van Quantum tot Molecuul

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

Lecture notes, current version: March 7, 2019

Contents

van Quantum tot Molecuul

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

1 HC1+HC2: Introduction to Quantum Theory

General introduction to quantum theory
Based on Atkins and de Paula, *Physical Chemistry* 10th edition.

In this lecture we present the mathematical framework of quantum theory. We introduce the main postulates of quantum mechanics and study the main physical consequences. We introduce the quantum wave-function and study the equation that determine its behaviour: the Schroedinger equation. We also introduce a number of mathematical tools which are required to compute phenomena in the quantum formalism. We also discuss Heisenberg's uncertainty principle, and show that it entails a fundamental limitation about the physical knowledge that we can have about quantum systems.

Learning goals of these two lectures:

- Understanding that the state of a quantum system is fully described by the quantum wave-function, which is a solution of the time-independent Schroedinger equation.
- Becoming familiar with the mathematical formalism of quantum physics, including Hermitian operators, eigenvalue equations, and orthogonality.
- Determining how experimental observables can be calculated from the wave-function, and perform sample calculations in specific systems.
- Understanding the probabilistic interpretation associated to the quantum wave-function.
- Computing the expectation values of physical quantities in selected quantum states.

Introduction. In this first two lectures of the course (HC1 and HC2), we will present a concise overview of the basic principles of quantum theory, the mathematical language relevant for the description of small objects such as atoms, molecules, and electrons.

We will state these central principles of quantum theory as established facts, avoiding (due to lack of time) the historical discussion of how these principles were developed or what is the experimental evidence that validates them. The interested student can find a brief introduction to the historical developments that lead to the formation of quantum theory on the slides that I have posted in Canvas:

`https://canvas.vu.nl/courses/35399/files/folder/Historical%20Introduction%20to%20Quantum%20Theory`

In the following, we present the central concepts of quantum theory, focusing on those which are of special relevance for this course. When required, we will also briefly review the corresponding mathematical methods that are needed to deal with the idea of quantum theory.

1.1 Wave-particle duality

Every object, from a plane to a proton or an electron, will under specific conditions behave as a *wave*: this concept is known as *wave-particle duality*. Therefore, this object will experience a behaviour usually associated with waves, such as diffraction when crossing a narrow slit and the formation of interference patters.

This of course sounds quite anti-intuitive: we don't see every day cars interfering between them or diffracting as they cross a tunnel. The reason that the wave-like character of everyday objects is that quantum theory tells us that it should becomes apparent only when the length scales involved (say the width of an aperture) are of the same order or smaller than their so-called *De Broglie wavelength*, defined as

$$\lambda \equiv \frac{h}{p} = \frac{h}{mv} \,, \tag{1.1}$$

where $h = 6.26 \times 10^{-36}$ $\text{m}^2\text{kg/s}$ is a numerical constant known as *Planck's constant*, ubiquitous in the quantum theory, and $p = mv$ is the linear momentum of the object with m being its mass and v its velocity. Planck's constant is one of the defining features of quantum theory, and in particular it determines when genuinely quantum effects will be important.

Let us illustrate under which conditions the wave-like nature of two very different objects will become relevant: an electron and a car:

First of all let us consider an electron that travels at 1% of the speed of light. This electron will have associated a linear momentum of $p = m \cdot v = 2.73 \times 10^{-24}\,\text{kg m/s}$, and therefore its De Broglie wavelength, that is, the distance scale for which an electron should exhibit a wave-like character, is

$$\lambda = \frac{h}{p} \simeq 0.5 \times 10^{-10}\,\text{m} \,, \tag{1.2}$$

which is the typical separation between atoms in an crystal. Therefore one expects to observe the wave character of electrons when crossing a crystalline lattice.

On the other hand, for a car of mass $m = 1000$ kg moving at a velocity of $v = 100$ km/h, such that its momentum is $p = 2.8 \times 10^4$ kg m/s, we can associate a De Broglie wavelength of

$$\lambda = \frac{h}{p} \simeq 2 \times 10^{-38}\,\text{m} \,, \tag{1.3}$$

implying that the wave character of a car is unobservable: a car would diffract only when crossing a tunnel of width $\sim \lambda$, much smaller than an atom itself.

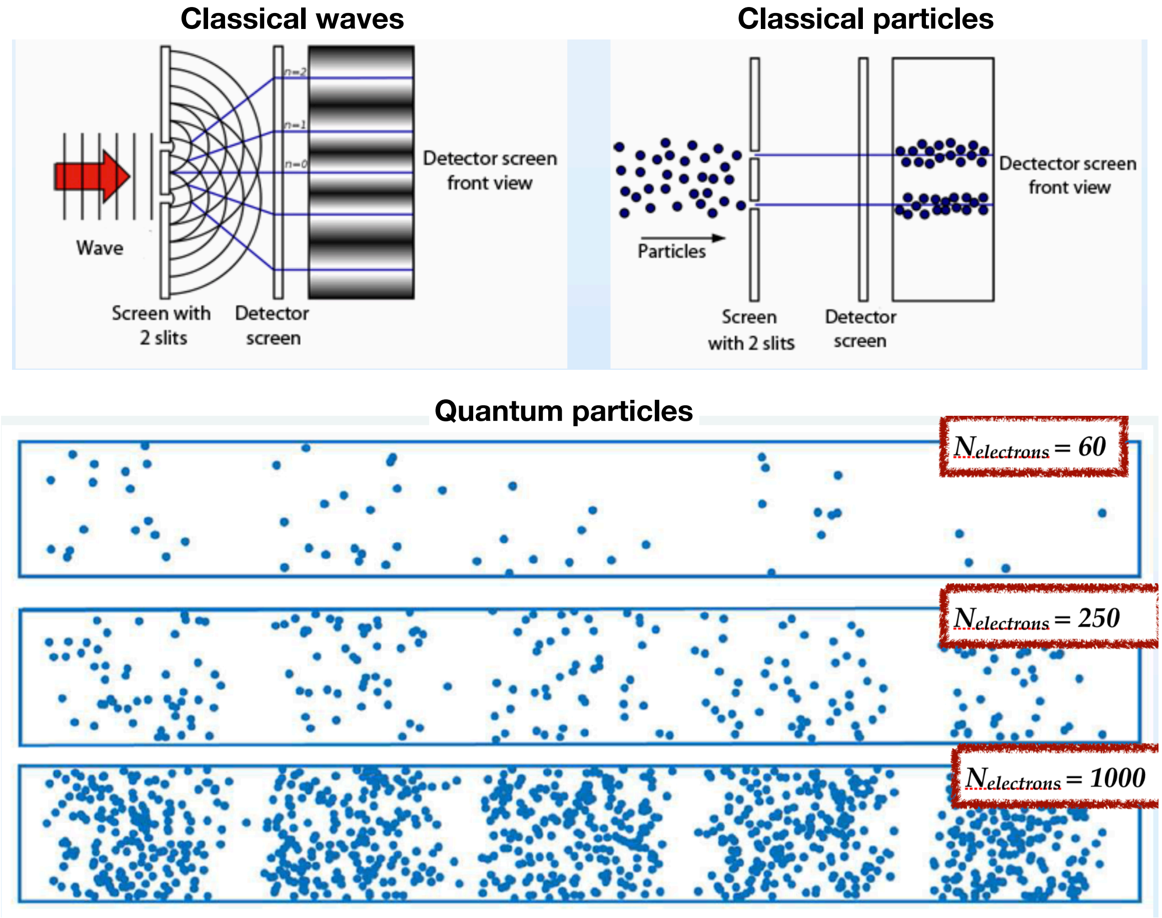


Figure 1.1: In classical theory, the behaviour of waves and particles in the double slit experiment is very different: waves lead to an interference pattern, while particles lead to just two blobs in the detector screen (upper diagrams). In quantum particle, the situation is very different (lower diagrams): if the number of incident particles is large enough, a clear interference patter also appears, just as if the particles behaved as waves.

Therefore, in quantum theory, when particles interact with objects of length similar or smaller than their de Broglie wavelength, their wave-like character will become apparent. In general particles such as protons and electrons will exhibit phenomena that we classically associate with either particles or waves depending on the specific circumstances.

The remarkable wave-particle behaviour of quantum theory is illustrated nicely by the famous double slit experiment, schematically represented in Fig. 1.1. In this experiment, electrons are beamed through a wall that has only two narrow slits, and then the position of the electrons is recorded in a detector screen. In classical theory, the behaviour of waves and particles in the double slit experiment is very different: waves lead to an interference pattern, while particles lead to just two blobs in the detector screen (upper diagrams). In quantum particle, the situation is very different (lower diagrams): if the number of incident particles is large enough, a clear interference patter also appears, just as if the particles behaved as waves. On the other hand, each particle leads to a single dot in the detector screen, so its wave-like character becomes clear only when the number of incident electrons is very large.

Interestingly, if one of the slits depicted in Fig. 1.1 is covered, then the wave-like interference patter formed in the detector screen disappears and the pattern of impacts is the same as for classical particles. The double slit experiment is perhaps the most iconic of those foundational experiments that highlight how the quantum world is radically difference from our everyday experience.

1.2 The quantum wave-function

In classical mechanics, one can obtain all the possible information about a system composed by N particles by specifying their positions $\vec{x}_i(t)$ and linear momenta $\vec{p}_i(t)$ at any given time t. One can in principle measure positions and velocities to any given precision, and therefore the state of the system can be completely determined.

The picture is quite different in quantum theory. There, all the *information* that can be obtained about a given quantum system is encoded in its so-called *wave-function* $\Psi(\vec{x})$. What is the physical interpretation of this wave function? Quantum theory tells us that the *probability* of finding the particle in a small region $d\vec{x}$ around the position $\vec{x}$ is given by the square of the wave function:

$$P(\vec{x}) = |\Psi(\vec{x})|^2 d\vec{x} \,. \tag{1.4}$$

In other words, the square of the wave function $|\Psi(\vec{x})|^2$ represents the *probability density* of finding the particle in a specific region of space and at given time upon its measurement. Knowledge of the wave-function specifies *all* the physical information that we can access for an specific quantum system: there is no additional "hidden" information on the system that is not already present in its wave function.

Since the wave function squared, $|\Psi(x)|^2$, represents the probability density in the coordinate x for this particle, we have that regions in x where $|\Psi(x)|^2$ is higher correspond to those regions where the particle has a higher likelihood to be found when a measurement is performed. This interpretation is known as the Born interpretation of the wave-function, and is further illustrated in Fig. 1.2. Those values of x where $|\Psi(x)|^2$ is maximal (minimal) have the highest (lowest) likelihood of being those where the particle is found upon a measurement.

Note that in general the quantum wave function is a complex function, so that the square is computed as

$$|\Psi(\vec{x})|^2 = \Psi^*(\vec{x})\Psi(\vec{x}) \,, \tag{1.5}$$

where the * sign indicates the *complex conjugate*. To be more precise, we have that

$$|\Psi|^2 = \Psi^*\Psi = \left[\mathrm{Re}(\Psi) - i\mathrm{Im}(\Psi)\right]\left[\mathrm{Re}(\Psi) + i\mathrm{Im}(\Psi)\right] \,, \tag{1.6}$$

in terms of the real and of the imaginary components of the wave function.

Let me give an specific example of a quantum wave function that arises in a relatively simple system - we will explain later how it has been computed. As we well see in HC3, for a particle that is confined in a one-dimensional box of dimensions $0 \le x \le L$ from an infinite potential of the form

$$\begin{aligned} V(x) &= 0 \,, \quad 0 \le x \le L \,, \\ V(x) &= \infty \,, \quad x \le 0 \quad \text{or} \quad x \ge L \,, \end{aligned} \tag{1.7}$$

one possible wave function allowed is:

$$\begin{aligned} \Psi(x) &= \sqrt{\frac{2}{L}} \sin\left(\frac{\pi x}{L}\right) \,, \quad 0 \le x \le L \,, \\ \Psi(x) &= 0 \,, \quad x \le 0 \quad \text{or} \quad x \ge L \,, \end{aligned} \tag{1.8}$$

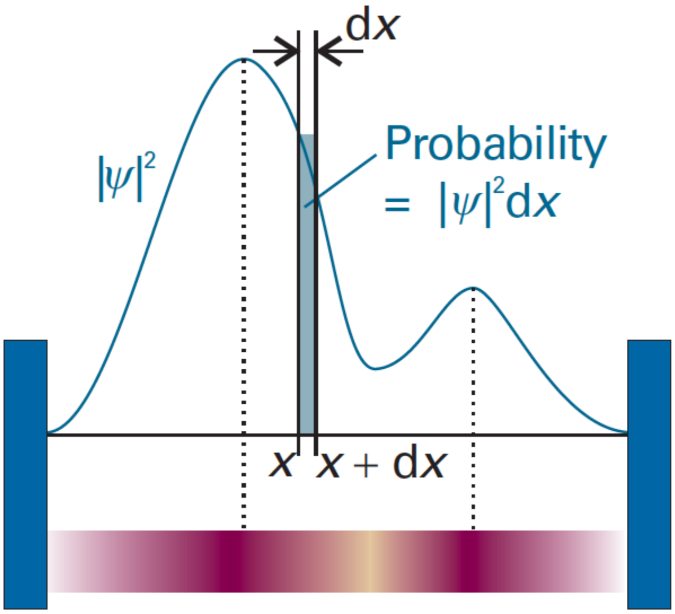


Figure 1.2: The physical interpretation of the quantum wave-function $\Psi(x)$ is that its square $|\Psi(x)|^2$ is the probability density associated to measuring the particle in a position x in space. Therefore, the probability to find the particle at position x in the range dx will be given by $|\Psi(x)|^2 dx$.

We can use this wave function to illustrate some important properties of the quantum wave function:

- The wave function is *normalised* to ensure that the probability of finding the particle *anywhere* in space is unity:

  $$\int_{-\infty}^{\infty} dx\, P(x) = \int_{-\infty}^{\infty} dx\, |\Psi(x)|^2 = 1\,, \tag{1.9}$$

  as we can check with our specific example:

  $$\int_0^L dx \frac{2}{L} \sin^2\left(\frac{\pi x}{L}\right) = \frac{2}{L}\frac{L}{\pi}\left[\frac{x}{2} - \frac{1}{4}\sin(2x)\right]_0^{\pi} = 1 \tag{1.10}$$

  which is appropriately normalised as expected. In computing this integral, we have used the indefinite integral Eq. (**??**) from the appendix, which reads

  $$\int \sin^2 x\, dx = \frac{x}{2} - \frac{\sin(2x)}{4}\,. \tag{1.11}$$

  This normalisation condition ensures the conservation of probability in quantum systems.

  If a given wave function $\widetilde{\Psi}(x)$ is not normalised, one can easily construct a normalised version as $\Psi(x) = N\widetilde{\Psi}(x)$, where the normalisation factor N is constructed as

  $$N = \left(\int dx |\widetilde{\Psi}(x)|^2\right)^{-1/2}, \tag{1.12}$$

  Note that here we assume that the wave function is always *square-integrable*, namely that integrals of the form Eq. (1.12) are always well defined (else the probabilistic interpretation would be impossible).

- The wave function is always *continuous*, specifically across boundaries. In the example we are considering one has two boundaries, at $x = 0$ and $x = L$. At these points the wave function is clearly continuous, since
$$\sin\left(\frac{\pi x}{L}\right)\Big|_{x=0,L} = 0\,. \tag{1.13}$$
Note that this continuity condition does not necessarily extend to derivatives of the wave function, for example the derivative of the wave function Eq. (1.8) is
$$\begin{aligned}\Psi'(x) &= \sqrt{\frac{2\pi}{L^2}}\cos\left(\frac{\pi x}{L}\right)\,, \quad 0 \le x \le L\,, \\ \Psi'(x) &= 0\,, \quad x \le 0 \quad \text{or} \quad x \ge L\,,\end{aligned} \tag{1.14}$$
which is not continuous at $x = 0$ or $x = L$, as you can easily check.

  The condition that the first derivative of the quantum wave function can be discontinuous applies only to boundaries where the potential $V(x)$ that a particle experiences becomes infinite (as is the case here, since the particle is confined in a box). In more realistic cases, where the potential $V(x)$ is finite, one can show that also the *first derivative of the quantum wave function is also continuous* and single-valued everywhere in space.

  So for quantum systems where $V(x)$ is finite everywhere, then both the quantum wave function $\Psi(x)$ and its derivative $d\Psi(x)/dx$ should be continuous for all values of x.

- In general there is not a unique answer to the question of where exactly a quantum particle is at a given time: all we know is what is the probability density of all possible positions x. However we can compute what is the *most likely value* of a given property to find upon measurement. For example, if we want to compute the expectation value (the mean) of an observable $f(x)$ that depends on the particle position x, we can use the following expression:
$$\langle f(x)\rangle \equiv \int_{-\infty}^{+\infty} dx\, f(x)\, P(x) = \int_{-\infty}^{+\infty} dx\, \Psi^*(x)\, f(x)\, \Psi(x)\,. \tag{1.15}$$
In the case of our example wave function Eq. (1.8), we can compute the most likely value of the position x of our particle as follows:
$$\langle x\rangle = \int_0^L dx\, x\, \frac{2}{L}\sin^2\left(\frac{\pi x}{L}\right) = \frac{2L}{\pi^2}\int_0^\pi dy\, y\, \sin^2(y) = \frac{2L}{\pi^2}\frac{\pi^2}{4} = \frac{L}{2}\,, \tag{1.16}$$
where we have used the integral in Eq. (**??**). This result implies that if we measure the position of the particle characterised by this wave function, the most likely result is that we find it in the middle of the box where is it confined, at $x = L/2$.

  Later in this lecture we will present the general formalism that allows the calculation of expectation values of arbitrary physical observables in quantum theory.

- In addition to studying the mathematical properties of the quantum wave function, it is often useful to draw it to gain some intuition of these properties. For instance, the graphical representation of the wave function $\Psi(x)$ in Eq. (1.8) and its corresponding probability density $|\Psi(x)|^2$ are shown in Fig. 1.3, assuming that $L = 1$. From there we can see how both $\Psi(x)$ and $|\Psi|^2(x)$ peak at $x = 1/2$, the midpoint of the box, consistent with the result above indicating that the most likely value for the position of the particle is $\langle x\rangle = L/2$.

As an example, the `Python` code used to produce Fig. 1.3 is the following:

```
npts=100
for i in range(0,npts+1):
    x[i] = i/npts
    phi[i] = (2)**1/2 * math.sin(math.pi*x[i])
    phi2[i] = phi[i]**2
import matplotlib.pyplot as plt
plt.plot(x,phi,color="blue",label=r"$\phi(x)$")
plt.plot(x,phi2,color="red",label=r"$\phi^2(x)$")
plt.xlabel(r"$x$")
```

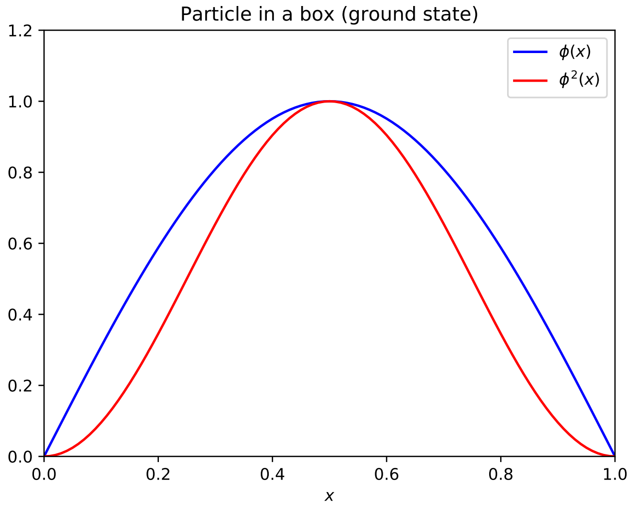


Figure 1.3: Graphical representation of the wave function $\Psi(x)$ in Eq. (1.8) for the ground state of a particle in a box and its corresponding probability density $|\Psi(x)|^2$. Here we assume that $L = 1$.

The Hilbert space.(*) To be more accurate, the wave function is quantum theory an element of an abstract vector space called the *Hilbert space*. We can project elements of this vector space into specific bases to construct different representations of the wave function. In particular, the notation $\Psi(x)$ that we have used so far is associated to the wave function of an specific *representation*, namely the *position representation*.

Since the wave function is an abstract quantity living in the *Hilbert* vector space, we can denote it using the so-called *braket* notation: $|\Psi\rangle$ (introduced by Paul Dirac). Its complex conjugate is denotes by $(|\Psi\rangle)^* = \langle\Psi|$, and their inner (scalar) product is written as $\langle\Psi|\Psi\rangle$. As we have discussed before, the wave function is normalised so that the probability of finding a particle *somewhere* must be one, and in this abstract notation this condition reads $\langle\Psi|\Psi\rangle = 1$. In order to express the wave function in specific representation, we need to project the Hilbert space elements into an specific basis. For example we can construct the position-space wave function as

$$\Psi(x) = \langle x|\Psi\rangle \,, \tag{1.17}$$

and similar for other presentations, such as the *momentum representation*.

Another important property of the quantum wave function, that we can express in this abstract Hilbert space notation, is known as the *superposition principle.* This means that is a given quantum system can exist in n different states, which we label as

$$|k\rangle\,, \quad k = 1, \ldots, n\,, \tag{1.18}$$

which k being the *quantum number* that classifies the elements of this vector space, then *any linear superposition* of the possible states of a quantum system is also a well-defined quantum state itself. In the above case, this means that a general quantum state of this system will be given by

$$|\Psi\rangle = \sum_{k=1}^{n} a_k |k\rangle\,, \tag{1.19}$$

with $\{a_k\}$ arbitrary numerical coefficients. As we will see in this course, this is a very important principle that in particular allows to construct approximate solutions for complex quantum systems where finding the exact solution is very challenging.

1.3 The Schroedinger equation

In classical physics, the motion of a free particle is described by the second Newton equations, which in one dimension reads

$$ma = m\frac{d^2x(t)}{dt} = F_{\rm tot}\,, \tag{1.20}$$

with a being the particle acceleration, $x(t)$ its position at time t, and $F_{\rm tot}$ is the total force acting on the particle. For a *free particle* $F_{\rm tot} = 0$ and the solution is simply

$$x(t) = x_0 + v(t - t_0)\,, \tag{1.21}$$

where x_0 and v_0 are the initial conditions for the position and velocity of the particle at $t = 0$.

As we discussed above, for a quantum system all information is provided by its wave function $\Psi(x)$ (now back to the position representation). So we need a different equation that determines the time evolution of the wave function, namely the analog of Newton's Eq. (1.20) for the quantum case. The corresponding dynamical equation in quantum theory is known as the *Schroedinger equation.* Here we will restrict ourselves to the time-independent Schroedinger equation, which determines the wave function of a quantum system in terms of its properties such as the local potential energy. In the simplest case of motion in one dimension, the time-independent Schroedinger equation read

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi(x)}{dx^2} + V(x)\Psi(x) = E\Psi(x)\,, \tag{1.22}$$

where $V(x)$ is the potential experienced by the particle when at position x and E is its total energy (the sum of kinetic plus potential energy). In this equation we find $\hbar \equiv h/2\pi = 1.054 \times 10^{-34}$, which is known as the *reduced Planck's constant.* In the case of a particle moving in three dimensions, the corresponding form of the time-independent Schroedinger equation reads as follows:

$$-\frac{\hbar^2}{2m}\left(\frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \frac{d^2}{dz^2}\right)\Psi(\vec{x}) + V(x, y, z)\Psi(\vec{x}) = E\Psi(\vec{x})\,. \tag{1.23}$$

A general quantum system will be defined by the masses m_i of the particles that compose it and the potential $V(x_1, x_2, \ldots)$ that these particles experience, as well as by suitable boundary conditions. By *boundary conditions* we mean requirements that the wave function must obey in some regions of space, for example, that the wave function should be zero in regions where the potential is infinite, $V(x) = \infty$. Solving a quantum system then means finding the wave functions that satisfy the Schroedinger equation subject to these conditions.

The free particle. The simplest quantum system is a free particle moving in one dimension. For such free particle the potential vanishes, $V(x) = 0$, and therefore the equation that we need to solve is:

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi(x)}{dx^2} = E\Psi(x)\,. \tag{1.24}$$

This equation is a second-order ordinary differential equation (ODE). To solve it, one can try a test solution of the form

$$\Psi(x) = \sin(Bx)\,, \tag{1.25}$$

which is a solution of Eq. (1.24) provided that

$$\begin{aligned} \frac{d}{dx}\Psi(x) &= B\cos(Bx)\,, \\ -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\Psi(x) &= -\frac{\hbar^2}{2m}\left(-B^2\sin(Bx)\right) = \frac{\hbar^2B^2}{2m}\sin(Bx)\,, \\ E &= \frac{\hbar^2B^2}{2m} \quad \to \quad B = \pm\sqrt{\frac{2mE}{\hbar^2}}\,. \end{aligned} \tag{1.26}$$

The more general solution is therefore of the plane-wave form:

$$\Psi(x) = A\cos\left(\sqrt{\frac{2mE}{\hbar}}x\right) + B\sin\left(\sqrt{\frac{2mE}{\hbar}}x\right)\,. \tag{1.27}$$

The coefficients A and B in Eq. (1.27) are free parameters to be fixed by the specific boundary conditions of the problem. This result shows that under exactly the same conditions, a particle propagates rather differently in classical theory, Eq. (1.20) as compared to quantum theory, Eq. (1.27).

In Fig. 1.4 we show the same graphical representation as in Fig. 1.3 now for the wave function of the free particle, Eq. (1.27), for $A = 0$ and $B = 1$, with an energy chosen so that $2mE/\hbar^2 = 1$. We can see clearly the periodic nature of the wave function, with the probability of finding the particle at the position x, $P(x) = |\Psi(x)|^2$, peaking at fixed separations. You might have noticed that the free-particle wave function Eq. (1.27) cannot be normalised: the integral of $|\Psi(x)|^2$ between $-\infty$ and ∞ is clearly infinite. This means that a free particle is not really a physical state. This fact is related to the Heisenberg uncertainty principle, which we discuss next.

1.4 Heisenberg's uncertainty principle

In quantum theory, certain pairs of physical variables, named *conjugated variables*, cannot be measured simultaneously with arbitrarily good precision. This limitation is not a matter of using measurement apparatus that have a finite resolution. It is actually a matter of principle: this statement holds true even for an ideal measurement apparatus without any experimental uncertainty associated. This condition restricts the amount of information that we can access for a specific quantum system, and is known as the *Heisenberg's uncertainty principle*. This principle is one of the building blocks of quantum mechanics.

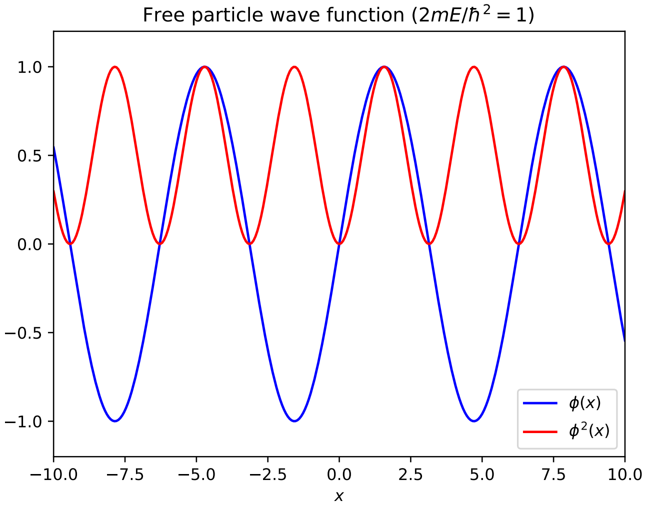


Figure 1.4: Same as Fig. 1.3 for the wave function of the free particle, Eq. (1.27), for $A = 0$ and $B = 1$, with an energy chosen so that $2mE/\hbar^2 = 1$. Note that this wave function cannot be normalised, highlighting that these solutions do not correspond to physical systems.

Perhaps the most important incarnation of the Heisenberg's uncertainty principle is the one that relates the position and the linear momentum of a particle. If we denote by Δx the uncertainty associated to the position of a given particle, and Δp_x the uncertainty associated to its linear momentum (in the same direction), quantum theory tells us that both Δx and Δp_x cannot be arbitrarily small at the same time. Instead, they need to satisfy the following inequality:

$$\Delta x \cdot \Delta p_x \geq \hbar \,, \tag{1.28}$$

in terms of the reduced Planck's constant. This means, for example, that if we aim to measure the momentum of a particle with resolution Δp, then we will not be able to determine its position with a resolution better than

$$\Delta x \geq \frac{h}{2\pi} \frac{1}{\Delta p_x} \,. \tag{1.29}$$

We emphasize again that Eq. (1.28) is not merely a restriction due to the imprecision of our experimental measurement apparatus: this is a fundamental limitation about the information that can be extracted from a quantum system even with a perfect measurement. The same principle holds for other variables, such as energy and time, where we have that

$$\Delta E \cdot \Delta t \geq \hbar \,. \tag{1.30}$$

This inequality implies that in order to measure the energy of a particle with perfect precision, the measurement would take infinite time, since $\Delta t \geq (\hbar/\Delta E)$.

It is important to emphasize at this point that not all physical variables are conjugate among them: for example, while there is a limitation in how well we can measure simultaneously the position and momentum of a given particle in the same direction,

$$\Delta y \cdot \Delta p_y \geq \hbar \,, \quad \Delta z \cdot \Delta p_z \geq \hbar \,, \tag{1.31}$$

this restriction is not present to measure the position in a given direction and its linear momentum in an orthogonal direction, for example we can have that

$$\Delta y \cdot \Delta p_x \geq 0\,, \quad \Delta x \cdot \Delta p_z \geq 0\,. \tag{1.32}$$

meaning that we can measure the position of a particle in one direction at the same time that its linear momentum in an orthogonal direction with arbitrarily good precision.

Heisenberg's principle is a radical departure from the classical picture, where the limits in which how well we can measure arbitrary physical quantities are only determined by the resolution of our measurement apparatus. How come then that classically we can measure the properties of any given object to any precision?

Because the restriction imposed by Heisenberg's uncertainty principle is only relevant when either distances or momenta are *very small*, else it is inconsequential. For example, let's consider the same car as in the previous example, with $m = 1000$ kg and $v = 100$ km/h, so that $p = 27.8 \times 10^3\ \mathrm{kg} \cdot \mathrm{m/s}$. So if one measures this momentum with precision better than *one part in a billion*, with $\Delta p = 10^{-12} \times p$, then Heisenberg's uncertainty principle implies that the uncertainty in the knowledge of the position of the car will be

$$\Delta x = \hbar/\Delta p = \left(1.05 \times 10^{-34}\,\mathrm{m^2 kg/s}\right) / \left(27.8 \times 10^{-8}\,\mathrm{kg}\cdot\mathrm{m/s}\right) = 3.8 \times 10^{-28}\,\mathrm{m}\,, \tag{1.33}$$

which certainly is irrelevant from the practical point of view.
On the other hand, for the electron with momentum $p = m \cdot v = 2.73 \times 10^{-24}\,\mathrm{kg\,m/s}$, if we measure its momentum with a 10% uncertainty, then the position of the electron will have an uncertainty of

$$\Delta x = \hbar/\Delta p = \left(1.05 \times 10^{-34}\,\mathrm{m^2 kg/s}\right) / \left(2.73 \times 10^{-25}\right) = 3.8 \times 3.8^{-10}\,\mathrm{m}\,, \tag{1.34}$$

which is of the same order of the size of an atom, thus certainly relevant in practice.

Heisenberg's uncertainty principle is also useful to understand what happens with the free-particle wave function. We saw that this wave function, Eq. (1.27), could not be normalised. The reason is that this wave function corresponds to a particle with well defined kinetic energy E and thus well-defined linear momentum $p_x = \sqrt{2mE}$ (since for a free particle $E = mv^2/2 = (mv)^2/2m = p^2/2m$). Therefore, the free-particle wave function corresponds to a quantum state with a well-definite momentum, and therefore characterised by $\Delta p_x = 0$. However, in this case Heisenberg's uncertainty principle implies that the uncertainty in the position of the particle will be

$$\Delta x \geq \hbar/\Delta p_x \geq \infty \tag{1.35}$$

so the particle is maximally delocalised: this is why it is not possible to normalise its wave function.

Commutation relations. Heisenberg's uncertainty principle illustrates an important concept of quantum theory: the *complementarity between physical observables.* We define two observables Ω_1 and Ω_2 to be *complementary* if the sequential application of the corresponding operators $\hat{\Omega}_1$ and $\hat{\Omega}_2$ onto a generic wave function Ψ does not *commute*, that is,

$$\hat{\Omega}_1\left(\hat{\Omega}_2\Psi\right) \neq \hat{\Omega}_2\left(\hat{\Omega}_1\Psi\right)\,, \tag{1.36}$$

or in other words, the results of applying the two operators $\hat{\Omega}_1$ and $\hat{\Omega}_2$ to a generic wave-function depends on the *order in which they have been applied*. It can be demonstrated that for each pair of complementary observables there will be a relation of the form of the Heisenberg uncertainty principle, Eq. (1.28), implying that a simultaneous measurement of the physical quantities Ω_1 and Ω_2 with arbitrary precision is not possible. As mentioned above, two examples of complementary variables are position and momentum in the same direction, $\Omega_1 = x$ and $\Omega_2 = p_x$, as well as time and total energy, $\Omega_1 = t$ and $\Omega_2 = E$.

The result that in quantum theory the subsequent application of operators associated to physical observables to a quantum wave function does not necessarily commute can be represented introducing another useful mathematical concept, the *commutator* between two operators. This commutation operator is defined as

$$\left[\hat{\Omega}_1, \hat{\Omega}_2\right] \equiv \hat{\Omega}_1\hat{\Omega}_2 - \hat{\Omega}_2\hat{\Omega}_1 \, , \tag{1.37}$$

where note that the commutator of two operators *is another operator*, that is, it should be understood as being applied to a wave-function. With this definition, for the position and momentum operators we find that their commutator is given by

$$[\hat{x}, \hat{p}_x] = i\hbar \, , \tag{1.38}$$

as is easy to verify explicitly. Indeed, since

$$\hat{x}\hat{p}_x\Psi = x\left(\frac{\hbar}{i}\frac{d}{dx}\right)\Psi = \frac{\hbar}{i}x\frac{d\Psi}{dx} \, , \tag{1.39}$$

$$\hat{p}_x\hat{x}\Psi = \left(\frac{\hbar}{i}\frac{d}{dx}\right)(x\Psi) = \frac{\hbar}{i}\Psi + \frac{\hbar}{i}x\frac{d\Psi}{dx} \, , \tag{1.40}$$

then their commutator gives

$$[\hat{x}, \hat{p}_x]\,\Psi = -\frac{\hbar}{i}\Psi = i\hbar\Psi \, , \tag{1.41}$$

which is the sought-for result Eq. (1.38). Therefore, x and p are *complementary observables* (their operators do not commute) and thus they cannot be measured simultaneously with arbitrary precision. This is another way of representing Heisenberg's uncertainty principle.

In general, for any two operators representing physical quantities, if one finds that

$$\left[\hat{\Omega}_1, \hat{\Omega}_2\right] \neq 0 \, , \tag{1.42}$$

that is, that the two operators do not commute, then the physical observables to which these two operators are associated are complementary and have associated a Heisenberg uncertainty relation.

1.5 Quantum entanglement (*)

A remarkable consequence of the basic principles of quantum theory, particularly at odds with our intuition that any system composed by many particles can be understood from the sum of the properties of individual particles, is that there exists many-particle systems of many particles which *cannot be expressed as a combination of single-particle states*. This is the opposite as with the classical view of the world, where any system, no matter how complex it is, can always be decomposed in all of its components, each one obeying well-defined dynamical equations. In quantum theory this is not necessarily true, and some states (actually the wide majority) do not admit a description in terms of quantum states of individual particles. Those states are said to be *entangled*.

Let us consider the case in which a particle can occupy only two states, so our basis vectors will be labelled as $|0\rangle$ and $|1\rangle$. It does not really matter what $|0\rangle$ and $|1\rangle$ stand for: could be two different values of the energy, or two different orientations of the internal angular momentum of a particle, the following argument does not depend on which physical system we consider. We will write the state composed by two of such particles as follows

$$|ab\rangle \equiv |a\rangle_1 |b\rangle_2 \,, \tag{1.43}$$

so the state $|ab\rangle$ is such particle 1 one is in the state $|a\rangle$ and particle 2 in the state $|b\rangle$. Is clear that some of the possible states admit a description in terms of single-particle states, for example, one can construct

$$|\Psi\rangle = |01\rangle = |0\rangle_1 |1\rangle_2 \,, \tag{1.44}$$

but however other states do not admit such description, for example, the so-called *Bell state*:

$$|\Psi\rangle_B = \frac{1}{\sqrt{2}} \left(|00\rangle + |11\rangle \right) \,. \tag{1.45}$$

This state cannot be decomposed into something like "state of particle 1" combined with "state of particle 2": it has no classical analog.

To demonstrate this interesting property, note that if the quantum state $|\Psi\rangle_B$ could be decomposed in terms of single particle states, one should be able to write it (using the superposition principle) as a linear combination of quantum states with well-defined values for $|a\rangle_1$ and $|b\rangle_2$, namely

$$|\Psi\rangle = \left(a|0\rangle_1 + b|1\rangle_1 \right) \left(c|0\rangle_2 + d|1\rangle_2 \right) = ac|00\rangle + ad|01\rangle + bc|10\rangle + db|11\rangle \,, \tag{1.46}$$

but you can easily convince yourself that Eq. (1.45) does not admit a decomposition of the form of Eq. (1.46). For example, either a or d must be zero since the ket $|01\rangle$ is absent, but this is inconsistent with the kets $|00\rangle$ and $|11\rangle$ being present. Therefore, in the quantum state Eq. (1.45) it is not possible to describe separately the individual particles: they are fundamentally *entangled* with each other.

There exist many ways of quantifying the *entanglement* that is present in a given quantum system. For example, our two-particle quantum state can be generically written as

$$|\Psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle \,, \tag{1.47}$$

where $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$ due to normalization. Now, we can define a measure of entanglement in terms of its *concurrence* as follows

$$\Delta \equiv 2 \left| \alpha\delta - \beta\gamma \right| \,, \tag{1.48}$$

By construction, $0 \le \Delta \le 1$, where the extreme cases $\Delta = 0$ (1) correspond to a product (maximally entangled) state. You can convince yourselves that Eq. (1.45) is not only an entangled state, but actually it corresponds to a *maximally entangled* state with $\Delta = 1$. Likewise, any state that can be decomposed into single particle states such as Eq. (1.46) has $\Delta = 0$, so exactly zero entanglement.

Entanglement is a very powerful feature of quantum theory. For instance, the exponential speed up in computational efficiency that *quantum computers* exhibit over the classical ones arises from the manipulation and exploitation of entanglement. Entanglement also plays a crucial role in many theories of elementary particle physics.

1.6 Observables and operators in quantum theory

As we have discussed, the wave-function Ψ of a quantum system contains the complete physical information on the system. In particular, the various properties of the system should be calculable only from the knowledge of the wave-function, without any additional inputs required. In quantum theory, we denote as an *observable* a given property of a quantum system (energy, momentum, angular momentum, electric charge, ...) that can be extracted from a *experimental measurement* of this system. Mathematically, the tools that we will use to compute the expected outcome of a measurement of a quantum system are known as *operators*, which can be expressed in the language of *eigenfunctions* and *eigenvectors*. So before discussing the role that operators have in quantum theory, let's briefly review the concept of an eigenvalue system with matrices.

Eigenvalue equations with matrices. As you have seen in linear algebra courses, an eigenvalue equation is a mathematical equation of the form

$$A\mathbf{v} = \lambda\mathbf{v}. \tag{1.49}$$

where A is an square matrix of dimensions $n \times n$ and $\mathbf{v}$ is a column vector with n dimensions. The scalar λ is known as the *eigenvector* of the equation. So the key feature of Eq. (1.56) is that applying a matrix A to the vector $\mathbf{v}$ returns the original vector with an overall rescaling, $\lambda\mathbf{v}$. In order to compute the eigenvalues of a matrix, we need to evaluate the solutions of the *characteristic equation* of the matrix A, given for

$$\det\left(A - \lambda \cdot I\right) = 0\,, \tag{1.50}$$

where I is the identity matrix of dimensions $n \times n$, and det is the determinant.

To illustrate how to compute eigenvalues, let's consider the following matrix

$$A = \begin{pmatrix} 1 & -3 \\ 3 & -5 \end{pmatrix} \tag{1.51}$$

which has associated the following characteristic equation

$$\det\left(A - \lambda \cdot I\right) = \begin{vmatrix} 1-\lambda & -3 \\ 3 & -5-\lambda \end{vmatrix} = \lambda^2 + 4\lambda + 4 = 0\,. \tag{1.52}$$

We can easily solve this quadratic equation finding that the two eigenvalues associated to the matrix A are $\lambda_1 = \lambda_2 = -2$. So in this case the two eigenvalues are the same.

Once we know the eigenvalues λ_i associated to a given matrix A, we can compute the corresponding eigenvectors $\mathbf{v}_i$, defined as the vectors that satisfy

$$A\mathbf{v}_i = \lambda\mathbf{v}_i. \tag{1.53}$$

which is an n–dimensional system of linear equations. In our example case, the equation that needs to be solved is

$$\begin{pmatrix} 1 & -3 \\ 3 & -5 \end{pmatrix}\begin{pmatrix} v_{1,1} \\ v_{1,2} \end{pmatrix} = \lambda_1 \begin{pmatrix} v_{1,1} \\ v_{1,2} \end{pmatrix} \tag{1.54}$$

$$\begin{aligned} v_{1,1} - 3v_{1,2} &= -2v_{1,1} \\ 3v_{1,1} - 5v_{1,2} &= -2v_{1,1} \end{aligned} \tag{1.55}$$

which is solved to find $\mathbf{v}_i = (1,1)$, as you can check. 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, is also an eigenvector.

You can also solve eigenvalue equations using numerical software. For instance, using `Python` you can easily find the eigenvalues and eigenvectors of an arbitrary matrix A using the following instructions:

```
from numpy import linalg as LA
A = [[1, −3],
    [3, −5]]
# Solve the eigensystem
w, v = LA.eig(A)
# Print the results
print("eigenvalues = ",w)
print("eigenvectors = ",v)
```

Eigenvalue equations in quantum systems. Formally, an *operator* is a way to represent the action of a specific mathematical function onto the wave-function Ψ. For instance, Schroedinger's equation Eq. (1.23) can be expressed in operator form as follows

$$\hat{H}\Psi(x) = E\Psi(x)\,, \tag{1.56}$$

where the "hat" ($\hat{}$) symbol indicates an operator, in this case the *Hamiltonian operator*, defined as

$$\hat{H} \equiv -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\,. \tag{1.57}$$

Operators in quantum theory act on wave function and return either numbers or functions. An operator equation of the form of Eq. (1.56) has the same form as the eigenvalue equations that we saw before in Eq. (1.56), replacing the matrix A by the Hamiltonian operator $\hat{H}$ and the eigenvector $\mathbf{v}$ by the quantum wave function $\Psi(x)$. Actually the analogy is more transparent for those quantum systems where the number of quantum states is finite, and where the Hamiltonian operator can be written as a square matrix as we will see.

In the case of Eq. 1.56, the application of the Hamiltonian operator $\hat{H}$ to the wave-function returns the total energy E of the quantum state. This implies that $\hat{H}$ is the operator that represents the total energy of a given quantum system. Therefore, we can determine the energy of a quantum system by applying the operator $\hat{H}$ to its wave function. Note that not all wave functions are eigenvalues of $\hat{H}$, meaning that one can construct wave functions for which

$$\hat{H}\Psi(x) \neq E\Psi(x)\,. \tag{1.58}$$

In general, solving a quantum system corresponds to determining the eigenvalues and eigenvectors of Eq. 1.56, which is the equivalent to solving the Schroedinger equation.

In general, an eigenvalue equation in quantum theory will take the general form

$$\hat{\Omega}\,\Psi_k = \omega_k\,\Psi_k\,, \quad k = 1, \ldots, n \tag{1.59}$$

where we have:

- $\hat{\Omega}$ represents a generic operator associated to the observable property Ω (such as the energy or momentum) of the quantum system,
- Ψ_k is the k-th eigenfunction, corresponding to the k-th eigenvalue ω_k.
- n is the number of eigenvalues that this equation has.

Note that the eigenvalues ω are in general complex and can be either constant or functions of other variables, such as the position x.

Let us provide two examples of eigenvalue equations in quantum theory.

(a) Is $\Psi(x) = e^{\alpha\, x}$ an eigenfunction of the operator $\hat{\Omega} \equiv d/dx$? To verify if this is the case, we apply the operator to the wave-function to find

$$\hat{\Omega}\Psi = \frac{d}{dx} e^{\alpha\, x} = \alpha e^{\alpha\, x} = \alpha \Psi \,, \tag{1.60}$$

so indeed in this case $\Psi(x)$ is an eigenfunction of $\hat{\Omega}$, with α being the associated eigenvalue.

(b) Is $\Psi(x) = e^{\alpha x} + e^{\alpha\, x^2}$ an eigenfunction of the same operator? Let us check it in the same way as before:

$$\hat{\Omega}\Psi = \frac{d}{dx}\left(e^{\alpha x} + e^{\alpha\, x^2}\right) = \alpha e^{\alpha x} + 2\alpha x e^{\alpha\, x^2} \neq \omega \Psi \,, \tag{1.61}$$

thus now the wave function $\Psi(x)$ *is not* an eigenfunction, since it cannot be written as the product of an eigenvalue and the same eigenfunction once the operator $\hat{\Omega}$ acts upon it.

Now going back to the Schroedinger equation as an written eigenvalue equation, Eq. (1.56), we can write its solutions as

$$\hat{H}\Psi_k(x) = E_k \Psi_k \,. \tag{1.62}$$

Using the quantum superposition principle, we know that we have function

$$\varphi(x) = \sum_k c_k \Psi_k(x) \,, \tag{1.63}$$

with c_k complex numbers, is also a possible valid quantum state for the system. However, crucially, $\varphi(x)$ is not itself a solution of Schroedinger's equation. You can check this easily using the operator form of the equation

$$\hat{H}\varphi(x) = \sum_k c_k \hat{H}\Psi_k(x) = \sum_k c_k E_k \Psi_k(x) \neq \omega\varphi(x), \tag{1.64}$$

showing how there exist physically allowed quantum systems that are not solutions of Schroedinger's equations. The reason is that these quantum states do not have a definite energy.

Observables and operators. Therefore in quantum theory any observable property Ω of a quantum system will be represented by an operator $\hat{\Omega}$ built up. We can determine the value of this property associated to the specific quantum system Ψ by solving the corresponding eigenvalue equations $\hat{\Omega}\Psi_k = \omega_k \Psi_k$. We now list the most important cases for this correspondence between observables and operators (for simplicity in one dimension only).

Observable	**Operator**
Position: x	$\hat{x} = x$
Linear momentum: $p_x = mv_x$	$\hat{p}_x = \frac{\hbar}{i}\frac{d}{dx}$
Kinetic energy: $T = \frac{1}{2}mv_x^2 = p_x^2/2m$	$\hat{T} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}$
Potential energy: $V = V(x)$	$\hat{V} = V(x)$
Total energy: $E = T + V = p_x^2/2m + V(x)$	$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)$

Later in the course we will see how this correspondence between observables and operators in quantum theory can be extended to other cases, such as the angular momentum.

The free particle revisited. We can now revisit the free-particle solution of the Schroedinger's equation, Eq. (1.27). We want to compute which is the linear momentum associated to this quantum state. What we need to do is to apply the linear momentum operator to the corresponding wave function (for $B = 0$):

$$\hat{p}_x \Psi(x) = \frac{\hbar}{i}\frac{d}{dx}\left[A\cos\left(\sqrt{\frac{2mE}{\hbar}}x \right) \right] = -\frac{\hbar}{i}A\sqrt{\frac{2mE}{\hbar^2}}\sin\left(\sqrt{\frac{2mE}{\hbar^2}}x \right) \neq p_x\Psi(x)\,, \tag{1.65}$$

so we find that a plane wave solution of the form of Eq. (1.27) with $B = 0$ is not an eigenvector of the linear momentum operator: we cannot associate a well-defined value of p_x to this quantum state.

On the other hand, if we impose that $B = iA$ we have that the wave function reads

$$\Psi(x) = A\cos\left(\sqrt{\frac{2mE}{\hbar^2}}x \right) + iA\sin\left(\sqrt{\frac{2mE}{\hbar^2}}x \right) = A\exp\left(i\sqrt{\frac{2mE}{\hbar^2}}x \right)\,, \tag{1.66}$$

and if now we apply the momentum operator to this wave function we get

$$\hat{p}_x \Psi(x) = \frac{\hbar}{i}\frac{d}{dx}\left[A\exp\left(i\sqrt{\frac{2mE}{\hbar^2}}x \right) \right] = \sqrt{2mE}\Psi(x) \tag{1.67}$$

so the momentum along the x direction p_x that can be associated to the free particle with wave function Eq. (1.66) is $p_x = \sqrt{2mE}$. One can easily show that the solution where $B = -iA$ is also an eigenvalue of the linear momentum operator with eigenvector $p_x = -\sqrt{2mE}$. Therefore, the general solution of the Schroedinger equation for the free particle is a linear combination of plane waves moving in the positive direction with well defined momentum and the corresponding plane waves moving in the negative direction.

This calculation indicates that not all the solutions of the Schroedinger's equation for a free particle are eigenvectors of the linear momentum operator. In other words, one can have states that have well-defined energy E but not well-defined linear momentum p_x. This is a consequence of the fact that the linear momentum has sign and direction, while the kinetic energy depends only on its modulus, $E \propto |\vec{p}|^2$.

So does this mean that free particles always have well-defined values of their linear momentum p_x? Does this imply that the position of a free particle is always maximally uncertain? No, if we take into account that physically, from Eq. (1.28), there should be an uncertainty associated also to the expectation value of $\hat{p}$, the momentum of the quantum state, if we want to have the particle localised in a finite region of space. A wave function for a free particle with a *finite momentum resolution* (and therefore a certain amount of spatial localisation) can be constructed by means of a *superposition* of free-particle solutions, each with *different*

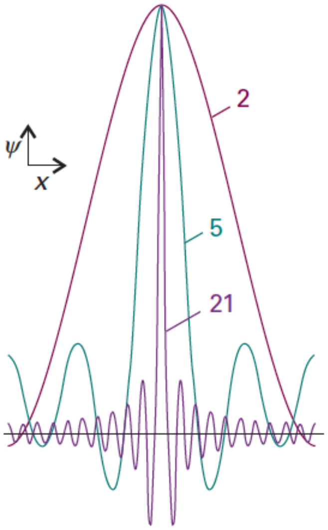


Figure 1.5: The wave function Eq. (1.69) corresponding to the superposition of N plane-wave free particle solutions, each with a different value of $k = n\pi$. We show the results from the addition of $N = 2, 5, 21$ terms in Eq. (1.69), showing that the more terms we include, the better the localization of the free particle (since the higher is the indetermination of its momentum p_x).

wave numbers k, namely

$$\Psi(x) = \sum_{k}^{N} c_k e^{ikx} \,. \tag{1.68}$$

The more the number of contributions N to the above sum, the higher the uncertainty on the momentum of the particle, and thus we obtain an increased localization of the particle in x. This can be seen by plotting the free-particle wave-function superposition:

$$\Psi(x) = \sum_{n=1}^{N} \cos(n\pi x) \,, \tag{1.69}$$

which is shown in Fig. 1.5: the more terms we add in the sum, the more values the momentum $\hbar k$ of the particle can take, hence Δp is higher (more uncertain) and Δx smaller (increased localization in space).

Hermiticity and orthogonality. In quantum mechanics, observable properties of a system have associated operators, whose eigenvalues correspond to the values than these observables can take when the measurement is performed in a quantum state described by an eigenfunction of this operator. However, not all operators will lead to physically sensible observables. In particular, only operators which are *Hermitian* are physically allowed. Hermitian operators have important properties, in particular in this course we will exploit the fact that:

- Hermitian operators have always associated *real eigenvalues* as opposed to complex ones.
- The eigenfunctions of Hermitian operators are always *orthogonal* among them: the overlap integral between two of such eigenfunctions vanishes.

What is the defining characteristic of an Hermitian operator? We say that an operator $\hat{\Omega}$ is Hermitian provided that, given two arbitrary wave-functions Ψ_i and Ψ_j, it satisfies the following property:

$$\int dx \Psi_i^* \hat{\Omega} \Psi_j = \left(\int dx \Psi_i^* \hat{\Omega} \Psi_j \right)^{\dagger} , \tag{1.70}$$

where the "dagger" symbol † indicates to to interchange i and j (that is, taking the transpose matrix), and then taking the complex conjugate, that is

$$\left(\int dx \Psi_i^* \hat{\Omega} \Psi_j \right)^{\dagger} = \int dx \left(\Psi_j^* \hat{\Omega} \Psi_i \right)^* . \tag{1.71}$$

One can verify that all the operators that we have introduced so far, such as the position $\hat{x}$ and the linear momentum $\hat{p_x}$ operators are Hermitian. For example, in the latter case we have

$$\int dx \Psi_i^* \, \hat{p} \, \Psi_j = \int dx \Psi_i^* \left(\frac{\hbar}{i} \frac{d}{dx} \right) \Psi_j = \frac{\hbar}{i} \left(\Psi_i^* \Psi_j \Big|_{-\infty}^{+\infty} - \int dx \Psi_j \frac{d}{dx} \Psi_i^* \right) = \int dx \Psi_j \left(-\frac{\hbar}{i} \right) \frac{d}{dx} \Psi_i^* \tag{1.72}$$

where we have used integration by parts, and exploited the facts that wave-functions must vanish at infinity (since they must be square-integrable). Therefore we have demonstrated that

$$\int dx \Psi_i^* \, \hat{p} \, \Psi_j = \int dx \Psi_j \left(-\frac{\hbar}{i} \right) \frac{d}{dx} \Psi_i^* = \left(\int dx \Psi_i^* \, \hat{p} \, \Psi_j \right)^{\dagger} . \tag{1.73}$$

and thus the linear momentum operator $\hat{p}$ is Hermitian.

In quantum theory, all physical observables (associated to quantities that can be accessed experimentally) are represented by Hermitian operators. This ensures that the expectation values of these observables in physical systems (as well as the outcome of any individual measurement) are real instead of complex (which could not be otherwise, since *i.e.* a physical system cannot have a complex energy or a complex position).

The key property that the eigenvalues of an Hermitian operator $\hat{\Omega}$ are real can be demonstrated from its definition Eq. (1.70) as follows. Let us assume that $\Psi_i = \Psi_j = \psi$ is an eigenfunction of this operator with eigenvalue ω, that is, the relation $\hat{\Omega}\psi = \omega\psi$ holds. Then we have that

$$\int dx \psi^* \Omega \psi = \int dx \psi^* \omega \psi = \omega \int dx \psi^* \psi = \omega \, , \tag{1.74}$$

since the wave-function is normalized to one, and that also we have that the "dagger" version of the above expression gives

$$\left(\int dx \psi^* \Omega \psi \right)^{\dagger} = \left(\omega \int dx (\psi^* \psi)^* \right) = \omega^* \tag{1.75}$$

and since $\hat{\Omega}$ is Hermitian, then it follows that $\omega^* = \omega$. In other words, ω is a real number irrespective of the specific operator that we use provided that it is Hermitian.

It is also possible to show that, for an Hermitian operator, its eigenfunctions are *orthogonal* among them. In other words, they constitute a *set of linearly independent basis functions* in the sense of a vector space in linear algebra. Mathematically, the orthogonality property implies that if $\Psi_i(x)$ and $\Psi_j(x)$ are two eigenfunctions of the operator $\hat{\Omega}$, then the following relation holds:

$$\int dx\, \Psi_i^*(x)\, \Psi_j(x) = \delta_{ij}\,, \tag{1.76}$$

that is, the integral vanishes unless the eigenfunctions are the same, $i = j$, in which case the integral equals one because of the normalization condition (the eigenvectors can always be normalised to ensure that Eq. (1.76) is satisfied). This important property allows to project a general quantum state as a *linear superposition* of the eigenfunctions of an Hermitian operator, which form a complete basis of the corresponding vector space.

1.7 Expectation values of physical quantities

The wave-function Ψ can be used to determine the *expectation value* of physical observables of a given quantum system, that is, the expected average over many measurements of this quantity. This is a familiar concept from probability theory: if we measure a given quantity b n times, say b_i, then the mean value over the measurements is

$$\langle b_i \rangle = \frac{1}{n} \sum_{i=1^n} b_i\,, \tag{1.77}$$

and represents our best estimate of the *true value* of the observable b. In quantum theory, we cannot predict in general the outcome of individual measurements, but we can predict what will be the average over a large number of measurements.

In quantum theory, if we have an observable Ω associated to an operator $\hat{\Omega}$, we can compute its expectation value in a quantum state characterised by the wave function $\Psi(x)$ as follows

$$\langle \Omega \rangle = \int dx\, \Psi^* \hat{\Omega} \Psi\,, \tag{1.78}$$

Here we can have two possible situations:

- The first one is that the wave function Ψ is already an eigenvector of the operator $\hat{\Omega}$, namely $\hat{\Omega}\Psi = \omega\Psi$ In that case we have that the expectation value is given by

  $$\langle \Omega \rangle = \int dx \Psi^* \hat{\Omega} \Psi = \int dx \Psi^* \omega \Psi = \omega \int dx \Psi^* \Psi = \omega\,, \tag{1.79}$$

  so it coincides with the corresponding eigenvalue. This is the simplest possible situation: if we measure the observable Ω in the system characterised by this wave function, then the outcome of the measurement will always be the same: ω.

- In general however the wave function $\Psi(x)$ will not be an eigenvector of the operator $\hat{\Omega}$. In this case, we can use the fact that the eigenvectors of $\hat{\Omega}$, defined by the condition

  $$\hat{\Omega}\Psi_k = \omega_k \Psi_k\,, \quad k = 1, \ldots, n \tag{1.80}$$

  form a complete basis in the Hilbert vector space. This means that I can always expand the original wave function $\Psi(x)$ as a linear combination of the eigenvectors Ψ_k, in the same way that if I have a basis in a vector space I can always expand a general vector in terms of a linear combination of the elements of the basis. In other words, I can always construct an expansion of the form

$$\Psi(x) = \sum_{k=1}^{n} c_k \Psi_k(x) \, . \tag{1.81}$$

with c_k being complex numerical coefficients. The coefficients c_k can be determined by exploiting the condition that the eigenvectors of an Hermitian operator form a complete basis:

$$\int dx \, \Psi^*(x)\Psi_j(x) = \int dx \left(\sum_{k=1}^{n} c_k \Psi_k(x) \right)^* \Psi_j(x) = \sum_{k=1}^{n} c_k^* \int dx \, \Psi_k^*(x)\Psi_j(x) = c_j^* \, , \tag{1.82}$$

since the $\{\Psi_k\}$ eigenvectors are orthonormal among them.

Therefore, in full generality, we can always expand a quantum wave function as in Eq. (1.81) in terms of the eigenvectors of the operator associated to the physical observable that we want to measure. In this case, the expectation value $\langle \Omega \rangle$ will be given by the weighted sum of eigenvalues, with weight $|c_k|^2$:

$$\langle \Omega \rangle = \int dx \left(\sum_k c_k \Psi_k \right)^* \hat{\Omega} \left(\sum_j c_j \Psi_j \right) = \sum_{k,j} c_k^* c_j \omega_j \int dx \Psi_k^* \Psi_j = \sum_{k,j} c_k^* c_j \omega_j \delta_{ij} = \sum_k |c_k|^2 \omega_k \, , \tag{1.83}$$

where we have use the orthogonality property of the eigenfunctions of an Hermitian operator Eq. (1.76). Note that in this derivation we assume that the coefficients of the linear superposition Eq. (1.81) are canonically normalized, that is, they satisfy

$$\sum_k |c_k|^2 = 1 \, , \tag{1.84}$$

as implied by the condition that the original wave function $\Psi(x)$ is appropriately normalised.

Therefore, if we know the expansion of a given wave function in terms of the eigenvectors of an Hermitian operator, Eq. (1.81), we can readily compute the expectation value of this operator as a weighted sum of its eigenvalues as indicated by Eq. (1.83).

1.8 Radioactivity (*)

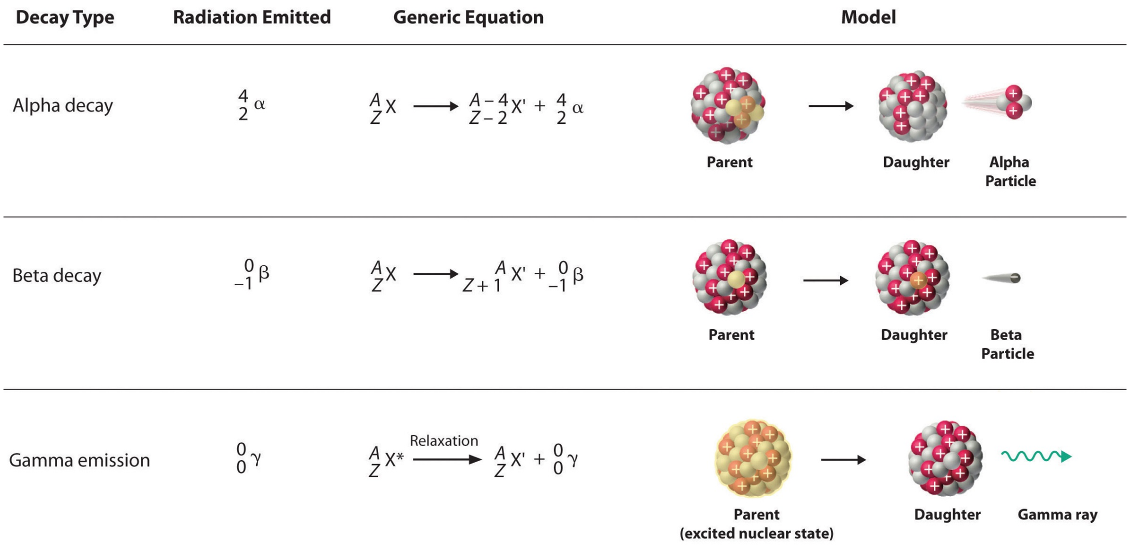
The concepts of *radioactivity* and *radiation* are intrinsically linked to quantum theory, and will appear frequently along the various topics covered in this course. By radioactivity we understand the spontaneous emission of energetic particles from unstable atoms. This phenomenon is a direct consequence of quantum theory. Radioactivity is an very important phenomenon in nuclear and particle physics, which has lead to many applications such as nuclear energy as well as to important medical imaging applications.

There are three main types of radioactive processes:

- α *radiation*: the emission of a He nucleus (a bound state of two protons and two neutrons).

  This radioactive decay process is mediated by the strong nuclear force. If the initial unstable element is ${}^{A}_{Z}X$, where Z is the *atomic number* (number of protons) and A is the *mass number* (numbers of protons plus neutrons), then the daughter element will be ${}^{A-4}_{Z-2}Y$, given the proton and neutron content of an α particle

- β *radiation*: the emission of an electron (or its anti-particle, the positron), together with a neutrino.



Decay Type	Radiation Emitted	Generic Equation	Model
Alpha decay	${}^{4}_{2}\alpha$	${}^{A}_{Z}\mathrm{X} \longrightarrow {}^{A-4}_{Z-2}\mathrm{X}' + {}^{4}_{2}\alpha$	Parent → Daughter Alpha Particle
Beta decay	${}^{0}_{-1}\beta$	${}^{A}_{Z}\mathrm{X} \longrightarrow {}_{Z+1}^{A}\mathrm{X}' + {}^{0}_{-1}\beta$	Parent → Daughter Beta Particle
Gamma emission	${}^{0}_{0}\gamma$	${}^{A}_{Z}\mathrm{X}^* \xrightarrow{\text{Relaxation}} {}^{A}_{Z}\mathrm{X}' + {}^{0}_{0}\gamma$	Parent (excited nuclear state) → Daughter Gamma ray

Figure 1.6: The three main types of radioactive decay processes. In alphas decay, mediated by the strong interaction, a He nucleus (2 protons and 2 neutrons) is emitted. In beta decay, mediated by the weak interaction, either an electron or a positron are emitted (together with a neutrino). In gamma decay, mediated by electromagnetism, a high energy photon is emitted.

This radioactive decay process is mediated by the weak interaction. If the initial unstable element is ${}^{A}_{Z}X$, for this process the final element will be ${}^{A}_{Z\pm1}Y$, depending on whether an electron or a positron has been emitted.

- γ radiation: the emission of a high energy photon γ.

  This radioactive decay process is mediated by the *electromagnetic interaction*. In this case, the values of the atomic Z and mass A numbers will be the same for the initial and final element: actually this transition consist of the relaxation of an excited atomic state to a state with less energy.

Therefore each of these three radioactive processes is mediated by a different fundamental interaction. The mean features of these three kinds of radioactive processes are summarized in Fig. 1.6.

You might ask yourselves: what causes a specific atom to decay? And how does exactly the atom know at which time it has to decay? Since radioactive decays are determined by quantum theory, at this point you might have already figured out that they are intrinsically probabilistic in nature. In other words, quantum theory allows us to predict when average a large collection of atoms will decay, but we cannot know when a given specific atom will decay. That is, in quantum theory we can compute the expectation value of the time it takes for a given type of atoms to decay radioactively.

If the number of atoms in a given radioactive sample is $N(t)$, and we have $N_0 = N(t=0)$ atoms to begin with, the infinitesimal variation dN in a time interval dt will be proportional to the number of atoms themselves, so it will be given by $-\lambda N(t)dt$, with λ known as the *decay constant* of this specific radioactive process. Solving the associated differential equation we see that the number $N(t)$ of radioactive atoms at any given time will be

$$dN = -\lambda N(t)dt \quad \rightarrow \quad N(t) = N_0 e^{-t/\tau}\,, \tag{1.85}$$

with $\tau = 1/\lambda$ representing the *average lifetime* of the atoms, that is, the average time they take to decay. To validate this interpretation, note that the average of a stochastic variable t with probability density $N(t)$

is given by

$$\langle t\rangle = \frac{\int_0^\infty dt\, t\, N(t)}{\int_0^\infty dt\, N(t)} = \tau\,. \tag{1.86}$$

Therefore, we see from Eq. (1.85) that the number of nuclei of a given sample that undergo radioactive decays decreases exponentially with time. One sometimes also defines the activity of a radioactive sample, defined as (minus) the rate of change of $N(t)$ per unit time,

$$A(t) \equiv -\frac{dN(t)}{dt} = \frac{N_0}{\tau}\exp\left(-\frac{t}{\tau}\right)\,. \tag{1.87}$$

The higher the value of the activity $A(t)$, the faster the radioactive decay process takes place.

The value of the decay constant λ is different for each atomic element and isotope. This provides among other things a rather useful method to determine the age of objects, the so-called *radiocarbon dating method*. This method is based on the fact that one of the isotopes of carbon, ^{14}C, has a decay constant of a few thousand years, while the more frequent isotope ^{12}C is stable. This means that if we know somehow the initial abundance of ^{14}C in a given sample, and measure the current abundance, since we know τ we can determine t from Eq. (1.85). In other words, we can determine in a rather reliable way the age of this specific sample. The original abundance of ^{14}C can be determined for instance by comparing with the current one of ^{12}C, since the ratio between the abundances of the two isotopes in various systems is rather well understood.

Let us illustrate the radiocarbon dating method with an example. The mean-life of the unstable isotope of Carbon, ^{14}C, is $\tau = 8267$ years. If we assume that a given sample had the same ratio $^{14}C/^{12}C$ as the one we have now in the atmosphere, we can compute the value N_0 of ^{14}C atoms in the original sample. By comparing with the current value of ^{14}C atoms in the sample, $N(t)$, we can then determine t.

For instance, assume that we find that a sample with only 10% of ^{14}C atoms as compared to its original abundance. By applying Eq. (1.85), we have that

$$0.1\, N_0 = N_0 e^{-t/(8267\,\mathrm{yr})} \qquad \to \qquad t = 19000\,\mathrm{yr}\,, \tag{1.88}$$

illustrating how radioactive dating can be used to determine the age of a given object.

1.9 Summary

We can now recapitulate what we have learned in these two lectures about the basic postulates and concepts of quantum mechanics. These postulates will be extensively used in the rest of the course, and applied to a wide variety of different quantum systems.

I/ All the physical information about a quantum system is contained in its wave-function $|\Psi\rangle$, which lives in an abstract vector space called the Hilbert space.

II/ In this course we will work mostly with the position representation of the wave function, $\Psi(x) = \langle x|\Psi\rangle$.

III/ The square of the wave-function $|\Psi(x)|^2$ represents the probability density for finding the particle at the position x upon a measurement.

IV/ The wave-function $\Psi(x)$ must be continuous, have a continuous first derivative (unless the potential is infinite at the boundary), be single valued, and be squared-integrable.

V/ Physical observables Ω are represented by Hermitian operators $\hat{\Omega}$ built upon the position $\hat{x}$ and momentum $\hat{p}$ operators, that act on the wave function of a quantum system. Examples of this are the total energy $\hat{H}$ and the linear momentum $\hat{p}_x$.

VI/ For time-independent systems with defined energy E_k, the wave function is the solution of Schroedinger's equation, which in operator form reads $\hat{H}|\Psi_k\rangle = E_k|\Psi_k\rangle$.

VII/ By means of the superposition principle, we can always express a general wave function as a combination of eigenvectors of Schroedinger's equation,

$$|\Psi\rangle = \sum_n c_k|\Psi_k\rangle \, . \tag{1.89}$$

The same type of expansion holds for the eigenvectors of any other Hermitian operator.

VIII/ Some physical observables are complementary, that is, they can not be measured simultaneously with arbitrary precision (even for a perfect measurement apparatus). Mathematically, their associated observables do not commute. Complementary observables obey relations of the form of Heisenberg's uncertainty principle.

IX/ The knowledge of the wave function of a system allows us to evaluate the expectation value $\langle\Omega\rangle$ of any observable quantity Ω by computing matrix elements of its associated operator $\hat{\Omega}$.

van Quantum tot Molecuul

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

2 HC3: The quantum mechanics of simple systems

In the previous lectures we have studied the solutions of the Schroedinger equation of the simplest possible quantum system: the free particle in one dimension. In this lecture, we discuss quantum mechanics applied to relatively simple systems such as a particle in a confining box and the quantum harmonic oscillator. We will study some remarkable phenomena that appear in these systems that do hot have an analog in classical mechanics, such as that the *the quantization of energies* and the *quantum tunneling* effect.

The **learning goals** of this lecture are:

(a) Learning how to solve the Schroedinger equation for simple quantum systems.

(b) Understanding that confinement of a quantum particle in space leads to the quantization of its energy levels and to a non-vanishing energy for the ground state.

(c) Interpreting the physical content of the wave-functions from the solutions of the Schroedinger equation.

(d) Understanding and applying the mathematical technique of separation of variables in differential equations.

(e) Understanding and applying the phenomenon of quantum tunneling that arise in the presence of finite potential barriers.

2.1 The particle in a box and energy quantization

In the previous lecture (HC2), when we solved the Schroedinger equation for a free particle in absence of boundary conditions, we found that its energy E could take any value. Therefore, for a free particle allowed to move anywhere in space, its energy levels *are not quantized*: there is no restriction on the values they can take. We now will see how once the particle is confined into a limited region of space, energy levels become automatically quantized. Perhaps the simplest system for which energy quantization arises is for the so-called *particle in a box* system.

This system is defined by a single particle moving under the effects of a potential of the form

$$\begin{aligned} V(x) &= 0 \quad \text{for} \quad 0 \le x \le L\,, \\ V(x) &= +\infty \quad \text{for} \quad x < 0 \quad \text{and} \quad x > L\,. \end{aligned} \tag{2.1}$$

In other words, the particle undergoes free motion in the region $0 \le x \le L$, but cannot move outside this range because it is *confined* by the infinite potential barrier Eq. (2.1) (since the particle would need an infinite energy to overcome that potential barrier).

Inside the region limited by the confining potential, $0 \le x \le L$, the solution of Schroedinger's equation will be the same as for a free particle (since the potential vanishes there) and thus we have

$$\Psi_k = Ae^{ikx} + B^{-ikx} = (A+B)\cos(kx) + (A-B)i\sin(kx) \equiv D\cos(kx) + C\sin(kx)\,, \tag{2.2}$$

where we have expanded the exponentials using $e^{ix} = \cos(x) + i\sin(x)$ and then redefined the (arbitrary) integration coefficients for reason that will become apparent below. For the time being, k is a real constant that can take any value.

Since for $x > L$ and $x < 0$ we have that $V(x) = \infty$, the particle cannot travel to this region (since it would require infinite energy) and thus the wave-function will be zero there: $\Psi(x) = 0$ for $x > L$ and $x < 0$. From this it follows that, since as discussed in Sect. 1.2 the wave function is continuous, we have that the following *boundary conditions* must be satisfied:

$$\Psi_k(x=0) = 0\,, \quad \Psi_k(x=L) = 0\,. \tag{2.3}$$

Therefore, the continuity requirement of the wave function leads to boundary conditions that can be used to fix the values of the integration constants that appear in Eq. (2.2):

$$\Psi_k(x=0) = D \to D = 0\ , \tag{2.4}$$

$$\Psi_k(x=L) = C\sin(kL) = 0 \to k = \frac{n\pi}{L}\ , \tag{2.5}$$

where n is an arbitrary *integer number.* The value of the overall coefficient C can be determined from requiring the normalization of the wave-function:

$$1 = \int_{-\infty}^{\infty} dx\, |\Psi_k(x)|^2 = C^2 \int_0^L dx\, \sin^2(n\pi x/L) \to C = \sqrt{\frac{2}{L}}\,. \tag{2.6}$$

Therefore, combining all the results we find that:

For a free particle confined to a box, the quantum wave-functions and energies are given by

$$\Psi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right)\ , \quad E_n = \frac{\hbar^2\pi^2 n^2}{2mL^2}\ . \tag{2.7}$$

Therefore, we now find that the *energies of the particle are quantized*, and labeled by an integer number n. This can be compared to the case of the unbounded free-particle, where there were no such restrictions on the values that E could take.

Remarkably, we find that the lowest energy that a particle can have in this system is not zero, but rather

$$E_{\rm min} = E_1 = \hbar^2\pi^2/2mL^2 \neq 0\,, \tag{2.8}$$

since $n = 0$ is not a valid quantum state since the wave function $\Psi_0(x) = 0$ everywhere. This is known as the *zero-point energy*, and is a consequence of the fact that a quantum particle in a confining potential cannot be completely at rest. Indeed, if the particle was at rest (vanishing kinetic energy), we would then know its momentum with arbitrary precision (if $p_x = 0$ then $\Delta p_x = 0$), contradicting Heisenberg's uncertainty principle: we have some information about the position of the particle, namely that it will be found between $x = 0$ and $x = L$.

The qualitative results found from the particle in a box system are fully general and apply to other systems: in quantum mechanics, *energy quantization* arises from the wave nature of the wave-function in the presence of boundary conditions, including periodic ones.

Let us now evaluate the expectation value of the momentum p_x for this system. As we have shown, the wave-function that solves the Schroedinger equation accounting for the boundary conditions of the system is given by

$$\Psi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{\pi n x}{L}\right)\,, \tag{2.9}$$

and if we compute the expectation value of p_x using Eq. (1.78) we find that

$$\langle p_x\rangle = \frac{2}{L}\int_0^L dx \sin\left(\frac{\pi n x}{L}\right)\left(\frac{\hbar}{i}\frac{d}{dx}\right)\sin\left(\frac{\pi n x}{L}\right) = \frac{2\hbar n\pi}{iL^2}\int_0^L \sin\left(\frac{\pi n x}{L}\right)\cos\left(\frac{\pi n x}{L}\right) = 0\,, \tag{2.10}$$

since the integral vanishes for any value of n, as can be checked using trigonometric identities:

$$\int dx \sin(ax)\cos(ax) = \frac{1}{2a}\sin^2(ax)\,. \tag{2.11}$$

So therefore we find that the expectation value of the momentum for the particle in a box is $\langle p_x\rangle = 0$. This result can be understood if we expand the solution Eq. (2.9) in terms of exponentials:

$$\Psi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{\pi n x}{L}\right) = \sqrt{\frac{1}{2L}}\frac{1}{i}\left(e^{i\pi n x/L} - e^{-i\pi n x/L}\right)\,, \tag{2.12}$$

which correspond to the superposition (with equal amplitude) of a plane wave moving in the positive direction with momentum $p_x = \hbar\pi n/L$ and another moving in the opposite direction with $p_x = -\hbar\pi n/L$, hence when averaging the two components of the wave-function, they cancel among them leading to $\langle p_x\rangle = 0$.

In the $n \to \infty$ limit, the energies of the particle become *effectively continuous*, as expected in classical physics. This is the realization of the so-called *correspondence principle* of quantum theory: for large values of the quantum numbers, the behaviour of the quantum theory becomes effectively classical.

2.2 Particle in a two-dimensional box

The next system that we will consider is similar than the previous one, but now the box has *two dimensions* in space, which we will denote by x and y. Therefore, the confining 2D potential of this system will take the following form:

$$\begin{aligned} V(x,y) &= 0 \quad \text{for} \quad 0 \le x \le L_x \quad \text{and} \quad 0 \le y \le L_y \,, \\ V(x,y) &= +\infty \quad \text{for} \quad y < 0\,, \quad y > L_y\,, \quad x < 0\,, \quad x > L_x\,, \end{aligned} \tag{2.13}$$

where we allowed the dimensions of the box to be different, $L_X \neq L_Y$. Inside the box, the Schroedinger equation is the same as that of the free particle of the previous example but now in two dimensions, namely:

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\Psi(x,y) = E\Psi(x,y) \;, \tag{2.14}$$

where note that now the wave-function depends on two variables, x and y, and thus the derivatives that appear are *partial derivatives* rather than total derivatives.

To solve this differential equation, we need to adopt the method of *separation of variables.* In this method, one assumes that the full wave-function $\Psi(x,y)$ can be expressed as a product of two functions, each depending separately on x and y only, that is, one tries a solution of the form:

$$\Psi(x,y) = \Psi_X(x)\Psi_Y(y)\,. \tag{2.15}$$

If we plug this *ansatz* on the two-dimensional Schroedinger equation, Eq. (2.14), we obtain the following expression

$$-\frac{\hbar^2}{2m}\Psi_Y(y)\left(\frac{\partial^2}{\partial x^2}\right)\Psi_X(x) - \frac{\hbar^2}{2m}\Psi_X(x)\left(\frac{\partial^2}{\partial y^2}\right)\Psi_Y(y) = E\Psi_X(x)\Psi_Y(y)\,, \tag{2.16}$$

and now, if we divide each side of the equation by $\Psi(x,y)$, we find that

$$-\frac{\hbar^2}{2m}\frac{1}{\Psi_X(x)}\left(\frac{\partial^2}{\partial x^2}\right)\Psi_X(x) - \frac{\hbar^2}{2m}\frac{1}{\Psi_Y(y)}\left(\frac{\partial^2}{\partial y^2}\right)\Psi_Y(y) = E\,. \tag{2.17}$$

In Eq. (2.17), the right-hand side (RHS) is independent of both x and y, and in the left-hand side (LHS) we have the sum of two pieces, the first one depending *only on* x and the second one depending *only on* y. Therefore, the only way the equation can be true is if each piece is separately equal to a *constant.* If these two constants are denoted respectively by E_X and E_Y respectively, we find

$$-\frac{\hbar^2}{2m}\frac{1}{\Psi_X(x)}\left(\frac{\partial^2}{\partial x^2}\right)\Psi_X(x) = E_X \;, \tag{2.18}$$

$$-\frac{\hbar^2}{2m}\frac{1}{\Psi_Y(y)}\left(\frac{\partial^2}{\partial y^2}\right)\Psi_Y(y) = E_Y \;, \tag{2.19}$$

which are of course nothing but *two separate Schroedinger equations*, one for the x component of the wave function, $\Psi_X(x)$, and another for the y component of the wave function, $\Psi_Y(y)$. The total energy of the system is then $E = E_X + E_Y$, which justifies our choice of notation for the integration constants.

From the discussion above, we see that the solution of the Schroedinger equation for a particle in a 2D box will be given by the product of solutions to the same equation in a 1D box. That is, we will have that the x- and y-components of the wave functions are

$$\Psi_x(x) = \sqrt{\frac{2}{L_x}} \sin\left(\frac{n_x \pi x}{L_x}\right) , \tag{2.20}$$

$$\Psi_y(y) = \sqrt{\frac{2}{L_y}} \sin\left(\frac{n_y \pi y}{L_y}\right) , \tag{2.21}$$

and thus the *quantum state of the system* is now being defined by *two independent integer numbers* (n_x, n_y) (the two quantum numbers of the system), and therefore the total wave-function is

$$\Psi(x,y) = \Psi_x(x)\Psi_x(y) = \sqrt{\frac{4}{L_x L_y}} \sin\left(\frac{n_x \pi x}{L_x}\right) \sin\left(\frac{n_y \pi y}{L_y}\right) . \tag{2.22}$$

The total energy of a given quantum state of the system will be specified by the quantum numbers of this state, n_x and n_2, and thus if we take into account the result for the one-dimensional case, Eq. (2.7), we have that

$$E_{n_x,n_y} = \frac{\hbar^2}{2m}\left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2}\right) . \tag{2.23}$$

Note that in the limit in which $L_x = L_y = L$ there will be a *degeneracy* of quantum states, meaning that different quantum states, defined by different pairs of quantum numbers (n_x, n_y) will have associated the *same total energy.* In this limit indeed the total energy becomes

$$E_{n_x,n_y} = \frac{h^2}{8m}\left(\frac{n_x^2 + n_y^2}{L^2}\right) , \tag{2.24}$$

so for example the quantum state $(n_x = 1, n_y = 2)$ will correspond to a state with the same energy as that with $(n_x = 2, n_y = 1)$. Degeneracy is a generic property of quantum systems: in general, many different states can have associated the same total energy.

2.3 Quantum tunneling

In classical physics, when we have a particle with total energy E moving inside a conservative potential $V(x)$, the particle will *confined* to the region defined by $E \geq V(x)$. Indeed, from energy conservation we have that the sum

$$E = T + V(x) = \frac{1}{2}mv^2 + V(x) , \tag{2.25}$$

where T is the particle's kinetic energy, is a constant of motion and must hold for all values of x. Therefore, we have that

$$v^2 = \frac{2}{m}\left(E - V(x)\right) \tag{2.26}$$

can only be satisfied if $E \geq V(x)$, else the velocity would be an unphysical complex number. Therefore, the particle cannot move in the region of x for which $V(x) > E$: we know that this region is *classically forbidden.*

However, in quantum physics this is *not* necessarily the case: a particle can cross a potential barrier even when its kinetic energy is smaller than the potential energy of the barrier.

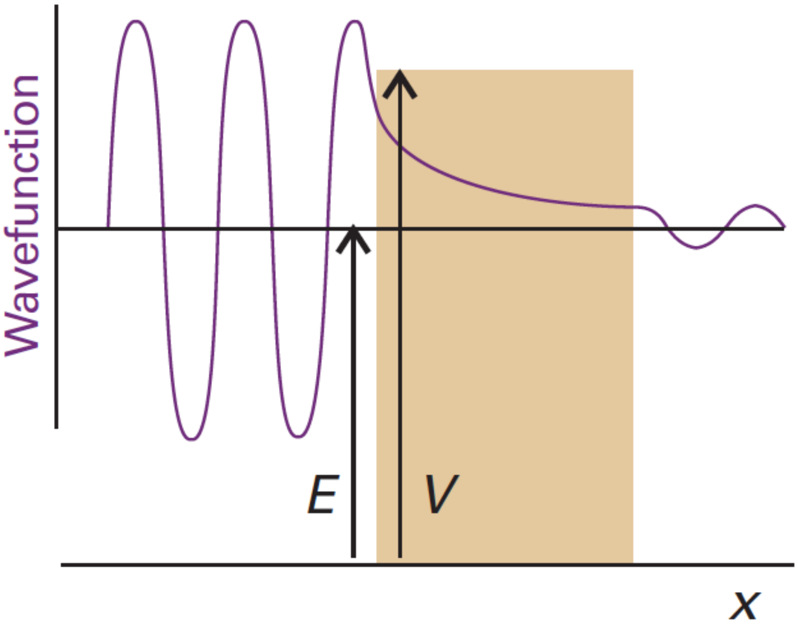


Figure 2.1: Schematic representation of the *quantum tunneling effect*: the wave-function of a particle with kinetic energy E_k is non-zero inside a barrier with potential energy $V > E_k$, and therefore has a finite probability (non-zero wave-function) to be found at the other side of the barrier.

This remarkable phenomenon is known as the *quantum tunneling* effect, and is schematically represented in Fig. 2.1, where we show how the wave function of a particle is non-zero even in the classically forbidden region with $V > E_k$, and thus leads to a finite probability of finding the particle at the right of the potential barrier.

The quantum tunneling effect is a direct consequence of the wave-like nature of the wave function. We can now quantify and compute explicitly the value of the wave-function inside and on the other side of the potential barrier. As indicated in Fig. 2.1, at the left of the barrier we have $V = 0$, and thus the solution of the Schroedinger equation there is the usual free-particle solution, namely

$$\Phi_k(x) = Ae^{ikx} + Be^{-ikx} \,, \tag{2.27}$$

where the kinetic energy is $E_k = \hbar^2 k^2/2m$ and thus the linear momentum is $p_x = k\hbar = \sqrt{2mE_k}$. In the region *inside the barrier*, the Schroedinger equation is instead given by:

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi}{dx^2} + V\Psi = E_k\Psi \,, \quad \text{with} \quad V \geq E_k \,. \tag{2.28}$$

Note that here I have identified the total energy of the particle E with the kinetic energy at the left side of the barrier, $E = E_k$, in order to *energy conservation* to be satisfied. Moreover, since the potential V is constant, this equation can be rewritten as

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi}{dx^2} = -(V - E_k)\Psi \,, \tag{2.29}$$

which has the form of a free-particle equation, but this time with a *negative energy*, $\widetilde{E} = -(V - E_k)$. Using the same method as solution as for a free particle, we get that the wave-function inside the barrier is now

$$\Psi = C\, e^{\kappa x} + D\, e^{-\kappa x}\,, \quad \kappa \equiv \sqrt{2m\,(V - E_k)}/\hbar \geq 0\,, \tag{2.30}$$

which as an *exponential solution*, rather than the oscillatory solution of the free particle equation. Therefore, the wave-function is non-zero inside the barrier, despite $V > E_k$ and thus of being forbidden in classical physics: it was a finite probability of being found in the classically forbidden region.

Finally, in the region right to the barrier, we have again a free-particle solution for a particle moving in the positive x direction, that is

$$\Psi = A' e^{ikx}\,, \quad k = \sqrt{2mE_k}/\hbar\,, \tag{2.31}$$

with equal momentum and energy as in the left side of the barrier due to the principle of energy conservation. In order to determine the values of the five integration constants introduced above, A, B, C, D, A', we need exploit two properties of the wave-function: it is *continuous everywhere*, and its *derivative is also continuous* for any value of x in the presence of finite potentials.[1] If we label as $x = 0$ and $x = L$ the start and end points of the potential barrier, continuity of the wave-function there implies that

$$\begin{aligned} A + B &= C + D\,, \\ Ce^{\kappa L} + De^{-\kappa' L} &= A' e^{ikL}\,, \end{aligned} \tag{2.32}$$

while the continuity of the first derivative of the wave-function at the same endpoints implies that

$$\begin{aligned} ikA - ikB &= \kappa C - \kappa D\,, \\ \kappa C e^{\kappa L} - \kappa D e^{-\kappa L} &= ikA' e^{ikL}\,. \end{aligned} \tag{2.33}$$

Note that we have five unknowns and four equations, and therefore we can express four of these unknowns in terms of a single integration constant, say A, that can then be fixed from the overall normalisation condition.

Transmission and reflection coefficients. As we saw in the free-particle case, for the solution in the left side of the barrier, Eq. (2.27) we could make the interpretation that the Ae^{ikx} component of the wave function can be associated with the *incident wave* (since its momentum was $p_x = +\hbar k$), while the Be^{-ikx} instead would be the *reflected wave* (with momentum $p_x = -\hbar k$ pointing in the negative x direction). Therefore, we can define a ratio $T = A'/A$ which physically can be interpreted as the ratio of the amplitude of the *transmitted* wave over the *incident* wave, see Fig. 2.2. This *transmission coefficient* T can be computed using the values of the integration coefficients B, C, D, A' obtained as explained above, resulting in the following expression:

$$T\left(\kappa L, \epsilon\right) = \left(1 + \frac{\left(e^{\kappa L} - e^{-\kappa L}\right)^2}{16\epsilon(1-\epsilon)}\right)^{-1}\,, \quad \epsilon \equiv E/V\,, \tag{2.34}$$

and where κ has been defined in Eq. (2.30). The transmission amplitude T has a number of important limiting cases. When taking the various limits, note that κ depends implicitly on ϵ as well, since

$$\kappa\hbar = \sqrt{2m(V - E_k)} = \sqrt{2mV}\sqrt{1-\epsilon}\,. \tag{2.35}$$

With this caveat, it is possible to derive the following important properties of the transmission amplitude T:

---

[1] The first derivative of the wave function can be discontinuous in the case of infinite potential wells.

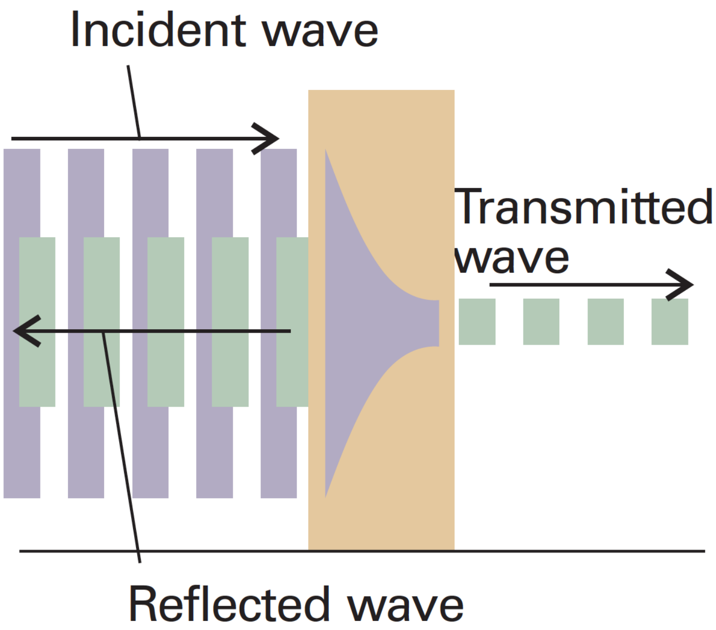


Figure 2.2: The physical interpretation of the quantum tunneling effect: an incident wave Ae^{ikx} left to the barrier is partially transmitted to the other side of the barrier with wave function $A' = e^{ikx}$ and partially reflected, with momentum in the opposite direction as compared to the incident wave, Be^{-ikx}. The ratio $T = A'/A$ of the transmitted over the incident amplitude is known as the *transmission coefficient*.

- In the limit $L \to 0$ for fixed κ, then the transmission coefficient $T \to 1$.

  This limit corresponds either to *very short barriers* L. In this two cases, it makes sense physically that the probability of tunneling becomes very high (and the amplitude of the *reflected way* conversely very small).

- In the limit $\kappa \to 0$ ($\epsilon \to 1$) for fixed L, then the transmission coefficient goes to

  $$T \to \left(1 + \frac{mVL^2}{\hbar}\right)^{-1} , \tag{2.36}$$

  so it does *not* tend to one even if $E_k \lesssim V$ (only in the case of very short barriers $L \to 0$ then $T \to 1$).

- For $E_k \ll V$, or what is the same $\epsilon \to 0$, we find that $T \to 0$.

  This can be physically understood from the fact that for a steep enough barrier, eventually the probability of transmission will become vanishingly small, in agreement with the *classical expectation*.

- As $E_k \to V$ ($\epsilon \to 1$), the value of the transmission amplitude increases monotonically, until the limiting value Eq. (2.36) is achieved.

- in the limit $\kappa L \gg 1$ the transmission amplitude Eq. (2.34) becomes

  $$T \simeq 16\epsilon(1-\epsilon)e^{-2\kappa L} . \tag{2.37}$$

  This limit corresponds to either very steep ($\kappa \to \infty$) or very long ($L \to \infty$) barriers, or the two at the same time. In this case we intuitively expect that the transmission probability will be small, and what Eq. (2.37) indeed shows is that T is *exponentially small* in this limit. We also note that $T \sim e^{-2L\sqrt{2mV}/\hbar}$, and thus that *lighter particles* will have a higher probability of tunneling that heavier particles.

The fact that the transmission coefficient Eq. (2.34) is different from zero is a striking deviation of quantum theory with respect to classical physics. The quantum tunneling effect indicates that for instance a naive particle picture of electrons or other quantum particles is far from adequate to describe the phenomena of the microcosm.

To conclude this discussion of the tunneling effect, recall that in HC2 we mentioned the *correspondence principle* of quantum theory, namely that in the appropriate limits the quantum behaviour should become effectively classical. In the case of the quantum tunneling effect, since we have that

$$\kappa = \frac{1}{\hbar}\sqrt{2m(V - E_k)}\,, \tag{2.38}$$

we find that, for fixed values of V and E_k, if $\kappa \to \infty$ then from Eq. (2.34) we see that $T \to 0$. This limit can be realized if

$$\hbar \ll \sqrt{2m(V - E_k)}\,. \tag{2.39}$$

Therefore, in this limit (where Planck's constant can be set to zero) we find that the tunneling probability goes to zero and this recover the classically expected behaviour.

2.4 The quantum harmonic oscillator

In classical mechanics, an *harmonic oscillator* is defined as particle moving under the effects of a *quadratic potential*, that is, a potential $V(x)$ of the form

$$V(x) = \frac{1}{2}\kappa x^2\,, \tag{2.40}$$

where κ is known as the *spring constant* or *Hooke's constant*. Though physically this potential is usually associated to a system based on a frictionless body attached to a flexible spring, the form Eq. (2.40) is fully general and applies to many other potentials. In particular, any potential can be approximated by Eq. (2.40) in the region near local minima (as can be seen by doing a Taylor expansion), so harmonic oscillators are ubiquitous in quantum mechanics.

The harmonic potential Eq. (2.40) vanishes at $x = 0$, the *equilibrium* position of the particle. Note that this potential is *confining*, since it increases quadratically as $|x|$ increases, and thus a particle in this potential would never be able to completely escape from it, no matter how large its energy is. From this potential, at the classical level we can compute the force that the particle will experience, namely

$$F = -\frac{dV(x)}{dx} = -\kappa x\,, \tag{2.41}$$

which is known as *Hooke's law*: in an harmonic oscillator (or in general, for a quadratic potential) the force is attractive and proportional to the deviation with respect to the equilibrium position.

Let us now study the behaviour of a quantum particle under the effects of the potential Eq. (2.40), that is, a *quantum harmonic oscillator*. The Schroedinger equation associated to this system will now be:

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi}{dx^2} + \frac{1}{2}\kappa x^2\Psi = E\Psi\,. \tag{2.42}$$

Solving this equation is beyond the scope of this course, though for completeness let me show here the explicit form of the solutions for the wave function:

$$\Psi_n(x) = \frac{1}{\sqrt{2^n n!}} \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/2\hbar} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) , \quad (2.43)$$

where $H_n(x)$ are as special family of orthogonal polynomials known as *Hermite polynomials*, and we have defined the *frequency* of the oscillator as $\omega \equiv \sqrt{\kappa/m}$, in analogy with the classical treatment. The solutions of Eq. (2.42) are labeled by the quantum number n, which takes only integer values $n = 0, 1, 2, 3, \ldots$.

As a consequence of the potential barrier that confines the particle, the energies of the quantum harmonic oscillator are *quantized*, and it is possible to show that they are given by

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) . \quad (2.44)$$

There are two important consequences of this result. First of all, we note that the difference in energy between two adjacent quantum levels n and $n + 1$ is constant, that is

$$E_{n+1} - E_n = \hbar\omega , \quad (2.45)$$

independently of the value of n. Secondly we find that, in the same way as for the particle confined in a box,

the energy of the *ground state* (that is, the *vacuum*) of the system, the eigenfunction with the smallest associated energy, is *different from zero*. Indeed we find that for $n = 0$ we get

$$E_0 = \frac{1}{2}\hbar\omega , \quad (2.46)$$

which is known as the *zero-point energy* of a quantum harmonic oscillator.

This is a direct consequence of Heisenberg's uncertainty principle: if the energy and momentum were zero, the particle would be at rest, and we would know p_x with arbitrary precision and thus $\Delta p_x = 0$. But then the indetermination on its position x would be maximal, as in the free particle case, and this is not possible due to the confining harmonic potential.

For small values of n, the Hermite polynomials that appear in the quantum wave-function Eq. (2.43) take relatively simple forms, and we list the first few of those:

$$\begin{aligned} H_0(x) &= 1 , \\ H_1(x) &= 2y , \\ H_2(x) &= 4y^2 - 2 , \\ H_3(x) &= 8y^3 - 12y , \end{aligned} \quad (2.47)$$

and so on. Therefore, for the ground state of the system, $n = 0$, the wave-function is

$$\Psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/2\hbar} , \quad (2.48)$$

which is known as a *Gaussian function*. It is easy to check explicitly that these wave-function is correctly

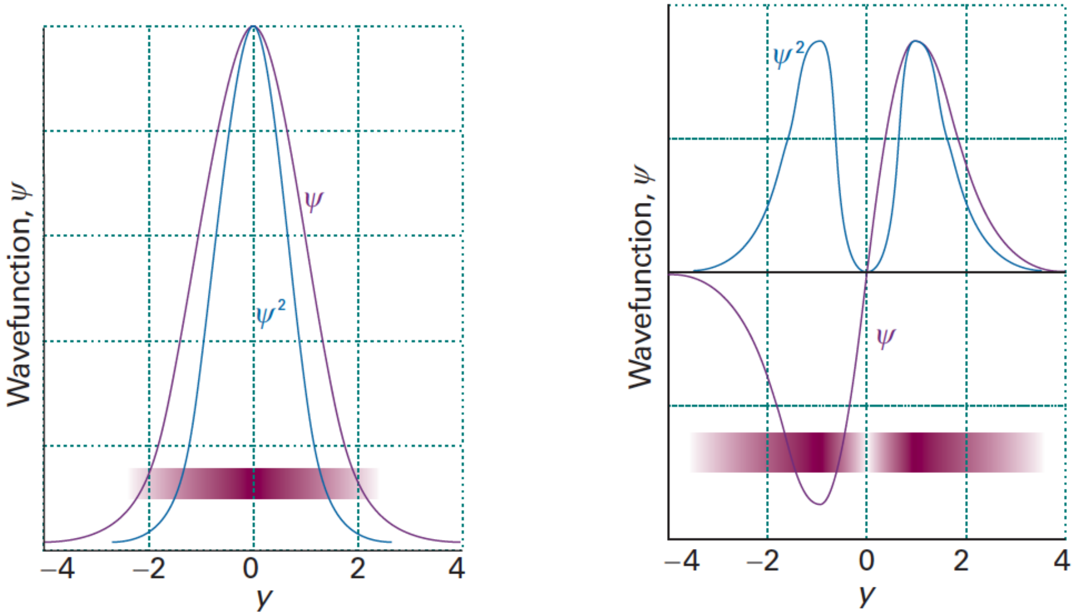


Figure 2.3: The wave-function $\Psi_n(x)$ and its square $|\Psi_n(x)|^2$ (which represents the probability density for the position x) in the quantum harmonic oscillator for the first two eigenstates, $n = 0$ (left plot) and $n = 1$ (right plot). We observe that while for $n = 1$ (the ground state) the maximum probability $|\Psi|^2$ is found for $x = 0$, for the excited state $n = 1$ it is rather more likely to find the particle far from the equilibrium position.

normalized (as is the case for other values of n), by verifying that

$$\int_{-\infty}^{\infty} dx\, |\Psi_0(x)|^2 = 1\,. \tag{2.49}$$

To show this, first one should make the change of variable $y \equiv \sqrt{m\omega/\hbar}x$ and then use the result for the Gaussian integral that

$$\int_{-\infty}^{\infty} dx\, e^{-x^2} = \sqrt{\pi}\,. \tag{2.50}$$

It is interesting to compare the probability densities for the position x for the first two energy levels, namely $|\Psi_0(x)|^2$ with $|\Psi_1(x)|^2$, where

$$\Psi_1(x) = \sqrt{2}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\sqrt{\frac{m\omega}{\hbar}}\,x\, e^{-m\omega x^2/2\hbar}\,, \tag{2.51}$$

This comparison is shown in Fig. 2.3, where we show the wave-function $\Psi_n(x)$ and its square $|\Psi_n(x)|^2$ (which remember represents the probability density for the position x) in the quantum harmonic oscillator for the first two eigenstates, $n = 0$ (left plot) and $n = 1$ (right plot).

We observe that while for $n = 0$ (the ground state) the maximum of the probability density $|\Psi|^2$ is found for $x = 0$ (the classical equilibrium position), for $n = 1$ (the first excited state) it is rather more likely to find the particle far from the equilibrium position. This property holds for other excited states: the higher the value of n, the more likely is to find the particle far from $x = 0$. This behaviour is consistent with the classical theory, where the harmonic oscillator spends more time in the turning points than in the equilibrium point because its velocity is the smallest in the former positions.

The exponential suppression in the wave function of the quantum harmonic oscillator, Eq. (2.43) implies that it goes to zero for $x \to \pm\infty$, no matter the value of n (that is, of how energetic is the particle). This is because the particle has a oscillatory behaviour in the region $E > V(x)$, but then decays exponentially in the classically forbidden region $E < V(x)$, which is eventually reached no matter how large is E (because of the form of the potential).

Properties of the quantum harmonic oscillator. Since we have the wave-functions for all values of n, Eq. (2.43), we know that we have a complete knowledge of this quantum system, and thus we can compute the expectation values of arbitrary physical observables. First of all, we can show that the expectation value of the position x of the harmonic oscillator is, for any value of the quantum number n,

$$\langle x \rangle = 0 \, , \tag{2.52}$$

in other words, the particle has a symmetric distribution of positions around the classical equilibrium position $x = 0$. This can be easily shown by noting that

$$\langle x \rangle = \int_{-\infty}^{\infty} dx \, x \, |\Psi(x)|^2 \sim \int_{-\infty}^{\infty} dx \, x \, e^{-m\omega x^2/\hbar} \left[H_n \left(\sqrt{\frac{m\omega}{\hbar}} x \right) \right]^2 = 0 \, , \tag{2.53}$$

since the *integrand is an odd function*, $f(x) = f(-x)$, and the integration range is symmetrical. To see this, note that $H_n(-x) = \pm H_n(x)$ for any values of n.

Having established that in the quantum harmonic oscillator the particle can be found with equal probability at the right and at the left of the equilibrium position, it is perhaps more interesting to now compute its *mean square displacement*, given by

$$\left\langle x^2 \right\rangle = \left(n + \frac{1}{2} \right) \frac{\hbar}{\sqrt{m k_f}} \, . \tag{2.54}$$

Therefore, the *standard deviation* of the position x is

$$\Delta x \equiv \sqrt{\langle x^2 \rangle - \langle x \rangle^2} = \sqrt{\left(n + \frac{1}{2} \right) \frac{\hbar}{(m k_f)^{1/2}}} \tag{2.55}$$

which for large values of n grows like $\Delta x \sim \sqrt{n}$: the likelihood of finding the particle at a greater distance from $x = 0$ increases as the square root of the quantum number n, despite the fact that on average the expectation value will still be $\langle x \rangle = 0$. These results are consistent with the previous discussion above, based on the behaviour of the wave functions Ψ_n. From Heisenberg's uncertainty principle, Eq. (1.28) we can determine the standard deviation associated to measurements of the linear momentum of the quantum harmonic oscillator in this limit,

$$\Delta p_x \sim \frac{\hbar}{\Delta x} \sim \frac{\hbar}{n^{1/2}} \, , \tag{2.56}$$

so the higher the value of n, the better the momentum of the harmonic oscillator can be predicted (in the *correspondence limit*, we recover the classical expectation that p_x can be determined with vanishingly small uncertainty.)

There are other properties of this quantum system that are useful to compute. The expectation value of the potential energy is given by

$$\langle V\rangle = \frac{1}{2}k_f\left\langle x^2\right\rangle = \frac{1}{2}\left(n+\frac{1}{2}\right)\hbar\left(\frac{k_f}{m}\right)^{1/2} = \frac{1}{2}\left(n+\frac{1}{2}\right)\hbar\omega \ , \tag{2.57}$$

where we have used the fact that the expectation value of an operator is *linear*, and the result for $\left\langle x^2\right\rangle$ just computed above. Therefore, given that the complete energy of the quantum harmonic oscillator was Eq. (2.44), we find that $\langle V\rangle = E_n/2$, and thus for the kinetic energy $\langle E_k\rangle = E_n/2$. So in the quantum harmonic oscillator, the energy is *equally shared* between kinetic and potential, for any value of n. This is actually a consequence of a deep principle called the *equipartition theorem.* In other words, we have that for this quantum system

$$\langle E\rangle = \langle E_k\rangle + \langle V\rangle = \frac{1}{2}\langle E\rangle + \frac{1}{2}\langle E\rangle \ . \tag{2.58}$$

so on average half the total energy corresponds to the kinetic energy and the other half to the potential energy.

2.5 Particle in a box with finite potential walls

The next quantum system that one could study is a modification of the particle-in-a-box system discussed above, with the difference that the potential barriers V at the two sides of the box now are *finite* (as opposed to infinite as before). In other words, the potential function in this system now reads

$$\begin{aligned} V(x) &= 0 \quad \text{for} \quad 0 \le x \le L \,, \\ V(x) &= +|V| \quad \text{for} \quad x < 0 \quad \text{and} \quad x > L \,. \end{aligned} \tag{2.59}$$

Although we will not work out this case explicitly in the lectures, the solution to the Schroedinger equation here follows quite directly the tunneling derivation that was presented above. Indeed, one has to solve the Schroedinger equation both inside the potential well (where we have the free-particle, oscillatory, solution) and in the classically forbidden regime *inside* the left and the right barriers, where we find a exponentially decaying solution. The integration constants of this problem can be uniquely determine by the conditions that the wave-function and its derivate should be continuous at both sides of the potential barrier, $x = 0$ and $x = L$, as well as by the overall normalization of the wave function.

In Fig. 2.4 we show in the left plot the schematic representation of the particle-in-a-box quantum system, characterized by finite potential barriers V at the two sides of the box. In the right plot, we show the wave function for the first two quantum states of the system, $n = 1$ and $n = 2$, which exhibit oscillatory behaviour inside the box but *decay exponentially fast* in the classically forbidden region with $E < V$. Note that, as in the quantum tunneling case discussed above, the wave-function is non-zero even inside the potential barrier, where $V > E_k$ with E_k being the kinetic energy of the particle inside the box, but quickly becomes very small if $V \gg E_k$.

Summary

To summarize, some of the important concepts that we have learned in this lecture are the following:

- The boundary conditions on the wave-function induced by a confining potential lead to the *quantization of the allowed energy levels.*

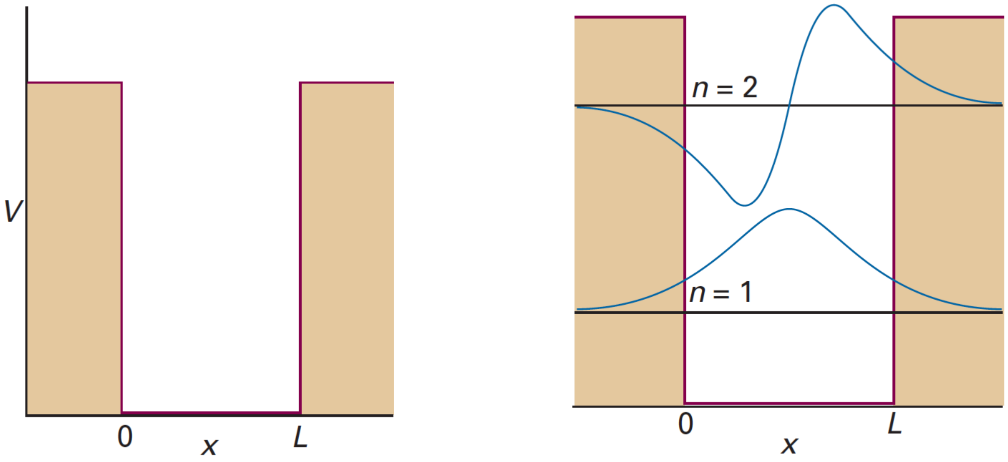


Figure 2.4: Schematic representation of the particle-in-a-box quantum system, characterized by finite potential barriers V at the two sides of the box (left plot). In the right plot we show the wave function for the first two quantum states of the system, $n = 1$ and $n = 2$, which exhibit oscillatory behaviour inside the box but decay exponentially fast in the classically forbidden region with $E < V$.

- This energy quantisation is a generic property of quantum systems in the presence of boundary conditions, as shown also in other systems such as the particle in a box with finite barriers and the quantum harmonic oscillator.
- In several quantum systems, the energy of the ground state is different to zero, unlike in classical physics. We denote this effect as the *zero-point energy*, and it is a direct consequence of Heisenberg's uncertainty principle.
- Quantum particles have a non-zero probability of being measured *within classically forbidden regions*, and to tunnel potential barriers even when their kinetic energy is smaller than the energy of the barrier.
- In some circumstances, quantum states can be *degenerate*, meaning that different states, characterized by different quantum numbers, can have associated the *same total energy*.
- In the *correspondence limit*, usually associated to high values of the quantum numbers of the system, quantum theory predictions should *reproduce their classical counterparts*.

Further reading

Further reading material about the topics covered in this lecture can be found in the following sections of the course textbook (*Physical Chemistry*, Atkins and De Paula, 10th edition):

- **8A.2**: Confined motion in one dimension.
- **8A.3** Confined motion in two or more dimensions.
- **8A.4**: Tunneling.
- **8B.1**: The harmonic oscillator.
- **8B.2**: The properties of oscillators

van Quantum tot Molecuul

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

3 HC4: Rotational motion in quantum mechanics

In this lecture, we discuss the quantum mechanics of systems characterised by rotational symmetry. In particular, we solve the Schroedinger equation for particles confined to a ring and to the surface of a sphere. We study the role that angular momentum plays for this systems. These systems are important in order to understand the quantum mechanics of the hydrogen atom as well as more complex atoms and molecules.

The **learning goals** of this lecture are the following:

(a) Becoming familiar with the concept of angular momentum in classical mechanics.

(b) Demonstrating the quantization of both the total angular momentum and of its individual components in systems characterised by rotational symmetry.

(c) Solving Schroedinger's equation for a free particle in a ring.

(d) Solving Schroedinger's equation for a free particle in a sphere.

Introduction. Quantum systems characterised by rotational symmetry are of great importance for the study of the physics of atoms and molecules. The obvious example is the hydrogen atom, a system composed by a proton and an electron bound together by the electrical force: since the Coulomb force depends only on the modulus of the separation $\vec{r}$ between the electron and the proton

$$F_C = -k\frac{Q_eQ_p}{|\vec{r}|^2} \tag{3.1}$$

then the system is characterised by an invariance under rotations that determines many of its most important properties, such as which specific quantum states are allowed.

With this motivation, both in this chapter and in the following ones we will study the implications of quantum mechanics for systems exhibiting rotational symmetry, both in two and in three spatial dimensions. For these systems the use of Cartesian coordinates is not practical, so first of all we will have to introduce more suitable coordinate systems that fully exploit the benefits of rotational symmetry.

3.1 Polar and spherical coordinates

The solution of Schroedinger's equation in systems that exhibit rotational symmetry is greatly facilitated by using polar or spherical coordinates rather than Cartesian coordinates. The use of spherical coordinates makes also more transparent the physical interpretation of the corresponding results. With this motivation, we begin this lecture with a review of polar and spherical coordinates.

Polar coordinates. Let us start with polar coordinates, that are relevant for problems where the particles in the system are confined to the two-dimensional (x, y) plane. As illustrated in Fig. 3.2 (left), a point in the two-dimensional (x, y) plane can be represented in terms of its distance from the coordinate origin, r, and the angle with the x-axis, φ. The explicit expression for the coordinate transformation between polar and Cartesian coordinates is given by

$$\begin{aligned} x &= r\cos\varphi\,, \\ y &= r\sin\varphi\,, \end{aligned} \tag{3.2}$$

where the radius r satisfies $x^2 + y^2 = r^2$. The inverse coordinate transformation is given by

$$\begin{aligned} r &= \sqrt{x^2+y^2}\,, \\ \varphi &= \arctan\left|\frac{y}{x}\right|\,, \end{aligned} \tag{3.3}$$

for φ lying in the first quadrant.

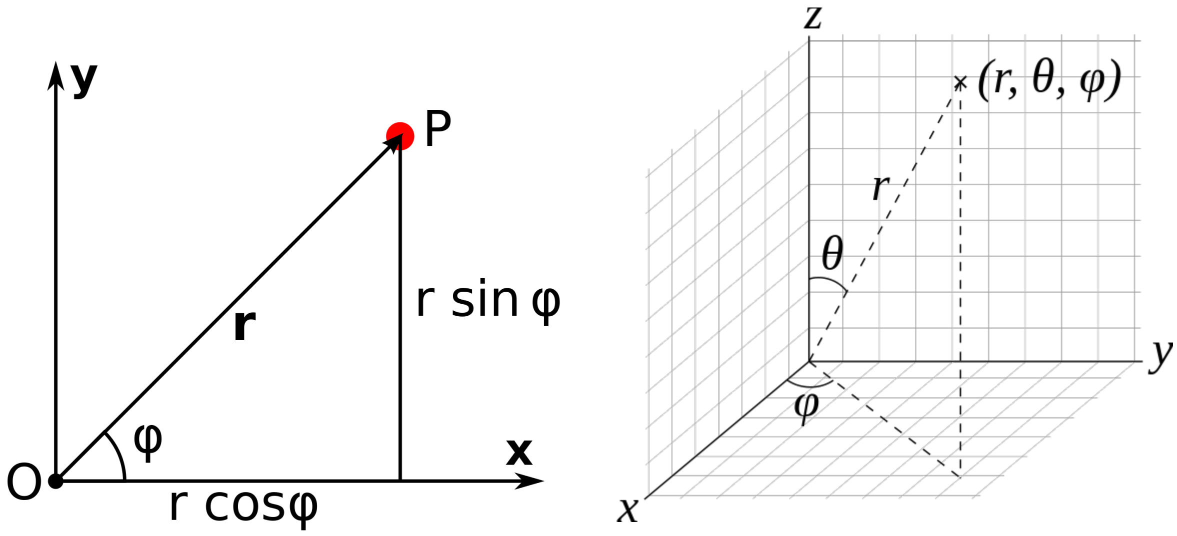
When computing transformations between Cartesian and spherical coordinates, it is useful to know the *Jacobian matrix* of the transformation, namely the set of partial derivatives between one set of coordinates with respect to the others. In this case, the Jacobian matrix is given by the following:

$$\begin{aligned} \frac{\partial r}{\partial x} &= \frac{\partial}{\partial x}\sqrt{x^2+y^2} = \frac{x}{\sqrt{x^2+y^2}} = \frac{x}{r} = \cos\varphi\,, \\ \frac{\partial r}{\partial y} &= \frac{\partial}{\partial y}\sqrt{x^2+y^2} = \frac{y}{\sqrt{x^2+y^2}} = \frac{y}{r} = \sin\varphi\,, \\ \frac{\partial \varphi}{\partial y} &= \frac{\partial}{\partial x}\arctan\left|\frac{y}{x}\right| = \frac{1}{1+y^2/x^2}(-y/x^2) = -\frac{y}{r^2} = -\frac{\sin\varphi}{r}\,, \\ \frac{\partial \varphi}{\partial y} &= \frac{\partial}{\partial y}\arctan\left|\frac{y}{x}\right| = \frac{1}{1+y^2/x^2}(1/x) = \frac{x}{r^2} = \frac{\cos\varphi}{r}\,, \end{aligned} \tag{3.4}$$

which can be used when doing the conversion between the two sets of coordinates, as will be illustrated below.

Spherical coordinates. Analogously to the definition of polar coordinates, for a particle moving in the three-dimensional space (x, y, z), one can define *spherical coordinates* as represented schematically in Fig. 3.2. In addition to r and φ (which in this context is referred to as the *azimuthal angle*) we now have a second angle called the *polar angle* θ. The equations for the coordinate transformation between Cartesian and spherical coordinates in three dimensions are given by

$$\begin{aligned} x &= r\sin\theta\cos\varphi\,, \\ y &= r\sin\theta\sin\varphi\,, \\ z &= r\cos\theta\,, \end{aligned} \tag{3.5}$$



Polar coordinates ***Spherical coordinates***

Figure 3.1: Left: in two dimensions, a point in the (x, y) plane can be represented by the polar coordinates (r, φ). Right: in three-dimensions, a point in the (x, y, z) space can be represented by the spherical coordinates (r, θ, φ). See text for more details.

where analogously to the case of polar coordinates we have the relation $r^2 = x^2 + y^2 + z^2$. The spherical coordinates have the following ranges:

$$0 \le r \le \infty \,, \qquad 0 \le \theta \le \pi \,, \qquad 0 \le \varphi \le 2\pi \,. \tag{3.6}$$

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

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2} \,, \\ \theta &= \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) = \arccos\left(\frac{z}{r}\right) \,, \\ \varphi &= \arctan\frac{y}{x} \,. \end{aligned} \tag{3.7}$$

The usefulness of spherical coordinates for problems with rotational symmetry, as we will see, stems for the simplification of the corresponding Schroedinger equation when using variables that reflect the underlying symmetries of the problem. In particular, the three-dimensional Schroedinger equation for a system with rotational symmetry can be *separated* as a modified one-dimensional equation for the radial coordinate r and a universal differential equation for the angular variables θ and φ, as we will show later.

We can now use the equations of the polar and spherical coordinates to solve Schroedinger's equation for two important quantum systems: a particle that is confined to move in a *ring*, and a particle which is confined to move in the *surface of a sphere*.

3.2 Quantum mechanics of a particle in a ring

In the previous lecture (HC3), we have applied Schroedinger's equation to determine the wave function $\Psi(x)$ of a variety of quantum systems. In particular, we considered the quantum mechanics of a particle confined in a box. There we found that the quantization of both the particle energy and of the allowed wave functions was a direct consequence of the boundary conditions required at the walls of the box due to the infinite potential barrier present.

Now we move to consider another quantum system of great importance for the understanding of the behaviour of atoms and molecules, namely the quantum mechanics of a particle undergoing rotational motion. We will start with the two-dimensional case (particle in a ring) and then move to the three-dimensional case (particle confined in the surface of the sphere). As we will see, also for these systems both the energy and the allowed wave functions are quantized. Here the underlying reason for the quantization will be not the presence of an infinite potential barrier, but rather the *periodic boundary conditions* that the wave function should satisfy due to the rotational symmetry.

Classical description and angular momentum. Let us first review the classical system. We consider here a particle following circular motion with a fixed radius R. As you have seen in the Mechanics course, the *angular momentum* associated to the motion of this particle is given by

$$\vec{J} = \vec{r} \times \vec{p} \, , \tag{3.8}$$

where $\vec{p} = m\vec{v}$ is the linear momentum, m is the mass, and $\vec{r}$ is the position vector with respect the origin of coordinates. In this equation, $\times$ stands for the cross-product between two vectors $\vec{v}_1$ and $\vec{v}_2$, defined as

$$\vec{v}_1 \times \vec{v}_2 = \left(|\vec{v}_1||\vec{v}_2| \sin\theta \right) \vec{n} = \left(v_{1,y}v_{2,z} - v_{1,z}v_{2,y} \right) \hat{\imath} + \left(v_{1,z}v_{2,x} - v_{1,x}v_{2,z} \right) \hat{\jmath} + \left(v_{1,x}v_{2,y} - v_{1,y}v_{2,x} \right) \hat{k} \, , \tag{3.9}$$

where θ is the angle between the two vectors in the plane that contains them, $\vec{n}$ is a unit vector perpendicular to the same plane, and $\hat{\imath}, \hat{\jmath}, \hat{k}$ are the unit vectors in the x, y, and z directions respectively.

From this definition of the cross-product of two vectors, the angular momentum $\vec{J}$ of a particle undergoing a circular motion will be perpendicular both to the position vector $\vec{r}$ and the linear momentum $\vec{p}$, as illustrated in Fig. 3.2. For example, in the case of circular motion in the (x, y) plane, the angular momentum points in the z direction, $\vec{J} = (0, 0, J_z)$ where $J_z = \pm R|\vec{p}|$ and R is the radius of the circular motion. Here the overall sign of J_z depends on whether the circular motion is clockwise or counterclockwise.

In classical mechanics, the kinetic energy for a particle undergoing circular motion is given by

$$E_{\rm kin} = \frac{J}{2mR^2} = \frac{J^2}{2I} \, , \tag{3.10}$$

with $I = mR^2$ is the *moment of inertia* and $J \equiv |\vec{J}|$. This can be compared with the kinetic energy for a particle undergoing linear motion, where we have the usual expression of

$$E_{\rm kin} = \frac{1}{2}mv^2 = \frac{p^2}{2m} \, . \tag{3.11}$$

Note that of course the two expressions are equivalent for circular motion, where $J = Rp$ and therefore Eq. (3.10) reads $E_{\rm kin} = J^2/2I = R^2p^2/2mR^2 = p^2/2m$.

Note that in the classical case both the kinetic energy E and the total angular momentum $\vec{J}$ of the particle in a ring can take any values. As we will see now, this is not true in the quantum case where both physical quantities will be *quantised*.

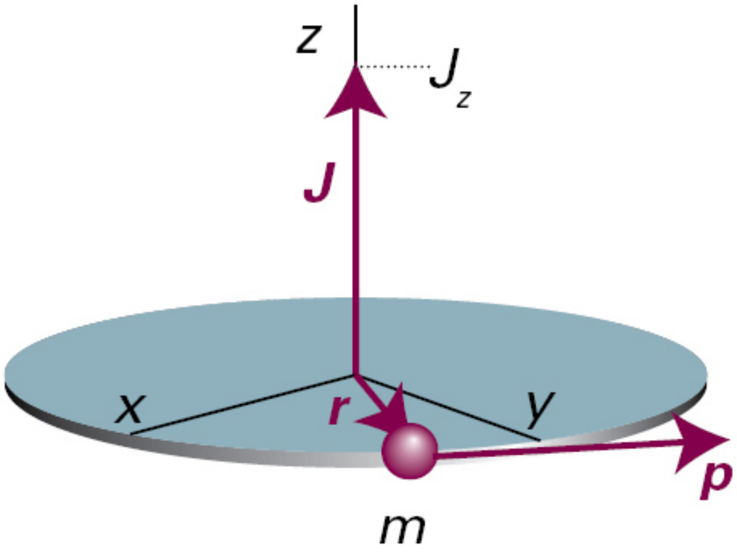


Figure 3.2: The angular momentum $\vec{J}$ of a particle undergoing a circular motion in the (x, y) plane is perpendicular both to the position vector $\vec{r}$ and the linear momentum $\vec{p}$, and therefore it points into the z direction, $\vec{J} = (0, 0, J_z)$.

Quantum description. Let us now study the particle in a ring from the point of view of quantum mechanics. For a particle moving in the (x, y) plane, the time-independent Schroedinger's equation reads

$$\widehat{H}\Psi(x, y) = E\Psi(x, y)\,, \tag{3.12}$$

where the Hamiltonian operator for a free particle in two dimensions is given by

$$\widehat{H} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\,. \tag{3.13}$$

In order to exploit the rotational symmetry of the system, we will perform a coordinate transformation from Cartesian to polar coordinates, Eq. (3.2). In polar coordinates, it can be shown that the Hamiltonian will be given by

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \varphi^2}\,, \tag{3.14}$$

and therefore Schroedinger's equation in polar coordinates reads

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \varphi^2}\right)\Psi(r, \varphi) = E\Psi(r, \varphi)\,. \tag{3.15}$$

To obtain the expression of the Hamiltonian in polar coordinates, one needs to apply the chain rule of derivation and use the Jacobian matrix for the change of coordinates, Eq. (3.4). We now need to solve this equation to determine the allowed values of the energy and the wave function for a particle confined to move in a ring.

In order to solve this differential equation, first of all we note since since for a particle in a ring the radial coordinate, $r = R$, is constant, then the wave function can depend only on the polar angle: $\Psi = \Psi(\varphi)$. Therefore, the derivatives of the wave function with respect to the radial coordinate r will vanish. With this simplification, the general expression of Schroedinger's equation in polar coordinates, Eq. (3.15), reads

$$-\frac{\hbar^2}{2I}\frac{d^2}{d\varphi^2}\Psi(\varphi) = E\Psi(\varphi)\,, \tag{3.16}$$

where we have used the fact that $mR^2 = I$ is the moment of inertia of the particle. Formally, Eq. (3.16) is the same type of differential equation as that of a one-dimensional free-particle, Eq. (1.24). Therefore, it admits the same form of solutions, namely a linear combinations of exponentials of the form:

$$\Psi(\varphi) = A\exp\left(i\sqrt{\frac{2IE}{\hbar^2}}\varphi\right) + B\exp\left(-i\sqrt{\frac{2IE}{\hbar^2}}\varphi\right)\,, \tag{3.17}$$

where A and B are integration constant to be determined by the boundary conditions of the problem as well as by the normalization condition of the wave function.

In Eq. (3.17) the particle energy E can take in principle any value. However, we need to make sure that the wave function satisfied the boundary conditions of the problem. For a particle restricted to move in a ring, the wave function should satisfy the periodicity condition

$$\Psi(\varphi + 2\pi) = \Psi(\varphi)\,, \tag{3.18}$$

given that physically the angles φ and $\varphi + 2\pi$ are identical and thus the quantum state should be the same. What are the implications of these periodicity boundary condition? For the wave function Eq. (3.17) to be invariant over the transformation Eq. (3.18), we need to have that

$$\exp\left(i\sqrt{\frac{2IE}{\hbar^2}}(\varphi + 2\pi)\right) = \exp\left(i\sqrt{\frac{2IE}{\hbar^2}}\varphi\right) \quad \to \quad \exp\left(2\pi i\sqrt{\frac{2IE}{\hbar^2}}\right) = 1\,, \tag{3.19}$$

which in turn implies that only specific values of the energy E will lead to physically acceptable solutions, namely those that satisfy

$$\sqrt{2IE_m}/\hbar = m \tag{3.20}$$

with m an integer number. Therefore we find that

The energy of a free particle confined to a ring is *quantized* due to the periodic boundary conditions. This means that we can label the physically allowed wave functions as $\Psi_m(x)$, with m being an angular quantum number that labels each eigenvector of the Hamiltonian.

Therefore, putting everything together, the solutions of the Schroedinger equation for a particle confined to a ring will be of the form of Eq. (3.17) subjet to the condition Eq. (3.20), that is, we will have

$$\Psi_m(\varphi) = A\exp\left(\pm im\varphi\right) \tag{3.21}$$

where m is the angular quantum number labelling the solutions, each with an associated kinetic energy with value

$$E_m = \frac{m^2\hbar^2}{2I}\,. \tag{3.22}$$

Moreover, using the relation between energy and angular momentum, $E = J_z^2/2I$, we find that also the angular momentum is quantised,

$$J_{z,m} = m\hbar\,, \tag{3.23}$$

and therefore the quantization of the energy of the particle implies that also its angular momentum is quantized. We observe here that $\hbar$ plays the role of the *quantum* of angular momentum: for all the states of the particle in a ring, the angular momentum is an integer multiple of $\hbar$.

Note that in general the solutions in Eq. (3.21) will not be normalised. We can however easily normalise them by applying the normalisation condition,

$$\int_0^{2\pi} d\varphi\, \Psi_m^*(\varphi)\Psi_m(\varphi) = A^2 \int_0^{2\pi} d\varphi\, e^{\mp im\varphi}e^{\pm im\varphi} = A^2 2\pi = 1 \quad \to \quad A = \frac{1}{\sqrt{2\pi}}\,, \tag{3.24}$$

where note how we integrate over all the allowed values that the polar angle φ can take.

We have seen explicitly how for the particle in a ring the periodic boundary conditions that the wavefunction must satisfy imply that both the energy and the angular momentum are quantised. Before moving forward, let's study in a bit more detail the role than angular momentum plays in quantum mechanics.

3.3 Angular momentum in quantum mechanics.

The analysis of the quantum mechanics of a particle undergoing circular motion in a ring has demonstrated that the angular momentum is quantized. Let us now take a closer look at the interpretation of these results. At the classical level, for a particle following a circular motion in the (x, y) plane, its angular momentum is given, using the definition of the cross-product between two vectors, by

$$\vec{J} = \vec{r} \times \vec{p} = (x, y, 0) \times (p_x, p_y, 0) = (0, 0, xp_y - yp_x) = (0, 0, J_z)\ . \tag{3.25}$$

In quantum theory, physical observables have associated operators that act on the wave function. What is therefore the operator $\hat{J}_z$ associated to the z component of the angular momentum J_z?

To determine the expression of $\hat{J}_z$, we need to exploit the fact that J_z is expressed in terms of the position $\vec{x}$ and linear momentum $\vec{p}$ vectors, for which we know which are the corresponding operators. Thus we need to start with the classical expression

$$J_z = xp_y - yp_x\,, \tag{3.26}$$

and implement the following replacements:

Observable	**Operator**
x	$\hat{x} = x$
y	$\hat{y} = y$
p_x	$-i\hbar\frac{\partial}{\partial x}$
p_y	$-i\hbar\frac{\partial}{\partial y}$

and this way we can identify what is the expression for the angular momentum operator in quantum mechanics:

$$\widehat{J}_z = \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\,, \tag{3.27}$$

where note that we use partial rather than total derivatives, since the wave function depends now on the two spatial coordinates x and y.

Angular momentum in polar coordinates. As we have seen in the discussion of Sect. 3.2, the quantum description of the motion in a particle in a system with rotational symmetry is most transparent when using polar coordinates. So we want to express the angular momentum operator Eq. (3.27) in polar coordinates rather than in Cartesian coordinates. For this we will use the chain rule of the derivative, so that for a general function $f(x,y)$ one has:

$$\begin{aligned}\frac{\partial}{\partial x} f(x,y) &= \frac{\partial}{\partial x} f\left(r(x,y),\varphi(x,y)\right) = \frac{\partial f}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial f}{\partial \varphi}\frac{\partial \varphi}{\partial x} = \cos\varphi \frac{\partial f}{\partial r} - \frac{\sin\varphi}{r}\frac{\partial f}{\partial \varphi}\,, \\ \frac{\partial}{\partial y} f(x,y) &= \frac{\partial}{\partial y} f\left(r(x,y),\varphi(x,y)\right) = \frac{\partial f}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial f}{\partial \varphi}\frac{\partial \varphi}{\partial y} = \sin\varphi \frac{\partial f}{\partial r} + \frac{\cos\varphi}{r}\frac{\partial f}{\partial \varphi}\,,\end{aligned} \tag{3.28}$$

where we have exploited the expressions for the Jacobian matrix associated to the change of coordinates, Eq. (3.4). Therefore, once we include also the contribution from the position operator we find

$$x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} = r\cos\varphi\left(\sin\varphi\frac{\partial}{\partial r} + \frac{\cos\varphi}{r}\frac{\partial}{\partial\varphi}\right) - r\sin\varphi\left(\cos\varphi\frac{\partial}{\partial r} - \frac{\sin\varphi}{r}\frac{\partial}{\partial\varphi}\right) = \frac{\partial}{\partial\varphi}\,, \tag{3.29}$$

where note that the contribution proportional to the partial derivative with respect to r vanishes. Therefore, we find that for a particle moving in the (x,y) plane, its angular momentum operator Eq. (3.27) has the following expression in polar coordinates:

$$\widehat{J}_z = \frac{\hbar}{i}\frac{\partial}{\partial\varphi}\,. \tag{3.30}$$

What are the eigenvalues and eigenvectors that we can associate to $\hat{J}_z$? First of all, we note that since we have constructed it in terms of $\hat{\vec{x}}$ and $\hat{\vec{p}}$, then $\hat{J}_z$ is automatically an Hermitian operator and thus it satisfies all the properties that were presented in Sect. 1.6. Second, if we construct the eigenvalue equation associated to $\hat{J}_z$ we have

$$\hat{J}_z\Psi_m(\varphi) = \frac{\hbar}{i}\frac{\partial}{\partial\varphi}\Psi_m(\varphi) = j_m\Psi_m(\varphi)\,, \tag{3.31}$$

where Ψ_n and j_n are the eigenfunctions and eigenvectors of the angular momentum operator respectively, labelled by the quantum number m. What are the solutions of this eigenvalue equation? We can try again a solution of the plane wave form

$$\Psi(\varphi) = A\exp(\pm iB\varphi)\,, \quad \to \quad \frac{\hbar}{i}\frac{\partial}{\partial\varphi}\Psi(\varphi) = \frac{\hbar}{i}A\,(\pm iB)\exp(\pm iB\varphi)\,, \tag{3.32}$$

which is indeed a solution of Eq. (3.31) with eigenvalue $j = \pm\hbar B$. In addition, as in the case of the Schroedinger equation, these wave functions should satisfy the periodicity boundary condition $\Psi(\varphi+2\pi) = \Psi(\varphi)$, which is satisfied provided that $B = m$ is an integer number. Putting everything together, we find that the eigenvectors and eigenvalues of the angular momentum operator in polar coordinates, $\hat{J}_z$ in Eq. (3.30), are labelled by the quantum number m and are given by

$$\Psi_m(\varphi) = A\exp\left(\pm im\varphi\right)\,, \quad j_m = \pm m\hbar\,. \tag{3.33}$$

By comparing with Eq. (3.21), we find that

> for a free particle in a ring the eigenfunctions of the angular momentum and Hamiltonian (Schroedinger equation) operators are identical, and therefore, these quantum states have well defined values both of the kinetic energy and of the angular momentum.

We also note that the quantisation conditions for the angular momentum that we found when solving Schroedinger equation for the free particle, Eq. (3.23) are the same than when solving the eigenvalue equation for the angular momentum operator.

Relation to Heisenberg's uncertainty principle. From our discussion of the free particle in one-dimension in Sect. 1.6, we found that the eigenvectors of its Hamiltonian were states with well-defined momentum, $\Delta p_x = 0$, and therefore a maximal uncertainty in its position, $\Delta x \to \infty$, and that this was a direct consequence of Heisenberg's uncertainty principle.

A similar situation arises from the free particle confined in a ring. We find that for the eigenvectors of Schroedinger equation Ψ_m, both their energy E_m and angular momentum j_m are well-defined and thus have zero uncertainty. What are the consequences for the localisation of the particle moving in a ring? If we compute the probability density in the polar angle φ, we find that

$$P(\varphi)\, d\varphi = \Psi^*_m(\varphi)\Psi_m(\varphi)\, d\varphi = \left(\frac{1}{\sqrt{2\pi}}e^{\mp im\varphi}\right)\left(\frac{1}{\sqrt{2\pi}}e^{\pm im\varphi}\right) d\varphi = \frac{1}{2\pi}d\varphi\,, \tag{3.34}$$

which is *independent* of the value of the polar angle φ, and therefore the particle is *maximally delocalised* along the ring: it has the same probability of being found anywhere in the ring.

The reason for this, as you might have guessed by now, is that the angular momentum J_z and the polar angle φ are *complementary variables* and therefore satisfy a Heisenberg uncertainty relation. To see this, we can start from the Heisenberg uncertainty principle satisfied by the linear momentum and the position,

$$\Delta x\, \Delta p_x \geq \frac{\hbar}{2}\,. \tag{3.35}$$

Then we can use that in polar coordinates, for a small change in the polar angle and using the fact that the radial coordinates r is constant for the particle in a ring we have

$$\delta x = r\left(\cos\varphi_2 - \cos\varphi_1\right) = -2r\sin\frac{\varphi_2+\varphi_1}{2}\sin\frac{\varphi_2-\varphi_1}{2} \simeq -r\,\delta\varphi\,. \tag{3.36}$$

Now using the relationship between angular and linear momentum for the particle in a ring, $p = J_z/r$ we have that $\Delta p = \Delta J_z/r$, and thus we can write Eq. (3.35) as follows

$$\Delta x\, \Delta p_x = r\,\Delta\phi\,\Delta J_z/r = \Delta\varphi\,\Delta J_z \geq \frac{\hbar}{2}\,. \tag{3.37}$$

Therefore, we conclude that the polar angle φ and the angular momentum J_z are *complementary variables* and thus have associated a Heisenberg uncertainty relation. This explains why for a particle in a ring its position is maximally delocalised: since we have states with well-defined angular momentum, then $\Delta J_z = 0$ and therefore the uncertainty on its angular position is $\Delta\varphi \to \infty$, much as in the case of the free particle in one dimension.

3.4 Quantum mechanics of a particle in spherical motion

Following this discussion of the quantum mechanics of the particle in a ring, and in particular of the role played by the angular momentum, we now move to discuss the case of a quantum particle confined to move in the surface of a sphere of constant radius $r = R$. Such configuration is important to describe the electronic wave functions of atoms and molecules, as we will see in the subsequent lectures. As in the case of the particle in a ring, it is advantageous to use a coordinate system that exploits the rotational symmetry of the

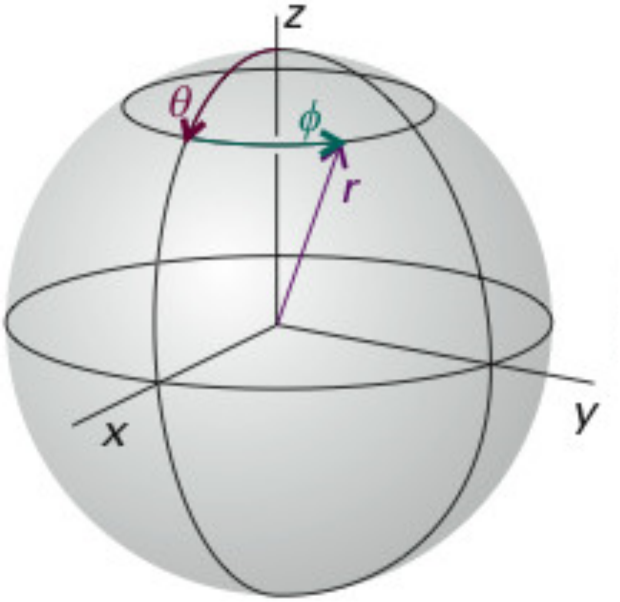


Figure 3.3: The motion of a particle confined to the surface of a sphere with fixed radius r are best described using spherical coordinates, see also Fig. 3.2.

problem, and therefore here we will adopt spherical coordinates as represented in Fig. 3.3 and defined by the coordinate transformations of Eq. (3.5).

Once we transform the free-particle Hamiltonian in three space dimensions from Cartesian to spherical coordinates, the Schroedinger equation for a free particle reads

$$-\frac{\hbar^2}{2m}\left[\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\left(\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}+\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\right)\right]\Psi(r,\theta,\varphi)=E\Psi(r,\theta,\varphi)\,. \tag{3.38}$$

This equation can be simplified by taking into account the fact that the wave function cannot depend on r since it is restricted to move in the surface of the sphere defined by $r = R$. Therefore, the differential equation that needs to be solved is:

$$-\frac{\hbar^2}{2mR^2}\left(\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}+\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\right)\Psi(\theta,\varphi)=E\Psi(\theta,\varphi)\,. \tag{3.39}$$

In order to solve this equation, we will use the method of *separation of variables*. In this method, one *assumes* that the full wave function is the product of a function of the polar angle θ only, $\Psi_\theta(\theta)$, and another function of the azimuthal angle φ only, $\Psi_\varphi(\varphi)$, namely we have that our *ansatz* reads

$$\Psi(\theta,\varphi)=\Psi_\theta(\theta)\Psi_\varphi(\varphi)\,, \tag{3.40}$$

and now we need to try if this *ansatz* satisfies Eq. (3.39). By plugging this ansatz it into the Schroedinger equation, one finds that

$$\frac{\Psi_\theta}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\Psi_\varphi+\frac{\Psi_\varphi}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\Psi_\theta=-\frac{2mER^2}{\hbar^2}\Psi_\theta\Psi_\varphi\,, \tag{3.41}$$

which can be simplified to read

$$\frac{1}{\Psi_\varphi}\frac{\partial^2}{\partial\varphi^2}\Psi_\varphi+\frac{\sin\theta}{\Psi_\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\Psi_\theta=-\frac{2mER^2}{\hbar^2}\sin^2\theta\,. \tag{3.42}$$

In the above equation, the first term of the LHS depends only on φ, while the second term of the LHC as well as the RHS depend only on θ. Therefore, the only way that this equation can be satisfied is if both the φ-dependent piece and the θ-dependent piece are equal to the same constant value, which we will denote for convenience m_l^2 for reasons that will become clear later. For the time being, m_l^2 just stands for a real positive constant, with no other restrictions about the values that it can take.

Therefore, we end up with two separate ordinary differential equations, each depending on a single variable, that we need to solve:

$$\frac{1}{\Psi_\varphi}\frac{\partial^2}{\partial\varphi^2}\Psi_\varphi=-m_l^2\,, \tag{3.43}$$

$$\frac{\sin\theta}{\Psi_\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\Psi_\theta+\frac{2mER^2}{\hbar^2}\sin^2\theta\;=m_l^2. \tag{3.44}$$

For simplicity, in the following we will define

$$\epsilon\equiv\frac{2mR^2}{\hbar^2}E\,, \tag{3.45}$$

which is now a dimensionless number proportional to the total energy of the particle E. Our goal is to separately determine the expressions of $\Psi_\theta(\theta)$ and $\Psi_\varphi(\varphi)$ so that we can reconstruct the whole wave function Eq. (3.40).

Let us now solve these two independent differential equations. First of all, we want to solve the equation for the azimuthal angle, which can be rearranged to read

$$\frac{\partial^2}{\partial\varphi^2}\Psi_\varphi(\varphi)=-m_l^2\Psi_\varphi(\varphi) \tag{3.46}$$

but this is nothing but the Schroedinger equation for a free particle in a ring, Eq. (3.16). Therefore we know that m_l will be quantised, in the sense that it can only adopt integer numbers, and that the wave function will be of the form

$$\Psi_{\varphi,m_l}(\varphi)=\frac{1}{\sqrt{2\pi}}\exp(\pm i m_l\varphi)\,, \tag{3.47}$$

where m_l can only take only integer values and the wave function is normalised.

Now we need to solve the second differential equation, the one that determines the dependence of the wave function on the polar angle θ:

$$\frac{\sin\theta}{\Psi_\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\Psi_\theta+\epsilon\sin^2\theta\;=m_l^2. \tag{3.48}$$

This differential equation is of a type called *Associated Legendre* equations. It is beyond the scope of the course to determine its solutions, and here we will just take them as given. As in the case of Ψ_φ, Ψ_θ must also obey periodic boundary conditions, in this case we have that

$$\Psi_\theta(\theta+2\pi)=\Psi_\theta(\theta)\,, \tag{3.49}$$

and this leads to introducing another quantum number related to angular momentum, which we denote by l. The values that these two quantum numbers, l and m_l, can take, are not independent, and it can be shown that the only allowed values are

$$\begin{aligned} l &= 0, 1, 2, 3, \ldots \\ m_l &= -l, -l+1, -l+2, \ldots, l-2, l-1, l \,, \end{aligned} \tag{3.50}$$

so that for a quantum state characterised by the quantum number l, then m_l can only take $(2l+1)$ values, and one has that $|m_l| \leq l$. As the values of the quantum numbers l and m_l increases, the expressions for the spherical Harmonics becomes more complicated.

Putting everything together, we find that the solutions of the Schroedinger equation for a particle in a sphere, Eq. (3.39), are given by

$$\Psi(\theta, \varphi) = \Psi_{\theta,l,m_l}(\theta)\Psi_{\varphi,l}(\varphi) \equiv Y_{l,m_l}(\theta, \varphi) \,, \tag{3.51}$$

where the functions $Y_{l,m_l}(\theta,\varphi)$ are called the *spherical harmonics*. The explicit expressions for the first few spherical harmonics are given by:

l	m_l	$Y_{l,m_l}(\theta,\varphi)$
0	0	$\left(\frac{1}{4\pi}\right)^{1/2}$
1	0	$\left(\frac{3}{4\pi}\right)^{1/2}\cos\theta$
1	±1	$\mp\left(\frac{3}{8\pi}\right)^{1/2}\sin\theta e^{\pm i\varphi}$
2	0	$\left(\frac{5}{16\pi}\right)^{1/2}\left(3\cos^2\theta - 1\right)$

It can be checked that all spherical harmonics are appropriately normalised, as corresponds to a quantum wave function. Note that for the ground state, the spherical harmonic with values $(l, m_l) = (0, 0)$ is just a constant.

Let us discuss some important properties of the solution of Schroedinger equation for the particle in the sphere, that follow from the solution of the corresponding differential equations subject to the appropriate boundary conditions.

- Quantum states are labelled by a pair of quantum numbers (l, m_l), which have to satisfy the restrictions of Eq. (3.50).

- A quantum state characterised by the quantum numbers (l, m_l) has associated a kinetic energy E_l and *total angular momentum* J_l given by

  $$E = l(l+1)\frac{\hbar^2}{2mR^2} \,, \qquad J_l = \sqrt{l(l+1)}\hbar \,, \tag{3.52}$$

  as can be checked by recalling that $E = J^2/2mR^2$. Note that both J_l and E_l are independent of m_l. The total angular momentum $J_l \equiv |\vec{J}_l|$ should not be mixed with its component in the z direction J_z, since in general for a particle moving in three dimensions as in this case these can be different.[2]

---

[2] For a particle moving in two dimensions, then J_l and J_z coincide since they represent the same physical quantity.

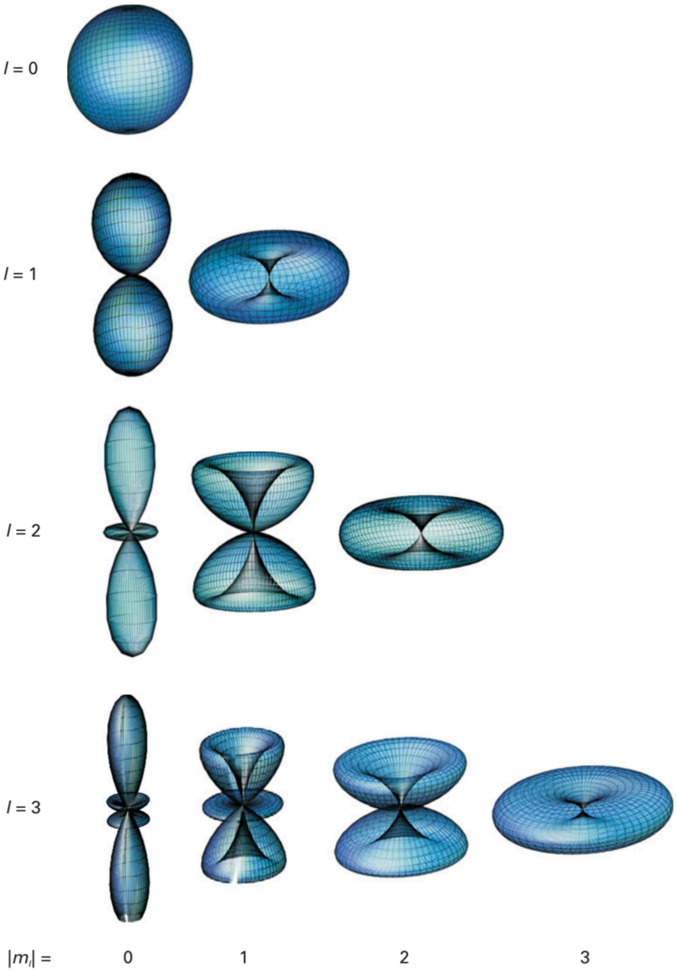


Figure 3.4: Schematic representation of the wave function of a particle in a sphere (the spherical harmonics defined in Eq. (3.51)) for the $l = 0, 1, 2, 3$. The distance of a point on the surface from the origin is proportional to the square modulus of the amplitude of the wave-function at that point.

- States with larger l correspond to greater angular momentum and kinetic energy E_l and J_l, as well as shorter de Broglie wavelengths.

  In Fig. 3.4 we show a schematic representation of the wave function of a particle in a sphere (the spherical harmonics defined in Eq. (3.51)) for the $l = 0, 1, 2, 3$. The distance of a point on the surface from the origin is proportional to the square modulus of the amplitude of the wave-function at that point.

- The value of m_l determines the *component of the angular momentum in the z direction*, as opposed to its total magnitude,:

  $$J_z = m_l \hbar \,, \tag{3.53}$$

  that can take $2l+1$ possible values. Therefore we find that both the *total magnitude* and the *orientation* of the angular momentum $\vec{J}$ are quantised for this system.

  In Fig. 3.5 we illustrate the allowed orientations of angular momentum vector $\vec{J}$ when $l = 2$, and only the values $m_l = -2, 1, 0, 1, 2$ are allowed. Recall that m_l determines the value of $J_z = m\hbar$, and that the total magnitude of the vector is fixed, $J = \sqrt{l(l+1)}\hbar = \sqrt{6}\hbar$ for $l = 2$.

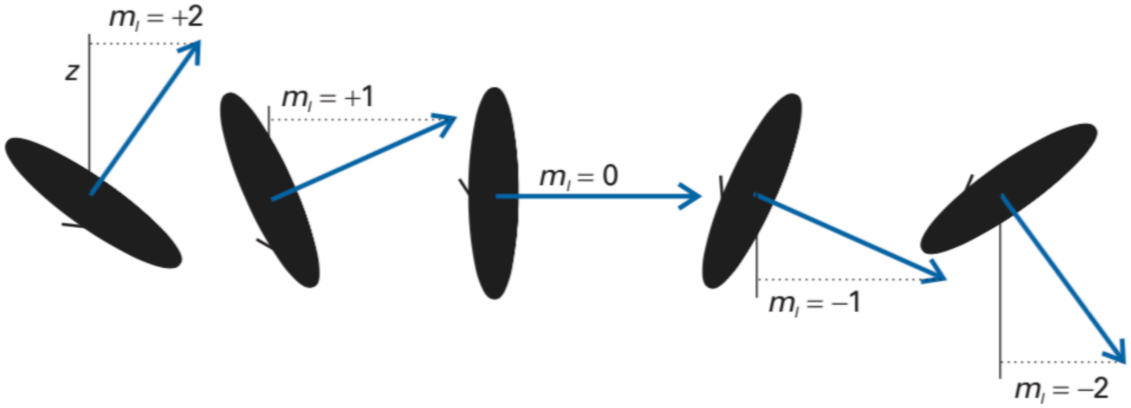


Figure 3.5: The permitted orientations of angular momentum vector $\vec{J}$ when $l = 2$, and only the values $m_l = -2, 1, 0, 1, 2$ are allowed. Recall that m_l determines the value of $J_z = m\hbar$, and that the total magnitude of the vector is fixed, $J = \sqrt{l(l+1)}\hbar = \sqrt{6}\hbar$ for $l = 2$.

Angular momentum in the x and y directions. So far, we have considered only the total magnitude and the value in the z component of the angular momentum $\vec{J}$. What about the components in the other directions? Using the same reasoning as in Eq. (3.27), one can show that operators associated to J_x and J_y are given by

$$\widehat{J}_x = \frac{\hbar}{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) , \tag{3.54}$$

$$\widehat{J}_y = \frac{\hbar}{i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) , \tag{3.55}$$

Moreover, one can demonstrate from their operator definition that the three components of the angular momentum do not commute with each other, and indeed one has that

$$\begin{aligned}\left[\hat{J}_x, \hat{J}_y\right] &= i\hbar\hat{J}_z \ , \\ \left[\hat{J}_y, \hat{J}_z\right] &= i\hbar\hat{J}_x \ , \\ \left[\hat{J}_z, \hat{J}_x\right] &= i\hbar\hat{J}_y \ ,\end{aligned} \tag{3.56}$$

These results implies that the three components of the angular momentum are *complementary variables* among them, and that therefore cannot be simultaneously determined with arbitrary precision. In other words, they satisfy relations of the Heisenberg uncertainty principle type: it is not possible for example to fully determine at the same time J_z and J_x. This means that for states such as the spherical harmonics of Eq. (3.51), which have associated well-defined values of J_z, the corresponding values of J_x and J_y will be undetermined.

On the other hand, the total angular momentum operator squared, which can be constructed as as

$$\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 \, , \tag{3.57}$$

can be shown to *commute* with the individual components, namely one has that

$$\left[\hat{J}^2, \hat{J}_x\right] = \left[\hat{J}^2, \hat{J}_y\right] = \left[\hat{J}^2, \hat{J}_z\right] = 0 \tag{3.58}$$

and therefore we can conclude that

The total magnitude of the angular momentum vector, $|\vec{J}|$, and its component in the z direction, J_z, can be simultaneously determined in a quantum system since the associated operators commute.

This result could have been predicted from the wave functions of the particle in a sphere, the spherical harmonics of Eq. (3.51): there we saw that each state had a well-defined value of the total angular momentum, determined by l, namely $|\vec{J}| = (l(l+1))^{1/2}\hbar$, as well as of its component in the z direction, determined by the value of m_l, with $J_z = m_l\hbar$.

Summary

To summarize, some important concepts that we have learned in this lecture about the application of quantum theory to a particle moving either in a circular or an spherical motion are the following:

- Imposing cyclical boundary conditions leads to the quantization of the angular momentum associated to the rotating particle.
- In the same way as Heisenberg's uncertainty principle prevents the simultaneous determination of the position and linear momentum of a particle, it is not possible to determine simultaneously the angular momentum and the angular position of a particle.
- Only the total magnitude of the angular momentum $|\vec{J}|$ and its component in a given direction J_i can be simultaneously determined. It is not possible to determine simultaneously the values of two different components of $\vec{J}$.
- The wave function of a particle in a sphere is given by the spherical harmonics, whose shape is determined by the values of the angular quantum numbers (l, m_l) associated to each state.

In the next lecture, we will extensively use these concepts to the study of the solutions of the Schroedinger equation applied to the hydrogen atom.

Further reading

Further reading material about the topics covered in this lecture can be found in the following sections of the course textbook (*Physical Chemistry*, Atkins and De Paula, 10th edition):

- **8C.1**: Rotation in two dimensions.
- **8C.2**: Rotation in three dimensions.

van Quantum tot Molecuul

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

4 HC5: The hydrogen atom

In this lecture, we present the quantum mechanics of the hydrogen atom, one of the most important systems for our understanding of atomic and molecular phenomena. A crucial property of this system is that its Schroedinger equation can be solved analytically, and this allows us to understand in a detailed way the resulting properties of the wave functions of the hydrogen atom. This knowledge will represent the building block for our subsequent studies, first of more complicated atoms composed by more than one electron, and then of molecules composed by two atoms or more.

The **learning goals** of this lecture are:

(a) Solving the Schroedinger equation for the hydrogen atom.

(b) Interpreting physically the radial and angular wave functions of the hydrogen atom.

(c) Understanding how the concept of probability density is modified in spherical coordinates as compared to Cartesian coordinates.

(d) Being able to draw and identify the main qualitative features of the electronic orbitals of the hydrogen atom.

Introduction. The hydrogen atom plays a central role in this course for a number of reasons:

- It is the *simplest possible atom* that exists, composed by only one proton and one electron.
- The Schroedinger equation of the hydrogen atom can be *solved exactly*, so we can study its properties in great detail without the need of any numerical approximations.
- The main features of the atomic orbitals of the hydrogen atom can then be applied to the description of other more complex atoms, in particular of *multi-electron atoms* (HC6).
- The hydrogen atom forms the basis for our understanding of *molecular* bonds and *molecular structure*.

The hydrogen atom is composed by one proton with electric charge $Q_p = +e$ and one electron with the opposite charge $Q_e = -e$ bound by the electric force attraction, the Coulomb force. Within a hydrogen atom, the electrostatic potential that the electron experiences is given by

$$V(r) = -\frac{1}{4\pi\epsilon_0}\frac{e^2}{r}\,, \tag{4.1}$$

where r is the distance that separates the electron and the proton, and ϵ_0 is the dielectric constant of the vacuum. Given that the mass of the proton is much larger than the mass of the electron, $m_p \simeq 1800\, m_e$, one can consider to very good approximation that the proton is at rest in this system.

An important feature of the Coulomb potential Eq. (4.1) is that it exhibits spherical symmetry: it depends *only* on the distance r between the electron and the proton, but not on the values of the polar θ and azimuthal φ angles of the distance vector $\vec{r}$ that connects them. We will therefore now be able to exploit everything that we learned in the previous chapter concerning the quantum mechanics of systems with spherical symmetry, in particular the solutions of the Schroedinger equation for a particle confined in the surface of a sphere that we presented in Sect. 3.4.

4.1 The Schroedinger equation for the hydrogen atom

The Hamiltonian for the hydrogen atom, which as usual corresponds to the operator associated to the total energy of the system, takes the following form:

$$\begin{aligned} \hat{H} &= \hat{E}^p_{\rm kin} + \hat{E}^e_{\rm kin} + \hat{V}(r) \\ &= -\frac{\hbar^2}{2m_p}\nabla^2_p - \frac{\hbar^2}{2m_e}\nabla^2_e - \frac{e^2}{4\pi\epsilon_0 r}\,, \end{aligned} \tag{4.2}$$

where the *Laplacian operator* is defined as the sum of the second derivatives with respect to the particle coordinates,

$$\nabla^2 \equiv \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\,, \tag{4.3}$$

and we denote by ∇^2_p the Laplacian acting on the proton spatial coordinates and ∇^2_e on those of the electron.

In the same way as in the *two-body problem in classical mechanics*, the problem is simplified if we express the coordinates of the proton and the electron relative to the center of mass of the system. Since $m_p \gg m_e$, the position of the center of mass coincides basically with the position of the proton, which can therefore be considered to be stationary (to be more precise, the center-of-mass of the system is freely moving since no forces act upon it). In the coordinates with respect the center of mass of the system, the Schroedinger equation reads

$$-\frac{\hbar^2}{2\mu}\nabla^2\Psi(r,\theta,\varphi) - \frac{e^2}{4\pi\epsilon_0 r}\psi = E\Psi(r,\theta,\varphi)\,, \tag{4.4}$$

where we have introduced the reduced mass of the proton-electron system,

$$\frac{1}{\mu} = \frac{1}{m_e} + \frac{1}{m_p} \simeq \frac{1}{m_e}\,, \tag{4.5}$$

given the fact that the proton is much more massive than the electron. From the numerical point of view, using m_e instead of μ is in general a very good approximation. Note that Eq. 4.4 is being expressed in the spherical coordinates defined in Sect. 3.1: as was shown there, the problem simplifies when adopting these coordinates that reflect the underlying symmetries of the system. Crucially, note that Eq. (4.4) is now one-particle differential equation, as compared to the two-particle equation that we started with.

In order to solve the differential equation Eq. (4.4), we will try the same method of *separation of variables* as the one used in HC4. That is, we want to express the total wave function of the system, $\Psi(r,\theta,\varphi)$, as a product of functions that depend separately on r, θ, and φ. Note that, as opposed to the problem that we discussed in Sect. 3.4, the electron now is *not confined* to the surface of a sphere, so we also need to account here for the radial dependence of the wave function. Therefore we adopt the following *ansatz* for the electron wave function:

$$\Psi(r,\theta,\varphi) = R(r)Y(\theta,\varphi)\,, \tag{4.6}$$

with $R(r)$ representing the radial part of the wave function and the angular dependence is contained on the *spherical harmonics* $Y(\theta,\phi)$, which as we know correspond to the solution of the Schroedinger equation for a particle confined in the surface of a sphere. Our choice of the spherical harmonics is Eq. (4.6) is justified since, as we will show, now the angular part of the wave function satisfies the same differential equation as the particle in the surface of a sphere of Sect. 3.4.

Using this ansatz, we obtain that the Schroedinger equation in spherical coordinates reads

$$-\frac{\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\Lambda^2\right)RY + V(r)RY = ERY\,, \tag{4.7}$$

where we have used the expression of the Laplacian in spherical coordinates, and where we have defined

$$\Lambda^2 \equiv \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}+\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\,. \tag{4.8}$$

Next, we use the fact that the radial function $R(r)$ is independent of the angular variables and that the spherical harmonics $Y(\theta,\varphi)$ do not depend on the radius r, so that the above equation can be written as

$$-\frac{\hbar^2}{2\mu}\left(Y\frac{\partial^2 R^2}{\partial r^2}+\frac{2Y}{r}\frac{\partial R}{\partial r}+\frac{R}{r^2}\Lambda^2 Y\right) + V(r)RY = ERY\,, \tag{4.9}$$

which can be simplified by multiplying both sides by r^2/RY, leading to

$$-\frac{\hbar^2}{2\mu R}\left(r^2\frac{d^2R}{dr^2}+2r\frac{dR}{dr}\right) + V(r)r^2 - \frac{\hbar^2}{2\mu Y}\Lambda^2 Y = Er^2\,. \tag{4.10}$$

From this expression, we can determine that our ansatz Eq. (4.6) with the separation of variables has been successful: the last term on the left hand side of Eq. (4.10) depends only on the angular variables, while the rest of the equation depends only on r or on constants. Therefore, the only way in which this condition can be satisfied is provided that

$$-\frac{\hbar^2}{2\mu R}\left(r^2\frac{d^2R}{dr^2}+2r\frac{dR}{dr}\right) \quad + \quad Vr + A = Er^2\,, \tag{4.11}$$

$$-\frac{\hbar^2}{2\mu Y}\Lambda^2 Y \quad = \quad A\,, \tag{4.12}$$

which A some unspecified constant, that will be fixed later. Therefore, we end up with two independent differential equations: one for the radial part of the wave function $R(r)$ and another for the angular part $Y(\theta,\varphi)$, which must be solved separately.

Angular dependence. From our study of the quantum mechanics of a particle confined in the surface of a sphere, we know how to solve the angular equation:

$$-\frac{\hbar^2}{2\mu Y(\theta,\varphi)}\Lambda^2 Y(\theta,\varphi) = A\,, \tag{4.13}$$

since using the definition of Λ^2 we have that it is written as

$$-\frac{\hbar^2}{2\mu}\left(\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\right)Y(\theta,\varphi) = AY(\theta,\varphi)\,, \tag{4.14}$$

but this is nothing but Schroedinger's equation for a particle confined to the surface of a sphere, Eq. (3.39), with mass μ. From this identification, it follows that $A = Er^2$, with r the coordinate, and then the solutions of this equation are indeed the spherical harmonics $Y_{l,m_l}(\theta,\varphi)$, labelled by the two angular quantum numbers (l, m_l). Therefore, since we know that the kinetic energy of the particle in a sphere is given by Eq. (3.52), we can conclude that the constant A must be quantised and is given by

$$A = l(l+1)\frac{\hbar^2}{2\mu}\,, \tag{4.15}$$

where l is the principal angular quantum number. Due to the periodic boundary conditions applied to the wave function, this quantum number can only take positive integer values: $l = 0, 1, 2, 3, \ldots$. Note that Eq. (4.15) automatically implies that the radial solutions of the hydrogen atom will also be quantised and at least will be labelled by angular quantum number l.

Radial dependence and the effective potential. We can now turn to solve the radial part of Schroedinger's equation for the hydrogen atom. Using Eq. (4.15), the radial differential equation reads

$$\frac{\hbar^2}{2\mu R}\left(r^2\frac{d^2R}{dr^2} + 2r\frac{dR}{dr}\right) + V(r)r^2 + \frac{\hbar^2 l(l+1)}{2\mu} = Er^2\,. \tag{4.16}$$

This differential equation can be written in a physically more transparent way using the rescaling $R(r) = u(r)/r$, which makes possible to simplify the above equation using

$$\left(r^2\frac{d^2R}{dr^2} + 2r\frac{dR}{dr}\right) = r\frac{d^2u(r)}{dr^2}\,, \tag{4.17}$$

which in turn leads to a simpler differential equation

$$-\frac{\hbar^2}{2\mu}\frac{d^2u}{dr^2} + V_{\rm eff}(r)u = Eu\,, \tag{4.18}$$

where we have defined an *effective interaction potential* $V_{\rm eff}(r)$ by

$$V_{\rm eff}(r) \equiv -\frac{e^2}{4\pi\epsilon_0 r} + \frac{l(l+1)\hbar^2}{2\mu r^2}\,, \tag{4.19}$$

which includes the contributions both from the Coulomb force as well as from the angular momentum of the electron (namely the centrifugal force that grows with l). After these manipulations, we can see how Eq. (4.18) formally corresponds to the Schroedinger equation for a particle of mass μ moving in one dimension r under the effects of the potential $V_{\rm eff}(r)$.

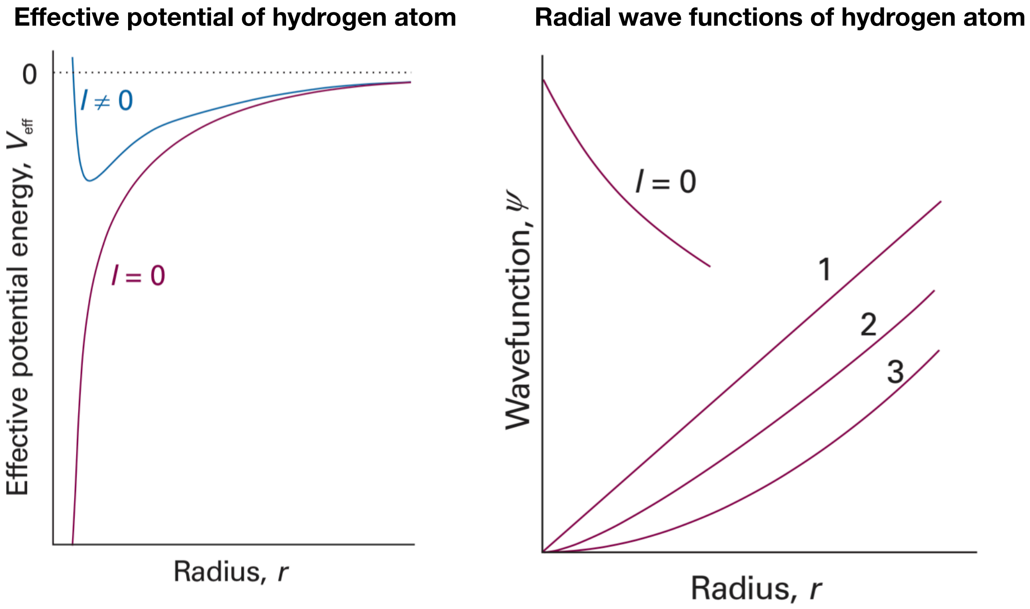


Figure 4.1: Left: the effective interaction potential of the hydrogen atom $V_{\text{eff}}(r)$, Eq. (4.19) includes contributions from the Coulomb potential and from the centrifugal force which is proportional to the electron angular momentum. Here we show its values both for $l \neq 0$ (finite angular momentum) and $l = 0$ (vanishing angular momentum). Right: schematic representation of the radial part of the wave function of the hydrogen atom for different values of the angular quantum number l.

The effective potential Eq. (4.19) has different behaviours when $l = 0$ (vanishing angular momentum) and when $l > 0$ (finite angular momentum), as shown in Fig. 4.1. Indeed, for $l = 0$ we recover the Coulomb *attractive* potential, that tends to $V_{\text{eff}} \to -\infty$ when $r \to 0$. However, in the presence of angular momentum $l \neq 0$, the effective potential becomes *repulsive* at small distances, $V_{\text{eff}} \to +\infty$ when $r \to 0$, therefore disfavouring small values of r in the energy budget of the problem. At large values of r, $r \to \infty$, the Coulomb potential always dominates since it is characterised by a gentle fall-off with r.

The physical interpretation of the effective potential $V_{\text{eff}}(r)$ Eq. (4.19) is the following:

- First of all, we see that the term that depends on the angular momentum l has the opposite sign as the Coulomb potential, implying that it corresponds to a *repulsive effective force* rather than an attractive one. So it has the opposite effect than the Coulomb interaction.
- Second, we observe that when the distance becomes small $V(r) \propto +r^{-2}$, and therefore the electron and the proton cannot get too close to each other.
- This effective potential also appears in the classical description of the two body problem, and its second term corresponds to the contribution of the *centrifugal force*.
- Indeed, while the radial equation Eq. (4.18) describes an effective one dimensional motion, the original physical particle is three-dimensional, and the remnants of this rotating motion are present in the effective potential though the contribution of the centrifugal force.
- We see that in the case that the angular momentum quantum number vanishes, $l = 0$, then the effective potential boils down to the Coulomb potential.

- For small values of the radial coordinate r, then $V_{\text{eff}} \simeq l(l+1)r^{-2}$, meaning that the larger the value of the angular momentum l the farther than on average the electron will be from the nucleus (because of the bigger centrifugal force).
- We see that for large separations r the effective potential goes to zero, with both the Coulomb attraction and the centrifugal force becoming vanishingly small, with the former dominating as discussed above.

Introducing the effective potential therefore highlights the different qualitative behaviours that can be expected from the radial solutions of the Schroedinger equation for the hydrogen atom.

4.2 The radial wave functions of the hydrogen atom

The previous discussion indicates that the radial wave functions $u(r)$ of the hydrogen atom can be derived by solving Eq. (4.18). It is beyond the scope to show how to explicitly solve this differential equation, and here we will limit ourselves to present its solutions and discuss their main features. First of all, when solving Eq. (4.18) one finds that the allowed energy levels are *quantised*, and that they are given in terms a new quantum number n, called the *principal quantum number*, as follows:

$$E_n = -\frac{\mu e^4}{32\pi^2\epsilon_o^2\hbar^2}\frac{1}{n^2}\,, \tag{4.20}$$

where n is a positive integer number: $n = 1, 2, 3, \ldots$. Note that $n = 0$ is not allowed since it corresponds to an unphysical solution.

As in previous quantum systems, the quantisation of the wave function and thus of the allowed energy states follows from *imposing boundary conditions* that need to be respected by the system from physical considerations. In this case, the quantisation of the energy levels indicated in Eq. (4.20) can be shown to be a consequence of the boundary condition that require that $R(r) \to 0$ for $r \to \infty$, namely that the radial component of the wave function must vanish at infinity. This boundary condition is necessary since else the probability distribution associated with the electron would not be well defined, in particular it would not be possible to assign to it a well-defined *probabilistic interpretation*.

Another important observation related to Eq. (4.20) is the fact that $E_n \le 0$ reflects that the electron and the proton in a hydrogen atom form a *bound state*. That is, the energy of the electron+proton system together is smaller than the separate energies of the electron and the proton, and therefore the formation of an hydrogen atom is energetically favoured. For this reason, E_n is also known as the *binding energy* of the electron, namely the energy that needs to be provided to free the electron from this bound state.

The quantum states with energies Eq. (4.20), labelled by the principal quantum number n, can be shown to have associated the following radial wave functions:

$$R_{n,l}(r) = N_{n,l}\rho^l L_{n+1}^{2l+1}(\rho)\exp\left(-\rho/2\right)\,, \tag{4.21}$$

where we have defined the following variables:

$$\rho \equiv \frac{2r}{na}\,, \qquad a = \frac{m_e}{\mu}a_0\,, \qquad a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2}\,. \tag{4.22}$$

Therefore, we see that each of the radial solutions $R_{n,l}(r)$ is labelled in terms of two quantum numbers: (n,l). Each pair of values of (n,l) then corresponds to a different quantum state, though note that for a given value of n the states with different angular quantum number l will be *degenerate* since they have the same energy, see Eq. (4.20).

In the context of atomic and molecular physics, these quantum states are known as *electronic orbitals*.

Note that since $m_e \simeq \mu$, then $a \simeq a_0$ to very good approximation. In Eq. (4.22) we have defined a new constant a_0 with units of length, known as the *Bohr radius*, and whose physical interpretation will become transparent now. One important property is that the quantisation conditions imply that the quantum numbers n and l are not independent, but that that have to satisfy the following condition:

For an orbital characterised by the principal quantum number n, the angular quantum number l can only take the following values: $l = 0, 1, 2, \ldots, n-1$. For example, if $n = 3$, the second excited states of the hydrogen atom, then l can be either 0,1, or 2. Therefore a given principal quantum number n has associated $n-1$ states each with a different value of l.

Let us now discuss the general structure and properties of these solutions, Eqns. (4.21) and (4.22):

- The values of the overall constants $N_{n,l}$ are fixed by the normalisation conditions of the wave functions.
- The radial wave functions $R_{n,l}(r)$ depend on both the radial n and the angular l quantum numbers. Note that the dependence in l arises from the effective potential Eq. (4.19) in the differential equation that determines $R(r)$.
- For small values of r, the wave functions can be shown to scale as $R_{n,l}(r) \simeq r^l$. Therefore, we see that the radial wave function vanishes at the origin except for $l = 0$ (for a system with vanishing angular momentum).

  We also see that the larger the value of l, the smaller the probability of finding the electron close to the proton. This behaviour is a direct consequence of the centrifugal component of the effective potential $V_{\text{eff}}(r)$, which tends to push the electron away for the proton if l is large.

  The small-r behaviour of the radial wave functions is schematically represented in Fig. 4.1. For $l = 1$, the radial wave function grows linearly, $R_1 \simeq r$, for $l = 2$, it grows quadratically, $R_1 \simeq r^2$, and so on. For a system with vanishing angular momentum instead, $l = 0$, we have that the wave function is constant for $r = 0$ and then decreases as r is increased.
- At large distances, the radial wave function falls off exponentially

  $$R_{n,l}(r) \simeq \exp\left(-r/na\right) , \tag{4.23}$$

  ensuring that the wave function *vanishes at infinity* as required by the boundary conditions. Recall that an exponential dominates over any fixed-order polynomial.

  Note that here the product na provides a measure of the *size of the hydrogen atom*: given the steeply-falling character of the exponential, the electron is confined to the region with $r \lesssim na$. We also note that the effective size of this atom increases with the main quantum number n.

Therefore, we can identify a_0, namely the Bohr radius, with the *radius of the hydrogen atom* when the electron occupies the ground state ($n = 1$, the state with smallest energy).

- The interpolation between the limiting small-r and large r limit is determined by a series of special functions called the *Associated Laguerre polynomials* $L_{n+1}^{2l+1}(\rho)$. These functions depend on the two quantum numbers n and l. For integer values of n and l, these polynomials are smooth functions of ρ, for instance the first few ones are given by

$$\begin{aligned} L_0^k(x) &= 1\,, \\ L_1^k(x) &= -x + k + 1\,, \\ L_2^k(x) &= \frac{1}{2}\left[x^2 - 2(k+2)x + (k+1)(k+2)\right]\,. \end{aligned} \tag{4.24}$$

Note in particular that these associated Laguerre polynomials $L_p^k(x)$ tend to a constant as $x \to 0$ and then grows at most as a power p, x^p, as $x \to \infty$. Therefore, at large values of the radius r, the asymptotic behaviour of the radial wave function is dominated by the exponential in Eq. (4.23).

Putting together Eq. (4.21) with the expressions for the Associated Laguerre polynomials, one can construct the explicit form of radial wave functions for any value of n and l. For instance, the three lowest energy states (labelled by $n = 1$ and $n = 2$) will have associated the following wave functions:

$$\begin{aligned} R_{1,0}(\rho) &= 2\left(\frac{1}{a}\right)^{3/2} e^{-\rho/2}\,, \\ R_{2,0}(\rho) &= \frac{1}{\sqrt{8}}\left(\frac{1}{a}\right)^{3/2} (2-\rho)e^{-\rho/2}\,, \\ R_{2,1}(\rho) &= \frac{1}{\sqrt{24}}\left(\frac{1}{a}\right)^{3/2} \rho e^{-\rho/2}\,. \end{aligned} \tag{4.25}$$

In particular, $R_{1,0}(r)$ represents the radial wave function associated to the ground state of the hydrogen atom. One can check that all these wave functions are normalised by integrating their squares over the allowed ranges for the r, θ, and φ coordinates. Note also that, as mentioned above, these wave functions vanish at $\rho = 0$ unless $l = 0$, where they take constant values. For instance, for $n = 1$ and $n = 2$, the value of the $l = 0$ wave functions at the origin are given by

$$\begin{aligned} R_{1,0}(r=0) &= 2\left(\frac{1}{a}\right)^{3/2}\,, \\ R_{2,0}(r=0) &= \frac{2}{\sqrt{8}}\left(\frac{1}{a}\right)^{3/2}\,. \end{aligned} \tag{4.26}$$

As will be shown shortly, this does not mean that the electron has a non-zero probability of being found at $r = 0$ in the same position as the proton. This would be problematic since quantum mechanics would not be able to describe such configuration.

Generalisation to hydrogen-like atoms This discussion can be generalised to *hydrogen-like atoms*, which are defined as atoms that have Z protons and $A - Z$ neutrons but only a single electron. One example of this would be the Helium atom if one of its two electrons has been stripped off. For hydrogen-like atoms, the only difference would be to replace e^2 by Ze^2 in the effective potential of the radial wave function, Eq. (4.19). Therefore the only difference at the level of radial wave functions would be to replace Bohr's radius a_0 by

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e Z e^2} \tag{4.27}$$

or in other words, replacing a_0 in the solution of the hydrogen atom for a_0/Z in the case of hydrogen-like atoms. For the three lowest energy states, the expression for the radial wave functions of hydrogen-like atoms are given by

$$\begin{aligned} R_{1,0}(\rho) &= 2\left(\frac{Z}{a}\right)^{3/2} e^{-Z\rho/2}\,, \\ R_{2,0}(\rho) &= \frac{1}{\sqrt{8}}\left(\frac{Z}{a}\right)^{3/2} (2-\rho)e^{-Z\rho/2}\,, \\ R_{2,1}(\rho) &= \frac{1}{\sqrt{24}}\left(\frac{Z}{a}\right)^{3/2} \rho e^{-Z\rho/2}\,. \end{aligned} \tag{4.28}$$

On the other hand, the angular part of the wave function, given by the spherical harmonics $Y_{l,m_l}(\theta,\varphi)$, is the same for hydrogen and hydrogen-like atoms.

In hydrogen-like atoms, the atomic size now will be smaller than the one of the hydrogen atom by a factor $1/Z$. This can be understood from the fact that having now Z protons instead of one creates a stronger electrical attraction with the electron, and therefore the electron is more likely to be found closer to the atomic nucleus.

Allowed electronic transitions in the hydrogen atom. We can express the values of the energy levels of the hydrogen atom Eq. (4.20) in a more physically transparent way as follows:

$$E_n = -\frac{hc}{n^2}\widetilde{R}_h\,, \tag{4.29}$$

where we have defined

$$\widetilde{R}_h = \frac{\mu e^4}{32\pi^2\epsilon_0\hbar^2}\,, \quad \widetilde{R}_h \equiv \frac{\mu}{m_e}\widetilde{R}_\infty\,, \tag{4.30}$$

and where the so-called *Rydberg constant* is given as

$$\widetilde{R}_\infty = \frac{m_e e^4}{8\epsilon_0^2 h^3 c}\,. \tag{4.31}$$

From Eq. (4.29) is clear that this Rydberg constant, neglecting proton mass effects, is minus the energy associated to the ground state of the hydrogen atom. In other words, it is a measure of the binding energy of the electron in the $n = 1$ state of the hydrogen atom.

If we define the *ionisation energy of the hydrogen atom*, I, as the energy that needs to be supplemented to such an atom in its ground state in order to ionize it, in terms of the Rydberg constant it will be given by

$$I = hc\widetilde{R}_h = 13.6\,\text{GeV}\,, \tag{4.32}$$

which is consistently also denoted as *one Rydberg.* So the Rydberg is a unit suitable to describe ionisation energies in hydrogen-like atoms.

Note that from the distribution of energy levels of the hydrogen atom, Eq. (4.29), one finds that the difference between two neighboring energy levels decreases as n is increased. Indeed, if one has two adjacent

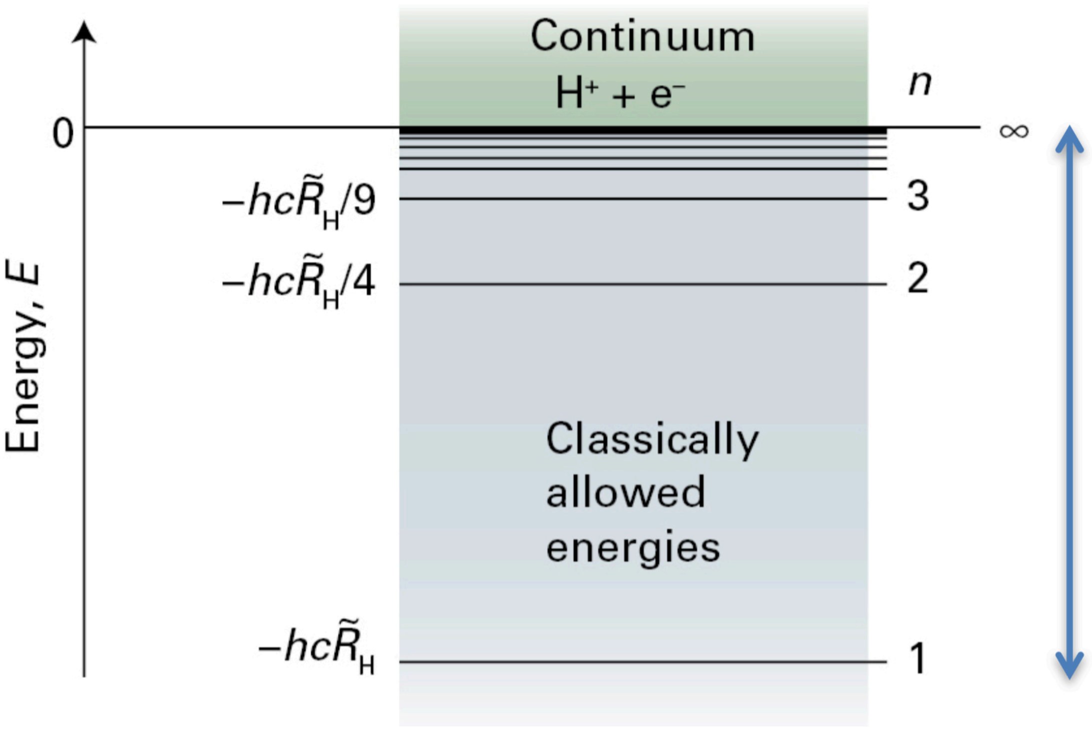


Figure 4.2: Schematic representation of the energy levels of the hydrogen atom. The ground state ($n = 1$) has energy $E_1 = -hc\widetilde{R}_H$, the first excited state ($n = 2$) has energy $E_2 = -hc\widetilde{R}_H/4$ and so on, expressed in terms of the Rydberg constant. As n keeps increasing, the difference between energy levels becomes smaller and smaller. If the atoms finds itself in the energy level n and absorb energy (say from a photon) with $E \geq -E_n$, then the hydrogen atom will ionize and dissociate into a proton and a electron, not bounded by the electromagnetic interaction anymore. Recall that the fact that $E_n < 0$ indicate that the electron and the proton form a *bound state*.

levels E_n and E_{n-1}, then their energy difference is given by

$$\Delta E_n = E_{n-1} - E_n = -\frac{hc}{(n+1)^2}\widetilde{R}_h + \frac{hc}{n^2}\widetilde{R}_h = hc\widetilde{R}_h \frac{2n+1}{(n^2+n)^2} , \tag{4.33}$$

and therefore in the limit of a highly excited state, $n \to \infty$, one as that this energy difference ΔE_n becomes vanishingly small.

This behaviour is illustrated in Fig. 4.2, where we display the schematic representation of the energy levels of the hydrogen atom in terms of the Rydberg constant. The ground state ($n = 1$) has energy $E_1 = -hc\widetilde{R}_H$, the first excited state ($n = 2$) has energy $E_2 = -hc\widetilde{R}_H/4$ and so on. As n keeps increasing, the difference between energy levels becomes smaller and smaller. If the atoms finds itself in the energy level n and absorbs energy (say from a photon) with $E \geq -E_n$, then the hydrogen atom will ionize and dissociate into a proton and a electron, not bounded by the electromagnetic interaction anymore. The higher the value of n, the smaller the energy of the photon that will be needed to dissociate an hydrogen atom. As mentioned above, for $n = 1$ this energy is known as the ionisation energy I of the hydrogen atom.

As is clear from Fig. 4.2 the fact that $E_n < 0$ indicate that the electron and the proton form a *bound state*. Positive energies of the unbounded electron+proton system can take any value (since they are not subjected to the quantisation conditions) and we say that they form a *continuum*.

4.3 The orbital structure of the hydrogen atom

We now have all the ingredients that we need in order to take a closer look at the orbital structure of the hydrogen atom. Recall that by *electronic orbitals* we mean simply the allowed quantum states for the electrons in this system, labelled by the pair of quantum numbers (n, l). First of all, we note that the values of the energy E_n are independent of the values of the angular quantum numbers l and m_l, and depend only on the values of the principal (radial) quantum number n. This means that, in general, there will be several orbitals sharing the same value of the energy E_n: in particular, for a given n, all orbitals with different values of l and m_l will have the same energy. Recall that for a fixed value of n, the allowed values of l are

$$l = 0, 1, 2, \ldots, n-1\,, \tag{4.34}$$

while in turn for a given value of l, the allowed values of m_l are given by

$$m_l = -l, -l+1, \ldots, l-1, l\,, \tag{4.35}$$

where the value of m_l determines the orientation of the angular momentum with respect to the z axis. So for each value of n there exist a number of orbitals that share the same energy E_n, and this number can be computed as

$$N_n = \sum_{l=0}^{n-1}(2l+1) = n^2 - n + n = n^2\,, \tag{4.36}$$

where we have used the expression for the arithmetic sum

$$\sum_{k=1}^{n} k = \frac{1}{2}n(n+1)\,. \tag{4.37}$$

Therefore, we have $N_1 = 1$ orbitals for the ground state, $N_2 = 4$ orbitals for the first excited state, $N_3 = 9$ orbitals for the second excited state, and so on.

For historical reasons, the electronic energy levels $n = 1, 2, 3, 4, \ldots$ in the hydrogen atom are denoted as the $K, L, M, N, \ldots$ *electronic shells*. In turn, orbitals with angular momentum $l = 0, 1, 2, 3, 4, \ldots$ are labelled by the letters $s, p, d, f, g, \ldots$. For instance, an s orbital in the L shell corresponds to $(n, l) = (2, 0)$, while a p orbital in the M shell corresponds to $(n, l) = (3, 1)$.

Following this convention, here we will use the following notation to denote the orbitals of the hydrogen atom in terms of their radial n and angular n quantum numbers:

n	l	notation
1	0	$1s$
2	0	$2s$
2	1	$2p$
3	0	$3s$
3	1	$3p$
3	2	$3d$

A graphical representation of the distribution of the orbitals of the hydrogen atom in shells (labelled by the value of n) and sub-shells (labelled by the value of l) is schematically represented in Fig. 4.3. Each shell

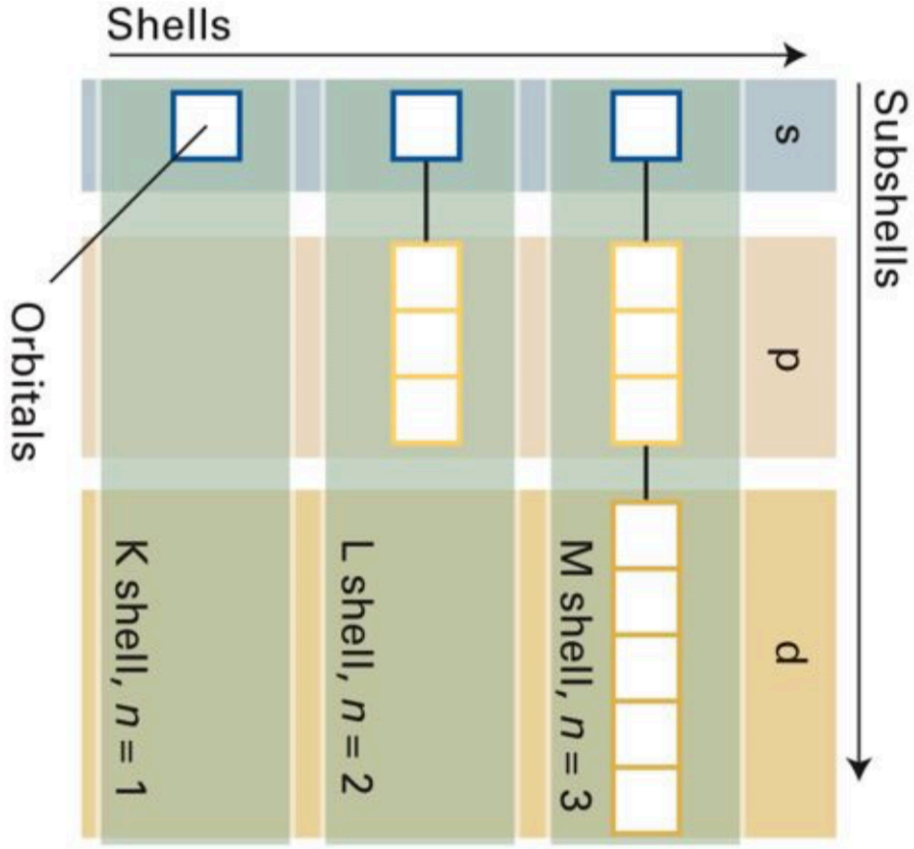


Figure 4.3: The distribution in shells and sub-shells of the orbitals of the hydrogen atom. Each shell corresponds to a different main quantum number n: the K shell is $n = 1$, the L shell is for $n = 2$, and so in. Within each shell, there is a division in subshells labelled by the value of the angular quantum number l. For example, the M shell ($n = 3$) has three subshells: s ($l = 0$), p ($l = 1$), and d ($l = 2$). A subshell with quantum number l contains $(2l - 1)$ orbitals. The total number of orbitals in a shell with principal quantum number n is n^2.

corresponds to a different principal quantum number n: the K shell is $n = 1$, the L shell is for $n = 2$, and so in. Within each shell, there is a division in subshells labelled by the value of the angular quantum number l. For example, the M shell ($n = 3$) has three subshells: s ($l = 0$), p ($l = 1$), and d ($l = 2$). For each subshell there are $2l - 1$ orbitals, corresponding to the possible values of the quantum number m_l can take in that subshell. The total number of orbitals in a shell with principal quantum number n is n^2. Both in the hydrogen and in hydrogen-like atoms, there is a single electron that can occupy any of the orbitals shown in Fig. 4.3.

4.4 Radial probability distributions

An important feature of using spherical coordinates in the description of the hydrogen atom is that one needs some care when constructing the radial probability distributions out of the corresponding radial wave functions. To illustrate this point, let us consider first of all the ground state of the hydrogen atom (orbital $1s$, with $n = 1$ and $l = 0$). The corresponding radial wave function is given by

$$R_0(r) = \frac{2}{a_0^{3/2}} e^{-r/a_0} \,, \tag{4.38}$$

and the associated angular part of the wave function is given by

$$Y_{0,0}\left(\theta,\varphi\right) = (4\pi)^{-1/2}\,, \tag{4.39}$$

which can be combined to construct the total electronic wave function of the ground state of the hydrogen atom:

$$\Psi_0(r,\theta,\varphi) = R_0(r)Y_{0,0}\left(\theta,\varphi\right) = \frac{1}{\sqrt{\pi a_0^3}}e^{-r/a_0}\,. \tag{4.40}$$

This wave function has a maximum for $r = 0$ and then decays exponentially for $r > 0$. Note that this wave function does not depend on the angles θ and φ, and is thus *spherically symmetric.*

As has been discussed above, the fact that the wave function has a maximum at the origin is characteristic of the electronic orbitals of type s, that is, those quantum states for which the angular momentum quantum number vanishes ($l = 0$). As we will now show, this property does *not* imply that the electron has a finite probability of being found at $r = 0$ (which would mean that the proton and the electron are at the same position), as a consequence of the properties of spherical coordinates. Indeed, the probability to find a electron at a position (x, y, z) specified in Cartesian coordinates is given by

$$P(x,y,z) = |\Psi(x,y,z)|^2 dxdydz\,, \tag{4.41}$$

where $dxdydz$ represent the spatial volume where the electron is allowed to be found upon measurement. To obtain the corresponding expression in the case of spherical coordinates, one needs to take into account the *Jacobian matrix* for the change of variables given by

$$dxdydx = \left|\frac{\partial(x,y,z)}{\partial(r,\theta,\varphi)}\right| drd\theta d\varphi = r^2 dr \sin\varphi d\varphi d\theta\,. \tag{4.42}$$

Exploiting this information, we can now compute the probability density associated to the wave function for the ground state of the hydrogen atom, Eq. (4.38), finding

$$\begin{aligned} P_0(r)dr &= \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta |\Psi_0(r)|^2 r^2 dr \sin\theta \\ &= \frac{1}{\pi a_0^3}\int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta d\theta r^2 e^{-2r/a_0} dr = \frac{4}{a_0^3} r^2 e^{-2r/a_0} dr\,, \end{aligned} \tag{4.43}$$

and therefore the radial probability density $P_0(r)$ associated to the ground state of the hydrogen atom is:

$$P_0(r) = \frac{4}{a_0^3} r^2 e^{-2r/a_0}\,, \tag{4.44}$$

where we have integrated over the angular variables, exploiting the fact that Ψ_0 does not depend on the angles (in statistics, we call this process *marginalisation*). When integrating over the angular variables, the following relation is useful:

$$\int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \sin\theta = 4\pi\,, \tag{4.45}$$

which is known as the total *solid angle* in three dimensions.

The result Eq. (4.43) indicates the probability of finding the electron in an interval of size dr at a distance r from the proton. Note that while the wave function itself peaked at $r = 0$, $\Psi_0(r = 0) \neq 0$, the associated probability vanishes at the origin, $P_0(r = 0) = 0$. This implies that the electron has *a vanishing probability* of being found at the same position as the nucleus of the hydrogen atom.

By weighting with the probability Eq. (4.43), one can compute various average properties of the ground state of the hydrogen atom. For instance, one can evaluate the *expectation value* for the distance r between the electron and the proton in the ground state by means of the weighted average:

$$\langle r \rangle = \int_0^\infty dr\, r\, P_0(r) = \frac{4}{a_0^3} \int_0^\infty dr\, r^3 e^{-2r/a_0} dr = \frac{4}{a_0} \times \frac{6a_0^4}{16} = \frac{3a_0}{2} \, . \tag{4.46}$$

This result justifies why the Bohr radius a_0 is good estimate for the average size of the hydrogen atom when this occupies its ground state. To compute this integral, we have used the following result:

$$\int_0^\infty dx\, x^n \exp(-bx) = \frac{n!}{b^{n+1}} \, . \tag{4.47}$$

It is also interesting to compute which is the *most likely value* of r, denoted by r^*, where the electron will be found. Note that in general this value will be *different* from the expectation value $\langle r \rangle$. To determine this, we need to determine where the radial probability density $P_0(r)$ has a maximum by imposing the condition:

$$\frac{dP_0(r)}{dr} = 0 = \frac{4}{a_0} \left(2r^* + r^{*2}(-2/a_0) \right) e^{-2r^*/a} = \frac{8r^*}{a_0}(1 - r^*/a_0)e^{-2r^*/a} \, , \tag{4.48}$$

which is satisfied by $r^* = a_0$. Therefore we find that the Bohr radius a_0 is *the most likely value* where the electron will be found in the ground state of the hydrogen atom.

It is illustrative to plot the radial probability distributions $P(r)$ Eq. (4.43) for different electronic orbitals of the hydrogen atom as a function of r/a_0. These are shown in Fig. 4.4. There are a number of interesting features that can be observed in this graph:

- The higher the value of n, the more likely it will be to find the electron far from the atomic nucleus.
- For a given n, electrons occupying orbitals with $l = 0$ have a larger probability of being found near the nucleus. Note that in all cases the probability distributions vanish at the origin of coordinates, $P(r = 0) = 0$.
- s-type orbitals peak at a distance r *larger* than that of the p-type and other types of orbitals. This means that the higher the value of l, the less spread a given electronic orbital is. In other words, d-type orbitals are less spread than p-type orbitals which in turn are less spread than s-type orbitals.

Using a similar strategy as the one used for the ground state, it is possible to determine the radial coordinate expectation value $\langle r \rangle$ and the most likely radius r^* for any (n, l) orbital of the hydrogen (or hydrogen-like) atoms using the corresponding radial probability densities.

4.5 General properties of the p- and d-type orbitals

Since we have devoted some attention to the discussion of the s-type orbitals of the hydrogen atom, including the ground state of the system, let us now discuss some of the general properties of the p-type and d-type orbitals, starting with the former, and then continuing with the later.

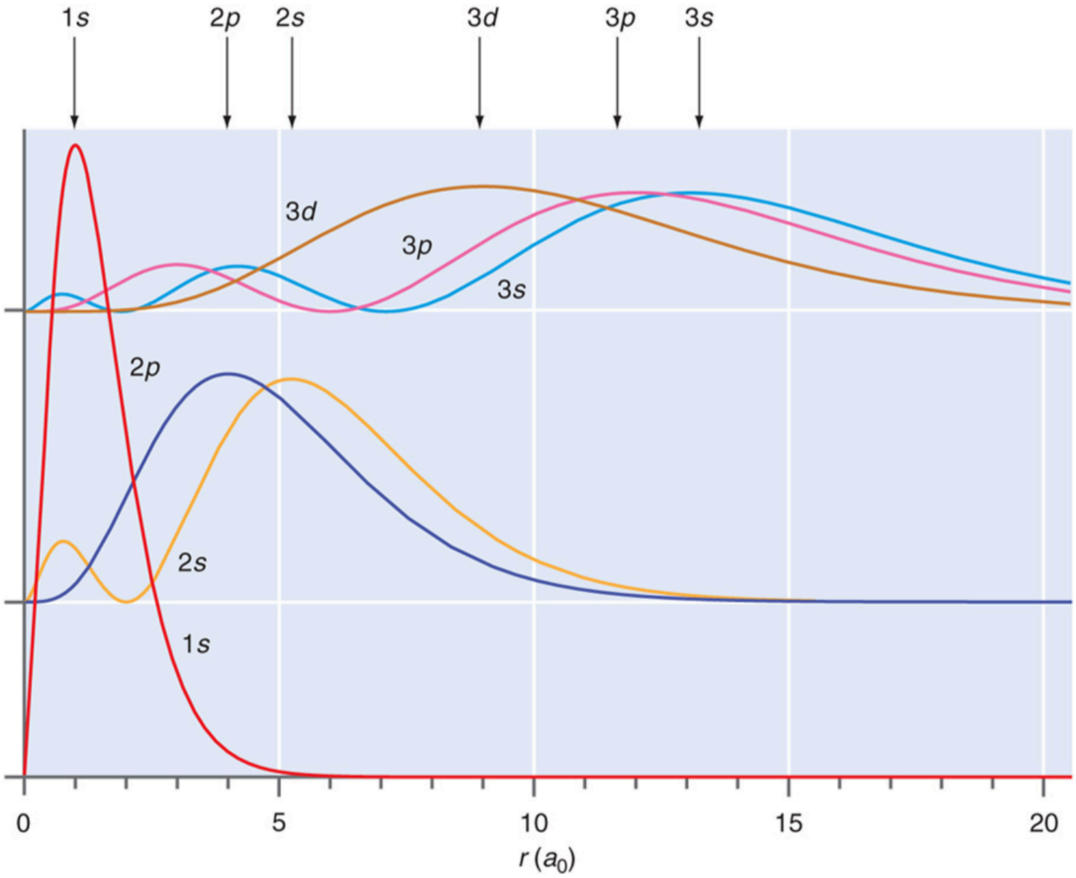


Figure 4.4: The radial probability distributions $P(r)$, Eq. (4.43), for different electronic orbitals of the hydrogen atom as a function of r/a_0. We find that the higher the value of n, the more likely it will be to find the electron far from the atomic nucleus. We also find that for a given n, electrons occupying orbitals with $l = 0$ have a larger probability of being found near the nucleus. All probability distributions vanish at the coordinate origin $P(r = 0) = 0$.

***p*-type orbitals.** The p-type orbitals are defined by an angular quantum number $l = 1$, and therefore a p-type subshell contains three orbitals, with $m_l = -1, 0$ and 1 respectively. The corresponding wave functions can be constructed by combining the radial $R_{n,l}(r)$ and angular $Y_{l,m_l}(\theta, \varphi)$ components of the wave function. For example, for the $2p$ orbital, defined by the $n = 2$ and $l = 1$ quantum wave numbers, in the case that $m_l = 0$ we have that

$$R_{2,1}(\rho) = \frac{1}{24^{1/2}} \left(\frac{1}{a}\right)^{3/2} \rho e^{-\rho/2} \tag{4.49}$$

$$Y_{1,0}(\theta, \varphi) = \left(\frac{3}{4\pi}\right)^{1/2} \cos\theta \,, \tag{4.50}$$

so here the wave function is symmetric with respect to the azimuthal angle φ. Combining the radial and angular components, we have the following wave function for the $2p_0$ orbital of the hydrogen atom.

$$\Psi_{2p_0}(r, \theta) = R_{2,1}(r) Y_{1,0}(\theta) = \frac{1}{4(2\pi)^{1/2}} \left(\frac{1}{a_0}\right)^{5/2} r \cos\theta e^{-r/2a_0} = r \cos\theta f(r) = z f(r) \,, \tag{4.51}$$

where we have used that in spherical coordinates $z = r\cos\theta$, and we have defined the function $f(r)$ that depends only on the radial coordinate $f(r)$ as follows:

$$f(r) \equiv \frac{1}{4(2\pi)^{1/2}} \left(\frac{1}{a_0}\right)^{5/2} e^{-r/2a_0} \, . \tag{4.52}$$

Note that denote the orbital with quantum numbers $(n, l, m_l) = (2, 1, 0)$ as $2p_0$, which is a more compact notation. So we find that the wave function can be written as a radial function $f(r)$ multiplied by z.

As discussed above, the wave functions for hydrogen-like atoms can be readily obtained from the wave functions of the hydrogen atom by means of the replacement $a \to a/Z$. To illustrate this, the wave function for the $2p_0$ orbital of an hydrogen-like atom with atomic number Z would be given by

$$\Psi_{2p_0}(r,\theta) = R_{2,1}(r)Y_{1,0}(\theta) = \frac{1}{4(2\pi)^{1/2}} \left(\frac{Z}{a_0}\right)^{5/2} r\cos\theta e^{-Zr/2a_0} = r\cos\theta f(r) = zf(r) \, . \tag{4.53}$$

The wave function of the $2p_0$ orbital, Eq. (4.51), vanishes everywhere in the plane $z = 0$, which is for this reason called a *nodal plane*. We also see that the wave function is positive for $z > 0$ and negative for $z < 0$, and that moreover it satisfies the reflection property that $\Psi_{2p_0}(r,z) = -\Psi_{2p_0}(r,-z)$.

From its mathematical structure, we also see that for a given value of z, the wave function of the $2p_0$ orbital will be symmetric with respect to the distance from the origin of coordinates r (the distance to the atomic nucleus), with

$$f(r) \propto e^{-r/2a_0} \, , \tag{4.54}$$

and that the wave function vanishes exponentially for $r \gg 2a_0$. Putting together all these properties, we can represent graphically the $2p_0$ orbitals of the hydrogen atom as in Fig. 4.5. To be precise, there we represent the *boundaries* or the orbitals, defined as the volumes that contain a given fraction of the total electron probability. Formally, the electronic orbitals extend to arbitrary high r.

We can observe the following properties of the $2p_0$ orbitals displayed in Fig. 4.5:

- The orbitals vanish in the nodal plane $z = 0$.
- The orbitals are symmetric with respect to this nodal plane, since $|\Psi_{2p_0}(r,z)|^2 = |\Psi_{2p_0}(r,-z)|^2$.
- The orbitals vanish for large r.
- The orbitals exhibit polar symmetry for a given value of z: they depend only on r but not on the angular variables.

Similar symmetry properties apply for the other $2p$ orbitals. For orbitals with quantum numbers $m_l = 1$ and $m_l = -1$, the corresponding wave functions for the hydrogen atom will be given by:

$$\Psi_{2p_{\pm1}} = R_{2,1}(r)Y_{1,\pm1}(\theta,\varphi) = \mp\frac{1}{8\pi^{1/2}} \left(\frac{1}{a_0}\right)^{5/2} r\sin\theta e^{\pm i\varphi} e^{-Zr/2a_0} \propto \mp r\sin\theta e^{\pm i\varphi} f(r) \, , \tag{4.55}$$

These wave functions are now complex quantities, so they cannot be plotted directly. To facilitate their interpretation, we note that since Ψ_{2p_1} and $\Psi_{2p_{-1}}$ are two degenerate eigenfunctions (meaning that they have associated the same energy), then by the properties of quantum mechanics we known that any linear

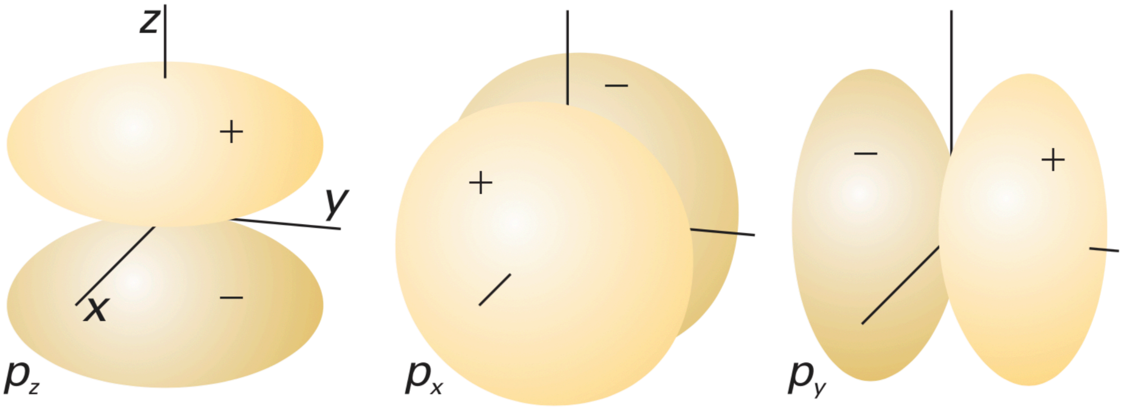


Figure 4.5: Schematic representation of the $2p_0$ (also denoted as $2p_z$), $2p_x$, and $2p_y$ orbitals of the hydrogen atom. Note that each of these orbitals exhibits a nodal plane for $z = 0$, $x = 0$, and $y = 0$ respectively.

combination of them will also be an eigenfunction of the Schroedinger equation. With this motivation, we define the following two linear combinations:

$$\Psi_{2p_x} \equiv -\frac{1}{\sqrt{2}}\left(\Psi_{2p_1} - \Psi_{2p_{-1}}\right) = r\sin\theta\cos\varphi f(r) = x f(r)\,, \tag{4.56}$$

$$\Psi_{2p_y} \equiv \frac{i}{\sqrt{2}}\left(\Psi_{2p_1} + \Psi_{2p_{-1}}\right) = r\sin\theta\sin\varphi f(r) = y f(r)\,, \tag{4.57}$$

again using the definition of spherical coordinates. Therefore we observe that these $2p_x$ and $2p_y$ orbitals share the same symmetry property as the $2p_0$ ($2p_z$) orbitals, namely having a nodal place for which the wave function vanishes. Moreover, this nodal plane is also a *symmetry plane*: the wave function is symmetric (up to an overall sign) with respect to this nodal plane. Therefore, as shown in Fig. 4.5, the p_x and p_y orbitals are nothing but a rotated version of the p_z orbitals (left panel in the Figure).

As opposed to the s-type orbitals, the p-type orbitals are not rotationally symmetric. It can be shown that instead all the orbitals of the hydrogen atom on the s-type orbitals are rotationally symmetric.

General results for the nodal planes. It can be shown that the radial part of the electron wave function $R_{n,l}(r)$ has $n - l - 1$ nodal planes, that is, planes where the wave function vanishes. This follows from the mathematical structure of the wave functions of the hydrogen atom. For example, p orbitals with $l = 1$, will have associated $n - 1$ nodal planes, and indeed we have seen that the $n = 2$ p-type orbitals have a single nodal plane associated. On the other hand, the angular wave function Y_{l,m_l} has l nodal planes. So in total, we see that an orbital defined by the quantum numbers n and l has $(n - l - 1) + (l) = n$ nodal planes in its wave function. You can verify explicitly this property with the explicit expressions of the wave functions that we have presented in this section.

Summary

To summarize, some important concepts that we have learned in this lecture are the following:

- The allowed quantum states for the electrons in a hydrogen (or hydrogen-like) atom are called *electronic orbitals*, and are characterised by three quantum numbers: (n, l, m_l).
- The principal quantum number n determines the energy of the orbital and the average distance $\langle r \rangle$ with respect to the nucleus. It fixes the main electronic shell structure.
- Despite the fact that for $l = 0$ the radial part of the electron wave function vanishes for $r = 0$, the associated probability density is $P(r = 0) = 0$ due to the Jacobian of the transformation from Cartesian to spherical coordinates.
- The angular momentum quantum numbers l and m_l determine the subshell structure. Within each shell defined by n, there will be l subshells each filled by $2l + 1$ orbitals.
- The s-type orbitals ($l = 0$) exhibit rotational symmetry, while the p-type orbitals are instead symmetric with respect to a nodal plane where the wave function vanishes.

In the next lecture, we will study the electronic structure of the rest of the atoms that define the other elements of the periodic table: we will move from single-electron atoms to multi-electron atoms.

Further reading

Further reading material about the topics covered in this lecture can be found in the following sections of the course textbook (*Physical Chemistry*, Atkins and De Paula, 10th edition):

- **9A.1**: The structure of hydrogenic atoms.
- **9A.2**: Atomic orbitals and their energies.

van Quantum tot Molecuul

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

5 HC6: Multi-electron atoms

In this lecture, we discuss how to apply the concepts we learned in HC5 about the quantum mechanics of the hydrogen atom to more complex elements, in particular to multi-electron atoms. Since moving beyond the two-body problem the Schroedinger equation does not admit an analytical solution, we will need to adopt a number of approximations. We will start by Helium, composed by two electrons, and then study atoms with three electrons or more.

The **learning goals** of this lecture are:

(a) Applying the quantum mechanics of the hydrogen atom to the understanding of the electronic structure of multi-electron atoms.

(b) Using well-motivated approximations to express the orbitals of multi-electron atoms in terms of those of the hydrogen system.

(c) Becoming familiar with the concepts of spin, Pauli exclusion principle, and fermions and bosons, and the role they play in the quantum mechanics of the electrons.

(d) Using the building up principle and Hund's role to determine the electronic structure of multi-electron atoms.

Introduction. In the previous lecture, we have studied the quantum mechanics of simplest possible of all chemical elements: the hydrogen atom, composed only by one proton and one electron. We have derived the wave functions for this system as well as for hydrogen-like atoms, that is, heavier elements with only one electron. Now we move to study more complicated atoms, in particular atoms with multiple electrons. We will start with the helium atom, which is the next-to-simplest atom since it contains only two electrons, and then apply the same ideas to atoms with more electrons. As we will show, we will base our study heavily on applying the same ideas and concepts that have appeared in the case of the hydrogen atom, and construct the wave functions of multi-electron atoms using the hydrogen atom wave functions as building blocks, in terms of the so-called *orbital approximation*.

5.1 The helium atom

We start this lecture with the next-to-simplest atom after the hydrogen atom: the helium (He) atom. A helium atom is composed by a nucleus with two protons and two neutrons (hence $Z = 2$ and $A = 4$) and two electrons. Since we always assume the atomic nucleus to be point-like (the nuclear radius is much smaller than the atomic radius), we have to solve in principle a *three-body problem*: the coupled motion of the two electrons and the atomic nucleus. As opposed to the two-body problem (the hydrogen atom) where a fully analytical solution was possible, for the case of the three-body problem it will become necessary to introduce a number of approximations.

The associated Schroedinger equation for the helium atom is given by

$$\widehat{H}\Psi = \left(-\frac{\hbar^2}{2m_e}\nabla_1^2 - \frac{2e^2}{4\pi\epsilon_0 r_1} - \frac{\hbar^2}{2m_e}\nabla_2^2 - \frac{2e^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 r_{12}} \right) \Psi = E\Psi \,, \tag{5.1}$$

where r_1 is the distance between the nucleus and the first electron, r_2 is the distance between the nucleus and the second electron, and r_{12} is the distance between the two electrons. The kinetic terms ∇_1^2 and ∇_2^2 act on the coordinates of the first ($\vec{r}_1$) and second ($\vec{r}_2$) electrons respectively, see Eq. (4.3) for their definition. In addition to the kinetic terms for the two electrons, we also have the two Coulomb potentials representing the attraction between each of the two electrons and the atomic nucleus, as well as the repulsive Coulomb potential between the two electrons (since they are both negatively charged). Note that we have used $Z = 2$ in the Coulomb potential between each of the electrons and the He nucleus.

The main difference between the Schroedinger equations for the hydrogen and for the helium atoms, Eq. (5.1), is that while the former admits an analytical solution, this is not possible for the latter, and therefore we are forced to adopt a number of approximations. The method that we will adopt in these lectures is the so-called *orbital approximation*:

In the orbital approximation, one assumes the wave function of the two-electron system can be expressed as the product of the wave function of the individual electrons, namely one writes

$$\Psi(\vec{r}_1, \vec{r}_2) = \Psi(\vec{r}_1)\Psi(\vec{r}_2) \,. \tag{5.2}$$

This assumption is based on neglecting the repulsive interaction between the two electrons.

Within this approximation, the Schroedinger equation Eq. (5.1) for this system simplifies to:

$$\widehat{H}\Psi = \left(-\frac{\hbar^2}{2m_e}\nabla_1^2 - \frac{2e^2}{4\pi\epsilon_0 r_1} - \frac{\hbar^2}{2m_e}\nabla_2^2 - \frac{2e^2}{4\pi\epsilon_0 r_2} \right) \Psi = E\Psi \,, \tag{5.3}$$

which now can be solved analytically. To see this, note that we can write using Eq. (5.2)

$$\widehat{H}\Psi = \left(\widehat{H}_1 + \widehat{H}_2 \right) \Psi(\vec{r}_1, \vec{r}_2) = \left(\widehat{H}_1 + \widehat{H}_2 \right) \Psi_1(\vec{r}_1)\Psi_2(\vec{r}_2) \,, \tag{5.4}$$

where $\widehat{H}_1$ is the Hamiltonian of a hydrogen-like atom with $Z = 2$ for the first electron and $\widehat{H}_2$ the same for the second electron. In this situation, it is clear that then $\Psi_1(\vec{r}_1)$ will be the wave functions of the hydrogen atom, since Eq. (5.4) can be written in that case as

$$\widehat{H}\Psi = \left(\widehat{H}_1 \Psi_1(\vec{r}_1) \right) \Psi_2(r_2) + \Psi_1(\vec{r}_1) \left(\widehat{H}_2 \Psi_2(r_2) \right) = (E_1 + E_2)\, \Psi_1(\vec{r}_1)\Psi_2(r_2) \,. \tag{5.5}$$

Therefore, we have shown that if we neglect the repulsive Coulomb interaction between the two electrons of the helium atom, we can derive an analytic solution for the wave function in terms of the wave functions of the hydrogen atom, with the total energy of the system being given by

$$E = E_1 + E_2 \,, \tag{5.6}$$

namely the sum of the individual energies of the two electrons. However, one still needs to account in some form for the effect that each electron has on the other, and this is achieved by means of a concept known as the *effective electric charge*.

The effective electric charge $Z_{\rm eff}$. The basic assumption that underlies the orbital approximation that we have used in Eq. (5.4) is that the motion of the electrons is *uncorrelated* between them, in other words, 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 basic idea is that 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.

With this motivation, one 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.

For the helium atom, 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. Putting everything together, combining the orbital approach for the Helium wave function in Eq. (5.2) with the concept of effective charge of the atomic nucleus, we can construct our approximation for the electronic wave functions of the helium atom. For instance, if the two electrons occupy the $1s$ orbital, then the wave function of the helium atom will be given by

$$\begin{aligned} \Psi(\vec{r}_1, \vec{r}_2) &= \Psi_{1s}(\vec{r}_1) \times \Psi_{1s}(\vec{r}_2) = \frac{Z_{\rm eff}^{3/2}}{(\pi a_0^3)^{1/2}} e^{-Z_{\rm eff} r_1/a_0} \times \frac{Z_{\rm eff}^{3/2}}{(\pi a_0^3)^{1/2}} e^{-Z_{\rm eff} r_2/a_0} \\ &= \frac{Z_{\rm eff}^3}{\pi a_0^3} e^{-Z_{\rm eff}(r_1+r_2)/a_0} \,, \end{aligned} \tag{5.7}$$

with $Z_{\rm eff} = 1.62$. 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.

Before we can discuss how to generalise these ideas to more complex atoms such as Li ($Z = 3$, composed by three electrons), we need to discuss in some detail one of the foundational concepts of quantum mechanics: *spin*. The reason is that we have stated above as an obvious fact that two electrons can *occupy the same quantum orbital*, but actually this is not possible unless we introduce a new property, spin, that does not have an analog in classical theory.

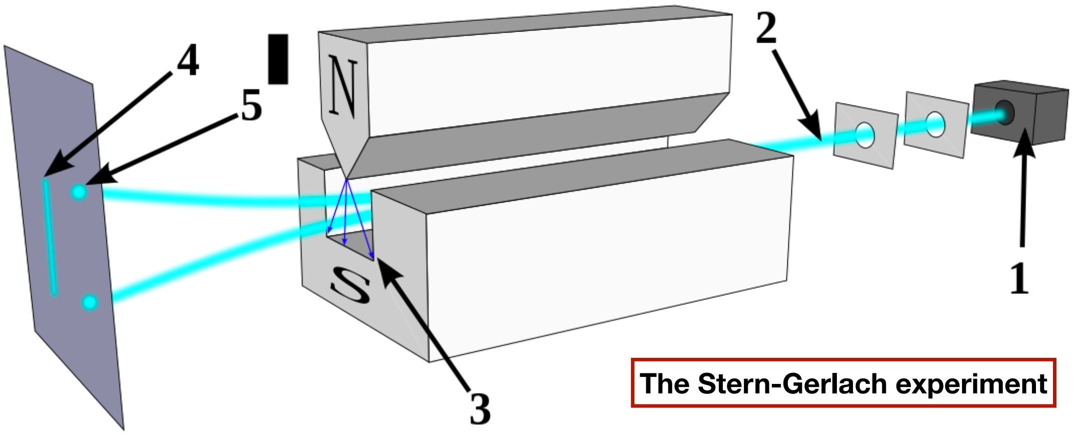


Figure 5.1: Schematic representation of the Stern-Gerlach experiment. A beam of atoms passes through a magnetic field perpendicular to their direction of motion. The beam is then deflected in two possible ways, indicating that the silver atoms are characterised by some form of internal angular momentum that can take only two values of same size but opposite sign, rather than a continuum of values as one could expect from classical theory.

5.2 Spin in quantum mechanics

Many introductions to quantum mechanics trace back the discovery of quantum property of *spin* to the *Stern-Gerlach experiment*. In this experiment, a bunch of silver atoms was beamed through a magnetic field $\vec{B}$ perpendicular to their direction of motion, as represented schematically in Fig. 5.1, therefore being deflected before impacting on a detector screen. Classically, the atoms could be deflected in principle in any direction, since the force that the magnetic field exerts on the silver atoms. The reason is that the torque acting on the magnetic moment $\vec{\mu}$ of the silver atoms in the presence of a external magnetic field $\vec{B}$ is

$$\vec{\tau} = \vec{\mu} \times \vec{B} \,. \tag{5.8}$$

where $\vec{\mu} \propto \vec{J}$ proportional to the total angular momentum of the silver atoms. Classically, the angular momentum of the silver atoms can take any value so one expects a continuum values for the torque and thus for the direction of deflection.

However, what the experiment found was that the silver atoms where deflected only in two directions (as opposed to a continuum), symmetrically with respect their original direction of motion. This result implies the silver atoms are characterised by some form of *angular momentum* that can take only *two values of same size but opposite sign*.

It was later understood that this angular momentum associated to silver atoms arises from the *intrinsic magnetic moment* of the electron in the outer atomic shell. We denote by *spin* this intrinsic angular momentum of the electron. We can naively think of spin as the angular momentum arising from the rotation of the electron along some axis, but this classical picture cannot really be true in the case of point particles and quantum mechanics. It is better to think of the spin as *another type of charge of the electron*, not unlike the electric charge. Therefore one can state that in quantum mechanics:

- Particles such as electrons and protons have associated a *new quantum number* denoted by spin s, which corresponds to a form of intrinsic angular momentum.
- For a particle characterised by a spin quantum number s, its spin can be aligned in $2s - 1$ different

directions (in the same way as for the orbital angular momentum), given by $-s, -s+1, \ldots, s-1, s$.

- Therefore, since an electron has spin quantum number $s = \hbar/2$, the component of the spin vector in the z direction can be either $+\hbar/2$ and $-\hbar/2$. As we will see, in the same way as that orbital angular momentum had associated the (l, m_l) pair of quantum numbers, also for spin we have the (s, m_s) combination, where for an electron $m_s = \pm 1/2$, since $s_z = \pm\hbar/2$.

This discussion about the properties of the electron spin will provide useful when constructing the electronic structure of multi-electron atoms: in addition to the the three quantum numbers (n, l, m_l) that define the wave functions of the electron atom, we will need now to specific also the value of spin quantum number m_s of the electron (that is, the projection of its spin along the z direction) in order to uniquely identify a given electronic orbital. Indeed, the existence of the electron spin is an important factor to determine the electronic orbital structure of multi-electron atoms.

Classical interpretation.(*) We can show how the naive interpretation of *spin* as arising from the internal rotation of an electron understood as a small sphere is not consistent. For this, one should recall that in classical mechanics objects spinning around a given rotation axis can have associated a conserved quantity called *angular momentum*, which is the analog of linear momentum in the case of rotational motion. For instance, in the case of a sphere rotating around an axis that goes through its center, we have that its angular momentum L will be given by

$$L = \left(\frac{2}{5} M R^2\right) \omega \,, \tag{5.9}$$

where M is its mass, R its radius, and ω the angular velocity of the sphere's rotation. Angular momentum is an important quantity since under some circumstances, specifically in a configuration with rotational symmetry, it will be *conserved*, just like linear momentum of a system is conserved in the absence of external forces.

From an intuitive point of view, we can view spin as the intrinsic angular momentum of electrons. To illustrate this analogy between classical and quantum versions of angular momentum, in Fig. 5.2 we show how in classical theory, objects rotating round an axis have associated a quantity known as angular momentum L, and that in quantum theory, elementary particles have associated a new type of intrinsic angular momentum known as spin. In the case of electrons, their total spin is $\hbar/2$, and therefore electrons exist in states with positive spin with respect a given axis, say the z axis, $L_z = +\hbar/2$ and in states with negative spin with respect to the same axis, $L_z = -\hbar/2$. Note that the superposition principle of quantum mechanics implies that an electron can also exist in a superposition of *up* and *down* spin states, for example systems with the wave function

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle + |\downarrow\rangle\right) \,. \tag{5.10}$$

are also physically allowed.

One should be careful however to not push this analogy between classical angular momentum and spin too much, since the latter is an intrinsic property of quantum theory with no classical counterpart. To begin with, classical angular momentum can take any value, while in the quantum case is quantised. Moreover, a given classical system will have associated a unique angular momentum, while quantum system can exists as superpositions of states with different angular momentum such as Eq. (5.10). And there is an easy way to show that thinking of the electron spin in the classical sense quickly leads us to contradictions.

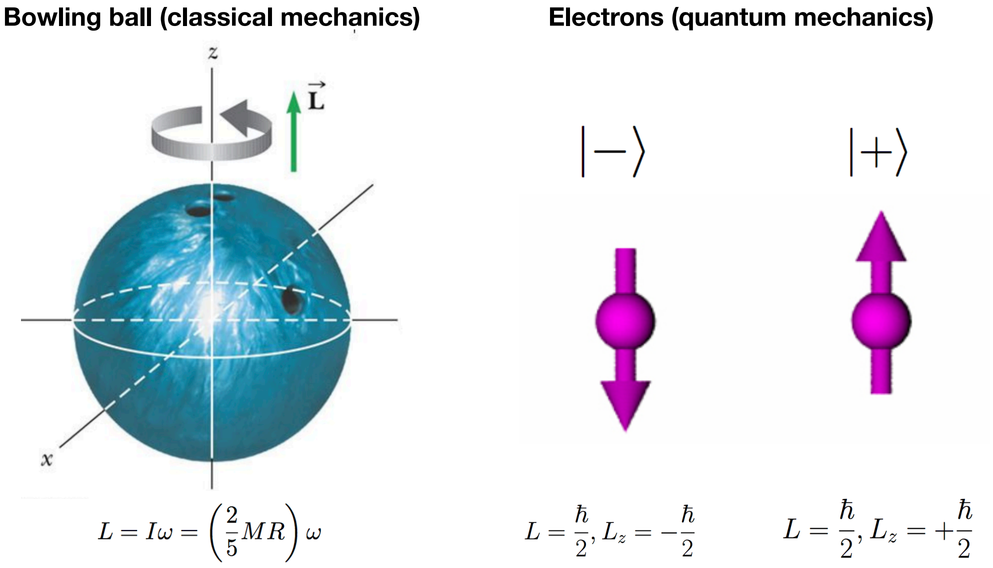


Figure 5.2: Left: in classical theory, objects rotating round an axis have associated a quantity known as angular momentum L. Right: in quantum theory, elementary particles have associated a new type of intrinsic angular momentum known as spin. The values that spin can take for a given particle are quantized. For instance, in the case of electrons, their total spin is $\hbar/2$, and therefore electrons exist in states with positive spin around a given axis (say the z axis), $L_z = +\hbar/2$ and in states with negative spin around the same axis, $L_z = -\hbar/2$.

It is easy to see that interpreting the spin of the electron from a classical point of view, that is, arising from the angular momentum of a tiny rotating sphere, is meaningless. To see this, try to equate the expression for the angular momentum of a rotating sphere, Eq. (5.9), with the value of the spin of an electron,

$$\frac{2}{5}MR\omega = \frac{\hbar}{2}\,, \tag{5.11}$$

and using in this expression the current upper bound on the electron radius, $R < 10^{-18}$ m, and its mass, $M = 9.11 \times 10^{-31}$ kg, we can determine its angular velocity,

$$\omega = \frac{5}{4}\frac{\hbar}{MR^2} = 9.1 \times 10^{32}\,\mathrm{rad/s} \tag{5.12}$$

which corresponds to a velocity of the electron's surface of $v = \omega R \sim 10^{15}$ m/s, more than a million times larger than the speed of light c, and thus inconsistent with special relativity.

The vector model of spin. As illustrated in Fig. 5.3, one can represent the spin of a $s = \hbar/2$ particle as a vector $\vec{s}$ restricted to a cone along the z direction. This model is defined by the following properties:

- For a particle with spin s, then m_s can take $2s+1$ values, which means 2 values for an electron (since $s = 1/2$ in units of $\hbar$).
- The projection of $\vec{s}$ along the z axis is always either $s_z = +\hbar/2$ or $s_z = -\hbar/2$, depending on whether the corresponding *spin quantum number* m_s is $m_s = 1/2$ or $m_s = -1/2$.
- The size of this vector is $|\vec{s}| = \sqrt{s(s+1)}\hbar$, which corresponds to the total magnitude of spin angular momentum of the electron. For a spin 1/2 particle this is $|\vec{s}| = \sqrt{3/4}\hbar$.

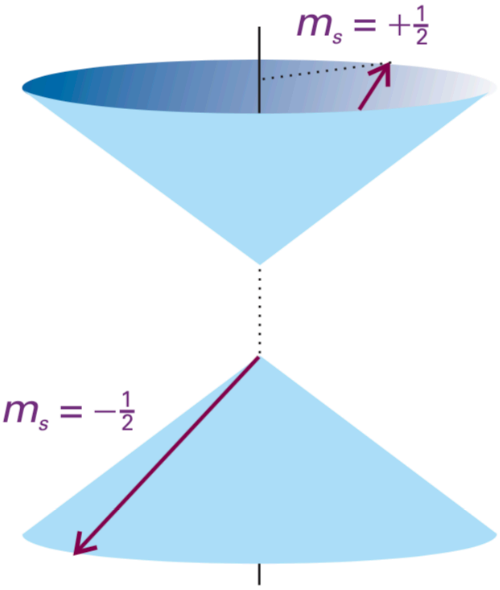


Figure 5.3: One can represent the spin of a $s = \hbar/2$ particle as a vector $\vec{s}$ sitting in a cone along the z direction. The projection of $\vec{s}$ along the z axis is always either $s_z = +\hbar/2$ or $s_z = -\hbar/2$, depending on whether $m_s = 1/2$ or $m_s = -1/2$. The magnitude of this vector is $|\vec{s}| = \sqrt{s(s+1)}\hbar$, which corresponds to the total spin angular momentum of the electron. Note that the projections along the x and y axis are not defined, and would be determined at random upon the corresponding measurement: even if m_s is fixed, the spin vector $\vec{s}$ can still take any position in the corresponding cone.

(*) This property can be understood as follows. Any measurement of one of the three components of the spin vector will always return either $+\hbar/2$ or $-\hbar/2$. Therefore, if we add them together, the total magnitude of the spin vector can be obtained from using Pythagoras's theorem, which gives:

$$|\vec{s}| = \left((\hbar/2)^2 + (\hbar/2)^2 + (\hbar/2)^2\right)^{1/2} = \left(3 \times (\hbar/2)^2\right) = \sqrt{3/4}\hbar\,, \tag{5.13}$$

as we wanted to demonstrate.

- Note that in this model the projections along the x and y axis are not defined, and would be determined at random upon the corresponding measurement: even if m_s is fixed, the spin vector $\vec{s}$ can still take any position in the corresponding cone.

Visualizing spin as in Fig. 5.3 is not completely rigorous, but will prove to be rather useful in the following.

5.3 Fermions and bosons: spin and statistics

The spin of a particle determines several of its most important properties. As mentioned above, in terms of their spin, all particles can be divided into *two classes*:

- *Fermions*: particles with *half-integer spin*: $\hbar/2$, $3\hbar/2$, ...
- *Bosons*: particles with *integer spin*: 0, $\hbar$, ...

Concerning the particles that will be relevant for this course, we have that both electrons, protons, and neutrons are *fermions* with spin $s = \hbar/2$. On the other hand, the total spin of an atom will be determined by the vectorial sum of the spin of its component electrons, leading to the fact that atoms can be either bosons (integer spin) or fermions (half-integer spin) depending on the specific electron configuration. In most cases, the configurations of the outermost electronic orbitals are the ones that determine the overall value of the spin of an atom.

The main difference between fermions and bosons is related to the fact of whether or not *two identical particles* can *occupy the same quantum state*. To illustrate this property, let us consider a simple system composed by two *identical particles*, say two electrons, one at position x_1 and the other at position x_2. For this system composed by two particles, its wave function can be separated in terms of the individual wave functions:

$$\Psi_{\rm tot}(x_1, x_2) = \Psi_1(x_1)\Psi_2(x_2)\,. \tag{5.14}$$

Now let us exchange the positions of the two particles $x_1 \to x_2$ and $x_2 \to x_1$. The new wave function of the system will be

$$\widetilde{\Psi}_{\rm tot}(x_1, x_2) = \Psi_1(x_2)\Psi_2(x_1)\,. \tag{5.15}$$

However, since the particles are identical, all physical information that I can extract from the system should be the same as before exchanging their positions. This requirement implies square of the wave function (which determines the probability of finding the system in a given quantum state) should be unchanged, that is,

$$|\Psi_{\rm tot}(x_1, x_2)|^2 = |\widetilde{\Psi}_{\rm tot}(x_1, x_2)|^2\,, \tag{5.16}$$

which implies that the wave function itself can only vary up to a complex phase,

$$\widetilde{\Psi}_{\rm tot}(x_1, x_2) = e^{i\phi}\Psi_{\rm tot}(x_1, x_2)\,. \tag{5.17}$$

Moreover, if we exchange back again the positions of the two particles, one is back to the starting configuration Eq. (5.14), and then the wave function should be the same as before. This implies that

$$e^{2i\phi} = 1 \qquad \to \qquad e^{i\phi} = \pm 1\,. \tag{5.18}$$

This exercise tells us that identical quantum particles can behave only on two ways if they exchange positions (or more in general, if they exchange quantum states) within a system:

- either the wave function of the system remains unchanged,
- or else it gets a minus sign.

We denote the first class of particles as *bosons*. Therefore, the wave function of a system composed by bosons is *symmetric* upon the exchange of two identical particles:

$$\widetilde{\Psi}_{\rm bosons}(x_1, x_2) = \Psi_{\rm bosons}(x_1, x_2)\,, \tag{5.19}$$

while we denote the second class of particles as *fermions*, for which under the same operation the wave function is instead *antisymmetric*:

$$\widetilde{\Psi}_{\rm fermions}(x_1, x_2) = -\Psi_{\rm fermions}(x_1, x_2)\,. \tag{5.20}$$

Note that there is no other option: upon the interchange of two identical particles, the wave function of the system is either *symmetric* (for bosons) or *antisymmetric* (for fermions).

The Pauli exclusion principle. The most important implication of this difference between fermions and bosons is the so-called *Pauli exclusion principle*. Its derivation follows directly from the previous discussion. consider the wave function of a system composed by two identical fermions that occupy the same position (again, more in general, that they occupy the same quantum state). We know that the total wave function will be antisymmetric upon the exchange of these two fermions, and therefore one finds that

$$\psi_{\rm fermions}(x_1, x_1) = -\psi_{\rm fermions}(x_1, x_1) = 0\,, \tag{5.21}$$

implying that:

Two fermions cannot occupy simultaneously the same quantum state.

This is the Pauli exclusion principle: a given quantum state, characterised by a set of quantum numbers, can only be occupied by *one fermion at a time*. Since electrons are fermions, this principle has important consequences for the electronic structure of multi-electron atoms. On the other hand, bosons can occupy the same quantum state without no limitations, since they have a symmetric wave function upon the exchange of two identical particles

$$\psi_{\rm bosons}(x_1, x_1) = +\psi_{\rm bosons}(x_1, x_1) \neq 0\,. \tag{5.22}$$

In the context of the orbital structure of multi-electron atoms, the Pauli exclusion principle tells us that two electrons cannot occupy the same quantum state. This implies that a given orbital, defined by the quantum numbers (n, l, m_l), can contain *at most two electrons*, one with spin quantum number $m_s = +1/2$ and the other with spin quantum number $m_s = -1/2$. It follows that if two electrons occupy the same orbital then they must have spins pointing in opposite direction.

We define as *paired electrons* two electrons with opposite spin that occupy the same quantum state. *Unpaired electrons* are those with spins pointing in the same direction, and that therefore due to the Pauli exclusion principle cannot occupy the same electronic orbital.

5.4 Spin and two-electron systems

From the Pauli exclusion principle, it follows that each electronic orbital can at most be occupied by two electrons at the same time. Therefore, the two-electron system plays an important role in the study of multi-electron atoms. Here we now discuss the various ways in which we can combine two electrons taking into account the contributions from their spin.

In the following, we indicate by $|+\rangle$ the quantum state corresponding to an electron with spin pointing in the positive direction along the z axis (thus $m_s = +1/2$), and by $|-\rangle$ the same electron pointing in the negative direction (thus $m_s = -1.2$). Within the vector model of the electron spin, represented schematically in Fig. 5.3, any state for which the spin vector lies in the upper cone will have associated a $|+\rangle$ while if we vector lies in the lower cone they will have associated a $|-\rangle$. With this notation, we can then construct the following combinations for the two-electron system

- Two electrons pointing upwards: $|+\rangle_1|+\rangle_2 \equiv |{+}{+}\rangle$.
- Two electrons pointing downwards: $|-\rangle_1|-\rangle_2 \equiv |{-}{-}\rangle$.

- One electron upwards, another downwards: $|+\rangle_1|-\rangle_2 \equiv |+-\rangle$ or $|-\rangle_1|+\rangle_2 \equiv |-+\rangle$.

Recall that quantum particles such as electrons are *indistinguishable*: we cannot really label one electron as "1" and the other as "2", since it is not possible to tell them apart. Therefore the spin configurations $|+-\rangle$ and $|-+\rangle$ are not physically meaningful, since they assume that we can tell one electron apart from the other. We should instead construct linear combinations of these which have a more transparent physical interpretation. With this motivation we define:

$$\begin{aligned} \sigma_+ &\equiv \frac{1}{\sqrt{2}}\left(|+-\rangle + |-+\rangle\right) , && (5.23) \\ \sigma_- &\equiv \frac{1}{\sqrt{2}}\left(|+-\rangle - |-+\rangle\right) . && (5.24) \end{aligned}$$

The rationale between choosing these combinations is the following. The first combination, σ_+, is *symmetric* under the exchange of the two particles of the system. The second combination, σ_-, is instead *antisymmetric* with respect to such exchange. Note that in these combinations specifying which ones is "electron 1" and "electron 2" is irrelevant: the same property holds irrespective of how one does the assignment. The overall factor $1/\sqrt{2}$ has been added to ensure the normalisation of the two-electron wave function.

Summarizing, the four possible spin configurations that can be physically constructed for a two electron system are the following:

$$\begin{aligned} & |++\rangle \\ & |--\rangle \\ \sigma_+ &= \frac{1}{\sqrt{2}}\left(|+-\rangle + |-+\rangle\right) , && (5.25) \\ \sigma_- &= \frac{1}{\sqrt{2}}\left(|+-\rangle - |-+\rangle\right) . \end{aligned}$$

One should emphasize that the fact that we can construct these four combinations does not mean that all of them can be realised in an specific physical system. Indeed, out of these four configurations, only σ_- is antisymmetric under the exchange of the two electrons as required by the Pauli exclusion principle. Therefore, we obtain the following important result:

> For two electrons that occupy the same atomic orbital Ψ_{n,l,m_l}, only one specific combination of radial and spin wave functions can be constructed, $\Psi_{n,l,m_l}(\vec{r_1})\Psi_{n,l,m_l}(\vec{r_2})\sigma_-$ in s way that satisfies the Pauli exclusion principle. Indeed, we see that despite the fact that the spatial part of the wave function, $\Psi_{n,l,m_l}(\vec{r_1})\Psi_{n,l,m_l}(\vec{r_2})$, is symmetric under the exchange of the two electrons, the total wave function is antisymmetric as required for a system of fermions since $\sigma_- \to -\sigma_-$ upon the exchange of the two electrons. On the other hand, for electrons that occupy different orbitals this restriction does not apply, and for example $\Psi_{n_1,l,m_l}(\vec{r_1})\Psi_{n_2,l,m_l}(\vec{r_2})\sigma_+$ with $n_1 \neq n_2$ is an allowed wave function.

Depending on the relative orientation of the spins of a two electron system, we can divide them as:

- A system of *two paired electrons* is called a *singlet state*, it has total spin $S = 0$, and the spin component of the wave function is σ_-, Eq. (5.24). Note we use upper case notation, S, to denote the total spin of the two electron system.

- A system of *two unpaired electrons* is called a *triplet state*. Such triplet state has a total spin of $S = 1$.

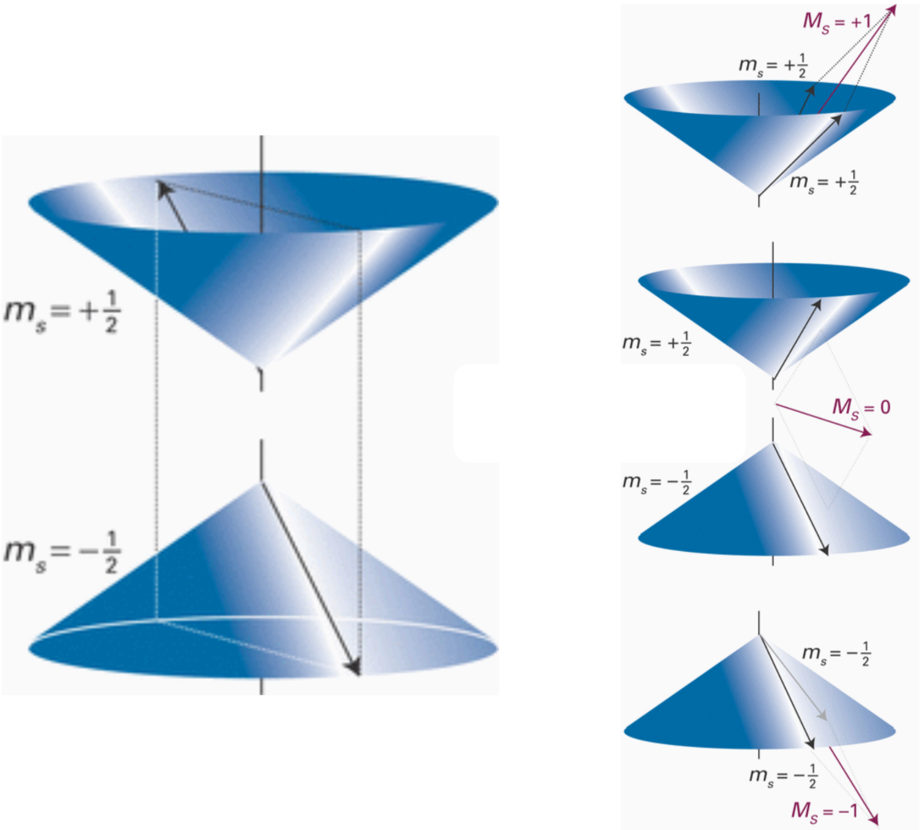


Figure 5.4: Schematic representations of the different possible spin configurations in a two-electron system in terms of the vector model for the spin. In the left plot, we show the *singlet state* of paired electrons, where the total spin of the system is $S = 0$. The first electron has $m_s = +1/2$ (thus lying in the upper cone) and the second electron has $m_s = -1/2$ (thus lying in the lower cone). Note that also their spin components in the x and y directions are the opposite. In the right plot of the figure, we show the three possible *triplet configurations*, each with total spin $S = 1$. They differ in the values of the spin projections along the z axis, s_z. From top to bottom we have the configurations with $m_s = +1$, $m_s = 0$, and $m_s = -1$ respectively. Note how the total value of m_s of the two electron system is the sum of that of the individual electrons.

> The spin configurations of the triplet states $|{+}{+}\rangle, \sigma_+$, and $|{-}{-}\rangle$ have associated a value of the total spin component in the z direction of $S_Z = +1, 0, -1$, respectively.

The way how the total spin S and its z-projection S_z are constructed from the spins of the two individual electrons for the various configurations is represented schematically in Fig. 5.4. There we use the same representation of $\vec{s}$ as a vector restricted to lie in the surface of a cone along the z direction as was used in Fig. 5.3 and in the corresponding discussion. In the left plot, we show the *singlet state* of paired electrons, where the total spin of the system is $S = 0$. The first electron has $m_s = +1/2$ (thus lying in the upper cone) and the second electron has $m_s = -1/2$ (thus lying in the lower cone). Note that also their spin components in the x and y directions are the opposite. In the right plot of the figure, we show the three possible *triplet configurations*, each with total spin $S = 1$. They differ in the values of the spin projections along the z axis, s_z. From top to bottom we have the configurations with $M_S = +1$, $M_S = 0$, and $M_S = -1$ respectively. Note how the total value of M_s of the two electron system is the sum of that of the individual electrons. For example, for the upper configuration, we have $M_S = m_s + m_s = 1/2 + 1/2 = +1$.

5.5 Multi-electron atoms and the building-up principle

The above discussion on the electronic spin completes our study of the helium atom, the next-to-simplest chemical element with $Z = 2$. We now to discuss how the electrons distribute themselves among the electronic orbitals of more complex elements, with $Z \geq 3$. In this discussion, we will keep using the *orbital approximation*, which assumes that the electronic orbitals of multi-electron atoms can be written as a combination of the orbitals of hydrogen-like atoms with suitable modifications such as that the the *effective electric charge*.

Shielding in multi-electron atoms. We have already introduced the concept of *effective electric charge* $Z_{\rm eff}$ in multi-electron atoms in Sect. 5.1. The idea is that in an multi-electron atoms with Z protons, the effective charge that one electron experiences is $Z_{\rm eff} < Z$ due to the combined effect of the smearing from the other electrons. This same idea can be expressed in terms of *shielding*: namely, each of the electrons in an atom partially shields the positive electric charge of the atomic nucleus for all other electrons. We therefore define the shielding constant σ as follows:

$$Z_{\rm eff} = Z - \sigma \,, \tag{5.26}$$

where the specific value of σ is different element by element and also varies within orbitals.

To illustrate this point, we know already that s-type orbitals have a higher likelihood to be found close to the nucleus at $r = 0$ (see Fig. 4.4 and the corresponding discussion), 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

$$\sigma_{3,s} \leq \sigma_{3,p} \leq \sigma_{3,l} \,. \tag{5.27}$$

- 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

$$\sigma_{1,s} \leq \sigma_{2,s} \leq \sigma_{3,s} \,. \tag{5.28}$$

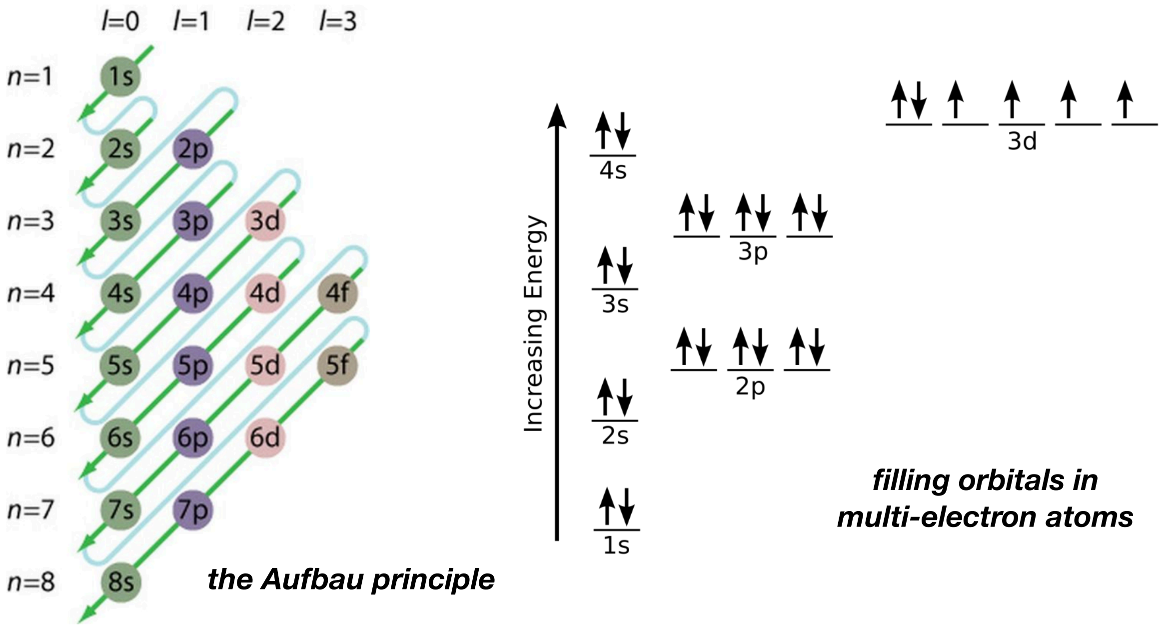


Figure 5.5: Left panel: the building up principle determines the ordering of the filling of electronic orbitals in multi-electron atoms. This order is determined by the values of the quantum numbers n and l, and ensures that the free orbitals with lower energies are always filled first. Right panel: example of the filling of atomic orbitals for iron (Fe), with $Z = 26$. Note that within a given sub-shell, electrons first occupy empty orbitals before starting to pair within the same orbital. For a subshell with angular quantum number l, there are $2l + 1$ orbitals that can be occupied (each by at most two electrons).

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

The building up (aufbau) principle. In order to fill the electron orbitals of a given multi-electron atom, one must follow the so-called *building up* or *aufbau principle*, which determines the filling order of the orbitals as a function of their energy. The building up principle states that:

> In a multi-electron atom, electrons will always fill the free orbital which has the lowest energy. Once this orbital is completely filled, electrons will start to occupy the next free orbital with the lowest energy, and so on.

We should be careful since this is *not a fundamental principle* but rather a rule of thumb, and the order of the filling of electronic orbitals in some atoms differs from that indicated by the aufbau principle. The specific order in which electronic shells should be filled according to the building up principle is represented in Fig. 5.5. When expressing orbitals in terms of their quantum numbers (n, l), we see that the filling order will be:

$$1s \to 2s \to 2p \to 3s \to 3p \to 4s \to 3d \to 4p \to \ldots . \tag{5.29}$$

Note that depending on the value of l, some orbitals with $n_2 > n_1$ will be filled before the n_1 shell is complete. For example, the $4s$ orbital should be filled before the $3d$ orbital according to the building up principle. Recall also that the number of electrons that fill a given subshell depends on the value of l: for $s-$type orbitals we can fit two electrons, for p-type orbitals we can fit six electrons and so on. Orbitals with angular quantum number l can accommodate up to $4l + 2$ electrons once we take their spin into account.

It is instructive to show how the building up principle is applied for different elements. As we will show, the results of this filling of electronic shells also provides useful information to understand the chemical behaviour of the different elements.

- The electronic structure of lithium (Li, $Z = 3$) is $1s^2 2s^1$.
- The electronic structure of beryllium (Be, $Z = 4$) is $1s^2 2s^2$.
- The electronic structure of nitrogen (N, $Z = 7$) is $1s^2 2s^2 2p^3$.
- The electronic structure of neon (Ne, $Z = 10$) is $1s^2 2s^2 2p^6$.
- The electronic structure of phosphorus (P, $Z = 15$) is $1s^2 2s^2 2p^6 3s^2 3p^3 = [\mathrm{Ne}] 3s^2 3p^3$.
- The electronic structure of sulphur (S, $Z = 16$) is $1s^2 2s^2 2p^6 3s^2 3p^4 = [\mathrm{Ne}] 3s^2 3p^4$.

where [Ne] represents the electronic structure of Ne with its $n = 1$ and $n = 2$ filled shells. We observe that both He and Ne have all shells filled ($n = 1$ for He, and $n = 1$ and $n = 2$ for Ne). This explains why these two elements, which belong to the family of *noble gases*, have a very low chemical reactivity: their electronic structure is complete and thus very stable.

In Fig. 5.5 we show as an example the filling of atomic orbitals for iron (Fe), with $Z = 26$. Note that within a given sub-shell, electrons first occupy empty orbitals before starting to pair within the same orbital - this is called the *Hund rule* and will be discussed below. This diagram illustrates how the electronic orbitals are filled in order of increasing energy, starting from the ground state ($1s$), then filling the first excited state ($2s$), and then subsequently the next free orbital with the lowest energy.

Hund's rule. The *Hund's rule* for the filling of atomic orbitals states that:

> When filling orbitals, electrons first occupy *the empty degenerate orbitals* before they occupy an orbital where another electron is already present.

This effect is illustrated in Fig. 5.5 for the case of the Fe atom ($Z = 26$). Note that, following Pauli's exclusion principle, two electrons that occupy the same orbital must have their spins *paired* (that is, pointing in opposite directions). A direct consequence of this rule is that the *ground state of an atom* with a given value of Z is the one that contains *the maximum number of unpaired electrons*, since this is the one that has associated the smallest total energy. An explanation of this rule is provided by the concept of *spin correlations*: electrons with unpaired spin (pointing in the same direction) stay away from each other, since this reduces their mutual repulsion and thus decreases the total energy of the system.

Let us see what are some of the practical consequences of the application of Hund's rule. According to this rule, the electrons that occupy the p-type orbitals in C, N, and O atoms should be distributed as follows:

Element	Atomic structure	Orbital filling
C ($Z = 6$)	$1s^2 2s^2 2p^2$	$1s^2 2s^2 2p_x^1 2p_y^1$
N ($Z = 7$)	$1s^2 2s^2 2p^3$	$1s^2 2s^2 2p_x^1 2p_y^1 2p_z^1$
O ($Z = 8$)	$1s^2 2s^2 2p^4$	$1s^2 2s^2 2p_x^2 2p_y^1 2p_z^1$

So we see that, to begin with, degenerate orbitals within a given subshell (in this case the $2p_x$, $2p_y$, and $2p_z$ orbitals) are occupied by only one electron. Only when these three degenerate orbitals contain one electron each, they will start to accommodate the second electron in a way that their spins are paired.

Qualitative explanation for Hund's rule (*). From the qualitative point of view, Hunds's rule can be understood as follows. We start by denoting $\Psi_a(r_1)$ and $\Psi_b(r_2)$ the wave functions for two electrons corresponding to two different orbitals a and b. There exist two possibilities for the spatial part of the combined wave function of the two-electron system, namely

$$\Psi_\pm = \frac{1}{\sqrt{2}} \left[\Psi_a(r_1)\Psi_b(r_2) \pm \Psi_b(r_1)\Psi_a(r_2) \right] . \tag{5.30}$$

The total wave function of the two-electron system, including the spin part, must be antisymmetric upon the exchange of the two electrons, due to the Pauli exclusion principle. Therefore we have in total four possibilities for the total wave function of the two-electron system (which recall occupies different orbitals), including also the contribution from the spin degrees of freedom:

$$\Psi_+\sigma_- , \quad \Psi_-\sigma_+ \quad \Psi_-|++\rangle , \quad \Psi_-|--\rangle . \tag{5.31}$$

Now let us consider what happens once the two electrons occupy the same orbital, that is, when $r_1 = r_2$. We see that while Ψ_- vanishes, this is not the case for Ψ_+. Therefore, if electrons occupy the Ψ_- quantum state, they have a vanishing probability of being found in the same position. This means that the quantum state Ψ_- is characterized by an *effectively smaller attraction* between the two electrons, and thus with a smaller binding energy than the state Ψ_+. Since the spatial wave function Ψ_+ has associated the singlet spin state σ_-, we have demonstrated that unpaired electrons in degenerate orbitals tend to appear in triplet spin configurations rather than in the singlet ones, and that provided a formal explanation for Hund's rule.

Summary

To summarize, some important concepts that we have learned in this lecture about

- Within the orbital approximation, the electronic structure of multi-electron atoms can be understood in therms of that of the hydrogen atoms with suitable modifications, such as the effective electric charge.
- Spin is a fundamental property of particles such as electrons and protons that does not have classical analog. It can be understood by means of the vector model of spin.
- The Pauli exclusion principle indicates that two fermions cannot occupy the same quantum state, and has important implication to understand the electronic structure of multi-electron atoms.
- The building-up principle and Hunds's rule determine the order in this the electronic shells will be filled in multi-electron atoms.

This lecture completes our study of the electronic structure of individual chemical elements. We have started from the simplest possible element, hydrogen, and then studied the electronic structure of multi-electron atoms first with Helium and then for atoms with $Z > 3$. In the second part of the course, we will exploit and generalize these concepts to the study of the electronic structure of *molecules*, bound states composed by two or more atoms.

Further reading

Further reading material about the topics covered in this lecture can be found in the following sections of the course textbook (*Physical Chemistry*, Atkins and De Paula, 10th edition):

- **9B.1**: *The orbital approximation.*
- **9B.2**: *The building-up principle.*

van Quantum tot Molecuul

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

6 HC7: Chemical bonds and molecules I

In the previous lectures, we have studied the quantum mechanics of individual atoms, starting from the simpler ones (hydrogen and helium), and then moving towards more complex elements. Now, in the following three lectures, we move to study the implications of quantum mechanics for molecules, that is, for bound states of different atoms. In particular, we will be interested in quantifying the properties of the chemical bonds that keep the elements of a molecule together. Here we will study mostly diatomic molecules (composed by two atoms), but we will also take a brief look at poly-atomic molecules (composed by more than two atoms). The two main frameworks that will be used in this lecture to explain chemical bonding will be *Valence Bond theory* and *Molecular Orbitals theory*.

The **learning goals** of this lecture are:

(a) Becoming familiar with the main features of two different formalisms that explain chemical bonds in molecules: Valence Bond theory and Molecular Orbitals theory.

(b) Understanding the concept of hybrid orbitals, and the role they play in Valence Bond theory.

(c) Constructing molecular orbitals starting from the electronic orbitals of individual atoms.

(d) Identifying the main geometrical features of the molecular orbitals that appear both in Valence Bond theory and Molecular Orbitals theory.

Introduction. The description of chemical bonds is a quantum mechanical problem involving the interactions between at least three particles. Even for the simplest molecule, H_2^+ (composed of two protons and one electron), the Schroedinger equation of the system does not admit an analytical, closed-form solution. Therefore, when modelling the chemical bonds in molecules, we will need to resort to a number of different approximations, not unlike what was done in the case of multi-electron atoms. In the three lectures of the course devoted to chemical bonding in molecules (HC7, HC8, and HC9), we will use various approximations to this problem, each with their own advantages and disadvantages, and being suitable to different situations.

In the left plot of Fig. 6.1 we show the simplest molecule is composed by two atomic nuclei, A and B, which interact with a single electron. For example, in the case of the H_2^+ molecule, the nuclei A and B

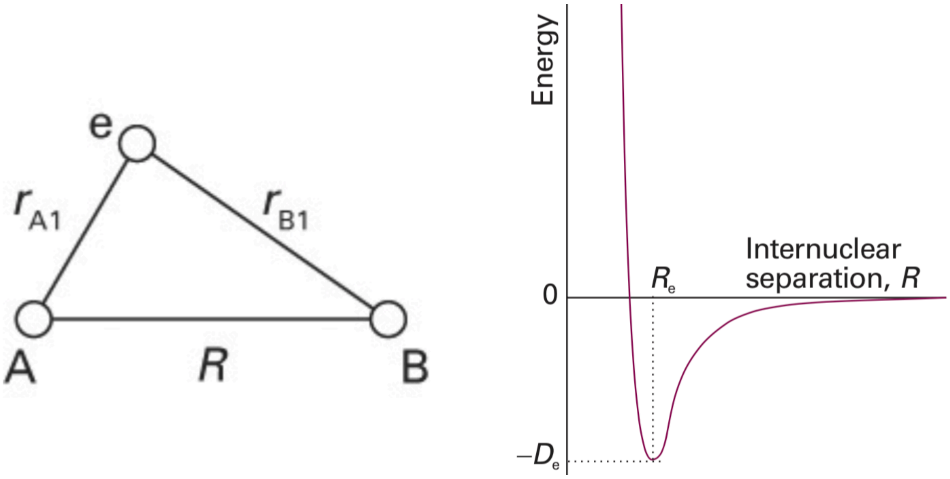


Figure 6.1: Left: the simplest molecule is composed by two atomic nuclei, A and B, which interact with a single electron. For example, in the case of the H_2^+ molecule, the nuclei A and B would two protons. We denote by R the inter-nuclear distance, and r_{A1} (r_{A2}) the distance between the electron and the nucleus A (B). Right: the total energy of the molecule is a function of the inter-nuclear separation R, with R_e being the equilibrium radius which corresponds to the molecular ground state.

would be two protons. In this context, we denote by R the inter-nuclear distance, and r_{A1} (r_{A2}) the distance between the electron and the nucleus A (B).

In the right panel of Fig. 6.1 we show the total energy of the molecule as a function of the inter-nuclear separation R. One can point out the following important features:

- We denote by R_e the equilibrium radius, which corresponds to the molecular ground state (the state with lowest total energy).
- We denote by D_e the value of $-E$ at the classical equilibrium position R_e: this is the *molecular dissociation energy*. If we provide to the molecule an energy $\geq D_e$ (for example with photons) then it will break apart into its constituent atoms A and B.
- For large values of the inter-nuclear separation, $R \to \infty$, there is no molecule at all and thus its energy vanishes. This configuration corresponds to that of the separated atoms A and B.
- For small values of the inter-nuclear separation, $R \to 0$, then the repulsive electrical interaction between the atomic nuclei leads to a fast ($\simeq R^{-1}$) growth of the energy. This means that A and B cannot be too close to each other.

The ultimate goal of applying quantum mechanics to molecules is to determine properties such as the equilibrium radius R_e in Fig. 6.1, which determine its defining features such as its size and its overall geometric configuration. The quantum modeling of molecules that we will adopt in these lectures will rely heavily on the

Born–Oppenheimer approximation: the assumption that the motion of the electrons that compose a molecule can be treated separately from that of the corresponding nuclei.

This approximation is justified by the fact that atomic nuclei are much heavier than electrons (since $m_p \gg m_e$), and therefore, from the point of view of the electrons, the atomic nuclei can be considered to be essentially static. Therefore, one can treat separately the motion of the electrons from the degrees of freedom associated to the atomic nuclei, such as its vibrational and rotational motion. In the language of quantum mechanics, the Born–Oppenheimer approximation can be translated to the statement that the total molecular wave function can be separated into its electronic and nuclear components:

$$\Psi_{\rm total} = \Psi_{\rm electronic} \otimes \Psi_{\rm nuclear} \, , \tag{6.1}$$

each obeying different types of Schroedinger equations that are essentially *decoupled* from each other.

Taking as starting point the Born–Oppenheimer approximation, in this lecture we will present two framework to describe chemical bonding in molecules. The first one is known as the *Valence Bond theory*, and the second as *Molecular Orbital theory*. Both frameworks can be used to describe

- *homonuclear* molecules, composed by the same element, and
- *heteronuclear* molecules, composed by different elements.

Moreover, these two frameworks can be applied both to

- *diatomic* molecules, composed by only two atoms), and
- *poly-atomic* molecules, composed by more than two atoms.

6.1 Chemical bonding in valence Bond theory

The *Valence Bond theory* is one of the existing frameworks aiming to explain chemical bonding in molecules. The basic ingredients of this theory are:

- Each electron can be assigned to their respective (original) atom within a molecule, rather than being delocalised within the whole molecule.
- The chemical bond in a molecule arises from the interactions between two electrons from neighboring atoms with *paired spin* (that is, with opposite spin orientations).
- The constructive interference between the two-electron wave function ensures a non-zero electron probability density in the region between the two atoms.

The fact that there is a finite probability of finding the electrons in the region between the two atoms is then responsible for the creation of the chemical bond. The electrons that are involved in the formation of this chemical bond are usually the electrons occupying the *outermost electronic shells* and are known as the *valence electrons*, hence explaining the name of the Valence Bond theory.

To illustrate the main features of Valence Bond theory, let us consider a simple diatomic molecule of the form of that in Fig. (6.1), composed by two atomic nuclei A and B but now also with two electrons, denoted by 1 and 2. Let us try to construct a physically sensible wave function that can describe this configuration. One possible option is that electron 1 occupies an atomic orbital of A, while electron 2 occupies an atomic orbital of B. In this case the molecular wave function will be

$$\Psi(1,2) = \Psi_A(1)\Psi_B(2) \, . \tag{6.2}$$

However, we cannot really tell apart electron 1 from electron 2 (being indistinguishable particles) so the

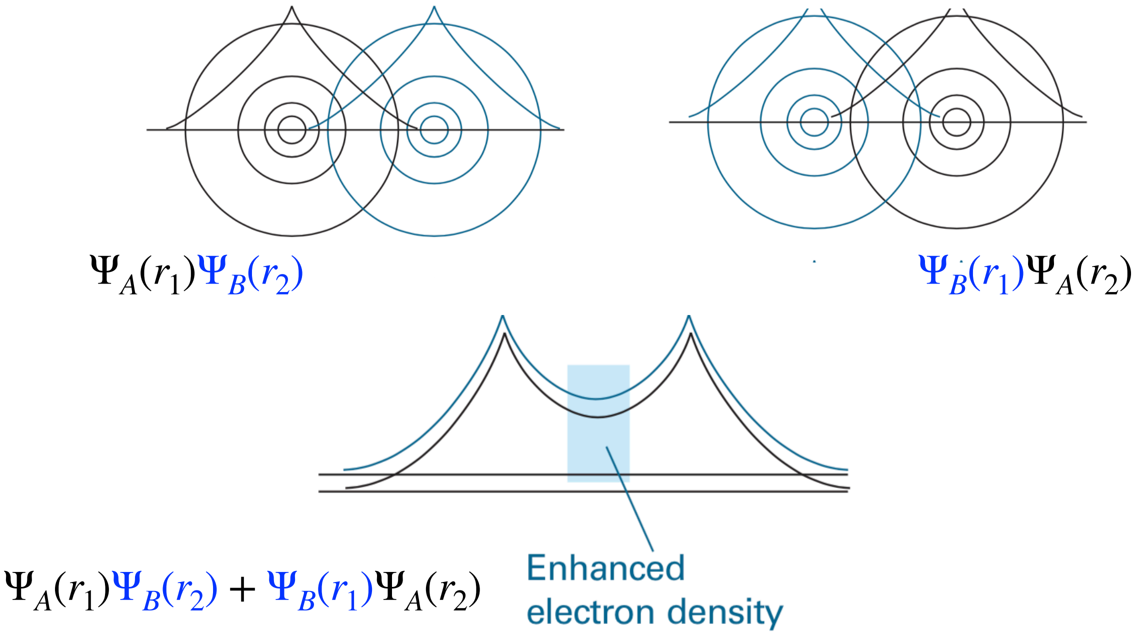


Figure 6.2: Consider a diatomic molecule with two valence electrons involved in the chemical bond, denoted by "1" and "2". The upper left plot represents the probability density if the first electron is on atom A (in black) and the second on atom B (in blue). The upper right plot corresponds to the case when electron "1" is on atom B and electron "2" in atom A. The positive interference of the wave functions in this configuration (bottom plot) leads to a region between the two atoms characterised by an enhanced electron density, resulting in the formation of a molecular bond.

physical wave function of the molecule must read

$$\Psi(1,2) = \left[\Psi_A(1)\Psi_B(2) \pm \Psi_B(1)\Psi_A(2)\right] . \tag{6.3}$$

It can be shown that the combination with a "+" sign has lower energy, and therefore will correspond to the molecular ground state. Since in this case the spatial part of the wave function is symmetric under the exchange of the two electrons, to make sure the Pauli exclusion principle is satisfied, the spin part of the wave function must be antisymmetric. Therefore, it corresponds to the singlet configuration $\sigma_-(1,2)$ in Eq. (5.25) where the two electrons are paired.

Putting together all this information, we obtain that in Valence Bond theory wave function of the two-electron system (that is, the molecular wave function) will be given by:

$$\Psi(1,2) = \left[\Psi_A(1)\Psi_B(2) + \Psi_B(1)\Psi_A(2)\right] \sigma_-(1,2) , \tag{6.4}$$

where the presence of the singlet spin wave function $\sigma_-(1,2)$ guarantees the antisymmetry of the total wave function. This procedure is represented schematically in Fig. 6.2. The upper left plot represents the probability density if the first electron is on atom A (in black) and the second on atom B (in blue), corresponding to Eq. (6.2) The upper right plot corresponds to the case when electron "1" is on atom B and electron "2" in atom A. The positive interference of the wave functions in this configuration (bottom plot) leads to a region between the two atoms characterised by an enhanced electron density, resulting in the formation of a molecular bond, see Eq. (6.4).

From this example, we can summarise how in the framework of Valence Bond theory chemical bonding in molecule arises from. Its main features are:

- The chemical bond arises from the overlap (positive interference) between the wave functions of the valence electrons from each of the atoms that compose a molecule.
- These electrons need to be paired, that is, they must have opposite spin.
- The geometrical properties of the resulting bonds will follow from the geometrical properties of the constituent atomic orbitals.

In the following we discuss the different types of chemical bonds in molecules that can be constructed according to Valence Bond theory.

6.2 Molecular bonds in diatomic molecules

Following this introduction to Valence Bond theory, let us know describe the different types of chemical bonds that can arise in this framework. Recall that, in this theory, chemical bonds arise from the interactions of paired electrons that occupy different orbitals in each of the atoms that form a molecule. Here we will limit ourselves to discuss the case of diatomic molecules. We will show how geometrical properties of the chemical bonds formed in Valence Bond theory follow from the corresponding ones of the constituent atomic orbitals. The most important types of molecular bonds in this context are known as the σ-bonds and the π-bonds.

The σ-bonds. One of the bonds that arises in Valence Bond theory is known as the *σ-bond*, whose main property is that it exhibits *cylindrical symmetry* around the axis that connects the two atoms, which we will denote by z axis for convenience. This bond can arise from configurations that are characterised by cylindrical symmetry such as for example:

- from the overlap of two $1s$ atomic orbitals, or
- from the overlap of two $2p_z$ orbitals (recall that the z axis coincides with the inter-nuclear axis).

The defining conditions to create a σ bond are thus *i)* cylindrical symmetry and *ii)* constructive interference between the electronic wave functions of the constituents atoms. Provided these two conditions can be satisfied, a molecular σ-bond can be formed.

In Fig. 6.3 we represent schematically some of the ways which these cilindrically symmetric σ-bonds can be formed:

- from the overlap between two $1s$ atomic orbitals,
- from the overlap of one $2p_z$ and $1s$ orbital, and
- from the overlap of two $2p_z$ orbitals with paired spins.

Note how in all these cases there is constructive interference between the wave functions of the valence electrons that are involved in the formation of this chemical bond.

The π-bonds. Another important type of molecular orbital that arises in Valence Bond theory are the so-called *π-bond*. These chemical bonds arise from the pairing between p-type electron orbitals which are *perpendicular* to the z axis that connects the two atoms. This condition implies that that π-bonds arise from the constructive interference between p_x and p_y orbitals of neighbouring atoms. Recall that we have assigned the z axis to the axis that separates the two atoms in a diatomic molecule, and therefore p_z orbitals

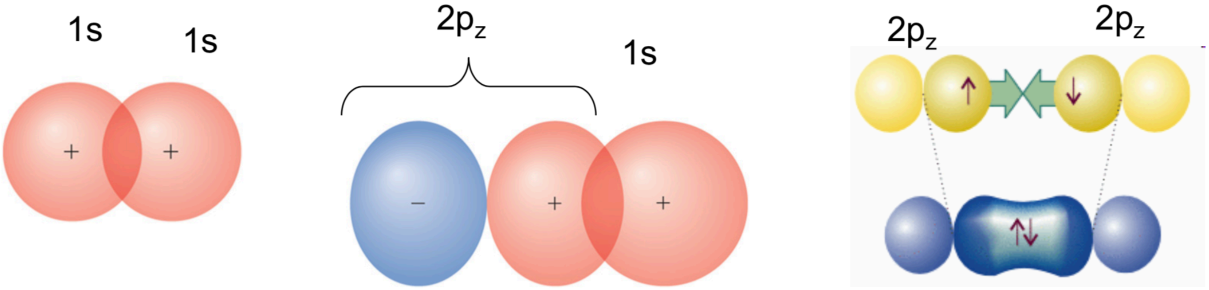


Figure 6.3: A σ-bond arises from the constructive interference of atomic orbitals. Its main characteristic feature is that of cylindrical symmetry around the axis that connect the two atoms (which here we take to be the z axis). There are several ways in which these cilindrically symmetric σ-bonds: *(i)* from the overlap between two $1s$ atomic orbitals, *(ii)* from the overlap of one $2p_z$ and $1s$ orbital, and *(iii)* from the overlap of two $2p_z$ orbitals with paired spins.

are parallel to the z axis and cannot be involved in the formation of a π-type molecular bond. As in the case of the σ-bonds, a definition condition for the π-bonds is that there is constructive interference between the constituent electronic orbitals, else the formation of this chemical bond is not possible.

A molecule with two π-bonds is represented schematically in the left panel of Fig. 6.4. There, a π-bond arises from orbital overlap and spin pairing between electrons in p-type orbitals with their axes perpendicular to the inter-nuclear axis (which by convention has been identified by the z axis). As can be seen, a molecular π-bond is characterised by two overlapping lobes of electron density separated by a nodal plane. Recall that the overlap between the p_x and p_y orbitals that lead to the formation of π-bond should lead to a constructive interference, else the formation of the bond is not possible.

The formation of π-bonds is shown in the right panel of Fig. 6.4 for the specific case of the nitrogen molecule N_2. According to the concepts presented in the previous chapter, see Sect. 5.5, we know that the electronic structure of nitrogen ($Z = 7$) taking into account the building up principle and Hund's rule will be given by:

$$1s^2\, 2s^2\, 2p_x^1\, 2p_y^1\, 2p_z^1 = [\mathrm{He}]\, 2s^2\, 2p_x^1\, 2p_y^1\, 2p_z^1 \,, \tag{6.5}$$

so in each of the three $2p$ orbitals only a single electron will be present. This configuration is a direct consequence of Hund's rule, that tells us that the state of minimum energy of a multi-electron atom is the one with the *highest number of unpaired electrons*.

Taking into account this configuration, as shown in Fig. 6.4, the two $2p_z$ orbitals from each of the two nitrogen atoms will form a (cilindrically symmetric) σ bond, while the two $2p_x$ and the two $2p_y$ orbitals will form one π-bond each respectively. In total, this will lead to a triple bond in molecular nitrogen, which we denote by $\mathrm{N{\equiv}N}$. This discussion is important to emphasize the different nature of the chemical bonds that keep together the nitrogen molecule: in general, more than one type of bond will be present in a given molecule. To summarise, the chemical bonding of the N_2 molecule will be composed by:

- one σ-bond, cilindrically symmetric along the z axis, and
- two π-bonds, composed by p_x and p_y orbitals,

therefore leading to a triple molecular bond (which is characterised by a high degree of chemical stability).

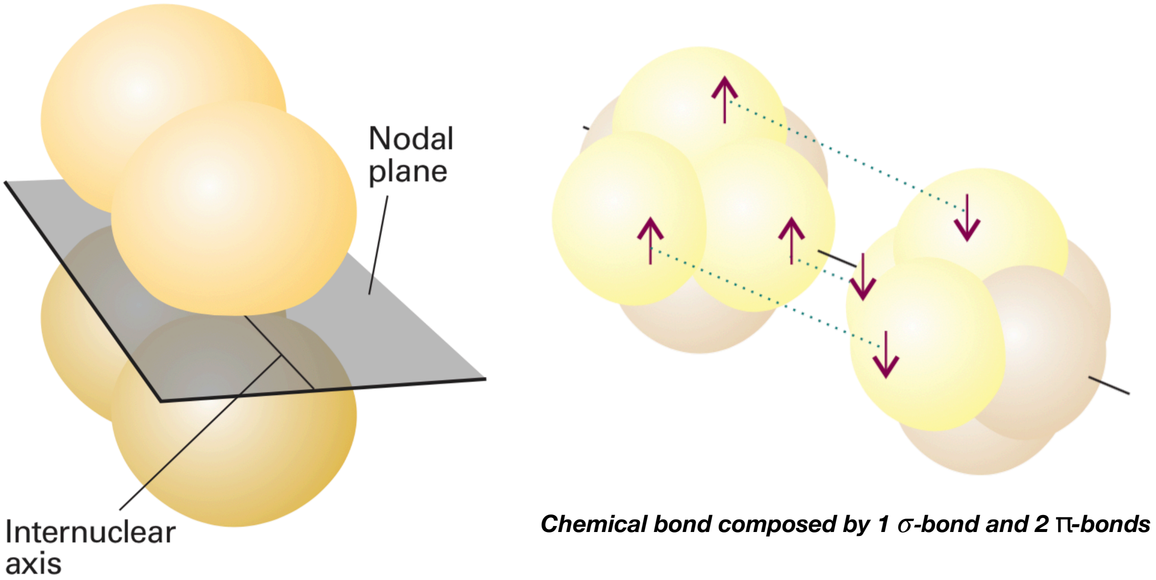


Figure 6.4: Left: A π-bond arises from orbital overlap and spin pairing between electrons in p-type orbitals with their axes perpendicular to the inter-nuclear axis (which by convention has been identified by the z axis). As can be seen, a molecular π-bond is characterised by two overlapping lobes of electron density separated by a nodal plane. Right: the chemical bonding in the nitrogen molecule N_2 is composed by one σ-bond (cilindrically symmetric along the z axis) and two π-bonds (composed by p_x and p_y orbitals), leading to a triple chemical bond.

Resonance in valence bond theory. Another useful concept that appears in Valence Bond theory is that of a *resonance*. Let us illustrate the physical meaning of this concept with the explicit case of the the HCl molecule. The components of this molecule have the following electronic structure:

- H: $1s^1$
- Cl: $[\mathrm{Ne}]3s^2\,3p^5$, where the closed-shell electronic structure of Ne is [He] $2s^2\,2p^6$ and that of Helium is $1s^2$. Therefore the complete electronic structure of Cl is $1s^2\,2s^2\,2p^6\,3s^2\,3p^5$. as dictated by the building-up principle Eq. (5.29).

Taking into account this electronic configuration, and that in Valence Bond theory the chemical bond is created from the interactions between paired electrons, each coming from one of the constituent atoms of the molecule, we see that for this molecule the $1s^1$ electron from the H atom and the unpaired $3p$ electron from the Cl atom should be the ones involved in the formation of the chemical bond.

For these reasons, one could consider writing the spatial part of the wave function of the HCl molecule as follows:

$$\Psi = \Psi^{\mathrm{H}}_{1s}(1)\Psi^{\mathrm{Cl}}_{3p}(2) + \Psi^{\mathrm{H}}_{1s}(2)\Psi^{\mathrm{Cl}}_{3p}(1)\,, \tag{6.6}$$

where the notation indicates that one electron comes from the $1s$ orbital of the H atom and the other electron comes from the $3p$ orbital of the Cl atom. Since electrons are indistinguishable, we have symmetrised the wave function following the procedure also outlined above, see the discussion around Eq. (6.4).

However, it can be shown that the electronic configuration represented in Eq. (6.6) is physically unlikely. The reason is that the electric charge associated to the Cl atom is much higher than that of the H one, $Z_{\mathrm{Cl}} \gg Z_{\mathrm{H}}$, and therefore one expects the electrons to have a larger probability density near the Cl atom occupying the $\Psi^{\mathrm{Cl}}_{3p}(1)\Psi^{\mathrm{Cl}}_{3p}(2)$ orbitals than near the H atom occupying the $\Psi^{\mathrm{H}}_{1s}(1)\Psi^{\mathrm{H}}_{1s}(2)$ orbitals. Note that in principle the electronic structure of Cl is such that it could incorporate the electron from H into a $3p$ orbital and become a Cl^- ion.

For this reason, it is more appropriate to express the combined wave function of the H+Cl quantum system as a superposition of a covalent bond (as described by the Valence Bond formalism) and an ionic bond, where the Cl has absorbed the electron from the H. In other words, one should write:

$$\Psi_{\mathrm{HCl}} = \Psi_{\mathrm{H}-\mathrm{Cl}} + \lambda \Psi_{\mathrm{H^+Cl^-}} \,, \tag{6.7}$$

that us, as a sum of the wave function for the H — Cl molecule bound together by the covalent bond, and the wave function for the ionized system where H and Cl interact via an *ionic bond* due to the fact that the Cl atom has attracted the electron from the hydrogen and thus ionized it. Note that the electronic structure of the Cl atom, $1s^2\, 2s^2\, 2p^6\, 3s^2\, 3p^5$, is such that by incorporating a new electron the Cl^+ ion has now a complete shell, representing a specially stable configuration.

We denote the first component of the wave function in Eq. (6.7) as the *covalent* component and the second as the *ionic* component. In general this configuration also holds for other molecules, and we can write their wave function as

$$\Psi = \Psi_{\mathrm{covalent}} + \lambda \Psi_{\mathrm{ionic}} \,, \tag{6.8}$$

with λ a numerical constant that depends on the specific molecule. The value of λ^2 then provides a measure of the ionic character of this specific molecular bound. This concept is known as *ionic-covalent resonance*. If λ^2 is small it means that a given molecule is characterised by a mostly covalent bound, while if λ^2 is large it means that the ionic character dominates.

6.3 Orbital hybridisation

In the context of Valence Bond theory, we refer to *orbital hybridisation* as the idea that individual atomic orbitals can be combined to form *hybrid orbitals*. In general, these new hybrid orbitals will be characterised by different energies and shapes than the original orbitals. The motivation to construct such hybrid orbitals is then to be able to pair electrons to form chemical bonds. For example, it is possible to form hybrid orbitals, as we will see, from the combination of s-type and p-type atomic orbitals.

The concept of orbital hybridisation is particularly well illustrated with the chemical bonds formed by Carbon (C). This element has $Z = 6$ and its electronic structure is therefore

$$\mathrm{C} \,:\, 1s^2 2s^2 2p_x^1 2p_y^1 \,, \tag{6.9}$$

so with two electrons in the $2p$ orbitals. In principle one might think that carbon can form only two chemical bonds, one for each of the unpaired electrons in the $2p$ orbitals. However, it turns out that this configuration is not possible, but that instead C can form *four independent chemical bonds* within a tetrahedral geometry. The reason for this is precisely the phenomenon of orbital hybridisation: in Valence Bond theory, sometimes the atomic orbitals that constitute the chemical bounds are hybrid orbitals as opposed to the original orbitals.

The reason for this behaviour is that the electronic structure by itself if not enough to fully explain all the possible types of molecular bonds that can be formed in the context of Valence Bond theory. In the specific case of carbon, it is possible that one of the electrons in the $2s$ orbital is promoted to the $2p$ orbital, so that the resulting electronic structure is the following:

$$\mathrm{C} \,:\, 1s^2\, 2s^1\, 2p_x^1\, 2p_y^1\, 2p_z^1 \,, \tag{6.10}$$

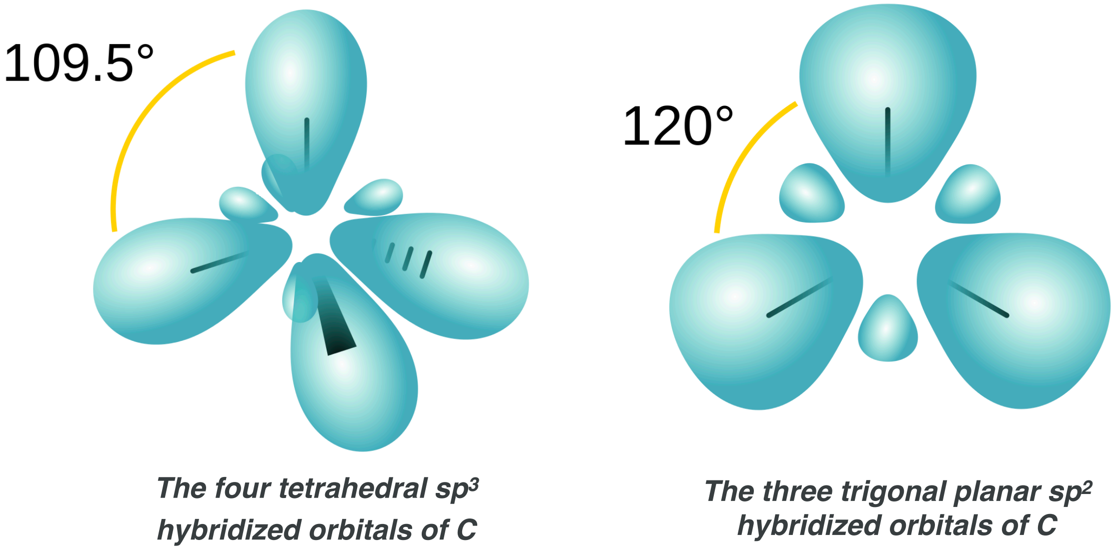


Figure 6.5: In Valence Bond theory, the chemical bonds of carbon are explained by means of the concept of the hybridization of orbitals. In the left plot, we show the four sp^3 hybridized orbitals that lead to the characteristic tetrahedral configuration. In the right plot, we show the corresponding three planar orbitals resulting from the sp^2 hybridisation.

so now we have four unpaired electrons in each of the $n = 2$ shell orbitals. This configuration is advantageous to form chemical bonds since now there are four electrons that can be paired with another four electrons from other atoms, creating four-atom molecules. While this configuration is in principle disfavoured energetically (as can be seen from the building-up principle), this is compensated by the fact that the final molecular bond is composed by four bonds as compared to the original two, and thus exhibits a higher binding energy in total. So overall the molecular binding energy is larger if four bonds are formed starting from the configuration Eq. (6.10) than two bonds with the seemingly lower-energy configuration Eq. (6.9).

In the Carbon atom, the four electronic orbitals hosting unpaired electrons in the configuration Eq. (6.10) can be combined among them to form hybrid orbitals, denoted by h_i. Specifically, one can construct the following linear combinations of orbitals:

$$\begin{aligned} h_1 &= s + p_x + p_y + p_z \,, \\ h_2 &= s - p_x - p_y - p_z \,, \\ h_3 &= s - p_x + p_y - p_z \,, \\ h_4 &= s + p_x - p_y - p_z \,, \end{aligned} \tag{6.11}$$

where note that the hybridisation process leaves unchanged the total number of orbitals. The way these hybridized orbitals are defined determines the geometrical properties of the corresponding molecular bonds, and in this case the characteristic tetrahedral structure of the carbon bonds. Since we end up having sp^3 hybrid orbitals, this process is known as sp^3 hybridisation. The structure of these sp^3 hybrid orbitals is ultimately determined by the constructive and destructive interference between the $2s$ and the $2p$ orbitals.

Several important organic molecules exhibit this phenomenon of the hybridisation of orbitals. One example of the sp^3 hybridisation is methane, CH_4, where each of the sp^3 hybrid orbitals of the C atom leads to a σ-bond (therefore with cylindrical symmetry) with a $1s$ orbital of the H atoms. Another example is ethane, C_2H_6, where both C atoms have sp^3 hybrid orbitals, leading to a σ bond between the two C atoms as well as between the C atoms and the H ones.

The sp^3 hybridisation defined in Eq. (6.11) is not the only possible type of orbital hybridisation. Another possibility is known as the sp^2 hybridisation, which is constructed from a linear combination of the $2s$ orbital with the $2p_x$ and $2p_y$ orbitals, while the unoccupied $2p_z$ orbital is not modified by this process. In this case, the specific linear combinations of atomic orbitals are the following:

$$\begin{aligned} h_1 &= s + \sqrt{2}p_y \,, \\ h_2 &= s + \left(\frac{3}{2}\right)^{1/2} p_x - \left(\frac{1}{2}\right)^{1/2} p_y \,, \\ h_3 &= s - \left(\frac{3}{2}\right)^{1/2} p_x - \left(\frac{1}{2}\right)^{1/2} p_y \,. \end{aligned} \tag{6.12}$$

From the geometric point of view, this configuration corresponds to a flat structure in the (x, y) plane, with angles of 120 degrees separating the three sp^2 hybrid orbitals. The $2p_z$ orbital is unchanged, and remains perpendicular to the (x, y) plane where the three hybridised orbitals lie.

This discussions illustrates how, in the context of Valence Bond theory, the chemical bonds of carbon are explained by means of the concept of the hybridization of orbitals, as illustrated in Fig. 6.5 in the case of Carbon. In the left plot, we show the four sp^3 hybridized orbitals that lead to the characteristic tetrahedral configuration. In the right plot, we show the corresponding three planar orbitals resulting from the sp^2 hybridisation. The unaffected $2p_z$ orbitals are perpendicular to the plane defined by the three sp^2 hybrid orbitals, and contribute to define the overall three-dimensional shape of the resulting molecule.

An example of a molecule exhibiting this sp^2 hybridisation is etheen, C_2H_4, which is defined by the following molecular bonds:

- A σ-bond between the two C atoms, involving hybrid sp^2 orbitals.
- A π-bond between the two C atoms, involving this time $2p_y$ orbitals.
- Four σ-bonds between the C and H atoms, involving the sp^2 orbitals of C with the $1s$ orbital of H.

Finally, another possible type of hybridisation is the one known as sp-hybridisation, which is based on the following linear combination between the $2s$ and $2p_z$ orbitals:

$$h_1 = s + p_z \,, \tag{6.13}$$

$$h_2 = s - p_z \,, \tag{6.14}$$

which lead to two sp orbitals with an angle of 180 degrees among them parallel to the z axis. For this configuration, the $2p_x$ and $2p_y$ orbitals are left unchanged. An example of a molecule which exhibits this type of hybridisation is ethyn, C_2H_2, which is characterised by the following molecular bonds:

- A σ-bond between the two C atoms, involving sp orbitals.
- Two π-bond between the two C atoms, involving the $2p_x$ and $2p_y$ orbitals.
- Two σ-bonds between the C and H atoms, involving the sp orbitals of C with the $1s$ orbital of H.

In Table (6.6) we summarize the various types of the C — C molecular bond types, which appear in molecules containing at least two carbon atoms, that we have been discussing so far. For each bond, we indicate its hybridisation character, its s/p ratio, the angle between equivalent σ bonds, and the length of the carbon-to-carbon bond.

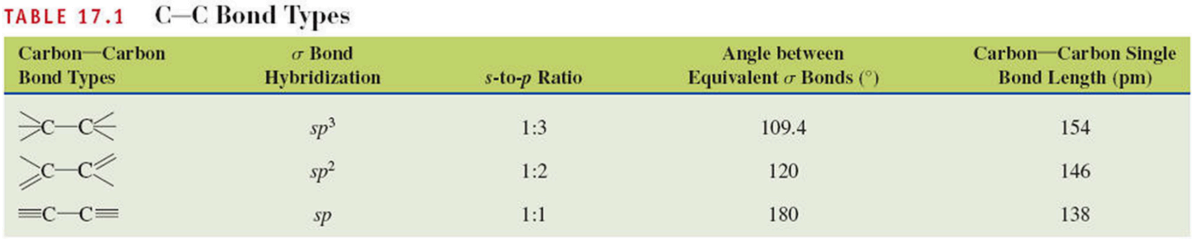


TABLE 17.1 **C—C Bond Types**

Carbon—Carbon Bond Types	σ Bond Hybridization	s-to-p Ratio	Angle between Equivalent σ Bonds (°)	Carbon—Carbon Single Bond Length (pm)
>C—C<	sp^3	1:3	109.4	154
>C—C=<	sp^2	1:2	120	146
≡C—C≡	sp	1:1	180	138

Figure 6.6: The different types of carbon bonds. For each bond, we indicate its hybridisation character, its s/p ratio, the angle between equivalent σ bonds, and the length of the carbon-to-carbon bond.

The same hybridisation of atomic orbitals takes place in other very important elements for organic molecules, such O and N. An important difference with respect the C case is that now we can have *lone electron pairs* being part of the hybridised orbital. Using the same strategy as in the case of carbon, one can determine for example the geometric properties of the hybridised orbitals that lead to the formation of an H_20 molecule.

6.4 Molecular Orbital theory

In the previous discussion, we have studied the chemical bonding in molecules in the framework of the Valence Orbital theory. Now we move to consider a more modern framework to describe chemical bonding, known as *Molecular Orbital theory*. To illustrate how this theory works, and in which aspects does it differ from the Valence Orbital theory, let us consider specific case of the the H_2^+ molecule. This is the simplest of all possible stable molecules, composed only by two protons and one electron, making it suitable to highlight the main features of this new theory.

As before, we will work in the Born-Oppenheimer approximation, and assume that the protons can be considered at rest as compared to the electron. To be more precise, we assume that the motion of the protons takes place at much longer timescales than the motion of the electron. Within this approximation, the Hamiltonian of the system is given by the following expression:

$$\widehat{H} = -\frac{\hbar^2}{2m_e}\nabla_1^2 + V(r_{A1}, r_{B1}, R)\,, \tag{6.15}$$

where the first term corresponds to the kinetic energy of the electron (with ∇_1^2 the Laplacian operator), and the electrostatic potential between the two protons and the electron is given by

$$V(r_{A1}, r_{B1}, R) = -\frac{e^2}{4\pi\epsilon_0}\left(\frac{1}{r_{A1}} + \frac{1}{r_{A2}} - \frac{1}{R}\right)\,, \tag{6.16}$$

where r_{1A} (r_{B1}) is the distance between the first (second) proton and the electron and R is the distance between the two protons that compose the H_2^+ molecule. Note that while the interaction between the protons and the electrons is attractive, the interaction between the protons is repulsive. We can observe from Eq. (6.15) that there is no kinetic term for the protons, since we are working on the assumption that there are at rest.

The basic principle underlying Molecular Orbital theory is that:

the electrons in a molecule cannot be assigned to individual atomic orbital, but rather they occupy *molecular orbitals* that are defined by the solution of the full molecular Schroedinger equation such as Eq. (6.15). In this framework, the chemical bonds are explained in terms of these molecular orbitals, and the distinction between individual atomic orbitals is lost.

Therefore, in order to find the molecular orbitals corresponding to the H_2^+ molecule, we need to find the eigenvectors and eigenvalues of Eq. (6.15). While in this case a exact solution exists, its expression is rather cumbersome and it is more instructive to derive an approximate solution. To achieve this, we need to take into account the following considerations:

- In the limit in which the separation between the two protons R becomes very large, $R \to \infty$, then the molecular orbital should be reduced to the sum of the individual electronic wave functions for the constituent atoms A and B.
- As the two H atoms move closer, below a certain separation threshold, $R < R_e$ with R_e the molecular equilibrium radius defined in Fig. 6.1, the electron can move freely (hop) between the two atoms, as shown schematically in Fig. 6.7.
- Therefore, for $R < R_e$, the concept of two individual, separate, electronic orbitals loses meaning and we have to deal with a single molecular orbital.

In the schematic representation of Fig. 6.7, one indicates how while an electron with (ground state) energy E_{H1s} is not able to hop between protons in the molecule, once R is reduced below the threshold R_e then an electron with energy E_g (that is, the energy of the molecular ground state) is now able to move freely between the two constituent atoms A and B of the molecule. In this schematic representation, the curves in blue represents the electronic wave functions and that in red the electrostatic potential. Recall that for the ground state of hydrogen the electronic wave function is given by $\Psi_{1,0,0}(r) \propto e^{-r/a_0}$.

An important point to mention in this context is that the strength of such molecular bonds is rather weaker than that of the original atomic bonds. To illustrate this, we can perform the following comparison:

- The sum of binding energies between two electrons and two protons, corresponding to two separate H atoms, is $E_b \simeq 2600$ kJ/mol.
- The binding energy between the two H atoms that constitute the H_2 molecule is $E_b \simeq 400$ kJ/mol, so around 6 times smaller.

Therefore we find that the molecular binding energy is much smaller than the atomic binding energy. This explains why to first approximation the atomic orbitals are unchanged when atoms form part of molecules: the molecular bond is only a *correction* (in terms of the total energy budget of the system) of the original atomic bonds. This result also provides, *a posteriori*, a justification of the assumptions underlying Valence Bond theory.

As mentioned above, in order to evaluate the molecular wave function of the H_2^+ molecule, we need to solve Eq. (6.15). To achieve this, we will make the assumption that the wave function describing the molecular orbital can be written as

$$\Psi_\pm = N\left(\Psi_A(r_A) \pm \Psi_B(r_B)\right) \,, \tag{6.17}$$

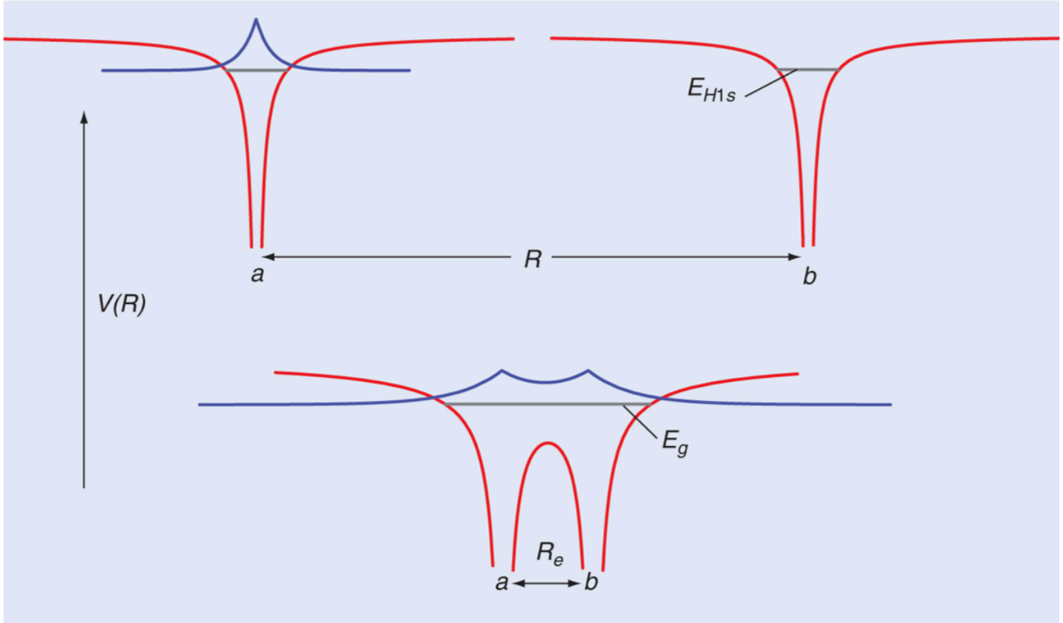


Figure 6.7: Schematic representation of the process leading to the formation molecular orbital of the H_2^+ molecule. If the separation between the two protons R is large, we basically have one electron occupying the original atomic orbital of the hydrogen atom, which in the ground state is $\Psi_{1,0,0}(r) \propto e^{-r/a_0}$, while the second atomic orbital is empty. But at the interatomic separation R is reduced, once the threshold R_e is passed the electron is free to move between the two protons, generating effectively a molecular orbital. Recall that R_e is the molecular equilibrium radius. In this scheme, the line in blue represents the electronic wave functions and that in red the electrostatic potential.

with $\Psi_A(r_A)$ and $\Psi_B(r_B)$ being the wave functions of the 1s orbital of the hydrogen atom, which are given, following the discussion of Sect. 4.2, by

$$\Psi_A(r_A) = \frac{e^{-r_A/a_0}}{\left(\pi a_0^3\right)^{1/2}}\,, \qquad \Psi_B(r_B) = \frac{e^{-r_B/a_0}}{\left(\pi a_0^3\right)^{1/2}}\,, \tag{6.18}$$

with again r_A (r_B) being the distance between the only electron in the H_2^+ molecule and the A (B) protons, and N being an overall normalisation constant. Therefore the full molecular wave function in this approximation will be given by

$$\Psi_\pm(r_1, r_2) = N(R)\left(\frac{e^{-r_A/a_0}}{\left(\pi a_0^3\right)^{1/2}} \pm \frac{e^{-r_B/a_0}}{\left(\pi a_0^3\right)^{1/2}}\right)\,, \tag{6.19}$$

with N being an overall normalisation constant which depends in general on the inter-nuclear separation system R, and is a function of the overlap between the atomic orbitals as discussed below.

The ansatz Eq. (7.9) implies that the electron wave function is delocalised around the H_2^+ molecule, rather than being assigned to any of the two hydrogen nuclei. This approach is called the *Linear Combination of Atomic Orbitals* (LCAO), and is frequently used to construct molecular orbitals.

If we evaluate the probability density associated to the molecular orbital Ψ_+, Eq. (7.9), one finds that

$$|\Psi_+(r_1,r_2)|^2 = N(R)^2\left(\Psi_A^2+\Psi_B^2+2\Psi_A\Psi_B\right) = \frac{N(R)^2}{\pi a_0^3}\left(e^{-2r_A/a_0}+e^{-2r_B/a_0}+2e^{-(r_A+r_B)/a_0}\right). \quad (6.20)$$

This can be compared to the probability density associated to two H_{1s} electrons which are very far from each other, which is given by

$$|\Psi_A(r_1)|^2 + |\Psi_B(r_2)|^2 = \frac{1}{\pi a_0^3}\left(e^{-2r_A/a_0}+e^{-2r_B/a_0}\right) \quad (6.21)$$

Therefore, we find that as compared to the individual atomic orbitals, there is an now additional term $2\Psi_A\Psi_B \propto e^{-(r_A+r_B)/a_0}$ which arises from their overlap (interference) within the molecule.

Thanks for the constructive interference in the probability density of Eq. (6.22), one has a molecular orbital where the probability to find the electron in the region between A and B is higher than if one considers the individual atomic orbitals separately. To show this, we can consider the value that Eq. (6.22) takes in the z axis (which corresponds to the inter-nuclear axis) in the case that A is located at $z=0$, where one has

$$|\Psi_+(z)|^2 = \frac{N(R)^2}{\pi a_0^3}\left(e^{-2|z|/a_0}+e^{-2|z-R|/a_0}+2e^{-(|z|+|z-R|)/a_0}\right), \quad (6.22)$$

This probability density is shown in Fig. 6.8, together with that associated to $\Psi_-(r_1,r_2)$ and to the individual orbitals Eq. (6.21). We can observe how in the case of $\Psi_+(r_1,r_2)$ there is a constructive interference of the wave functions leading to an enhanced probability density in the region between A and B (in this case $0 \le r \le R$), while for the case of $\Psi_-(r_1,r_2)$ we have instead destructive interference in the same region This constructive interference in the case of$\Psi_+(r_1,r_2)$ is beneficial from the energetic point of view: the electron has an *optimal* interaction with the two atomic cores. For this reason, the wave function Ψ_+ is called a *bonding molecular orbital.* One can verify that it corresponds to an σ-type orbital, cilindrically symmetric around the axis that connects the two protons and without any angular momentum around this axis. Specifically, this molecular orbital turns out to be a 1σ orbital, which is the σ-bond orbital with the lowest energy.

It should be clear from Fig. 6.8 that the the normalization constant N in Eq. (6.17) of the molecular orbital does also depend on the inter-nuclear separation R. Intuitively, the degree of overlap of the individual electronic orbitals will increase if R is reduced, affecting the overall wave function normalisation. The expression of $N(R)$ can be computed by integrating $|\Psi_+|^2$ over the whole phase space, and requiring that this integral is 1. For a general diatomic molecule, it is given by

$$N = \left(2(1+S)\right)^{-1/2}, \qquad S \equiv \int d\tau \Psi_A(r_A)\Psi_B(r_B), \quad (6.23)$$

where S quantifies the *overlap* between the two atomic wave functions, which depends on the separation R between the two protons: the smaller the value of R, the greater the overlap will be. Using the explicit expressions for the atomic wave functions ($1s$ orbitals of Hydrogen) one obtains that this overlap integral Eq. (6.23) is given by

$$S(R/a_0) = \left(1+\frac{R}{a_0}+\frac{1}{3}\left(\frac{R}{a_0}\right)^2\right)e^{-R/a_0}, \quad (6.24)$$

which is a function of R/a_0, the ratio between Bohr's radius a_0 (which is a measure of the size of the H atom in the ground state) and the interatomic separation in the molecule R. We observe that the overlap S falls off exponentially for $R \gg a_0$, so therefore it is only significant for $R \lesssim a_0$. In other words, the overlap is only

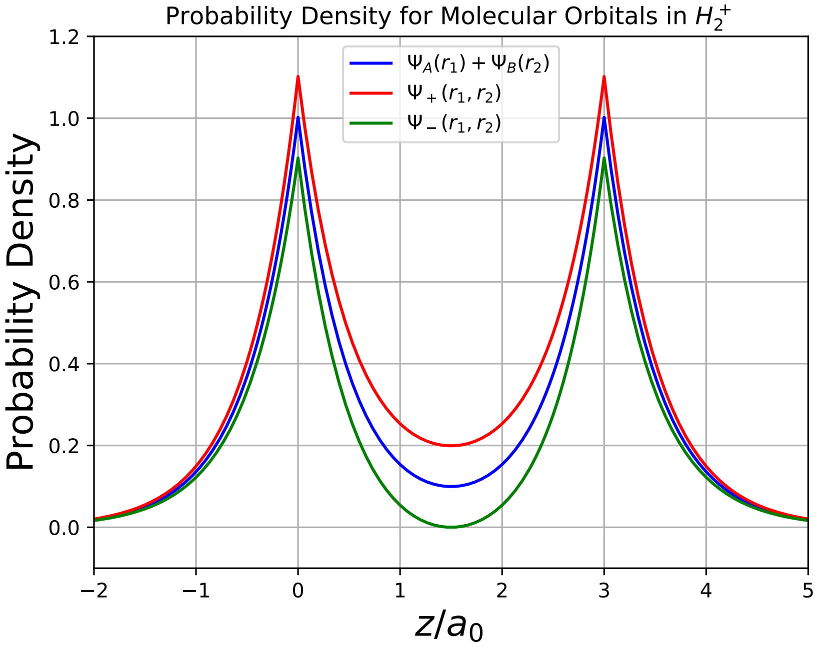


Figure 6.8: The probability density along the inter-nuclear axis (the z axis) for the H_2^+ molecule for the case in which $R = 3a_0$ and where the position of A and B is $z = 0$ and $z = 3a_0$ respectively. The show three cases: the separate atomic orbitals, and the two LCAO molecular orbitals $\Psi_\pm(r_A, r_B)$ defined in Eq. (7.9) We can observe how in the case of $\Psi_+(r_1, r_2)$ there is a constructive interference of the wave functions leading to an enhanced probability density in the region between A and B, while for the case of $\Psi_-(r_1, r_2)$ we have instead destructive interference.

relevant when the distance between the two atoms is comparable to their own size, else it can be neglected. In Fig. 6.9 we show the value of the overlap integral $S(R/a_0)$, Eq. (6.24), for the H_2^+ molecule. We see for example that for $R \geq 3a_0$ the overlap integral is already less than 40% as compared to the maximum overlap possible.

Binding energy in Molecular Orbital theory. By solving the molecular Schroedinger equation using the Hamiltonian in Eq. (6.15), one can determine the energy associated to the Ψ_+ bonding molecular orbital:

$$E_{1\sigma} = E_{H_{1s}} + \frac{j_0}{R} - \frac{J+K}{1+S}\,, \tag{6.25}$$

in terms of the following components:

- $E_{H_{1s}}$ is the energy corresponding to the $1s$ orbital of H, which therefore does not depend on the inter-nuclear distance R.
- The term j_0/R corresponds to the potential energy associated to the repulsive interaction between the two protons separated by a distance R, where one has defined $j_0 = e^2/4\pi\epsilon_0$.
- The last term is given in terms of overlap integrals in a similar way as in Eq. (6.23), which for the H_2^+ molecule are:

$$J \equiv j_0 \int \Psi_A^2 d\tau = \frac{j_0}{R}\left[1 - \left(1 + \frac{R}{a_0}\right)e^{-2R/a_0}\right]\,, \tag{6.26}$$

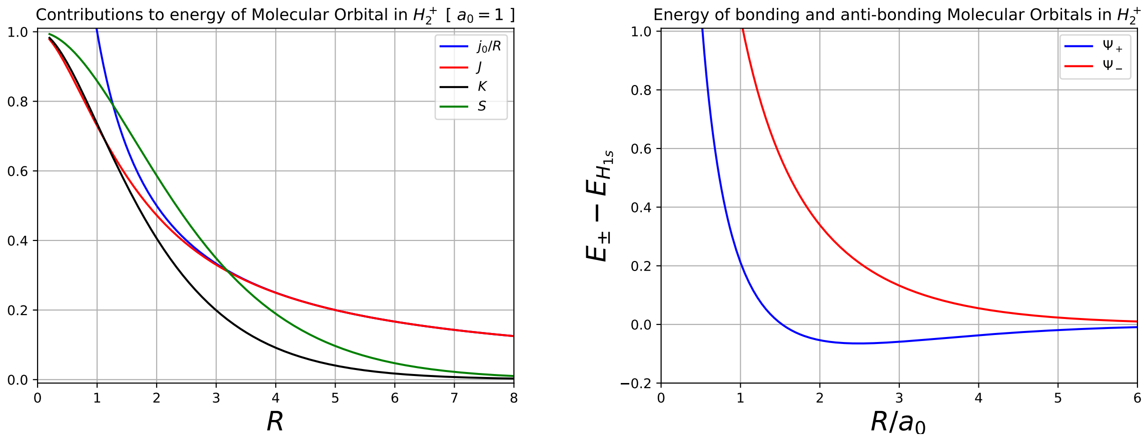


Figure 6.9: Left: The different contributions to the total energy of the molecular bonding orbital Ψ_+ of the H_2^+ molecule, using the notation of Eq. (6.25), including the overlap integral $S(R/a_0)$. Right: the energy associated to the bonding and anti-bonding molecular orbitals in H_2^+ once the energy of the original atomic orbital has been subtracted, Eq. (6.30).

$$K \equiv j_0 \int d\tau \frac{\Psi_A \Psi_B}{r_B} = \frac{j_0}{a_0}\left(1 + \frac{R}{a_0}\right) e^{-R/a_0} \, . \tag{6.27}$$

Let us discuss some important properties of Eq. (6.25), the energy associated to the Ψ_+ bonding molecular orbital:

- Both overlap integrals K and S are positive-definite, but J can be negative.
- The three overlap integrals S, K, and J, are all exponentially suppressed for large values of the interatomic distance R when $R \gg a_0$, so become negligible if the two atoms are very far from each other. In this case the Eq. (6.25) reduces to $E_{1\sigma} \simeq E_{H_{1s}}$, namely the energy of the original electronic orbital of H_{1s}.
- The term J is a measure of the interaction between a given nucleus and the electron density from the other one.
- The term K is a measure for the interaction between one nucleus and the additional electron density between the nuclei arising from the overlap of orbitals.

In the left panel of Fig. 6.9 we show different contributions to the total energy of the molecular bonding orbital Ψ_+ of the H_2^+ molecule, using the notation of Eq. (6.25), in the units where $a_0 = 1$. In the corresponding right panel, we show the total energy, once that of the original electronic orbital of H_{1s} has been subtracted

The fact that in Fig. 6.9 one has that $(E_+ - E_{H_{1s}}) \leq 0$ indicates that the formation of the chemical bond by means of a molecular orbital is energetically favoured.

The other possible linear combination of atomic orbitals in Eq. (6.17) corresponds to the so-called *anti-bonding molecular orbital* Ψ_-. If we compute the corresponding probability density, one fins

$$\Psi_-^2 = N^2 \left(\Psi_A^2 + \Psi_B^2 - 2\Psi_A\Psi_B\right) \, , \tag{6.28}$$

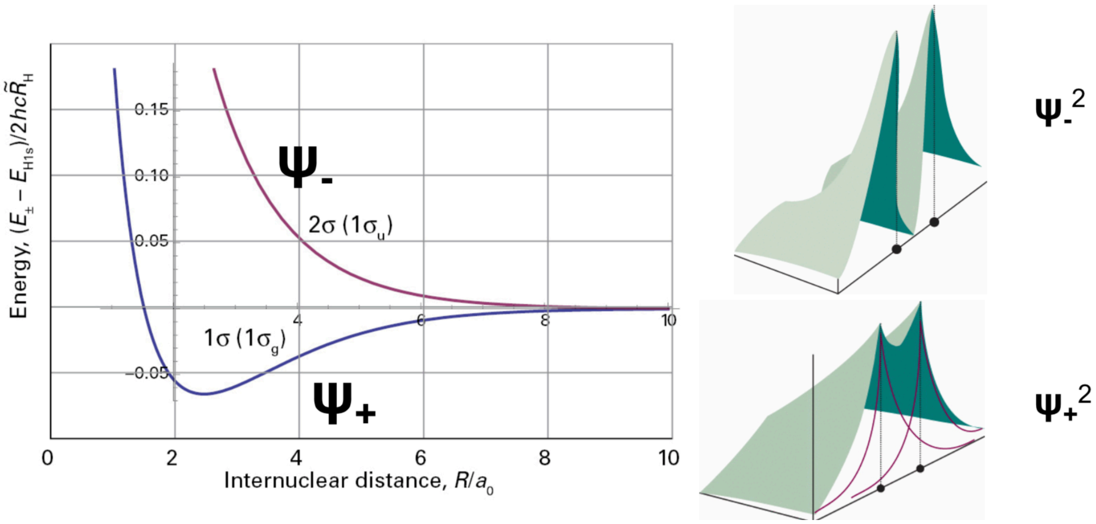


Figure 6.10: Left plot: the energy of the bonding Ψ_+ and anti-bonding Ψ_- orbitals as a function of the inter-nuclear distance R/a_0. Right plot: the probability densities associated to the the anti-bonding (upper) and bonding (lower) molecular orbitals. The fact that this probability is large in the region between the two atoms in the latter case reflects the bonding character of this molecular orbital.

which as shown in the right plot of Fig. 6.10, as well as in Fig. 6.8 leads to a smaller electron density in the region between the hydrogen nuclei due to *destructive* interference. This molecular orbital turns out to be a 2σ-type orbital, which is also a σ-bond but with a higher energy than the one associated to the Ψ_+ orbital. The energy of this orbital is now given by

$$E_{2\sigma} = E_{H1s} + \frac{j_0}{R} - \frac{J-K}{1-S} \, , \tag{6.29}$$

In Fig. 6.10 (see also Fig. 6.9) we compare the energies of the bonding and anti-bonding orbitals, $E_{1\sigma}$ and $E_{2\sigma}$ respectively, as a function the inter-nuclear distance R/a_0, once we have subtracted the energy of the original atomic orbitals E_{H1s}. That is, we are plotting

$$\left(E_\pm - E_{H1s}\right)/2hc\widetilde{R}_h = \frac{1}{2hc\widetilde{R}_h}\left(\frac{j_0}{R} - \frac{j \pm k}{1 \pm S}\right) \, . \tag{6.30}$$

One can see how for the bonding orbital the total energy is smaller than that of the separate electronic orbitals, justifying the formation of the molecular bond in the H_2^+ molecule.

Consistently with the structure of Eq. (6.30), we see from Fig. 6.10 that both for the bonding and anti-bonding orbitals this energy goes to zero for large internucleear distances. We also find that the absolute strength of the 1σ bonding orbital is weaker than the 2σ anti-bonding orbital, a fact that will become important later in the course.

Parity transformation. Another important property of these molecular wave functions is their symmetry properties upon a reflecting transformation, what is known as a *parity transformation.* In particular, the Ψ_+ orbital is symmetric (*gerade g*), meaning that upon a parity transformation $\vec{r} \to -\vec{r}$ it transforms as

$$\Psi_+(x,y,z) = \Psi_+\left(-x,-y,-z\right) \, , \tag{6.31}$$

while instead the Ψ_- orbital is anti-symmetric (*ungerade u*), meaning that upon the same transformation we have

$$\Psi_+(x,y,z) = -\Psi_-\left(-x,-y,-z\right) \ , \tag{6.32}$$

Due to this property, it is customary to denote in this approach a bonding molecule orbital as σ_g while an anti-bonding one as σ_u. In addition, this anti-bonding character is often indicated with a σ_u^* sign. This notation will be useful to discuss more advanced aspects of the molecular chemical bonds to be presented in the following lectures.

Summary

To summarize, some important concepts that we have learned in this lecture about electronic orbitals in molecules and chemical bonding are the following:

- The equilibrium radius R_e corresponds to the value of the inter-nuclear distance R for which the molecule is in the ground state. Its value can be computed in the framework of the Molecular Orbital theory, from the minimisation of the energy associated to the molecular orbital with lowest energy.
- Within Valence Bond theory, chemical bonds are formed from the constructive interference of the wave functions of valence electrons whose spins are paired and that are localised in specific atoms of the molecule.
- Atomic orbitals can hybridise to form chemical bounds in molecules if the resulting configuration is energetically favourable.
- Within the Molecular Orbital approach, the electronic orbitals of molecules are constructed from linear combinations of atomic orbitals; electrons are delocalised in the whole molecule and cannot be ascribed to specific atoms.

In the next lecture we will continue our study of the implications of quantum theory for the description of the properties of molecules.

Further reading

Further reading material about the topics covered in this lecture can be found in the following sections of the course textbook (*Physical Chemistry*, Atkins and De Paula, 10th edition):

- **10A.1**: *Diatomic molecules.*
- **10A.2**: *Polyatomic molecules.*
- **10B.1**: *Linear combination of atomic orbitals.*

van Quantum tot Molecuul

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

7 HC8: Chemical bounds and molecules II

In this lecture, we continue our study of the chemical bonds in molecules, following the previous general discussion of the main features of Valence Bond theory and Molecular Orbital theory. From now on, we will focus on the more modern Molecular Orbital framework, and here we will apply it first to determine the electronic structure of homo-nuclear diatomic molecules, and then of hetero-nuclear diatomic ones. We will present the corresponding version of the building-up principle of multi-electron atoms in the molecular case, and develop a variational principle to construct the molecular orbitals for hetero-nuclear molecules.

The **learning goals** of this lecture are:

(a) Determining the filling order of molecular orbitals for various types of diatomic molecules.

(b) Computing the binding order of diatomic molecules, which determines the degree of (in)stability of the corresponding chemical bond.

(c) Constructing molecular orbitals suitable for the description of hetero-nuclear diatomic molecules.

(d) Exploiting the variational principle to determine the coefficients of molecular orbitals for hetero-nuclear molecules.

We therefore start by further applying the principles of Molecular Orbital theory for the description of chemical bonding in homo-nuclear diatomic molecules, composed by two atoms of the same element.

7.1 Chemical bonding in homo-nuclear diatomic molecules

Let us start our discussion by considering homo-nuclear diatomic molecules, such as O_2, N_2, and F_2, which are composed by two atoms of the same element. For these molecules:

only the valence electrons will involved in the chemical bonding, while the electrons in the (completed) inner shells will not play any role and can be safely ignored.

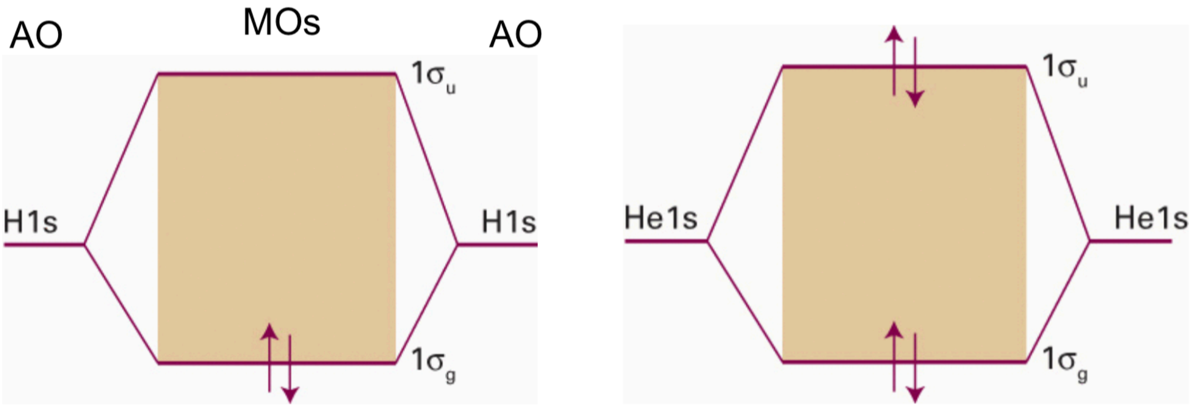


Figure 7.1: The electronic structure of molecular hydrogen H_2 (left) and helium He_2 (right plot). The individual atomic orbitals (of the $1s$ type) merge into two molecular orbitals: a bonding orbital $1\sigma_g$ and an anti-bonding orbital $1\sigma_u$. In the case of the H_2 molecule, the two electrons occupy the bonding orbital, leading to an stable molecule. In the case of the He_2 molecule instead, two electrons in addition also occupy the anti-bonding orbital, resulting into an unstable configuration from the energetic point of view. (see also Fig. 6.10).

For instance, in the H_2 molecule, only the $1s$ valence electrons are involved in the formation of the chemical bounds. Here we will work in the framework of the Molecular Orbital theory presented in Sect. 6.4, where the electrons are delocalised within the molecule and occupy electronic molecular orbitals.

For these molecules, a modified version of the building-up principle discussed in the Sect. 5.5 also holds. Recall that the building up principle determines the order in which electronic shells are filled in multi-electron atoms, based on their energy: shells that have lower energy must be filled first. Moreover, Hund's rule also tells us that degenerate orbitals are filled first with single electrons before each orbital is completely filled with paired electrons. In the case of molecular orbitals, the corresponding version of the building up principle arises from similar physical considerations and is based on the following set of rules:

- Electrons will occupy *empty molecular orbitals* in a sequential way to ensure that the total energy of the molecule is minimised.

- No more than two electrons can occupy a given molecular orbital, as a consequence of the Pauli exclusion principle.

- Electrons occupy first of all the unpaired individual degenerate orbitals (orbitals with the same associated energy), in order to minimize the energy induced by the electron-electron repulsive interaction. This is the same as Hund's rule in the case of multi-electron atoms, see the discussion in Sect. 5.5.

- The above condition implies that two electrons occupying a degenerate orbital will have parallel spins, since this leads to a state with a smaller total energy.

As an example, let us take a look at the electronic structure of molecular hydrogen, represented in Fig. 7.1. As we can see, the two individual atomic orbitals (of the $1s$ type) merge into two molecular orbitals, one with bonding character, $1\sigma_g$, and the other anti-bonding character, $1\sigma_u$. The two electrons occupy the bonding orbital (since it has the lowest energy of the two molecular orbitals), so we can define the electronic structure of H_2 as $(1\sigma_g)^2$. This means that the H_2 molecule is stable, due to the fact that all electrons involved in the chemical bond occupy a bonding orbital, which is an energetically-favoured configuration.

The situation on the other hand is quite different for molecular Helium, He_2, also shown in Fig. 7.1 (right panel). This molecule has four electrons, so therefore two electrons occupy the bonding orbital $1\sigma_g$

while the other two electrons occupy instead the anti-bonding orbital $1\sigma_u^*$. Therefore the resulting electronic configuration would be $(1\sigma_g)^2(1\sigma_u^*)^2$. This is a crucial difference as compared to the H_2 case: the anti-bonding orbital, as was shown in Fig. 6.10, is stronger as compared to the bonding orbital. In other words, one has that

$$|E_+ - E_{H_{1s}}| < |E_- - E_{H_{1s}}| \, , \tag{7.1}$$

implying that the total binding energy is positive and therefore the He_2 molecule is not a stable chemical configuration. This illustrates that in order to achieve chemical bonds one needs to populate with electrons the appropriate type of molecular orbitals, those that minimise the overall energy of the molecule.

Building upon the discussion of H_2 and He_2, let us now turn to discuss the electronic structure of homonuclear diatomic molecules with $Z > 2$ such as Li_2, Be_2, B_2, C_2, N_2, O_2, F_2. All these elements are characterised by having their valence electrons in the $2s$ and $2p$ atomic orbitals, with the inner $1s$ shell completed and thus not playing any role in the chemical binding. In other words, for these molecules, the electrons that occupy the $1s$ orbitals do not count anymore from the point of view of the molecular bonds. Indeed, following the building up principle and Hund's rule, the electronic structures of these elements are:

- Li $(Z = 3)$: [He] $2s^1$
- Be $(Z = 4)$: [He] $2s^2$
- C $(Z = 6)$: [He] $2s^2 2\,p_x^1\, 2p_y^1$
- N $(Z = 7)$: [He] $2s^2 2\,p_x^1\, 2p_y^1\, 2p_z^1$

and so on. Concerning the geometry of the resulting molecular orbitals, recall that:

> an important property of the $2s$ atomic orbitals is that they are *rotationally symmetric*. Recall from the discussion in Sect. 4.3 that s-type orbitals have $l = m_l = 0$, and therefore the angular dependence of their wave function will be given by $Y_{0,0} = (4\pi)^{-1/2}$, that is, constant. So all directions in the (θ, φ) surface have the same probability and therefore s-type orbitals are rotationally symmetric.

Therefore, in the context of Molecular Orbital theory, the $2s$ atomic orbitals lead to the formation of both a (cilindrically symmetric) $1\sigma_g$ bonding molecular orbital as well as a $1\sigma_u^*$ anti-bonding molecular orbital. In addition to this, we also know that the $2p_z$ atomic orbitals can also lead to the formation of a molecular orbital of the σ-type (recall that the z-axis coincides with the inter-nuclear axis). Specifically, the formation of a σ-type molecular orbital is then possible for the following configurations:

- $2s - 2s$,
- $2s - 2p_z$,
- $2p_z - 2p_z$.

Taking into account the energies associated to each molecular orbital, the $2s - 2s$ and $2p_z - 2p_z$ configurations turn out to be the most favoured ones. The $2s - 2s$ combination leads to the formation of a $1\sigma_g/1\sigma_u^*$ pair of molecular orbitals, while the $2p_z - 2p_z$ combination leads to a $2\sigma_g/2\sigma_u^*$ pair instead (therefore with higher energy). In the latter case, the $2p_z - 2p_z$ combination leads to a bonding molecular orbital ($2\sigma_g$) while the $2p_z + 2p_z$ combination instead leads to an anti-bonding molecular orbital ($2\sigma_u^*$).

In addition to these σ-type orbitals, in homo-diatomic molecules also the π-type molecular orbitals play an important role. One of these orbitals can be constructed as

$$1\pi_u \,:\, 2p_x + 2p_x \,, \tag{7.2}$$

which leads to a *bonding molecular orbital*. This orbital is odd under coordinate inversion (parity transformation), the opposite as the corresponding case of the σ bonds. A closely related orbital is

$$1\pi_g \,:\, 2p_x - 2p_x \,, \tag{7.3}$$

which leads instead to destructive interference and thus corresponds to an *anti-bonding orbital*. This orbital is even (*gerade*) under a coordinate inversion transformation. Likewise, we will have another $1\pi_u$ ($1\pi_g$) bonding (anti-bonding) molecular orbital from the $2p_y + 2p_y$ ($2p_y - 2p_y$) linear combination of atomic orbitals. Therefore, we have the important result that

> The $1\pi_u$ and $1\pi_g^*$ molecular orbitals are degenerate, and can in total be occupied by up to four electrons (that is, two per orbital).

Finally, one can demonstrate that it is not possible to construct molecular orbitals from the combination of a $2s$ with either a $2p_x$ or a $2p_y$ orbital, the reason being that the corresponding overlap integral S vanishes. To summarise, for diatomic molecules where the $n = 2$ electrons are involved in the formation of the chemical bonds, one has the following molecular orbitals to be occupied by electrons:

- *Bonding orbitals*: $1\sigma_g$ ($2\ e^-$), $2\sigma_g$ ($2\ e^-$), $1\pi_u$ ($4\ e^-$).
- *Anti-bonding orbitals*: $1\sigma_u^*$ ($2\ e^-$), $2\sigma_u^*$ ($2\ e^-$), $1\pi_g$ ($4\ e^-$).

A molecule will be stable if more electrons occupy bonding orbitals than anti-bonding orbitals.

Concerning the specific electronic configuration of these molecules, as mentioned above we have a version of the building-up principle for molecular orbitals. As in the atomic case, electrons start to fill the molecular orbitals starting from the first orbital available with the lowest energy. This is illustrated in Fig. 7.2, where we indicate the energy of the molecular orbitals for homo-different diatomic molecules (composed by the same element) for elements from $Z = 3$ (Li) to $Z = 9$ (F). In each case, we indicate how the $2(Z - 2)$ electrons present in the molecule fill the various orbitals (recall that the 2 electrons in the $1s$ atomic shells do not participate in the bonding). As mentioned above, the $1\pi_u$ and $1\pi_g$ molecular orbitals are degenerate, and can in total be occupied by four electrons, as happens for C_2 and heavier molecules.

From Fig. 7.2 we see that for molecules with $Z \le 7$, the associated building-up principle has the following structure:

$$1\sigma_g\,,\ 1\sigma_u\,,\ 1\pi_u\,,\ 2\sigma_g\,,\ 1\pi_g\,,\ 2\sigma_u\,. \tag{7.4}$$

Note however that is ordering not universal, for example for molecules with $Z \ge 8$ the corresponding building up principle is given instead by

$$1\sigma_g\,,\ 1\sigma_u\,,\ 2\sigma_g\,,\ 1\pi_u\,,\ 1\pi_g\,,\ 2\sigma_u\,, \tag{7.5}$$

where the $2\sigma_g$ bonding orbital is filled earlier than the $1\pi_u$ orbital.

The molecular binding order b. In order to illustrate this filling mechanism, let us consider the electronic structure of molecular nitrogen N_2. Nitrogen has $Z = 7$ electrons, or which 5 participate in the chemical

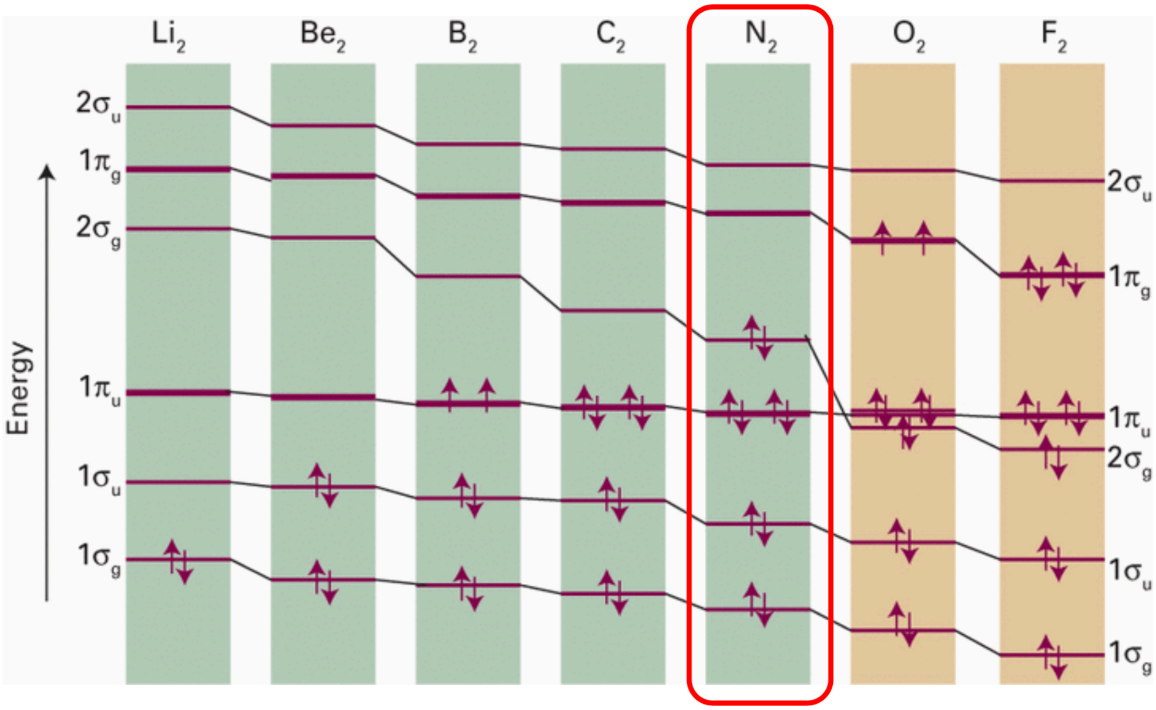


Figure 7.2: The energy of the molecular orbitals for homo-different diatomic molecules (composed by the same element) for elements from $Z = 3$ (Li) to $Z = 9$ (F). In each case, we indicate how the $2(Z-2)$ electrons present in the molecule fill the various orbitals (recall that the 2 electrons in the $1s$ atomic shells do not participate in the bonding). The $1\pi_u$ and $1\pi_g$ molecular orbitals are degenerate, and can in total be occupied by four electrons, as happens for C_2 and heavier molecules.

bonding from the $n = 2$ shell. Therefore, the molecular orbitals will host 10 electrons. Following the filling order indicated in Fig. 7.2, the electronic structure of the N_2 molecule will be given by

$$1\sigma_g^2\,1\sigma_u^{*2}\,1\pi_u^4\,2\sigma_g^2\,, \tag{7.6}$$

where recall that we are using a notation such that:

- $\sigma_{g(u)}$ indicate bonding (anti-bonding) molecular orbitals
- $\pi_{g(u)}$ indicate anti-bonding (bonding) molecular orbitals.

An important question in this context is what factors determine the overall stability of a given molecule:

The stability of a given molecular chemical bond can be characterised by its *binding order* b:

$$b \equiv \frac{1}{2}\left(N - N^*\right)\,, \tag{7.7}$$

with N (N^*) the number of electrons occupying bonding (anti-bonding) molecular orbitals.

Let us compute the value of the binding order b in the case of the N_2 molecule. From the molecular electronic structure of Eq. (7.6), we see that 8 electrons occupy bonding orbitals ($1\sigma_g^2$, $1\pi_u^4$, $2\sigma_g^2$) while 2 electrons occupy anti-bonding orbitals ($1\sigma_u^{*2}$). We have thus that $b = (8-2)/2 = 3$, and therefore this molecule is characterised by a triple binding. It is clear that the higher the value of the binding order b, the stronger a given molecular chemical bond will be. Molecules with $b \leq 0$ are thus unstable, as in the case of example of molecular helium, for which $b = 0$ (since as shown in Fig. 7.1) the same number of electrons occupy bonding and anti-bonding orbitals.

Chemical bonding in O_2. As another illustrative example, let us consider the electronic structure of molecular oxygen O_2, a molecule which is characterized by having a total of 16 $(= 2Z)$ electrons, of which 12 play a role in the chemical bonding (since the four electrons in the $1s^2$ shells are not relevant here). The electronic structure of the oxygen molecule is illustrated in Fig. 7.3. As we can see, each of the two oxygen atoms has a $2s$ and three $2p$ orbitals, which are combined among them when forming the molecular bonds. The two $2s$ orbitals combine into a $1\sigma_u$ and a $1\sigma_g$ orbitals, such as in the case of molecular nitrogen. The three $2p$ orbitals combine into $2\sigma_g$, $2\sigma_u$, $1\pi_g$ and $1\pi_u$ orbitals. Note that the $2\sigma_g$ orbital has a lower energy that the $1\pi_u$ orbital, as opposed to the case of N_2. Note also an important property that always hold when constructing molecular orbitals:

> The total number of electrons that can occupy the molecular orbitals must be the same as the total number of electrons that occupy the original atomic orbitals. In the case of O_2, we have that up to 12 electrons can occupy the $2p$ orbitals of both O atoms, which the same number that can occupy the corresponding molecular orbitals: 2 in $2\sigma_g$, 4 in $1\pi_u$, 4 in $1\pi_g$, and 2 in $2\sigma_u$.

Given that O_2 has 16 electrons, it is easy to see that all molecular orbitals up to the $1\pi_u$ one are fully filled. On the other hand, there are only two electrons in the $1\pi_g^*$ anti-bonding molecular orbital, which is twice degenerated. Following Hund's rule, we know that these two electrons will occupy different $1\pi_g^*$ orbitals. Therefore, the two electrons in the different $1\pi_g^*$ orbitals will be unpaired, leading to a triplet spin state. Putting all this information together, we find that the binding order of molecular oxygen is

$$b_{O_2} = (8-4)/2 = 2 > 0\,, \tag{7.8}$$

thus forming a double chemical bond, and demonstrating that the formation of this molecule is energetically favourable.

7.2 Heteronuclear diatomic molecules

Following this discussion of the electronic structure of homonuclear diatomic molecules, now we turn to study those molecules which are composed by two different atoms, known as *heteronuclear molecules*. In order to start the discussion, we must first of all take a step back and revisit our discussion of the H_2 molecule. There we saw that the molecular orbitals, Eq. (6.17), were constructed from a linear combination of the $1s$ atomic orbitals of H and were given by

$$\Psi_\pm = N\left(\Psi_A \pm \Psi_B\right)\,, \tag{7.9}$$

with the same normalization factor N for the two components. In other words, the two atomic orbitals Ψ_A and Ψ_B contribute with the same weight to the molecular orbital. The resulting electronic molecular configuration was $(1\sigma_g)^2$, which corresponds to a bonding orbital (the Ψ_+ combination).

Can now now apply the ideas of Molecular Orbital theory for the case of heteronuclear diatomic molecules.

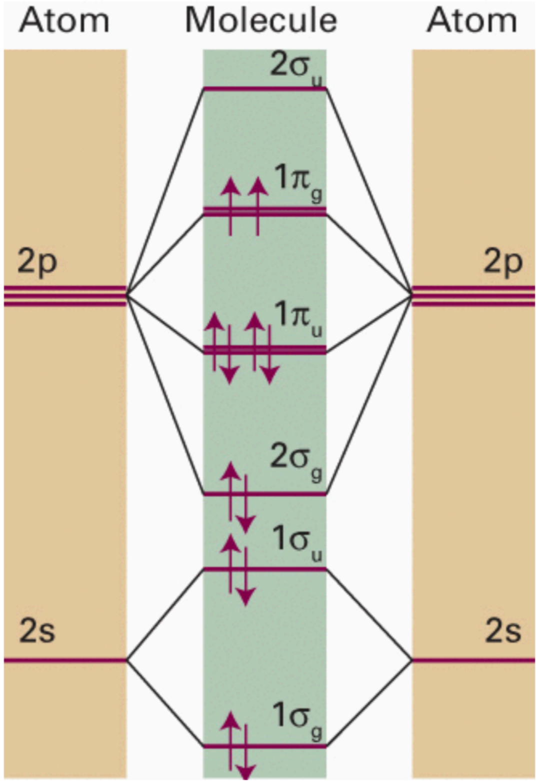


Figure 7.3: The electronic structure of the O_2 molecule. The 8 electrons from the $n = 2$ shells are involved in the formation of the molecular orbitals and the corresponding chemical binding. The molecule is stable because the overall binding energy in this case is dominated by that of the bonding orbitals rather than by the anti-bonding orbitals. Indeed, one can check that the binding order is $b = (8 - 4)/2 = 2 > 0$ and thus the formation of this molecule is energetically favourable.

In this case the molecular electronic wave function will be of the form:

$$\Psi = c_A \Psi_A + c_B \Psi_B \, , \tag{7.10}$$

with the main difference as compared to homo-nuclear molecules are that:

- Since in general $c_A \neq c_B$, the molecular orbitals will not have well-defined symmetries under parity transformation (see the discussion in Sect. 6.4).
- For the same reason, the corresponding atomic orbitals Ψ_A and Ψ_B used to construct the molecular orbital Ψ in Eq. (7.10) will now have different energies.
- The fact that $c_A^2 \neq c_B^2$ also implies that the probability distribution of the electrons involved in the molecular bond will not be symmetric with respect inter-nuclear axis, but will be peaked either close to A or close to B (depending if which of the two coefficients is bigger).

The latter property in particular has the consequence that that the resulting chemical binding will be *polarised*, that is, it will exhibit a non-zero electric dipole moment. For specific the case of an *ionic binding*, one of the coefficients c_A or c_B will be set to zero.

Let us illustrate these concepts with an specific example, namely the HF molecule. The molecular wave

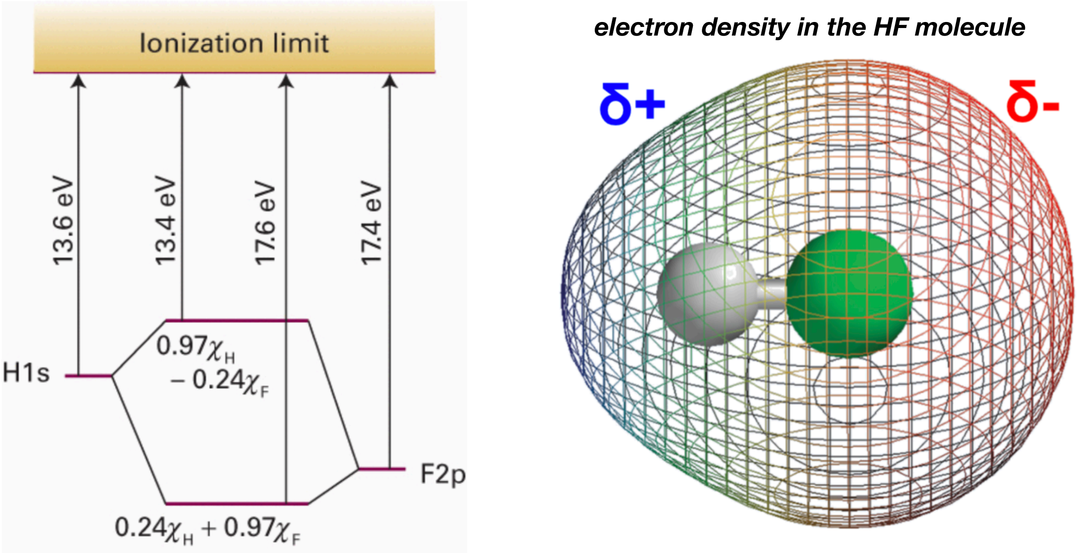


Figure 7.4: Left: the energies of the molecular orbitals of HF, compared to that of the original atomic orbitals. The ground state is dominated by the $2p$ orbital of F, as shown by the fact that they have similar ionisation energies. Right: the electron density in the HF molecule is larger closer to the F nucleus than to the hydrogen, inducing a net electric dipole moment in the molecule: the regions with a higher likelihood of finding the electrons will be on average more negatively charged than the regions with lower likelihood of finding them.

function will be given by

$$\Psi = c_H \Psi_H + c_F \Psi_F \, , \tag{7.11}$$

where with Ψ_H we indicate the $1s$ orbital of H while Ψ_F stands for the $2p_z$ atomic orbital of F. These are the two atomic orbitals that are involved in this chemical bond. It is possible to determine that the values of the coefficients are $c_H = 0.24$ and $c_F = 0.97$ for the ground molecular state and $c_H = 0.97$ and $c_F = -0.24$ for the first excited molecular state, as represented schematically in Fig. 7.4. From these values for the coefficients c_A and c_F we can determine that:

- In both cases the molecular wave functions will be normalized since $c_F^2 + c_H^2 = 1$ (and the individual atomic orbitals are normalised by construction).

- The molecular ground state,

  $$\Psi = 0.24\, \Psi_H + 0.97\, \Psi_F \, , \tag{7.12}$$

  is dominated by the $2p_z$ orbital of F, and therefore the electrons in this orbital will be closer to F than to H, as illustrated schematically in Fig. 7.4. This is a consequence of the fact that the total positive effective charge of F, $Z_{\rm eff} \gg 1$, is rather larger than that of H.

- The electron density is asymmetric with respect to the axis that connects the two nuclei, given that the electron is more likely to be found close to the F nucleus. This implies that HF is a *polar molecule* with a non-zero electron dipole moment.

In general, in a heteronuclear diatomic molecule, if the element A is more likely to attract elements (for instance if its effective electric charge $Z_{A,\rm eff}$ is larger) than one will have that $|c_A|^2 > |c_B|^2$.

Electro-negativity and electro-affinity.(*) The polar character of a given molecule can be assessed in terms of the *electro-negativity* χ of the component atoms. This electro-negativity χ of an atom quantifies the tendency of an atom to attract the molecular electron density towards itself. Comparing the values of χ for some important elements we find for example that

$$\chi_F > \chi_O > \chi_N > \chi_C > \chi_H \, . \tag{7.13}$$

This implies in particular that within organic molecules composed by elements such as H, C, N, and O, the electron density will be larger closer to the O and N nuclei than closer to the H and C nuclei.

A formal definition of of the electro-negativity χ was provided by Pauling, and is given by

$$\chi_P(A,B) = \left[D_0(AB) - \frac{1}{2}\left(D_0(AA) + D_0(BB)\right) \right]^{1/2} \, , \tag{7.14}$$

where D_0 indicates the *dissociation energy* of the system in the parenthesis.[3] Another definition is the one provided by Mulliken, given by

$$\chi_M(A,B) = \frac{1}{2}\left(1 + E_{ea}\right) \, , \tag{7.15}$$

with I being the ionisation energy of the system and E_{ea} the *electro-affinity*, defined as the is amount of energy released or spent when an electron is added to an atom or molecule. The two definitions can be shown to be related among them as follows:

$$\chi_P = 1.35\sqrt{\chi_M} - 1.37 \, , \tag{7.16}$$

and they have associated the same qualitative interpretation.

7.3 The variational principle for diatomic molecules

In the context of the quantum mechanics of molecules, the variational principle represents a systematic way of determining the degree of polarity and the value of the coefficients of a given molecular orbital constructed as in Eq. (7.10). This variational principle is based the following general idea:

> A randomly chosen molecular wave function, Ψ, has associated an energy E higher than the energy for the "true" molecular wave function Ψ_0, that is, $E \le E_0$, since Ψ_0 corresponds to the molecular ground state.

Therefore, the procedure underlying the variational principle is the following:

- one starts by constructing an specific wave function, $\Psi = c_A\Psi_A + c_B\Psi_B$, as a linear combination of atomic orbitals, with arbitrary values values of c_A and c_B,
- then one varies the values of c_A and c_B until the total energy of the molecule E has been minimised: we identify the resulting values of c_A and c_B as those that define the *molecular ground state.*

From the mathematical point of view, the variational principle corresponds to finding the solution of an optimisation problem of the form

$$\frac{\partial E}{\partial c_A} = \frac{\partial E}{\partial c_B} = 0 \, , \tag{7.17}$$

---

[3] In chemistry, the bond-dissociation energy D_0 is one possible measure of the strength of a given chemical bond.

where total energy of the molecule, E, is the expectation value of the corresponding Hamiltonian:

$$E(c_A, c_B) = \frac{\int \Psi^* \widehat{H} \Psi d\tau}{\int \Psi^* \Psi d\tau}\,, \tag{7.18}$$

with $\hat{H}$ the electronic Hamiltonian of the molecule, see for example Eq. (6.15) for the case of the H_2^+ molecule, and as usual $d\tau$ indicates integration over all phase space. The variational principle is widely used in quantum mechanics, in particular to find the ground states and energies of quantum systems for which an analytical treatment is not available.

Let us now compute the energy Eq. (7.18) in the specific case of an heteronuclear diatomic molecule, where the molecular orbital is expressed as a linear combination of atomic orbitals as

$$\Psi = c_A\, \Psi_A \,+\, c_B\, \Psi_B\,. \tag{7.19}$$

We start by computing the denominator in Eq. (7.18). Using the fact that Ψ_A and Ψ_B are normalized atomic orbitals, one has that

$$\int \Psi^* \Psi d\tau = c_A^2 + c_B^2 + 2 c_A c_B S\,, \qquad S = \int \Psi_A^*\, \Psi_B\, d\tau \tag{7.20}$$

with S being the usual *overlap integral* between the atomic orbitals. The corresponding expression for the numerator of Eq. (7.18) is given by

$$\int \Psi^* \widehat{H} \Psi d\tau = c_A^2 \alpha_A + c_B^2 \alpha_B + 2 c_A c_B \beta\,, \tag{7.21}$$

where we have defined the following terms:

$$\alpha_A \equiv \int \Psi_A^* \widehat{H} \Psi_A d\tau\,, \tag{7.22}$$

$$\alpha_B \equiv \int \Psi_B^* \widehat{H} \Psi_B d\tau\,, \tag{7.23}$$

$$\beta \equiv \int \Psi_B^* \widehat{H} \Psi_A d\tau\,. \tag{7.24}$$

In the equations above, α_A and α_B correspond to the expectation values of the molecular Hamiltonian operator $\hat{H}$ for the atomic orbitals Ψ_A and Ψ_B respectively, namely the *energies* of these two atomic orbitals. Here β is known as the *resonance* integral, is typically negative, and vanishes in the case where the atomic orbitals do not overlap (for example when A and B are far from each other). Note that S and β differ because in the latter case one has to account for the effect of the Hamiltonian before evaluating the overlap of the wave functions.

Combining all this information together, we find that the energy of the molecule as a function of the coefficients c_A and c_B will be given by

$$E(c_A, c_B) = \frac{c_A^2 \alpha_A + c_B^2 \alpha_B + 2 c_A c_B \beta}{c_A^2 + c_B^2 + 2 c_A c_B S}\,. \tag{7.25}$$

We can now impose the variational principle, Eq. (7.17), to determine the values of c_A and c_B which minimise the total energy of the molecule. Applying these conditions one finds:

$$\begin{aligned} \frac{\partial E}{\partial c_A} &= \frac{(\alpha_A - E)c_A + (\beta - SE)c_B}{c_A^2 + c_B^2 + 2c_Ac_BS} = 0\,, & (7.26) \\ \frac{\partial E}{\partial c_B} &= \frac{(\alpha_B - E)c_B + (\beta - SE)c_A}{c_A^2 + c_B^2 + 2c_Ac_BS} = 0\,, & (7.27) \end{aligned}$$

so therefore the coefficients c_A and c_B must satisfy the following requirements:

$$\begin{aligned} (\alpha_A - E)\,c_A + (\beta - ES)\,c_B &= 0\,, \\ (\beta - ES)\,c_A + (\alpha_B - E)\,c_B &= 0\,. \end{aligned} \tag{7.28}$$

The solution of the above system of linear equations, as can be shown explicitly, is the one that satisfies the following equation:

$$E^2(1 - S^2) + E(2\beta S - (\alpha_A + \alpha_B)) + \left(\alpha_A\alpha_B - \beta^2\right) = 0\,, \tag{7.29}$$

which in turn leads to two solutions for the molecular energy E, corresponding to the *bonding* and the *anti-bonding* orbitals respectively:

$$E_\pm = \left(\alpha_A + \alpha_B - 2\beta S \mp \left[(\alpha_A + \alpha_B - 2\beta S)^2 - 4(1 - S^2)(\alpha_A\alpha_B - \beta^2)\right]^{1/2}\right)/2(1 - S^2)\,. \tag{7.30}$$

Let us study this result for some important cases.

Diatomic homonuclear molecules. First, in the case of homonuclear molecules, we have $\alpha_A = \alpha_B = \alpha$, and the solution simplifies as follows

$$(\alpha_A + \alpha_B - 2\beta S)^2 - 4(1 - S^2)(\alpha_A\alpha_B - \beta^2) = (2\alpha - 2\beta S)^2 - 4(1 - S^2)(\alpha_A^2 - \beta^2) = 4(S\alpha - \beta)^2\,, \tag{7.31}$$

so therefore we have that the general solution is written down as

$$E_\pm = \frac{(2\alpha - 2\beta S) \mp 2(S\alpha - \beta)}{2(1 - S^2)} = \frac{\alpha(1 \mp S) \pm \beta(1 \mp S)}{(1 - S)(1 + S)}\,, \tag{7.32}$$

from which we can obtain the final result:

$$E_\pm = \frac{\alpha \pm \beta}{1 \pm S}\,, \tag{7.33}$$

where the (equal) coefficients in the bonding orbital are

$$c_A = c_B = \left(2(1 + S)\right)^{1/2}\,, \tag{7.34}$$

while for the anti-bonding orbital with energy E_- we have instead

$$c_A = -c_B = \left(2(1 - S)\right)^{1/2}\,. \tag{7.35}$$

This is consistent with the results that we derived in the case of diatomic homonuclear molecules in the context of molecular orbital theory in Sect. 6.4, see in particular Eqns. (6.25) and (6.29). So the variational principle reproduces the general results of Molecular Orbital theory in the case of homo-nuclear molecules.

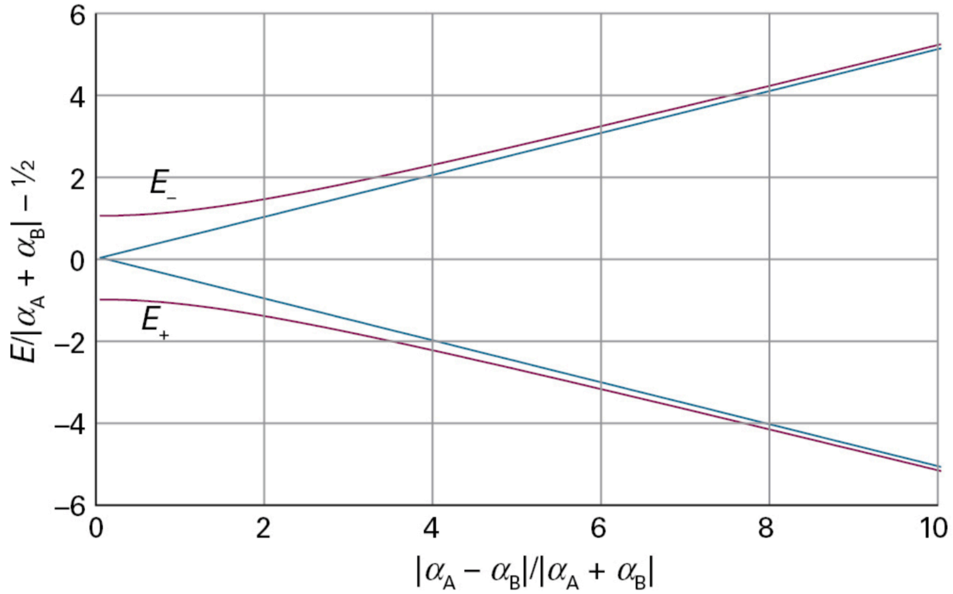


Figure 7.5: The variation of the energies of the molecular orbitals as the energy difference of the contributing atomic orbitals is changed. The plots are for $\beta = -1$; the blue lines correspond to the case where there is no mixing and the resonance integral vanishes, $\beta = 0$ (this is the case known as the zero-orbital approximation).

The zero-overlap approximation. In the more complicated case of heteronuclear molecules, we can simplify the solution by setting the overlap integral to zero, that is, $S = 0$. This is the so-called *zero overlap approximation*. In this case the molecular energies are found to be

$$E_\pm = \frac{1}{2}\left(\alpha_A + \alpha_B\right) \mp \frac{1}{2}|\alpha_A - \alpha_B| \left[1 + \left(\frac{2\beta}{\alpha_A - \alpha_B}\right)^2\right]^{1/2} . \tag{7.36}$$

It is illustrative to plot this solution as a function of $|\alpha_A - \alpha_B|$, which is a measure of how different are the energies of the two atoms that compose the molecule. With this motivation in Fig. 7.5 we plot the following quantity

$$\frac{E_\pm}{|\alpha_A + \alpha_B|} - \frac{1}{2} = \pm\frac{1}{2}\frac{(\alpha_A - \alpha_B)}{|\alpha_A + \alpha_B|}\left[1 + \left(\frac{2\beta}{\alpha_A - \alpha_B}\right)^2\right]^{1/2} \tag{7.37}$$

as a function of $|\alpha_A + \alpha_B|/|\alpha_A - \alpha_B|$. We see that Eq. (7.37) increases (decreases) for E_- (E_+) more or less linearly for large values of its argument. We also see that for large values of the argument of Eq. (7.37), the contribution from the resonance integral β is markedly suppressed. It is also interesting to point out that in the limit $|\alpha_A - \alpha_B| \to 0$, that is, when $\alpha_A = \alpha_B = \alpha$, we have that:

$$\frac{E_\pm}{2\alpha} - \frac{1}{2} = \pm\frac{\beta}{2\alpha} \quad \to \quad E_\pm = \alpha \pm \beta\, . \tag{7.38}$$

The state with energy $E_+ = \alpha - |\beta|$ corresponds to a bonding orbital, while for the state with energy $E_- = \alpha + |\beta|$ corresponds instead to an anti-bonding orbital. Therefore, in this limit we recover again the results of the homonuclear case for the limit in which the overlap integral $S = 0$ vanishes.

Small resonance integral limit. Another interesting configuration corresponds to the limit where $|\alpha_A - \alpha_B| \gg 2|\beta|$. This is the limit in which there is a big energy difference between the two atomic orbitals, as measured in the units of the resonance integral β. In this limit we can perform a Taylor expansion and find the following values for the energies:

$$\begin{aligned} E_+ &= \min(\alpha_A, \alpha_B) - \frac{\beta^2}{|\alpha_A - \alpha_B|}\,, \\ E_- &= \max(\alpha_A, \alpha_B) + \frac{\beta^2}{|\alpha_A - \alpha_B|}\,, \end{aligned} \tag{7.39}$$

from where we see that the absolute values of the energies $|E_\pm|$ is maximal for two similar atomic orbitals $\alpha_A \simeq \alpha_B$. In this approximation, and recalling that we are working in the zero overlap limit, we find the following values for the c_A and c_B coefficients:

$$c_A = \left(1 + \left(\frac{\alpha_A - E_\pm}{\beta}\right)^2\right)^{-1/2}\,, \qquad c_B = \pm\left(\frac{\alpha_A - E_\pm}{\beta}\right) c_A\,. \tag{7.40}$$

which can be evaluated given the values of the energies of the atomic orbitals α_A and α_B and of the resonance integral β.

Example: the HF molecule. In order to illustrate the applications of the variational principle for the calculation of molecular orbitals of heteronuclear diatomic molecules, we can use again the HF molecule from Fig. 7.4. This molecule is characterised by the following numerical values:

- Energy of the atomic orbital A (H_{1s}): $\alpha_A = -13.6$ eV.
- Energy of the atomic orbital B (F_{2p_z}): $\alpha_B = -17.4$ eV.
- Value of the resonance integral: $\beta \simeq -1$ eV.
- Value of the orbital integral: $S \simeq 0$.

Therefore, the difference between the energies of the two molecular orbitals, $\alpha_A - \alpha_B = 3.8$ eV is rather larger than the resonance integral $\beta = -1$ eV and then we can use the "small resonance integral" and "zero overlap" approximations leading to Eq. (7.39), which leads to

$$E_+ \simeq -17.4\text{ eV} - \frac{1\text{ eV}^2}{3.8} \simeq -17.6\text{ eV}\,, \tag{7.41}$$

$$E_- \simeq -13.6\text{ eV} + \frac{1\text{ eV}^2}{3.8} \simeq -13.4\text{ eV}\,, \tag{7.42}$$

consistent with the values for the energies of the molecular orbitals quoted in Fig. 7.4.

Concerning corresponding values of the coefficients c_A and c_B, using the results of the variational principle and this two approximations one has for the ground state (molecular bonding orbital):

$$c_A = \left(1 + \left(\frac{\alpha_A - E_+}{\beta}\right)^2\right)^{-1/2} = \left(1 + \left(\frac{-13.6 + 17.6}{1}\right)^2\right)^{-1/2} = 0.243\,, \tag{7.43}$$

$$c_B = -\left(\frac{\alpha_A - E_+}{\beta}\right) c_A = +\left(\frac{-13.6 + 17.6}{1}\right) 0.243 = 0.97\,. \tag{7.44}$$

which again are the same values as those quoted in Fig. 7.4. Therefore we have used the variational principle to construct the ground state of the HF molecule, and it is given by

$$\Psi_{\mathrm{HF},+} = 0.24\,\Psi_{H_{1s}} \, + 0.97\,\Psi_{F_{2p}} \, . \tag{7.45}$$

You can check that the correct results are also reproduced by the anti-bonding orbital $\Psi_{\mathrm{HF},-}$ of this molecule.

In summary, the variational method is a powerful tool to

Summary

To summarize, some important concepts that we have learned in this lecture about the electronic structure of homo-nuclear and hetero-nuclear molecules are the following:

- The filling order of molecular orbitals in diatomic molecules follows similar rules as in the case of the filling order of multi-electron atoms, and in particular it has associated an specific version of the building-up principle (and Hund's rule also applies).
- The difference between the number of bonding and anti-bonding orbitals determines the chemical stability of a molecule, as well as the strength of the associated chemical bond.
- In the case of hetero-nuclear diatomic molecules, the coefficients c_A and c_B that multiply the two atomic orbitals Ψ_A and Ψ_B will in general be different due to the different electro-affinity of the two elements.
- The variational principle can be used to determine the coefficients of the linear combination of atomic orbitals composing an heteronuclear diatomic molecule.

In the next lecture we complete with our study of the chemical bonding of molecules in the framework of Molecular Orbital theory by considering more complex molecules, in particular molecules composed by more than two atoms.

Further reading

Further reading material about the topics covered in this lecture can be found in the following sections of the course textbook (*Physical Chemistry*, Atkins and De Paula, 10th edition):

- **10A.2**: *Polyatomic molecules.*
- **10B.2**: *Orbital notation.*
- **10C.1**: *Electron configurations.*
- **10D.1**: *Polar bonds.*
- **10D.2**: *The variation principle.*

van Quantum tot Molecuul

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

8 HC9: Chemical bounds and molecules III

In this lecture, we complete our study of chemical bonding in molecules by considering more complex molecules beyond the diatomic configuration. Specifically, we will consider here poly-atomic molecules, defined as molecules composed by more than two atoms (in general of a different type). We will present calculational techniques allowing us to construct molecular orbitals and determine their energies for some specific types of complex molecules.

The **learning goals** of this lecture are:

(a) Extending the methods of Molecular Orbital theory to poly-atomic molecules.

(b) Expressing the solutions of the variational principle as a linear algebra matrix problem.

(c) Applying the Huckel approximation to the description of molecular energies of complex molecules such as benzene.

(d) Determine the main physical properties of the weak molecular bond known as hydrogen bridges.

8.1 Molecular orbitals for poly-atomic molecules with the Huckel method

We start this lecture by presented how we can construct the molecular orbitals for poly-atomic molecules (composed by more than two atoms). Building upon the same approach that we developed previously in our study of molecular orbitals for the case of diatomic molecules, here we will also assume that the molecular wave function can be constructed as a *linear combination of atomic orbitals* (LCAO) of the form:

$$\Psi = \sum_{k=1}^{N} c_k \Psi_k \,, \tag{8.1}$$

with N the number of atoms that compose a molecule, and Ψ_k the corresponding atomic orbitals. One difficulty with this approach is that it becomes very cumbersome for $N \gg 2$, and in addition it is not straightforward to generalize the variational principle to poly-atomic molecules.

In this context, it is useful to introduce the so-called *Huckel method*. This approach is a variation of the LCAO method to construct molecular orbitals for a specific type of poly-atomic molecules:

> The Huckel method can be applied to the π-orbitals of *conjugated hydrocarbon molecules*. By conjugated molecules we denote molecules which contain a system of connected p orbitals with delocalised electrons, and hydrocarbons are those molecules composed entirely of C and H atoms.

The basic idea of the Huckel method is that the π-type molecular orbitals should be treated independently from the σ-type ones. Recall here that the σ-type orbitals are such that they determine the overall shape of the molecule.

Let us illustrate how the Huckel approach works with an explicit example, namely the C_2H_4 molecule (ethene). Recall than in the context of Valence Bond theory (Sect. 6.3), this molecular orbital was explained from a sp^2-type hybridisation leading to a σ-type C-C bond and a π-type C-C bond. Therefore, the two carbon atoms of the ethene molecule are linked by means of a double bond. In the context of Molecular Orbital theory, we can now construct the following wave function

$$\Psi = c_A \Psi_A + c_B \Psi_B \, , \tag{8.2}$$

corresponding to the π orbitals, so Ψ_A and Ψ_B correspond here to the $2p$ orbitals of the two C atoms. We can now use the variational principle, see Sect. 7.3, in order to determine the energies of the molecular orbitals, using the fact that $\alpha_A = \alpha_B$ (since the bond involves two identical atoms). As we saw there, see Eq. (7.28), the values of the coefficients c_A, c_B and thus of the energies $E_\pm$ are determined by the solution of the following equation:

$$\begin{vmatrix} \alpha - E & \beta - ES \\ \beta - ES & \alpha - E \end{vmatrix} = 0 \, , \tag{8.3}$$

where S was the overlap integral and β the resonance integral, Eq. (7.24). Eq. (8.3) is a quadratic equation that can be solved analytically, but the corresponding equations involving ≥ 3 atoms are rather more complicated. Therefore if we want to construct molecular orbitals for more complex molecules we need to introduce a number of approximations.

Assumptions in the Huckel method. The Huckel method to solve Eq. (8.3) is based on the following set of approximations, which as we will show then generalise to more complex molecules:

- Set to zero all the overlap integrals, $S = 0$.
- Set to zero all the resonance integrals, $\beta = 0$, for all the atoms that are *not nearest-neighbors* among them.
- The resonance integrals for neighboring atoms take always the same value, β, irrespective of the specific type of atom.

While these are quite drastic approximations, they still provide a useful way to estimate the structure and energies of the molecular orbitals of poly-atomic molecules without too much calculational effort.

By means of these approximations, the energies of the ethene molecule computed with the variational principle Eq. (8.3) simplify to the following”

$$\begin{vmatrix} \alpha - E & \beta - ES \\ \beta - ES & \alpha - E \end{vmatrix} = 0 \quad \Rightarrow \quad \begin{vmatrix} \alpha - E & \beta \\ \beta & \alpha - E \end{vmatrix} = (\alpha - E)^2 - \beta^2 = 0\,, \tag{8.4}$$

which leads to the result that we also found in the case of the diatomic molecules, namely

$$E_\pm = \alpha \pm \beta\,, \tag{8.5}$$

see Eq. (7.33) and the fact that we are imposing that the overlap integral vanishes, $S = 0$. If one recalls that since the resonance integral β is negative ($\beta < 0$), E_+ corresponds to the *bonding* orbital while E_- corresponds instead to the *anti-bonding* orbital. Note that the Huckel approach involves a fair degree of simplification: all diagonal elements become equal to $\alpha - E$, all the non diagonal elements between neighboring atoms become β, and all other elements in the matrix are set to zero. This kind of linear algebra systems are rather easier to solve than the original systems. Within this approximation, the molecular orbitals of ethene are given by

$$\Psi_\pm = \frac{1}{\sqrt{2}}\left[\Psi_{2p_z}(1) \pm \Psi_{2p_z}(2)\right]\,. \tag{8.6}$$

The usefulness of the Huckel approximation is that it allows to compute the molecular orbitals of poly-atomic molecules in a simplified way.

8.2 Matrix formulation of Huckel theory

In full generality, the Huckel theory to construct molecular orbitals can be framed as a linear algebra matrix problem. Let us show how this is the case. Recall that the variational principle (Sect. 7.3) lead to the following system of equations for the coefficients c_A and c_B for a heteronuclear diatomic molecule:

$$\begin{aligned} &(\alpha_A - E)\,c_A + (\beta - ES)\,c_B = 0\,, \\ &(\beta - ES)\,c_A + (\alpha_B - E)\,c_B = 0\,. \end{aligned} \tag{8.7}$$

Now, let us change notation and express these equations as follows:

$$\alpha_J = H_{JJ}\,, \qquad \beta = H_{AB}\,, \qquad S_{AB} = S\,, \qquad ES_{JJ} = E\,, \tag{8.8}$$

where the last substitution applies only to the diagonal entries. Note that here we are simply multiplying by one, since the diagonal overlap integral satisfies

$$S_{JJ} = \int \Psi_J^* \Psi_J = 1\,, \tag{8.9}$$

due to the normalisation of the atomic orbitals. Using this notation, the system of equations determined from the variational principle now reads

$$\begin{aligned} &(H_{AA} - ES_{AA})\,c_A + (H_{AB} - ES_{AB})\,c_B = 0\,, \\ &(H_{BA} - ES_{BA})\,c_A + (H_{BB} - ES_{BB})\,c_B = 0\,. \end{aligned} \tag{8.10}$$

The rationale for these substitutions is that the underlying symmetries of the system of equations are now more manifest. There are two sets of solutions of the above system of equations, which we denote by

(c_{1A}, c_{1B}) and (c_{2A}, c_{2B}). With this notation, this system of equations can be written as a matrix equation in a rather compact way:

$$\begin{pmatrix} H_{AA} - E_i S_{AA} & H_{AB} - E_i S_{AB} \\ H_{BA} - E_i S_{BA} & H_{BB} - E_i S_{BB} \end{pmatrix} \begin{pmatrix} c_{i,A} \\ c_{i,B} \end{pmatrix} = 0 \, , \tag{8.11}$$

which can be written as the following matrix equation

$$\widehat{H} c_i = \widehat{S} c_i E_i \tag{8.12}$$

where we have defined the matrices

$$\widehat{H} \equiv \begin{pmatrix} H_{AA} & H_{AB} \\ H_{BA} & H_{BB} \end{pmatrix} , \qquad \widehat{S} \equiv \begin{pmatrix} S_{AA} & S_{AB} \\ S_{BA} & S_{BB} \end{pmatrix} , \qquad c_i = \begin{pmatrix} c_{i,A} \\ c_{i,B} \end{pmatrix} . \tag{8.13}$$

Note that both $\widehat{S}$ and $\widehat{H}$ are symmetric matrices.

We can express this condition in a yet more compact way by defining the following matrix of coefficients:

$$\widehat{c} = \begin{pmatrix} c_{1,A} & c_{2,A} \\ c_{1,B} & c_{2,B} \end{pmatrix} , \tag{8.14}$$

as well as the matrix of energy values

$$\widehat{E} = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} . \tag{8.15}$$

With these definitions, we can write Eq. (8.12) in terms of only matrices:

$$\widehat{H} \, \widehat{c} = \widehat{S} \, \widehat{c} \, \widehat{E} \, . \tag{8.16}$$

At this point we are ready to simplify this linear algebra problem by using the Huckel assumptions presented in Sect. 8.1, namely:

- Set to zero all overlap integrals $S = 0$.
- Set to zero the resonance integrals, $\beta = 0$, except for atoms that are nearest neighbors.
- The resonance integral β is taken to be the same for all neighboring atoms.

In matrix notation, these approximations correspond to setting $H_{AB} = H_{BA} = \beta$ and $\widehat{S}$ to the unity matrix. Therefore Eq. (8.16) simplifies to

$$\widehat{H} \, \widehat{c} = \widehat{c} \, \widehat{E} \quad \Rightarrow \quad \widehat{c}^{-1} \, \widehat{H} \, \widehat{c} = \widehat{E} \, , \tag{8.17}$$

where we have multiplied by $\widehat{c}^{-1}$ the equation from the left, and used the fact that $\widehat{c}^{-1} \, \widehat{c} = \mathbb{1}$. Eq. (8.17) has the form of an *eigenvalue problem*: if I apply the Hamiltonian operator for my system $\widehat{H}$ to the matrix of coefficients $\widehat{c}$ I recover the same multiplied by the energies of the molecular orbitals $\widehat{E}$. Therefore, the solutions for $\widetilde{E}$ are determined from the transformation that diagonalises the Hamiltonian matrix $\widehat{H}$. Since our goal is to find the energies $\widehat{E}$ that characterize the molecular orbitals of poly-atomic molecules, we need to diagonalise first the Hamiltonian of the molecule expressed in matrix form as indicated by Eq. (8.17).

8.3 Huckel theory for poly-atomic orbitals

Although we have derived Eqns. (8.16) and (8.17) in the case of a molecule with two orbitals, the same approach can be used for more complex molecules composed by more than two atoms. To illustrate this point, let us consider the butadiene molecule C_4H_6, schematically represented in Fig. 8.1. Here we have four carbon atoms, and therefore the molecular Hamiltonian will be a 4×4 matrix. Within the approximations of Huckel's theory, we can simplify this Hamiltonian matrix as follows:

$$\widehat{H} = \begin{pmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} & H_{34} \\ H_{41} & H_{42} & H_{43} & H_{44} \end{pmatrix} \simeq \begin{pmatrix} \alpha & \beta & 0 & 0 \\ \beta & \alpha & \beta & 0 \\ 0 & \beta & \alpha & \beta \\ 0 & 0 & \beta & \alpha \end{pmatrix} = \alpha \mathbb{1} + \beta \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} . \tag{8.18}$$

Note that the entries of the Hamiltonian matrix are ordered sequentially: say the first row is the leftmost carbon atom, while the last row corresponds to the rightmost carbon atom. In Fig. 8.1 we show the graphical correspondence with Eq. (8.1), showing how in the Hamiltonian matrix of this molecule we will have four factors of α and six factors of β. Note that the left-most and right-most C atom carry a single factor of β since they have bonds with only one other C atom.

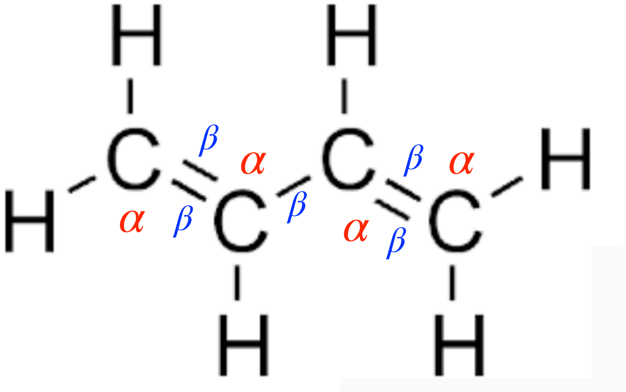


Figure 8.1: The chemical structure of the butadiene molecule, C_4H_6. Each carbon atom has associated a factor α corresponding to the atomic orbital energy. Each bond between two carbon atoms has associated a factor β from the resonance integral. Therefore in the Hamiltonian matrix of this molecule we will have four factors of α and six factors of β, as can be seen in Eq. (8.18).

The diagonalisation of the matrix in Eq. (8.1) then leads to the values for the energies of the four molecular orbitals of butadiene (recall that the C — H bonds are treated separately). One can show either analytically or numerically that the diagonalisation of this Hamiltonian matrix leads to the following solutions for the orbital energies:

$$E = \begin{pmatrix} \alpha + 1.62\beta & 0 & 0 & 0 \\ 0 & \alpha + 0.62\beta & 0 & 0 \\ 0 & 0 & \alpha - 0.62\beta & 0 \\ 0 & 0 & 0 & \alpha - 1.62\beta \end{pmatrix} , \tag{8.19}$$

where remember that $\beta \leq 0$ and thus the first entry of the matrix corresponds to the molecular orbital with the lowest energy (in other words, to the molecular ground state). This molecular electronic structure of the butadiene molecule C_4H_6 in the framework of Huckel theory is represented in Fig. 8.2. An example of a `Python` program to diagonalise the Hamiltonian matrix is

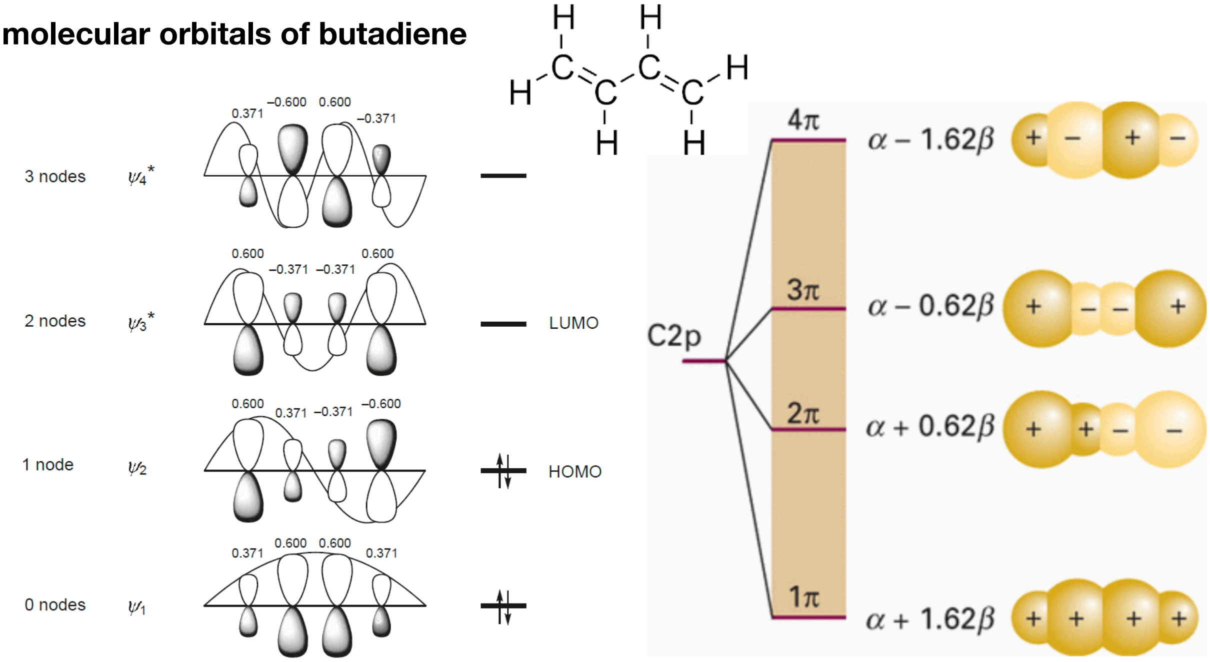


Figure 8.2: The molecular electronic structure of the butadiene molecule, C_4H_6, in the framework of Huckel theory, see text for more details. The π-type molecular orbitals are constructed from the $2p$ orbitals of carbon. Since the resonance integral $\beta < 0$, the molecular orbital with energy $\alpha + 1.62\,\beta$ corresponds to the ground state. The lower the orbital energy (1π orbital), the more constructive the interference between atomic orbitals is. For the anti-bonding orbital 4π, the interference between atomic orbitals is maximally destructive.

```
from numpy import linalg as LA
# define matrix of resonance integrals
beta = [[0, 1, 0, 0],
        [1, 0, 1, 0],
        [0, 1, 0, 1],
        [0, 0, 1, 0]]
# Solve the eigensystem
w, v = LA.eig(beta)
# print eigenvalues and eigenvectors
print("\n eigenvalues = ")
print(w)
print("\n eigenvectors = ")
print(v)
```

The matrix $\widehat{c}$ that diagonalises the Hamiltonian then contains the coefficients that allow constructing the wave functions of the molecular orbitals. You can check from the explicit expression of the matrix $\widehat{c}$ (composed by the eigenvectors of the system) that the resulting molecular orbitals are given by, in order of increasing energy E_i, by the following linear combinations of atomic orbitals:

$$\begin{aligned}
\Psi_1 &= 0.372\Psi_A + 0.602\Psi_B + 0.602\Psi_C + 0.372\Psi_D \,, \\
\Psi_2 &= 0.602\Psi_A + 0.372\Psi_B - 0.372\Psi_C - 0.602\Psi_D \,, \\
\Psi_3 &= 0.602\Psi_A - 0.372\Psi_B - 0.372\Psi_C + 0.602\Psi_D \,, \\
\Psi_4 &= -0.372\Psi_A + 0.602\Psi_B - 0.602\Psi_C + 0.372\Psi_D \,,
\end{aligned} \tag{8.20}$$

which are represented schematically in Fig. 8.2, and where A, B, C, D indicate the four carbon atoms of butadiene. We find that in the case of the 1π molecular orbital, which corresponds to that with smallest energy, all the p-type atomic orbitals interfere in a constructive way. In the case of the 4π anti-bonding molecular orbital, there is instead a negative interference.

This discussion allows us to gain more insight of the electronic structure of the butadiene molecule. We have found that this molecule contains 4×3 σ-type bonds (between the C — H system) and 4 π-type bonds (delocalised along the four C atoms). Therefore, the resulting electronic configuration for the latter can be written as $1\pi^2 2\pi^2$. In this context one often uses the following terminology:

- HOMO stands for *Highest Occupied Molecular Orbital*. In the case of the butadiene molecule, the HOMO is the 2π bonding orbital.
- LUMO stands for *Lowest Unoccupied Molecular Orbital*. In the case of the butadiene molecule, the LUMO is the 3π anti-bonding orbital.

If we compute the total energy of the occupied π bonds of the butadiene molecule, taking into account that each orbital fits two electrons, we find that it is given by

$$E_\pi = 2\left(\alpha + 1.62\beta\right) + 2(\alpha + 0.62\beta) = 4\alpha + 4.48\beta < 4\alpha \,, \tag{8.21}$$

which is smaller than the energy of the original atomic orbitals, thus explaining the stability of the underlying chemical bonds (recall that the resonance integral satisfies $\beta < 0$): the molecular configuration of benzene leads to a smaller total energy than the four C atoms separately.

The case of benzene. Let us consider, as another example of the applications of Huckel theory, the benzene molecule C_6H_6, illustrated in Fig. 8.3. In this molecule, the σ-type molecular bonds can be explained by the valence bond theory, while the π-type bonds will be described by Huckel theory. From the geometric (cyclic) structure of the benzene molecule, see Fig. 8.3, we know that the Hamiltonian, after applying the Huckel approximations, will be given by

$$\hat{H} = \begin{pmatrix} \alpha & \beta & 0 & 0 & 0 & \beta \\ \beta & \alpha & \beta & 0 & 0 & 0 \\ 0 & \beta & \alpha & \beta & 0 & 0 \\ 0 & 0 & \beta & \alpha & \beta & 0 \\ 0 & 0 & 0 & \beta & \alpha & \beta \\ \beta & 0 & 0 & 0 & \beta & \alpha \end{pmatrix} = \alpha \mathbb{1} + \beta \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} , \tag{8.22}$$

where now for instance we have a β in the last column of the first row since the C atoms in benzene form a closed structure (since it is a cyclic molecule). The diagonalisation of the matrix that multiplies the resonance integral β has as eigenvalues ± 2 and ± 1, where the latter pair of eigenvalues is repeated. Therefore the energies for the six molecular orbitals will be given by:

$$E = \alpha \pm 2\beta\,, \quad E = \alpha \pm \beta\,, \quad E = \alpha \pm \beta\,. \tag{8.23}$$

This result indicates, as illustrated also in Fig. 8.3, that the a_{2u} and the e_{1g} orbitals are completely filled, and thus the total energy associated with the π orbitals is

$$E_\pi = 2(\alpha + 2\beta) + 4(\alpha + \beta) = 6\alpha + 8\beta\,, \tag{8.24}$$

to be compared to the corresponding result for three ethene molecules (C_2H_4) where the energy contained in the π molecular orbitals would be

$$3 \times (2\alpha + 2\beta) = 6\alpha + 6\beta\,, \tag{8.25}$$

showing that the formation of the benzene molecule is energetically favourable.

Electron delocalisation. In the context of *conjugated molecules*, one denotes the *delocalisation* of the electrons in given molecule as the fact that electrons in π orbitals are free to move around and are not localised in a fixed bond. The greater the degree of delocalisation of a molecule, the greater their stability. So in the case of the benzene molecule its delocalisation energy is given by the difference in the π orbitals of benzene as compared to ethene, namely

$$E_{\rm delocalisation} = 2\beta \simeq -460\,{\rm kJ/mol}\,. \tag{8.26}$$

The characteristic ring-like structure of benzene shown in Fig. 8.3 indicates that it is an *aromatic* molecule. In general, an organic molecule is said to be aromatic if it contains a planar ring formed by $4n+2$ π-electrons with $n \geq 0$. As another example, in the butadiene molecule represented in Fig. 8.1 the electrons belonging to the π-type orbitals are free to move along the chain defined by the four C atoms.

To summarise the discussions in this section, Huckel theory is:

> a semi-empirical method to calculate the molecular orbitals of the π type in various types of molecules. It is based on rather dramatic simplifications of the Hamiltonian of the systems, which neglects important effects but allows to explain quantitatively the main features of these bonds.

8.4 Hydrogen bonds

We conclude this discussion of the chemical bonds in molecules by explaining the important case of *hydrogen bonds*, also known as *hydrogen bridges*. They arise due to several factors:

- the electrostatic attraction between two polar groups,
- that occurs when a hydrogen atom covalently bound to a highly electro-negative atom such as nitrogen or oxygen,
- experiences the electrostatic field of another highly electro-negative atom nearby.

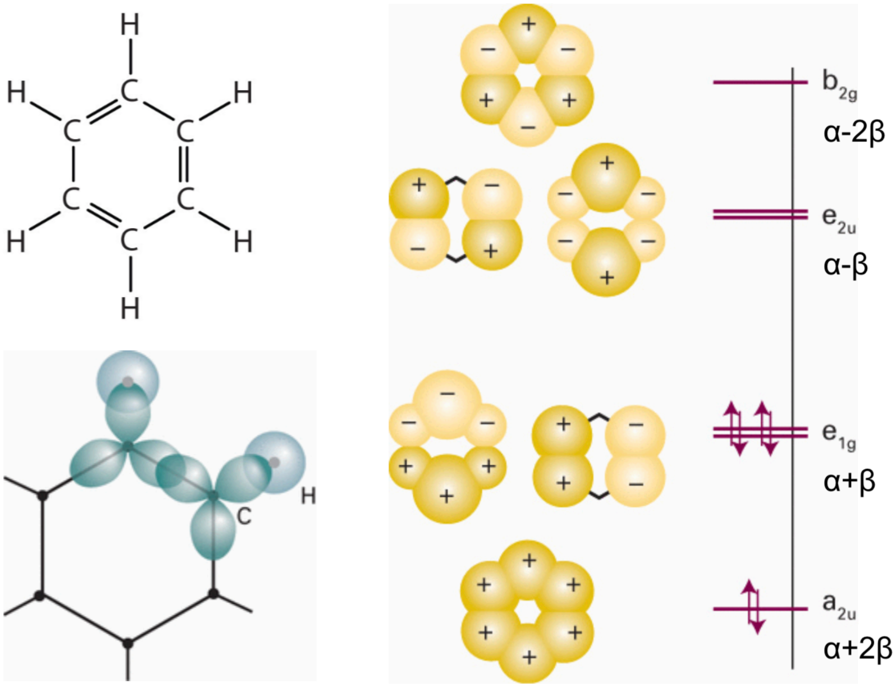


Figure 8.3: The chemical structure of the benzene molecule, C_6H_6. This molecule is composed by σ bonds between the C and H atoms and π bonds between the p-type atomic orbitals in C. In the right plot we show the resulting electronic molecular structure according to Huckel theory.

It is an attractive type of interaction between the A-H molecule and B, where A and B are electro-negative and B has *a lone pair*.

We define as *lone pairs* as a pair of electrons that form part of an atomic orbital not involved in chemical bonding. An important example are the lone pairs in O within an H_2O molecule, indicated in blue in Fig. 8.4. These are four electrons that occupy two of the sp^3 hybridised orbitals that do not participate in the bonding between O and H.

Examples of electro-negative elements (that is, with a high power to attract electrons) are N, O, F. These are therefore elements that that construct molecules that exhibit hydrogen bridges. In terms of its strength, a hydrogen bond is rather weaker than a covalent bond: $\simeq 20$ kJ/mol as compared to $\simeq 400$ kJ/mol respectively. It is still stronger that thermal fluctuations, where $k_BT \simeq 2.5$ kJ/mol under ambient conditions, and therefore it is energetically favourable. Note also that the hydrogen bond is a di-polar interaction: the partially positive H atom binds with the partially negative B atom. Some examples of hydrogen bonds between different elements are shown in the left panel of Fig. 8.5. For example, an O — H bond can form a hydrogen bridge with another O atom, as does happen in water.

These hydrogen bonds (or hydrogen bridges) can be described by means of the Molecular Orbital theory. Following the same notation as in the rest of the lectures, we will denote as

- Ψ_A is the atomic orbital of element A (which has high electro-negativity).
- Ψ_H is the $1s$ orbital of hydrogen,

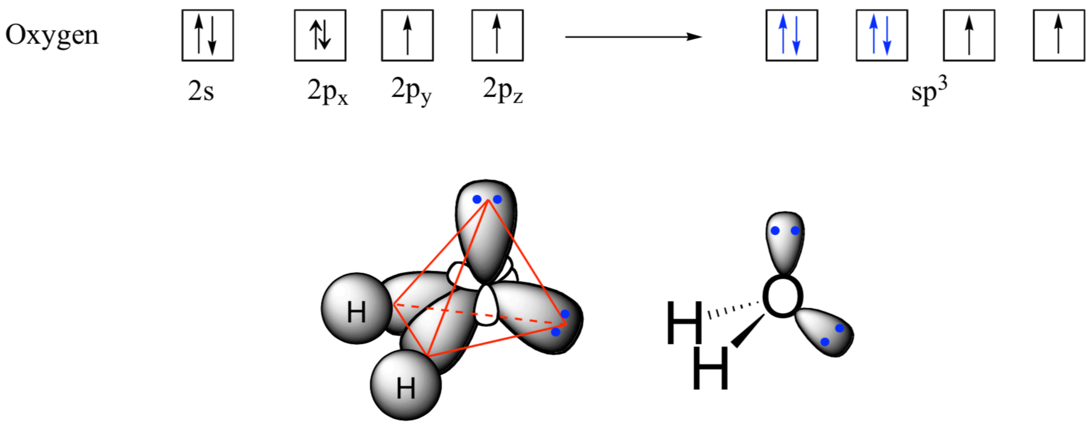


Figure 8.4: We define as *lone pairs* as a pair of electrons that form part of an atomic orbital not involved in chemical bonding. An important example are the lone pairs in O within an H_2O molecule, indicated in blue. These are four electrons that occupy two of the sp^3 hybridised orbitals that do not participate in the bonding between O and H.

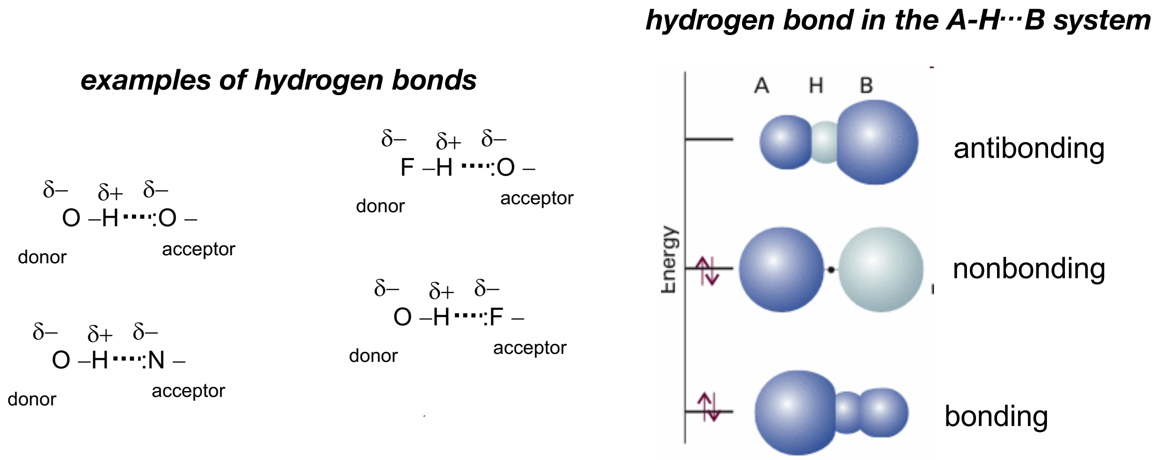


Figure 8.5: Left plot: examples of hydrogen bonds establishes between different elements. In all cases the *acceptor* atom must have both high electro-negativity (indicated by the δ^- sign) and contain a lone pair, see also Fig. 8.4 Right plot: the typical energies of the molecular orbitals involved in the formation of an hydrogen bridge. We find three orbitals: the bonding, the non-bonding, and the anti-bonding orbital.

- Ψ_B is the orbital corresponding to the lone pair of B, which is involved in the formation of the hydrogen bridge.

The wave function of the hydrogen bridge will then be written as a linear combination of the atomic orbitals involved in the bond, namely

$$\Psi = c_1\Psi_A + c_2\Psi_H + c_3\Psi_B \, . \tag{8.27}$$

As indicated in Fig. 8.5, we will have four electrons in the molecular orbitals corresponding to this hydrogen bond. There will be in particular three orbitals: the bonding, the non-bonding, and the anti-bonding orbital. Of the four electrons that participate in the hydrogen bond, two come from the A-H molecule and 2 corresponds to the lone pair of the B element. The bonding and non-bonding molecular orbitals are then completely filled with electrons, leading to an overall bonding interaction (although much weaker than the covalent bond).

The hydrogen bridge can be understood as a contact interaction between the electron clouds of the A-H and B systems. Therefore, for this bond to have an appreciable effect the A — B system must be reasonably close to the atom B. In a covalent bond, the typical distance between the atoms involved in the bind is around 1.3 Armstrong. But in a hydrogen bond (or hydrogen bridge), the distance between A — B and B is larger, typically around 3 Armstrong. If the separation between A — B and B becomes larger than 3.5 Armstrong, then no hydrogen bridge can be created. So therefore only we need a relatively dense medium in order that the rate of formation of hydrogen bounds becomes significant enough.

Summary

To summarize, some important concepts that we have learned in this lecture about the electronic structure of heteronuclear molecules.

- The methods of Molecular Orbital theory can be extended to poly-atomic molecules and be expressed as a linear algebra matrix problem.
- However the expressions become quickly cumbersome and one needs to introduce approximations to make them solvable. In the Huckel approximation, overlap integrals are set to zero and also the resonance integrals except for atoms which are nearest neighbours.
- The Huckel method can be applied to the calculation of of molecular energies for complex molecules such as benzene.
- Hydrogen bridges are weak molecular bounds that are formed between an A — H system, where A has high electro-negativity, and B, where B also has high electro-negativity and in addition contains a lone pair.

Further reading

Further reading material about the topics covered in this lecture can be found in the following sections of the course textbook (*Physical Chemistry*, Atkins and De Paula, 10th edition):

- **10E.1**: The Huckel approximation.
- **10E.2**: Applications.

Having successfully completed our study of chemical bonding in molecules from the point of view of quantum theory, in the rest of the lectures we will study the interaction of molecules and nuclei with different types of electromagnetic radiation. This will include a discussion of other types of dynamics that can take place in molecules beyond the motion of the electrons, in particular the rotational and vibrational spectra of atomic nuclei within molecules.

9 HC10: Molecular and vibrational spectroscopy

van Quantum tot Molecuul

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

HC10 is the first of the last three lectures of the course. In this final part, we will discuss the *interaction of molecules and nuclei with different types of electromagnetic radiation*, in what is often called light-matter interaction. First of all, in HC10 we will discuss the general properties of *molecular spectroscopy and vibrational spectra*, where by *spectroscopy* we understand the study of the interaction between electromagnetic radiation and matter using different experimental techniques. Then in HC11 we will study electronic transitions in molecules. Finally, in HC12 we will illustrate the underpinnings of Nuclear Magnetic Resonance (MNR), highlighting how quantum physics is central to one of the most widely used methods of medical imaging.

The *learning goals* of this lecture are:

- Understanding and applying the basic properties of absorption and emission of electromagnetic radiation in molecules.
- Becoming familiar with the vibrational states and the corresponding transitions among them for di- and poly-atomic molecules.
- Determining how vibrational transitions lead the interactions with infrared light via Raman scattering, and learn how to exploit this knowledge for practical applications.

Prelude: the electromagnetic wave spectrum. Since in this lectures we are going to study different types of electromagnetic radiation, it is useful to recap the main features of the electromagnetic wave spectrum, summarised in Fig. 9.1. At the quantum level, electromagnetic waves are represented by photons, the quanta of light. Photons with short (long) wavelengths λ correspond to photons with large (small) frequencies ν and energies E_γ, since $E_\gamma = h\nu = hc/\lambda$. The visible light occupies just a small region of the full EM spectrum. Radio waves correspond to very long wavelengths (meters or kilometers) and low energies, while gamma rays have the highest energies and shortest wavelength (at the level of interatomic distances).

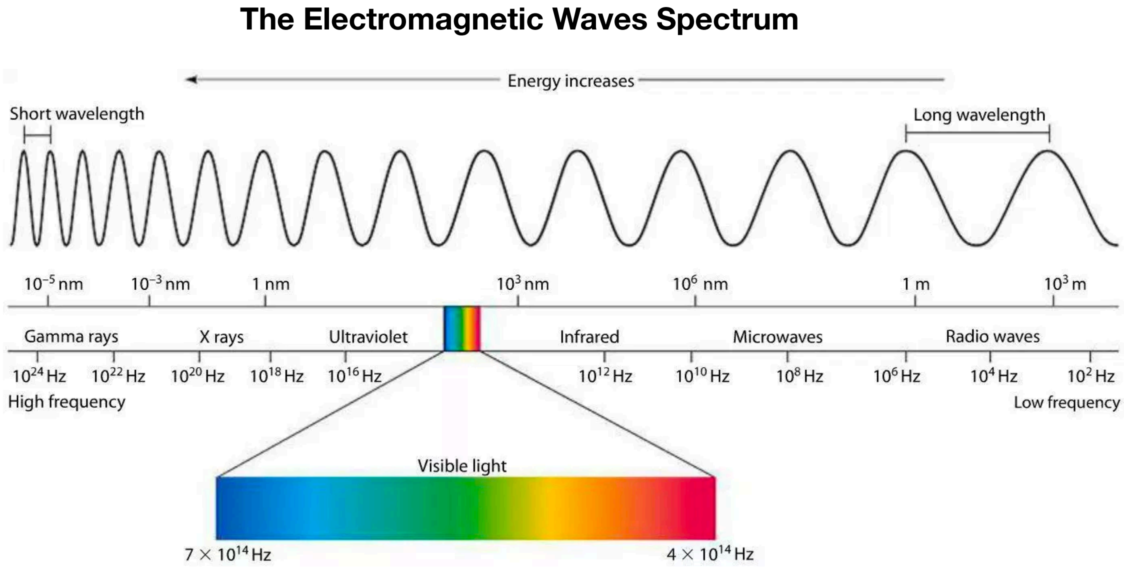


Figure 9.1: The electromagnetic wave spectrum. Photons with short (long) wavelengths λ correspond to photons with large (small) frequencies ν and energies E_γ, since $E_\gamma = h\nu = hc/\lambda$. The visible light occupies just a small region of the full EM spectrum.

9.1 Basics of light-matter interaction: emission and absorption

We start this lecture by presenting the basic aspects of the interactions between light and molecules. As mentioned above, the term *spectroscopy* stands for the study of the interactions between electromagnetic radiation (photons) and matter. This interaction can take place in two ways:

- by *absorption* of electromagnetic radiation by matter,
- or by the complementary process, the *emission* of electromagnetic radiation. This emission can be either *spontaneous* or *stimulated*, as we explain below.

Let us start by considering a *two-level system*, represented in Fig. 9.2. This is a quantum system in which a given particle, say an electron in a molecular orbital, can only occupy *two quantum states*, one with energy E_i, called the *ground state*, and the other with energy E_f with $E_f > E_i$, called the *excited state*. As discussed in first lecture, the interactions between light and matter are mediated by:

> *photons*, which are the quanta of electromagnetic radiation. At the quantum level, the interactions between light and electrons/molecules always implies the emission and absorption of photons of a given energy. Note that in these processes energy must always be conserved.

In this two-level system, one can have the following processes from the interaction between light and matter:

- The *spontaneous emission* of electromagnetic radiation (in the absence of external radiation) will take place via the transition $E_f \to E_i$, that is, where the system moves from the excited state to the ground state. By energy conservation, the energy of the emitted photon will be

  $$E_\gamma = h\nu = E_f - E_i\,, \quad \text{and thus} \quad \nu = (E_f - E_i)/h \tag{9.1}$$

  will the the frequency of the emitted radiation by this system.

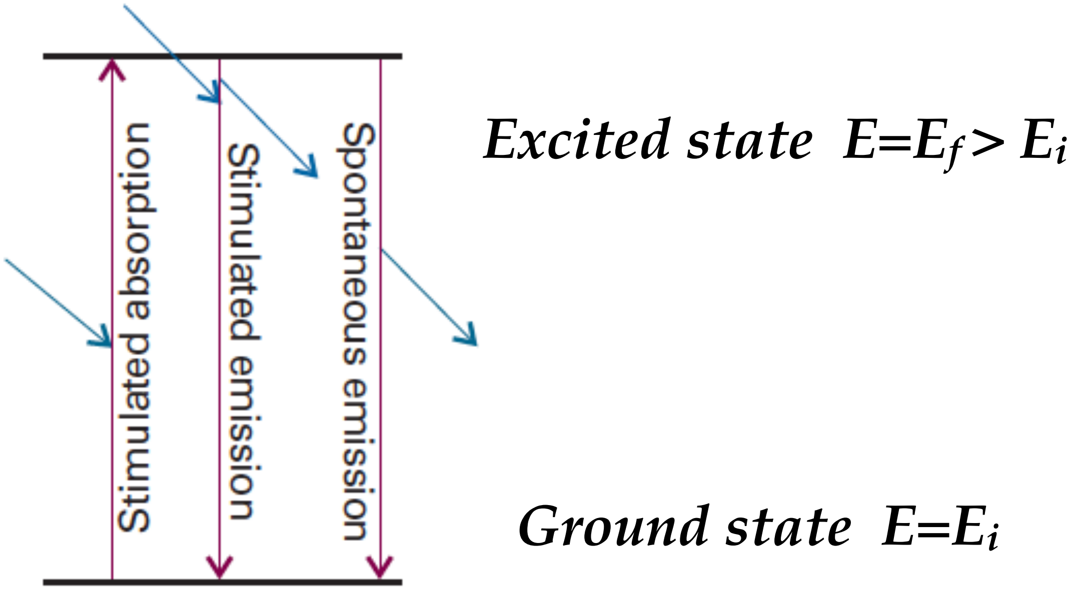


Figure 9.2: Schematic representation of the possible interactions between light and matter for a two-level quantum system, defined by a ground state with energy E_i and an excited state with energy $E_f > E_i$. One can have (i) stimulated absorption of external radiation, (ii) the stimulated emission following absorption of radiation, and (iii) the spontaneous emission of radiation.

- The complementary process to spontaneous emission is that of the *stimulated absorption* of a photon with energy $E_\gamma = (E_f - E_i)$, which excites the system from the ground state E_i to the excited state E_f. This is process has an associated *transition rate* given by

  $$w_{f\leftarrow i} = B_{fi}\,\rho(\nu)\,, \tag{9.2}$$

  where $w_{f\leftarrow i}$ stands for the number of transitions $i \to f$ which take place per unit time and per molecule in the system, B_{fi} is known as the *Einstein coefficient* for stimulated absorption, and $\rho(\nu)$ is the *spectral energy density* of the isotropic radiation field at the frequency ν of the transition, in other words, the density of states available for a photon of frequency ν. From Eq. (9.2), we can derive that the *total number of stimulated absorptions* $W_{f\leftarrow i}$ that take place in our system is

  $$W_{f\leftarrow i} = N_i\,w_{f\leftarrow i} = N_i\,B_{fi}\,\rho(\nu)\,, \tag{9.3}$$

  where N_i is the number (or its density) of molecules in the system in the ground state.

- The spontaneous emission of photons is not the only possible way to induce a $f \to i$ transition in our system. In the presence of external electromagnetic radiation (incoming photons) in our two-level system, as illustrated in Fig. 9.3, an *stimulated emission* of a photon will also take place, and the associated transition rate is

  $$w_{f\to i} = B_{if}\,\rho(\nu)\,, \tag{9.4}$$

  where B_{if} is the *Einstein coefficient for stimulated emission.* Note that the difference between spontaneous and stimulated emission is that in the latter case the system of molecules is already in the presence of electromagnetic radiation with frequency Eq. (9.1). In contrast, *spontaneous emission* of a

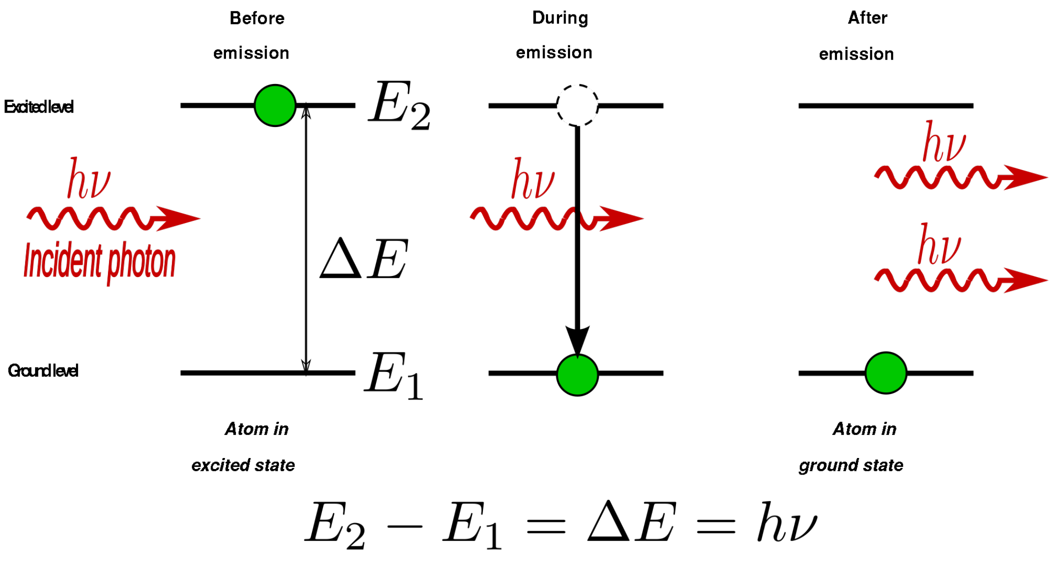


Figure 9.3: Stimulated emission of electromagnetic radiation in a two-level quantum system. The system is found in the excited state E_2, and as result of the interaction with an incident photon with energy $E_\gamma = h\nu = E_2 - E_1$, the system ends up in the ground state and *two photons* with energy E_γ are emitted from the system.

photon corresponds to the case where no external electromagnetic radiation is present in our system.

These three processes: (i) stimulated absorption, (ii) spontaneous emission, and (iii) stimulated emission are enough to explain the majority of light-matter interaction phenomena.

The black body revisited. We will show below that the Einstein coefficients are related by $B_{fi} = B_{if}$. If this is the case, one finds that in *thermal equilibrium*, where the transition rate $i \to j$ must be the same as that of $j \to i$ (else the system would not be in equilibrium), and ignoring the effects of spontaneous emission, one has that

$$W_{f\leftarrow i} = W_{i\leftarrow f} \quad \to \quad N_i = N_f \,, \tag{9.5}$$

that is, the population of the ground and the excited state is the same. This is in contradiction with thermodynamics, since according to the Boltzmann distribution the occupation number of a state with energy E at temperature T should be

$$N(E) \sim \exp\left(-\frac{E}{k_B T}\right) \,, \tag{9.6}$$

with k_B being Boltzmann constant, and then the ratio between the population of a state with energy E_f and another one with density E_i should be classically be given by

$$\frac{N_f}{N_i} = \exp\left(-\frac{(E_f - E_i)}{k_B T}\right) \,, \tag{9.7}$$

which is inconsistent with Eq. (9.5).

The incorrect assumption in the derivation above has been neglecting the effects of spontaneous emission, which are important in this context. The rate of this process is given by A, the *Einstein coefficient for spontaneous emission*, which is independent of ρ (since spontaneous emission is an stochastic atom-by-atom process). Therefore, the total rate for electromagnetic emission in our system per molecule will be given by the sum of the rates for spontaneous and stimulated emission,

$$w_{f\to i} = A + B_{if}\rho \; . \tag{9.8}$$

Therefore, in thermal equilibrium, where populations of the two levels, N_i and N_f, do not vary with time, we find that

$$N_i B_{fi}\rho(\nu) = N_f \left(A + B_{if}\rho\right)(\nu) \, , \tag{9.9}$$

and therefore we obtain the following condition on the density of states for photons of frequency ν,

$$\rho(\nu) = \frac{N_f A}{N_i B_{fi} - N_f B_{if}} = \frac{A/B_{fi}}{N_i/N_f - B_{if}/B_{fi}} = \frac{A/B_{fi}}{e^{h\nu/k_B T} - B_{if}/B_{fi}} \, , \tag{9.10}$$

where in the last step we have used the condition that N_f/N_i should be determined by the Boltzmann distribution Eq. (9.7).

For electromagnetic radiation off a black body, or in other words, for electromagnetic radiation in thermal equilibrium at temperature T, we know that the density of states for photons of energy ν is given by *Planck's distribution*,

$$\rho(\nu) = \frac{8\pi h\nu^3/c^3}{e^{h\nu/k_B T} - 1} \, , \tag{9.11}$$

and then the comparison with Eq. (9.10) allows us to identify $B_{fi} = B_{if} \equiv B$ as well as to determine the Einstein coefficient for spontaneous emission, which is given by

$$A = \left(\frac{8\pi h\nu^3}{c^3}\right) B \, . \tag{9.12}$$

There are a number of important implications of this derivation:

- The Einstein coefficients for stimulated absorption and emission are *identical.* In other ways, the stimulated absorption and emission are symmetric processes.
- The Einstein coefficient A for spontaneous emission is proportional to B, that of stimulated emission. In other words, spontaneous and stimulated photon emission take place at rates that are proportional to each other in a given system.
- Since $A \sim \nu^3$, we find that *spontaneous emission becomes important for high frequencies*, *i.e.* is important for light emission but much less for radio emission.

Selection rules for photon emission and absorption. In general, not all transitions involving the emission and absorption of photons that are consistent with energy conservation will be allowed in a given quantum system. There are two main conditions for photon emission and absorption to take place for a specific quantum system, and these are:

(a) *Resonance*: the energy of the photon must be the same as the difference in energy between ground level and the excited state, $h\nu = E_f - E_i$.

(b) *Selection rules*: conditions on the quantum numbers of the states f and i which restrict the allowed transitions, arising for instance from the conservation of angular momentum (since the photon is a spin-1 particle).

The origin of these selection rules can be traced back to the formalism of quantum theory. The interactions between matter and electromagnetic radiation can be described in quantum mechanics using the *time-dependent perturbation theory*, since the EM field oscillates in time. Classically, for a molecule to interact with an electromagnetic field and emit or absorb a photon with frequency ν, it must posses, even if for a short amount of time, a *dipole* oscillating at that frequency. In quantum theory, the corresponding Hamiltonian for this interaction is

$$\widehat{H} = \widehat{H}^{(0)} + \widehat{H}^{(1)}(t) \quad \widehat{H}^{(1)}(t) = -\mu_z \mathcal{E} \cos(\omega t)\,, \tag{9.13}$$

where $\widehat{H}^{(0)}$ is the time-independent (static) Hamiltonian of the system and $\widehat{H}^{(1)}(t)$ is the time-dependent interaction term between the oscillating electric field of frequency ω and amplitude $\mathcal{E}$ and the molecular dipole moment μ_z.

Assuming that the perturbation is switched on at $t = 0$, it is possible to show that the rate of change of the population of the quantum state Ψ_f due to transitions from Ψ_i induced by $\widehat{H}^{(1)}(t)$ (where Ψ_f and Ψ_i are eigenstates of the time-independent Hamiltonian $\widehat{H}^{(0)}$) is given by

$$w_{f\leftarrow i} \propto |H^{(1)}_{fi}|^2 \quad \text{where} \quad H_{fi} \equiv \int \Psi_f^* \widehat{H}^{(1)}(t) \Psi_i d\tau\,, \tag{9.14}$$

which in the case of electromagnetic radiation corresponds to

$$w_{f\leftarrow i} \propto |\mu_{fi}|^2 \mathcal{E}^2 \quad \mu_{fi} = \int \Psi_f^* \widehat{\mu}_z \Psi_i d\tau\,, \tag{9.15}$$

and therefore the rate of this transition is proportional to the square of the *transition dipole moment* μ_{fi}. For $\mu_{fi} = 0$, the rate vanishes and thus such specific transition is not allowed. This is the main condition that underlies the selection rules for electronic transitions in a molecule. Recall that for atoms we have that

$$\widehat{\vec{\mu}} = -e\vec{r}\,, \tag{9.16}$$

so that its components are $\mu_x = -ex$ and so on, and a similar expression holds for molecules but summing over all its individual components. In summary, the main consequence of the selection rules is that

> electronic transitions in molecules involving the emission or absorption of photons can only take place if the corresponding dipole matrix element is different from zero, $\mu_{fi} \neq 0$.

Qualitatively, $\mu_{fi} \neq 0$ corresponds of a modification in the *geometric distribution of the electric charge* within a given atomic state. Let us illustrate this concept with two examples:

- The transition $1s \to 2s$ does not have associated any charge shape redistribution, since both atomic orbitals are spherically symmetric, therefore in this case $\mu_{fi} = 0$ and this transition is not allowed.
- For the transition $1s \to 2p$ we move from a spherically symmetric orbital ($1s$) to another which is not ($2p$), implying that there is a charge redistribution, so $\mu_{fi} \neq 0$ and that this transition is allowed.

This behaviour is illustrated in Fig. 9.4, which represents the changes in the electric charge distribution in the $1s \to 2s$ transition (forbidden) and in the $1s \to 2p$ transition (allowed).

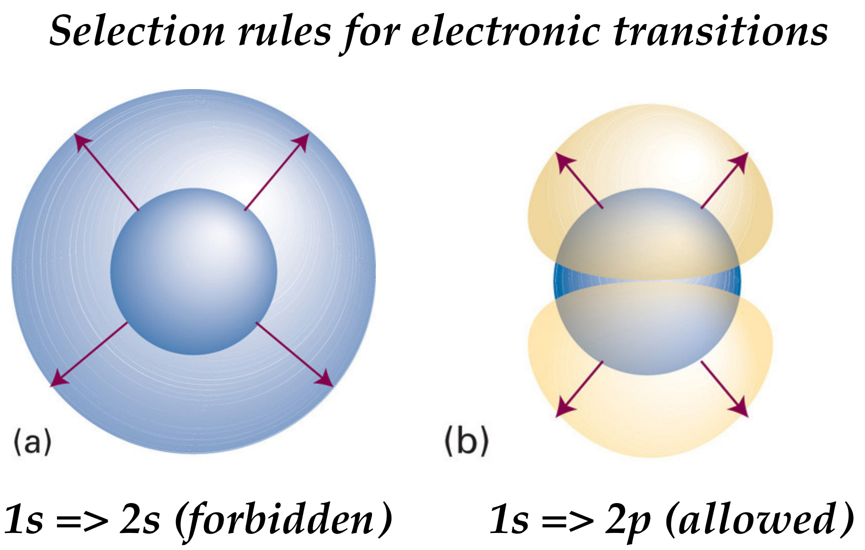


Figure 9.4: In electronic transitions in molecules involving the emission and absorption of photons, only those transitions which involve an electric charge redistribution, implying non-zero value of the electric dipole moment matrix element $\mu_{fi} \neq 0$, are allowed by the selection sum rules. For instance, the $1s \to 2s$ transition (left plot) does not involve a change of shape of the charge distribution, and thus is forbidden, unlike the $1s \to 2p$ transition (right plot) which is instead allowed since the shape of the electric charge distribution changes.

The Beer-Lambert law (*) From the *macroscopic point of view*, absorption of electromagnetic radiation can be described by the so-called *Beer-Lambert law*, which states that

$$I = I_0 \, 10^{-\epsilon [J] L} \,, \tag{9.17}$$

where I_0 and I are the intensities of the electromagnetic radiation beam before and after crossing a material, $[J]$ is the concentration of absorbing molecules, ϵ is the *molar absorption coefficient*, also now as *extinction coefficient*, and L is the *optical path length* (which differs from the geometrical path length in mediums other than vacuum). In other words, the Beer-Lambert law states that absorption increases exponentially with the optical patch length L, the molar absorption coefficient ϵ (which is specific of the material) and the concentration of absorbing molecules. It then follows that the *absorbance* of the sample is given by

$$A \equiv \log \frac{I_0}{I} = \epsilon [J] L \,. \tag{9.18}$$

Similarly, a transmission coefficient $T = I/I_0$ can be defined, which related to the absorbance as $A = -\log T$.

This law can be derived as follows: consider the change of intensity of the incident electromagnetic radiation over an infinitesimal length,

$$dI = -\kappa [J] I dx \,, \tag{9.19}$$

where κ is an undetermined proportionality constant. Integrating over the total length L, and redefining $\epsilon = \kappa/10$, the sought-for result Eq. (9.17) is recovered. The molar absorption coefficient ϵ depends on the wavelength of the incident light. Thus for some applications it is advantageous to define an *integrated absorption coefficient* $\mathcal{A}$ as follows

$$\mathcal{A} = \int \epsilon(\nu) d\nu \,, \tag{9.20}$$

which takes into account the contribution to the total absorption from all relevant frequencies.

Summary In this part of the lecture we have seen that:

- The Einstein coefficients A, B_{fi} and B_{if} describe the (spontaneous and stimulated) absorption and emission of electromagnetic radiation (photons) in a two-level quantum system.
- The Einstein coefficients for stimulated absorption and emission are equal $B_{fi} = B_{if} = B$.
- The Einstein coefficient for spontaneous emission A is proportional to both B and to ν^3, and thus spontaneous emission can be neglected for long wavelengths.
- The selection rules determine which transitions between atomic levels are allowed. Only those transitions for which the dipole matrix element $\mu_{fi} \neq 0$ are possible. The condition $\mu_{fi} \neq 0$ can be understood as a redistribution of the geometrical charge configuration.
- The Beer-Lambert law can be used to describe macroscopically the phenomenon of the absorption of electromagnetic radiation

9.2 Vibrational modes and light interaction in diatomic molecules

We now turn to discuss the important topic of *vibrational modes* and their interaction with electromagnetic radiation in molecules, a topic known as *vibrational spectroscopy.* To begin with, we will restrict ourselves to *diatomic molecules*, and then move to more complex poly-atomic molecules.

As a representative diatomic molecule, we can consider ionised molecular hydrogen, H_2^+. The molecular orbitals of this molecule have been studied in some detail in Sect. 6.4. There we worked in the *Born-Oppenheimer* approximation, assuming that the position of the two protons was fixed. Now we consider the degrees of freedom associated to the vibrational models of these two protons in the H_2^+ molecule (assumed to be decoupled from the electron motion).

The potential energy of ionised molecular hydrogen can be represented by a potential $V(r)$ of the form of Fig. 9.5, where r denotes the distance between the two protons that compose the molecule. This potential is characterized by an equilibrium distance r_e where it has an absolute minimum (thus $V'(r = r_e) = 0$), so that for smaller or larger values of the inter-atomic separation r the potential energy increases. This potential is known as the *Morse* potential, and can be described by the equation

$$V(r) = D_e \left(1 - e^{-a(r-r_e)}\right)^2 , \tag{9.21}$$

where r is the inter-atomic distance, r_e the equilibrium distance, a is a parameter that determines the width of the potential well, and D_e is the *dissociation energy*, the energy above which the molecule separates into its two constituents atoms separately (in other words, $r \to \infty$). When $r \to 0$ the potential becomes highly repulsive, due to the electric repulsion between the two positive charges of the protons.

For small deviations of r with respect the equilibrium position, the Morse potential reduces to the *harmonic oscillator potential.* Indeed, starting from Eq. (9.21) we find that that

$$\begin{aligned} V(r = r_e) &= 0 & (9.22)\\ \frac{dV(r)}{dr} &= 2D_e \left(1 - e^{-a(r-r_e)}\right)(-ae^{-a(r-r_e)}) = 0 \to r = r_e & (9.23)\\ \frac{d^2V(r)}{dr^2}(r = r_e) &= 2a^2 D_e \,, & (9.24) \end{aligned}$$

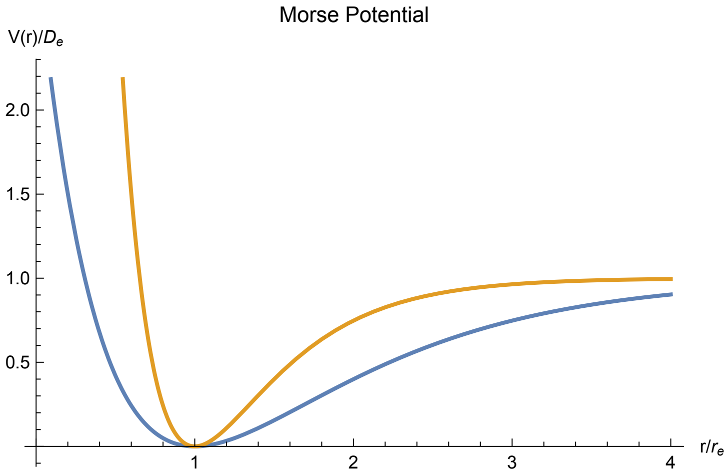


Figure 9.5: The Morse potential $V(r)/D_e$ Eq. (9.21) as a function of r/r_e, for two values of the parameter a. We note that when the separation is large, $r \gg r_e$, the Morse potential $V(r)$ tends to the dissociation energy D_e: even small energy fluctuations can break the molecule in this limit. For small values of r the Morse potential becomes highly repulsive because of the two positive electric charges of the proton.

and therefore the Taylor expansion of the Morse potential up to second order reads

$$V(r) \simeq a^2 D_e \left(r - r_e\right)^2 + \mathcal{O}\left(\left(r - r_e\right)^3\right) , \tag{9.25}$$

which can be identified with an harmonic oscillator potential, such as the one we studied in some detail in Sect. 2.4, with *spring constant* now set to $k_f = 2a^2 D_e$. In other words, the larger the value of a, the *stiffer* the potential between the two atoms that form the diatomic molecule.

Therefore, the vibrational modes of a diatomic molecule can be described to good approximation by the states of a quantum harmonic oscillator with spring constant $k_f = 2a^2 D_e$.

In Sect. 2.4, we showed that the Schroedinger equation for a quantum harmonic oscillator was given by

$$-\frac{\hbar^2}{2m_{\rm eff}}\frac{d^2\Psi}{dx^2} + \frac{1}{2}k_f x^2 \Psi = E\Psi , \tag{9.26}$$

where since vibrations are with respect the center of mass of the diatomic system, the effective mass is given by the reduced mass of the system, defined as

$$m_{\rm eff} = \frac{m_1 m_2}{m_1 + m_2} . \tag{9.27}$$

The energies of the quantum harmonic oscillator are quantized and given by

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega , \quad \omega = \sqrt{\frac{k_f}{m_{\rm eff}}} , \tag{9.28}$$

and thus the quantum states are equally spaced in energy among them. Eq. (9.28) can be rewritten as

$$E_n = hc\widetilde{G}_n , \quad \widetilde{G}_n \equiv \left(n + \frac{1}{2}\right)\widetilde{\nu} , \quad \widetilde{\nu} = \frac{1}{2\pi c}\left(\frac{k_f}{m_{\rm eff}}\right)^{1/2} , \tag{9.29}$$

which is sometimes useful since now G_n has units of *wave numbers*, that is, units of inverse length. The angular frequency of the oscillations is given by ω.

Infrared spectroscopy. What is the order of magnitude for frequencies associated to these transitions between different vibrational states in a diatomic molecule? In the case of the HCl molecule for instance, we have that the effective spring constant is found to be $k_f = 520$ N/m, and then it is possible to show that the typical wavelength associated to vibrational transitions in HCl is $\lambda \simeq 3.4\ \mu$m, which corresponds to the *infrared region* of the electromagnetic spectrum, see Fig. 9.1. Therefore, molecular vibrational modes can be explored by means of *infrared (IR) spectroscopy.*

Selection rules. As in the case of transitions between molecular electronic states, also for the transitions between *molecular vibrational states* that involve the emission or absorption of a photon there exist a number of *selection rules* that restrict the allowed transitions. As in the case of individual atoms, the allowed transitions will be determined by the dipole moment matrix element

$$\vec{\mu}_{fi} = \int \Psi_i^* \,\widehat{\vec{\mu}}\, \Psi_i \, d\tau \tag{9.30}$$

where in the case of molecules the dipole moment operator reads

$$\widehat{\vec{\mu}} = \sum_{i=1}^{N} q_i \vec{r}_i \,, \tag{9.31}$$

and the sum runs over the components of the system.

For individual atoms, only transitions for which $\vec{\mu}_{fi} \neq 0$, and thus involving a charge redistribution, were allowed. For infrared transitions between vibrational states of a molecule, the condition for allowed transitions is that the electric dipole moment of the molecule changes when the atoms are displaced relative to each other. From the wave functions harmonic oscillator, it can be demonstrated that transitions between molecular vibrational states involving the emission or absorption of photons must satisfy:

$$\Delta n = \pm 1 \,, \tag{9.32}$$

with n the principal quantum number, which involves a change in wave number of

$$\Delta \widetilde{G}_n = \widetilde{G}_{n+1} - \widetilde{G}_n = \widetilde{\nu} \,, \tag{9.33}$$

which is independent of n.

At room temperatures we have that $KT/hc \simeq 200$ cm^{-1}, so this will be the typical wave number of infrared absorptions by vibrational modes. Since from thermodynamics we know that the ground state $n = 0$ will be the one with higher occupation, this implies that infrared absorption by molecules will be dominated by the fundamental transition $1 \leftarrow 0$.

Anharmonicity. While for small values of n the harmonic approximation Eq. (9.25) to the Morse potential Eq. (9.21) is reasonable, for excited states with high values of the quantum number n eventually the quadratic approximation will break down. To see this, let us first of all rewrite the Morse potential as

$$V(r) = hc\widetilde{D}_e \left(1 - e^{-a(r-r_e)}\right)^2 \,, \quad a = \left(\frac{m_{\rm eff}\omega^2}{2hc\widetilde{D}_e}\right)^2 \,. \tag{9.34}$$

Then, as n is increased, the energy levels become more densely packed, as opposed to the result with the quadratic potential where $\Delta E_n = E_{n+1} - E_n = \hbar\omega$ was constant and independent of n. It can be shown that the energy of the harmonic oscillator for larger values of n is better approximated by

$$G_n = \left(n + \frac{1}{2}\right)\widetilde{\nu} - \left(n + \frac{1}{2}\right)^2 x_e\widetilde{\nu}\,, \quad x_e \equiv \frac{a^2\hbar}{2m_{\rm eff}\omega} = \frac{\widetilde{\nu}}{4\widetilde{D}_e}\,, \tag{9.35}$$

and x_e is defined as the *anharmonicity constant*, which quantifies the deviations of the energy of the vibrational modes of a diatomic molecule with respect to the quadratic approximation. What this means is that now the differences in *wavenumber* between two adjacent quantum states will be given by

$$\Delta\widetilde{G}_n \equiv \widetilde{G}_{n+1} - \widetilde{G}_n = \widetilde{\nu} - 2(n+1)x_e\widetilde{\nu} + \mathcal{O}(n^2)\ , \tag{9.36}$$

so we see that for high values of n the separation between adjacent energy levels will be reduced due to these anharmonicity effects. We also note that anharmonicity also leads to the *partial lifting of the* $\Delta n = \pm 1$ *selection rule* for vibrational modes, since that rule was derived assuming an harmonic potential.

9.3 Vibrational modes and light interaction in poly-atomic molecules

Up to here, the discussions concerning the vibration modes have been restricted to the simplest possible molecules, namely diatomic molecules. Next we turn to discuss the effects of vibrational models in more complex molecules composed by more than two atoms. The main differences as compared to the diatomic case are that:

- In molecules composed by two atoms, there can only be *one possible vibration mode*, corresponding to vibrations with respect the equilibrium position of the inter-atomic potential.
- For a molecule composed by N atoms, in general we can have up to $3N - 5$ *vibrational modes.*

Of particular interest are the so called *normal modes* of vibration, which is another name for the *independent vibration modes.* The excitation of a normal vibration mode, which affects the motion of a number of atoms, does not lead to the excitation of any other normal mode or group of atomic motions.

An suitable example of a poly-atomic molecule is carbon dioxide, CO_2. Since this molecule has three atoms, $N = 3$, it will be characterized by $3N - 5 = 4$ normal modes. An interesting question here is: can the first and the third atom in the molecule vibrate in a fully independent way? As illustrated in Fig. 9.6, there are various possible vibration models of the CO_2 molecule. The first two modes are known as *symmetric stretch* and *antisymmetric stretch* respectively, while the two bottom ones are the *bending modes.* We also indicate the indicate the associated value of the wave number associated to each mode. Every normal mode behaves as a *fully independent harmonic oscillator*, characterized by their own set of vibrational energy levels with wave numbers

$$\widetilde{G}_q(n) = \left(n + \frac{1}{2}\right)\widetilde{\nu}_q \qquad \widetilde{\nu}_q = \frac{1}{2\pi c}\left(\frac{k_q}{m_q}\right)^{1/2}\ . \tag{9.37}$$

Given that the selection rules require a *change in the electric dipole moment* to activate a given vibration mode, we see that:

- The *symmetric stretch* vibrational mode is infrared inactive, since the electric dipole moment (the electric charge distribution of the molecule) is unchanged.
- The other three normal modes of the CO_2 molecule are IR active, since the electric dipole moment (the electric charge distribution) varies during the vibrational motion.

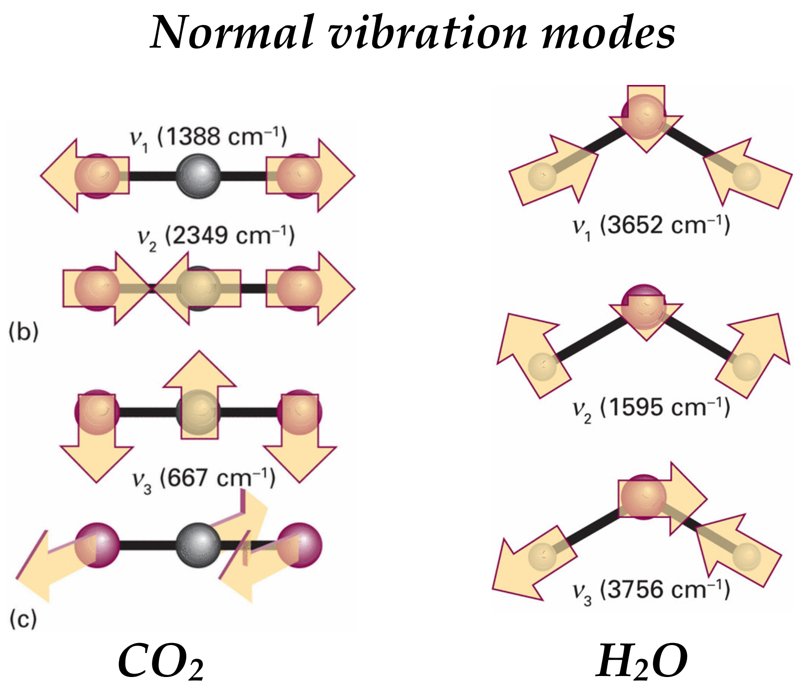


Figure 9.6: Some of the normal vibration modes that appear in CO_2 (left plot) and in H_2O (right plot) molecules, where we also indicate the associated value of the wave number associated to each mode. In the case of CO_2, the first two modes are known as *symmetric stretch* and *antisymmetric stretch* respectively, while the two bottom ones are the *bending modes*.

Therefore, not all possible vibrational modes of a molecule can be probed with infrared spectroscopy.

Raman spectroscopy (*). Raman spectroscopy is an spectroscopic method based on the scattering via molecules of incident monochromatic light. With this technique the vibrational and rotational modes of a molecule can be observed. The inelastic scattering of the incoming Raman photons with these vibrational states changes the energy of the scattered photon, which then can be detected providing information on the vibrational spectrum of the molecule (for example allowing to identify which kind of molecules we have in our system). The vibrational transitions underlying Raman scattering are represented in Fig. 9.7. If the outgoing photon has a higher energy than the incoming radiation, the associated transition is called a *anti-Stokes* transition, and in the opposite case the have a Stokes transition.

Also in Raman scattering only transitions that comply with a set of *selection rules* will be physically allowed. For instance, in general if the molecular vibrational state changes due to the Raman scattering the polarizability of the molecule will be also modified. As a rule of thumb, weak infrared transitions will lead to strong Raman transitions, and conversely. We know that in the quantum harmonic oscillator the selection rule $\Delta n = \pm 1$ should hold. Transitions with $\Delta n = +1$ are the Stokes transitions and those with $\Delta n = -1$ are the anti-Stokes one, that in general will have associated low intensities since higher vibrational states are scarcely populated at room temperatures.

What is the main difference then between Raman spectroscopy and IR spectroscopy? That in the latter case the motion corresponding to a normal mode should be accompanied by a change of the electric dipole moment, while in the former case we require instead a change in the molecule's *polarizability* as it changes its vibrational state. Therefore we can probe complementary vibration modes in Raman as compared to IR spectroscopy. And here a useful *exclusion rule* follows: if a molecule has a centre of symmetry, then no

Raman spectroscopy

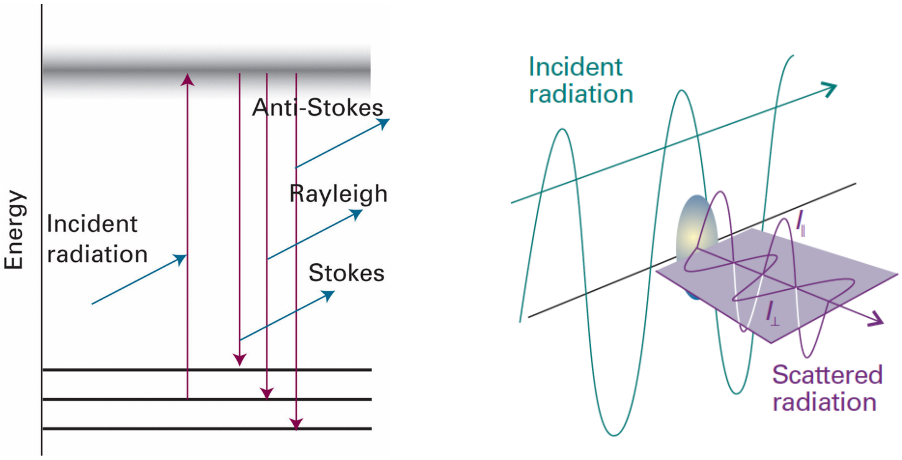


Figure 9.7: Schematic representation of Raman spectroscopy: incident monochromatic radiation undergoes inelastic scattering with the vibrational energy levels of a molecule, and thus in general the outgoing photon will have both different energy and polarization as the incoming photon (as shown in the right plot).

normal modes can be at the same time infrared and Raman active.

A specific example of Raman spectroscopy is known as *resonant Raman spectroscopy*, where the incident radiation has a frequency corresponding to almost the actual electronic excitation of the molecule. In these conditions, a photon is emitted when the excited state returns to a state close to the ground state.

Summary

To summarize, some of the important concepts that we have learned in this lecture are the following:

- The interactions between molecules and light can be described by three different processes: spontaneous emission, stimulated emission, and stimulated absorption.
- The quantum vibrational states of molecules can be modeled using the harmonic oscillator with a given effective mass and effective spring constant k_f upon approximating the inter-atomic potential by a quadratic expansion.
- There are *selection rules* which restrict the vibrational transitions, and only those that change the electric dipole moment of the molecule are allowed.
- For excited vibrational states, deviations of the quadratic potential become important, and are quantified by x_e, the anharmonicity constant.
- An a poly-atomic molecule, the normal modes represent the independent motion of groups of atoms, and each normal mode can be treated as fully independent harmonic oscillator.
- Raman spectroscopy can be used to probe some of the infrared transitions between vibrational modes of a molecule that cannot be accessed by IR spectroscopy.

Further reading

Further reading material about the topics covered in this lecture can be found in the following sections of the course textbook (*Physical Chemistry*, Atkins and De Paula, 10th edition):

- **12A.1**: The absorption and emission of radiation.
- **12B.2**: The rotational energy levels.
- **12C.1**: Microwave spectroscopy.
- **12C.2**: Rotational Raman spectroscopy.
- **12D.1**: Vibrational motion.
- **12D.2**: Infrared spectroscopy.
- **12D.3**: Anharmonicity.

10 HC11: Molecular spectroscopy and electronic transitions

van Quantum tot Molecuul

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

In this lecture, we continue our study of molecular spectroscopy and electronic transitions in molecules focusing on *p-electron conjugate systems*, the Franck-Condon principle, the physics underlying the phenomena of *fluorescence* and *phosphorescence*, as well as the basic principles of operation of a *laser*. We will also present an explicit application of these ideas in the medical context, namely the *photo-dynamic therapy* (PDT).

The **learning goals** of this lecture are:

(a) Understanding and applying π-π^* electronic transitions in molecules.

(b) Understanding which electronic transitions can take place for either absorption or emission of a photon while the atomic nuclei are at rest (the Franck-Condon principle), including the mathematical description and implications of absorption and emission spectra.

(c) Applying the principles of fluorescence and phosphorescence.

(d) Demonstrating basic principles that underlie the operation of a laser, such as the concept of population inversion.

10.1 Electronic transitions in molecules

In Sect. 9.2 we studied the transitions between different vibrational states of a molecule, mediated by the emission of absorption of photons. We now consider instead the *electronic transitions* that take place between the different electronic states of a molecule, such as those described by Molecular Orbital theory. As opposed to vibrational transitions, which were associated to *infrared frequencies*, electronic molecular transitions take place mainly in the *visible and ultraviolet* regions of the electromagnetic spectrum, and therefore, among the many important properties of these transitions, they are also responsible of the *color* of materials and substances.

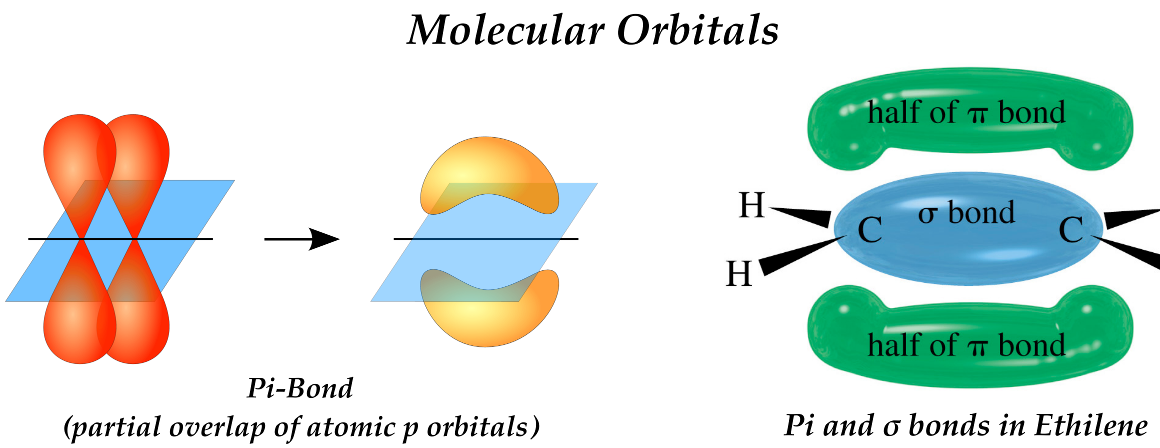


Figure 10.1: Left plot: a π-bond in a molecule is formed by the partial overlap between atomic p-type orbitals. Right plot: the bonding structure of ethylene, C_2H_4, showing a σ bond between the two C atoms (rotationally symmetric with respect to the bond axis) and the two halves of the π bond arising from the p atomic orbitals.

Molecular orbitals. It is useful at this point to recall some important properties of *molecular orbitals*, first presented in Sect. 6. We know that that electronic orbitals in atoms can have different orbital angular momentum quantum numbers: s $(l = 0)$, p $(l = 1)$, d $(l = 2)$ and so on, each one with different geometric properties. In molecules, under the right circumstances, the electrons might become *de-localized* by means of the overlap between individual atomic orbitals, that is, they cannot be assigned to any of the specific components of the molecule. This is the basic principle of Molecular Orbital theory.

The strongest type of *chemical covalent bonds* in molecules are known as σ-bonds, formed by a head-on overlapping between atomic orbitals leading to a molecular orbital. The key property of this bond is that the resulting orbitals are *rotationally symmetric with respect to the bond axis.* Another important type of molecular bonds are the so-called π-bonds, covalent chemical bonds where two lobes of one atomic orbital overlap with the two lobes of another atomic orbital, resulting in the bonding between the two atoms. In Fig. 10.1 we show how a π-bond in a di-atomic molecule is formed by the partial overlap between atomic p-type orbitals. In the same figure we also show the bonding structure of the ethylene molecule, C_2H_4, showing a σ bond between the two C atoms (rotationally symmetric with respect to the bond axis) and the two halves of the π bond arising from the p atomic orbitals.

In this lecture we will focus on *conjugate p-electron systems*, which denote molecules where atomic p-type orbitals are connected by means of de-localized electrons, thus giving place to π-type covalent bonds. These molecules are very important in organic chemistry and biology, and the wide majority of light-absorbing molecules, known as *chromophores*, that appear biology are indeed p-electron conjugate systems.

The electronic transitions within a conjugate p-electron molecular system are known as the $\pi \to \pi^*$ transitions. This is a specific type of molecular electronic transitions, others include the $\sigma \to \sigma^*$ and the $n \to \sigma^*$ transitions. Here by π^* we indicate an *anti-bonding molecular orbital* of the π-type, that is, a molecular orbital that when occupied by electrons *weakens* the bond between the constituents atoms of the molecule and thus increases (rather than decreases) the energy of the molecule relative to the separated individual atoms. These $\pi \to \pi^*$ transitions, involving de-localized p-orbitals, can be described by *Huckel theory* (Sect. 8.1) and are characterized by *strong light absorption*. Recall that *Huckel theory* can be used to construct linear combinations of atomic orbitals (LCAO) into molecular orbitals to determine the energies of π electrons in general conjugated molecules.

For molecular orbitals, we often make the distinction between *HOMO*, the *highest occupied molecular*

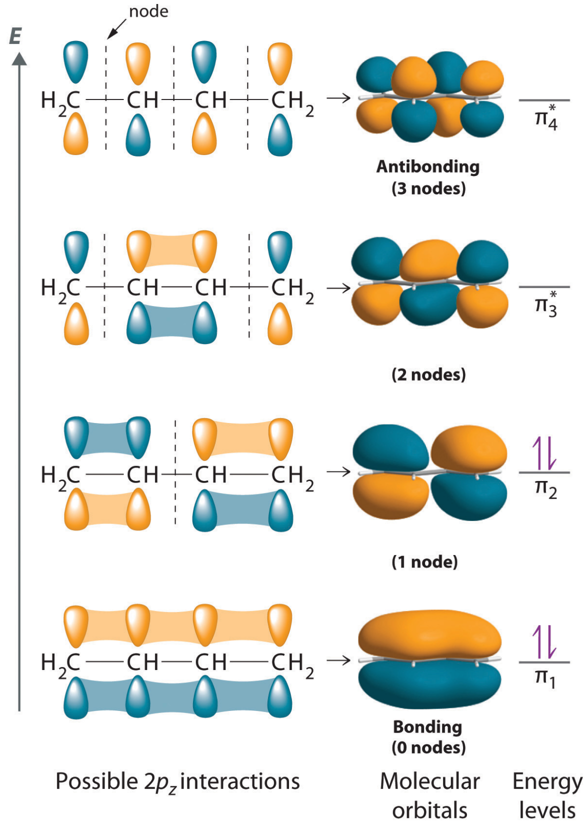


Figure 10.2: Schematic representation of the molecular orbitals of the butadiene molecule, C_4H_8. The two bottom orbitals are *bonding* π molecular orbitals, while the top two ones are instead *anti-bonding* π^* molecular orbitals. The highest occupied molecular orbital (HOMO) is here π_2, while the lowest unoccupied molecular orbital (LUMO) is instead π_3^*. Note that in the HOMO the two electrons are paired (opposite spins).

orbital and *LUMO*, the *lowest unoccupied molecular orbital*. Let us consider the representative case of the butadiene molecule C_4H_6. In Fig. 10.2 we show a schematic representation of the molecular orbitals of butadiene. The two bottom orbitals are *bonding* π molecular orbitals, while the top two are instead *anti-bonding* π^* molecular orbitals. The highest occupied molecular orbital (HOMO) is here π_2, while the lowest unoccupied molecular orbital (LUMO) is instead π_3^*. Note that in the HOMO the two electrons are paired, that is, they have opposite spins.

Allowed molecular transitions. In order to determine which electronic molecular transitions are possible, we need to take into account the corresponding *selection rules*. In the case of individual atoms and of vibrational transitions in molecules, these selection rules have been reviewed in HC10. Similarly as in the case of the selection rules for electronic transitions in individual atoms, also for electronic transitions in molecules, only if the dipole moment matrix element satisfies $\mu_{fi} \neq 0$ a given transition will be allowed.

Therefore, in the case of electronic transitions between the molecular orbitals of butadiene represented in Fig. 10.2, only those that are characterized by a non-zero electric dipole matrix element, $\mu_{fi} \neq 0$, will be physically allowed. Using the wave function for the molecular orbitals of butadiene, it it can be shown that

$$\int \Psi^*_{\pi_3^*}\, \hat{\mu}\, \Psi_{\pi_2} d\tau \neq 0 \,, \tag{10.1}$$

and therefore a transition between the HOMO and LUMO levels, that is, a $\pi \to \pi^*$ transition, is allowed. This transition will have associated the absorption of visible light by the butadiene molecule. It can also be shown that is x is the molecular bonding axis, only $\mu_{x,fi} \neq 0$, while instead $\mu_{y,fi} = \mu_{z,fi} = 0$. This has the consequence that the optical transition $\pi_2 \to \pi_3^*$ will only be possible for *light polarized along the x axis.* On the other hand, the absorption of a photon by the HOMO-1 level, that is, the electronic state immediately below HOMO, labeled π_1 in Fig. 10.2, resulting into a transition to the LUMO level π_3^* turns out not to be allowed by the selection rules, since the corresponding dipole moment matrix element vanish,

$$\int \Psi^*_{\pi_3^*} \, \hat{\mu} \, \Psi_{\pi_1} d\tau = 0 \, . \tag{10.2}$$

By computing the energy of allowed electronic molecular transitions, it can be shown that often organic molecules actually absorb light in the *ultraviolet range*, for instance ethylene absorbs at $\lambda = 163$ nm while butadiene at $\lambda = 220$ nm.

More about electronic molecular transitions. We can now present some additional facts that are important for the understanding of the physics underlying electronic transitions between molecular orbitals:

- In molecular orbitals, the *conjugation length* is defined as the length of the orbital resulting from the partial overlap of atomic p-orbitals. The more p-orbitals that contribute to a molecular orbital, the larger its conjugation length will be. This means that an electron that belongs to this orbital will be *more de-localized* the larger the conjugating length of that orbital, since it will be able to *hop* between more atoms that compose the molecule. As a general rule, the greater the conjugation length of a molecular orbital, and this the greater the amount of de-localization of the electrons that belong to this orbital, the smaller its energy, or in other words, the *stronger its contribution to the molecular bonding.*

  Recall that a similar behaviour was observed when studying the particle in a box system using the Schroedinger equation in Sect. 2.1. There we saw that the energy of the allowed quantum states of the particle scaled with the length of the box L as $E \sim L^{-2}$, in other words, the greater the box (and thus the de-localization) the smaller the energy of the particle, qualitatively the same behaviour as for molecular orbitals.

- Another important type of electronic molecular transitions are the $n \to \pi^*$ optical transitions. Recall that by n we denote the molecular orbital of an *auxochrome*, a functional group of atoms with free (unbounded) electron pairs that when attached to a *chromophore* (that is, a light-absorbing molecule) modifies both the wavelength and intensity of absorption. In some circumstances, an electron might transition between a n orbital of an *auxochrome* to the π^* orbital of a *chromophore.* In this transitions, the larger the conjugation length, the larger the *extinction coefficient* that determines how strongly this specific molecule will absorb light of a given wavelength.

- It can be shown that there is a direct connection between the value of the dipole moment transition matrix element μ_{fi} and the extinction coefficient $\epsilon(\nu)$ for light of a given frequency, that determines how strongly this molecule does absorb light of this specific frequency. Intuitively, one could expect that the greater the value of the electric dipole moment transition matrix element μ_{fi}, the stronger the absorption of light with wavelengths corresponding to this specific electronic transition. And indeed, the integral over all frequencies of $\epsilon(\nu)$ is proportional to the square of μ_{fi}, in other words

$$\int \epsilon(\nu) d\nu \propto |\mu_{fi}|^2 \, . \tag{10.3}$$

Therefore, for very *narrow* transitions, the extinction coefficient ϵ is simply proportional to $|\mu_{fi}|^2$.

10.2 The Franck-Condon principle

We now turn to discuss the *Franck-Condon principle*, which describes the transitions in which a *simultaneous* change in the vibrational and electronic states of a molecule takes place, due to the emission or absorption of a photon of the appropriate energy. This principle is important to explain the *vibrational fine structure* in the optical absorption spectrum. Recall that vibrational transitions (in the *infrared* range of the EM spectrum) have associated much smaller energies than optical transitions (in the *visible* range), and thus appear in the optical absorption spectrum as small splittings of the main absorption lines. In general, we denote as a *vibronic transition* a transition that involves simultaneous modifications in the vibrational and electronic energy states of a given molecule.

The Franck-Condon (FC) principle states that since nuclei are so much more massive than electrons, $m_e \ll m_N$, *electronic transitions* between different quantum states take place at much shorter time-scales that those for which the nuclei can respond, therefore the *vibrational transitions* take place at slower rates. In other words, the probability of an *electronic molecular transition* is the highest where the *relative separation between two atomic nuclei* is unchanged. This phenomenon is thus also known as a *vertical transition*, since in a (r, E) plane the transition takes place vertically along the y axis, as shown in Fig. 10.3. As in general in quantum mechanics, the transition probability will be proportional to the overlap between vibrational wave functions in the ground electronic and in an excited electronic states, and it can be shown that this overlap is maximal when the relative separation between nuclei r is unchanged.

The physical interpretation of the Franck-Condon principle is summarized in Fig. 10.3, where we show a schematic representation of the molecular potential energies for the ground and the excited electronic states in a molecule. This molecular potential energy is nothing by the Morse potential that we discussed in HC10, Eq. (9.21). According to the Franck-Condon principle, the most intense vibronic transitions take place between the ground vibrational state in the electronic ground state to to the vibrational state *lying immediately above it* in the electronic excited state. Transitions to other vibrational states also occur, but with lower intensities.

In the quantum mechanical description of the FC principle, the molecule undergoes a transition to the upper vibrational state that most closely resembles the vibrational wave-function of the vibrational ground state of the lower electronic state. The two wave functions shown here exhibit the greatest overlap and hence the matrix element for the transition probability among them is the highest.

Dipole matrix element for vibronic transitions. In order to compute the matrix element for the electric dipole moment of such a vibronic transition, we need to take into account the sum of electronic and of nuclear contributions to the total molecular dipole moment operator, that is

$$\hat{\vec{\mu}} = -e \sum_j \vec{r}_j + e \sum_I Z_I \vec{R}_I \, , \tag{10.4}$$

where j runs over the electrons and I over the nuclei in our molecule, and Z_I is the total positive charge of the nucleus I. In order to compute the dipole transition matrix element, for both the initial and final states

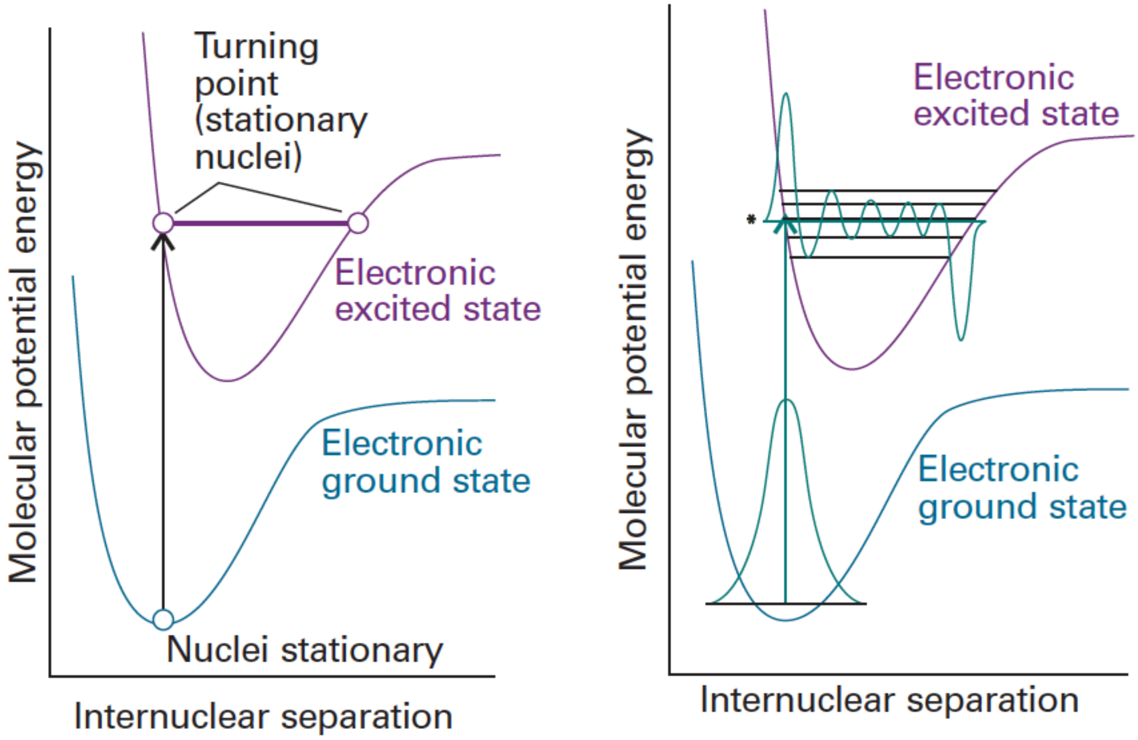


Figure 10.3: Schematic representation of the molecular potential energy (Morse potential Eq. (9.21)) for the ground and the excited electronic states in a molecule. Left plot: according to the Franck-Condon principle, the most intense vibronic transitions takes place between the ground vibrational state in the electronic ground state to to the vibrational state lying immediately above it in the electronic excited state, in a way that the *inter-nuclei separation* is kept constant. Transitions to other vibrational states also occur, but this lower intensities. Right plot: in the quantum mechanical description of the FC principle, the molecule undergoes a transition to the upper vibrational state that most closely resembles the vibrational wave-function of the vibrational ground state of the lower electronic state. The two wave functions shown here exhibit the greatest overlap, and hence the matrix element for the transition probability among them is the highest, corresponding to a transition with fixed inter-nuclear separation.

we separate the wave function into its electronic Ψ_ϵ and vibrational Ψ_ν components. This way the dipole matrix element for a given vibronic transition between an initial $\Psi_i = \Psi_{\epsilon_i}\Psi_{\nu_i}$ and final $\Psi_f = \Psi_{\epsilon_f}\Psi_{\nu_f}$ final state, will be given by

$$\vec{\mu}_{fi} = \int \Psi^*_{\epsilon_f}\Psi^*_{\nu_f}\left[-e\sum_j \vec{r}_j + e\sum_I Z_I\vec{R}_I\right]\Psi_{\epsilon_i}\Psi_{\nu_i}\,d\tau\,, \tag{10.5}$$

which can be rearranged as follows:

$$\vec{\mu}_{fi} = -e\sum_j \int \Psi^*_{\epsilon_f}\vec{r}_j\Psi_{\epsilon_i}d\tau_\epsilon \int \Psi^*_{\nu_f}\Psi_{\nu_i}d\tau_\nu + e\sum_I Z_I \int \Psi^*_{\epsilon_f}\Psi_{\epsilon_i}d\tau_\epsilon \int \Psi^*_{\nu_f}\vec{R}_I\Psi_{\nu_i}d\tau_\nu\,, \tag{10.6}$$

where we have used that the vibrational wave functions Ψ_ν depend only on the inter-nuclei separation $\vec{R}_I$ while the electronic wave functions Ψ_{ϵ_i} depend only on the position of the electrons $\vec{r}_i$.

The second term in Eq. (10.6) vanishes since the electronic wave functions are *orthogonal among them*,

$$\int \Psi^*_{\epsilon_f}\Psi_{\epsilon_i}d\tau_\epsilon = 0 \quad \text{for} \quad i \neq i\,, \tag{10.7}$$

but note that this does not apply to the *vibrational wave functions* necessarily, they correspond to different electronic states so they do not need to be orthogonal among them. Taking this property into account, we find that the dipole matrix element can be written as

$$\vec{\mu}_{fi} = -e \sum_j \int \Psi^*_{\epsilon_f} \vec{r}_j \Psi_{\epsilon_i} d\tau_\epsilon \int \Psi^*_{\nu_f} \Psi_{\nu_i} d\tau_\nu \equiv \mu_{\epsilon_f \epsilon_i} S(\nu_f, \nu_i) \, , \tag{10.8}$$

where we have introduced the following definitions:

$$\mu_{\epsilon_f \epsilon_i} \equiv -e \sum_j \int \Psi^*_{\epsilon_f} \vec{r}_j \Psi_{\epsilon_i} d\tau_\epsilon \, . \tag{10.9}$$

$$S(\nu_f, \nu_i) \equiv \int \Psi^*_{\nu_f} \Psi_{\nu_i} d\tau_\nu \, . \tag{10.10}$$

The first of these expressions, $\mu_{\epsilon_f \epsilon_i}$, corresponds to the *electronic dipole transition matrix element* between two molecular electronic states. The second of these expressions, $S(\nu_f, \nu_i)$, corresponds instead to the *overlap between the vibrational wave functions in the ground and in the excited states.* Therefore, the rate for a vibronic transition that involves simultaneously the change in the electronic and vibrational states of a molecule will be proportional to the *Franck-Condon factor*,

$$|S(\nu_f, \nu_i)|^2 \, , \tag{10.11}$$

which states quantitatively what was qualitatively illustrated in Fig. 10.3, namely that the rate for vibronic transitions is maximal when the vibrational wave functions in the ground and excited electronic states resemble the most.

10.3 Fluorescence and phosphorescence

We now describe to important phenomena related to light emission by molecules:

- *Phosphorescence* is the *slow emission of light* in molecules, taking place in an interval of *microseconds or more.*
- *Fluorescence* is the *fast emission of light* in molecules, taking place within a time interval of *nanoseconds* or less.

Therefore, these two phenomena related to light emission in molecules have time scales that differ by several orders of magnitude, and now we explain the reason for this behaviour.

Fluorescence. Let us start by discussing the phenomenon of *fluorescence.* From our discussion of HC10, we can recall a number of important properties of the electronic transitions within the two-level system, represented in Figs. 9.2 and 9.3, in particular we found that:

- The Einstein coefficient for stimulated absorption B_{fi} was proportional to the integral of the extinction coefficient over the relevant absorption band. It is therefore an intrinsic property of a given molecule.
- The Einstein coefficients for stimulated absorption and emission are identical, $B_{fi} = B_{if} \equiv B$.

- The Einstein coefficient for spontaneous emission A was proportional to both B and ν^3, with ν being the frequency of the emitted light:
$$A = \left(\frac{8\pi h\nu^3}{c^3}\right) B \,, \tag{10.12}$$
implying that spontaneous emission become relevant only at relatively high frequencies.

With these properties, we can compute rate of variation in time of the population of molecules in in the excited electronic state f due to emission an absorption as follows

$$\frac{dN_f}{dt} = -N_f A - N_f B_{if}\rho(\nu) + N_i B_{fi}\rho(\nu) \,. \tag{10.13}$$

In normal circumstances $\rho(\nu)$, the density of states of the EM field is small (unless a very powerful illumination is used) and can be neglected. This leaves a simple differential equation which depends only on the probability of spontaneous emission A,

$$\frac{dN_f}{dt} = -N_f A \quad \to \quad N_f(t) = N_0 e^{-At} = N_0 e^{-t/\tau_R} \,, \tag{10.14}$$

where $\tau_R = 1/A$ is the *radiative lifetime* of this specific electronic state of the molecule. The higher the rate for spontaneous emission, the smaller the radiative lifetime of the molecule will be. In the absence of other processes that decrease the population of the state f, the radiative lifetime τ_R is equal to the total lifetime of the excited state.

In general however there are other contributions to this process that need to be taken into account. Indeed, the total life time of an excited state f will receive contributions from various mechanisms, in addition to spontaneous emission. The most important ones are:

- The internal conversion (IC) from an excited vibrational state to the vibrational ground state of a given electronic level, with constant rate k_{IC}.

- An inter-system crossing (ISC) from a singlet to a triplet state, with constant rate k_{ISC}

- Other processes such as *electron transfer* , with constant rate k_Q.

To study these electronic transitions within molecules, in general it is useful to use a special type of diagrams known as *Jablonski diagram*, which schematically indicates the transitions that can take place starting from an excited state. In Fig. 10.4 we show the schematic representation of the fluorescence and phosphorescence phenomena by means of a Jablonski diagram. In this diagram, A indicates the absorption of a photon, F and P represent the fluorescence and phosphorescence transitions respectively, S and T denote singlet and triplet states, and finally IC and ISC label the internal conversion and the inter-system crossing respectively.

Let us discuss the Jablonski diagram of Fig. 10.4. First of all, an electron in the electronic ground state absorbs an incoming photon and is promoted to an excited electronic state, as well as to an excited vibrational state. From the internal conversion (IC) process, this electron loses energy via non-radiative (vibrational) transitions until it is found in the ground vibrational state of the first excited electronic state S_1. From there on, two things can happen: either decay to the electronic ground state (*fluorescence*) or a inter-system crossing (ISC) to a triplet state, which eventually gives rise to a much slower light emission (*phosphoresce*) since the triplet to singlet transition is suppressed. We therefore see that an important component of the fluorescence process involves the molecule relaxing from an excited vibrational state to the lowest vibrational state via non-radiative transitions.

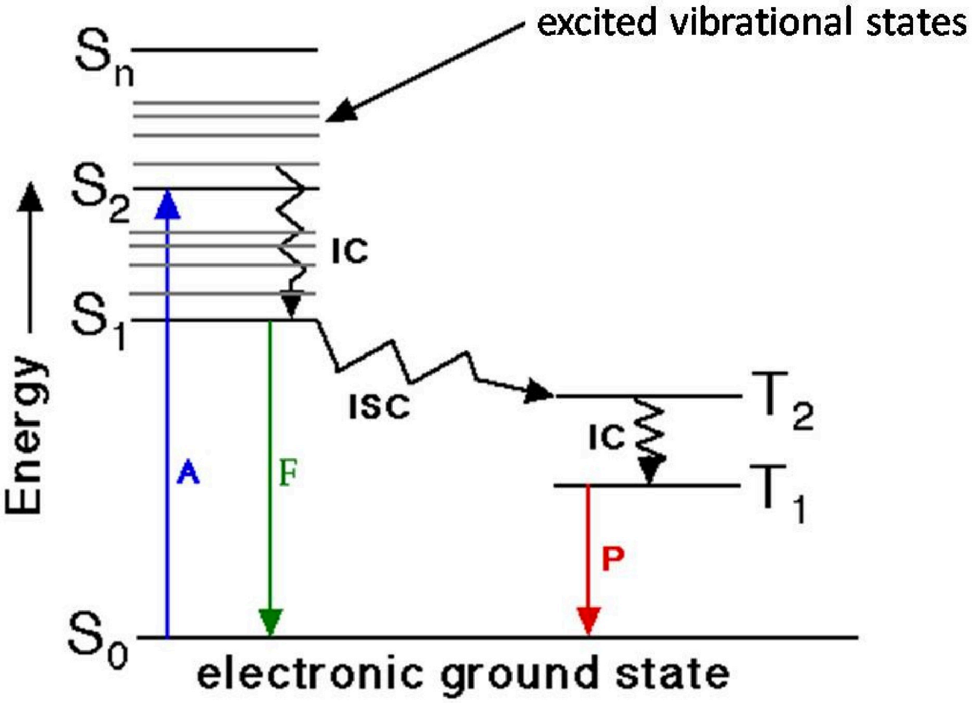


Figure 10.4: Schematic representation of the fluorescence and phosphorescence phenomena by means of the *Jablonski diagram*. In this diagram, A indicates the absorption of a photon, F and P represent the fluorescence and phosphorescence transitions respectively, S and T denote singlet and triplet states, and finally IC and ISC label the internal conversion and the inter-system crossing respectively.

As a result of the competition between the various processes that de-populate the excited state f, its total lifetime τ will be in general shorter than the radiation lifetime τ_R. The *total lifetime of the excited state* is called the *fluorescence lifetime* τ_F. Therefore, in general the radiative lifetime will be different (larger) from the fluorescence lifetime. Let us make this statement more quantitative. In the presence of these additional processes, the *occupancy number* of the electronic state f, N_f, will vary with time as follows:

$$\frac{dN_f}{dt} = -N_f \left(A + k_{IC} + k_{ISC} + k_Q \right) \quad \to \quad N_f(t) = N_0 e^{-(A+k_{IC}+k_{ISC}+k_Q)t} \, , \tag{10.15}$$

from where we readily see that the fluorescence lifetime τ_F will be determined by the sum of the rates of all processes that de-populate f, namely,

$$1/\tau_F = A + k_{IC} + k_{ISC} + k_Q \, , \tag{10.16}$$

In the fluorescence process, one usually introduces the *quantum yield* Φ_F defined as the number of emitted photons divided by the number of absorbed photons

$$\Phi_F \equiv \frac{\#\ \text{Emitted Photons}}{\#\ \text{Absorbed Photons}} = \frac{A}{A + k_{IC} + k_{ISC} + k_Q} \, , \tag{10.17}$$

Let us know take a look at the rates for these various processes in a representative molecule, in this case *clorofyll*. In this molecule, the rates for the various processes that de-populate the quantum state f are

$$\begin{aligned} k_R = A &= 0.5 \times 10^8 \text{ s}^{-1} \quad \rightarrow \quad \tau_R = 20\,\text{ns} \\ k_{ISC} &= 1.0 \times 10^8 \text{ s}^{-1} \quad \rightarrow \quad \tau_{ISC} = 10\,\text{ns} \\ kIC &= 1.5 \times 10^8 \text{ s}^{-1} \quad \rightarrow \quad \tau_{IC} = 16.6\,\text{ns} \end{aligned} \tag{10.18}$$

from where the fluorescence lifetime is $\tau_F = 3.3$ ns, and the fluorescence quantum yield $\Phi_F = 0.166$, meaning that for each 100 absorbed photons by the molecule, there will be 16 emitted photons.

As can be seen from the Jablonski diagram in Fig. 10.4, in general fluorescence radiation will correspond to longer wavelengths (that is, smaller energies) than the absorbed light, since the electrons undergo internal transitions losing energy before a photon is radiated. This difference is as known as the *Stokes shift*: the difference, either in wavelength or in wave number, between the absorption and fluorescence emission maxima.

Phosphorescence. In the *phosphorescence* radiation phenomenon, the molecule goes from a *singlet* excited state to a *triplet* excited state, via the so-called inter-system crossing. as illustrated in Fig. 10.4 (see also the left plot of Fig. 10.7). The resulting triplet states are long-lived, with lifetimes that can range from microseconds to seconds, since a transition to the ground state is forbidden by the selection rules (Pauli exclusion principle). Recall that a *singlet to triplet conversion* is a transition from paired spin to unpaired spin in a given molecular electronic state.

Therefore, as shown in Fig. 10.5, the origin of a *phosphorescence* transition is a transition from a ground singlet state (with paired spins) to the excited LUMO state where the electron spin remains the same. From there, a transition from the excited singlet state to the excluded triplet state via inter-system crossing will take place, since the latter has lower energy due to Hunds's rule. This inter-system crossing is enhanced by the spin-orbit coupling, and thus mostly occurs when heavy nuclei are nearby, such as S, Fe or Mg. The longevity of the *phosphorescence* radiation is then explain since the relaxation of a triplet to a singlet ground state is spin-forbidden.

The key property of phosphorescence is that, following light absorption, the subsequent re-emission takes place at a lower intensity for up to several hours after the original excitation. It is thus a very slow transition.

10.4 Lasers

The use of lasers is ubiquitous in spectroscopy, as well as in many forms of microscopy and in a wealth of medical applications. Therefore, it is important to understand how a laser works. The name *laser* was originally an acronym of *Light Amplification by Stimulated Emission of Radiation*, since it is a device that *amplifies the intensity* of the incident electromagnetic radiation. Let us now show how this is possible.

To illustrate how a laser works, we consider first of all the usual two-level system, such as the one depicted in Fig. 9.2. Since the Einstein coefficient for stimulated emission is B_{if}, the rate of $f \rightarrow i$ transitions leading to light *stimulated emission* will be given by $N_f B_{if} \rho(\nu)$, as we have demonstrated in HC10. Now, a necessary condition for *lasing* is achieving *population inversion* in the system, that is, a situation such that in thermal equilibrium the population of excited states is higher than the population of lower energy states. Only under this condition it will be possible to achieve more stimulated emission than stimulated absorption, and this achieve amplification of the incident radiation intensity. Otherwise, given that the Einstein coefficients for stimulated absorption and emission are the same, $B_{fi} = B_{if}$, and that in thermal equilibrium the ratio of

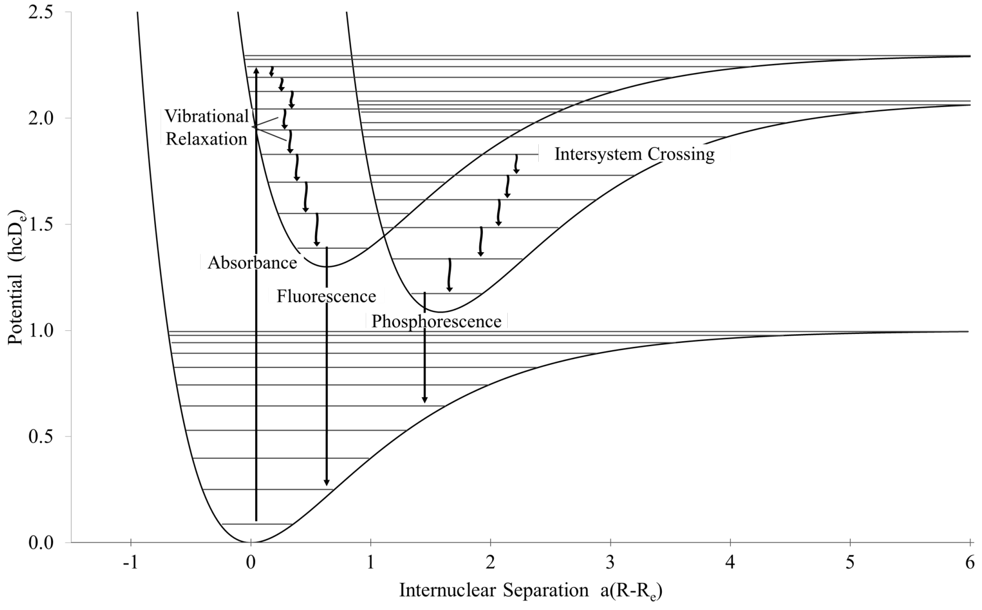


Figure 10.5: In *phosphorescence transitions*, the molecule undergoes an inter-system crossing from a singlet to a triplet state, and the resulting state has a very long lifetime since the relaxation from the excited triplet to the singlet ground state is not allowed by Pauli's exclusion principle.

populations of the ground state and the excited state are given by the Boltzmann distribution

$$\frac{N_f}{N_i} = \exp\left(-\frac{E_f - E_i}{k_B T}\right) , \tag{10.19}$$

it would be impossible to achieve more emission than absorption.

The concept of population inversion in the simple case of a two-level system is illustrated in Fig. 10.6: starting from a system in thermal equilibrium, obeying Boltzmann statistics and thus where the population of the excited state is smaller than that of the ground state, $N_f < N_i$, a *pumping system* increases the population of the excited state until it becomes larger than that of the ground state. Under these conditions, the overall rates for stimulated emission will be higher than those of the stimulated absorption, and lasing (light amplification) will take place. The crucial point for the laser operation is therefore how to best implement this pumping to achieve population inversion.

It is easy to realize is that actually *lasing is impossible in a two level system.* To see this, note that due to the equality of the Einstein coefficients, $B_{if} = B_{fi}$, the population of N_f can only be as large as 50% of the total population of the system, $N_T = N_i + N_f$, no matter how hard we pump. To implement the lasing concept one needs at least a system with three quantum states or even four. In Fig. 10.6 we show the schematic operation of a three-level and four-level laser, where we indicate the specific transition that is responsible for the lasing. Of all the possible transitions in the system, the one that is responsible for the lasing effect (and that requires population inversion) is $A \to X$ for the three-level laser and $A \to A'$ for the four-level laser, and therefore the pumping should be such that in equilibrium $N_A \geq N_X$ ($N_A \geq N_{A'}$) for the three (four) level laser. The best molecules to be used for lasing are thus for which the rates of inter-system transitions are those that facilitate the most achieving population inversion.

The basic mechanism of light amplification in a laser arises from the *snowball effect* applied to stimulated emission: each time that a stimulated emission takes place, additional photons with the correct energy are

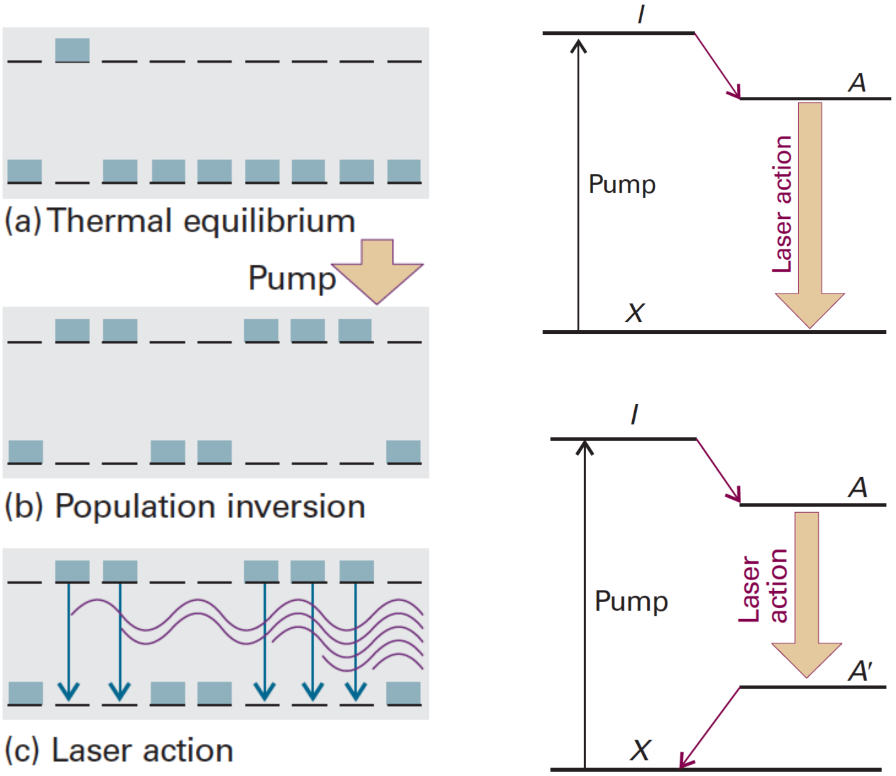


Figure 10.6: Left plot: the operation of a laser is based on the principle of *population inversion*: starting from a system in thermal equilibrium, obeying Boltzmann statistics and thus where the population of the excited state is smaller than that of the ground state, a *pumping system* increases the population of the excited state until it becomes larger than that of the ground state. In these conditions, stimulated emission will dominate over absorption and this the laser will indeed enhance light intensity. Right plot: The schematic operation of a three-level (upper diagram) and four-level (lower diagram) laser, where we indicate the specific transition that is responsible for the lasing.

produced, inducing yet further additional simulated emissions, with an exponential grow of the intensity of the incident electromagnetic radiation.

10.5 Photodynamic Therapy (PDT) (*)

Photodynamic therapy, or PDT for short, is a representative example of a medical application which is based on the optical transitions between different molecular electronic states. PDT, also known as *photochemotherapy*, is a form of photo-therapy involving light and a photosensitizing chemical substance, which in conjunction with molecular oxygen can eliminate cancerous tissue.

PDT works as follows. We start from a molecular singlet state, with two paired electrons (opposite spin) in the same electronic quantum level. This level will be the HOMO, the Higher Occupied Molecular Orbital, as illustrated in Fig. 10.7. Using laser radiation, an electron is excited to first the LUMO orbital and then via an inter-system crossing it ends up in its first excited triplet state ^3P. The resulting state photosensitizes the formation of an excited singlet state of O_2, $^1O_2$, which are extremely reactive and destroy cellular components. Therefore, the photochemical cycle that leads to the shrinkage of diseased tissue is the

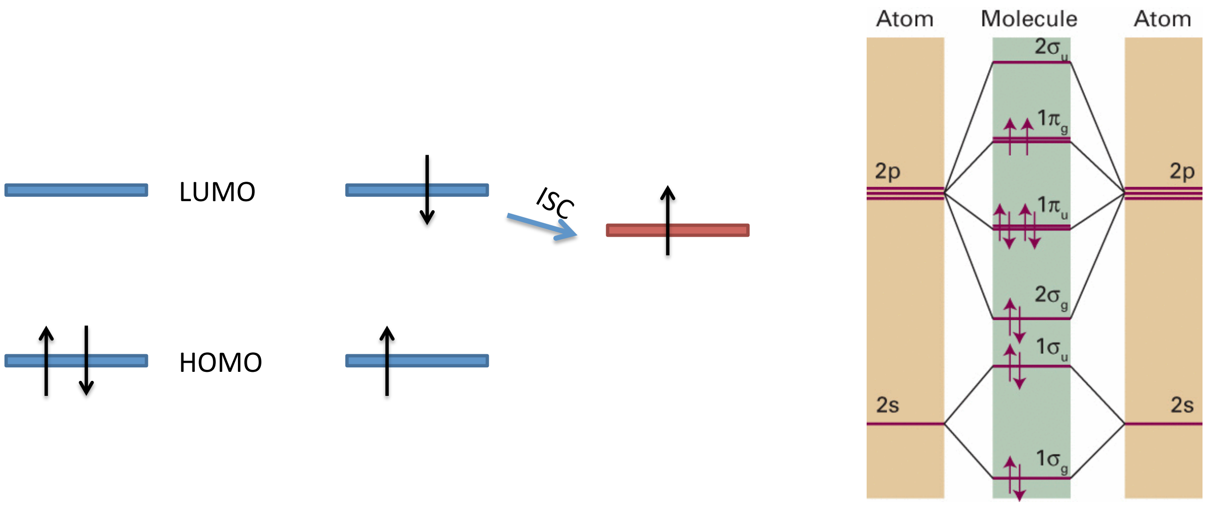


Figure 10.7: Left plot: the basic molecular electronic transition in *Photodynamic Therapy* is the one from a singlet state with paired electrons to a triplet state where the two electrons have unpaired spin via an inter-system crossing. Right plot: schematic representation of the molecular orbitals in the O_2 molecule, where the $2s$ orbitals form two σ-bonds and the $2p$ orbitals form two σ and two π bonds.

following:

$$\begin{aligned}
\text{Absorption} &\rightarrow \quad P + h\nu \rightarrow P^* \\
\text{Inter} - \text{system crossing} &\rightarrow \quad P^* \rightarrow {}^3P \\
\text{Photosensitization} &\rightarrow \quad {}^3P + {}^3O_2 \rightarrow P + {}^1O_2 \qquad (10.20) \\
\text{Oxidation reactions} &\rightarrow \quad {}^1O_2 + \text{reactants} \rightarrow \text{products}
\end{aligned}$$

Note that the energy in the triplet state is lower than that of the singlet state due to Hunds's rule, as illustrated in Fig. 10.7. In the right plot of Fig. 10.7 we show an schematic representation of the molecular orbitals in the O_2 molecule, where the $2s$ orbitals form two σ-bonds and the $2p$ orbitals form two σ and two π bonds. As we mentioned above, an inter-system crossing transition is enhanced by large spin-orbit couplings, and therefore occurs mostly where heavy nuclei such as S, Fe and Mg are nearby. Since the relaxation from the triplet state to the singlet ground state is forbidden by Pauli exclusion system (which forbids two electrons with the same spin in the same quantum state), the electronic configuration displayed in Fig. 10.7 has a remarkable longevity, which facilitates the photosensitization of a large number oxygen molecules. This is the same mechanism that explained the very long lifetime of phosphorescence.

Therefore, PDT is based on the photosensitization of O_2 molecules to turn them into extremely reactive and destroy cancerous tissue. In this technique, first of all the patient is injected with a photosensitive molecule, and one waits until this substance accumulates in the tumor. Then the tumor is illuminated with laser light, induces the singlet to triplet transition in the molecule and which leads to the selective formation of reactive ${}^1O_2$ and the restriction of cancerous tissue. This technique is mostly used for cancer present in body cavities such as mouth, throat or bladder. A summary of the PDT mechanism is shown in Fig. 10.5.

Summary

Some of the important lessons that we have learned in this lecture are the following:

- The absorption of light (visible and UV) in biological molecules takes place predominantly by means

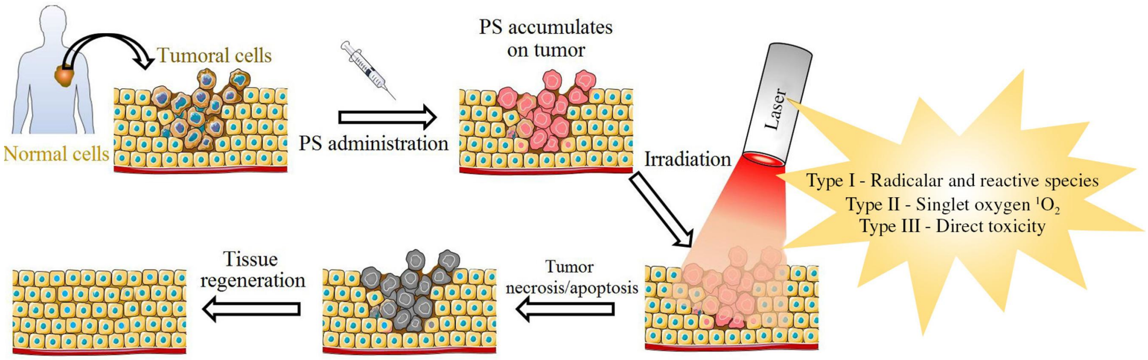


Figure 10.8: In photo-dynamic therapy, a photo-sensitive (PS) molecule is injected on the patient and accumulates on tumor. The subsequent irradiation activates a chemical reaction which creates reactive oxygen molecules $^1O_2$ which destroy the cancerous tissue.

of $\pi \to \pi^*$ transitions in p-conjugated electron systems.

- A larger p-electron conjugation length leads to absorption at longer wavelengths, that is, at smaller photon energies. This is explained by the reduced electron confinement for large conjugation lengths.
- Symmetry considerations are important when determining whether or not an given optical transition is possible within a specific molecule.
- The Franck-Condon principle explains the transitions in which a *simultaneous* change in the vibrational and electronic states of a molecule takes place, the so-called *vibronic transitions*.
- The Einstein coefficient for spontaneous emission determines the radiation lifetime of a molecule τ_R, which in general is different (larger) than its total lifetime.
- The fluorescence lifetime τ_F is usually shorter than the radiative lifetime τ_R due to competing processes that de-populate the excited states.
- The phosphorescence phenomenon is explained by the long-lived emissions from a triplet excited state to a singlet ground state.
- The light amplification achieved in a laser is based on the concept of *population inversion*.
- The Photodynamic Therapy is an representative example of a medical application based on optical transitions in biomolecules.

Further reading

Further reading material about the topics covered in this lecture can be found in the following sections of the course textbook (*Physical Chemistry*, Atkins and De Paula, 10th edition):

- **12D.4**: Vibration-rotation spectra.
- **12E.2**: infrared absorption spectra.
- **12E.3**: Vibrational Raman spectra.
- **13A.1**: Diatomic molecules.
- **13B.1**: Fluorescence and phosphorescence.
- **13C.1**: Population inversion.

11 HC12: Magnetic Nuclear Resonance

van Quantum tot Molecuul

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

In this final lecture of the course, we will explore the important topic of Magnetic Nuclear Resonance and show it is the basis for one of the most important applications of quantum theory in medical therapy, namely *Magnetic Resonance Imaging*.

The **learning goals** of this lecture are:

(a) Understand that the atomic nucleus has a magnetic moment, and that upon interaction with a magnetic field its energy levels split for different values of z component of the nuclear spin.

(b) Understand and be able to interpret Nuclear Magnetic Resonance (NMR) spectra.

(c) Understand how pulse techniques can be employed in order to measure NMR spectra.

(d) Become familiar with the basic principles of Magnetic Resonance Imaging (MRI).

11.1 Nuclear spin

Classically in the presence of an external magnetic field $\vec{B}$ a particle with a magnetic moment $\vec{\mu}$ has associated an energy $E = -\vec{\mu} \cdot \vec{B}$. In quantum theory, promoting observables to operators, we find that the Hamiltonian that describes this interaction is

$$\hat{H} = -\hat{\vec{\mu}} \cdot \vec{B} \,. \tag{11.1}$$

An electron moving around an atomic nucleus has associated an associated *angular momentum* $\vec{l}$ and thus a magnetic moment, so in this case the Hamiltonian operator reads

$$\hat{H} = -\gamma_e \hat{\vec{l}} \cdot \vec{B} \,, \quad \gamma_e \equiv -\frac{e}{2m_e} \,, \tag{11.2}$$

where γ_e is the *magnetogyric ratio of the electron*, also known as gyro-magnetic ratio.

Considering without loss of generality a magnetic field with magnitude B_0 pointing in the z direction, the Hamiltonian will be $\hat{H} = -\gamma_e B_0 \hat{l}_z$. From our previous discussion of angular momentum in quantum

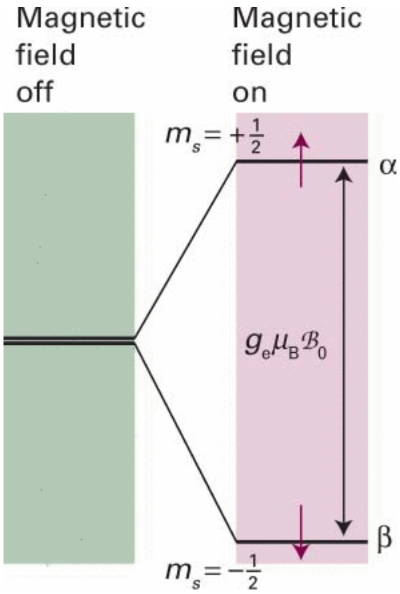


Figure 11.1: In the presence of an external magnetic field B_0, a given electronic energy state splits into two different states due to the different contribution to the $\vec{s}\cdot\vec{B}$ coupling to the total energy. These two states are separated by an energy shift of $\Delta E = g_e\mu_B B_0$.

mechanics, we know that this Hamiltonian will have as eigenvalues:

$$E_{m_l} = -\gamma_e m_l \hbar B_0 \equiv \mu_B m_l B_0 \,, \quad \mu_B \equiv \frac{e\hbar}{2m_e} = 9.27\times 10^{-24}\,\mathrm{J\,T^{-1}}\,, \tag{11.3}$$

where the constant μ_B is known as *Bohr's magneton*, and m_l is the quantum number of angular momentum in the z direction.

The result above is the contribution from the electron energy from the interaction of its *orbital angular momentum* with an external magnetic field. However, an electron also has an *intrinsic angular momentum*, its spin, which will also need to be accounted for. In this case the corresponding Hamiltonian will be

$$\hat{H} = -g_e\gamma_e\vec{B}\cdot\hat{\vec{s}}, \tag{11.4}$$

with $g_e \simeq 2.002319$ is known as the *electron's g-factor*, which deviates from 2 due to relativistic quantum effects, and $\hat{\vec{s}}$ is the spin operator. As in the case of the angular momentum, for a magnetic field in the z direction, the contribution to the total electron energy of the interaction between the electron spin and the magnetic field will be:

$$E_{m_s} = g_e\,\mu_B\,m_s\,B_0\,, \tag{11.5}$$

where of course now $m_s = \pm 1/2$. Therefore, for $B_0 \neq 0$ a given energy level will split due to this effect by an amount $\Delta E = g_e\mu_B B_0$, as shown in Fig. 11.1.

The interaction between a magnetic moment and an external field can also be represented within the *vector model of the spin*, illustrated in Fig. 11.2. In this picture, the external magnetic field induces a *precession* of the electron spin, defined in general as the change in the orientation of the rotational axis of a rotating body. The magnetic moment of the electron will experience a torque $\vec{\Gamma} = \vec{\mu}\otimes\vec{B}$, and thus the

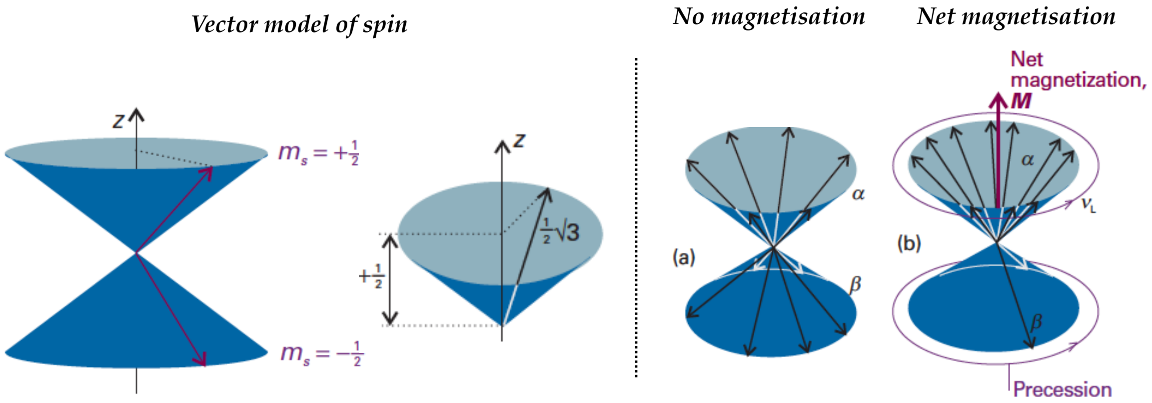


Figure 11.2: Left plot: in the vector model of electron spin, spin can be represented by a vector of length $\sqrt{3}/2$ whose component in the z direction is $\pm 1/2$. Right plot: in the absence of an external field, there is no net magnetization of the electron spin. For an external magnetic field, the electron spin develops a net magnetization, represented as a precession around the z axis.

magnetic moment will rotate around the axis defined by the magnetic field with frequency

$$\nu_L = \frac{|\gamma_e B_0|}{2\pi}\,, \tag{11.6}$$

which is known as the *Larmor frequency*, and play an important role in the following discussion.

The splitting of energy levels due to the interaction between a magnetic field and an angular momentum can be used for the so-called *magnetic resonance*. Let us first discuss we case of electrons, and then we will turn to nuclei. Since the magnetic field induces an energy splitting of $\Delta E = g_e \mu_B B_0$, a transition between these two energy levels will have associated a frequency

$$\nu = \frac{g_e \mu_B B_0}{h} = \frac{g_e e B_0}{4\pi m_e}\,, \tag{11.7}$$

where we have use the definition of Bohr's magneton. Eq. (11.7) is known as the *resonance frequency* for electrons: absorption of electromagnetic radiation for this value of ν will thus be favored. For a reasonably strong magnetic field of $B = 0.3$ T, we find that this frequency is $\nu \simeq 10$ GHz, in the domain of *microwave radiation*. Note that the Boltzmann distribution predicts that the state with lower energy will have higher population, and this the absorption of EM radiation at these frequencies will be facilitated. Note also that this transition can take place only with the condition of paired electrons between the initial and final levels.

Now, *any* particle that has associated an angular momentum will exhibit similar interactions with an external electric field. What about the *atomic nucleus* then? Atomic nuclei are composed by protons and neutrons, each of which are fermions, that is, *spin 1/2* particles. Their angular momenta will add up and in general the nuclear spin I will be different from zero. And from there on the usual properties of angular momentum follow, for instance, a nucleus with spin I will have associated an intrinsic angular momentum of $\sqrt{I(I+1)}\hbar$, and its components $m_I\hbar$ along a given axis will take the values $m_I = I, I-1, \ldots, -I+1, -I$.

The resulting splitting of the energy levels will be similar to the electron case, with the only difference that now the atomic nucleus has a different magnetic moment. Therefore, we will have that the energies associated to the nuclear spin in a external magnetic field are given by

$$E_{m_I} = -\gamma\hbar B_0 m_I\,, \quad \gamma\hbar \equiv g_I \mu_N \quad \mu_N = \frac{e\hbar}{2m_p} = 5.051 \times 10^{-27}\,\mathrm{J\,T^{-1}}\,, \tag{11.8}$$

where we have introduced the *nuclear magneton* μ_N and the nuclear g-factor g_I. In the specific case of a $I = 1/2$ nucleus, the energy levels will split into two with a energy difference $\gamma\hbar B_0$, and now the resonance frequency Eq. (11.7) for $B = 12$ T will be $\nu \simeq 500$ MHz, in the regime of *radio-waves.* This very significant change in the resonant frequency of nuclei as compared to electrons is explained from the fact that $\mu_N \ll \mu_e$ because of the much larger nucleon mass $m_N \gg m_e$. Therefore, nuclear magnetic resonance can be achieved for radiation *much smaller energies* (longer wavelengths) than those of the electron case.

11.2 The chemical shift

In practice, the interaction between atomic nuclei and external magnetic fields will be partially shielded due to the surrounding electrons (atomic and/or molecular), and therefore nuclei experience a smaller magnetic field,

$$\widetilde{B}_0 = B_0 - \delta B \equiv B_0 \left(1 - \sigma\right) \, , \tag{11.9}$$

where σ is the so-called *shielding constant*, and its value depends very sensitively on the specific chemical structure surrounding the nucleus. As a consequence of this shielding, for a fixed value of the external magnetic field B_0, nuclei of different chemical identify will have associated *different resonant frequencies.* We call this *chemical shift* the measure of the differences in the resonance frequency due to the shielding effects of the electrons, which allows to discriminate not only between different chemical elements but also between the same element in different locations within a molecule.

Given that the local magnetic field experienced by the nucleus is $B_0 \left(1 - \sigma\right)$, the associated Larmor frequency will also be reduced, and will be given by

$$\nu_L = (1 - \sigma)\frac{\gamma B_0}{2\pi} \, . \tag{11.10}$$

Then we can define the *chemical shift* δ as the shift of the Larmor frequency relative to a reference value ν_0, that is

$$\delta \equiv \frac{\nu_L - \nu_{L,0}}{\nu_{L,0}} \times 10^6 \, , \tag{11.11}$$

which is roughly independent of the specific value of B_0, which cancels partially in the ratio. Therefore, if for example we have a reference Larmor frequency of 500 MHz for a given field B_0, then a chemical shift of $\delta = 1$ ppm (parts per million) corresponds to a shift in the Larmor frequency ν_L as compared to the reference value of 500 Hz.

To illustrate the concept of chemical shift, in Fig. 11.3 we show a schematic representation of the ethanol molecule, CH_3CH_2OH. In the same figure we also show chemical shift δ associated to different hydrogen nucleus in the ethanol molecule. We observe that the closer to the oxygen nucleus, the larger the chemical shift. This can be explained since oxygen is the most electro-negative element of the molecule and thus attracts the most the molecular electron density, leading to an enhanced chemical shift of the neighboring hydrogen nucleus. From Fig. 11.3 we also see that there is a *fine structure* for the each nucleus in the ethanol molecule. This fine structure arises from the *spin-spin coupling*, where the spin from one nucleus contributes to the local magnetic field of a neighboring nucleus.

11.3 NMR spectroscopy

We now want to illustrate how we can measure NMR spectra and chemical shifts, as required for medical applications. This can be achieved by means of the *NMR spectrometer*, and requires introducing the concept of the *pulsed technique.*

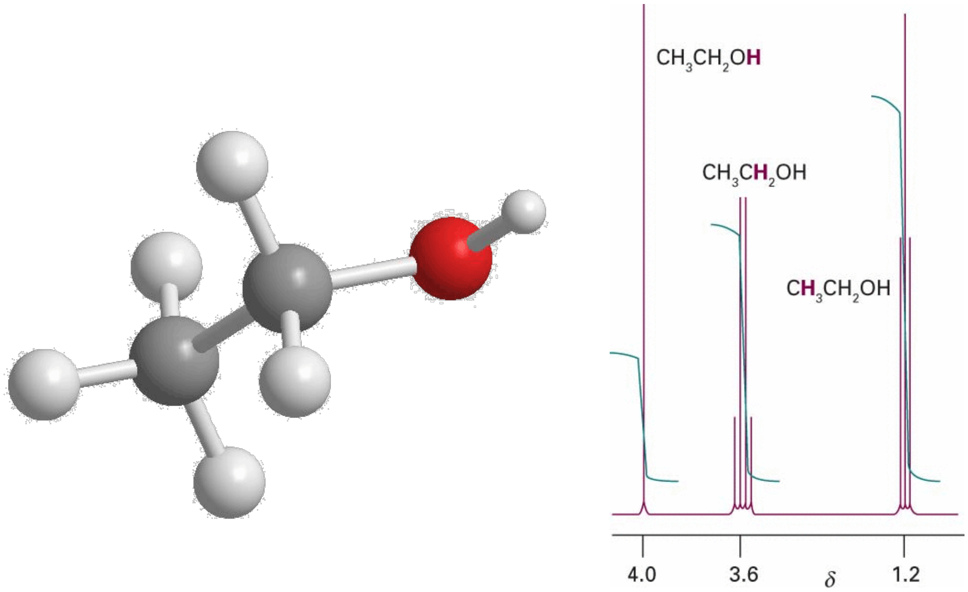


Figure 11.3: Right plot: schematic representation of the ethanol molecule, CH_3CH_2OH. Left plot: the chemical shift δ associated to different hydrogen nucleus in the ethanol molecule. We observe that the closer to the O nucleus, the larger the chemical shift.

The basic ingredient of NMR is the measurement of the *chemical shifts* for each atom in a molecule of the analysed sample, which then allows its identification by comparing with the reference spectra. One possibility would be to measure the *rate of absorption of radio waves* as a function of their frequency ν, which should be dominated by the Larmor frequencies of the nuclei in the sample. As illustrated by Fig. 11.1 (applied to the magnetic moment of nuclei), the value of the photon energy that induces the resonant transition $\beta \to \alpha$, $E_\gamma = \gamma B_0/2\pi$, lies in the radio part of the EM spectrum. This approach is however challenging, since first of all the population difference between α and β, the nuclear atomic states with different spin, is small, and second because such a frequency scan is a rather slow process. Indeed, the difference in population between α and β can be as small as $\mathcal{O}\left(10^{-5}\right)$.

The alternative is to use *pulse techniques*, which are much more efficient. To understand the pulse method works, we consider a vector model for the nuclear spin, such as that shown in Fig. 11.2 for the electron. Let us assume that the nuclear spin is $I = 1/2$ for simplicity. The length of this vector is $\sqrt{I(I+1)} = \sqrt{3}/2$ (in units of $\hbar$), and the projection in the z direction is fixed to be $I_z = \pm 1/2$. Recall that the operators $\hat{I}_x$, $\hat{I}_y$ and $\hat{I}_z$ do not commute among them.

For a vanishing external magnetic field, $B_0 = 0$, the energy of the two states α and β are the same, and so they have the same average population. Once we switch on the external magnetic field, $B_0 \neq 0$, the energy of the two levels splits, with $E_\beta < E_\alpha$ (see Fig. 11.1). Due to Boltzmann statistics, we know that the population of the lower energy state α will be larger than that of the higher energy one β, $N_\alpha < N_\beta$. This implies that a *net magnetization* is created in the analyzed material, proportional to $N_\beta - N_\alpha$.

The next step in the pulse method is that, in addition to the static magnetic field B_0, we also add an *oscillatory magnetic field* in the (x, y) plane, with Larmor frequency ν_L and circular polarization so that B_1 rotates in the (x, y) plane with frequency ν_L. Recall that the static magnetic field B_0 is applied along the z direction, and is thus perpendicular to the new oscillatory magnetic field. This additional field will induce the transition $\alpha \leftarrow \beta$, that is, from the lower to the higher energy nuclear magnetic states. If we now transform to the *rotating frame* with Larmor frequency ν_L, the nuclear spins will experience a *constant*

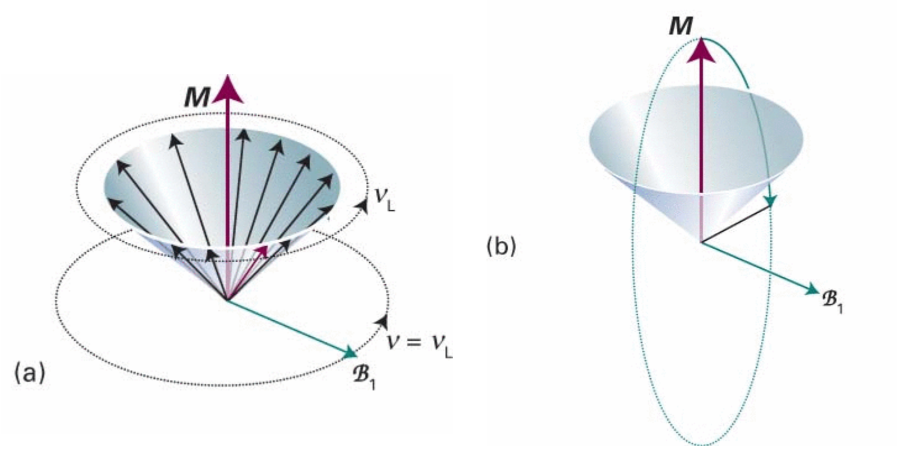


Figure 11.4: In the pulse method, a magnetic field polarized in the (x, y) plane oscillates with frequency ν_L (left plot), perpendicular to the nuclear spin. In the rotating reference frame, the nuclear spin experiences a constant magnetic field B_1 (right plot).

magnetic field B_1. Therefore, in this frame, the nuclear spins will exhibit a precession with a different value of ν_L, as illustrated in Fig. 11.4. This new Larmor frequency will be

$$\widetilde{\nu}_L = \frac{\gamma B_1}{2\pi}\,, \tag{11.12}$$

and thus the period of the new Larmor precession will be

$$T = \frac{2\pi}{\gamma B_1}\,. \tag{11.13}$$

In the pulse method, rather than having the oscillatory field B_1 activated the whole time, we have it on only for a small amount of time, the *pulses*. In particular the duration of this pulses should be one quarter of the total precession period, $(1/4)\cdot(2\pi)/\gamma B_1$, corresponding to am angular rotation of $\pi/2$, which typically corresponds to a duration of microseconds. Now, while in the rotating frame the magnetization M is constant in the (x, y) plane, in the *laboratory frame* it will rotate with Larmor frequency ν_L. This oscillatory magnetization can be detected in the *coil* of an *NMR spectrometer*, where it can be transformed into output electromagnetic radiation in the radio spectrum.

A characteristic signal of the measured MNR signal is that it will appear as a Free Induction Decay (FID), which oscillates with Larmor frequency ν_L but whose amplitude decays exponentially due to the finite duration of the length of the pulse. The oscillating signal detected in the MNR spectrometer can then be Fourier-transformed to determine the values of the Larmor frequencies for each of the nuclei in the studied sample, as shown in Fig. 11.5. This way, in a sample with different nuclei and/or nuclei with different chemical shifts, the Fourier transform will exhibit multiple peaks corresponding to each of the relevant Larmor frequencies.

The reason why the MNR signal in Fig. 11.5 decays exponentially with time can be undertood as follows. After a $\pi/4$ pulse, the magnetization of the atomic nuclei is not in thermal equilibrium with its surroundings. In particular, since β has a higher population than α, the pulse will affect more the former level and thus contradict the expected Boltzmann populations. Indeed, after the $\pi/4$ pulse the population of spins in the two levels is similar. After a some time, following the pulse, the system will be back to thermal equilibrium

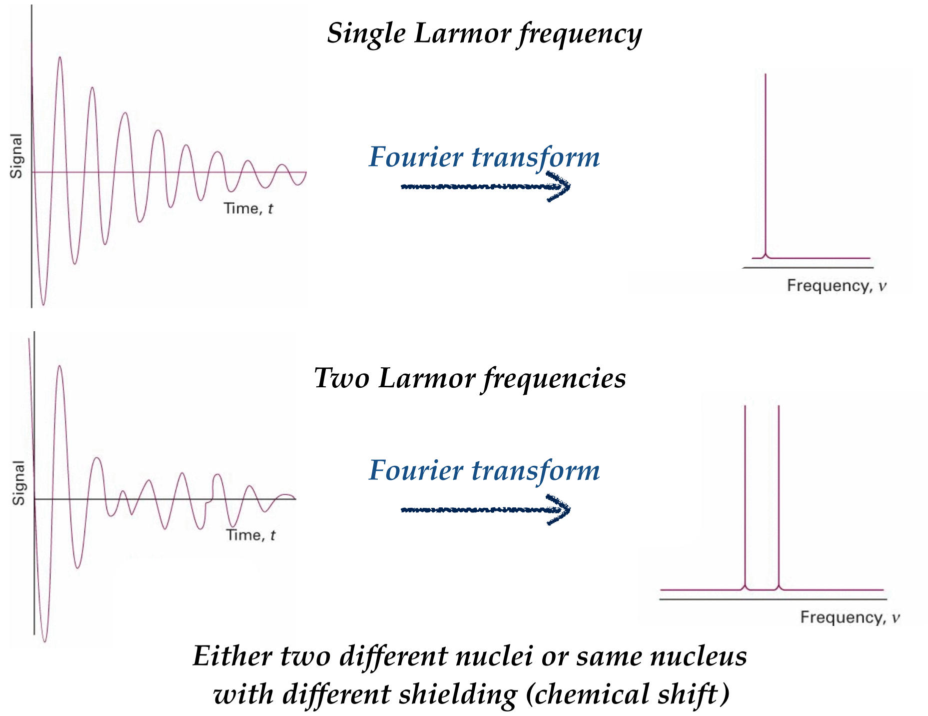


Figure 11.5: The oscillating signal measured in the MNR spectrometer can be Fourier transformed to determine the values of the Larmor frequencies present for the studied sample. In a sample with different nuclei and/or nuclei with different chemical shifts, the Fourier transform will exhibit multiple peaks corresponding to each of the relevant Larmor frequencies.

where

$$\frac{N_\alpha}{N_\beta} = e^{-h\nu_L/k_BT} \, . \tag{11.14}$$

This *longitudinal relaxation time* T_1 will be given by the time it takes for the system to recover its original magnetization M_0 along the z direction, namely

$$(M_z(t) - M_0) \propto e^{-t/T_1} \, , \tag{11.15}$$

where we see that for $t \gg T_1$ the magnetization will be $M_z(t) = M_0$, as was the case before the pulse.

Moreover, after the $\pi/4$ pulse there will also be a new magnetization M_y along a given direction in the (x, y) plane, that is, the phase of individual nuclear spins along this plane (in the vector model of the atomic spin) will be common. Some time after the pulse, this common phase will be lost and the associated macroscopic magnetization in the transverse direction M_y will disappear. We can then define the *transverse relaxation time* T_2 as

$$M_y(t) \propto e^{-t/T_2} \, , \tag{11.16}$$

so that for $t \gg T_2$ we have that $M_y \to 0$, as was the case before the pulse. The longitudinal relaxation mechanism also affects the transversal one, so typically $T_1 \geq T_2$, of the order of a few seconds. The main impact of the T_2 relaxation is the *broadening of spectral lines.* Indeed, it can be shown that the *width at half-maximum* of a MNR line is given by

$$\nu_{1/2} = \frac{1}{\pi T_2} \, . \tag{11.17}$$

The experimental apparatus for the measurement of NMR spectra is illustrated in Fig. 11.6. First of

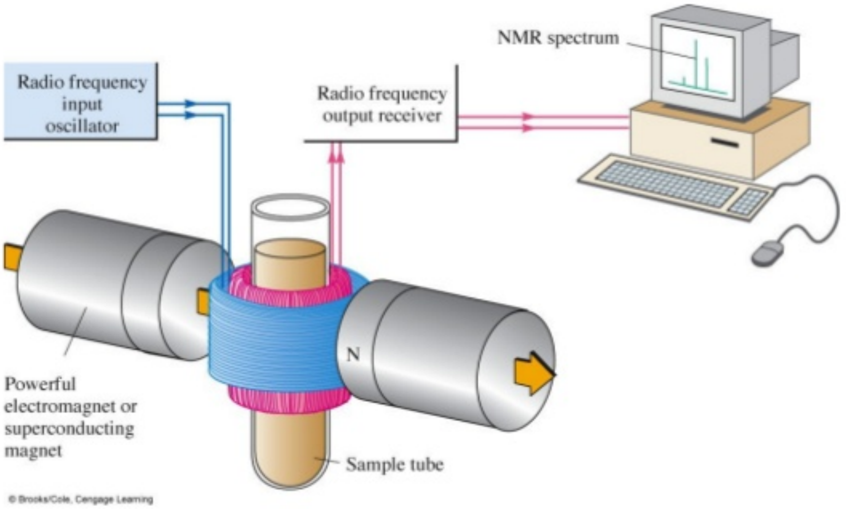


Figure 11.6: In an NMR spectrometer, a powerful magnet is used to apply a strong magnetic field to the studied sample. Then, the sample is irradiated with pulsed electromagnetic radio waves and the response of the system under this radiation (which frequencies are absorbed) is monitored to determine the resultant NMR spectra.

all, in an NMR spectrometer, a powerful magnet is used to apply a strong magnetic field B_1 to the studied sample. Then, *pulsed electromagnetic radio waves* and the response of the system under this radiation (which frequencies are absorbed) is monitored to determine the resultant NMR spectra. The frequency of the pulsed radio waves is tuned to reproduce the Larmor frequencies of specific magnetic nuclear resonances, as explained above. The Fourier transform of the output radio frequency, as indicated in Fig. 11.5, indicates the Larmor frequencies of the various atomic components of the sample.

11.4 Magnetic resonance imaging (MRI)

Magnetic Nuclear Resonance is the key principle underlying *Magnetic Resonance Imaging.* This imaging method is based on the deliberate application of an *inhomogeneous magnetic field*, in other words, a *magnetic field gradient.* As we have discussed, each molecule of the imaged sample, in the presence of a external magnetic field, will have associated a Larmor frequency of

$$\nu_L(\vec{r}) = (1-\sigma)\frac{\gamma B_0(\vec{r})}{2\pi} \, , \tag{11.18}$$

which now will be different for each position $\vec{r}$ of the sample, due to the inhomogeneity of the magnetic field. By comparing the observed Larmor frequencies with the expected values for some molecules and taking into account the magnetic field gradient applied to the sample, we can reconstruct a *three-dimensional* image of the chemical components of the sample.

The typical MRI application is based on the proton resonance, therefore the largest signal will come from those body parts with the highest concentration of water. To enhanced the contract, it is possible to use that the longitudinal and transverse relaxation times T_1 and T_2 depend on the time of local environment, for example they are different in membranes than in bulk water. This way, it is possible to image separately using MRI different types of soft tissue such as gray matter, white matter or spinal fluid. In Fig. 11.7 we illustrate how by means of Molecular Resonance Imaging it is possible to obtain three-dimensional images of soft tissue, in this case white and grey matter inside the skull.

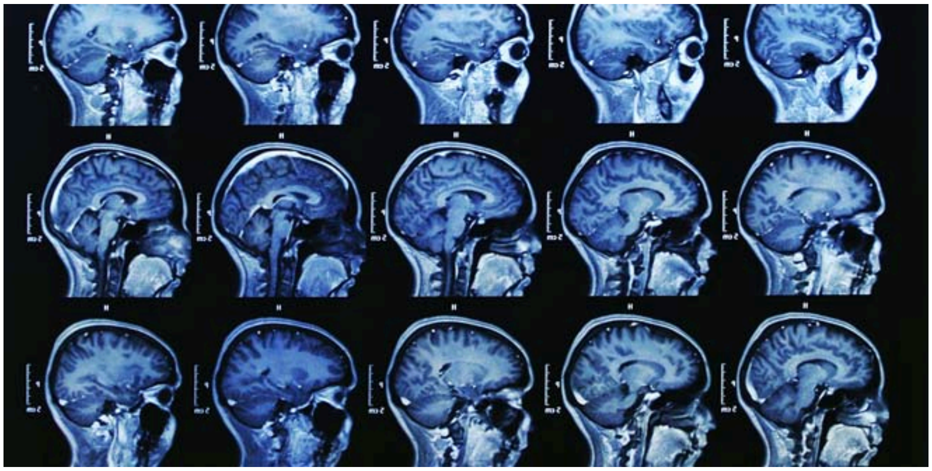


Figure 11.7: By means of Molecular Resonance Imaging, it is possible to obtain three-dimensional images of soft tissue, in this case white and grey matter inside the skull.

Summary

Let us summarize some of the important topics that have been covered in this lecture:

- A moving electric charge has an associated *angular momentum* $\vec{l}$, which can then interact with a external magnetic field $\vec{B}$.
- The interaction of the electron spin with an external magnetic field induces an energy splitting that can be proved by resonant absorption at microwave frequencies.
- Atomic nucleus also have an intrinsic angular momentum, spin, built upon the spin of its components, protons and neutrinos. Depending on the specific nucleus, this spin can be zero, half-integer or integer.
- The nuclear magneton μ_N is much smaller than the electron magneton μ_B due to the mass suppression of the atomic nuclei as compared to the electron mass.
- Magnetic resonance absorption in nuclei takes place at radio frequencies.
- The chemical shift quantifies the degree of shielding of nucleus by core electrons in the presence of external magnetic field.
- Magnetic nuclear resonances can be efficiently probed using the *pulse method*, were short pulses of $\pi/4$ angle are used to excite the magnetic resonances.
- The MNR signal decays exponentially die to the longitudinal and transverse mechanism of spin relaxation.
- Magnetic nuclear resonance is the key for Magnetic Resonance Imaging, where a magnetic field gradient can be used to obtain a three-dimensional image of soft tissues.

Further reading

Further reading material about the topics covered in this lecture can be found in the following sections of the course textbook (*Physical Chemistry*, Atkins and De Paula, 10th edition):

- **14A.1**: Nuclear magnetic resonance.
- **14B.1**: The chemical shifts.
- **14B.2**: The origin of shielding constants.
- **14C.1**: The magnetisation vector.