, \tag{3.76}$$

which is stronger the higher the electron density of the material.

The existence of such quantum mechanical pressure, sometimes referred to as **degeneracy pressure**, provides some input to the question of why a **cold solid does not collapse**, demonstrating that what (partially) ensures its “solidity” is not thermal effects (which have been ignored) nor the interactions of the electrons among them or with the atomic nuclei (which we have neglected). A cold solid does not collapse because it is composed of identical fermions, and fermions do not like to occupy the same quantum state!

Note that it is quite remarkable how we have managed to produce a reasonably good prediction for the properties of a solid using exclusively the fact that **electrons are fermions**, without saying a single word about their interactions. This demonstrates that many of the fundamental properties of solids are dictated by basic quantum mechanical considerations, rather than from the details of the specific interactions present within in the system.

3.3.2 Band structure in solids

The main limitation of the free electron gas model is that it considers a solid as a collection of non-interacting electrons restricted to the material volume, but ignores the fact that these electrons arise from the valence orbitals of atoms that are distributed in a more or less **regular manner** within the solid. Thus this model neglects the interactions of these valence electrons with the atoms that compose the solid. In order to improve the free electron model, we need to account for, at least in an approximate way, the **interactions between the atoms in the solid and the electrons**.

As in the case of the free-electron gas, it can be shown that the specific **details of the atom-electron interactions** are not relevant, and that many important properties of solids can be recovered just by taking into account the fact that the atomic nuclei are stationary and **distributed regularly** within the solid. Hence, the defining feature of the model we are going to present now is that the regular arrangement of the atoms within a crystalline solid gives rise to **periodic potential**, which repeats itself in the three spatial directions.

The discussion of quantum systems characterised by periodic potentials is greatly simplified by the powerful **Bloch's Theorem**. Consider a one-dimensional quantum system characterised by a periodic potential, namely a potential that repeats itself after a distance a:

$$V(x+a) = V(x)\,, \tag{3.77}$$

which of course implies that $V(x+2a) = V(x+a) = V(x)$ or in general that

$$V(x+na) = V(x) \quad \text{for} \quad n = 1, 2, 3, \ldots\,. \tag{3.78}$$

The wave function for a particle propagating through this potential energy will satisfy the one-dimensional Schroedinger equation,

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) + V(x)\psi(x) = E\psi(x)\,. \tag{3.79}$$

Bloch's Theorem tells us that, given the periodic nature of the potential defined by Eq. (3.78), the solution of the one-dimensional Schroedinger equation, must satisfy:

$$\psi(x+a) = e^{iqa}\psi(x)\,, \tag{3.80}$$

with q some constant number independent of the position x. To demonstrate this property, note that if $\psi(x)$ satisfies Eq. (3.79) then it will also satisfy

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x+a) + V(x+a)\psi(x+a) = E\psi(x+a)\,, \tag{3.81}$$

where I have rescaled $x \to x+a$. But we are dealing with a **periodic potential** as indicated by the periodicity condition Eq. (3.78), and hence:

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x+a) + V(x)\psi(x+a) = E\psi(x+a)\,, \tag{3.82}$$

and if now we replace

$$\psi(x+a) = e^{iqa}\psi(x) \tag{3.83}$$

we reproduce the original equation Eq. (3.79), demonstrating Bloch's Theorem. The overall prefactor e^{iqa} does not depend on x and thus cancels out in the Schroedinger equation.

Periodicity of the electron probability distribution

For the time being q is a general complex number, and thus it could change the normalisation of the wave function. In short we will demonstrate that $q \in \mathbb{R}$ and thus the e^{iqa} prefactor is a **complex phase**. This is a reassuring result, given the periodicity of the potential heavily suggests that the **probability distribution** associated to the electron should also be **periodic**, and hence that

$$|\psi(x+a)|^2 = |\psi(x)|^2\,, \tag{3.84}$$

which is indeed the case thanks to Bloch's Theorem, Eq. (3.80), given that e^{iqa} is a complex phase. Therefore in a crystalline solid with a periodic potential the probability distribution associated to its valence electrons will be periodic as well.

While any real solid is finite, finite-size effects associated to its boundaries can be neglected for macroscopic solids as compared to the dominant bulk (volume) effects. We can ignore them formally by imposing the **periodic boundary condition**

$$\psi(x+Na)=\psi(x)\,, \tag{3.85}$$

where $N=\mathcal{O}\left(10^{23}\right)$ is a macroscopic number of periods, which corresponds to a macroscopic number of atoms. Taking into account the periodicity properties of the wave function, Eq. (3.80), we find that imposing this periodic boundary conditions leads to

$$\psi(x+Na)=e^{iqNa}\psi(x)=\psi(x) \quad \to \quad qNa=2\pi n\,, \tag{3.86}$$

and thus the constant q is **real and quantised**:

$$q_n=\frac{2\pi n}{Na}\,, \quad n=0,\pm1,\pm2,\pm3,\ldots \tag{3.87}$$

Note that $(1/Na)\simeq 10^{-13}$ for a typical interatomic separation of $a=\mathcal{O}\left(10^{-10}\right)$, and hence the values of q_n will form a **quasi-continuum**. Note that if we had not imposed the periodic boundary conditions then q would indeed take continuous values, but the difference between quasi-continuum actual continuum is irrelevant in practice.

The usefulness of Bloch's Theorem

The crucial relevance of Bloch's theorem is that we have reduced the problem of solving the Schroedinger equation on the whole solid to that of **solving it in a unit cell**, for example the cell defined by $0\leq x\leq a$. Once the wave function for this unit cell is known, then Eq. (3.80) guarantees that we can evaluate it for any other cell within the solid.

Armed with the powerful Bloch Theorem, we can now present our model for the band structure of solids. Still restricting ourselves to the one-dimensional case, let us assume that our periodic potential is a linear combination of **Dirac delta functions**:

$$V(x)=\alpha\sum_{j=0}^{N-1}\delta(x-ja)\,. \tag{3.88}$$

That is, we have infinite spikes localised at the nodes $x=ja$ of the lattice with j being an integer number. The specific details of the potential are not crucial (one could have used finite square barrier potentials, for example), the really important property is their **periodicity**. Indeed, while this potential is far from the real physical situation (we should have Coulomb-type attractive interactions between the electrons and the positive atomic cores), it allows us to evaluate the physical consequences of a periodic potential in the resulting properties of solids.

The potential energy defined by Eq. (3.88) is clearly a periodic potential, since $V(x+a)=V(x)$, as you can check as follows

$$V(x+a)=\alpha\sum_{j=0}^{N-1}\delta(x-ja+a)=\alpha\sum_{j=0}^{N-1}\delta(x-(j-1)a)\,, \tag{3.89}$$

but now you can redefine $j' = j - 1$ to have

$$V(x+a) = \alpha \sum_{j=0}^{N-1} \delta(x-(j-1)a) = \alpha \sum_{j'=-1}^{N-2} \delta(x - j'a) = V(x) \tag{3.90}$$

since we are imposing **periodic boundary conditions** for the potential, Eq. (3.85) such that the lattice node at $j' = -1$ in the sum is identified with the node with $j' = N-1$. Indeed, is we plug in $x = -a$ in Eq. (3.85) we get

$$\psi(-a+Na) = \psi(-a) \quad \to \quad j' = N-1 \quad \text{identified with} \quad j' = -1\,, \tag{3.91}$$

and hence the potential is indeed periodic as required.

By virtue of Bloch's Theorem, we only need to solve the Schroedinger equation for this potential in a single "crystal" cell, which we take to be $0 \le x \le a$. The delta function potential Eq. (3.88) is clearly zero for $0 < x < a$, and it is only non-zero for $x = 0$ and $x = a$. Therefore, for $0 \le x \le a$ we recover the usual free-particle wave functions

$$\psi(x) = A\sin(kx) + B\cos(kx)\,, \quad k = \sqrt{2mE}/\hbar\,, \quad 0 < x < a\,. \tag{3.92}$$

We can now use Bloch's Theorem,

$$\psi(x+a) = e^{iqa}\psi(x) \quad \to \quad \psi(x) = e^{-iqa}\psi(x+a)\,, \tag{3.93}$$

to compute the wave function in the cell to the left of this one, $-a < x < 0$, and we find

$$\psi(x) = e^{-iqa}\left(A\sin(k(x+a)) + B\cos(k(x+a))\right)\,, \quad k = \sqrt{2mE}/\hbar\,, \quad -a < x < 0\,. \tag{3.94}$$

In order to determine the integration constants A and B, we need to impose the relevant **local boundary conditions** on the wave function of this system. We will this way find the conditions to be satisfied by the physical wave functions of this periodic system.

- The wave function must be **continuous everywhere**, and hence for $x = 0$ we must impose:

  $$B = e^{-iqa}\left(A\sin(ka)) + B\cos(ka)\right) \tag{3.95}$$

  which is a transcendental equation similar to the one that arises in the finite potential barrier and related problems (which can only be solved numerically).

- The derivative of the wave function is **discontinuous across infinite potentials**, such as the Dirac delta function of this system, with the value of the discontinuity given by

  $$\left.\frac{d\psi(x)}{dx}\right|_{+\epsilon} - \left.\frac{d\psi(x)}{dx}\right|_{x=-\epsilon} = \frac{2m}{\hbar^2}\int_{-\epsilon}^{\epsilon} dx\, V(x)\psi(x)\,. \tag{3.96}$$

  where I am assuming that the potential $V(x)$ is infinite at $x_0 = 0$. Our potential at $x = 0$ is a delta function and hence the RHS reads

  $$\frac{2m}{\hbar^2}\int_{-\epsilon}^{\epsilon} dx\, V(x)\psi(x) = \frac{2m\alpha}{\hbar^2}\int_{-\epsilon}^{\epsilon} dx\, \delta(x)\psi(x) = \frac{2m\alpha}{\hbar^2}\psi(0) = \frac{2m\alpha}{\hbar^2}B\,. \tag{3.97}$$

Now in the original cell, $x = +\epsilon$, the derivative is given by

$$\frac{d\psi(x)}{dx} = Ak\cos(kx) - Bk\sin(kx)\,, \quad \left.\frac{d\psi(x)}{dx}\right|_{+\epsilon} = kA \tag{3.98}$$

while instead for $x = -\epsilon$ the corresponding derivative is evaluated to be:

$$\begin{aligned} \frac{d\psi(x)}{dx} &= e^{-iqa}\left(Ak\cos(k(x+a)) - Bk\sin(k(x+a))\right), \\ \left.\frac{d\psi(x)}{dx}\right|_{-\epsilon} &= e^{-iqa}\left(Ak\cos(ka)) - Bk\sin(ka)\right) \end{aligned} \tag{3.99}$$

and thus the boundary condition for the derivate of the wave function reads:

$$\left[kA - e^{-iqa}k\left(A\cos(ka)) - B\sin(ka)\right)\right] = \frac{2m\alpha}{\hbar^2}B\,. \tag{3.100}$$

Eqns. (3.95) and (3.100) provide two relations that constrain the integration constants A and B and the allowed values of the energy k of the particles propagating in this periodic potential.

To solve this system of equations, we can first use Eq. (3.95) to determine A in terms of B

$$A = B\frac{\left(e^{iqa} - \cos(ka)\right)}{\sin(ka)} \tag{3.101}$$

and then insert this relation into Eq. (3.100), where the dependence on B cancels out. This indicates that B is a normalisation constant fixed by the wave function normalisation requirement. Doing some algebra, we end up with the following relation between the quasi-continuous parameter q and the particle wave vector k as a function of the strength α of the potential:

$$\cos(qa) = \cos(ka) + \frac{m\alpha}{\hbar^2 k}\sin(ka)\,. \tag{3.102}$$

This relation illustrates how for given values of the potential strengths α and of the **lattice parameter** a the energies of the electrons $E = \hbar^2k^2/2m$ in the system are related to the values of $q = 2\pi n/Na$, and hence they will be **quantised**. Indeed, for given values of q, a. and α, only electron energies such that

$$E = \frac{\hbar^2k^2}{2m} \quad \text{satisfying} \quad \cos(qa) = \cos(ka) + \frac{m\alpha}{\hbar^2 k}\sin(ka)\,, \tag{3.103}$$

are allowed solutions of the Schroedinger equation of this periodic potential. However, there is now a marked difference with previous systems that we have considered: the fact that q is essentially a continuous variable (since N is macroscopic) implies that instead of discrete energy levels, we will have **continuous energy bands** separated by **gaps** without any allowed states.

To determine which are the allowed values of the energy E of the electrons in this system, we can write $z = ka$ and $\beta = ma\alpha/\hbar^2$ such that Eq. (3.102) becomes

$$\cos\left(\frac{2\pi n}{N}\right) = \cos(z) + \frac{\beta}{z}\sin(z) = f_\beta(z)\,. \tag{3.104}$$

For a given value of β, which is proportional to the strength of the interaction between the electrons and the atomic nuclei, for each value of $n = 1, 2, 3, 4, \ldots$ there will be a solution for z.

However, N is a huge number, so for any value of $\cos(qa)$ between -1 and +1 it is essentially guaranteed that we will find a value of n that fulfills this condition. Hence, any value of $k = z/a$ such that $|f_\beta(z)| \leq 1$ corresponds to an **allowed energy state** for my system: we therefore find that the valence electrons moving in the periodic potential of a solid cluster into **bands** of allowed energy separated by **energy gaps** (for $|f(z)| > 1$). This is a rather significant difference as compared to *e.g.* the hydrogen atom, or to any other quantum systems: in periodic potentials, the allowed levels of the electron energies cluster into **continuum bands** separated by finite gaps where no solution is possible.

Fig. 3.1 displays the function $f_\beta(z)$, Eq. (3.104) as a function of z for $\beta = 30$ and $\beta = 4$. We indicate with dashed horizontal lines the region $|f_\beta(z)| \leq 1$ which defines the allowed electronic bands. In the bottom panels, we indicate with a solid grey rectangle the location of these allowed electronic bands in the two scenarios. Note how the width of the band increases as z is increased, and eventually becomes a continuum without any gaps breaking the energy band. Note also that smaller the value of β, the earlier the energy continuum kicks in (since in $\beta \to 0$ limit the strength of the periodic potential vanishes).

Note also that in the limit $\beta \to 0$ we can always find a solution to Eq. (3.104): this means that in this limit the energies are not quantised and can take any possible value. This is not unexpected, since for $\beta \to 0$ we recover the free particle system (subject to periodic boundary condition) and there we know that any positive value of the energy is physically allowed.

Summary

We can now recapitulate what have we learned in this chapter concerning the quantum mechanics for systems composed by **identical particles**:

I/ Quantum mechanics deals with truly indistinguishable particles: an electron is identical in all respects to every other electron in the Universe.

II/ Systems composed by identical particles in quantum mechanics behave in a very different way as compared to systems composed by distinguishable particles.

III/ Depending on their spin, particles can be classified into fermions and bosons. Bosons behave as if they experienced some attractive interaction and tend to cluster together in the same quantum state. Fermions, on the contrary, behave as if they experienced some repulsive interaction and tend to get far for each other.

IV/ The Pauli exclusion principle for fermions tells us that a given quantum state can be occupied at most by one fermion at the same time.

V/ The free electron gas model allows us to explain some important properties of solids, and shows how there exists a quantum mechanical pressure from the fermion degeneracy that partially explains the stability of solids.

VI/ The existence of continuous energy bands for the valence electrons in solids can be explained by the presence of a periodic potential, where the electrons move, generated by the atomic cores sitting at regularly-spaced locations within the crystalline lattice of a solid.

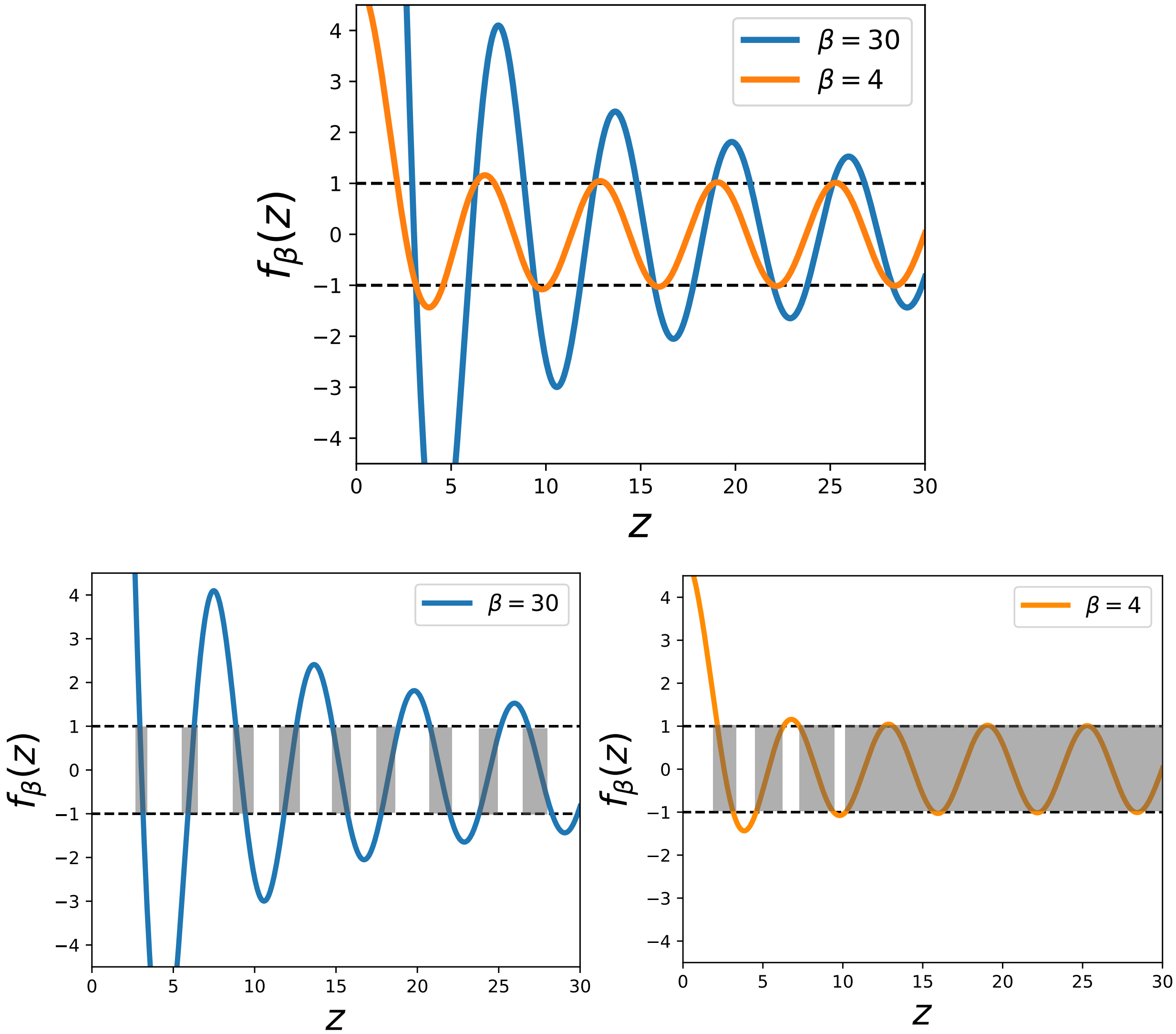

Figure 3.1: The function $f_\beta(z)$, Eq. (3.104) as a function of z for $\beta = 30$ and $\beta = 4$. We indicate with dashed horizontal lines the region $|f_\beta(z)| \leq 1$ which defines the allowed electronic bands. In the bottom panels, we indicate with a solid grey rectangle the location of these allowed electronic bands in the two scenarios. Note how the width of the band increases as z is increased, and eventually becomes a continuum without any gaps breaking the energy band. The smaller the value of β, the earlier the energy continuum kicks in.

Quantum Mechanics 2

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

Current version: **March 2, 2021**

4 Chapter 7: Time-independent perturbation theory

Learning Goals

- To determine in a systematic manner the **higher-order corrections** to the energies and state vectors of an unperturbed quantum system by means of perturbation theory.
- To assess the criteria that determine the conditions whereby a certain perturbative expansion can be expected to be **convergent**.
- To identify the conditions of applicability of the perturbative techniques in the case of **degenerate spectra.**
- To compute the first non-trivial corrections to the **energy levels of the hydrogen atom**, arising from relativistic effects and from the spin-orbit interactions.

In this section of the lecture notes we present the main concepts discussed in **Chapter 7 ("Time-Independent Perturbation Theory")** 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

- **7.1**: Non-degenerate Perturbation Theory.
- **7.2**: Degenerate Perturbation Theory.
- **5.3**: The Fine Structure of Hydrogen.

The number of quantum mechanical systems that can be solved in an **analytic closed form** is unfortunately rather small. With a few exceptions, in general we will need to carry out some approximation in order to solve a given problem in quantum mechanics. The virtue of **perturbation theory** is that, starting from the known solutions to a simpler (“unperturbed) problem, it makes possible determining systematically and in a fully controllable manner the corrections that modify the unperturbed results. In other words, perturbation theory is a systematic procedure for obtaining approximate solutions to the perturbed problem, by building on the exact known solutions to the unperturbed case. The higher the number of terms we include on this series expansion, the better our estimate of the full energies and wave functions will be. Perturbation theory is extremely valuable in quantum mechanics, and here we will illustrate its use with different examples including the calculation of the **fine structure of hydrogen**.

4.1 Non-degenerate perturbation theory

Let us assume that we have some quantum system that we have been able to solve **exactly**. Some instances of exactly-solvable systems that you have encountered so far are the infinite square and spherical well potentials, the quantum harmonic oscillator in 1D and 3D, and the hydrogen atom. By solving the system we mean finding the eigenvalues and eigenvectors of the corresponding Schroedinger equation, which in full generality can be expressed as

$$\hat{H}^{(0)}|\psi_k^{(0)}\rangle = E_k^{(0)}|\psi_k^{(0)}\rangle \,, \tag{4.1}$$

where the superindices $^{(0)}$ indicate that these are the **exact solutions to the unperturbed Hamiltonian** $\hat{H}^{(0)}$. As usual, the eigenvectors of this unperturbed Hamiltonian will be orthonormal,

$$\langle\psi_k^{(0)}|\psi_l^{(0)}\rangle = \delta_{kl} \,. \tag{4.2}$$

The exact solutions of the unperturbed Hamiltonian Eq. (4.1) are a necessary starting point of the perturbative method: if these are not available, the perturbative expansion is not applicable.

Degenerate vs non-degenerate problems

Although it might not seem a big deal at this point, the derivation that follows is only guaranteed to work if the unperturbed Hamiltonian $\hat{H}^{(0)}$ has associated a **non-degenerate spectra** where if $j \neq k$ then $E_k^{(0)} \neq E_j^{(0)}$. The presence of degeneracies in the energy spectrum introduces problems that we will discuss how to tackle later in this chapter.

Now consider that we modify the original potential in the unperturbed Hamiltonian Eq. (4.1) in the following manner

$$V(x) = V^{(0)}(x) + \delta V(x) \,, \tag{4.3}$$

where $\delta V(x)$ is some **small correction** with respect to the original potential energy $V^{(0)}(x)$. We will quantify soon what exactly do we mean by this “smallness” requirement. We want to solve the new Schroedinger equation in terms of the new eigenvalues and eigenvectors

$$\hat{H}|\psi_k\rangle = E_k|\psi_k\rangle \quad \text{where} \quad \hat{H} = \hat{H}^{(0)} + \delta V(x) \tag{4.4}$$

where we have removed the superindices to indicate that now eigenvalues and eigenvectors refer to the full Hamiltonian. We will do this by using **perturbation theory** under the assumption that $\delta V(x)$ will only modify in a moderate manner the original eigenvalues and eigenvectors (of course, if you know how to solve exactly Eq. (4.4), perturbation theory is irrelevant!).

With this motivation, let us express the new, perturbed Hamiltonian as

$$\hat{H} = \hat{H}^{(0)} + \lambda \hat{H}' \,, \tag{4.5}$$

where λ is a **small dimensionless parameter**. Note that the kinetic term is not modified by the change in the potential energy, and hence we will have that $\lambda\hat{H}' = \delta V(x)$. Since λ is a small dimensionless parameter, it seems sensible to express both the new wave function and the corresponding energies in terms of a **series expansion** in λ, starting from the original (unperturbed) values:

$$\begin{aligned} \big|\psi_n\big\rangle &= \big|\psi_n^{(0)}\big\rangle + \lambda\big|\psi_n^{(1)}\big\rangle + \lambda^2\big|\psi_n^{(2)}\big\rangle + \mathcal{O}\left(\lambda^3\right) \,, \\ E_n &= E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \mathcal{O}\left(\lambda^3\right) \,. \end{aligned} \tag{4.6}$$

Some comments about this expansion are relevant now:

- $\big|\psi_n^{(1)}\big\rangle$ and $E^{(1)}$ are known as the **first-order corrections** to the state vector and the energies, $\big|\psi_n^{(2)}\big\rangle$ and $E^{(2)}$ are known as the **second-order corrections** and so on.
- Each subsequent term in this series is **suppressed** by a factor of the small expansion parameter λ as compared to the previous one.[14]
- The more terms in the expansion in λ we include, in principle the better estimate to the full result that we can expect. The number of required terms to obtain a reliable approximation depends, among other things, on the value of λ.
- If λ is small enough, this series expansion is guaranteed to represent a good approximation: in the $\lambda \to 0$ limit then $\hat{H} \to \hat{H}^{(0)}$ and the wave function and energies reduced to the original (unperturbed) values.

We can now try to insert these perturbative expansions, Eq. (4.6) into the original Schroedinger equation, we thus obtain:

$$\left(\hat{H}^{(0)} + \lambda\hat{H}'\right)\left(\big|\psi_n^{(0)}\big\rangle + \lambda\big|\psi_n^{(1)}\big\rangle + \lambda^2\big|\psi_n^{(2)}\big\rangle + \ldots\right) \tag{4.7}$$

$$= \left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots\right)\left(\big|\psi_n^{(0)}\big\rangle + \lambda\big|\psi_n^{(1)}\big\rangle + \lambda^2\big|\psi_n^{(2)}\big\rangle + \ldots\right) \tag{4.8}$$

The key idea underlying perturbation theory is to collect terms at the **same order** in the expansion in the small parameter λ, and solve the corresponding equations in a sequential manner: first we solve the corrections to the energies and the state vectors at $\mathcal{O}\left(\lambda\right)$, then at $\mathcal{O}\left(\lambda^2\right)$ and so on. Recall that $\mathcal{O}\left(\lambda^0\right)$ corresponds to the unperturbed result which we assume has already been solved/

By collecting the terms proportional to λ^0, λ^1, and λ^2 we find the following relations:

$$\hat{H}^{(0)}\big|\psi_k^{(0)}\big\rangle = E_k^{(0)}\big|\psi_k^{(0)}\big\rangle \qquad \left[\text{order } \lambda^0\right] \tag{4.9}$$

$$H^{(0)}\big|\psi_k^{(1)}\big\rangle + H'\big|\psi_k^{(0)}\big\rangle = E_k^{(0)}\big|\psi_k^{(1)}\big\rangle + E_k^{(1)}\big|\psi_k^{(0)}\big\rangle \qquad \left[\text{order } \lambda^1\right] \tag{4.10}$$

$$H^{(0)}\big|\psi_k^{(2)}\big\rangle + H'\big|\psi_k^{(1)}\big\rangle = E_k^{(0)}\big|\psi_k^{(2)}\big\rangle + E_k^{(1)}\big|\psi_k^{(1)}\big\rangle + E_k^{(2)}\big|\psi_k^{(0)}\big\rangle \qquad \left[\text{order } \lambda^2\right] \tag{4.11}$$

---

[14] Actually in many cases the perturbative series are only **asymptotic**, meaning that at some point including higher-order terms in the expansion degrades, rather than improves, the agreement with the exact result.

The first of these relations, Eq. (4.9) was our starting point (the Schroedinger equation for the unperturbed system, we we assume to be solvable and known). Solving Eq. (4.10) will allow us to evaluate the first-order corrections to the unperturbed state vectors and the energies.

Notice also the neat property of the perturbative relations: the equation for $E_k^{(1)}$ and $\left|\psi_k^{(1)}\right\rangle$ relies only on the knowledge of $E_k^{(0)}$ and $\left|\psi_k^{(0)}\right\rangle$, while the relation for $E_k^{(2)}$ and $\left|\psi_k^{(2)}\right\rangle$ relies on the knowledge of both $E_k^{(0)}$ and $\left|\psi_k^{(0)}\right\rangle$ as well as $E_k^{(1)}$ and $\left|\psi_k^{(1)}\right\rangle$. This structure indicates that the perturbation theory technique must the applied in a **sequential manner**: unless I compute the first-order corrections, there is no way I may attempt to evaluate the second-order corrections.

4.1.1 First-order perturbation theory

Given that the leading order (unperturbed) problem is solved, let us now attempt to solve Eq. (4.10) and determine the $\mathcal{O}(\lambda)$ corrections to the energy $E_k^{(1)}$ and the state vector $|\psi_k^{(1)}\rangle$, and use them to improve the unperturbed $\mathcal{O}(\lambda^0)$ results. The starting point is Eq. (4.10):

$$H^{(0)}\left|\psi_k^{(1)}\right\rangle + H'\left|\psi_k^{(0)}\right\rangle = E_k^{(0)}\left|\psi_k^{(1)}\right\rangle + E_k^{(1)}\left|\psi_k^{(0)}\right\rangle \tag{4.12}$$

We can now multiply from the left by the bra $\left\langle\psi_k^{(0)}\right|$ to obtain the following scalar equation

$$\left\langle \psi_k^{(0)} \middle| \hat{H}^{(0)}\psi_k^{(1)} \right\rangle + \left\langle \psi_k^{(0)} \middle| \hat{H}'\psi_k^{(0)} \right\rangle = E_k^{(0)}\left\langle \psi_k^{(0)} \middle| \psi_k^{(1)} \right\rangle + E_k^{(1)}\left\langle \psi_k^{(0)} \middle| \psi_k^{(0)} \right\rangle \tag{4.13}$$

The braket in the last term gives just unity since the eigenvectors are normalised. Furthermore, we can use the property that the Hamiltonian is an Hermitian operator to write

$$\left\langle \psi_k^{(0)} \middle| \hat{H}^{(0)}\psi_k^{(1)} \right\rangle = \left\langle \hat{H}^{(0)}\psi_k^{(0)} \middle| \psi_k^{(1)} \right\rangle = E_k^{(0)}\left\langle \psi_k^{(0)} \middle| \psi_k^{(1)} \right\rangle \tag{4.14}$$

and therefore we can express the relation in Eq. (4.13) as

$$E_k^{(0)}\left\langle \psi_k^{(0)} \middle| \psi_k^{(1)} \right\rangle + \left\langle \psi_k^{(0)} \middle| \hat{H}'\psi_k^{(0)} \right\rangle = E_k^{(0)}\left\langle \psi_k^{(0)} \middle| \psi_k^{(1)} \right\rangle + E_k^{(1)}\,, \tag{4.15}$$

which gives us the result for the **first-order corrections to the energy**

$$E_k^{(1)} = \left\langle \psi_k^{(0)} \middle| \hat{H}'\psi_k^{(0)} \right\rangle\,, \tag{4.16}$$

that is, the matrix element of the perturbed Hamiltonian evaluated by means of the unperturbed (original) eigenvectors. Note we since we know both the explicit expressions of the eigenvectors $\left|\psi_k^{(0)}\right\rangle$ and of $\hat{H}'$, we can always evaluate this quantity since we don't need to solve any equation before. We thus find that the **first-order correction to the energy** is the **expectation value of the Hamiltonian perturbation in the original states**.

Perturbative corrections to the infinite well potential. Let us illustrate this important result with an specific example. Assume that we have a particle confined to the infinite square well of width L. We have extensively studied this system in the present course, and at this point we probably know by heart the expression for its eigenfunctions and energies:

$$\psi_n^{(0)}(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right)\,, \quad E_n^{(0)} = \frac{\pi^2 n^2 \hbar^2}{2mL^2}\,, \quad n = 1, 2, 3, \ldots\,. \tag{4.17}$$

Consider now that we modify the original infinite potential well $V(x)$ with a sinusoidal correction of the form

$$\delta V(x) = \lambda V_1 \sin\left(\frac{\pi x}{L}\right) . \tag{4.18}$$

This correction has the same shape as the **ground state** of the system, and in particular vanishes at the edges of the well and has a maximum for $x = L/2$. In this perturbed potential, λ is some **dimensionless parameter**, while V_1 has units of energy and is responsible for the strength of the perturbation. For perturbation theory to work, λ must be small in some sense, and we will quantify below what this implies.

By means of Eq. (4.16), one can compute the first-order correction to the energy levels of this quantum state in terms of the following trigonometric integrals:

$$E_n^{(1)} = \left\langle \psi_n^{(0)} \middle| \hat{H}' \psi_n^{(0)} \right\rangle = \left\langle \psi_n^{(0)} \middle| V_1 \sin\left(\frac{\pi x}{L}\right) \psi_n^{(0)} \right\rangle = \frac{2V_1}{L} \int_0^L dx \, \sin^2\left(\frac{n\pi x}{L}\right) \sin\left(\frac{\pi x}{L}\right) \tag{4.19}$$

By using known results for these trigonometric integrals, we can evaluate the first-order corrections to the first energy levels and find, in terms of strength of the perturbed potential V_1, and hence to obtain that at $\mathcal{O}(\lambda)$ in the perturbative expansion our estimate for the full energies of the system will be:

$$\begin{aligned} E_1^{(1)} &= \frac{8V_1}{3\pi} \simeq 0.85\, V_1 \quad \rightarrow \quad E_1 = \frac{\pi^2\hbar^2}{2mL^2} + \lambda \frac{8V_1}{3\pi} , \\ E_2^{(1)} &= \frac{32V_1}{15\pi} \simeq 0.68\, V_1 \quad \rightarrow \quad E_2 = \frac{4\pi^2\hbar^2}{2mL^2} + \lambda \frac{32V_1}{15\pi} , \\ E_3^{(1)} &= \frac{72V_1}{35\pi} \simeq 0.66\, V_1 \quad \rightarrow \quad E_3 = \frac{9\pi^2\hbar^2}{2mL^2} + \lambda \frac{72V_1}{35\pi} . \end{aligned} \tag{4.20}$$

You can evaluate the first-order perturbative corrections for higher excited states using the same approach.

At this point we are ready to answer an important question: in which respect λ is supposed to be small for perturbation theory to work? We have found that in this system, at first-order in the perturbative expansion, the energies of the eigenfunction are modified as

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \mathcal{O}\left(\lambda^2\right) = \frac{\pi^2 n^2 \hbar^2}{2mL^2} + \lambda F_n V_1 + \mathcal{O}\left(\lambda^2\right) , \tag{4.21}$$

with F_n being the numerical values defined above from the overlap integral,

$$F_n \equiv \frac{2}{L} \int_0^L dx \, \sin^2\left(\frac{n\pi x}{L}\right) \sin\left(\frac{\pi x}{L}\right) = \mathcal{O}(1) , \tag{4.22}$$

Perturbative convergence criterion

For this perturbative expansion to make sense, the **first-order correction should be smaller than the unperturbed result**. Therefore, the **convergence condition** of the perturbative series is that

$$\frac{\lambda F_n V_1}{\pi^2 n^2 \hbar^2 / 2mL^2} \ll 1 \tag{4.23}$$

and therefore the condition that λ must satisfy to be considered **small** is

$$\lambda \ll \frac{\pi^2 n^2 \hbar^2}{2mL^2 F_n V_1} . \tag{4.24}$$

In particular, the stronger the size of the perturbed potential, V_1, the smaller than λ needs to be for the perturbative expansion to make sense. We also note that the **convergence condition** for the parameter λ depends on the principal quantum number n, and for very excited states the condition is less stringent than for the ground state. This makes sense, since we have seen that this potential induces approximately a constant correction to the energy levels irrespective of n but the unperturbed energy levels increase with n^2.

Finally, one can formally define the validity of the perturbative regime by the condition $\lambda = 1$, and in this case we have that

$$V_1 = \frac{\pi^2 n^2 \hbar^2}{2mL^2 F_k} \,. \tag{4.25}$$

For a given value of the principal quantum number n, if the strength of the perturbed potential V_1 is larger than this value then perturbation theory is completely unreliable and one needs to attempt to solve this problem by using other methods.

First-order corrections to the state vectors. Eq. (4.16) provided us with the result for the **first-order corrections to the energies**. We want now to evaluate the first-order correction to the state vectors using a similar strategy, and hence complete the first-order perturbative calculation for a general quantum system. The starting point will be the same relation that used for the calculation of the first-order corrections to the energies:

$$\hat{H}^{(0)}|\psi_n^{(1)}\rangle + \hat{H}'|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(1)}\rangle + E_n^{(1)}|\psi_n^{(0)}\rangle \tag{4.26}$$

which can be also be expressed as

$$\left(\hat{H}^{(0)} - E_n^{(0)}\right)|\psi_n^{(1)}\rangle = \left(E_n^{(1)} - \hat{H}'\right)|\psi_n^{(0)}\rangle \,. \tag{4.27}$$

Note that the RHS of Eq. (4.27) is known, and formally what we are dealing with is a second-order inhomogeneous differential equation for $|\psi_n^{(1)}\rangle$.

Instead of trying to solve this equation directly, I will use the fact that the unperturbed eigenvectors form a **complete basis** in this Hilbert space, and thus I know for sure that I can expand $|\psi_n^{(1)}\rangle$ as a linear combination of the $\{|\psi_j^{(0)}\rangle\}$ in the usual manner:

$$|\psi_n^{(1)}\rangle = \sum_j c_j^{(n)}|\psi_j^{(0)}\rangle \,, \tag{4.28}$$

where we have explicitely indicated that for each value of the principal quantum number n we will have a different set of coefficients $\{c_j^{(n)}\}$. This way, we have transformed the problem of determining the correction to the quantum state $|\psi_n^{(1)}\rangle$ to that of finding the corresponding coefficients $c_j^{(n)}$ of its linear expansion.

If we insert this linear expansion into Eq. (4.27), we obtain

$$\left(\hat{H}^{(0)} - E_n^{(0)}\right)\left[\sum_j c_j^{(n)}|\psi_j^{(0)}\rangle\right] = \left(E_n^{(1)} - \hat{H}'\right)|\psi_n^{(0)}\rangle \,, \tag{4.29}$$

but we know how to evaluate the LHS, given that the $|\psi_j^{(0)}\rangle$ are the eigenvectors of the unperturbed Hamiltonian with eigenvalues $E_n^{(0)}$ and hence

$$\left(\hat{H}^{(0)} - E_n^{(0)}\right)\left[\sum_j c_j^{(n)}|\psi_j^{(0)}\rangle\right] = \sum_j c_j^{(n)}\left(\hat{H}^{(0)} - E_n^{(0)}\right)|\psi_j^{(0)}\rangle = \sum_j c_j^{(n)}\left(E_j^{(0)} - E_n^{(0)}\right)|\psi_j^{(0)}\rangle \,. \tag{4.30}$$

A crucial observation, as we will discover soon, is that the term $j = n$ does not contribute to the sum since then $\left(E_j^{(0)} - E_n^{(0)}\right) = 0$, and hence we can write

$$\left(\hat{H}^{(0)} - E_n^{(0)}\right)\left[\sum_j c_j^{(n)} |\psi_j^{(0)}\rangle\right] = \sum_{j\neq n} c_j^{(n)} \left(E_j^{(0)} - E_n^{(0)}\right) |\psi_j^{(0)}\rangle \, . \tag{4.31}$$

Taking this result into account, I find that Eq. (4.29) gives me:

$$\sum_{j\neq n} c_j^{(n)} \left(E_j^{(0)} - E_n^{(0)}\right) |\psi_j^{(0)}\rangle = \left(E_n^{(1)} - \hat{H}'\right) |\psi_n^{(0)}\rangle \, , \tag{4.32}$$

For reasons that will become obvious in a second, let me now multiply this expression from the left with the bra $\langle\psi_l^{(0)}|$, which gives me:

$$\sum_{j\neq n} c_j^{(n)} \left(E_j^{(0)} - E_n^{(0)}\right) \left\langle \psi_l^{(0)} \middle| \psi_j^{(0)} \right\rangle = \left\langle \psi_l^{(0)} \middle| \left(E_n^{(1)} - \hat{H}'\right) \psi_n^{(0)} \right\rangle . \tag{4.33}$$

Due to the orthonormality properties of the unperturbed eigenstates, the LHS will only be non-zero is $l = j$:

$$\sum_{j\neq n} c_j^{(n)} \left(E_j^{(0)} - E_n^{(0)}\right) \left\langle \psi_l^{(0)} \middle| \psi_j^{(0)} \right\rangle = \sum_{j\neq n} c_j^{(n)} \left(E_j^{(0)} - E_n^{(0)}\right) \delta_{lj} = c_l^{(n)} \left(E_l^{(0)} - E_n^{(0)}\right) \quad (l \neq n) \tag{4.34}$$

Using this result, Eq. (4.33) simplifies to

$$c_l^{(n)} \left(E_l^{(0)} - E_n^{(0)}\right) = \left\langle \psi_l^{(0)} \middle| \left(E_n^{(1)} - \hat{H}'\right) \psi_n^{(0)} \right\rangle = -\left\langle \psi_l^{(0)} \middle| \hat{H}' \psi_n^{(0)} \right\rangle \qquad (l \neq n) \, , \tag{4.35}$$

where in the last step we have used that $l \neq n$ and hence the term proportional to $E_n^{(1)}$ cancels out due to orthogonality.

But at this point we are basically done, since we have now determined the sought-for coefficients of the linear expansion of the first-order corrections to the state vectors defined in Eq. (4.28) as

$$c_l^{(n)} = \frac{\left\langle \psi_l^{(0)} \middle| \hat{H}' \psi_n^{(0)} \right\rangle}{E_n^{(0)} - E_l^{(0)}} \qquad (l \neq n) \, . \tag{4.36}$$

Note that we have all the information to evaluate them, since these are standard matrix elements of the perturbed Hamiltonian $\hat{H}'$ evaluated in the unperturbed states. Finally, we find that the **first-order corrections** to the state vectors are given by

$$|\psi_n^{(1)}\rangle = \sum_{j\neq n} c_j^{(n)} |\psi_j^{(0)}\rangle = \sum_{j\neq n} \frac{\left\langle \psi_j^{(0)} \middle| \hat{H}' \psi_n^{(0)} \right\rangle}{E_n^{(0)} - E_j^{(0)}} \left| \psi_j^{(0)} \right\rangle . \tag{4.37}$$

Note that in this expression the denominator is only non-zero (and hence the expression for $|\psi_n^{(1)}\rangle$ well defined) provided that the unperturbed spectrum is **non-degenerate**.

If instead the perturbed spectrum is degenerate, with two or more different eigenvectors $j \neq n$ sharing the same energy eigenvalue, $E_j^{(0)} = E_n^{(0)}$, then we are in trouble since the denominator will vanish for some elements of the sum rendering the calculation meaningless. We will discuss below how to tackle this situation, which will require the formulation of **degenerate perturbation theory**.

The perturbed infinite well revisited. Let me compute now the first-order perturbed wave functions corresponding to the perturbed infinite well potential that we considered above, and where the unperturbed results were given by Eq. (4.17). For simplicity, I will limit this calculation to the ground state, and hence what I need to compute is

$$|\psi_1^{(1)}\rangle = \sum_{j\neq 1} \frac{\left\langle \psi_j^{(0)} \middle| \hat{H}'\psi_1^{(0)} \right\rangle}{E_1^{(0)} - E_j^{(0)}} \left| \psi_j^{(0)} \right\rangle . \tag{4.38}$$

The energy difference is

$$E_1^{(0)} - E_j^{(0)} = \frac{\pi^2\hbar^2}{2mL^2} - \frac{\pi^2 j^2\hbar^2}{2mL^2} = -(j^2-1)\frac{\pi^2\hbar^2}{2mL^2} \tag{4.39}$$

and hence I can write, going back to the wavefunction notation which is more convenient to this problem,

$$\psi_1^{(1)}(x) = -\frac{2mL^2}{\pi^2\hbar^2}\sqrt{\frac{2}{L}} \sum_{j\neq 1} \frac{\left\langle \psi_j^{(0)} \middle| \hat{H}'\psi_1^{(0)} \right\rangle}{(j^2-1)} \sin\left(\frac{j\pi x}{L}\right) , \tag{4.40}$$

where I have moved all the prefactors outside the sum. The overlap integrals with the perturbed Hamiltonian are given by:

$$\left\langle \psi_j^{(0)} \middle| \hat{H}'\psi_1^{(0)} \right\rangle = \frac{2V_1}{L}\int_0^L dx\, \sin\left(\frac{j\pi x}{L}\right)\sin^2\left(\frac{\pi x}{L}\right) , \tag{4.41}$$

and these integrals can be computed, for example the first few terms in the sum give

$$\begin{aligned} \left\langle \psi_2^{(0)} \middle| \hat{H}'\psi_1^{(0)} \right\rangle &= 0 , \\ \left\langle \psi_3^{(0)} \middle| \hat{H}'\psi_1^{(0)} \right\rangle &= -\frac{8}{15\pi}V_1 , \\ \left\langle \psi_4^{(0)} \middle| \hat{H}'\psi_1^{(0)} \right\rangle &= 0 , \\ \left\langle \psi_5^{(0)} \middle| \hat{H}'\psi_1^{(0)} \right\rangle &= -\frac{8}{105\pi}V_1 . \end{aligned} \tag{4.42}$$

and hence we find that the first-order correction to the wave function of the ground state in this problem is given by:

$$\psi_1^{(1)}(x) = \frac{2mL^2V_1}{\pi^2\hbar^2}\sqrt{\frac{2}{L}}\left(\frac{1}{15\pi}\sin\left(\frac{3\pi x}{L}\right) + \frac{1}{3}\frac{1}{105\pi}\sin\left(\frac{5\pi x}{L}\right) + \ldots .\right) \tag{4.43}$$

If you represent graphically these results, you can verify that the sum over j converges rather quickly, and thus one does not need that many terms in order to produce a sensible estimate for $\psi_1^{(1)}(x)$.

4.1.2 Second-order perturbation theory

In the same manner, it is possible to compute the second-order corrections to the energy and to the eigenvectors (and eventually also at third order, then at fourth order, and so on). Let us show the explicit result for the second-order correction to the energy, $E_n^{(2)}$. The starting point will be Eq. (4.11),

$$\hat{H}^{(0)}\left|\psi_n^{(2)}\right\rangle + \hat{H}'\left|\psi_n^{(1)}\right\rangle = E_n^{(0)}\left|\psi_n^{(2)}\right\rangle + E_n^{(1)}\left|\psi_n^{(1)}\right\rangle + E_n^{(2)}\left|\psi_n^{(0)}\right\rangle m \tag{4.44}$$

and let me now multiply by the bra $\left\langle\psi_n^{(0)}\right|$ from the left:

$$\left\langle \psi_n^{(0)} \middle| \hat{H}^{(0)} \middle| \psi_n^{(2)} \right\rangle + \left\langle \psi_n^{(0)} \middle| \hat{H}' \middle| \psi_n^{(1)} \right\rangle = E_n^{(0)}\left\langle \psi_n^{(0)} \middle| \psi_n^{(2)} \right\rangle + E_n^{(1)}\left\langle \psi_n^{(0)} \middle| \psi_n^{(1)} \right\rangle + E_n^{(2)}\left\langle \psi_n^{(0)} \middle| \psi_n^{(0)} \right\rangle . \tag{4.45}$$

By exploiting the Hermiticity of $\hat{H}$, you can see how the first term in the LHS and in the RHS of the equation cancel with each other. Therefore we end up with the following relation

$$\left\langle \psi_n^{(0)} \middle| \hat{H}' \middle| \psi_n^{(1)} \right\rangle = E_n^{(1)} \left\langle \psi_n^{(0)} \middle| \psi_n^{(1)} \right\rangle + E_n^{(2)} \tag{4.46}$$

and hence the second-order correction to the unperturbed energies, $E_n^{(2)}$, is given by

$$E_n^{(2)} = \left\langle \psi_n^{(0)} \middle| \hat{H}' \middle| \psi_n^{(1)} \right\rangle - E_n^{(1)} \left\langle \psi_n^{(0)} \middle| \psi_n^{(1)} \right\rangle . \tag{4.47}$$

Crucially, the calculation of $E_n^{(2)}$ depends only on quantities that we have evaluated already, and hence it should be conceptually (though often not technically) straightforward to evaluate.

In particular, we can evaluate the matrix element $\left\langle \psi_n^{(0)} \middle| \psi_n^{(1)} \right\rangle$ based on the results we have obtained above in Eq. (4.37):

$$\left\langle \psi_n^{(0)} \middle| \psi_n^{(1)} \right\rangle = \left\langle \psi_n^{(0)} \middle| \sum_{j \neq n} \frac{\left\langle \psi_j^{(0)} \middle| \hat{H}' \psi_n^{(0)} \right\rangle}{E_n^{(0)} - E_j^{(0)}} \middle| \psi_j^{(0)} \right\rangle = 0 \tag{4.48}$$

since the only value of j which would make the matrix element non-zero, namely $j = n$, is not included in the sum. Therefore,

$$E_n^{(2)} = \left\langle \psi_n^{(0)} \middle| \hat{H}' \middle| \psi_n^{(1)} \right\rangle = \left\langle \psi_n^{(0)} \middle| \hat{H}' \sum_{j \neq n} \frac{\left\langle \psi_j^{(0)} \middle| \hat{H}' \psi_n^{(0)} \right\rangle}{E_n^{(0)} - E_j^{(0)}} \middle| \psi_j^{(0)} \right\rangle = \sum_{j \neq n} \frac{\left| \left\langle \psi_j^{(0)} \middle| \hat{H}' \psi_n^{(0)} \right\rangle \right|^2}{E_n^{(0)} - E_j^{(0)}} \tag{4.49}$$

is the sought-for result for the **second-order corrections to the energies**. Note that the overlap integrals are the same as those appearing in the first-order correction to the state vector, Eq. (4.37), and hence we should have evaluated them already!

The perturbed energies at $\mathcal{O}\left(\lambda^2\right)$

Putting together our calculations, we can express our estimate for the energy E_n of the full Hamiltonian of the problem up to second-order in the perturbative expansion as follows

$$E_n = E_n^{(0)} + \lambda \left\langle \psi_k^{(0)} \middle| \hat{H}' \psi_k^{(0)} \right\rangle + \lambda^2 \sum_{j \neq n} \frac{\left| \left\langle \psi_j^{(0)} \middle| \hat{H}' \psi_n^{(0)} \right\rangle \right|^2}{E_n^{(0)} - E_j^{(0)}} + \mathcal{O}(\lambda^3) . \tag{4.50}$$

Interestingly, note that the fact that perturbation theory works at the first order does not guarantee that it will also work at the second order. For example, if we have a **quasi-continuum energy spectrum** with many values of $E_j^{(0)}$ very close to $E_n^{(0)}$, then the $\mathcal{O}\left(\lambda^2\right)$ correction may receive an enhancement that could spoil the convergence of the perturbative expansion.

4.2 Degenerate perturbation theory

The perturbation theory results that we have derived so far have assumed that the unperturbed system (defined by the original Hamiltonian $H^{(0)}$) is **non degenerate**, that is, that each eigenvector of the Hamiltonian has associated a distinct energy. If this is not the case, we quickly get into problems, as illustrated by the expression for the first-order corrections to the state vector Eq. (4.37), which diverges if the energy spectrum has some degeneracy (even a two-fold degeneracy makes the previous calculation unusable).

However, we know that many important quantum systems exhibit some level of energy degeneracy: for example, a particle in a two-dimensional box, the hydrogen atom (states with the same n but different values of l and m are degenerate), or the situation where we have two distinguishable particles confided into a one dimensional potential well. This situation demands to formulate an alternative approach to perturbation theory which works also if the unperturbed system exhibits a degenerate spectrum. We will develop this framework now, which we will denote as **degenerate perturbation theory**. First of all we will present the explicit calculation for the case of a two-fold degeneracy, and then explain how this calculation can be extended to an n-fold degeneracy.

Two-fold degeneracy. Let us start with a relatively simple quantum system, characterised by a unperturbed Hamiltonian $H^{(0)}$ with two orthonormal eigenvectors that share the same energy. This system is therefore defined by the following relations:

$$\hat{H}^{(0)}\big|\psi_a^{(0)}\big\rangle = E^{(0)}\big|\psi_a^{(0)}\big\rangle\,, \quad \hat{H}^{(0)}\big|\psi_b^{(0)}\big\rangle = E^{(0)}\big|\psi_b^{(0)}\big\rangle\,, \quad \big\langle\psi_b^{(0)}\big|\psi_a^{(0)}\big\rangle = 0\,. \tag{4.51}$$

We denote such situation as a **two-fold degeneracy**, given that we have two distinct eigenvalues which share the same energy.

An important property of this configuration is that **any linear combination of the two eigenvectors** is also an eigenvector of the Hamiltonian, with of course the same energy. That is, if we have a general linear combination given by

$$\big|\psi^{(0)}\big\rangle \equiv c_a\big|\psi_a^{(0)}\big\rangle + c_b\big|\psi_b^{(0)}\big\rangle\,, \tag{4.52}$$

with c_a, c_b being arbitrary complex coefficients, this state will also be an eigenstate of the original unperturbed Hamiltonian, given that

$$\hat{H}^{(0)}\big|\psi^{(0)}\big\rangle = \hat{H}^{(0)}\left(c_a\big|\psi_a^{(0)}\big\rangle + c_b\big|\psi_b^{(0)}\big\rangle\right) = c_aE^{(0)}\big|\psi_a^{(0)}\big\rangle + c_bE^{(0)}\big|\psi_b^{(0)}\big\rangle = E^{(0)}\big|\psi^{(0)}\big\rangle\,, \tag{4.53}$$

with the same value of the energy $E^{(0)}$ as the two eigenstates.

Assume now that we add a small perturbation to our Hamiltonian (small, since else perturbation theory would not be applicable), which we denote as

$$\hat{H} = \hat{H}^{(0)} + \lambda\hat{H}'\,, \tag{4.54}$$

where, as in the case of non-degenerate perturbation theory, we explicitely factorise the small dimensionless parameter λ (such that $\hat{H}'$ still has units of energy). In general, we will find that adding $\hat{H}'$ to the Hamiltonian of the system **breaks the degeneracy** of the unperturbed system, in that the perturbed energies associated to the perturbed states $\big|\psi_a\big\rangle$ and $\big|\psi_b\big\rangle$ will become different, $E_a \neq E_b$ once the $\mathcal{O}(\lambda)$ corrections have been calculated.

Conversely, if we start from the perturbed state vectors $\big|\psi_a\big\rangle$ and $\big|\psi_b\big\rangle$ and take the limit $\lambda \to 0$, we will have that $E_a, E_b \to E^{(0)}$ but in general the resulting state vectors in this limit will be **different** from the unperturbed eigenvectors $\big|\psi_a^{(0)}\big\rangle$ and $\big|\psi_b^{(0)}\big\rangle$, and will rather be a linear combination of them. This is not unexpected since any linear combination Eq. (4.52) represents an equally good eigenvector as our original choice in the presence of degeneracy, so we have some **intrinsic ambiguity**.

To make sense of perturbation theory, we need first of all to identify the linear combinations Eq. (4.52) to which the perturbed state vectors $\big|\psi_a\big\rangle$ and $\big|\psi_b\big\rangle$ reduce in the $\lambda \to 0$ limit - else I do not know with respect to what I am evaluating the perturbative correction.

Finding the degenerate unperturbed state vectors

We define the **degenerate unperturbed state vectors** $|\psi_1\rangle$ and $|\psi_2\rangle$ as those state vectors, eigenvectors of the unperturbed Hamiltonian $\hat{H}^0$, which are the $\lambda \to 0$ limit of the associated perturbed state vectors. In other words, we have that

$$\lim_{\lambda\to 0} |\psi_a\rangle = |\psi_1\rangle \qquad \text{(with energy } E_a \to E^{(0)}) \,, \tag{4.55}$$

$$\lim_{\lambda\to 0} |\psi_b\rangle = |\psi_2\rangle \qquad \text{(with energy } E_b \neq E_a \to E^{(0)}) \,, \tag{4.56}$$

where both $|\psi_1\rangle$ and $|\psi_2\rangle$ have associated energies $E^{(0)}$ and are constructed as linear combinations following the procedure of Eq. (4.52). These degenerate unperturbed state vectors are referred as "good" states in the textbook, which is perhaps not the most self-explanatory terminology that could be used.

Clearly, the applicability of degenerate perturbation theory requires first of all to determine what are these **degenerate unperturbed state vectors** to then subsequently evaluate their perturbative corrections. Let us illustrate this concept with a explicit example.

The perturbed harmonic oscillator in 2D. Let us consider here a quantum harmonic oscillator in two dimensions. The unperturbed Hamiltonian is given by

$$\hat{H}^{(0)} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + \frac{1}{2}m\omega^2\left(x^2+y^2\right) \,. \tag{4.57}$$

We can now perturb this system by adding a small correction to the Hamiltonian,

$$\lambda\hat{H}' = \lambda m\omega^2 xy \,, \tag{4.58}$$

which induces as one can see a **coupling** between the oscillator modes in the x direction and those in the y direction, which in the regular 2D quantum harmonic oscillator are fully independent and can be treated separately as two 1D oscillators.

The ground state of the unperturbed system ($n_x = n_y = 0$) is non-degenerate and has energy $E = \hbar\omega$. However, the **first excited state** will be degenerate, since the two different eigenfunctions with ($n_x = 1, n_y = 0$) and ($n_x = 0, n_y = 1$) will have the same energy, $E = 2\hbar\omega$. Clearly, this is a situation for which non-degenerate perturbation theory breaks down.

One **possible basis** for these two degenerate states will be given by the product of the x-direction and y-direction wave functions with quantum numbers ($n_x = 1, n_y = 0$) and ($n_x = 0, n_y = 1$), namely

$$\psi_a^{(0)}(x,y) = \psi_0(x)\psi_1(y) = \sqrt{\frac{2}{\pi}}\frac{m\omega}{\hbar} y e^{-m\omega(x^2+y^2)/2\hbar} \,, \tag{4.59}$$

$$\psi_b^{(0)}(x,y) = \psi_1(x)\psi_0(y) = \sqrt{\frac{2}{\pi}}\frac{m\omega}{\hbar} x e^{-m\omega(x^2+y^2)/2\hbar} \,. \tag{4.60}$$

However, as mentioned above this is by no means a unique choice, and any pair of normalised orthogonal linear combinations

$$\psi_1^{(0)}(x,y) = c_a^{(1)}\psi_a^{(0)}(x,y) + c_b^{(1)}\psi_b^{(0)}(x,y)\,, \quad \psi_2^{(0)}(x,y) = c_a^{(2)}\psi_a^{(0)}(x,y) + c_b^{(2)}\psi_b^{(0)}(x,y)\,, \tag{4.61}$$

would represent an equally good basis, given that these linear combinations are also degenerate eigenstates of my Hamiltonian and have the **same energy** $E = 2\hbar\omega$. If we are to apply perturbation theory, we need to find the **degenerate unperturbed eigenfunctions**, that is, the combinations $\psi_1^{(0)}(x,y)$ and $\psi_2^{(0)}(x,y)$ which correspond to the $\lambda \to 0$ limit of the perturbed wavefunctions.

This example is particularly interesting in that the perturbed state can be **solved exactly** by rotating to a coordinate system such that

$$x' = \frac{x+y}{\sqrt{2}}\,, \quad y' = \frac{x-y}{\sqrt{2}}\,, \tag{4.62}$$

where you can show that the Hamiltonian is transformed as

$$\hat{H}^{(0)} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial y'^2}\right) + \frac{1}{2}m(1+\lambda)\omega^2\left(x'\right)^2 + \frac{1}{2}m(1-\lambda)\omega^2\left(y'\right)^2\,. \tag{4.63}$$

which is of course nothing but two independent (decoupled) harmonic oscillators with the frequency $\omega_\pm = \omega\sqrt{1\pm\lambda}$. Here again we see that λ must be a small number: for $\lambda \geq 1$ the system stops even being a system of harmonic oscillators and becomes something completely different. This is another way of visualizing the **convergence criterion** of the perturbative expansion.

Expressing the perturbed Hamiltonian in this manner, it is clear that the **exact solutions** for the two-body system wave functions are

$$\psi_{nl}(x',y') = \psi_n^+(x')\psi_l^-(y')\,, \tag{4.64}$$

where ψ_n^+ is the n-th eigenfunction of a 1D harmonic oscillator with frequency $\omega_+ = \omega\sqrt{1+\lambda}$ and ψ_l^- is the l-th eigenfunction of a 1D harmonic oscillator with frequency $\omega_- = \omega\sqrt{1-\lambda}$. The energies associated to these exact solutions to the perturbed problem are given by:

$$E_{nl} = \hbar\omega_+\left(n+\frac{1}{2}\right) + \hbar\omega_-\left(l+\frac{1}{2}\right) = \hbar\omega\left(\sqrt{1+\lambda}\left(n+\frac{1}{2}\right) + \sqrt{1-\lambda}\left(l+\frac{1}{2}\right)\right)\,. \tag{4.65}$$

Crucially, we see that the exact energies of the perturbed problem are now **non-degenerate**: every independent combination of the quantum numbers n, l will have associated a different value of the energy E_{nl}. Furthermore, in the limit $\lambda \to 0$ we recover the original unperturbed quantum system and in this case the energies as we knew are degenerate, since

$$\lim_{\lambda\to 0} E_{nl} = \hbar\omega\left(n+l+1\right)\,, \tag{4.66}$$

and hence any combination of n and l for which $n+l$ is constant will be degenerate and will have associated the same energy.

In Fig. 4.1 we display the exact energies of the perturbed 2D harmonic oscillator problem, Eq. (4.65), as a function of the expansion parameter λ for two combinations of quantum numbers that lead to degenerate energies in the unperturbed Hamiltonian, namely $(n,l) = (4,1)$ and $(n,l) = (1,4)$ (left) and $(n,l) = (3,2)$ and $(n,l) = (2,3)$ (right). One can observe how in the $\lambda \to 0$ limit one recovers the degenerate unperturbed energies of Eq. (4.66), and how the perturbative correction breaks this degeneracy for $\lambda > 0$. Actually, note that beyond the first excited level of the system the degeneracy is larger than the two-fold one we are

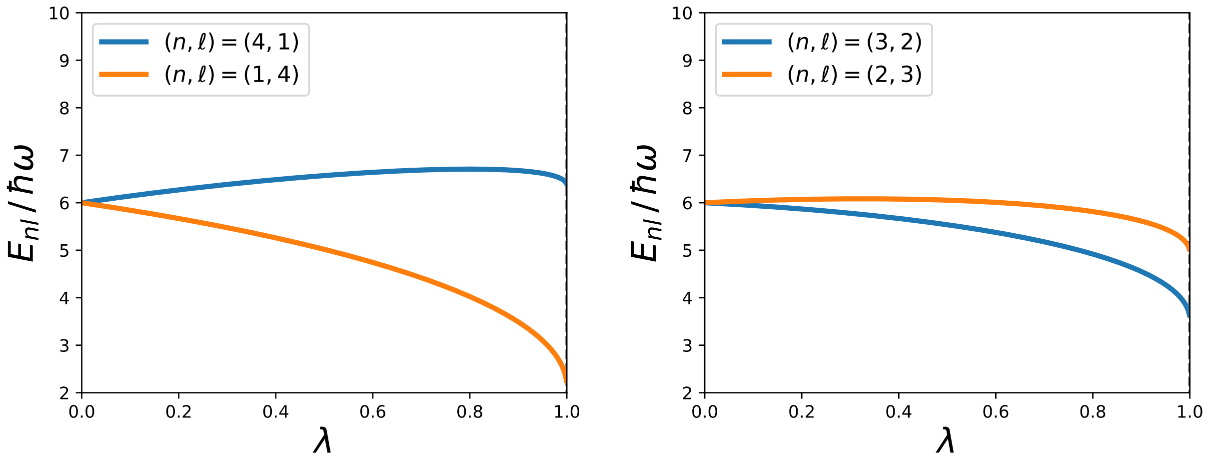

Figure 4.1: The exact energies of the perturbed 2D harmonic oscillator problem, Eq. (4.65), as a function of the expansion parameter λ for two combinations of quantum numbers that lead to degenerate energies in the unperturbed Hamiltonian, namely $(n,l)=(4,1)$ and $(n,l)=(1,4)$ (left) and $(n,l)=(3,2)$ and $(n,l)=(2,3)$ (right). One can observe how in the $\lambda\to 0$ limit one recovers the degenerate unperturbed energies of Eq. (4.66), and how the perturbative correction breaks this degeneracy for $\lambda>0$.

interested here, so we will focus here only on the first excited state.

Since it this case we know what are the exact wave functions that solve the perturbed problem, Eq. (4.64), we can take the $\lambda\to 0$ limit and find the relevant **degenerate unperturbed eigenfunctions** to be used for the perturbative calculation of the energies. Rotating back to the original coordinate system (x,y), we obtain the following result:

$$\begin{aligned}\lim_{\lambda\to 0}\psi_{01}(x,y) &= \lim_{\lambda\to 0}\psi_0^+\left(\frac{x+y}{\sqrt{2}}\right)\psi_1^-\left(\frac{x-y}{\sqrt{2}}\right) &(4.67)\\ &= \sqrt{\frac{2}{\pi}\frac{m\omega}{\hbar}}\frac{x-y}{2}e^{-m\omega(x^2+y^2)/2\hbar}=\frac{1}{\sqrt{2}}\left(-\psi_a^{(0)}+\psi_b^{(0)}\right) &(4.68)\end{aligned}$$

and likewise for the orthogonal combination:

$$\lim_{\lambda\to 0}\psi_{10}(x,y)=\frac{1}{\sqrt{2}}\left(\psi_a^{(0)}(x,y)+\psi_b^{(0)}(x,y)\right) \tag{4.69}$$

Therefore, we find that, for this specific problem, the “good” states of the unperturbed Hamiltonian, to be used in a perturbative calculation, are given by:

$$\psi_\pm^{(0)}(x,y)=\frac{1}{\sqrt{2}}\left(\psi_b^{(0)}(x,y)\pm\psi_a^{(0)}(x,y)\right)\,, \tag{4.70}$$

that is, these represent the $\lambda\to 0$ limit of the eigenstates of the **perturbed** Hamiltonian. Thus they are the correct baseline upon which to evaluate the perturbative corrections to the energy and the state vector of the system.

As will be shown below, if we know what are these “good” degenerate unperturbed eigenfunctions we can apply perturbation theory without the fear of obtaining unphysical results.

Degenerate perturbation theory. In general, we will not know in advance what are the suitable degenerate unperturbed eigenfunctions to be used in the perturbative calculation. For this reason, in the following we will keep the coefficients of the linear combination α and β as free parameters

$$\big|\psi^{(0)}\big\rangle = \alpha\big|\psi_a^{(0)}\big\rangle + \big|\beta\psi_b^{(0)}\big\rangle\,, \tag{4.71}$$

and show how we can determine their values from a direct calculation.

As in the case of non-degenerate perturbation theory, our goal is to take the perturbed Hamiltonian

$$\hat{H} = \hat{H}^{(0)} + \lambda\hat{H}'\,, \tag{4.72}$$

where λ is a small adimensional parameter, and express both the new state and the corresponding energies in terms of a **series expansion** in λ, starting from the original (unperturbed) values:

$$\begin{aligned} \big|\psi\big\rangle &= \big|\psi^{(0)}\big\rangle + \lambda\big|\psi^{(1)}\big\rangle + \lambda^2\big|\psi^{(2)}\big\rangle + \mathcal{O}\left(\lambda^3\right)\,, \\ E &= E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \mathcal{O}\left(\lambda^3\right)\,. \end{aligned} \tag{4.73}$$

Note that here we have suppressed the indices, since we are assuming the specific scenario of a **two-fold degeneracy** as mentioned above. While $E^{(0)}$ is degenerate, the perturbed Hamiltonian will break the degeneracy and we will obtain two distinct solutions for the perturbed energy E.

We can now insert the series expansions in λ in the Schroedinger equation, expand and collect terms that are proportional to the same power of the expansion parameter λ. The first non-trivial relation that we find is the one that determines the **first-order corrections to the energy**, that is

$$\hat{H}^{(0)}\big|\psi^{(1)}\big\rangle + \hat{H}'\big|\psi^{(0)}\big\rangle = E^{(0)}\big|\psi^{(1)}\big\rangle + E^{(1)}\big|\psi^{(1)}\big\rangle\,, \tag{4.74}$$

which is nothing but the same equation we had in the non-degenerate case, Eq. (4.12). Now, as opposed to the case of non-degenerate perturbation theory, we will multiply this expression from the left with the bra $\big\langle\psi_a^{(0)}\big|$ instead of by the original unperturbed eigenvector $\big\langle\psi^{(0)}\big|$ (for the reason that we don't even really know how to select $\big\langle\psi^{(0)}\big|$ at this point!)

By doing this, we end up with the following expression

$$\left\langle\psi_a^{(0)}\middle|\hat{H}^{(0)}\psi^{(1)}\right\rangle + \left\langle\psi_a^{(0)}\middle|\hat{H}'\psi^{(0)}\right\rangle = E^{(0)}\left\langle\psi_a^{(0)}\middle|\psi^{(1)}\right\rangle + E^{(1)}\left\langle\psi_a^{(0)}\middle|\psi^{(0)}\right\rangle \tag{4.75}$$

Now using the Hermiticity of $\hat{H}^{(0)}$, and introducing the expansion that defines the "good" eigenvectors (which remember, we are trying to determine here),

$$\left|\psi^{(0)}\right\rangle = \alpha\left|\psi_a^{(0)}\right\rangle + \beta\left|\psi_b^{(0)}\right\rangle\,, \tag{4.76}$$

we end up with the following result:

$$\alpha\left\langle\psi_a^{(0)}\middle|\hat{H}'\psi_a^{(0)}\right\rangle + \beta\left\langle\psi_a^{(0)}\middle|\hat{H}'\psi_b^{(0)}\right\rangle = \alpha E^{(1)} \tag{4.77}$$

This result can be express in a more compact manner in terms of the **matrix elements** of the perturbed Hamiltonian, defined as

$$W_{ij} \equiv \alpha\left\langle\psi_i^{(0)}\middle|\hat{H}'\psi_j^{(0)}\right\rangle\,, \quad i,j = a,b\,. \tag{4.78}$$

Taking into account this definition, and repeating the calculation now multiplying Eq. (4.12) from the left with the bra $\langle\psi_b^{(0)}|$, we end up with the following **matrix equation**:

$$\begin{pmatrix} W_{aa} & W_{ab} \\ W_{ba} & W_{bb} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E^{(1)} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} , \tag{4.79}$$

which is nothing but the **eigenvalue equation** for the matrix $\mathbf{W}$. By solving this eigenvalue equation we obtain both the first-order perturbative corrections to the energy, $E^{(1)}$ and well as the eigenvectors α and β which define the "good" states of the unperturbed Hamiltonian. Note that as expected we will have two solutions for $E^{(1)}$ and two for the vector (α, β) which will define the two sought-for "good" combinations.

The solution of this eigenvalue equation is provided by the **characteristic equation** of the eigenvalue problem, given by

$$\begin{vmatrix} W_{aa} - E^{(1)} & W_{ab} \\ W_{ba} & W_{bb} - E^{(1)} \end{vmatrix} = 0 \tag{4.80}$$

If we use that the matrix elements satisfy $W_{ba} = W_{ab}^*$ (because of the hermiticity of the Hamiltonian), then we find that the solutions of the resulting quadratic equation will be given as follows:

$$E_{\pm}^{(1)} = \frac{1}{2}\left[W_{aa} + W_{bb} \pm \sqrt{(W_{aa} - W_{bb})^2 + 4|W_{ab}|^2}\right] . \tag{4.81}$$

Again, recall that W_{aa} and W_{bb} are real because of hermiticity. Hence, provided we evaluate the matrix elements W_{ij}, we should able to evaluate the first order corrections to the energy and to determine the "good" states of the unperturbed Hamiltonian.

Degenerate perturbation theory

Eq. (4.81) is the fundamental results of two-fold degenerate perturbation theory: the two roots correspond to the two perturbed energies $E_{\pm}^{(1)}$. We hence see how the perturbed Hamiltonian **lifts the degeneracy** which is present at the unperturbed level. Note also that in the case where $W_{ab} = 0$ then the eigenvalues are trivially

$$E_{+}^{(1)} = W_{aa} = \left\langle \psi_a^{(0)} \middle| \hat{H}' \psi_a^{(0)} \right\rangle , \qquad E_{+}^{(2)} = W_{bb} = \left\langle \psi_b^{(0)} \middle| \hat{H}' \psi_b^{(0)} \right\rangle \tag{4.82}$$

which is the result obtained using **non-degenerate perturbation theory**, Eq. (4.16).

Clearly, if we can start from the "good" states of the unperturbed Hamiltonian, then we would have a nice head start since we could directly use non-degenerate perturbative calculations.

The perturbed 2D harmonic oscillator reloaded. Let us verify that these findings are consistent with our previous calculation for the perturbed 2D harmonic oscillator. To do this, we need to evaluate the matrix elements W_{ij} and then solve the corresponding eigenvalue equation. You can convince yourselves that

$$W_{aa} = \int\int \psi_a^{(0)}(x,y)\hat{H}'\psi_a^{(0)}(x,y)dxdy = \omega^2 \int dx\, x\, |\psi_0(x)|^2 \int dy\, y\, |\psi_0(y)|^2 = 0 \tag{4.83}$$

since the ground state wave functions are symmetric in their argument. For the same reasoning, $W_{bb} = 0$. The only non-trivial matrix element is the off-diagonal one:

$$W_{aa} = \int\int \psi_a^{(0)}(x,y)\hat{H}'\psi_b^{(0)}(x,y)dxdy = \omega^2 \int dx\, x\, \psi_0(x)\psi_1(x) \int dy\, y\, \psi_0(y)\psi_1(y) = \frac{\hbar\omega}{2} \,, \tag{4.84}$$

which can be evaluated for example using the explicit one-particle wave functions of the unperturbed Harmonic oscillator problem.

Hence the matrix which is to be diagonalised turns out to be

$$\begin{pmatrix} W_{aa} & W_{ab} \\ W_{ba} & W_{bb} \end{pmatrix} = \frac{\hbar\omega}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \tag{4.85}$$

which can be easily diagonalised and whose eigenvalues and eigenvectors turn out to be

$$|\psi_\pm^{(0)}\rangle = \frac{1}{\sqrt{2}}\left(|\psi_a^{(0)}\rangle \pm |\psi_b^{(0)}\rangle\right) \tag{4.86}$$

$$E_\pm^{(1)} = \pm\frac{\hbar\omega}{2} \,, \tag{4.87}$$

consistently with the results that we found before with the full analytic solution to this problem. To see this, take the exact solution for the energies, Eq. (4.65), and expand for the first order in λ:

$$\begin{aligned} E_{nl} &= \hbar\omega\left(\sqrt{1+\lambda}\left(n+\frac{1}{2}\right) + \sqrt{1-\lambda}\left(l+\frac{1}{2}\right)\right) . \\ &\simeq \hbar\omega\left(\left(1+\frac{\lambda}{2}\right)\left(n+\frac{1}{2}\right) + \left(1-\frac{\lambda}{2}\right)\left(l+\frac{1}{2}\right)\right) = \hbar\omega(n+l+1) + \lambda\hbar\omega\frac{(n-l)}{2} \,. \end{aligned} \tag{4.88}$$

Now, here we were considering the two-fold degeneracy associated to $(n,l) = (1,0)$ and $(n,l) = (0,1)$, so the correction is

$$\delta E = \lambda \times \left(\pm\frac{\hbar\omega}{2}\right) , \tag{4.89}$$

exactly as we wanted to demonstrate.

The “good” states theorem

A systematic strategy to find the “good” states to be used in degenerate perturbation theory applications is the following. Consider an hermitian operator $\hat{A}$ such that $\left[\hat{A}, \hat{H}^{(0)}\right] = \left[\hat{A}, \hat{H}'\right] = 0$. This operator is such that the degenerate eigenvectors of $\hat{H}^{(0)}$, namely $|\psi_a^{(0)}\rangle$ and $|\psi_b^{(0)}\rangle$ are also eigenvectors of this operator, but this time with distinct eigenvalues, that is

$$\hat{A}|\psi_a^{(0)}\rangle = \mu|\psi_a^{(0)}\rangle \,, \quad \hat{A}|\psi_b^{(0)}\rangle = \nu|\psi_b^{(0)}\rangle \qquad \mu \neq \nu \,. \tag{4.90}$$

Then $|\psi_a^{(0)}\rangle$ and $|\psi_b^{(0)}\rangle$ correspond to the **“good” states** to be used in the perturbative calculation.

Generalisation to higher order degeneracy. The previous derivation can be generalised to deal with an n-fold degeneracy using the same strategy. For example, let us consider the case of the **three-fold degeneracy**. Consider we have an unperturbed Hamiltonian $\hat{H}^{(0)}$ with three eigenvectors with identical energy, say $\left|\psi_a^{(0)}\right\rangle$, $\left|\psi_b^{(0)}\right\rangle$, and $\left|\psi_c^{(0)}\right\rangle$. To determine the first-order perturbative corrections to the unperturbed energy $E^{(0)}$, we need to evaluate the matrix elements of the perturbed Hamiltonian

$$W_{ij} \equiv \left\langle \psi_i^{(0)} \middle| \hat{H}' \psi_j^{(0)} \right\rangle , \quad i,j = a,b,c \,, \tag{4.91}$$

and then solve the corresponding **eigenvalue equation**:

$$\begin{pmatrix} W_{aa} & W_{ab} & W_{ac} \\ W_{ba} & W_{bb} & W_{bc} \\ W_{ca} & W_{cb} & W_{cc} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = E^{(1)} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} . \tag{4.92}$$

to determine the "good
states of the unperturbed Hamiltonian, defined in terms of the linear combination

$$|\psi^{(0)}\Big\rangle = \alpha|\psi_a^{(0)}\Big\rangle + \beta|\psi_b^{(0)}\Big\rangle + \gamma|\psi_c^{(0)}\Big\rangle \tag{4.93}$$

The same method generalizes to the case of an n**-fold degeneracy**.

4.3 The fine structure of hydrogen

We now have the tools to deploy perturbation theory to quantify the **fine structure** of the hydrogen atom. By this term, we refer to a number of subleading effects that modify the energy levels and the wave functions of the electron in the hydrogen atom with respect to those that were computed in Sect. 2.3, and that are referred as the **Bohr energy levels**:

$$E_n = -\left[\frac{m_e}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2\right]\frac{1}{n^2} = \frac{E_1}{n^2} = -\frac{13.6\,\mathrm{eV}}{n^2} \,, \tag{4.94}$$

where −13.6 eV indicate the ionisation energy of hydrogen. These corrections are numerically subleading but highly relevant for many applications in atomic physics, which require a precise prediction for the energy levels of the hydrogen atom.

This fine structure of the hydrogen atom is composed by two effects, both of which are suppressed by a factor α^2 as compared to the Bohr energies, where the **fine structure constant**, the fundamental parameter of the quantum theory of electromagnetism, is a **dimensionless parameter** defined as

$$\alpha \equiv \frac{e^2}{4\pi\epsilon_0\hbar c} \simeq \frac{1}{137.036} \,. \tag{4.95}$$

Note that in terms of the fine structure constant α, the Bohr energy levels of the hydrogen atom are given by

$$E_n = -\frac{\alpha^2 m_e c^2}{2n^2} \,, \tag{4.96}$$

and as we will show below, the effects that will evaluate now scale as $E^{(1)} = \mathcal{O}(\alpha^4)$.

The two component of this **fine structure** of the hydrogen atom are:

- The **relativistic correction**, which accounts for the the fact that electrons move at energies not that far from the speed of light, and therefore one needs to consider special relativity effects.

- The **spin-orbit coupling**, which accounts for the interaction of the electron spin with its orbital motion (technically, with the magnetic field generated by the proton which "orbits" around the electron from the viewpoint of the latter).

Here we will discuss these two components of the fine structure in turn. There exist even smaller corrections, such as the **Lamb shift** and the **hyperfine structure**, but we will not consider them in this course.

4.3.1 The relativistic correction

Let us start by the Hamiltonian of the hydrogen atom:

$$\hat{H}\psi(\mathbf{r}) = \left(-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 - \frac{e^2}{4\pi\epsilon_0 r}\right)\psi(\mathbf{r}) = E\psi(\mathbf{r})\,. \tag{4.97}$$

The first term in the Hamiltonian corresponds to the operator representation of the classical **kinetic energy** T. However, if the electron moves at a non-negligible fraction of the speed of light c, the classical expression for the kinetic energy will not be accurate and one instead should use the corresponding relativistic expression:

$$T_{\rm rel} = mc^2\left(\gamma - 1\right)\,, \tag{4.98}$$

where we have used the usual **dilation factor** defined in relativistic calculations

$$\gamma = \frac{1}{\sqrt{1-\beta^2}}\,,\quad \beta \equiv \frac{v}{c}\,. \tag{4.99}$$

We can estimate the relevance of relativistic effects of the electron in the hydrogen atom as follows. For an **electron in the ground state**, its kinetic energy is of the order of $T = \mathcal{O}(10)$ eV (recall that the ionisation energy is 13.6 eV). The ratio of kinetic over rest mass then determines the dilation factor

$$\frac{T_{\rm rel}}{E_{\rm rest}} = \frac{10~{\rm eV}}{0.511~{\rm MeV}} = \gamma - 1\,, \tag{4.100}$$

where we have entered the value of the rest mass energy of the electron. One then finds that $\beta \simeq 0.006$, hence the electron moves at a speed of **0.6% the speed of light**: enough for its relativistic correction to be measurable.

In order to find the operator representation of $T_{\rm rel}$, and hence determine what are the relativistic corrections to the Hamiltonian of the hydrogen atom, we need to express $T_{\rm rel}$ in terms of the linear momentum. Taking into account that in special relativity the total energy of a particle is the sum of its kinetic and rest energies, you can find that the corresponding expression is

$$T_{\rm rel} = mc^2\left[\sqrt{1+\left(\frac{p}{mc}\right)^2} - 1\right] \simeq \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \mathcal{O}\left(\frac{p^6}{m^5c^4}\right)\,. \tag{4.101}$$

It is perfectly fine to truncate this Taylor expansion at the first order: the expansion parameter is p/mc, and for an electron in the hydrogen atom this is a small number as demonstrated above: it is the ratio of its kinetic energy to its rest mass. Another way to see why this is the relevant expansion parameter is to express the previous equation as

$$T_{\rm rel} = \frac{p^2}{2m}\times\left(1 - \frac{p^2}{4m^2c^2} + \mathcal{O}\left(\frac{p^4}{m^4c^4}\right)\right)\,. \tag{4.102}$$

Henceforth, at first order in perturbation theory, the lowest-order **relativistic correction** to the Hamiltonian of the hydrogen atom is given by:

$$H_r' = -\frac{p^4}{8m^3c^2}\,. \tag{4.103}$$

We can thus now use **perturbation theory**, Eq. (4.16), to determine the first order corrections to the Bohr energies:

$$E_r^{(1)} = \langle H_r' \rangle = -\frac{1}{8m^3c^2}\left\langle \psi_n^{(0)} \middle| \hat{p}^4 \psi_n^{(0)} \right\rangle = -\frac{1}{8m^3c^2}\left\langle \hat{p}^2\psi_n^{(0)} \middle| \hat{p}^2 \psi_n^{(0)} \right\rangle. \tag{4.104}$$

Note that the size of the perturbative correction will depend (at least) on the principal quantum number n. In order to make progress with our calculation, we can use the fact that the Schroedinger equation for the unperturbed states of the hydrogen atom reads

$$\hat{p}^2 \left| \psi_n^{(0)} \right\rangle = 2m\left(E_n - V(r)\right)\left| \psi_n^{(0)} \right\rangle \tag{4.105}$$

to evaluate the first order correction in the following manner:

$$E_r^{(1)} = -\frac{1}{2mc^2}\left\langle (E_n - V(r))^2 \right\rangle = -\frac{1}{2mc^2}\left[E_n^2 - 2E_n\langle V \rangle + \left\langle V^2 \right\rangle\right], \tag{4.106}$$

which by inserting the **Coulomb potential** gives the following result:

$$E_r^{(1)} = -\frac{1}{2mc^2}\left[E_n^2 - \frac{E_n e}{2\pi\epsilon_0}\left\langle \frac{1}{r} \right\rangle + \frac{e^2}{16\pi^2\epsilon_0^2}\left\langle \frac{1}{r^2} \right\rangle\right] \tag{4.107}$$

By using the radial probability distribution $P_{nl}(r) = r^2|R_{nl}(r)|^2$ associated to the unperturbed electronic orbital $\psi_{nlm}(r,\theta,\phi)$, we can evaluate these expectation values

$$\left\langle \frac{1}{r} \right\rangle = \int_0^\infty dr\, r|R_{nl}(r)|^2 = \frac{1}{n^2 a}\,, \tag{4.108}$$

$$\left\langle \frac{1}{r^2} \right\rangle = \int_0^\infty dr\, |R_{nl}(r)|^2 = \frac{1}{(\ell + 1/2)\,n^3a^2}\,, \tag{4.109}$$

in terms of the Bohr radius a. We can insert these results in Eq. (4.107) and express the result in terms of the unperturbed energy $E_n^{(0)}$. The final result for the **relativistic correction** to the Bohr energy levels of the hydrogen atom is

$$E_r^{(1)} = -\frac{1}{2m_ec^2}\left(E_n^{(0)}\right)^2\left[\frac{4n}{\ell+1/2} - 3\right] = -\frac{1}{2m_ec^2}\left(\frac{\alpha^2 m_e c^2}{2n^2}\right)^2\left[\frac{4n}{\ell+1/2} - 3\right], \tag{4.110}$$

which can be simplified to read:

$$E_r^{(1)} = -\frac{\alpha^4 m_e c^2}{8n^4}\left[\frac{4n}{\ell+1/2} - 3\right], \tag{4.111}$$

which as discussed before is a correction of the order of $\mathcal{O}(E_r^{(1)}) = \alpha^4 mc^2$.

A few comments on this interesting results are in order:

- The use of perturbation theory appears to be **fully justified**: the **relative correction** due to relativistic effects is proportional to the ratio between the unperturbed (non relativistic) energy and the rest energy of the electron

$$\frac{E_r^{(1)}}{E_n^{(0)}} \propto \frac{E_n^{(0)}}{mc^2} \simeq 2 \times 10^{-5}\,. \tag{4.112}$$

Clearly, the next term in the perturbative expansion should be of the order $\mathcal{O}\left(10^{-10}\right)$ and hence it is perfectly fine to just keep the leading term.

- In the unperturbed hydrogen atom, electronic orbitals are **degenerate**: all the states ψ_{nlm} with common values of the principal number n and different ℓ and m share the same energy $E_n^{(0)}$.

  This degeneracy is **partially lifted** by the relativistic correction, and now at first order in perturbation theory the energies of the electronic orbitals of the hydrogen atom $E_{n,\ell}$ depend both on n and ℓ:

  $$E_{n,\ell} = -\frac{\alpha^2 m_e c^2}{2n^2} - \frac{\alpha^4 m_e c^2}{8n^4}\left[\frac{4n}{\ell+1/2} - 3\right] = -\frac{\alpha^2 m_e c^2}{2n^2}\left(1 + \frac{\alpha^2}{4n^2}\left[\frac{4n}{\ell+1/2} - 3\right]\right) . \tag{4.113}$$

  Note however that the $(2\ell+1)$-fold degeneracy in m remains and it is not affected by the relativistic correction.

- As a related point, one might have argued that the use of **non-degenerate perturbation theory** was not appropriate since, well, the electronic orbitals of the unperturbed hydrogen atom are **degenerate**. Fortunately, ψ_{nlm} can be shown to be "good" states for this problem, due to the rotationally symmetric nature of the perturbation, which depends only on r but not on θ and ϕ, as indicated by Eq. (4.107).

4.3.2 The spin-orbit coupling interaction

The second contribution to the fine structure of hydrogen, in addition to the relativistic correction, comes from the interaction of the **electron spin** with its own **orbital motion**. Recall from your electromagnetism courses that any orbiting electric charge generates a magnetic field $\mathbf{B}$. From the viewpoint of the electron, the positive charge of the proton **orbiting around it** is generating an "external" magnetic field which will interact with its spin.

Indeed, this magnetic field generated by the proton "orbital motion" will interact with the **magnetic dipole moment of the electron** that arises from its spin, leading to the Hamiltonian

$$H = -\mathbf{B}\cdot\boldsymbol{\mu} , \tag{4.114}$$

which tends to orient the spin of the electron in the same direction as its magnetic field (since this is the configuration that minimises the total energy of the system). We denote this effect as the **spin-orbit coupling** or interaction, and here we will compute how it modifies the unperturbed (Bohr) energies of the hydrogen atom. Since electrons orbiting in different directions and with spin pointing in different orientations will receive a different contribution from this perturbed Hamiltonian, we can expect that this spin orbit coupling **breaks further the degeneracy** present in the electronic orbitals of the hydrogen atom.

In order to construct the perturbed Hamiltonian responsible for the spin orbit interaction, we need to evaluate the magnetic dipole moment of the electron and the magnetic field generated by the proton. The calculation involves some interesting elements of classical electrodynamics combined with some corrections which are beyond the scope of this course, for example, the electron magnetic moment $\boldsymbol{\mu}_e$ can only be computed in a proper way in relativistic quantum mechanics, that is in **quantum field theory**, and the result is

$$\boldsymbol{\mu}_e = -\frac{e}{m}\boldsymbol{S} , \tag{4.115}$$

with $\boldsymbol{S}$ being the electron spin. Here we just quote the final result for the Hamiltonian responsible for this spin-orbit interaction in terms of $\boldsymbol{S}$ and the angular momentum $\boldsymbol{L}$:

$$H'_{\rm so} = \left(\frac{e^2}{8\pi\epsilon_0}\right)\frac{1}{m^2c^2r^3}\boldsymbol{S}\cdot\boldsymbol{L} , \tag{4.116}$$

where $\boldsymbol{L}$ is the angular momentum associated to the orbital motion of the proton around the electron, from the electron perspective. We can now use our perturbation theory toolbox to evaluate how this perturbed Hamiltonian modifies the Bohr energy levels of the hydrogen atom.

Adding the spin-orbit correction Eq. (4.116) to the unperturbed Hamiltonian of the hydrogen atom has an important consequence: the Hamiltonian no longer commutes with $\boldsymbol{L}$ and $\boldsymbol{S}$ separately, so spin and orbital angular momentum are **no longer conserved separately**. However, $H'_{\rm so}$ does commute with the total angular momentum $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$, so we can write

$$\left[\hat{H}'_{\rm so}, \boldsymbol{J}\right] = \left[\hat{H}'_{\rm so}, \hat{L}^2\right] = \left[\hat{H}'_{\rm so}, \hat{S}^2\right] = 0\,. \tag{4.117}$$

By recalling the generalised Ehrenfest theorem, hence we know that $\boldsymbol{J}$, S^2 and L^2 will be conserved quantities not affected by the spin-orbit interaction. This implies that once one accounts for spin-orbit coupling effects, one needs to change the quantum numbers that label the eigenstates of the full Hamiltonian as follows:

- **Unperturbed Hamiltonian** H^0: the eigenstates of the Hamiltonian can be constructed from the eigenstates of $\hat{L}^2$, $\hat{L}_z$, $\hat{S}^2$, $\hat{S}_z$.
- **Perturbed Hamiltonian** $H = H^0 + H'_{\rm so}$: the eigenstates of the Hamiltonian can be constructed from the eigenstates of $\hat{J}^2$, $\hat{J}_z$, $\hat{S}^2$, $\hat{L}^2$. So we have four quantum numbers as well, just they are different as compared to those in the unperturbed case.

Let us use this property to determine the sought-for first-order corrections to the energy levels of the hydrogen atom due to the spin-orbit interactions. First of all let us express $H'_{\rm so}$ in terms of those quantities that are conserved upon this interaction. We can write

$$J^2 = (\boldsymbol{L} + \boldsymbol{S})^2 = L^2 + S^2 + 2\boldsymbol{S}\cdot\boldsymbol{L}\,, \tag{4.118}$$

which implies that the spin-orbit coupling Hamiltonian will be proportional to:

$$\boldsymbol{S}\cdot\boldsymbol{L} = \frac{1}{2}\left(J^2 - L^2 - S^2\right) \tag{4.119}$$

Now, as we said before, the eigenstates of $\hat{H} = \hat{H}^0 + \hat{H}'_{\rm so}$ can be labelled by the following conserved quantum numbers

$$|n, l, s, j, m_j\rangle\,, \tag{4.120}$$

where m_j is the quantum number associated to J_z. Applying the $\boldsymbol{S}\cdot\boldsymbol{L}$ operator to these eigenstates returns

$$\boldsymbol{S}\cdot\boldsymbol{L}|n,l,s,j,m_j\rangle = \frac{1}{2}\left(J^2 - L^2 - S^2\right)|n,l,s,j,m_j\rangle = \frac{\hbar^2}{2}\left(j(j+1) - \ell(\ell+1) - s(s+1)\right)|n,l,s,j,m_j\rangle \tag{4.121}$$

so using that $s = 1/2$ we conclude that the eigenstates of the operator $\boldsymbol{S}\cdot\boldsymbol{L}$ are

$$\frac{\hbar^2}{2}\left(j(j+1) - \ell(\ell+1) - \frac{3}{4}\right)\,. \tag{4.122}$$

We are now ready to compute the expectation values of the perturbed Hamiltonian. By using non-degenerate perturbation theory (for the same reason as in the case of the relativistic correction), we can use that

$$E^{(1)}_{\rm so} = \langle H'_{\rm so}\rangle = \left(\frac{e^2}{8\pi\epsilon_0}\right)\frac{1}{m^2c^2}\left\langle\frac{\boldsymbol{S}\cdot\boldsymbol{L}}{r^3}\right\rangle \tag{4.123}$$

These perturbed eigenstates have well-defined values of $\langle \boldsymbol{S}\cdot\boldsymbol{L}\rangle$ computed above. Recall that in the hydrogen atom, the dependence of the wave function on the radial coordinate r **factorises** with the dependence on the angular variables (θ,ϕ) and the spin, and hence expectation values can be computed separately. By using the result of the expectation value of r^{-3}

$$\left\langle \frac{1}{r^3}\right\rangle = \frac{1}{\ell(\ell+1/2)(\ell+1)n^3a^3} \tag{4.124}$$

we end up with the following result

$$E^{(1)}_{\rm so} = \frac{e^2\hbar}{16\pi\epsilon_0 m^2c^2}\frac{(j(j+1)-\ell(\ell+1)-3/4)}{\ell(\ell+1/2)(\ell+1)n^3a^3} = \frac{\left(E_n^{(0)}\right)^2}{mc^2}\left[\frac{n\,(j(j+1)-\ell(\ell+1)-3/4)}{\ell(\ell+1/2)(\ell+1)}\right]\,, \tag{4.125}$$

and hence it is of the relative order, $\mathcal{O}\left(E_n^{(0)}/mc^2\right)$, sams as in the case of the relativistic correction. Note that Eq. (4.125) can be applied only to states with $\ell > 0$, since those for $\ell = 0$ have vanishing angular momentum and hence they do not experience the effects of the spin-orbit coupling.

Combining the relativistic correction with the spin-orbit correction, we end up with the following result for the **fine structure** of the hydrogen atom:

$$E^{(1)}_{\rm fs} = \frac{\left(E_n^{(0)}\right)^2}{2mc^2}\left(3-\frac{4n}{j+1/2}\right) \tag{4.126}$$

and the final result for the energies of the electronic orbitals of the hydrogen atom at $\mathcal{O}\left(\alpha^2\right)$ in perturbation theory are given by the following result:

$$E_{nj} = -\frac{(13.6\,{\rm eV})}{n^2}\left[1+\frac{\alpha^2}{n^2}\left(\frac{n}{j+1/2}-\frac{3}{4}\right)\right]\,. \tag{4.127}$$

Some comments about this foundational calculation:

- The *relative effect* of the fine-structure corrections is most important for the ground state of hydrogen $(n = 1)$, and then decreases as $1/n$.
- Likewise, the higher the value of the total angular momentum j, the less important the relative impact of the fine structure effects: for a given n, these corrections are the largest for $j = 1/2$.
- The degeneracy in ℓ is now broken but there is still **degeneracy in j:** all states with the same total angular momentum j will exhibit the same energy.
- The "good" quantum numbers to describe the electronic orbitals of the hydrogen atom are now n, ℓ, s, j, m_j, to be compared to those we used for the unperturbed orbitals: n, ℓ, s, m_ℓ, m_s.

Interestingly, let me point out that the **exact expression** for the fine structure of hydrogen can be derived from the Dirac equation that describes the relativistic quantum mechanics of fermions, it can be checked that expanding this exact result up to $\mathcal{O}\left(\alpha^2\right)$.

Summary

We can now recapitulate what have we learned in this chapter concerning the application of perturbation theory to quantum systems for which a closed form analytic solution of the corresponding Schroedinger equation does not exist.

I/ In the case of non-degenerate energy spectra, perturbation theory provides a systematic prescription to improve our estimate of the energies and state vectors of the system in the case of a modified Hamiltonian.

II/ This perturbative expansion is controlled by some small expansion parameter, which sets the convergence rate (or lack thereof) of this expansion.

III/ The perturbation theory method applied also to systems for which the exact solution is known, and verify that the correct expressions are obtained in terms of the power series expansion in the small parameter λ.

IV/ The presence of degenerate energy states can spoil the convergence of perturbation theory, unless we rotate the unperturbed states to a special basis where this problem is absent.

V/ Perturbation theory often breaks the degeneracies that affect the energy spectra of unperturbed systems, since they affect different unperturbed eigenvectors in a different manner.

VI/ Using perturbation theory we can evaluate the first non-trivial corrections to the Bohr energy levels of the hydrogen atom, which correspond to $\mathcal{O}\left(\alpha^2\right)$ terms in terms if the fine structure constant.

Quantum Mechanics 2

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

Current version: **March 2, 2021**

5 Chapter 8: The Variational Principle

Learning Goals

- To formulate the variational principle and apply it to different quantum systems to provide **upper bounds for their ground state energies**.
- To demonstrate that a **careful choice of test function** can lead to variational principle results that coincide with the exact solution of the Hamiltonian.
- To deploy the variational principle to construct a better approximation of the **ground state of the Helium** atom.
- To calculate from first principles the **effective electric charge** $Z_{\rm eff}$ of the helium atom, which previously was taken to be a phenomenological parameter extracted from data.

In this section of the lecture notes we present the main concepts discussed in **Chapter 8 ("The Variational Principle")** 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

- **8.1**: Theory
- **8.2**: The ground state of Helium

As we have seen through this course, unfortunately the number of quantum systems whose Hamiltonian admits an exact, analytic diagonalisation is quite scarce. This is why **approximate methods** such as the perturbation theory presented in the previous lecture are so hugely important in quantum mechanics. Here we are going to present another of these approximate methods, known as the variational principle. The main idea is to find some test wave function that provides an **upper bound** to the energy of the **ground state** of the system. By means of a clever choice of this test wave function, together with an optimisation of its adjustable parameters, we can obtain a very decent estimate of the ground state energy and wave functions of relatively complex systems without the need to solve the problem exactly.

5.1 Theoretical foundations

The key result underlying the variational principle method is the following. We would to determine the **ground state** of a given Hamiltonian $\hat{H}$, but are unable to solve its eigenvalue equation exactly for the ground state (or for any other state in general):

$$\hat{H}|\psi_{\rm gs}\rangle = E_{\rm gs}|\psi_{\rm gs}\rangle \, . \tag{5.1}$$

Else we need to resort to some approximate method to estimate the ground state energy $E_{\rm gs}$ of this quantum system. We can demonstrate that if we select at random some **normalised state vector** $|\psi\rangle$, then we can obtain an **upper bound** on the ground state energy by evaluating the following matrix element:

$$E_{\rm gs} \le \langle\psi|\hat{H}\psi\rangle = \langle H\rangle \ , \tag{5.2}$$

where the equality is only true if we use $|\psi_{\rm gs}\rangle$ as our test function (at least, assuming that the ground state of our Hamiltonian is non-degenerate). In other words, the **expectation value** of the Hamiltonian of the system on some test function is guaranteed to provide an upper bound on $E_{\rm gs}$. This result is of course trivial if $|\psi\rangle$ is one of the excited states of the Hamiltonian, but perhaps less so if this state vector has been genuinely chosen at random.[15]

Let us demonstrate that this result is indeed true. To show this, I will express my **test state** $|\psi\rangle$ as a linear combination of the eigenvectors of my Hamiltonian, that is

$$|\psi\rangle = \sum_n c_n|\psi_n\rangle \quad \text{where} \quad \hat{H}|\psi_n\rangle = E_n|\psi_n\rangle \, , \tag{5.3}$$

where as usual $\sum_n |c_n|^2 = 1$ due to the normalisation of the state vector, and the coefficients are computed as $c_n = \langle\psi_n|\psi\rangle$. I can now use this expansion to evaluate the expectation value of the Hamiltonian in the test state vector $|\psi\rangle$ as follows:

$$\langle\psi|\hat{H}\psi\rangle = \sum_{n,m} c_n^* c_m \langle\psi_m|\hat{H}|\psi_n\rangle = \sum_{n,m} c_n^* c_m E_n \langle\psi_m|\psi_n\rangle = \sum_n |c_n|^2 E_n \, . \tag{5.4}$$

Again, this is what we expect: $\langle\psi|\hat{H}\psi\rangle$ is nothing but the expectation value of the energy in our state $|\psi\rangle$. However, the ground state energy is by definition the **lowest of the energy eigenvalues** E_n, and hence

$$\langle\psi|\hat{H}\psi\rangle = \sum_n |c_n|^2 E_n \ge \sum_n |c_n|^2 E_{\rm gs} = E_{\rm gs} \, , \tag{5.5}$$

---

[15] In practical applications, as we will show, a clever choice of the test function improves significantly the accuracy of our bound state energy estimate.

thus demonstrating the main result of the variational principle.

As you might have guessed, clever choices of the test state vector will provide much better estimates to $E_{\rm gs}$ than genuinely random choices.

Why variational?

The name of Variational Principle stems from the fact that in general our test state $|\psi\rangle$ will be made to depend on some **tunable parameters**, that determine for example the shape of the wave function. Then, once $\langle\psi|\hat{H}\psi\rangle$ as been computed, we can optimise the result by minimising with respect to these tunable parameters, since any of their values is still providing an upper bound on $E_{\rm gs}$. In other words, we will determine the **optimal variation** of the test function parameters to achieve the lowest possible value for $\langle H\rangle$ and hence the **best possible estimate** for $E_{\rm gs}$.

The variational method for the quantum harmonic oscillator. Let us now demonstrate how this method works for one specific quantum state for which we know the exact solution, and hence we can check whether or not the estimate provided by the variational principle is reasonable. Let's consider the **quantum harmonic oscillator** and try a Gaussian test function

$$\psi(x,b) = Ae^{-bx^2}\,, \tag{5.6}$$

where A is determined by normalisation and b is some positive real constant, which plays the role of adjustable parameter in our test function. You can check that $A=(2b/\pi)^{1/4}$ by evaluating the usual definite Gaussian integral, so b is the only genuine **free parameter** of our test function for the ground state of the system.

Following now the prescription provided by the **variational principle**, we need to take the Hamiltonian associated to this system

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d}{dx^2}+\frac{1}{2}m\omega^2x^2\,, \tag{5.7}$$

and then compute its expectation value in the trial state defined in Eq. (5.13). We can evaluate separately the contributions from the kinetic and from the potential energy and find:

$$\langle T\rangle = -\frac{\hbar^2}{2m}\left(\frac{2b}{\pi}\right)^{1/2}\int_{-\infty}^{\infty}dx\,e^{-bx^2}\left(\frac{d^2}{dx^2}e^{-bx^2}\right)=\frac{\hbar^2 b}{2m}\,, \tag{5.8}$$

$$\langle V\rangle = \frac{1}{2}m\omega^2\left(\frac{2b}{\pi}\right)^{1/2}\int_{-\infty}^{\infty}dx\,e^{-bx^2}\left(x^2e^{-bx^2}\right)=\frac{m\omega^2}{8b}\,, \tag{5.9}$$

and combining the two contributions we find that an **upper limit** to the energy of the ground state of the harmonic operator Hamiltonian is given by

$$\langle H\rangle(b) = \left(\frac{\hbar^2 b}{2m}+\frac{m\omega^2}{8b}\right)\ge E_{\rm gs}\,. \tag{5.10}$$

Now, a crucial observation is that this result holds true for **any value of the free parameter** b, and hence the best estimate that we can have for the ground state energy is obtained by **minimizing** this expectation value of $\hat{H}$ with respect to the value of b:

$$\frac{d\langle H\rangle(b)}{db} = \frac{\hbar^2}{2m}-\frac{m\omega^2}{8b^2}=0 \quad\to\quad b=\frac{m\omega}{2\hbar}\,. \tag{5.11}$$

We can now collect our results: our best estimate for the wave function and energy of the ground state of the quantum harmonic oscillator and its energy obtained by virtue of the variational principle (that is, we did not have to solve any differential equation!) will be

$$\psi(x) = \left(\frac{2b}{\pi}\right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar}x^2\right) , \qquad E_{\rm gs} \le \frac{1}{2}\hbar\omega \, . \tag{5.12}$$

By comparing with the **exact solution** for the ground state wave function of the harmonic oscillator, we see that in this case the equality is satisfied and the resulting wave functions and energies are **identical** to those of the exact result. This is no coincidence, since we have chosen a test function which is Gaussian, which is the same as in the exact solution. If you would have tried a different test function, such as for example

$$\psi(x,b) = Ae^{-b|x|} \, , \tag{5.13}$$

we would not have saturated the upper bound on $E_{\rm gs}$, as you can check explicitely.

Choosing the test function

As we have seen from this example, the closer our test function $|\psi\rangle$ is to the actual ground state of the system $|\psi_{\rm gs}\rangle$, the better the estimate provided by the variational principle will be. Actually, if you choose a functional form which is consistent with $\psi_{\rm gs}(x)$ and then adjust the tunable parameters, you are **guaranteed** to recover the exact result for the ground state wave function $\psi_{\rm gs}(x)$ and the energy $E_{\rm gs}$ **without having to solve the Schroedinger equation.**

The variational principle for the infinite square well. Another illustrative example of the usefulness of the variational principle it to deploy it to estimate the ground state wave function and energies of the infinite square well potential (again, we know the exact results for this quantum system, so this facilitates the validation of the method).

In this case the exact ground state energy and wave function are

$$E_{\rm gs} = \frac{\hbar^2\pi^2}{2ma^2} \, , \quad \psi_1(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{\pi x}{a}\right) \, . \tag{5.14}$$

Let us see how well we can approximate the ground state energy by using a test wave function which is **not sinusoidal** (hence we will not recover this time the exact value of $E_{\rm gs}$). Since we know that the wave function must vanish at the edges of the potential well due to the wave function continuity, a possible choice for our test function will be

$$\psi(x) = \frac{2}{a}\sqrt{\frac{3}{a}}x \quad \text{for} \quad 0 \le x \le a/2 \tag{5.15}$$

$$\psi(x) = \frac{2}{a}\sqrt{\frac{3}{a}}(a-x) \quad \text{for} \quad a/2 \le x \le a \tag{5.16}$$

$$\psi(x) = 0 \quad \text{otherwise} \tag{5.17}$$

where the overall factor has been fixed by imposing the normalisation of the text wave function. Note that this function does not have any adjustable parameters, as opposed to the previous example.[16]

---

[16] One option to improve this test wave function is to have that $\psi(x,p) \propto x^p \; ((1-x)^p)$ and then optimise the estimate with respect to the tunable parameter p.

We can now compute the expectation value of $\hat{H}$ in this test wave function, which is reduced to the contribution of the **kinetic energy** inside the potential well (since the potential vanishes inside the well, and the wave function vanishes outside the well). We have to be careful here, since the test function $\psi(x)$ has a discontinuous first derivative, and hence its second derivate is expressed in terms of the Dirac delta function:

$$\frac{d^2\psi(x)}{dx^2} = \frac{2}{a}\sqrt{\frac{3}{a}}\left[\delta(x) - 2\delta(x-a/2) + \delta(x-a)\right] . \tag{5.18}$$

Equipped with this result, I can now evaluate the required value of the expectation value of the Hamiltonian,

$$\langle H \rangle = -\frac{\hbar^2}{2m}\int_0^a dx\, \psi(x)\frac{d^2\psi(x)}{dx^2} = \frac{12\hbar^2}{2ma^2} , \tag{5.19}$$

where the only delta function that contributes to the integral is the one at $x = a/2$. Hence our **upper bound to the ground state energy** for the infinite square well potential found with the variational principle is

$$\langle H \rangle = \frac{12\hbar^2}{2ma^2} \geq E_{\rm gs} = \frac{\hbar^2\pi^2}{2ma^2} \tag{5.20}$$

which is quite a decent approximation, with the variational principle estimate giving a value 17% higher than the actual result. We could have improved our estimate by having the slope of the test function a free parameter and then minimise $\langle H \rangle$ with respect to it, or by trying polynomial test wave functions that resemble the sinusoidal behaviour of the exact ground state.

5.2 The ground state of Helium

The two previous examples have been discussed with a purely pedagogical motivation, since we knew the exact results of the corresponding Hamiltonians. We are now going to use the variational principle to obtain a **genuinely new result**, namely the ground state and energy of the **Helium atom** taking into into account the electron repulsion. As we discussed in Sect. 3.2, the Schroedinger equation for the Helium atom cannot be solved exactly and it is thus mandatory to introduce some approximations.

Our starting point is the full Hamiltonian of the helium atom

$$H = -\frac{\hbar^2}{2m}\left(\boldsymbol{\nabla}_1^2 + \boldsymbol{\nabla}_2^2\right) - \frac{e^2}{4\pi\epsilon_0}\left(\frac{2}{r_1} + \frac{2}{r_2} - \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\right) . \tag{5.21}$$

Previously, we found a solution for the ground state and energy by **neglecting the electron repulsion term**, which we then improved phenomenologically using the idea of effective electric charge. Let us now use the variational principle to determine an alternative estimate of the ground state wave function which takes into account from the very beginning the electron repulsion.

As a **test function**, we will use the following function:

$$\psi_0\left(\mathbf{r}_1, \mathbf{r}_2\right) = \frac{8}{\pi a^3} e^{-2(r_1+r_2)/a} , \tag{5.22}$$

which is the solution for the ground state of the Helium atom assuming that the electrons are **non-interacting**, that is, when the electron repulsion is neglected. This choice is convenient because it is an eigenfunction of the Hamiltonian of the system minus the electron repulsion term, that is,

$$\langle H \rangle = 8E_1 + \langle V_{ee} \rangle \tag{5.23}$$

where $E_1 = -13.6$ eV is the **ionisation energy** of the hydrogen atom, and the factor 8 comes because the energies of hydrogen-like atoms scale as $E_n^{(Z)} = Z^2 E_n^{(\mathrm{H})}$ in terms of the hydrogen atom energies.

The expectation value of the electronic repulsion potential is given by

$$\langle V_{ee} \rangle \equiv \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \left(\frac{8}{\pi a^3} \right)^2 \int d\mathbf{r}_1 \, d\mathbf{r}_2 \, \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} e^{-4(r_1+r_2)/a} \tag{5.24}$$

If we are able to evaluate this integral, we will obtain the variational principle estimate for the ground state of the helium atom. The evaluation of the integral is presented in the textbook, and here we just quote its result

$$\langle V_{ee} \rangle = -\frac{5}{2} E_1 \simeq 34 \, \mathrm{eV} \, , \tag{5.25}$$

and hence we find that the variational principle tells us that an upper bound for the ground state energy of the Helium atom will be given by:

$$\langle H \rangle = 8E_1 + \langle V_{ee} \rangle = -75 \text{ eV} \geq E_{\mathrm{gs}} \, . \tag{5.26}$$

If we compare with the experimental result, $E_{\mathrm{gs}} = -78.975$ eV, we can observe that the estimate provided by the variational principle is much better than the one provided by neglecting the inter-electron repulsion, $8E_1 = -109$ eV. So this is a nice result, specially since we could **recycle** a known wave function which was already eigenfunction of most of the Hamiltonian to reduce the computational burden of the problem.

We have seen from our previous examples that it is convenient to endow our test wave function with some **free parameter** and then minimise the value of $\langle H \rangle$ with respect to it. This way we can further improve on our estimate for E_{gs} for the Helium atom. In Sect. 3.2 we introduced the idea of the **effective charge** to partially account for the screening that the electrons induce in the total positive change of the helium nucleus. There the value of Z_{eff} was fixed by some phenomenological analysis to reproduce the experimental results, here we will determine this effective charge from **first principles** using the variational method.

With this motivation, instead of the parameter-free test function Eq. (5.22) that we used before, we will adopt now the following test function

$$\psi_0 \left(\mathbf{r}_1, \mathbf{r}_2, Z \right) = \frac{Z^3}{\pi a^3} e^{-Z(r_1+r_2)/a} \, , \tag{5.27}$$

with Z being a free parameter to be optimised once we have evaluated the expectation value $\langle H \rangle$. It is clear that the result of this calculation will then be the **optimal value of the effective electric charge** which accounts for the electron repulsion. Let us know apply the variational principle to this system.

This wave function is an eigenstate of the "unperturbed" Hamiltonian (neglecting the electron repulsion) if it had Z in the Coulomb terms rather than 2. To account for this property, we rewrite Eq. (5.21) as follows:

$$\begin{aligned} H &= -\frac{\hbar^2}{2m} \left(\boldsymbol{\nabla}_1^2 + \boldsymbol{\nabla}_2^2 \right) - \frac{e^2}{4\pi\epsilon_0} \left(\frac{Z}{r_1} + \frac{Z}{r_2} \right) \\ &+ \frac{e^2}{4\pi\epsilon_0} \left(\frac{(Z-2)}{r_1} + \frac{(Z-2)}{r_2} - \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \right) \, . \end{aligned} \tag{5.28}$$

Hence, the expectation value of Eq. (5.21) using the test wave function Eq. (5.27) will be given by

$$\langle H \rangle = 2Z^2 E_1 + 2(Z-2) \left(\frac{e^2}{4\pi\epsilon_0} \right) \left\langle \frac{1}{r} \right\rangle + \langle V_{ee} \rangle \, , \tag{5.29}$$

where the expectation value of r^{-1} is computed using the ground state wave function of a hydrogen-like atom with electric charge Z:

$$\left\langle \frac{1}{r} \right\rangle = \frac{Z}{a}\,. \tag{5.30}$$

For the average of V_{ee} we also need to account for the new electric charge factor Z,

$$\langle V_{ee} \rangle = -\frac{5Z}{2} E_1\,, \tag{5.31}$$

and hence our final result for the expectation value of the (full) Hamiltonian of the Helium atom in the test wave function Eq. (5.27) will be given by

$$\langle H \rangle = \left[-2Z^2 + (27/4)Z \right] E_1\,. \tag{5.32}$$

The final step of the variational principle calculation is to optimise this result with respect to the value of the adjustable parameter Z, namely

$$\frac{d\langle H \rangle}{dZ} = \left[-4Z + (27/4) \right] E_1 = 0 \tag{5.33}$$

which leads to a **first principle calculation of the effective charge** of the Helium atom in the ground state, whose value is given by

$$Z = \frac{27}{16} = 1.69\,, \tag{5.34}$$

which is in good agreement with the phenomenological result, and that leads to an expectation value of

$$\langle H \rangle = -77.5\,\mathrm{eV}\,, \tag{5.35}$$

which is now really close to the experimentally measured value of $E_{\rm gs} = -78.975$ eV. As expected, the addition of a tunable parameter to the test wave function has lead to an improved estimate of the ground state energy as compared to the parameter-less test function Eq. (5.22).

This calculation demonstrates the usefulness and reliability of the variational method to estimate the ground state wave functions and energies of complex quantum systems for which the exact solutions are not available. By trying **more flexible test functions** and a **higher number of tunable parameters**, one can achieve each time more precise estimates of the value of $E_{\rm gs}$ in a systematic manner.

Summary

We can now recapitulate what have we learned in this chapter concerning the use of the variational principle to determine approximations to the energies and the wave functions of the ground states of quantum systems that cannot be solved exactly.

I/ The variational principle procedure guarantees an **upper bound** on the ground state of a general Hamiltonian, no matter how complex this is.

II/ By using test functions with more or one **tunable parameters**, we can improve in a systematic manner this estimate of the ground state energy.

III/ If we use a test function which has the same functional dependence of the exact ground state with an approximate number of tunable parameters, the variational principle calculation will return the **exact wave function and ground state energy** $E_{\rm gs}$ without having to solve the Schroedinger question.

IV/ By means of the variational principle, we can construct an approximation to the **ground state of the Helium atom** which is in very good agreement with the experimental data and carry out a **first principle calculations of the effective charge** accounting for the effects of electron repulsion in the system.