# An Introduction to Quantum Mechanics

*SERP-CHEM: 2011-16*

# Foreword

These are the lecture notes for an introductory course on quantum mechanics which took place at the University of Genova during the first semester of the 2012/15 academic years. The course has been designed for the students of the SERP-CHEM program and it will last for about 10 weeks. The students attending the course have had no previous exposure to the subject and their academic training has focused mostly on chemistry and material science.

There are a vast number of books, new and old, of different degrees of difficulty and mathematical sophistication, that provide a self-sufficient account of quantum mechanics. A very partial and, to a large extent, arbitrary list of some of these, selected from the relatively modern and less technical undergraduate textbooks, is presented in the Bibliography of these notes. These notes, therefore, are in no way intended to be a replacement for a standard textbook on quantum mechanics. Their purpose is different.

Ten weeks is certainly much too brief a span of time to give a comprehensive view of even the most elementary aspects and applications of quantum mechanics. The topics covered by most textbooks, including those listed in the Bibliography, provide material sufficient for a course of a minimum of 2 semesters in an undergraduate program in physics. Therefore, these lecture notes are intended primarily to serve as a list of those (few) topics that I have selected for discussion in class.

Moreover, I think that any introductory course in quantum mechanics faces a dilemma over whether to adopt a more or less historical approach to the topic — whether to explain and motivate the logical steps that lead to the change of paradigm that is inherent in the quantum mechanical description of physical phenomena, in other words, or, alternatively, to introduce from the start the new framework and derive from it the plethora of microscopic phenomena that quantum mechanics correctly accounts for.

I've opted for a sort of compromise between these two extremes. The first section of these notes, which deals with material that will occupy the first weeks of class, provides a condensed and simplified review of the so-called "crisis" of classical physics which led to the discovery of quantum mechanics. In the second part, which in principle could be read independently from the first, I introduce in a somewhat axiomatic way, but with no pretense of mathematical rigor, the rules of the quantum mechanical descriptions of nature. It is possible that this presentation is not completely identical to those found in the elementary textbooks and I therefore thought that it might be useful for the students to have it written down. In the last sections, I list a series of applications of quantum mechanics to atomic and molecular physics; some of these will be discussed in class and some will be assigned as exercises.

Finally, and most importantly, the text of these notes is interspersed with problems, of varying degrees of difficulty, which fill out the logic of the accompanying explanations and which the students are therefore strongly encouraged to solve between class meetings.

I should emphasize again that these notes are not self-sufficient: students must use the texts listed in the Bibliography for all the topics that I will present in a conventional fashion. The notes dwell on those specific topics that will either be discussed in a somewhat non-standard way or that I feel might require more detailed explanation for students without a background in physics.

While preparing these lecture notes, I realized that my presentation of the material is frequently inspired by the first class on quantum mechanics that I ever took, a course that was taught, many years ago, at the University of Pisa, by Luigi E. Picasso. The lecture notes of L. E. Picasso's course on quantum mechanics have since become a book, *Lezioni di Meccanica Quantistica, (Edizioni ETS, Pisa, 2000)*, which is, in my opinion, one of the best introductory textbooks of quantum mechanics ever written. Unfortunately for my SERP students, and for numerous others around the globe who do not read Italian, the book has only been published in this language. I would like to take this opportunity to express both my admiration for the clarity of his lessons and my gratitude for such an exciting introduction to this fascinating topic. I would also like to apologize to my former teacher for any superficiality, distortion or outright misunderstandings that may be found in the re-elaboration of his lessons that form these notes.

# Bibliography

General textbooks:

- Quantum Mechanics, Vol 1 e 2: C. Cohen-Tannoudji, B. Diu, F. Laloe, (John Wiley & Sons., 1977).
- The Feynman Lectures on Physics, Vol 3 (Quantum Mechanics): Richard Phillips Feynman, Robert B. Leighton and Matthew Sands (1966).
- Introduction to Quantum Mechanics: David J. Griffiths (Benjamin Cumming, 2004).
- Modern Quantum Mechanics: J. J Sakurai (Addison Wesley, 1993).
- Quantum Physics: Michel Le Bellac (Cambridge University Press, 2006)

An elementary introduction to quantum mechanics, relativity and statistical mechanics designed for undergraduate students of technical curricula:

- Introduction to the Basic Concepts of Modern Physics: C.M. Becchi and M. D'Elia (Springer, 2007).

An elementary introduction to quantum mechanics designed specifically for undergraduate students of chemistry, material science and biology:

- Introduction to Quantum Mechanics: S.M. Blinder (Elsevier, 2004).

# Contents

# 1 The Failure of Classical Physics

## 1.1 Atomic stability

The idea of atoms — and the idea that to each chemical element there corresponds an atom — was introduced to explain basic phenomena in chemistry. Other phenomena (electric conduction, light emission from matter, electrolysis, etc) show that atoms, although neutral, contain positive and negative electric charges. Negative charges inside atoms turn out to be all equals among themselves: they are called electrons. Their mass and charge were measured (Thomson, Millikan) to be

$$e = -1.6 \times 10^{-19}\, C \qquad m_e = 0.9 \times 10^{-27}\, g \tag{1.1}$$

Atomic masses are much bigger. For example, for hydrogen,

$$m_H \approx 1.7 \times 10^{-24}\, g \approx 1836\, m_e \tag{1.2}$$

The linear dimensions of atom turn out to be of the order of $10^{-8}\, cm \equiv 1$Å.

The atomic model which imposed itself after the first decade of the XX century is named after Rutherford (1911)[1]. Rutherford's hypothesis is that an atom is made of a nucleus which contains positive charges and whose size is much smaller than the size of the atom: negatively charged electrons orbit around the nucleus much like the planets of the solar system.

In this model the dimension of the atom is linked to the size of the electrons orbits which, in turn, is determined by the energies of the electrons. Take for example the equation of dynamics for the circular electronic orbits of the electron of the hydrogen atom:

$$|\vec{F}| = \frac{e^2}{4\,\pi\,\epsilon_0\, r^2} = m\,|\vec{a}| = \frac{m\, v^2}{r} \tag{1.3}$$

---

[1] An alternative model, formulated by Thomson, was seriously considered in the first decade of the XX century. Experiments involving the diffusions of $\alpha$ particles through sheets of gold appeared to invalidate Thomson model and favour Rutherford hypothesis.

where $\vec{a}$ is the acceleration and $m$ is the reduced mass[2] of the electron-proton system. Since the proton is much heavier than the electron, $m$ is approximately equal to the electron mass. From (1.3) one obtains

$$r = \frac{e^2}{4\,\pi\,\epsilon_0\,m\,v^2} \tag{1.4}$$

The energy $E$ of the electron is

$$E = \frac{1}{2}\,m\,v^2 - \frac{e^2}{4\,\pi\,\epsilon_0\,r} = -\frac{1}{2}\,m\,v^2 \tag{1.5}$$

Hence

$$r = \frac{e^2}{8\,\pi\,\epsilon_0\,|E|} \Leftrightarrow E = -\frac{e^2}{4\,\pi\,\epsilon_0}\,\frac{1}{2\,r} \tag{1.6}$$

This shows that the radius of the atom is only determined when the electron energy is given. This is of course a general feature of the Keplerian problem, which admits different orbits for different energies. But one knew from chemistry that atoms of a certain kind all have the same dimension. This implies, within the planetary model, that, for example, the electrons of all hydrogen atoms all rotate around orbits of the same energy: why should this be so? What should such "preferred" energy be? This is the problem of the stability of atomic dimensions.

**Problem 1.** Show that there is no length one can make starting from $e$ and $m$.

The second problem posed by Rutherford's model is that, according to the law of classical electromagnetism, atoms should collapse by emitting radiation. Indeed, an accelerated charge emits radiation and therefore loses energy according to the law

$$\begin{aligned}\frac{d\,E}{d\,t} &= -\frac{2}{3}\,\frac{e^2}{4\,\pi\,\epsilon_0\,c^3}\,|\vec{a}|^2 = -\frac{2}{3}\,\frac{e^6}{(4\,\pi\,\epsilon_0\,c)^3\,m^2}\,\frac{1}{r^4} \\ &= \frac{d\,E}{d\,r}\,\frac{d\,r}{d\,t} = \frac{e^2}{4\,\pi\,\epsilon_0}\,\frac{1}{2\,r^2}\,\frac{d\,r}{d\,t}\end{aligned} \tag{1.7}$$

---

[2] $m \equiv \frac{m_e\,m_p}{m_e + m_p}$.

and hence

$$\frac{d\,r}{d\,t} = -\frac{4}{3}\,\frac{e^4}{(4\,\pi\,\epsilon_0)^2\,c^3\,m^2}\,\frac{1}{r^2} \tag{1.8}$$

Integrating

$$\frac{(4\,\pi\,\epsilon_0)^2\,c^3\,m^2}{e^4}\,(r^3 - r_0^3) = -4\,t \tag{1.9}$$

where $r_0$ is the radius of the orbit at a certain initial time $t = 0$. Assuming $r_0 \approx 10^{-8}\,cm$, one concludes that the time it takes for an electron to fall onto the nucleus is

$$\tau = \frac{(4\,\pi\,\epsilon_0)^2\,c^3\,m^2}{4\,e^4}\,r_0^3 \tag{1.10}$$

Numerically this turns out to be

$$\tau \approx 1.\times 10^{-10}\,sec \tag{1.11}$$

i.e. an extremely short time, in contrast to the apparent longevity of atoms.

## 1.2 Classical phase space

Consider the classical equations of motion of a particle of mass $m$ moving in 3-dimensional space under the influence of a space dependent force $\vec{F}(\vec{x})$:

$$m\,\frac{d^2\vec{x}(t)}{d\,t} = \vec{F}(\vec{x}) \tag{1.12}$$

where $\vec{x} = (x_1, x_2, x_2)$ are the cartesian coordinates of the position of the particle. This force is said to be *conservative* if

$$\vec{F}(\vec{x}) = -\vec{\nabla}\,V(\vec{x}) \tag{1.13}$$

where $V(\vec{x})$ is a function of the coordinates $\vec{x}$, and $\vec{\nabla}\,V(\vec{x})$ is its *gradient*

$$\vec{\nabla}\,V(\vec{x}) = \big(\frac{\partial\,V(\vec{x})}{\partial x_1}, \frac{\partial\,V(\vec{x})}{\partial x_2}, \frac{\partial\,V(\vec{x})}{\partial x_{3})}\big) \tag{1.14}$$

In this case the function

$$E = T + V = \frac{1}{2}\,m\,\big(\frac{d\vec{x}(t)}{d\,t}\big)^2 + V(\vec{x}) \tag{1.15}$$

is a *constant of motion*: this means that $E$ when evaluated on a solution of the equation of motions (1.12) is independent of time. Indeed

$$\begin{aligned}\frac{dE}{dt} &= m\,\frac{d\vec{x}(t)}{d\,t}\cdot\frac{d^2\vec{x}(t)}{d\,t}+\vec{\nabla}\,V(\vec{x})\cdot\frac{d\vec{x}(t)}{d\,t}= \\ &= \frac{d\vec{x}(t)}{d\,t}\cdot\vec{F}(\vec{x})+\vec{\nabla}\,V(\vec{x})\cdot\frac{d\vec{x}(t)}{d\,t}=0 \end{aligned} \tag{1.16}$$

The classical equations of motion are second order differential equations for the trajectory $\vec{x}(t)$ of the particle. A trajectory is therefore completely specified by giving the *initial conditions* of the particle, i.e. its position $\vec{x}_0$ and its *velocity* $\vec{v}_0$ at a given time $t=t_0$

$$\vec{x}_0=\vec{x}(0) \qquad \vec{v}_0=\dot{\vec{x}}(0) \tag{1.17}$$

The space of all solutions of the classical equations of motion is therefore the space of all the initial conditions $(\vec{x_0},\vec{v}_0)$. This space is called the classical *phase space* $\mathcal{M}_{ph.s.}$ of the system.

It is important to appreciate that the pairs $(\vec{x_0},\vec{v}_0)$, i.e. positions and velocities, are not the most convenient coordinates on $\mathcal{M}_{ph.s.}$. It is more useful to consider the pairs $(\vec{x_0},\vec{p}_0)$, with $\vec{p}_0=\vec{p}(0)$, where

$$\vec{p}=(p_1,p_2,p_3) \tag{1.18}$$

are the (linear) *momenta* relative to the cartesian coordinates $\vec{x}=(x_1,x_2,x_3)$

$$p_i(t)=m\,\dot{x}_i(t) \qquad i=1,2,3 \tag{1.19}$$

The coordinates $(x_i,p_i)$ are one example of *canonically conjugate coordinates* on $\mathcal{M}_{ph.s.}$. If one re-writes the energy (1.15) in terms of the linear momenta one obtains a function of the canonical coordinates $(x_i,p_i)$

$$H(x_i,p_i)=\sum_{i=1}^{3}\frac{p_i^2}{2\,m}+V(x_1,x_2,x_3) \tag{1.20}$$

which is known as the *Hamiltonian* (function) of the system. Let us observe that the equations of motion (1.12) can be written in terms of the canonical pairs $(x_i,p_i)$ and the Hamiltonian function in the following way

$$\dot{p}_i=-\frac{\partial H(x,p)}{\partial x_i} \qquad \dot{x}_i=\frac{\partial H(x,p)}{\partial p_i} \qquad i=1,2,3 \tag{1.21}$$

This is the so-called Hamiltonian form of the classical equations of motion. The Hamiltonian equations are a system of differential equations of the *first order* for the pairs $(x_i, p_i)$. They are completely equivalent to the Newton second order equations (1.12) for the space coordinates $x_i$.

In many problems it is convenient to use coordinates different than cartesian ones. In these applications the Hamiltonian form (1.21) the equations of motion becomes quite useful.

To appreciate this point let us consider for example spherical coordinates to describe the motion in space of a particle of mass $m$:

$$x_1 = r\,\sin\theta\cos\phi \qquad x_2 = r\,\sin\theta\sin\phi \qquad x_3 = r\cos\theta \tag{1.22}$$

To obtain the equation of motions of the system in the spherical coordinates one could plug the relations (1.22) into the second order differential equations (1.12). This can be algebraically quite involved even in a simple case like this.

**Problem 2.** Derive the second order differential equations for $(r, \theta, \phi)$ which are obtained by plugging in (1.22) into the Newton equations (1.12) for a particle of mass $m$ moving in a potential $V(r, \theta, \phi)$.

We see that form of the second order differential equations depends on the coordinates chosen. For example the equations of motion for a particle in a potential do not depend on the first derivatives of the coordinates when written in cartesian coordinates

$$m\,\ddot{x}^i = -\frac{\partial V}{\partial x^i} \tag{1.23}$$

When written in spherical coordinates the equations of motions contains instead first order time derivatives of the spherical coordinates.

One nice thing about the Hamiltonian formulation is that the equations of motion in any coordinate system — for example in spherical coordinates (1.22) — can still be written exactly in the same form as in (1.21)

$$\begin{aligned} \dot{p}_r &= -\frac{\partial H}{\partial r} \qquad \dot{p}_\theta = -\frac{\partial H}{\partial \theta} \qquad \dot{p}_\phi = -\frac{\partial H}{\partial \phi} \\ \dot{r} &= \frac{\partial H}{\partial p_r} \qquad \dot{\theta} = \frac{\partial H}{\partial p_\theta} \qquad \dot{\phi} = \frac{\partial H}{\partial p_\phi} \end{aligned} \tag{1.24}$$

*once we find suitable momenta variables* $p_r, p_\theta$ and $p_\phi$, functions of the coordinates and their first time derivatives. $p_r, p_\theta$ and $p_\phi$ are said to be *canonically conjugate*, respectively, to $r, \theta$ and $\phi$.

To determine the momenta canonically conjugate to the spherical coordinates, let us consider the distance squared $\Delta\vec{x}^2$ between two points with spherical coordinates $(r,\theta,\phi)$ and $(r+\Delta r,\theta+\Delta\theta,\phi+\Delta\,\phi)$ with $(\Delta r,\Delta\theta,\Delta\phi)$ infinitesimally small :

$$\Delta\vec{x}^2 = \Delta r^2 + r^2\,\Delta\theta^2 + r^2\,\sin^2\phi\,\Delta\,\phi^2 \tag{1.25}$$

The energy (1.15) writes then

$$\begin{aligned} E &= \frac{1}{2}\,m\left(\frac{d\,\vec{x}}{dt}\right)^2 + V(x) = \\ &= \frac{1}{2}\,m\left[\dot{r}^2 + r^2\,\dot{\theta}^2 + r^2\,\sin^2\phi\,\dot{\phi}^2\right] + V(r,\theta,\phi) \end{aligned} \tag{1.26}$$

where $V(r,\theta,\phi)$ is the potential expressed in terms of spherical coordinates. Therefore if we define

$$p_r = m\,\dot{r} \qquad p_\theta = m\,r^2\,\dot{\theta} \qquad p_\phi = m\,r^2\,\sin^2\theta\,\dot{\phi} \tag{1.27}$$

the Hamiltonian function becomes

$$H = \frac{p_r^2}{2\,m} + \frac{p_\theta^2}{2\,m\,r^2} + \frac{p_\phi^2}{2\,m\,r^2\,\sin^2\theta} + V(r,\theta,\phi) \tag{1.28}$$

and the last three of the equations (1.24) are automatically satisfied

$$\dot{r} = \frac{\partial H}{\partial p_r} \qquad \dot{\theta} = \frac{\partial H}{\partial p_\theta} \qquad \dot{\phi} = \frac{\partial H}{\partial p_\phi} \tag{1.29}$$

The other three equations

$$\begin{aligned} \dot{p}_r &= -\frac{\partial H}{\partial r} = \frac{p_\theta^2}{m\,r^3} + \frac{p_\phi^2}{m\,r^3\,\sin^2\theta} - \frac{\partial V}{\partial r} \\ \dot{p}_\theta &= -\frac{\partial H}{\partial \theta} = \frac{\cos\theta\,p_\phi^2}{m\,r^2\,\sin^3\theta} - \frac{\partial V}{\partial \theta} \\ \dot{p}_\phi &= -\frac{\partial H}{\partial \phi} = -\frac{\partial V}{\partial \phi} \end{aligned} \tag{1.30}$$

are now precisely the equations of motion that one obtains by rewriting Newton equations (1.12) in spherical coordinates.

Summarizing, given any system of coordinates $Q^i(x_i)$, it is possible to find momenta $P^i$ canonically conjugate $Q^i$ such that equations of motions write in the universal form (1.21) [3]

$$\dot{P}^i = -\frac{\partial H(Q,P)}{\partial Q^i} \qquad \dot{Q}^i = \frac{\partial H(Q,P)}{\partial P^i} \qquad i = 1,2,3 \tag{1.31}$$

Once we describe the equations of motion in terms of canonical coordinates $(Q^i, P^i)$, one can even consider transformations which mix spatial coordinates and momenta

$$Q \to \tilde{Q} = \tilde{Q}(Q,P) \qquad P \to \tilde{P} = \tilde{P}(Q,P) \tag{1.32}$$

as long as the equations of motion continue to write in the canonical form (1.31). These wider set of coordinates and momenta transformations are called *canonical transformations*.

In conclusion, the classical phase space is the space of all solutions of the classical equations of motions. Canonical coordinates are convenient kinds of coordinates on phase space for which the equations of motion write in the form (1.31).

One property of canonical coordinates which is important for the application which will be discussed in the next section is the following. Consider a point $P = (x_i^{(0)}, p_i^{(0)})$ in phase space and a little hyper-cube $\Delta V$ based on it associated to small variations $(\Delta x_i, \Delta p_i)$ of the cartesian canonical coordinates:

$$\Delta V = \{(x,p) \big|\, x_i^{(0)} \leq x_i \leq x_i^{(0)} + \Delta x_i,\ p_i^{(0)} \leq p_i \leq p_i^{(0)} + \Delta p_i\} \tag{1.33}$$

The volume of $\Delta V$ is

$$\mathrm{Vol}(\Delta V) = \Delta x_1\, \Delta x_2\, \Delta x_3\, \Delta p_1\, \Delta p_2\, \Delta p_3 = \mathrm{Vol}(\Delta V_x)\, \mathrm{Vol}(\Delta V_p) \tag{1.34}$$

Let us now consider the *same* region $\Delta V$ as described in a different system of canonical coordinates. To be specific let us consider the spherical canonical

---

[3] There exists a systematic way to find the canonical momenta $P^i$ corresponding to generic coordinates $Q^i(x)$, which involves considering another function of the coordinates and of its first derivatives, which is called the *Lagrangian* function of the system. The Lagrangian function is related to the Hamiltonian by an operation known as Legendre transform. This systematic way is usually discussed in advanced courses on classical mechanics. We will limit ourselves to simple cases for which it will be usually easy to find canonical momenta $P^i$ by requiring that the last three of the equations (1.31) hold.

coordinates we discussed above. Then the volume of the spatial component $\Delta V_x$ writes

$$\Delta x_1\,\Delta x_2\,\Delta x_3 = r^2\,\sin\theta\,\Delta r\,\Delta\theta\,\Delta\phi \tag{1.35}$$

where $(r,\theta,\phi)$ are the spatial coordinates of $P$. The factor $r^2\sin\theta$ is called the Jacobian of the transformations (1.22):

$$r^2\,\sin\theta = \left|\frac{\partial(x_1,x_2,x_3)}{\partial(r,\theta,\phi)}\right| = \det\begin{pmatrix} \frac{\partial x_1}{\partial r} & \frac{\partial x_1}{\partial\theta} & \frac{\partial x_1}{\partial\phi} \\ \frac{\partial x_2}{\partial r} & \frac{\partial x_2}{\partial r} & \frac{\partial x_2}{\partial\phi} \\ \frac{\partial x_3}{\partial r} & \frac{\partial x_3}{\partial\theta} & \frac{\partial x_3}{\partial\phi} \end{pmatrix} \tag{1.36}$$

The interesting fact, which we will prove below, is that the volume of the momenta component $\Delta V_p$ is

$$\Delta p_1\,\Delta p_2\,\Delta p_3 = \frac{1}{r^2\,\sin\theta}\,\Delta p_r\,\Delta p_\theta\,\Delta p_\phi \tag{1.37}$$

Therefore

$$\mathrm{Vol}(\Delta V) = \Delta x_1\,\Delta x_2\,\Delta x_3\,\Delta p_1\,\Delta p_2\,\Delta p_3 = \Delta r\,\Delta\theta\,\Delta\phi\,\Delta p_r\,\Delta p_\theta\,\Delta p_\phi \tag{1.38}$$

We will show that this is true for any canonical system of coordinates

$$\Delta x_1\,\Delta x_2\,\Delta x_3\,\Delta p_1\,\Delta p_2\,\Delta p_3 = \Delta Q_1\,\Delta Q_2\,\Delta Q_3\,\Delta P_1\,\Delta P_2\,\Delta P_3 \tag{1.39}$$

Let us discuss an important implication of property (1.39). Consider a function on the phase space $f(x,p)$, written in cartesian canonical coordinates. Let us suppose $f$ is such that the integral over phase space

$$\bar f = \int_{\mathcal{M}_{ph.s.}} \prod_i dx^i dp^i\, f(x,p) \tag{1.40}$$

converges. The property (1.39) ensures that $\bar f$ is a number which is independent of the system of canonical coordinates chosen:

$$\bar f = \int_{\mathcal{M}_{ph.s.}} \prod_i dx^i dp^i\, f(x,p) = \int_{\mathcal{M}_{ph.s.}} \prod_i dQ^i dP^i\, \tilde f(Q,P) \tag{1.41}$$

In this formula $\tilde f(Q,P) = f(x(Q,P),p(Q,P))$ is the function on moduli space written in the $(Q^i,P^i)$ system of canonical coordinates.

In conclusion there is a *natural* way — by which we mean a way which is independent of choices of spatial coordinates — to compute the volume of a region in phase space. As we explained, points in phase spaces represents solutions of the classical equations of motion. Hence there is a natural way to "count" classical solutions: the "number" of classical solutions in a given region of phase space can be measured by its volume evaluated by means of a system of canonical coordinates[4]. We just showed the this volume is independent of which particular system we use.

As an application of this concept, let us consider a 1-dimensional particle of mass $m$ moving in the potential $V(x)$:

$$H = \frac{p^2}{2\,m} + V(x) \tag{1.42}$$

One might ask "how many" classical solutions there are with energy less than a given value $E$. We argued above that a measure of those solutions is the area of the region in phase space corresponding to the points $(x, p)$ for which $H(x, p) \leq E$. The boundary of this region is described in the $(x, p)$ plane by the equation

$$\frac{p^2}{2\,m} + V(x) = E \tag{1.43}$$

and therefore is given by the curves

$$p(x) = \pm\sqrt{2\,m(E - V(x))} \tag{1.44}$$

Let us suppose for simplicity that, given the energy $E$, there are only two points $x_1(E)$ and $x_2(E)$ which are the solutions of

$$V(x_i) = E \qquad i = 1, 2 \tag{1.45}$$

and that

$$V(x) \leq E \qquad \text{for } x_1(E) \leq x \leq x_2(E) \tag{1.46}$$

$x_i(E)$, with $i = 1, 2$ are the end points of the oscillatory motion of energy $E$. The area delimited by the two curves (1.44) is

$$N(E) = 2\int_{x_1(E)}^{x_2(E)} dx\,|p(x)| = \oint dx\,p(x) \tag{1.47}$$

---

[4] Of course this "number" is not integer, but a real number which has the dimensions of *Energy* × *Time* = *Mass* × *Length*$^2$/*Time*.

where $\oint$ denotes the integration of $p$ along the closed curve in the $(x, p)$ plane corresponding to the oscillatory motion. The function $N(E)$ measures the "number" of classical solutions with energy less than $E$. The solutions whose energy is between $E$ and $E + dE$ are therefore given by

$$dN(E) = \frac{dN(E)}{dE}\, dE \equiv \rho(E)\, dE \tag{1.48}$$

where $\rho(E)$ is the *density* of classical solutions as function of the energy. From (1.47) we obtain

$$\rho(E) = 2 \int_{x_1(E)}^{x_2(E)} \sqrt{\frac{m}{2}}\, \frac{dx}{\sqrt{E - V(x)}} \tag{1.49}$$

But since the equations of motion of the 1-dimensional particle

$$\frac{dx(t)}{dt} = \pm \sqrt{\frac{2\,(E - V(x))}{m}} \tag{1.50}$$

can be integrated to give

$$t - t_0 = \pm \int_{x_0}^{x} dx' \sqrt{\frac{m}{2\,(E - V(x))}} \tag{1.51}$$

We conclude that the integral in the right hand side of (1.49) is precisely the *period* $\tau(E)$ of the oscillatory motion of energy $E$

$$\rho(E) = 2 \int_{x_1(E)}^{x_2(E)} \sqrt{\frac{m}{2}}\, \frac{dx}{\sqrt{E - V(x)}} = \tau(E) \tag{1.52}$$

**Problem 3.** Use formula (1.52) to show that the density $\rho(E)$ of solutions for the harmonic oscillator $V(x) = \frac{1}{2}\, m\, \omega^2\, x^2$ is constant: $\rho(E) = \frac{2\,\pi}{\omega}$.

Let us prove (1.39). Let

$$x^i = x^i(Q) \tag{1.53}$$

be the relation between the cartesian coordinates and a generic system of coordinates $Q^i$ (spherical coordinates, for example). Then if

$$M_{ij} = \frac{\partial x^i}{\partial Q^j} \tag{1.54}$$

is the matrix of the partial derivatives, we have

$$\Delta x^i = \sum_j M_{ij}\, \Delta Q^j \tag{1.55}$$

Moreover

$$p^i = m\, \dot{x}^i = m \sum_j M_{ij}\, \dot{Q}^j \tag{1.56}$$

The distance squared between two points infinitesimally close in space is

$$\Delta s^2 = \sum_i \Delta x^i\, \Delta x^i = \sum_{i,k,l} M_{ik}\, M_{il}\, \Delta Q^k\, \Delta Q^l \tag{1.57}$$

Therefore the kinetic energy writes

$$T = \frac{1}{2}\, m \sum_{i,k,l} M_{ik}\, M_{il}\, \dot{Q}^k\, \dot{Q}^l \tag{1.58}$$

From this we deduce that the momenta $P^i$ canonically conjugate to $Q^i$ are

$$P^i = m \sum_{j,k} M_{ji}\, M_{jl}\, \dot{Q}^l \tag{1.59}$$

Hence

$$\sum_i M^{-1}_{ij}\, P^i = m \sum_l M_{jl}\, \dot{Q}^l \tag{1.60}$$

Substituting this into (1.56) we obtain

$$p^i = \sum_j M^{-1}_{ji}\, P^j \tag{1.61}$$

In conclusion

$$\frac{\partial p^i}{\partial P^j} = M^{-1}_{ji} \tag{1.62}$$

Hence

$$\begin{aligned} \Delta x_1\, \Delta x_2\, \Delta x_3 &= \det\big(\frac{\partial x^i}{\partial Q^j}\big)\, \Delta Q_1\, \Delta Q_2\, \Delta Q_3 = \\ &= \det(M)\, \Delta Q_1\, \Delta Q_2\, \Delta Q_3 \\ \Delta p_1\, \Delta p_2\, \Delta p_3 &= \det\big(\frac{\partial p^i}{\partial P^j}\big)\, \Delta P_1\, \Delta P_2\, \Delta P_3 = \\ &= \det(M^{-1})\, \Delta P_1\, \Delta P_2\, \Delta P_3 \end{aligned} \tag{1.63}$$

from which (1.39) follows.

Let us conclude this section by discussing some applications of spherical coordinates.

### 1.2.1 Central potentials

To start with, consider a potential $V(r)$ which does not depend on the angles $\theta$ and $\phi$. These kind of potential are called *central*. The equations of motions (1.30) for central potential become

$$\begin{aligned} \dot p_r &= \frac{p_\theta^2}{m\,r^3} + \frac{p_\phi^2}{m\,r^3\,\sin^2\theta} - \frac{\partial V}{\partial r} \\ \dot p_\theta &= \frac{\cos\theta\,p_\phi^2}{m\,r^2\,\sin^3\theta} \\ \dot p_\phi &= 0 \end{aligned} \tag{1.64}$$

The last equation says that the momenta canonically conjugate to $\phi$ is conserved, and it is a constant independent of time. $p_\phi$ is the component along $z$ of the angular momentum

$$p_\phi = m\,r^2\,\sin^2\theta\,\dot\phi = L_z \tag{1.65}$$

The second equation implies that

$$\begin{aligned} \frac{d}{dt}\big(p_\theta^2 + \frac{p_\phi^2}{\sin^2\theta}\big) &= 2\,\big(p_\theta\,\dot p_\theta - \frac{p_\phi^2}{\sin^3\theta}\,\cos\theta\,\dot\theta\big) = \\ &= 2\,\big(mr^2\,\dot\theta\,\frac{\cos\theta\,p_\phi^2}{m\,r^2\,\sin^3\theta} - \frac{p_\phi^2}{\sin^3\theta}\,\cos\theta\,\dot\theta\big) = 0 \end{aligned} \tag{1.66}$$

Therefore $p_\theta^2 + \frac{p_\phi^2}{\sin^2\theta}$ is also conserved: it is the square of the angular momentum

$$\vec L^2 = p_\theta^2 + \frac{p_\phi^2}{\sin^2\theta} \tag{1.67}$$

The radial equation of motion reduces to

$$\dot p_r = m\,\ddot r = \frac{\vec L^2}{m\,r^3} - \frac{\partial V}{\partial r} \tag{1.68}$$

This is equivalent to a 1-dimensional motion on the semi-axis $r > 0$ in the potential

$$V_{eff}(r) = \frac{\vec{L}^2}{2\,m\,r^2} + V(r) \tag{1.69}$$

The term $\frac{\vec{L}^2}{2\,m\,r^2}$ is called the *centrifugal* potential.

### 1.2.2 2-dimensional rotor

Spherical coordinates are also useful to describe the motion of a particle which is constrained to move on a surface of fixed radius $r = a$. In this case $r$ is not a dynamical variable, since $\dot{r} = 0 = p_r$. The Hamiltonian function (1.28) reduces therefore to

$$H = \frac{p_\theta^2}{2\,I} + \frac{p_\phi^2}{2\,I\,\sin^2\theta} + V(\theta,\phi) \tag{1.70}$$

where

$$I = m\,a^2 \tag{1.71}$$

is called the momentum of inertia of the system. The phase space of this system is 4-dimensional and $(\theta, \phi, p_\theta, p_\phi)$ are canonical coordinates for it.

For $V(\phi, \theta) = 0$ we obtain the Hamiltonian for a 2-dimensional *rotor*

$$H = \frac{p_\theta^2}{2\,I} + \frac{p_\phi^2}{2\,I\,\sin^2\theta} = \frac{L^2}{2\,I} \tag{1.72}$$

whose equations of motions are

$$\begin{aligned}\dot{\theta}^2 &= \frac{L^2}{I^2} - \frac{p_\phi^2}{I^2\,\sin^2\theta}\\ \dot{\phi} &= \frac{p_\phi}{I\,\sin^2\theta}\end{aligned} \tag{1.73}$$

By taking the square root of the first equation and dividing by the second we obtain the equations for the orbits

$$\frac{d\theta}{d\phi} = \pm\sin\theta\sqrt{\frac{L^2}{p_\phi^2}\,\sin^2\theta - 1} \tag{1.74}$$

from which

$$\phi + \phi_0 = \pm \arctan \frac{\cos\theta}{\sqrt{\frac{L^2}{p_\phi^2}\,\sin^2\theta - 1}} \tag{1.75}$$

where $\phi_0$ is an integration constant. This equation can be written as

$$(\frac{L^2}{p_\phi^2} - 1)\,\sin^2(\phi + \phi_0) = \cot^2\theta \tag{1.76}$$

and hence

$$p_\phi\,\cos\theta \pm \sqrt{L^2 - p_\phi^2}\,\sin(\phi + \phi_0) = 0 \tag{1.77}$$

This is the equation for a maximal circle on the sphere, sitting in the plane passing for the origin and orthogonal to the vector

$$\vec{L} = (\sqrt{L^2 - p_\phi^2}\,\cos\phi_0, \sqrt{L^2 - p_\phi^2}\,\sin\phi_0, p_\phi) \tag{1.78}$$

This vector has to be interpreted therefore as the conserved angular momentum.

### 1.2.3 Particle on a circle

Consider now a particle constrained to move on a circle of radius $a$ on a fixed plane. The corresponding Hamiltonian is obtained from (1.28) by imposing not only $r = a$, $p_r = m\dot{r} = 0$ but

$$\theta = \frac{\pi}{2} \qquad p_\theta = m\,a^2\dot{\theta} = 0 \tag{1.79}$$

as well. Thus

$$H = \frac{p_\phi^2}{2\,I} + V(\phi) \qquad I \equiv m\,a^2 \tag{1.80}$$

from which one deduce the equation of motion

$$\dot{\phi}^2 = \frac{2\,(E - V(\phi))}{I} \tag{1.81}$$

In the free case $V(\phi) = 0$, $p_\phi$ is a constant of motion

$$E = \frac{p_\phi^2}{2\,I} \tag{1.82}$$

and

$$\dot{\phi} = \frac{p_\phi}{I} \tag{1.83}$$

## 1.3 Specific heat capacities

Another area of classical physics where a difficulty emerged, both at the experimental and conceptual level, was statistical mechanics: specifically, its predictions regarding the specific heat capacities.

The starting point of statistical mechanics is Boltzman probability distribution, whose content is the following. Consider a thermodynamical system made of very large number $N$ of mechanical subsystems (e.g. molecules of a gas). Assume that the subsystems can exchange, through collisions, momentum and energy among them. The Boltzman law states that, when the system is at the thermodynamical equilibrium at temperature $T$, the probability $dP(E)$ that a subsystem be at energies between $E$ and $E+dE$ is proportional to

$$dP(E) \propto dE\, \rho(E)\, \mathrm{e}^{-\beta E} \tag{1.84}$$

where

$$\beta \equiv \frac{1}{k\,T} \tag{1.85}$$

The number

$$k = 1.38 \times 10^{-16}\, erg/K = \frac{1}{12000} eV/K \tag{1.86}$$

is the Boltzmann constant and $\rho(E)$ is the density of subsystem configurations (i.e. of classical solutions) with energy $E$.[5]

The probability distribution in (1.84) must be normalized to one

$$\int_0^\infty \frac{dP(E)}{dE}\, dE = 1 \tag{1.87}$$

Therefore, if one introduces the classical *partition function* of the system

$$Z(\beta) \equiv \int_0^\infty dE\, \rho(E)\, \mathrm{e}^{-\beta E} \tag{1.88}$$

---

[5] In Eq. (1.86) we introduced the *electronvolt*, a convenient unit of energy for atomic phenomena:

$$1\, eV = 1.6 \times 10^{-19}\, C \cdot 1\, Volt = 1.6 \times 10^{-12}\, erg = 1.6 \times 10^{-19}\, J$$

This is the energy that an electron, whose charge is $e \approx -1.6 \times 10^{-19}\, C$, acquires when accelerated by a difference of potential equal to 1 Volt.

one can rewrite Eq. (1.84) as follows

$$\frac{dP(E)}{dE} = \frac{\rho(E)\,\mathrm{e}^{-\beta\,E}}{Z(\beta)} \tag{1.89}$$

The average energy $\bar{E}$ of the subsystems is therefore

$$\begin{aligned}\bar{E} = \int_0^\infty dE\,\frac{dP(E)}{dE}\,E = \int_0^\infty dE\,E\,\frac{\rho(E)\,\mathrm{e}^{-\beta\,E}}{Z(\beta)} = \\ = -\frac{\partial_\beta\,Z(\beta)}{Z(\beta)} = -\partial_\beta\,\log Z(\beta)\end{aligned} \tag{1.90}$$

Let $q_1, \ldots q_f$ and $p_1 \ldots p_f$, be canonical coordinates and momenta of as subsystem, and $H(q,p)$ its Hamiltonian function. We will call $f$ the number of degrees of freedom of the (sub)system. We explained in the previous section that a measure of the “number” of classical solutions in a small region around a point $(q,p)$ of phase space is its canonical volume

$$dq_1 \ldots dq_f\,dp_1 \ldots dp_f \equiv d^f q\,d^f p \tag{1.91}$$

Therefore the partition function (1.88) can also be written as an integral over phase space

$$Z(\beta) = \int d^f q\,d^f p\,\mathrm{e}^{-\beta\,H(q,p)} \tag{1.92}$$

Let us now suppose that the Hamiltonian function $H(q,p)$ of the subsystem be a positive definite quadratic form of the $q_i$’s and $p_i$’s. Then $H(q,p)$ can be put, by a suitable canonical transformation, in the canonical (“diagonal”) form

$$H(q,p) = \sum_{i=1}^f (a_i\,q_i^2 + b_i\,p_i^2) \tag{1.93}$$

with $a_i, b_i \geq 0$. Recalling the formula for the gaussian integral

$$I(\alpha) = \int_{-\infty}^\infty dx\,\mathrm{e}^{-\alpha\,x^2} = \sqrt{\frac{\pi}{\alpha}} \qquad \alpha > 0 \tag{1.94}$$

we obtain for the partition function of the quadratic system (1.93)

$$Z(\beta) = \prod_{i|a_i\neq 0} \sqrt{\frac{\pi}{\beta\,a_i}}\,\prod_{i=1}^f \sqrt{\frac{\pi}{\beta\,b_i}}\,\prod_{i|a_i=0} L_i \tag{1.95}$$

where the $L_i$'s are the sizes of those canonical coordinates $q_i$ which do not appear in the Hamiltonian. For example, the cartesian coordinates $x_i$ do not appear in the Hamiltonian of a translational invariant system: the associated $L_i$ are in this case the sizes of the box containing the system. We can assume that the $b_i$'s are all non-vanishing, since otherwise the corresponding canonical coordinate $q_i$ would be "frozen" to a constant value and thus it would be non-dynamical. In conclusion

$$Z(\beta) = \beta^{-\frac{g}{2}}\, V\, K \tag{1.96}$$

where $K = \prod_{i:a_i\neq 0} \sqrt{\frac{\pi}{a_i}}\, \prod_{i=1}^{f} \sqrt{\frac{\pi}{b_i}}$ is a constant independent of the temperature, $V = \prod_{i:a_i=0} L_i$ is the "volume" of the system, and $g$ is the number of non-vanishing $b_i$ and $a_i$, i.e. $f \leq g \leq 2\, f$.

From the partition function one can compute the internal energy of the thermodynamical system

$$U = N\, \frac{\int d^f q\, d^f p\, H(q,p)\, \mathrm{e}^{-\beta\, H(q,p)}}{\int d^f q\, d^f p\, \mathrm{e}^{-\beta\, H(q,p)}} = -N\, \partial_\beta\, \log Z(\beta) \tag{1.97}$$

The molar heat capacity at constant volume is defined to be

$$C_V = \frac{\partial U}{\partial T}\Big|_V = -\frac{\beta}{T}\, \frac{\partial U}{\partial \beta}\Big|_V = k\, N_A\, \beta^2 \partial_\beta^2\, \log Z(\beta) \tag{1.98}$$

where in $U$ we took $N = N_A = 6.03 \times 10^{23}$ is the Avogadro number. For the quadratic system (1.93) we obtain therefore

$$U = \frac{g}{2}\, N\, k\, T \qquad C_V = k\, N_A\, \frac{g}{2} \equiv R\, \frac{g}{2} \tag{1.99}$$

where $R = 1.93\, cal/(K \times mol)$ is the ideal gas constant. This result goes under the name of "equipartion theorem" and it is sometimes stated by saying that each quadratic degree of freedom contributes $\frac{1}{2}\, kT$ to the internal energy.

Let us consider some examples. For a monoatomic "perfect" gas, one takes the free hamiltonian

$$H(p,q) = \frac{\vec{p}^2}{2\, m} \tag{1.100}$$

where $\vec{p}$ is the momentum of the molecule. Therefore $g = 3$ for a perfect monoatomic gas, and the classical prediction for its molar heat capacity is $C_V = \frac{3}{2}\, R$.

A diatomic gas can be schematized as a rigid rod to whose ends the two atoms forming the molecule are attached

$$H(p,q) = \frac{\vec{p}^2}{2\,m} + \frac{1}{2\,I}\Big(p_\theta^2 + \frac{p_\phi^2}{\sin^2\theta}\Big) \tag{1.101}$$

where $I$ is the momentum of inertia of the rod. This Hamiltonian is not, strictly speaking, of the form we considered since the coefficient of $p_\phi^2$ is not a constant but it depends on $\theta$.

**Problem 4.** Generalize the derivation of the equipartion theorem (1.99) to the Hamiltonian of the bi-atomic molecule (1.101).

The classical prediction for the bi-atomic molecule is thus $C_V = \frac{5}{2}\,R$.

A solid can be schematized as a set of $N$ three-dimensional oscillators, whose Hamiltonian is

$$H(p,q) = \frac{\vec{p}^2}{2\,m} + \frac{1}{2}\sum_{i=1}^{3} k_i\, q_i^2 \tag{1.102}$$

The classical prediction for the solid is hence $C_V = 3\,R$

The experimental situation is the following:

- The measured $C_V$'s of monoatomic gases ($He$, $Ne$, $Ar$,...) agree, for almost all temperature and when the perfect gas conditions are met, with the theoretical prediction $\frac{3}{2}\,R$.
- The measured $C_V$ for bi-atomic gases ($H_2$, $N_2$,...), under perfect gas conditions, is in good agreement with the theoretical prediction $\frac{5}{2}R$ only for temperatures sufficiently high (in the range $10 - 500K$). The observed value of $C_V$ decreases as temperature lowers.
- For metals only in few cases there is agreement at room temperature between observed $C_V$ and its theoretical value $C_V = 3\,R$. The experimental value is generally lower than the classical value, and $C_V \to 0$ when $T \to 0$.

The experimental situation can be described by saying that the effective number of (quadratic) degrees of freedom decreases (continuously!) as the temperature decreases. As a matter of fact, the very concept of "degree of freedom" has not an absolute *physical* meaning. How many degrees of freedoms a given physical system has, depends on the schematization, and therefore on the approximation, we use to describe it. Hence, in a sense, it is even surprising that there is, for certain range of temperatures, an agreement between experimental data and equipartion theorem. In short, the equipartion theorem, which is a direct result of classical kinetic theory, raises both conceptual and experimental difficulties.

The equipartition theorem holds for quadratic Hamiltonians. One might hope that by considering more general, non-quadratic, systems one could account for the discrepancy between the predictions of equipartition theorem and the experimental behaviors of specific heats at low temperatures. In the last part of this section we want to explain in detail why this hope is a vain one.

The physical reason for why non-quadratic Hamiltonians cannot be the solution out of the difficulty that we have been discussing is the following. At sufficiently low temperatures most classical subsystems of the statistical ensemble will find themselves near the minima of the potential. Therefore as the temperature lowers we expect that the quadratic approximation for the Hamiltonian of the subsystem be more and more accurate and the equipartition theorem be valid: for low temperatures the (molar) specific heat for *any* system — quadratic or not — should always tend to a non-vanishing constant, a multiple of $R/2$ determined by the number $g$ of quadratic degrees of freedom near the minimum of the potential. But it is precisely this low temperature prediction which is in conflict with the observation[6]: experimentally, specific heats for a number of systems tend to zero at low temperatures.

Let us verify analytically this physical picture for the case of a simple system, the 1-dimensional *anharmonic* oscillator

$$H = \frac{p^2}{2\,m} + \frac{1}{2} m\,\omega^2\,x^2 + \frac{1}{4}\,g\,x^4 \tag{1.103}$$

The classical partition function for this system is

$$\begin{aligned} Z_{anharm}(\beta) &= \int dx\,dp\,\mathrm{e}^{-\beta\,(\frac{p^2}{2\,m}+\frac{1}{2}m\,\omega^2\,x^2+\frac{1}{4}\,g\,x^4)} = \\ &= \sqrt{\frac{2\,\pi\,m}{\beta}} \int_{-\infty}^{\infty} dx\,\mathrm{e}^{-\beta\,(\frac{1}{2}m\,\omega^2\,x^2+\frac{1}{4}\,g\,x^4)} \end{aligned} \tag{1.104}$$

---

[6] And it is logically incoherent, as we explained above.

Putting

$$y \equiv \frac{\sqrt{\beta\, m}}{\sqrt{2}}\,\omega\, x \tag{1.105}$$

we obtain

$$Z_{anharm}(\beta) = \frac{\sqrt{4\,\pi}}{\beta\,\omega}\int_{-\infty}^{\infty} dy\, \mathrm{e}^{-y^2-\gamma\, y^4} \tag{1.106}$$

where we introduced the adimensional parameter

$$\gamma \equiv \frac{g}{m^2\,\omega^4}\,\frac{1}{\beta^2} \tag{1.107}$$

The integral in (1.106) cannot be expressed in terms of elementary functions. But we can compute an approximation of it valid when $\gamma$ is small, i.e. when

$$\gamma \ll 1 \Rightarrow \beta \gg \frac{\sqrt{g}}{m\,\omega^2} \tag{1.108}$$

which is the same as low temperatures

$$T \ll \frac{m\,\omega^2}{k\,\sqrt{g}} \tag{1.109}$$

In this regime it is legitimate Taylor-expand the anharmonic term

$$\begin{aligned} Z_{anharm}(\beta) &= \frac{\sqrt{4\,\pi}}{\beta\,\omega}\int_{-\infty}^{\infty} dy\, \mathrm{e}^{-y^2}\,\mathrm{e}^{-\gamma\, y^4} = \\ &= \frac{\sqrt{4\,\pi}}{\beta\,\omega}\int_{-\infty}^{\infty} dy\, \mathrm{e}^{-y^2}\big(1-\gamma\, y^4+\cdots\big) \end{aligned} \tag{1.110}$$

Recalling that

$$\int_{-\infty}^{\infty} dy\, \mathrm{e}^{-y^2}\, y^4 = \frac{3\,\sqrt{\pi}}{4} \tag{1.111}$$

we obtain

$$\begin{aligned} Z_{anharm}(\beta) &= \frac{2\,\pi}{\beta\,\omega}\big(1-\frac{3}{4}\,\gamma+\cdots\big) = \\ &= \frac{2\,\pi}{\omega\,\beta}\big(1-\frac{3\,g}{4\,m^2\,\omega^4}\,\frac{1}{\beta^2}+\cdots\big) \end{aligned} \tag{1.112}$$

The internal energy becomes

$$\begin{aligned} U_{anharm} &= -N\partial_\beta \, \log Z_{anharm}(\beta) = -N\,\partial_\beta \, \log \frac{2\,\pi}{\omega\,\beta} + \\ &\quad -N\,\partial_\beta \log\bigl(1 - \frac{3\,g}{4\,m^2\,\omega^4}\,\frac{1}{\beta^2} + \cdots\bigr) = \\ &\quad = N\,\beta - N\frac{3\,g}{2\,m^2\,\omega^4}\,\frac{1}{\beta^3} + O(\frac{1}{\beta^5}) \end{aligned} \tag{1.113}$$

The molar specific heat becomes

$$c_V(T) = R\,\bigl[1 - \frac{9\,g\,k^2\,T^2}{2\,m^2\,\omega^4} + O(T^4)\bigr] \tag{1.114}$$

This confirms that for low temperatures $T \ll \frac{m\,\omega^2}{k\,\sqrt{g}}$ we recover the result of equipartition theorem, up to a small temperature dependent correction, that we determined.

## 1.4 Blackbody radiation

Electromagnetic radiation trapped in a box of size $L$ whose sides are perfectly reflecting surfaces can be decomposed in an infinite number of harmonic oscillators of (angular) frequencies

$$\omega_{\vec{k}} = c\,|\vec{k}| \qquad \omega_{\vec{k}} \equiv 2\,\pi\,\nu_{\vec{k}} \tag{1.115}$$

where $\vec{k}$ are the discrete allowed wave vectors

$$\vec{k} = \frac{\pi}{L}\,\vec{n} \qquad \vec{n} = (n_1, n_2, n_3) \tag{1.116}$$

and $n_i$ are positive integers. Each harmonic oscillator is doubly degenerate, since for each frequency there exist two linearly independent polarizations of the associated electromagnetic wave.

Consider this as the subsystem of a thermodynamical ensemble in thermodynamical equilibrium at temperature $T$. The number of harmonic oscillators with frequency less than a given $\omega$ equals *twice* (because of light polarizations) the number of points of the lattice (1.116) contained in a sphere of radius $|\vec{k}| = \frac{\omega}{c}$. The volume of this sector of the sphere in $\vec{k}$ space is

$$\mathcal{V}_{\vec{k}} = \frac{\pi}{6}\,|\vec{k}|^3 = \frac{\pi\,\omega^3}{6\,c^3} \tag{1.117}$$

The volume of each cell of the lattice is $\frac{\pi^3}{L^3}$. The number of points of the lattice with $|\vec{k}| \leq \frac{\omega}{c}$ is therefore

$$N_\omega = \frac{L^3}{6\,\pi^2\,c^3}\,\omega^3 \tag{1.118}$$

The number of oscillators of frequency less than $\omega$ is twice this number.

Let us apply the equipartition theorem to these oscillators: we saw that each oscillator contribute $kT$ to the internal energy, thus the contribution of oscillators of frequency less than $\omega$ to the internal energy is

$$U(\omega) = \frac{L^3\,k\,T}{3\,\pi^2\,c^3}\,\omega^3 \tag{1.119}$$

and the spectral energy distribution

$$\frac{dU(\omega)}{d\,\omega} = \frac{L^3\,k\,T}{\pi^2\,c^3}\,\omega^2 \tag{1.120}$$

This is the Rayleigh (1900) and Jeans (1905) for the spectral energy distribution of the black body radiation. This result agrees with observations for low frequency $\omega$ but for high frequency it gives a result which is both paradoxical and in disagreement with experiments. From the historical point of view it was the search for a spectral distribution of black body radiation in agreement with experimental results which motivated M. Planck in 1900 to introduce his hypothesis regarding quantization of energy, which was later better specified by Einstein in his attempt to account for the photoelectric effect.

**Problem 5.** Eq. (1.120) gives the classical spectral distribution of the *energy density* of the black body radiation. Find the spectral distribution of the emitted energy flux, that is the spectral distribution of the energy per unit of area and unit of time.

## 1.5 The Photoelectric Effect

The photoelectric effect was first observed by H. Hertz in 1887 and then studied by many experimentalists in the the following years[7]: matter (metals and non-metallic solids, liquid, gases) emits electrons as consequence of the absorption of incident electromagnetic radiation whose frequency lies in the visible or the ultraviolet spectrum. Experimentally, the phenomenon exhibits the following behavior:

---

[7] Among others, J.J. Thomson in 1899, P. Lennard in 1900.

- The effect occurs only when the frequency $\nu$ of the incoming radiation is higher than a threshold frequency $\nu_0$, which depends on the material.
- The number of electrons emitted per unit of time is proportional to the intensity of the incoming radiation.
- The maximal kinetic energy $T_{max}$ of the emitted electrons satisfies the relation

$$T_{max} = h\,(\nu - \nu_0) \qquad h \approx 6.6 \times 10^{-34}\,J \times sec \tag{1.121}$$

It is difficult to explain these features within a model in which atoms are schematized as harmonic oscillators absorbing classical electromagnetic radiation. Let us briefly discuss why.

To start with let us consider the numbers involved. The frequencies of the electromagnetic waves involved in the effect are those of the visible spectrum

$$\nu_{visible} \approx (0.40 - 0.750) \times 10^{15} Hz \tag{1.122}$$

The energies acquired by the photo-electrons are of the order of magnitude of the electron-volt

$$E_{electron} \sim h\,\nu_0 \sim (1 - 10)\, \times 10^{-19}\,J \tag{1.123}$$

Electromagnetic energy fluxes of the order of magnitudes

$$\Phi \sim (10 - 10^2)\,watt/m^2 \tag{1.124}$$

give rise to detectable effect. The time $\tau$ within which the effect manifests itself is quite short

$$\tau \sim 10^{-8}\,sec \tag{1.125}$$

We know that electrons are attached to atoms of size $r_0 \sim 10^{-10}\,m$: the energy power that hits a surface of atomic size, orthogonal to the radiation flux, is therefore

$$W = \Phi\, \times r_0^2 \sim 10^{-19}\,watt \tag{1.126}$$

Hence, the total energy which irradiates a single atom during the time $\tau$ within which the effect occurs is

$$E_{radiation} = \tau \times W \sim 10^{-27}\,J \tag{1.127}$$

We, therefore, encounter a first puzzle: the typical energy $E_{electron}$ of the emitted photo-electrons is many orders of magnitude larger than the electromagnetic energy that has hit the atoms during the relevant time scale.

Let us analyze the problem in more detail. Let us schematize the electron as one-dimensional damped harmonic oscillator, of proper frequency $\omega_0$ and damping frequency $\Gamma = \frac{1}{\tau}$. The equation of motion of the oscillator under the force of the time-dependent electric field $E(t)$ associated to the electromagnetic wave is

$$\ddot{x} + \omega_0^2\, x + \frac{1}{\tau}\,\dot{x} = \frac{e\,E(t)}{m} \tag{1.128}$$

Assuming the electromagnetic wave has angular frequency $\omega = 2\,\pi\,\nu$

$$E(t) = E\,\cos(\omega\,t) = \Re(E\,\mathrm{e}^{-i\,\omega\,t}) \tag{1.129}$$

the stationary solution of (1.128) has the form

$$x(t) = \Re(x_\omega\,\mathrm{e}^{-i\,\omega\,t}) \tag{1.130}$$

where

$$x_\omega = \frac{e\,E}{m}\,\frac{1}{\omega_0^2 - \omega^2 - i\,\frac{\omega}{\tau}} \tag{1.131}$$

The amplitude of the forced oscillations is therefore

$$|x_\omega| = \frac{e\,E}{m}\,\frac{1}{\sqrt{(\omega_0^2 - \omega^2)^2 + \frac{\omega^2}{\tau^2}}} \tag{1.132}$$

For the electron to be extracted from the atom it is necessary that this amplitude be larger than the atomic radius $r_0$

$$|x_\omega| \gtrsim r_0 \tag{1.133}$$

This gives the critical value of electric field:

$$E_{crit} = \frac{r_0\,m}{e}\,\sqrt{(\omega_0^2 - \omega^2)^2 + \frac{\omega^2}{\tau^2}} \tag{1.134}$$

This is the *minimal* amplitude of the electric field for which the electromagnetic wave is sufficiently strong to pull out an electron from the atom. The energy flux of the electromagnetic wave depends quadratically on the modulus of the electric field

$$\Phi = c\,\epsilon_0\,|E|^2 \tag{1.135}$$

We conclude that the minimal electromagnetic energy flux which is needed to trigger the effect is

$$\Phi_{crit} = \frac{c\,\epsilon_0\,r_0^2\,m^2}{e^2}\big[(\omega_0^2 - \omega^2)^2 + \frac{\omega^2}{\tau^2}\big] \tag{1.136}$$

We see that there are two different frequency regimes: one is the regime far away from the resonance

$$\omega \gg \omega_0 \gg \frac{1}{\tau} \tag{1.137}$$

In this regime

$$\Phi_{crit}^{non-res} \sim \frac{c\,\epsilon_0\,r_0^2\,m^2\,\omega^4}{e^2} \qquad \text{far from resonance}$$

The resonant regime instead is the one for which

$$\omega \sim \omega_0 \tag{1.138}$$

and

$$\Phi_{crit}^{res} \sim \frac{c\,\epsilon_0\,r_0^2\,m^2\,\omega_0^4}{e^2}\,\frac{1}{\omega_0^2\,\tau^2} \qquad \text{at resonance}$$

Let us estimate the order of magnitude of these energy fluxes, for $\omega \sim 10^{15}\,Hz$

$$\begin{aligned}\Phi_{crit}^{non-res} &\sim \frac{3\,10^8 \times 10^{-20}\,0.9^2 \times 10^{-60}\,10^{60}}{4\,\pi \times\; 9\,10^9 \times (1.6)^2\,10^{-38}}\,watt/m^2 \approx \\ &\approx 10^{15}\,watt/m^2\end{aligned}$$

We see therefore that, according to the classical model, a gigantic (and completely unrealistic) energy flux would be required to trigger the effect away from the resonance. This is in contraddiction with the experiment. When resonant conditions are met, we have instead

$$\Phi_{crit}^{res} \sim 10^{15}\,\frac{1}{10^{14}}\,watt/m^2 = 10\,watt/m^2$$

which is a realistic value.

Hence, the classical model predicts the existence of a *threshold intensity* of the incident radiation for the effect to occur. But it does not explain the experimentally observed *threshold frequency*, Eq. (1.121). Let us also estimate the energy of the photo-electron according to the classical model: we expect the electron to acquire an energy of the order of

$$T \sim e\,E\,r_0 \tag{1.139}$$

which is the work done by the electric field $E$ to pull the charge $e$ out of an atom of size $r_0$. Note that this energy depends on the intensity of the incident radiation but not on its frequency, contrary to the experimental behavior Eq. (1.121). Moreover the electric field

$$|E|^2 = \frac{\Phi}{c\,\epsilon_0} \tag{1.140}$$

is, for the relevant incident powers $\Phi \sim 10\,watts/m^2$ of the following order of magnitude

$$|E| \sim \sqrt{\frac{4\,\pi\ \times 9\,10^9\ \times 10}{3\,10^8}}\,V/m \sim 10^2 V/m \tag{1.141}$$

For such values

$$T \sim e\,E\,r_0 \sim 1.6\,10^{-19} \times 10^2\,10^{-10}\,J \sim 10^{-27}\,J \sim 10^{-8}\,eV \tag{1.142}$$

This is much less than the typical measured energies of the photo-electrons, which, as indicated in Eq. (1.123), are of the order of (few) electron-volts.

Summarizing, the classical model does not account, even qualitatively, for the experimentally observed features of the photoelectric effect: it predicts a threshold for the electromagnetic flux but not for the frequency; and it predicts that the energy absorbed by the photo-electrons increases with the intensity but not with the frequency of the incident radiation. Both features are in marked contrast with the experiment. Moreover, quantitatively, the model fails to explain the high energies with which the electrons are emitted.

Einstein's proposal was to assume that electromagnetic energy is exchanged between radiation and electrons in "packets", all carrying the same energy related to the radiation frequency as follows

$$E_\nu = h\,\nu \tag{1.143}$$

Denoting by $W$ the energy necessary to pull out an electron from the metal, one deduces that the maximal kinetic energy of the emitted electron is

$$T_{\max} = h\,\nu - W = h\,\big(\nu - \frac{W}{h}\big) \tag{1.144}$$

in agreement with (1.121) with

$$\nu_0 = \frac{W}{h} \tag{1.145}$$

Einstein models also explains the proportionality of the number of emitted electrons to the intensity of incoming radiation, since the latter is proportional to the flux of electromagnetic "packets" hitting the matter sample. To these "packets" it was given the name of "photons". The constant $h$ appearing in Einstein relations (1.143) has the dimensions of an angular momentum, or of time×energy and it is known as *Planck's constant.* As of matter of fact, it had already made its appearance in Planck's formula for the spectrum of black body radiation.

Let us evaluate the energy of photons for light in the visible spectrum (1.122):

$$E_{visible} \approx (2.6 - 5.0) \times 10^{-19} J \approx (1.6 - 3)\, eV$$

Recalling the relation between frequency $\nu$ and wavelength $\lambda$ valid for electromagnetic radiation traveling in the vacuum,

$$\nu = \frac{c}{\lambda} \Rightarrow E = \frac{h\,c}{\lambda} \tag{1.146}$$

one derives the useful formula

$$E[eV] = \frac{12400}{\lambda[\text{Å}]}$$

where $\lambda$ is expressed in Angström, $1\text{Å} = 10^{-8}\,cm$ and the energy $E$ in electron-volts.

**Problem 6.** A lamp of 100 W emits light whose wave length is $5.890\,10^{-7} m$. How many photons are emitted in 1 second?

## 1.6 Compton effect

The particle, or "corpuscular", aspect of electromagnetic radiation which emerges from Einstein's explanation of the photoelectric effect also gives a simple description of the Compton "shift", a phenomenon observed by A. H. Compton (1923) by studying scattering of electromagnetic radiation of high frequency (X or gamma rays) by electrons.

The physical process underlying Compton effect is the following. Consider an electron at rest, hit by an incident photon of momentum $\vec{p}$ and energy $E = c\,|\vec{p}| = h\,\nu$, according to Einstein hypothesis. Let $\vec{p'}$ and $E'$ momentum and energy of the scattered photon and $\vec{P}$ the momentum of the electron after the scattering. Energy and momentum conservation state that

$$\vec{p} = \vec{p'} + \vec{P} \qquad E + m\,c^2 = E' + \sqrt{m^2\,c^4 + c^2\,\vec{P}^2} \tag{1.147}$$

Therefore

$$\begin{aligned}(E - E' + m\,c^2)^2 &= m^2\,c^4 + c^2\,(\vec{p}^2 + \vec{p'}^2 - 2\,\vec{p}\cdot\vec{p'}) = \\ &= m^2\,c^4 + E^2 + (E')^2 - 2\,E\,E'\,\cos\theta\end{aligned} \tag{1.148}$$

where $\theta$ is the scattering angle, i.e. the angle formed by $\vec{p'}$ and $\vec{p}$. From this we obtain

$$(E - E')\,m\,c^2 = E\,E'\,(1 - \cos\theta) \tag{1.149}$$

and hence

$$\lambda' - \lambda = \frac{h}{m\,c}(1 - \cos\theta) \tag{1.150}$$

Therefore, radiation which is diffused forward with $\theta = 0$ has the same wavelength as the incident radiation but the wavelength of radiation which is diffused at $\theta \neq 0$ is shifted by an amount which is proportional to the so-called *Compton length* of the electron

$$\lambda_c \equiv \frac{h}{m\,c} \approx \frac{12400}{0.510 \times 10^6}\,10^{-8}\,cm \approx 2.43 \times 10^{-10}\,cm \tag{1.151}$$

Classically, the diffused radiation should have the same frequency as the incident one. Compton scattering shows that one can attribute to photons not only energy but also momentum which are both conserved in scattering processes. This strengthen the "corpuscular" picture of light.

## 1.7 Bohr's quantum theory of atomic spectra.

Einstein's idea that the exchange of electromagnetic energy occurs in "quanta" was used by Bohr to explain atomic spectra.

It had been observed that both the absorption and the emission spectra of matter (gas) are composed of discrete frequencies $\{\nu_\alpha;\ \alpha = 1, 2, \ldots\}$ — the so-called *spectral lines*. Spectroscopists had succeeded to express emission frequencies of a given substance as differences of "spectroscopic terms":

$$\frac{1}{\lambda_{\alpha_{n,m}}} = \frac{\nu_{\alpha_{n,m}}}{c} = T(n) - T(m) \tag{1.152}$$

with $n, m \geq 0$, integers. To each $n$ there corresponds a *spectral series*, and it was also observed that the absorption spectra correspond to the first series. Assuming Einstein relation between energy and frequency, one derives from (1.152) the energy of the absorbed or emitted photon corresponding to a given spectral line:

$$E_{n,m} = h\,\nu_{\alpha_{n,m}} = h\,c\,\big(T(n) - T(m)\big) \tag{1.153}$$

By energy conservation, this should be the same as the difference of the energies of the electron which emits or absorbs the photon: Bohr then concluded that

$$E_n = -h\,c\,T(n) \tag{1.154}$$

should be interpreted as the "allowed" energies of the atomic electrons.

Let us consider the hydrogen atom. Spectroscopists had found a simple formula for the spectroscopic term of this substance

$$T_H(h) = \frac{R_H}{n^2} \qquad n = 1, 2, \ldots \tag{1.155}$$

where

$$R_H = 109677.6\,cm^{-1} \tag{1.156}$$

is known as the hydrogen Rydberg constant. Bohr's hypothesis leads therefore to

$$E_n = -\frac{h\,c\,R_H}{n^2} \tag{1.157}$$

Together with the classical formula (1.6) for the radius of the Rutherford atom, this implies that the radiuses of the "allowed" orbits are quantized

$$r_n = \frac{e^2}{8\,\pi\,\epsilon_0\,|E|} = \frac{e^2\,n^2}{8\,\pi\,\epsilon_0\,h\,c\,R_H} \equiv \frac{\alpha}{4\,\pi\,R_H}\,n^2 \tag{1.158}$$

where we introduced the *fine-structure constant*

$$\alpha = \frac{e^2}{2\,\epsilon_0\,h\,c} \equiv \frac{e^2}{4\,\pi\,\epsilon_0\,\hbar\,c} \approx \frac{1}{137} \qquad \hbar \equiv \frac{h}{2\,\pi} \tag{1.159}$$

For $n = 1$ one obtains for the smallest hydrogen orbit radius the value

$$a_B = r_1 \approx 0.53 \times 10^{-8}\,cm \tag{1.160}$$

which is called the Bohr radius. This value is in agreement with expected atomic size, suggesting that Bohr's hypothesis of energy quantization is on the right track.

To develop further his hypothesis Bohr sought to derive an expression of $R_H$ in terms of fundamental constants. Since $h$ has the dimension of angular momentum, one is motivated to compute the angular momentum of an electron on circular "allowed" orbits

$$L_n = m\,v_n\,r_n = \sqrt{\frac{e^2}{4\,\pi\,\epsilon_0}}\sqrt{r_n\,m} = n\,\sqrt{\frac{e^4}{(4\,\pi\,\epsilon_0)^2}\frac{m}{2\,h\,c\,R_H}} \tag{1.161}$$

Thus the angular momentum on the "allowed" orbits is quantized and it is an *integer* multiple of a minimal angular momentum: Bohr's proposal was to identify this minimal angular momentum with a multiple of $h$:

$$L = n\,\sqrt{\frac{e^4}{(4\,\pi\,\epsilon_0)^2}\frac{m}{2\,h\,c\,R_H}} = n\,\hbar \qquad \hbar \equiv \frac{h}{2\,\pi} \tag{1.162}$$

From this one can derive the Rydberg constant

$$R_H = \frac{e^4\,m}{(4\,\pi\,\epsilon_0)^2\,c\,\hbar^3}\,\frac{1}{4\,\pi} = \frac{\alpha^2}{2\,\lambda_c} \approx 109737 cm^{-1} \tag{1.163}$$

in excellent agreement with experimental data (1.156). The discrete energy levels of the hydrogen atom write therefore

$$E^H_n = -\frac{e^4\,m}{(4\,\pi\,\epsilon_0)^2\,\hbar^2}\,\frac{1}{2\,n^2} \tag{1.164}$$

The lowest energy level is the one corresponding to $n = 1$

$$E_1^H = -\frac{e^4 \, m}{(4 \, \pi \, \epsilon_0)^2 \, \hbar^2} \, \frac{1}{2} \approx -13.6 \, eV \tag{1.165}$$

and it should be identified with the hydrogen ionization energy. The Bohr radius, i.e. the radius of the smallest circular orbit, is

$$a_B = \frac{4 \, \pi \, \epsilon_0 \, \hbar^2}{e^2 \, m} \tag{1.166}$$

**Problem 7.** Classical electromagnetism predicts that a charge rotating on a circular orbit with angular orbital frequency $\omega$ emits electromagnetic radiation with the same frequency. Show that the frequency of the photon emitted by an electron of the Bohr hydrogen atom which goes from the level $E_{n+1}^H$ to the level $E_n^H$, is, for $n$ *large*, approximately equal to the frequency of the classical circular motion corresponding to that energy.

This result, which is one way to derive Bohr's formula (1.163) for the Rydberg constant, is an instance of the so-called *correspondence principle*.

Bohr (and Sommerfeld) extended the quantization condition (1.162) to mechanical systems more general than the hydrogen atom, by reformulating it as the following condition

$$\oint p \, dq = n \, h \tag{1.167}$$

where $(q, p)$ are canonical conjugate pair of coordinate and momentum, and the integral is taken over the "allowed" classical orbits.

**Problem 8.** Apply the Bohr-Sommerfeld quantization condition (1.167) to the unidimensional harmonic oscillator $H = \frac{p^2}{2\,m} + \frac{1}{2} \, m \, \omega^2 \, q^2$ to show that its allowed energy levels are: $E_n = \hbar \, \omega \, n$

**Problem 9.** Apply the Bohr-Sommerfeld quantization condition (1.167) to an unidimensional free particle moving on a segment of length $L$ to show that its allowed energy levels are: $E_n = \frac{h^2 \, n^2}{8 \, m \, L^2}$.

**Problem 10.** Apply the Bohr-Sommerfeld quantization condition (1.167) to an unidimensional free particle moving in a potential well $V(x)$ defined by

$$V(x) = \begin{cases} -V_0 \text{ for } 0 < x < L \\ 0 \quad\;\; \text{for } x < 0 \text{ and } x > L \end{cases} \tag{1.168}$$

where $V_0 > 0$ is the depth of the well. Determine the discrete energy levels. How many they are? What is the minimum value of $V_0$ for which there exist discrete levels?

**Problem 11.** Apply Bohr-Sommerfeld quantization condition to a satellite of mass $m = 1\ kg$ rotating around the earth. Determine the allowed radiuses in terms of the mass of the earth $M$ and the universal constant of gravitation $G$. Suppose the satellite is on on circular orbit of radius near the radius of the earth $R = 6400\ km$, with a given Bohr integer number $\bar{n}$. How much does the radius change if the satellite shifts to an orbit with $n = \bar{n} + 1$? Express the answer in meters.

**Problem 12.** Apply the Bohr-Sommerfeld quantization condition (1.167) to compute the energy levels of an unidimensional free particle moving in the potential

$$V(x) = c\,|x| \tag{1.169}$$

**Problem 13.** Apply the Bohr-Sommerfeld quantization condition (1.167) to compute the energy levels and the orbits of a particle of mass $m$ moving in circular orbits of radius $r$ in a potential $V(r) = \sigma\, r$.

**Problem 14.** Apply the Bohr-Sommerfeld quantization condition to derive energy levels and allowed orbit radiuses for a particle of mass $m$ moving in circular orbits of radius $r$ attracted toward the center of the orbit by an elastic force $F = -k\, r$.

**Problem 15.** Derive the numerical values for the Rydberg constants predicted by Bohr atomic model for the following hydrogenoids: hydrogen $H$, deuteron $D$ (whose nucleus has one proton and one neutron), $He^+$, $Li^{++}$ and, for the same elements, the wave lengths of the spectral lines corresponding to the transition between the first excited level and the fundamental level.

## 1.8 Quantum theory of gas specific heat capacities

### 1.8.1 Einstein and Debye theories of specific heat capacities of solids

Let us consider the implications of Bohr quantization condition for the specific heat capacities of solids. The classical partition function (1.92) should be replaced for a system with quantized energy levels $E_\alpha$ by

$$Z(\beta) = \sum_\alpha \mathrm{e}^{-\beta\, E_\alpha} \tag{1.170}$$

Following Einstein, let us first schematize the solid as a set of $N$ decoupled identical 3-dimensional oscillators of frequency $\omega$

$$H(\vec{x},\vec{p}) = \frac{\vec{p}^2}{2\,m} + \frac{1}{2}\,m\,\omega^2\,\vec{x}^2 \tag{1.171}$$

The levels obtained from Bohr quantization condition are[8]

$$E_{\vec{n}} = \hbar\big(n_1+n_2+n_3\big)\,\omega \qquad \vec{n} = \big(n_1,n_2,n_3\big) \qquad n_i = 1,2,\ldots \tag{1.172}$$

The quantum partition function becomes

$$Z(\beta) = \prod_{i=1,2,3}\sum_{n_i=0}^{\infty} \mathrm{e}^{-\hbar\,\beta\,\omega\,n_i} = \prod_{i=1,2,3}\frac{1}{1-\mathrm{e}^{-\beta\,\hbar\,\omega}} = \frac{1}{(1-\mathrm{e}^{-\beta\,\hbar\,\omega})^3} \tag{1.173}$$

Thus the internal energy is

$$U = N\,\frac{\sum_\alpha E_\alpha\,\mathrm{e}^{-\beta\,E_\alpha}}{Z(\beta)} = -N\,\partial_\beta\,\log Z(\beta) = \frac{3\,N\,\hbar\,\omega\,\mathrm{e}^{-\beta\,\hbar\,\omega}}{1-\mathrm{e}^{-\beta\,\hbar\,\omega}} \tag{1.174}$$

and the molar heat capacity

$$\begin{aligned} C_V &= \frac{\partial U}{\partial T}\Big|_V = -\frac{\beta}{T}\,\frac{\partial U}{\partial\beta}\Big|_V = k\,N_A\,\beta^2\partial_\beta^2\,\log Z(\beta) = \\ &= 3\,N_A\,k\,\frac{(\hbar\,\omega)^2}{(k\,T)^2}\,\frac{\mathrm{e}^{-\beta\,\hbar\,\omega}}{(1-\mathrm{e}^{-\beta\,\hbar\,\omega})^2} \end{aligned} \tag{1.175}$$

For $T$ large one recovers the classical prediction

$$C_V \approx 3\,N_A\,k\,\frac{(\hbar\,\omega)^2}{(k\,T)^2}\,\frac{1}{(\beta\,\hbar\,\omega)^2} = 3\,N_A\,k \qquad k\,T \gg \hbar\,\omega \tag{1.176}$$

For $T$ small on obtains instead

$$C_V \approx 3\,N_A\,k\,\frac{(\hbar\,\omega)^2}{(k\,T)^2}\,\mathrm{e}^{-\frac{\hbar\,\omega}{k\,T}} \to 0 \qquad\qquad k\,T \ll \hbar\,\omega \tag{1.177}$$

---

[8] Notice that these levels *differ* by a constant independent of $n$ from the correct levels dictated by quantum mechanics, as it will be discussed later on. This constant does not affect the computation which follows.

This does agree with experimental data at *qualitative*, but not *quantitative*, level. Experimentally $C_V$ vanishes with a power law for $T \to 0$

$$C_V \propto T^3 \qquad \text{as } T \to 0 \qquad \text{experimentally} \tag{1.178}$$

and not exponentially as predicted by the Einstein model.

A variant of this model is the Debye model, which, more realistically, schematizes the crystal with $3\,N$ decoupled 1-dimensional oscillators with *different* frequencies $\omega_a$ with $a = 1, 2, \ldots, 3\,N$. This is equivalent to assume that the original $N$ identical 3-dimensional oscillators are coupled among themselves. The $\omega_a$ are the frequencies of the so-called *normal modes* of the oscillators system. The internal energy of Debye model is therefore

$$U = \sum_{a=1}^{3\,N} \frac{\hbar\,\omega_a\,\mathrm{e}^{-\beta\,\hbar\,\omega_a}}{1 - \mathrm{e}^{-\beta\,\hbar\,\omega_a}} \tag{1.179}$$

and the molar heat capacity

$$C_V = k \sum_{a=1}^{3\,N_A} \frac{(\hbar\,\omega_a\,\beta)^2\,\mathrm{e}^{-\beta\,\hbar\,\omega_a}}{(1 - \mathrm{e}^{-\beta\,\hbar\,\omega_a})^2} \tag{1.180}$$

For $T \gg \hbar\,\omega_{max}$, where $\omega_{max}$ is the greatest of the frequencies $\{\omega_a\}$, the behavior of Debye model is the same as Einstein's model and reproduces the classical limit

$$C_V \approx 3\,k \sum_{a=1}^{3\,N_A} \frac{(\hbar\,\omega_a\,\beta)^2}{(\hbar\,\omega_a\,\beta)^2} = 3\,k\,N_A \qquad k\,T \gg \hbar\,\omega_{max} \tag{1.181}$$

For low temperature, i.e. for $\beta \to \infty$ the behavior of the Debye model is different. To study this let us approximate the sum in (1.179) with an integral

$$U = \int_{\omega_{min}}^{\omega_{max}} d\omega\,\frac{d\,n(\omega)}{d\,\omega}\,\frac{\hbar\,\omega\,\mathrm{e}^{-\beta\,\hbar\,\omega}}{1 - \mathrm{e}^{-\beta\,\hbar\,\omega}} \tag{1.182}$$

where $n(\omega)$ is the number of normal modes with frequency less than $\omega$. To compute $n(\omega)$ one needs to find the normal modes of the oscillators system: these are known as the "phonon" excitations of the solid. They are labeled by a discrete index $\vec{k}$, called *pseudo-momentum*, taking values in a lattice, very much like the normal modes (1.116) of the electromagnetic wave in a box

$$\vec{k} = \frac{\pi}{L}\,\vec{n} \qquad \vec{n} = (n_1, n_2, n_3) \qquad n_i \leq \frac{L}{\Delta} \tag{1.183}$$

The important difference with respect to photons is that the integers $n_i$ are limited by a maximal value $n_{max} = \frac{L}{\Delta}$, where $\Delta$ is the *lattice spacing* and $L$ is the size of the solid sample. The relation between $\omega_{\vec{k}}$ and $\vec{k}$ is called the *dispersion relation* of the given solid. For an isotropic solid and low pseudo-momenta the dispersion relation of phonons has the same form as that of photons

$$\omega_{\vec{k}} = v\,|\vec{k}| \qquad \vec{k} \to 0 \tag{1.184}$$

where $v$ is the speed of sound of the solid. With this assumption, we can repeat the computation which leads to (1.118) and obtain the number $N_\omega$ of pseudo-momenta corresponding to phonons of energy less than $\omega$:

$$n(\omega) = 3 \times \frac{L^3}{6\,\pi^2\,v^3}\,\omega^3 = \frac{L^3}{2\,\pi^2\,v^3}\,\omega^3 \tag{1.185}$$

where the factor of 3 takes into account that, unlike photons, phonons may be both transverse and longitudinal. Thus[9]

$$3\,N_A = \frac{L^3}{2\,\pi^2\,v^3}\,\omega_{max}^3 \Rightarrow \omega_{max} = v\left(\frac{N_A}{V}\right)^{\frac{1}{3}}\pi\left(\frac{6}{\pi}\right)^{\frac{1}{3}} \tag{1.186}$$

We have therefore for the internal energy

$$U = \frac{3\,L^3}{2\,\pi^2\,v^3}\int_0^{\omega_{max}} d\omega\,\frac{\hbar\,\omega^3\,\mathrm{e}^{-\beta\,\hbar\,\omega}}{1-\mathrm{e}^{-\beta\,\hbar\,\omega}}$$

It is convenient to define a temperature $T_D$ — called the *Debye temperature* of the given solid — as follows

$$T_D \equiv \frac{\hbar\,\omega_{max}}{k} \tag{1.187}$$

The molar heat capacity writes

$$\begin{aligned} C_V &= \frac{3\,k\,L^3}{2\,\pi^2\,v^3}\int_0^{\omega_{max}} d\,\omega\,\frac{(\hbar\,\omega\,\beta)^2\,\omega^2\,\mathrm{e}^{-\beta\,\hbar\,\omega}}{(1-\mathrm{e}^{-\beta\,\hbar\,\omega})^2} = \\ &= \frac{3\,k\,L^3}{2\,\pi^2\,v^3\,\hbar^3\,\beta^3}\int_0^{\frac{\hbar\omega_{max}}{k\,T}} \frac{dx\,x^4\,\mathrm{e}^{-x}}{(1-\mathrm{e}^{-x})^2} = \\ &= 9\,k\,N_A\,\frac{T^3}{T_D^3}\int_0^{\frac{T_D}{T}} \frac{dx\,x^4\,\mathrm{e}^{-x}}{(1-\mathrm{e}^{-x})^2} \end{aligned} \tag{1.188}$$

---

[9] We are assuming, *incorrectly*, that the relation (1.184), which is valid for low pseudo-momenta, can be extrapolated to any $\vec{k}$. For this reason the expression we are going to obtain agrees well with experimental data only in the low and high temperature limits.

**Problem 15.** Show that the Debye model predicts

$$\begin{aligned} C_V &\approx 3\,k\,N_A && T \gg T_D \\ C_V &\approx k\,N_A\,\frac{12\,\pi^4}{5}\,\frac{T^3}{T_D^3} = \frac{L^3}{v^3\,\hbar^3}\,\frac{2\,\pi^2}{5}\,k^4\,T^3 && T \ll T_D \end{aligned} \tag{1.189}$$

**Problem 16.** Suppose that (1.184) and (1.185) are only valid only for *low* pseudo-momenta. Assume therefore that the density of states grows quadratically with the frequency

$$\frac{dn(\omega)}{d\,\omega} = A\,\omega^2 + O(\omega^3) \quad \text{for } \omega \to 0 \tag{1.190}$$

only for *small* frequencies $\omega$. Assume $n(\omega)$ is an otherwise generic function of $\omega$ for bigger $\omega$ up to a certain given $\omega_{max}$, which defines a Debye temperature $T_D$ as in (1.187). Show that the predictions of the Debye model for the specific heats in the high and low temperature limits are given by the same expressions in (1.189), once the relation $A = \frac{3\,L^3}{2\,\pi^2\,v^3}$ we derived in (1.185) is used.

### 1.8.2 Specific heat capacities of gases

The Hamiltonian associated to mono-atomic gases

$$H = \frac{\vec{p}^2}{2\,m} \tag{1.191}$$

describes free particles moving in a box of size $L$. Bohr quantization condition gives for the energy levels

$$E_{\vec{n}} = \frac{h^2\,(n_1^2 + n_2^2 + n_3^2)}{8\,m\,L^2} \tag{1.192}$$

For a box of macroscopic size $L \approx 10^{-2}\,m$ and atomic masses $m \approx 10^{-27}\,kg$ one obtains for the lowest energy level

$$E_0 = \pi^2 \left(\frac{0.5 \times 10^{-10}}{10^{-2}}\right)^2 \times \frac{1}{2000} \times 13.6\,ev \approx 2 \times 10^{-18}\,eV \tag{1.193}$$

The separation between levels is for $n_i = n \gg 1$

$$\Delta E_n = E_{n+1} - E_n \sim n\,E_0 \tag{1.194}$$

Levels which are appreciably populated are those for which $E_{\vec{n}} \lesssim k\,T$, i.e.

$$n^2 \lesssim \frac{k\,T}{E_0} \tag{1.195}$$

For those levels

$$\Delta E_n \lesssim \sqrt{k\,T\,E_0} \tag{1.196}$$

The quantum thermodynamical properties of the gas will be significantly different than the classical ones for temperatures

$$k\,T \ll \Delta E_n \lesssim \sqrt{k\,T\,E_0} \Leftrightarrow \sqrt{k\,T} \ll \sqrt{E_0} \tag{1.197}$$

Since $T = 1\,K \approx \frac{1}{12000}\,\frac{eV}{k}$, we conclude that we expect quantum effect be relevant for monoatomic perfect gases for temperatures $T \lesssim 10^{-14}\,K$. Since gases become liquid at temperatures much higher than those, we conclude that the classical prediction $C_V = \frac{3}{2}\,R$ remains unchanged at quantum level.

**Problem 17.** Discuss the effect of quantization on the specific heat capacities of bi-atomic gases, knowing that: a) the rotational degrees of freedom around an axis orthogonal to the axis connecting the two atoms give rise to discrete energy levels with separation $\Delta E \sim 10^{-4} - 10^{-2}\,eV$; b) vibrational degrees of freedom have $\Delta E \sim 10^{-1}\,eV$; c) rotational degrees of freedom around the axis connecting the two atoms have $\Delta E \sim 10\,eV$.

## 1.9 Planck distribution of black body radiation

Let us reconsider the classical computation of black body radiation reviewed in section 1.4, where we considered the ensemble of harmonic oscillators associated to the modes of the electromagnetic field in a box of size $L$, whose frequencies are given in (1.116)

$$\omega_{\vec{k}} = c\,|\vec{k}| \qquad \vec{k} = \frac{\pi}{L}\,\vec{n} \qquad \vec{n} = (n_1, n_2, n_3) \tag{1.198}$$

The quantum internal energy is

$$U(\beta) = \sum_{\vec{k}} \frac{\hbar\,\omega_{\vec{k}}\,\mathrm{e}^{-\hbar\,\beta\,\omega_{\vec{k}}}}{1 - \mathrm{e}^{-\hbar\,\beta\,\omega_{\vec{k}}}} \tag{1.199}$$

We already computed the number of oscillators with frequency less than $\omega$ to be

$$2\,N_\omega = \frac{L^3}{3\,\pi^2\,c^3}\,\omega^3 \tag{1.200}$$

Therefore, transforming the sum into an integral, which is valid for $L \to \infty$, we have

$$U(\beta) = \frac{L^3}{\pi^2\,c^3}\,\int_0^\infty d\omega \frac{\omega^3\,\hbar\,\mathrm{e}^{-\hbar\,\beta\,\omega}}{1-\mathrm{e}^{-\hbar\,\beta\,\omega}}$$

The energy density distribution is therefore

$$\frac{du(\omega)}{d\omega} = \frac{\hbar}{\pi^2\,c^3}\,\frac{\omega^3}{\mathrm{e}^{\hbar\,\beta\,\omega}-1}$$

which is the celebrated Planck result.

## 1.10 De Broglie relation

Einstein's relation, $E = h\,\nu$, expressing the "particle-like" or corpuscular aspect of electromagnetic *waves*, involves the same Planck constant $h$ which appears in Bohr-Sommerfeld quantization condition (1.167) which applies instead to *particles*. This suggested to L. De Broglie that particles might in turn exhibit a wave nature.

Consider for example a 1-dimensional particle moving in a box of size $L$; one has from (1.167) that

$$p\,L = n\,h \Rightarrow L = n\,\frac{h}{p} \tag{1.201}$$

Remembering that for a photon $\frac{h}{p} = \lambda$, De Broglie (1923) proposed that a *wavelength* is to be associated to any *particle* of momentum $p$ according to the relation

$$\lambda = \frac{h}{p} \tag{1.202}$$

Bohr-Sommerfeld quantization condition can be restated, within De Broglie's proposal, by saying that the "allowed" classical orbits of a particle are those whose length is an integer multiple of the particle wavelength (1.202)

$$L = n\,\lambda \tag{1.203}$$

**Problem 18.** Compute the De Broglie wave length $\lambda$ of the electron in a circular orbit of radius $r$ of the hydrogen atom. Show that the condition that the length of the orbit be an integer number of wave length, i.e.

$$n\,\lambda = 2\,\pi\,r \tag{1.204}$$

is equivalent to Bohr's quantization condition.

For a photon De Broglie relation reduces to Einstein relation. For an electron of energy $E$, instead, it gives

$$\lambda = \frac{h}{\sqrt{2\,m\,E}} = 2\,\pi\,a_B\,\sqrt{\frac{|E_1^H|}{E}} \approx \frac{12.26\text{Å}}{\sqrt{E[ev]}} \tag{1.205}$$

**Problem 19.** Show that: a) for a macroscopic body wave (quantum) effects implied by De Broglie relation are negligible; b) for atoms in matter they show up at very low temperatures; c) for electrons in atoms or solid they are important.

The De Broglie wavelength associated to an electron of energy $E \sim 100\,ev$ is of the order of the Å, which is the wavelength of X rays. The wave nature of electrons at these energies should therefore show up in experimental contexts similar to those required to study X rays diffraction.

Diffraction patterns for X rays of wavelength $\lambda$ are produced by scattering on atomic lattices whose lattice spacing $d$ is of the same order of magnitude as $\lambda$. Davisson and Germer (1927) performed diffraction experiments on the atomic lattice of a crystal of nickel, by replacing X rays with monochromatic (i.e. mono-energetic) electrons. The diffused electrons turned out to follow similar diffraction pattern as X rays, with diffraction peaks at angles determined by the Bragg's law

$$2\,d\,\sin\theta = n\,\lambda \tag{1.206}$$

Davisson and Germer were able in this way to directly check De Broglie relation (1.202).

# 2 The Principles of Quantum Mechanics

## 2.1 The double-slit experiment, wave-particle duality and Heisenberg uncertainty principle

Read the bibliography.

## 2.2 The rules of Quantum Mechanics

### 2.2.1 Superposition principle

- States of a quantum system are represented by elements — *vectors* — of a *complex* vector space $\mathcal{H}$.
- Vectors differing by the multiplication of a complex number correspond to the *same* physical state:

$$v, v' \in \mathcal{H} \quad \alpha \in \mathbb{C} \quad v' = \alpha\, v \Leftrightarrow \quad v' \sim v \tag{2.1}$$

*Observations*

- The superposition principle expresses the fact that if $v_1$ and $v_2$ correspond to two physical states, the vector $v_1 + v_2$ corresponds to another physical state, which has "something" to do with both $v_1$ and $v_2$.
- The space of states $\mathcal{H}$ depends on the particular physical system under consideration. In general $\dim \mathcal{H} = \infty$, but specific aspects of a given system maybe be described by finite-dimensional vector spaces. It is part of the theory to specify the physical informations necessary to characterize a "state" of the system.

### 2.2.2 Observables, eigenstates and eigenvalues

*Some definitions*

- The *observables* are the physical quantities that can be measured on a system (E.g.: position, momentum, kinetic energy, angular momentum, etc.).
- The possible results of the measure of an observable $\mathcal{O}$ are real numbers $\{\lambda_1, \lambda_2, \ldots\}$. The $\lambda_i$'s are called the *eigenvalues* of the observable $\mathcal{O}$.

- The result of the measure of an observable $\mathcal{O}$ on a state $v$ is not *in general* determined a priori, but the theory predicts the *probabilities* of the outcome: we will denote by $P_i(v)$ the probability to obtain $\lambda_i$ for the measure of $\mathcal{O}$ on $v$. We have $\sum_i P_i(v) = 1$.

- The *eigenstates* of an observable $\mathcal{O}$ are those states for which the result of the measure of $\mathcal{O}$ is certain: $P_i(v) = \delta_{i,i_0}$. We will denote a vector representing such an eigenstate by either $|\lambda_{i_0}\rangle$ or $v_{\lambda_{i_0}}$, and we will say that $\mathcal{O}$ has a *definite* value $\lambda_{i_0}$ on it.

- $\lambda_i$ is said to be a *non-degenerate eigenvalue* if there exists a unique eigenstate with eigenvalue $\lambda_i$. $\mathcal{O}$ is said to be *non-degenerate* if all its eigenvalues are non-degenerate.

*Postulates*

- (Von Neumann) If $\lambda_i$ is the result of the measure of $\mathcal{O}$ on $v$, immediately after the measure the state of the system is an eigenstate of $\mathcal{O}$ with eigenvalue $\lambda_i$. (This defines an "ideal" measure, and it is postulated that for each observables there exist arbitrarily ideal measure devices.)

- Any vector $v$ is an eigenstate of some non-degenerate observable $\mathcal{O}_v$. $v$ can be "prepared" by measuring some observable.

### 2.2.3 Transition amplitudes and transition probabilities

- The probability $P(v \to w)$ to obtain a state $w$ after a measure on the state $v$ is called the *transition probability* from $v$ to $w$.

- There exists on $\mathcal{H}$ a positive definite hermitian product $\langle \cdot, \cdot \rangle$ such that

$$P(v \to w) = \frac{\left|\langle w, v\rangle\right|^2}{\langle v, v\rangle \langle w, w\rangle} \tag{2.2}$$

*Observations*

- It is convenient to work with normalized vectors $|v| = |w| = 1$. In this case $P(v \to w) = \left|\langle w, v\rangle\right|^2$ and $\langle w, v\rangle$ is called the *transition amplitude*.

- One calls "observable" both the physical quantity and the device used to measure it.

- The measure of an observable on a state $v$ "perturbs" the state $v \to v_{\lambda_{i_0}}$ except in the case in which $v$ is an eigenstate of $\mathcal{O}$.
- After the measure of a non-degenerate observable the system will find itself in a completely determined state.
- If $\mathcal{O}$ is non degenerate, $v$ is a state, then the transition probability $P(v \to |\lambda_i\rangle) = P_i(v)$.
- $P$ in (2.2) satisfies $0 \leq P \leq 1$ thanks to the Schwartz inequality: $|\langle v, w\rangle| \leq |v|\,|w|$.
- $P(v \to w) = 1 \Leftrightarrow v = \alpha\, w$, with $\alpha \in \mathbb{C}$.

## 2.3 Basis of eigenvectors

In this subsection we will show that given an observable $\mathcal{O}$ one can always find an *orthonormal basis* $\{|\lambda_i\rangle \equiv v_{\lambda_i}\}$ of eigenvectors of $\mathcal{O}$.

### 2.3.1 Non-degenerate observables

Let $\mathcal{O}$ be a non-degenerate observable and $|\lambda_i\rangle \equiv v_{\lambda_i}$ its eigenvectors. Let us choose $|v_{\lambda_i}| = 1$. By our definitions

$$P_i(|\lambda_j\rangle) = \delta_{ij} \Leftrightarrow \langle v_{\lambda_i}, v_{\lambda_j}\rangle = \delta_{ij} \tag{2.3}$$

*Theorem*

- $\{v_{\lambda_i}\}$ is a (orthonormal) basis of $\mathcal{H}$.

*Proof:* First notice that there is no non-vanishing $v$, such that $\langle v, v_{\lambda_i}\rangle = 0$ for any $i$. If this were the case, in fact, $P_i(v) = 0$ for any $i$, which is impossible since we know instead that $\sum_i P_i(v) = 1$.

Now, if $\{v_{\lambda_i}\}$ were not a basis, there would exists a (non-vanishing) $v$ which is not a linear combination of $v_{\lambda_i}$. Consider then

$$v' = v - \sum_i (v, v_{\lambda_i}) v_{\lambda_i} \tag{2.4}$$

$v'$ does not vanish otherwise $v$ would be a linear combination of the $\{v_{\lambda_i}\}$. Moreover $v'$ is orthogonal to all $v_{\lambda_i}$, i.e.

$$(v', v_{\lambda_j}) = (v, v_{\lambda_j}) - (v, v_{\lambda_j}) = 0 \qquad \forall j \tag{2.5}$$

But we just showed that there are no such $v'$. Therefore there are no $v$ which are not linear combinations of the $v_{\lambda_i}$'s. Q.E.D..

Then given any $v \in \mathcal{H}$ we have

$$v = \sum_i c_i \, v_{\lambda_i} \qquad c_i \in \mathbb{C} \tag{2.6}$$

Then

$$c_i = \frac{\langle v_{\lambda_i}, v \rangle}{\langle v_{\lambda_i}, v_{\lambda_i} \rangle} \tag{2.7}$$

and

$$v = \sum_i v_{\lambda_i} \, \frac{\langle v_{\lambda_i}, v \rangle}{\langle v_{\lambda_i}, v_{\lambda_i} \rangle} = \sum_i |\lambda_i\rangle \, \frac{1}{\langle \lambda_i | \lambda_i \rangle} \, \langle \lambda_i | v \rangle \tag{2.8}$$

The numbers $c_i$ are the transition amplitudes from $v \to v_{\lambda_i}$ and $|c_i|$ are the corresponding transition probabilities.

### 2.3.2 Degenerate observables

Let $\mathcal{O}$ be a degenerate observable and let $\mathcal{H}_{\lambda_i}$ the space of eigenvectors of $\mathcal{O}$ with eigenvalue $\lambda_i$.

**Problem 1.** Show that $\mathcal{H}_{\lambda_i}$ is a linear space.

$\mathcal{H}_{\lambda_i}$ is called the *eigenspace* associated to $\lambda_i$. Its dimension is called the *degeneracy* of the eigenvalue $\lambda_i$.

**Problem 2.** Show that eigenspaces $\mathcal{H}_{\lambda_i}$ and $\mathcal{H}_{\lambda_j}$ corresponding to different eigenvalues $i \neq j$ are orthogonal.
**Problem 3**. Show that $\mathcal{H} = \oplus_i \mathcal{H}_{\lambda_i}$.

It follows that, even if a *generic* basis of eigenvectors of $\mathcal{H}$ is *not* orthonormal, we can always choose a basis of $\mathcal{H}$ which is an orthonormal basis of eigenvectors of $\mathcal{H}$. Such a basis is not unique.

Summarizing, in both the degenerate and non-degenerate case, given a generic vector $v \in \mathcal{H}$, we can decompose it as

$$v = \sum_i v_{\lambda_i} \tag{2.9}$$

where $v_{\lambda_i} \in \mathcal{H}_{\lambda_i}$ are called the *orthogonal projection* of $v$ onto $\mathcal{H}_{\lambda_i}$. In the generic (degenerate) case, $v_{\lambda_i}$ are linear combinations of an orthonormal set of eigenvectors with the same eigenvalue $\lambda_i$

$$v_{\lambda_i} = \sum_\alpha c_{i;\alpha}\, v^{(\alpha)}_{\lambda_i} \qquad \langle v^{(\alpha)}_{\lambda_i}, v^{(\beta)}_{\lambda_i} \rangle = \delta_{\alpha\beta} \tag{2.10}$$

where

$$c_{i,\alpha} = \langle v^{(\alpha)}_{\lambda_i}, v \rangle \tag{2.11}$$

It is important to note that although $v_{\lambda_i}$ are orthogonal among themselves

$$\langle v_{\lambda_i}, v_{\lambda_j} \rangle = 0 \qquad \text{for} \quad i \neq j \tag{2.12}$$

their norm is not, in general, unity

$$\langle v_{\lambda_i}, v_{\lambda_i} \rangle = \sum_\alpha |c_{i;\alpha}|^2 \tag{2.13}$$

If we choose $|v| = 1$, this is precisely the probability that a measure of $\mathcal{O}$ on $v$ gives $\lambda_i$

$$P_{\lambda_i}(v) = \sum_\alpha P(v \to v^{(\alpha)}_{\lambda_i}) = \sum_\alpha |\langle v^{(\alpha)}_{\lambda_i}, v \rangle|^2 = \sum_\alpha |c_{i;\alpha}|^2 \tag{2.14}$$

Therefore we can refine the postulate of Von Neumann stated before, for the case of degenerate eigenvalues, as follows:

- If the measure of $\mathcal{O}$ on $v$ gives $\lambda_i$, the system immediately after the measure is in the state $v_{\lambda_i}$.

Indeed,

$$\begin{aligned} P(v \to v_{\lambda_i}) = \frac{|\langle v_{\lambda_i}, v \rangle|^2}{\langle v_{\lambda_i}, v_{\lambda_i} \rangle} = \frac{|\langle v_{\lambda_i}, v_{\lambda_i} \rangle|^2}{\langle v_{\lambda_i}, v_{\lambda_i} \rangle} = \\ = \langle v_{\lambda_i}, v_{\lambda_i} \rangle = P_{\lambda_i}(v) \end{aligned} \tag{2.15}$$

## 2.4 Operators and observables

A *linear operator* $\mathcal{O}_{op}$ is associated to an observable $\mathcal{O}$ as follows

$$\mathcal{O}_{op}\, v = \sum_i \lambda_i\, v_{\lambda_i} = \sum_i \lambda_i \sum_\alpha v^{(\alpha)}_{\lambda_i}\, \langle v^{(\alpha)}_{\lambda_i}, v\rangle \tag{2.16}$$

- The *average* of the results obtained by measuring $\mathcal{O}$ on $v$ is

$$(v, \mathcal{O}_{op}\, v) = \sum_i \lambda_i \langle v_{\lambda_i}, v_{\lambda_i}\rangle = \sum_i \lambda_i\, P_{\lambda_i} \tag{2.17}$$

- When $\mathcal{O}_{op}$ acts on an eigenvector $v^{(\alpha)}_{\lambda_i}$ with eigenvalue $\lambda_i$ it produces the same vector multiplied by $\lambda_i$:

$$\mathcal{O}_{op}\, v^{(\alpha)}_{\lambda_i} = \lambda_i\ v^{(\alpha)}_{\lambda_i} \tag{2.18}$$

- $\mathcal{O}_{op}$ is *hermitian*

$$\langle w, \mathcal{O}_{op}\, v\rangle = \langle \mathcal{O}_{op}\, w, v\rangle \Leftrightarrow \mathcal{O}_{op} = \mathcal{O}^\dagger_{op} \tag{2.19}$$

**Problem 4.** Show that if $\mathcal{O}_{op}\, v = \lambda_i\, v$, then $v$ is an eigenvector of $\mathcal{O}$ with eigenvalue $\lambda_i$.
**Problem 5.** Show that if $\hat{O}$ is an hermitian operator (i.e. if $\hat{O} = \hat{O}^\dagger$) and $v$ is such that $\hat{O}\, v = \lambda\, v$, then $\lambda \in \mathbb{R}$.
**Problem 6.** Show that if $\hat{O}$ is an hermitian operator, and $\hat{O}\, v = \lambda\, v$ and $\hat{O}\, v' = \lambda'\, v'$ with $\lambda \neq \lambda'$, then $\langle v, v'\rangle = 0$.

Given an hermitian linear operator $\hat{O}$, we are going to call vectors which satisfy $\hat{O}\, v = \lambda\, v$ for some $\lambda \in \mathbb{R}$ eigenvectors of that linear operator with eigenvalue $\lambda$. We are going to use therefore the same words "eigenvector" and "eigenvalue" that we introduced for observables also in relation to linear operators. This terminology is justified by what we proved above, i.e. that eigenvectors and eigenvalues of an observable are eigenvector and eigenvalues of the associated linear operator (and vice-versa, see *Problem 4*).

We have seen that the linear operators $\mathcal{O}_{op}$ which correspond to observables are hermitian and admit a basis of eigenvectors. We are going to use the same word "observable" that we have used for measurable quantities also for an hermitian operator with such a property, i.e. for hermitian operators which have enough eigenvectors to form a basis.

- An hermitian operator on a finite dimensional linear space is an observable.
- Not all hermitian operators on infinite dimensional linear spaces are observables

## 2.5 Compatible observables

In classical mechanics given the state, every observable has a determined value. For example, the state of a unidimensional particle is defined by the couple $(q, p)$ of coordinate e momentum. An observable is a real function $O(q, p)$ and it has a determined value given $(q, p)$. In quantum mechanics this is not so.

We have seen that given the state $v$, the value of an observable $\mathcal{O}$ is not determined in general. It is determined only if $v$ is an eigenstate with a given eigenvalue $\lambda$. Let us suppose this be the case, $v = v_\lambda$. Consider now a *different* observable $\mathcal{O}'$. In general $v_\lambda$ will not be an eigenstate of $\mathcal{O}'$: in this case $\mathcal{O}$ has a definite value on $v$ but $\mathcal{O}'$ does not. $\mathcal{O}'$ has a definite value on $v$ only if $v_\lambda$ is *also* an eigenvector of $\mathcal{O}'$. In other words $\mathcal{O}$ and $\mathcal{O}'$ have definite values on $v$ only if

$$\mathcal{O}_{op}\, v = \lambda\, v \qquad \mathcal{O}'_{op}\, v = \eta\, v \tag{2.20}$$

We will say that $v$ is a *simultaneous* eigenvector of $\mathcal{O}$ and $\mathcal{O}'$ and we will write $v = v_{\lambda;\eta}$.

**Definition**: When two observables have enough simultaneous eigenvectors to form a basis, they are said to be *compatible*.

**Problem 7.** Show that two observables $\mathcal{O}$ and $\mathcal{O}'$ are compatible if and only if the measure of $\mathcal{O}'$ on *any* eigenstate of $\mathcal{O}$ of eigenvalue $\lambda$ produce a state which is still an eigenstate of $\mathcal{O}$ with the same eigenvalue $\lambda$.

In other words two observables are compatible if they can be simultaneously measured on a state: the measure of one observable on an eigenstate of the other does not destroy the information regarding the first observable.

Let us check that the operators corresponding to compatible observables commute:

$$\big[\mathcal{O}^{(1)}_{op}, \mathcal{O}^{(2)}_{op}\big] \equiv \mathcal{O}^{(1)}_{op}\, \mathcal{O}^{(2)}_{op} - \mathcal{O}^{(2)}_{op}\, \mathcal{O}^{(1)}_{op} = 0 \tag{2.21}$$

Indeed let $\{v_{\lambda_i;\eta_\alpha}\}$ be a basis of simultaneous eigenstates. Then

$$\mathcal{O}^{(1)}_{op}\, v_{\lambda_i;\eta_\alpha} = \lambda_i\, v_{\lambda_i;\eta_\alpha} \tag{2.22}$$

Hence

$$\mathcal{O}^{(2)}_{op}\, \mathcal{O}^{(1)}_{op}\, v_{\lambda_i;\eta_\alpha} = \lambda_i\, \mathcal{O}^{(2)}_{op}\, v_{\lambda_i;\eta_\alpha} = \lambda_i\, \eta_\alpha\, v_{\lambda_i;\eta_\alpha} \tag{2.23}$$

In the same way

$$\mathcal{O}^{(1)}_{op}\, \mathcal{O}^{(2)}_{op}\, v_{\lambda_i;\eta_\alpha} = \eta_\alpha\, \mathcal{O}^{(1)}_{op}\, v_{\lambda_i;\eta_\alpha} = \eta_\alpha\, \lambda_i\, v_{\lambda_i;\eta_\alpha} \tag{2.24}$$

In conclusion

$$\left[\mathcal{O}^{(1)}_{op}, \mathcal{O}^{(2)}_{op}\right] v_{\lambda_i;\eta_\alpha} = 0 \tag{2.25}$$

Since any $v$ is a linear combination of $v_{\lambda_i;\eta_\alpha}$'s, by linearity

$$\left[\mathcal{O}^{(1)}_{op}, \mathcal{O}^{(2)}_{op}\right] v = 0 \tag{2.26}$$

for any $v$.

It is less obvious, but it is a central theorem of linear algebra, that the converse statement of (2.21) is true

**Theorem**: If $\left[\mathcal{O}^{(1)}_{op}, \mathcal{O}^{(2)}_{op}\right] = 0$, $\mathcal{O}$ and $\mathcal{O}'$ are compatible.

**Problem 8.** Prove the theorem above in the case when $\mathcal{O}$ and $\mathcal{O}'$ are non-degenerate.

This theorem is not immediate to prove in the degenerate case. The proof, for which we refer to the bibliography, makes use of the following lemma, which is important *per se* and which will be used repeatedly in this course:

**Lemma**: Suppose $\mathcal{O}_{op}$ and $\mathcal{O}'_{op}$ commute. Then the action of $\mathcal{O}'_{op}$ on an eigenvector of $\mathcal{O}_{op}$ produces an eigenvector of $\mathcal{O}_{op}$ with the same eigenvalue.

The proof of the lemma is obvious: suppose

$$\mathcal{O}_{op}\, v = \lambda\, v \tag{2.27}$$

Then

$$\mathcal{O}_{op}\left(\mathcal{O}'_{op}\, v\right) = \mathcal{O}'_{op}\, \mathcal{O}_{op}\, v = \mathcal{O}'_{op}\, \lambda\, v = \lambda\left(\mathcal{O}'_{op}\, v\right) \tag{2.28}$$

We can extend the notion of compatible observables to the case of more than two observables: we will say that the observables $\mathcal{O}, \mathcal{O}', \mathcal{O}'' \ldots$ are compatible if there exists a basis of simultaneous eigenvectors. It is easy to see that the operators corresponding to compatible observables all commute among themselves.

We can also introduce the concept of *complete* set of compatible observables. This will be a set of compatible observables $\mathcal{O}, \mathcal{O}', \mathcal{O}'' \ldots$ whose corresponding basis of simultaneous eigenvectors is non-degenerate. In other words a state on which all the observables $\mathcal{O}, \mathcal{O}', \mathcal{O}'' \ldots$ of a complete set have definite values is unique.

## 2.6 Uncertainty relations

Given an observable $\mathcal{O}$ and a state $v$ we saw that the average value of $\mathcal{O}$ on that state is

$$\langle \mathcal{O} \rangle \equiv \langle v, \mathcal{O}_{op}\, v \rangle \qquad |v|^2 = 1 \tag{2.29}$$

We can introduce the average of the square of the difference between $\mathcal{O}$ and its mean value

$$\Delta\, \mathcal{O} \equiv \langle v, \left(\mathcal{O}_{op} - \langle \mathcal{O}_{op} \rangle\right)^2 v \rangle \qquad |v|^2 = 1 \tag{2.30}$$

as a measure of the uncertainty of $\mathcal{O}$ on the state $v$.

**Problem 9.** (**Uncertainty relations**): Let $\Delta\mathcal{O}$ and $\Delta\mathcal{O}'$ the uncertainties of two observables $\mathcal{O}$ and $\mathcal{O}'$ on a given state. Then

$$\Delta\mathcal{O}\, \Delta\mathcal{O}' \geq \frac{1}{2} \langle v, i \left[\mathcal{O}_{op}, \mathcal{O}'_{op}\right] v \rangle \tag{2.31}$$

## 2.7 Canonical relations

Recall Heisenberg uncertainty relation for an unidimensional particle of coordinate $x$ and momentum $p$:

$$\Delta x\, \Delta p \geq \frac{\hbar}{2} \tag{2.32}$$

This relation states that the observables that in quantum mechanics correspond to position and momentum are not compatible. Therefore the corresponding operators $x_{op}$ and $p_{op}$ do not commute. The relation (2.31) suggests to *postulate* that their commutator is

$$\big[x_{op}, p_{op}\big] = i\,\hbar\,\mathbb{I} \tag{2.33}$$

where $\mathbb{I}$ is the identity operator on the state space $\mathcal{H}$. Because of Eq. (2.31), the *canonical commutation relation* (2.33) implies Heisenberg uncertainty principle (2.32).

For a particle in 3-dimension Eqs. (2.33) generalize as follows

$$\big[x^i_{op}, p^j_{op}\big] = i\,\hbar\,\delta^{ij}\,\mathbb{I} \qquad \big[x^i_{op}, x^j_{op}\big] = \big[p^i_{op}, p^j_{op}\big] = 0 \quad i, j, = 1, 2, 3 \tag{2.34}$$

In the following, to lighten the notation, we will write $\hat{x}$ and $\hat{p}$ for $x_{op}$ and $p_{op}$, or simply, when there is no ambiguity, $x$ and $p$.

The canonical commutation relations (2.34) state that coordinates and momenta of a particle are not compatible observables but that both the three coordinates $x^i$ and the three momenta $p^i$ are compatible among themselves. It is natural to ask if either of them form a complete set of observables. In other words one would like to know if, by specifying the three coordinates $x^i$ (or, alternatively, the three components $p^i$ of the momentum) of a *non-relativistic* particle, the corresponding physical state is completely determined. This is an *empirical* question. One infers from experiments that this is so only for special kind of particles (which are called *spinless*). It turns out that to fully determine the physical state of other kinds of particles — as the electrons, the protons, etc. — it is necessary to specify the value of additional compatible observables beyond coordinates (or momenta) These extra-observables are called *spin* variables. We will not discuss them until later in the course. For the time being we will assume that spin variables can be neglected and that either the coordinates or the momenta form a complete set of observables for a non-relativistic particle.

## 2.8 Representations

For our purposes, a representation is a choice of a complete set of compatible observables and of the associated basis. Let us denote such basis with $|\alpha\rangle$[10] where $\alpha$ is a collective index belonging to a set $I$, and let us suppose that this basis is orthonormal.

$$\langle\alpha|\beta\rangle = \delta_{\alpha\beta} \tag{2.35}$$

---

[10] In this section we will use the Dirac notation for vectors and hermitian products.

Any state vector $|v\rangle$ can be decomposed as follows

$$|v\rangle = \sum_{\alpha \in I} |\alpha\rangle \langle\alpha|v\rangle \tag{2.36}$$

We will say that, in this representation, the sequence of complex numbers $\{\langle\alpha|v\rangle, \alpha \in I\}$ *represents* $v$

$$|v\rangle \to \{\langle\alpha|v\rangle\} \tag{2.37}$$

Given two vectors $|v\rangle$ and $|w\rangle$ represented by the two sequences

$$|v\rangle \to \{\langle\alpha|v\rangle\} \qquad |w\rangle \to \{\langle\alpha|w\rangle\}, \tag{2.38}$$

their hermitian product can be written in terms of the same sequences

$$\langle w|v\rangle = \sum_{\alpha} \big(\langle\alpha|v\rangle\big)^* \langle\alpha|w\rangle \tag{2.39}$$

In particular the norm of a vector is

$$|v|^2 = \langle v|v\rangle = \sum_{\alpha} \big|\langle\alpha|v\rangle\big|^2 \tag{2.40}$$

An operator $\mathcal{O}$ acting on $v$ will produce another vector, which also can be decomposed in the same basis

$$\mathcal{O}\,|v\rangle = \sum_{\alpha \in I} \mathcal{O}|\alpha\rangle \langle\alpha|v\rangle = \sum_{\alpha \in I} \sum_{\beta \in I} |\beta\rangle \langle\beta|\mathcal{O}|\alpha\rangle \langle\alpha|v\rangle \tag{2.41}$$

Therefore the representative of the vector $\mathcal{O}\,|v\rangle$ is the sequence of complex numbers

$$\mathcal{O}\,|v\rangle \to \{\sum_{\beta \in I} \langle\alpha|\mathcal{O}|\beta\rangle \langle\beta|v\rangle\} \tag{2.42}$$

We see that the knowledge of the numbers $\langle\alpha|\mathcal{O}|\beta\rangle$ for all $|\alpha\rangle$ and $|\beta\rangle$ is sufficient to compute the action of $\mathcal{O}$ on any vector.

We will say that $\mathcal{O}$ is *represented* in the chosen representation by the *matrix*

$$\mathcal{O} \to \{\langle\beta|\mathcal{O}|\alpha\rangle\} \tag{2.43}$$

Consider now any of the observables $\mathcal{O}_D$ belonging in the complete set of compatible observables which defines our representation. By definition, the $|\alpha\rangle$'s are eigenvectors of $\mathcal{O}_D$

$$\mathcal{O}_D\,|\alpha\rangle = \lambda_D(\alpha)\,|\alpha\rangle \qquad \lambda_D(\alpha) \in \mathbb{R} \tag{2.44}$$

where $\lambda_D(\alpha)$ are the eigenvalues of $\mathcal{O}_D$[11]. Therefore the matrix which represents the observable $\mathcal{O}_D$ is *diagonal*

$$\mathcal{O}_D \to \{\langle\beta|\mathcal{O}_D|\alpha\rangle\} = \lambda_D(\alpha)\,\delta_{\alpha\beta} \tag{2.45}$$

Hence the action of $\mathcal{O}_D$ on a generic state represented by the sequence $\{\langle\alpha|v\rangle\}$ produces a vector represented by the sequence

$$\mathcal{O}_D|v\rangle \to \{\lambda_D(\alpha)\,\langle\alpha|v\rangle\} \tag{2.46}$$

We will refer therefore to such a representation as the representation which *diagonalizes* all the operators $\mathcal{O}_D$ of the complete set of compatible observables.

**Problem 10.** A system is described by a 2-dimensional space of states. Let $|1\rangle$ and $|2\rangle$ be the normalized eigenstates of an observable $\mathcal{O}$ with eigenvalues $\lambda = 1$ and $\lambda = 2$ respectively. Let $v = |1\rangle + i\sqrt{3}|2\rangle$ be a state of the system.

a) What is the probability that the measure of $\mathcal{O}$ on $v$ gives the result $\lambda = 2$?

b) Write the $2 \times 2$ matrix which represents $\mathcal{O}$ in the basis $\{|1\rangle, |2\rangle\}$.

c) Let $\mathcal{O}'$ be another observable, such that $\mathcal{O}'|1\rangle = |2\rangle$ and $\mathcal{O}'|2\rangle = |1\rangle$. Write the $2 \times 2$ matrix which represents $\mathcal{O}'$ in the basis $\{|1\rangle, |2\rangle\}$.

d) What are the eigenvalues of $\mathcal{O}'$?

e) What is the average of $\mathcal{O}'$ on the state $v$?

### 2.8.1 Schrödinger representation

Let us take as complete set of compatible observables of a *spinless non-relativistic particle* the cartesian coordinates $x^i$, with $i = 1, 2, 3$. The associated representation is called the *Schrödinger* representation.

On physical grounds, we expect that the eigenvalues of each coordinate observable $x^i$ — i.e. the positions in space where the particle can be found — be (all) the real numbers. Let us denote by $\hat{x}^i$ the operator corresponding

---

[11] We are using the same notation for the observable and the corresponding operator.

to the $i$-th cartesian coordinate and by $x^i$ the real number representing one of its eigenvalues. The basis $\{|x_1, x_2, x_3\rangle\}$ of simultaneous eigenvectors of $\hat{x}^i$ satisfies

$$\hat{x}^i\,|x_1, x_2, x_3\rangle = x^i\,|x_1, x_2, x_3\rangle \tag{2.47}$$

Therefore a generic state vector $|v\rangle$ will be represented by the collection of complex numbers

$$|v\rangle \to \{\langle x_1, x_2, x_3|v\rangle\} \equiv v(\vec{x}) \tag{2.48}$$

or, in other words, by a *complex valued function* of the 3 coordinates. This function is called the *wave function* corresponding to the state $v$.

The hermitian product of two states $|v\rangle$ and $|w\rangle$ represented by two functions $v(\vec{x})$ and $w(\vec{x})$ is, according to (2.39),

$$\begin{aligned}\langle w|v\rangle = \int dx_1\,dx_2\,dx_3\,w^*(x_1, x_2, x_3)\,v(x_1, x_2, x_3) \equiv \\ \equiv \int_{\mathbb{R}^3} d^3\vec{x}\,w^*(\vec{x})\,v(\vec{x})\end{aligned} \tag{2.49}$$

where we replaced the sum of $\alpha$ for an integral over the 3-coordinates. In particular the norm of a state vector whose wave function is $v(\vec{x})$ is

$$|v|^2 = \int_{\mathbb{R}^3} d^3\vec{x}\,|v(\vec{x})|^2 \tag{2.50}$$

Let us consider now the action of the operator $\hat{x}^i$ on the state represented by the wave function $v(\vec{x})$:

$$\hat{x}^i|v\rangle = \int_{\mathbb{R}^3} d^3\vec{x}\,\hat{x}^i|x_1, x_2, x_3\rangle\,\langle x_1, x_2, x_3|v\rangle = \int_{\mathbb{R}^3} d^3\vec{x}|x_1, x_2, x_3\rangle\,x^i\,v(\vec{x}) \tag{2.51}$$

Hence $\hat{x}^i|v\rangle$ is *represented* in the Schrödinger representation by the function

$$\hat{x}^i|v\rangle \to x^i\,v(\vec{x}) \tag{2.52}$$

Let us now try to find the representative $\hat{p}^i_S$ of the momenta $p^i$ in the Schrödinger representation

$$\hat{p}^i|v\rangle \to \hat{p}^i_S\left(v(\vec{x})\right) \tag{2.53}$$

The canonical commutation relations (2.34) require that

$$\hat{p}^i_S\left(x^j\,v(\vec{x})\right) - x^j\left(\hat{p}^i_S\left(v(\vec{x})\right)\right) = -i\,\hbar\,\delta^{ij}\,v(\vec{x}) \tag{2.54}$$

It is immediate to verify that a linear operator $\hat{p}_S$ which satisfies this equation is

$$\hat{p}^i_S = -i\,\hbar\,\frac{\partial}{\partial x^i} \tag{2.55}$$

**Problem 12.** Show that $\hat{p}^i_S$ is hermitian with respect to the hermitian product (2.49) of the Schrödinger representation if wave functions $v(\vec{x})$ vanish at $|\vec{x}| \to \infty$.

Eq. (2.55) is not the only solution of the equation (2.54) but one can show that all other solutions are physically equivalent.

Summarizing, in the Schrödinger representation, states are represented by wave functions $v(\vec{x})$ which are normalizable in the norm (2.50) and the action on $\hat{x}^i$ and $\hat{p}^i$ on them is

$$\hat{x}^i_S\,v(\vec{x}) = x^i \qquad \hat{p}^i_S\,v(\vec{x}) = -i\,\hbar\frac{\partial\,v(\vec{x})}{\partial x^i} \tag{2.56}$$

Let us now look for the wave-functions of the eigenstates of momenta in the Schrödinger representation: these wave-functions satisfy

$$\hat{p}^i_S\,v_{\vec{p}}(\vec{x}) = p^i\,v_{\vec{p}}(\vec{x}) \tag{2.57}$$

that is

$$-i\,\hbar\frac{\partial\,v_{\vec{p}}(\vec{x})}{\partial x^i} = p^i\,v_{\vec{p}}(\vec{x}) \tag{2.58}$$

The solutions of these differential equations are

$$v_{\vec{p}}(\vec{x}) = C_{\vec{p}}\,\mathrm{e}^{\frac{i}{\hbar}\,\vec{p}\cdot\vec{x}} \tag{2.59}$$

We see therefore that states with definite momentum $\vec{p}$ are represented in the Schrödinger representation by plane waves with wavelength $\lambda$

$$\lambda = \frac{2\,\pi\,\hbar}{|\vec{p}|} = \frac{h}{|\vec{p}|} \tag{2.60}$$

We recover in this way De Broglie relation (1.202)

### 2.8.2 Momentum representation

One can consider the representation which diagonalizes the momenta $\hat{p}^i$, which is defined by a relation analogous to (2.47)

$$\hat{p}^i \,|p_1, p_2, p_3\rangle = p^i \,|p_1, p_2, p_3\rangle \tag{2.61}$$

Therefore states in the momentum representation are represented by functions of the momenta

$$|v\rangle \to \{\langle p_1, p_2, p_3|v\rangle\} \equiv \tilde{v}(\vec{p}) \tag{2.62}$$

and the hermitian product becomes

$$\langle w|v\rangle = \int_{\mathbb{R}^3} d^3\vec{p}\, \tilde{w}^*(\vec{p})\, \tilde{v}(\vec{p}) \tag{2.63}$$

In this representation the momentum is diagonal

$$\hat{p}^i|v\rangle \to \hat{p}^i_{mom}\, \tilde{v}(\vec{p}) = p^i\, \tilde{v}(\vec{p}) \tag{2.64}$$

and the coordinates act as derivatives

$$\hat{x}^i|v\rangle \to \hat{x}^i_{mom}\, \tilde{v}(\vec{p}) = i\,\hbar\, \frac{\partial \tilde{v}(\vec{p})}{\partial p^i} \tag{2.65}$$

**Problem 13**. Show that the eigenstates of the position operators with eigenvalues $\vec{x}$ are represented in the momentum representation by wave functions $\tilde{v}_{\vec{x}}(\vec{p}) = \tilde{C}_{\vec{x}}\, \mathrm{e}^{-\frac{i}{\hbar}\vec{x}\cdot\vec{p}}$.

### 2.8.3 Unitary transformations

In our description, a change of representation is a change of basis. Let $\{|\alpha\rangle\}$ and $\{|a\rangle\}$ be two basis and $|v\rangle$ a generic vector

$$\begin{aligned} |v\rangle &= \sum_\alpha |\alpha\rangle\, \langle\alpha|v\rangle = \sum_a |a\rangle\, \langle a|v\rangle = \\ &= \sum_a \sum_\alpha |\alpha\rangle\, \langle\alpha|a\rangle\, \langle a|v\rangle \end{aligned} \tag{2.66}$$

The representatives of the same vector in the two basis are therefore related by

$$\langle\alpha|v\rangle = \sum_a \langle\alpha|a\rangle \, \langle a|v\rangle \tag{2.67}$$

The matrix

$$U \equiv \{\langle\alpha|a\rangle\} \tag{2.68}$$

allows one to compute the change of representative. Since

$$\sum_a \langle\alpha|a\rangle \, \langle a|\beta\rangle = \langle\alpha|\beta\rangle = \delta_{\alpha\beta} \tag{2.69}$$

and

$$\langle a|\beta\rangle = (\langle\beta|a\rangle)^* \tag{2.70}$$

the matrix $U$ is unitary

$$\sum_a U_{\alpha a} \, U^*_{\beta a} = \sum_a U_{\alpha a} \, (U^\dagger)_{a\beta} = \delta_{\alpha\beta} \tag{2.71}$$

**Problem 11.** On a spin 1/2 system the following 3 observables are defined

$$S_z = \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \qquad S_x = \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \qquad S_y = \frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} \tag{2.72}$$

a) Find the eigenvectors and eigenvalues of the 3 observables $\{S_x, S_y, S_z\}$.

b) Suppose the system is in the state described by the vector

$$v = \begin{pmatrix}1\\ 3\,i\end{pmatrix} \tag{2.73}$$

Find the probabilities to obtain the various possible values of $\{S_x, S_y, S_z\}$ on $v$. (You should in other words compute 6 probabilities, 2 for each of the 3 observables.)

c) Compute the 2x2 unitary matrix which connects the basis of eigenvectors of $S_z$ with the basis of eigenvectors of $S_y$.

d) Find the eigenvectors and eigenvalues of the observable

$$S_{\vec{n}} \equiv n_x \, S_x + n_y \, S_y + n_z \, S_z \tag{2.74}$$

where

$$\vec{n} = (n_x, n_y, n_z) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \tag{2.75}$$

$\theta$ and $\phi$ are polar coordinates: $0 \leq \phi \leq 2\pi$ and $0 \leq \theta \leq \pi$.

e) Compute the 2x2 unitary matrix which connects the basis of eigenvectors of $S_z$ with the basis of eigenvectors of $S_{\vec{n}}$.

### 2.8.4 Generalized observables and wave packets

A problem with both the Schrödinger and the momentum representation is the fact that the eigenvalues of the operators which define the representation form a set which is not denumerable but continuous. The consequence of this is that the corresponding eigenvectors are not normalizable.

To see this, consider, for example, the momentum eigenvectors (2.59) in the Schrödinger representations. Since

$$|v_{\vec{p}}|^2 = |C_{\vec{p}}|^2 \tag{2.76}$$

the norm of $v_{\vec{p}}(\vec{x})$ is a non-convergent integral, and therefore it is not well defined.

Strictly speaking therefore, *neither* the coordinates *nor* the momenta operators are observables: they do not have normalizable eigenvectors.

This is not a problem *per se*. Even if we assumed the existence of a basis of eigenvectors to construct Schrödinger and momentum representations, the representations can be defined without making reference to such basis. We can *define* the Schrödinger representation as the representation in which the linear space of states is the space of wave functions $v(\vec{x})$ normalizable with respect to the norm (2.50), on which coordinates and momenta act as in (2.56). We can define the momentum representation in an analogous way, by means of Eqs. (2.63) and (2.64) and (2.65), with no reference to the continuous basis $|\vec{p}\rangle$. One can construct out of $\hat{x}^i$ and $\hat{p}^j$ operators which do have a complete set of normalizable eigenvectors and thus are genuine observables.

The physical reason why coordinates or momenta (or any other physical variable with a *continuous spectrum*) are not genuine observables is that no physical device can measure coordinates (or momenta) with infinite precision. Physical devices can measure coordinates (or momenta) with finite precision. Since this precision can in principle be arbitrarily improved, it is possible to "approximate" the operators $\hat{x}^i$ arbitrarily well by means of genuine observables which correspond to physically realizable measuring instruments.

From the mathematical point of view, it is convenient to include in our formalism also those operators which, although not physical, can nevertheless be approximated with arbitrarily good precision by genuine observables. We will call such operators and their eigenvectors *improper* or *generalized* observables and eigenvectors.

To understand the kind of mathematics which is needed to include generalized observables in our formalism, consider an orthonormal basis

$$\langle\alpha|\beta\rangle = \delta_{\alpha\beta} \tag{2.77}$$

and the relation which holds for any vector $|v\rangle$

$$\langle\alpha|v\rangle = \sum_{\beta}\langle\alpha|\beta\rangle\,\langle\beta|v\rangle \tag{2.78}$$

In the case of the Schrödinger representations the relation above should write as follows

$$v(\vec{x}) = \int d^3\vec{y}\,\langle\vec{x}|\vec{y}\rangle\,v(\vec{y}) \tag{2.79}$$

Now it must be that

$$\langle\vec{x}|\vec{y}\rangle = 0 \qquad \text{for} \quad \vec{x} \neq \vec{y} \tag{2.80}$$

since eigenvectors corresponding to different eigenvalues are orthogonal. But then, whatever finite value $\langle\vec{x}|\vec{y}\rangle$ takes for $\vec{x} = \vec{y}$, the integral in the r.h.s. of (2.79) would vanish, if $\langle\vec{x}|\vec{y}\rangle$ were any "reasonable" function. In some non-rigorous sense, the only way in which both (2.79) and (2.80) could hold at the same time, is if $\vec{x} = \vec{y}$ becomes infinite for $x = y$. This again is a manifestation of the fact that $|x\rangle$ is not a normalizable state. One should not therefore think of $\langle\vec{x}|\vec{y}\rangle$ as an usual function but rather as an object that can be multiplied by a normalizable function and integrated: in other words we will take (2.79) as the *definition* of what $\langle\vec{x}|\vec{y}\rangle$ is.

We will write, following Dirac,

$$\langle\vec{x}|\vec{y}\rangle = \delta(\vec{x}-\vec{y}) \tag{2.81}$$

where $\delta(x-y)$ is *defined* by the following property

$$f(\vec{x}) \equiv \int d^3\vec{y}\,\delta(\vec{x}-\vec{y})\,f(\vec{y}) \tag{2.82}$$

for any normalizable function $f(x)$. Objects like these are called *distributions.*

In the same way we will write

$$\langle\vec{q}|\vec{p}\rangle = \delta(\vec{q}-\vec{p}) \tag{2.83}$$

for the basis of improper eigenstates $|\vec{p}\rangle$ of the momenta.

Let us now consider again the wave functions of the momentum eigenstates in the Schrödinger representation

$$v_{\vec{p}}(\vec{x}) = \langle\vec{x}|\vec{p}\rangle = C_{\vec{p}}\,\mathrm{e}^{\frac{i}{\hbar}\,\vec{p}\cdot\vec{x}} \tag{2.84}$$

and the wave functions of the position eigenstates in the momentum representation

$$\tilde{v}_{\vec{x}}(\vec{p}) = \langle\vec{p}|\vec{x}\rangle = C_{\vec{x}}\,\mathrm{e}^{-\frac{i}{\hbar}\,\vec{p}\cdot\vec{x}} \tag{2.85}$$

Since

$$\langle\vec{x}|\vec{p}\rangle = \langle\vec{p}|\vec{x}\rangle^* \tag{2.86}$$

we have

$$C_{\vec{p}} = C^*_{\vec{x}} \tag{2.87}$$

The l.h.s. of the equation above is a function of $\vec{p}$ while the r.h.s. is a function of $\vec{x}$: such an equation is satisfied only if this function is a constant independent of both $\vec{p}$ and $\vec{x}$

$$C = C_{\vec{p}} = C^*_{\vec{x}} \tag{2.88}$$

Then

$$\begin{aligned} v_{\vec{p}}(\vec{x}) &= C\,\mathrm{e}^{\frac{i}{\hbar}\,\vec{p}\cdot\vec{x}} \\ \tilde{v}_{\vec{x}}(\vec{p}) &= C^*\mathrm{e}^{-\frac{i}{\hbar}\,\vec{p}\cdot\vec{x}} \end{aligned} \tag{2.89}$$

We want to determine the normalization coefficient $C$, by demanding that (2.83) holds:

$$\langle \vec{q} | \vec{p} \rangle = \int_{\mathbb{R}^3} d^3\vec{x}\, v^*_{\vec{q}}(\vec{x})\, v_{\vec{p}}\,(\vec{x}) = |C|^2 \int_{\mathbb{R}^3} d^3\vec{x}\, \mathrm{e}^{\frac{i}{\hbar}(\vec{p}-\vec{q})\cdot\vec{x}} = \delta(\vec{q}-\vec{p}) \tag{2.90}$$

The integral which appears in this equation is not a convergent integral and the result of the integration is not a function in the usual sense: it is a distribution. As we said above, the meaning of equation (2.90) is made explicit by multiplying both sides of it by an arbitrary function $\tilde{v}(\vec{p})$ and integrating in $\vec{p}$

$$\tilde{v}(\vec{q}) = \int_{\mathbb{R}^3} d^3\vec{p}\, \tilde{v}(\vec{p})\, \delta(\vec{q}-\vec{p}) = |C|^2 \int_{\mathbb{R}^3} d^3\vec{p}\, \tilde{v}(\vec{p}) \int_{\mathbb{R}^3} d^3\vec{x}\, \mathrm{e}^{\frac{i}{\hbar}(\vec{p}-\vec{q})\cdot\vec{x}} \tag{2.91}$$

We must determine the value of $|C|^2$ which ensures the validity of this equation. Since the equation above must hold for any function $\tilde{v}(p)$, if we assume that such a $|C|^2$ exists,[12] then we can find it by computing both sides of the equation for a specific choice for $\tilde{v}(p)$. Let us choose, for example

$$\tilde{v}(\vec{p}) = \mathrm{e}^{-\alpha\,\vec{p}^2} \qquad \alpha > 0 \tag{2.98}$$

If $\alpha > 0$ this function is normalizable. Then the integral which appears on

---

[12] A mathematically more rigorous derivation of the value of $|C|^2$, which does not assume the validity of (2.91) for some $C$ but derives it, requires some results from the theory of Fourier transforms. Here I give a condensed account of these results. The Fourier transform $F(f)$ of a normalizable function $f(k)$ is defined to be

$$F(f)(x) \equiv \int_{-\infty}^{\infty} dk\, \mathrm{e}^{i\,k\,x}\, f(k) \tag{2.92}$$

The central propositions regarding Fourier transforms are the following:

- If $f(k)$ is normalizable $F(f)$ is normalizable too.
- If $F(f)(x)$ is the Fourier transform of $f(k)$ then

$$2\,\pi\, f(k) = \int_{-\infty}^{\infty} dx\, \mathrm{e}^{-i\,k\,x}\, F(f)(x) \tag{2.93}$$

Plugging (2.92) into (2.93), and changing the order of integration, one obtains

$$\begin{aligned} 2\,\pi\, f(k) &= \int_{-\infty}^{\infty} dx\, \mathrm{e}^{-i\,k\,x} \int_{-\infty}^{\infty} dk'\, \mathrm{e}^{i\,k'\,x}\, f(k') = \\ &= \int_{-\infty}^{\infty} dk'\, f(k') \int_{-\infty}^{\infty} dx\, \mathrm{e}^{-i\,k\,x}\, \mathrm{e}^{i\,k'\,x} \end{aligned} \tag{2.94}$$

Recalling the definition (2.82) of the Dirac delta function we conclude that

$$\int_{-\infty}^{\infty} dx\, \mathrm{e}^{-i\,k\,x}\, \mathrm{e}^{i\,k'\,x} = 2\,\pi\,\delta(k-k') \tag{2.95}$$

This result is immediately generalizable to three dimensions

$$\int_{\mathbb{R}^3} d^3\vec{x}\, \mathrm{e}^{-\frac{i}{\hbar}\,\vec{q}\cdot\vec{x}} \mathrm{e}^{\frac{i}{\hbar}\,\vec{p}\cdot\vec{x}} = (2\,\pi\,\hbar)^3\,\delta(\vec{p}-\vec{q}) \tag{2.96}$$

Comparing with (2.90) we conclude

$$|C|^2 = \frac{1}{(2\,\pi\,\hbar)^3} \tag{2.97}$$

the r.h.s. of (2.91) writes, after changing the order of the integrations

$$\begin{aligned}
&\int_{\mathbb{R}^3} d^3\vec{p}\,\mathrm{e}^{-\alpha\,\vec{p}^2} \int_{\mathbb{R}^3} d^3\vec{x}\,\mathrm{e}^{\frac{i}{\hbar}(\vec{p}-\vec{q})\cdot\vec{x}} = \int_{\mathbb{R}^3} d^3\vec{x}\,\mathrm{e}^{-\frac{i}{\hbar}\,\vec{q}\cdot\vec{x}} \int_{\mathbb{R}^3} d^3\vec{p}\,\mathrm{e}^{-\alpha\,\vec{p}^2}\,\mathrm{e}^{\frac{i}{\hbar}\,\vec{p}\cdot\vec{x}} = \\
&\quad = \int_{\mathbb{R}^3} d^3\vec{x}\,\mathrm{e}^{-\frac{i}{\hbar}\,\vec{q}\cdot\vec{x}} \int_{\mathbb{R}^3} d^3\vec{p}\,\mathrm{e}^{-\alpha\,(\vec{p}-\frac{i\,\vec{x}}{2\,\alpha\,\hbar})^2}\,\mathrm{e}^{-\frac{\vec{x}^2}{4\,\alpha\,\hbar^2}} = \\
&\quad = \int_{\mathbb{R}^3} d^3\vec{x}\,\mathrm{e}^{-\frac{i}{\hbar}\,\vec{q}\cdot\vec{x}} \int_{\mathbb{R}^3} d^3\vec{p}\,\mathrm{e}^{-\alpha\,\vec{p}^2}\,\mathrm{e}^{-\frac{\vec{x}^2}{4\,\alpha\,\hbar^2}} = \\
&\quad = \left(\frac{\pi}{\alpha}\right)^{\frac{3}{2}} \int_{\mathbb{R}^3} d^3\vec{x}\,\mathrm{e}^{-\frac{i}{\hbar}\,\vec{q}\cdot\vec{x}}\,\mathrm{e}^{-\frac{\vec{x}^2}{4\,\alpha\,\hbar^2}} = \left(\frac{\pi}{\alpha}\right)^{\frac{3}{2}} \int_{\mathbb{R}^3} d^3\vec{x}\,\mathrm{e}^{-\frac{(\vec{x}+2\,i\,\alpha\,\hbar\,\vec{q})^2}{4\,\alpha\,\hbar^2}}\,\mathrm{e}^{-\alpha\,\vec{q}^2} = \\
&\quad = \left(\frac{\pi}{\alpha}\right)^{\frac{3}{2}} \int_{\mathbb{R}^3} d^3\vec{x}\,\mathrm{e}^{-\frac{\vec{x}^2}{4\,\alpha\,\hbar^2}}\,\mathrm{e}^{-\alpha\,\vec{q}^2} = \left(\frac{\pi}{\alpha}\right)^{\frac{3}{2}} \left(4\,\alpha\,\hbar^2\,\pi\right)^{\frac{3}{2}}\,\mathrm{e}^{-\alpha\,\vec{q}^2} = \\
&\quad = (2\,\pi\,\hbar)^3\,\mathrm{e}^{-\alpha\,\vec{q}^2} = (2\,\pi\,\hbar)^3\,\tilde{v}(\vec{q}) \qquad (2.99)
\end{aligned}$$

where we have repeatedly used the formula for the gaussian integral (1.94). Plugging this result into Eq. (2.91) we obtain

$$\tilde{v}(\vec{q}) = |C|^2\,(2\,\pi\,\hbar)^3\,\tilde{v}(\vec{q}) \qquad (2.100)$$

that is

$$|C|^2 = \frac{1}{(2\,\pi\,\hbar)^3} \qquad (2.101)$$

Therefore if we normalize the wave functions of the eigenstates of momentum as follows

$$v_{\vec{p}}(\vec{x}) = \frac{1}{(2\,\pi\,\hbar)^{\frac{3}{2}}}\,\mathrm{e}^{\frac{i}{\hbar}\,\vec{p}\cdot\vec{x}} \qquad (2.102)$$

then indeed

$$\langle\vec{q}|\vec{p}\rangle = \int_{\mathbb{R}^3} d^3\vec{x}\,v^*_{\vec{q}}(\vec{x})\,v_{\vec{p}}\,(\vec{x}) = \delta(\vec{q}-\vec{p}) \qquad (2.103)$$

Let us use the result (2.103) to derive the relation between the Schrödinger and momentum representations. Consider the general formula (2.67) and specify $|\alpha\rangle$ and $|a\rangle$ to be, respectively, the coordinates and momenta (generalized) eigenvectors $|\vec{x}\rangle$ and $|\vec{p}\rangle$:

$$\langle\vec{x}|v\rangle = \int_{\mathbb{R}^3} d^3\vec{p}\,\langle\vec{x}|\vec{p}\rangle\,\langle\vec{p}|v\rangle \qquad (2.104)$$

We have $\langle\vec{x}|v\rangle = v(\vec{x})$, $\langle\vec{p}|v\rangle = \tilde{v}(\vec{p})$ and $\langle\vec{x}|\vec{p}\rangle = v_{\vec{p}}(\vec{x})$. Thus the *same* vector $|v\rangle$ is described in the Schrödinger and momentum representation by wave functions $v(\vec{x})$ and $\tilde{v}(\vec{p})$ related to each other by

$$v(\vec{x}) = \frac{1}{(2\,\pi\,\hbar)^{\frac{3}{2}}}\int_{\mathbb{R}^3} d^3\vec{p}\,\mathrm{e}^{\frac{i}{\hbar}\,\vec{p}\cdot\vec{x}}\,\tilde{v}(\vec{p}) = \int_{\mathbb{R}^3} d^3\vec{p}\,v_{\vec{p}}(\vec{x})\,\tilde{v}(\vec{p}) \tag{2.105}$$

This equation shows that a generic state $|v\rangle$ which is described in the Schrödinger representation by the normalizable wave-function $v(\vec{x})$ can be written as a linear superposition of *non-normalizable* plane waves $v_{\vec{p}}(\vec{x})$. The coefficients $\tilde{v}(\vec{p})$ of this linear superposition define the normalizable wave function of momenta which represents the same state $|v\rangle$ in the momentum representation.

In conclusion, although plane waves $v_{\vec{p}}(\vec{x})$ do not describe physically realizable states, normalizable linear superpositions of plane waves do. This justifies and makes convenient the inclusion in our formalism of generalized, non-normalizable, wave functions such as $v_{\vec{p}}(\vec{x})$. For example a state with definite momentum $\vec{p}_0$ is not physically realizable but it can be approximated by normalizable *packets* which are linear superpositions of plane waves with different momenta laying in a small interval around an average value $\vec{p}_0$

$$v_{\vec{p}_0,\Delta}(\vec{x}) = \int d^3\vec{p}\,v_{\vec{p}}(\vec{x})\,\phi_{\vec{p}_0,\Delta}(\vec{p}) \tag{2.106}$$

where $\phi_{\vec{p}_0,\Delta}(\vec{p})$ is a function which is significantly different than zero only in a small region of size $\Delta$ around the value $\vec{p} = \vec{p}_0$. By taking $\Delta$ arbitrarily small we can approximate the plane wave $v_{\vec{p}_0}(\vec{x})$ arbitrarily well. For example we can take

$$\phi_{\vec{p}_0,\Delta}(\vec{p}) = \frac{1}{(\pi)^{\frac{3}{4}}\,\Delta^{\frac{3}{2}}}\,\mathrm{e}^{-\frac{(\vec{p}-\vec{p}_0)^2}{2\,\Delta^2}} \tag{2.107}$$

**Problem 14.** Show that $v_{\vec{p}_0,\Delta}(\vec{x})$ with the choice (2.107) is a normalizable wave function of unit norm. Compute the average of $\vec{x}$, $\vec{p}$ $\Delta p^2$ and $\Delta x^2$ on the state described by the packet (2.106). Make use of the formula for gaussian integrals

$$\int_{-\infty}^{\infty} dx\,\mathrm{e}^{-\alpha\,x^2+\beta\,x} = \sqrt{\frac{\pi}{\alpha}}\,\mathrm{e}^{\frac{\beta^2}{4\,\alpha}} \tag{2.108}$$

where $\alpha > 0$ and $\beta$ is a generic *complex* number.

**Problem 15.** Let $\mathcal{F}$ denote the map which sends the (one-dimensional) wave function in the Schrödinger representation to the wave function in the momentum representation

$$\mathcal{F} : \psi(x) \to \tilde{\psi}(p) = \int_{-\infty}^{\infty} \frac{dx}{(2\,\pi\,\hbar)^{\frac{1}{2}}}\, \mathrm{e}^{-i\,\frac{p}{\hbar}\,x}\, \psi(x) \tag{2.109}$$

Show that

$$\mathcal{F}\,\mathcal{F}^{\dagger} = \mathbb{I} \qquad \mathcal{F}^{2} = \mathcal{I} \tag{2.110}$$

where $\mathbb{I}$ is the identity operator and $\mathcal{I}$ is the parity operator

$$\mathcal{I} : \psi(x) \to \psi(-x) \tag{2.111}$$

# 3 Simple systems

## 3.1 Free particle on the line

One of the simplest systems one can think of is a free non-relativistic, spinless particle of mass $m$ moving on the real axis. The hamiltonian of this system is

$$\hat{H} = \frac{\hat{p}^2}{2\,m} \tag{3.1}$$

Observe that $\hat{p}$ and $\hat{H}$ commute

$$\left[\hat{p}, \hat{H}\right] = 0 \tag{3.2}$$

Thus there exists a basis of simultaneous eigenvectors of $\hat{p}$ and $\hat{H}$. Moreover, since $\hat{p}$ is a non-degenerate (generalized) observable, its (generalized) eigenvectors $|p\rangle$ form a basis and are, therefore, states with definite energy

$$\hat{H}|p\rangle = \frac{p^2}{2\,m}\,|p\rangle \tag{3.3}$$

Hence the eigenvalues of $\hat{H}$ are

$$E_p = \frac{p^2}{2\,m} \tag{3.4}$$

and they cover the whole positive real semi-axis $E > 0$. Given $E > 0$, what is the eigenspace $\mathcal{H}_E$? Since $|p\rangle$ and $|-p\rangle$ both have the same energy, we have that for $E > 0$, $\mathcal{H}_E$ has dimension 2 and it is given by the linear combinations

$$v_E = \alpha\,|p\rangle + \beta\,|-p\rangle \qquad v_E \in \mathcal{H}_E \tag{3.5}$$

Thus, although all the eigenvectors of $\hat{p}$ are eigenvectors of $\hat{H}$ the reverse is not true.

Note the difference with the classical situation: in the classical case, there are just two states with a given energy $E > 0$, that with $p > 0$ (particle moving to the right) or that with $p < 0$ (particle moving to the left). In the quantum case, there is a whole 2-dimensional space of states with this property. States for which both $\alpha$ and $\beta$ are different than zero are states with definite energy but not with definite momentum: in these states the particle is moving both to the left and to the right.

Let us consider this same problem in the Schrödinger representation. The wave functions $\psi_E(x)$ of eigenvectors of $\hat{H}$ are given by

$$-\frac{\hbar^2}{2\,m}\psi''_E(x) = E\,\psi_E(x) \tag{3.6}$$

The general solution of this equation is

$$\psi_E(x) = \alpha\,\mathrm{e}^{\frac{i}{\hbar}\sqrt{2\,m\,E}\,x} + \beta\,\mathrm{e}^{-\frac{i}{\hbar}\sqrt{2\,m\,E}\,x} \tag{3.7}$$

for *any* $E$ real.

Observe that for $E < 0$ the solution (3.7) is the sum of two real exponentials

$$\psi_E(x) = \alpha\,\mathrm{e}^{\frac{\sqrt{2\,m\,|E|}}{\hbar}\,x} + \beta\,\mathrm{e}^{-\frac{\sqrt{2\,m\,|E|}}{\hbar}\,x} \qquad E < 0 \tag{3.8}$$

Therefore each one of the two exponentials *diverge* either for $x \to \infty$ or for $x \to -\infty$.

For $E > 0$ the solution (3.7) is instead the sum of two trigonometric function and it remains *bounded* for $x \to \infty$. Note that for $E > 0$ the general solution is not normalizable.

Since the analysis in the momentum representation told us that the physical spectrum has $E > 0$, this example suggests the following *general rule*:

*The physically "acceptable" generalized eigenfunctions of a generalized observable with a continuous spectrum must remain bounded for* $x \to \pm\infty$.

## 3.2 Free particle on a segment

Let us consider a free particle

$$\hat{H} = \frac{\hat{p}^2}{2\,m} \tag{3.9}$$

confined to the segment $x \in [0, L]$. Since the particle is confined to the segment, the probability to find it outside is zero. Therefore $\psi(x) = 0$ for $x > L$ and $x < 0$. By continuity

$$\psi(0) = \psi(L) = 0 \tag{3.10}$$

The equation for the eigenfunction of the energy is still Eq. (3.6). Among the solutions (3.7) we must choose those which satisfy the boundary conditions (3.10).

**Problem 1.** Show that the solutions of (3.7) which satisfy (3.10) are

$$\psi_n(x) = \sqrt{\frac{2}{L}}\, \sin(k_n\, x) \tag{3.11}$$

where

$$k_n = \frac{n\,\pi}{L} \tag{3.12}$$

and have energies

$$E_n = \frac{\hbar^2\, k_n^2}{2\, m} = \frac{\hbar^2\, \pi^2\, n^2}{2\, m\, L^2} \qquad n = 1, 2, \ldots \tag{3.13}$$

This system has therefore a discrete energy spectrum (which we had already derived by means of the Bohr-Sommerfeld quantization rules). The energy eigenfunctions are normalizable:

$$\int_0^L \psi_n^*(x)\, \psi_m(x) = \delta_{n,m} \tag{3.14}$$

and form a *basis* of the space of states, i.e. the space of the normalizable wave-functions which satisfy the boundary conditions (3.10). Notice that $\hat{p}$ has *no* eigenvectors on this space and therefore is not an observable, not even a generalized one.

## 3.3 Free particle in 3-dimensional box

The generalization of the previous problem to 3-dimensions is

$$\hat{H} = \frac{\vec{p}^2}{2\, m} \tag{3.15}$$

where the particle is confined in a box $0 \leq x^i \leq L_i$, $i = 1, 2, 3$ of sizes $L_1, L_2, L_3$. The normalizable energy eigenfunctions are

$$\psi_{n_1,n_2,n_3}(\vec{x}) = \prod_{i=1}^{3} \sqrt{\frac{2}{L_i}} \sin(k_{n_i}\, x^i) \qquad n_i = 1, 2, \ldots$$

$$k_{n_i} = \frac{\pi\, n_i}{L_i} \qquad E_{n_1,n_2,n_3} = \frac{\hbar^2\, \pi^2}{2\, m} \big(\frac{n_1^2}{L_1^2} + \frac{n_2^2}{L_2^2} + \frac{n_3^2}{L_3^2}\big) \tag{3.16}$$

**Problem 2.** Suppose $L_1 = L_2 \neq L_3$. Discuss the degeneracy of the energy levels. Answer the same question when $L_1 = L_2 = L_3$.

## 3.4 One-dimensional particle in a potential

A particle moving on the real line in a potential is described by the hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2\, m} + V(\hat{x}) \tag{3.17}$$

The equation for the energy eigenfunctions in the Schrödinger representation is called the *Schrödinger equation*

$$\big(-\frac{\hbar^2}{2\, m} \frac{d^2}{d\, x^2} + V(x)\big)\psi_E(x) = E\, \psi_E(x) \tag{3.18}$$

This is an ordinary linear differential equation of second order: as such, it has, for any $E$, *two* linearly independent solutions. According to our rule (section 3.1), however, not all solutions are physically acceptable. In the following we discuss at qualitative level, with no pretense of mathematical rigor, how the number and the kind of acceptable solutions of the Schrödinger equation depend on the energy $E$ and on the behavior of the potential at infinity. We refer to the bibliography for more detailed analysis.

Classically, the motion is confined to regions of the real axis for which

$$E - V(x) > 0 \qquad \text{classically allowed regions} \equiv \text{Type I} \tag{3.19}$$

Let us call these regions "of Type I". We will call the classically inaccessible regions "of Type II"

$$E - V(x) < 0 \qquad \text{classically forbidden regions} \equiv \text{Type II} \tag{3.20}$$

To understand if a solution is acceptable or not we have to study its behavior for $x \to \pm\infty$. We will consider the cases

- (a) The regions $x \to \pm\infty$ are both Type I
- (b) The regions $x \to \pm\infty$ are both Type II
- (c) The regions $x \to \pm\infty$ are one of Type I and one of Type II

These are *not* all the possible cases. There are situations in which at $x \to \pm\infty$ there is an infinite alternation of regions of Type I and of Type II. This is what happens for example with periodic potentials, an interesting case whose discussion we will postpone to a later section.

Let us remark that since $E$ is real, if $\psi_E(x)$ is a solution of the Schrödinger equation its complex conjugate is a solution of the same equation. Therefore both the real and the imaginary parts of $\psi_E(x)$ are solutions too. Moreover if a solution $\psi_E(x)$ is acceptable then both the real and the imaginary parts are acceptable. Because of this, in the following discussion we can suppose with no loss of generality that the wave functions are *real*.

Let us further observe that

$$\frac{\psi_E''(x)}{\psi_E(x)} = \frac{2\,m}{\hbar^2}\big(V(x) - E\big) \tag{3.21}$$

Therefore:

$$\begin{aligned} \frac{\psi_E''(x)}{\psi_E(x)} &< 0 \qquad \text{in Type I regions} \\ \frac{\psi_E''(x)}{\psi_E(x)} &> 0 \qquad \text{in Type II regions} \end{aligned} \tag{3.22}$$

The ratio $\frac{\psi_E''(x)}{\psi_E(x)}$ is linked to the *concavity* of the graph of the real function $\psi_E(x)$. Specifically:

- In regions of Type I the graph of the wave function has concavity directed towards the real axis: in other words the graph has positive (negative) concavity for negative (positive) values of $\psi_E(x)$.
- In regions of Type II the graph of the wave function has concavity directed away from the real axis: in other words the graph has positive (negative) concavity for positive (negative) values of $\psi_E(x)$.

This means that in the regions of Type I — the classical accessible regions — the wave function has an oscillatory behavior. Because of this, in case (a), if the potential is bounded, both solutions of the Schrödinger equation remains bounded for $|x| \to +\infty$ for any value of $E$. In this case, hence, the energy spectrum is continuous and doubly degenerate. We saw explicitly in section 3.2 that the spectrum of the free particle moving on the real line is indeed of this kind.

Consider now case (b). Because of the concavity property, if the whole real line is of type II, the wave function diverges either for $x \to +\infty$ or for $x \to -\infty$. In this case there are no solutions at all. If instead $E$ is greater than the minimum value of $V(x)$ then there is at least one region of type I in between the asymptotic regions of type II. In the asymptotic regions the wave function can remain bounded only if it tends to zero. For any value of $E$ one can choose a special linear combinations of the two independent solutions that vanishes at one infinity, for example for $x \to -\infty$. For generic values of $E$, however, this solution will diverge for $x \to +\infty$. Only for special values of $E$ it will happen that the linear combination which tends to zero for $x \to -\infty$ will do the same for $x \to +\infty$. We conclude that in this case the spectrum is discrete and non-degenerate.

In case (c), the same reasoning of the previous paragraph, ensures that for any $E$ we can find a special linear combination of solutions which tends to zero in the asymptotic region of type II. This solution will be oscillatory in the other asymptotic region, that of type I. Then, in case (c), the spectrum is continuous and non-degenerate.

## 3.5 Step potential

Let us consider an undimensional particle on the real line in the potential

$$V(x) = \begin{cases} 0 & \text{if } x < 0, \\ V_0 & \text{if } x > 0. \end{cases} \tag{3.23}$$

where $V_0 > 0$. Classical trajectories are of two types. Those with $0 < E < V_0$, which are confined to the negative semi-axis $x < 0$; they correspond to particles coming from the negative spatial infinity, bouncing off the potential step at $x = 0$ and travelling back to $x \to -\infty$ with the same kinetic energy and opposite momentum. The trajectories with $E > V_0$ correspond instead either to particles coming from $x \to -\infty$ with momentum

$$p = \sqrt{2\, m\, E} \tag{3.24}$$

and flying off to $x \to +\infty$ with reduced momentum

$$p' = \sqrt{2\,m\,(E - V_0)} \tag{3.25}$$

or to particles following the inverse trajectories, coming from $x \to +\infty$ with momentum $-p'$ and flying off to $x \to -\infty$ with momentum $-p$.

The discussion in the previous section teaches us that the quantum energy spectrum also must be positive. The general solution of the Schrödinger equation for $E > 0$ is

$$\psi_E(x) = \begin{cases} A\,\mathrm{e}^{\frac{i}{\hbar}\,p\,x} + B\,\mathrm{e}^{-\frac{i}{\hbar}\,p\,x} & \text{if } x < 0 \\ A'\,\mathrm{e}^{\frac{i}{\hbar}\,p'\,x} + B'\,\mathrm{e}^{-\frac{i}{\hbar}\,p'\,x} & \text{if } x > 0 \end{cases} \tag{3.26}$$

where

$$p = \sqrt{2\,m\,E} \qquad p' = \sqrt{2\,m\,(E - V_0)} \tag{3.27}$$

and $A, B, A', B'$ are complex constants.

### 3.5.1 Energy eigenfunctions of discontinuous potentials

Let us now consider the Schrödinger equation

$$-\frac{\hbar^2}{2\,m}\,\psi''_E(x) + V(x)\,\psi(x) = E\,\psi(x) \tag{3.28}$$

and let us integrate both sides of it on a small interval of length $\epsilon > 0$ centered around $x = 0$

$$-\frac{\hbar^2}{2\,m}\int_{-\frac{\epsilon}{2}}^{\frac{\epsilon}{2}} dx\,\psi''_E(x) = \int_{-\frac{\epsilon}{2}}^{\frac{\epsilon}{2}} dx\,\big(E - V(x)\big)\,\psi_E(x) \tag{3.29}$$

Since the potential remains bounded across the discontinuity at $x = 0$, we have

$$-\frac{\hbar^2}{2\,m}\,\big(\psi'_E(\frac{\epsilon}{2}) - \psi'_E(-\frac{\epsilon}{2})\big) = \epsilon\,\big(E - V(0)\big)\,\psi_E(0) + O(\epsilon^2) \tag{3.30}$$

Taking the limit for $\epsilon \to 0^+$ we conclude that

$$\lim_{\epsilon \to 0^+}\big(\psi'_E(\frac{\epsilon}{2}) - \psi'_E(-\frac{\epsilon}{2})\big) = 0 \tag{3.31}$$

This means that the first derivative of the wave function is *continuous* at $x = 0$. This also implies, *a fortiori*, that the wave function $\psi_E(x)$ itself is continuous in $x = 0$.

### 3.5.2 Energy eigenfunctions with $0 < E < V_0$

In this case $p'$ is purely imaginary

$$p' = i\,|p'| = i\,\sqrt{2\,m\,(V_0 - E)} \tag{3.32}$$

The request that the solution remains bounded for large $x$ forces us to take

$$B' = 0 \tag{3.33}$$

Therefore the solution of the Schrödinger equation becomes

$$\psi_E(x) = \begin{cases} A\,\mathrm{e}^{\frac{i}{\hbar}\,p\,x} + B\,\mathrm{e}^{-\frac{i}{\hbar}\,p\,x} & \text{if } x < 0, \\ A'\,\mathrm{e}^{-\frac{1}{\hbar}\,|p'|\,x} & \text{if } x > 0. \end{cases} \tag{3.34}$$

Imposing these conditions on the solution (3.34) we obtain

$$A + B = A' \qquad \frac{i}{\hbar}\,p\,(A - B) = -\frac{1}{\hbar}\,|p'|\,A' \tag{3.35}$$

from which

$$\begin{aligned} B &= \frac{p - i\,|p'|}{p + i\,|p'|}\,A = \mathrm{e}^{-2\,i\,\theta_E}\,A \qquad \theta_E \equiv \arctan\frac{|p'|}{p} = \sqrt{\frac{V_0}{E} - 1} \\ A' &= \frac{2\,p}{p + i\,|p'|}\,A \end{aligned} \tag{3.36}$$

and

$$\psi_E(x) = A \times \begin{cases} \mathrm{e}^{\frac{i}{\hbar}\,p\,x} + \mathrm{e}^{-2\,i\,\theta_E}\,\mathrm{e}^{-\frac{i}{\hbar}\,p\,x} & \text{if } x < 0, \\ \frac{2\,p}{p + i\,|p'|}\,\mathrm{e}^{-\frac{1}{\hbar}\,|p'|\,x} & \text{if } x > 0 \end{cases} \tag{3.37}$$

Without loss of generality we can choose the normalization of this eigenfunction such that

$$A = 1 \tag{3.38}$$

The spectrum is therefore continuous and non-degenerate for $0 < E < V_0$.

This solution describes for $x < 0$ the superposition of two waves, the incident one with wave number $+k = \frac{p}{\hbar}$ and the wave reflected by the potential step and which propagates in the opposite direction with wave number $-k$. The incident and the reflected waves are out of phase, the phase difference being equal to $2\,\theta_E$. We also observe that the wave function has an exponentially descreasing tail inside the classically forbidden region $x > 0$.

### 3.5.3 Energy eigenfunctions with $E > V_0$

In this case $p'$ is a *real* number. The form of the solution of the Schrödinger equation is

$$\psi_E(x) = \begin{cases} A\,\mathrm{e}^{\frac{i}{\hbar}\,p\,x} + B\,\mathrm{e}^{-\frac{i}{\hbar}\,p\,x} & \text{if } x < 0, \\ A'\,\mathrm{e}^{\frac{i}{\hbar}\,p'\,x} + B'\,\mathrm{e}^{-\frac{i}{\hbar}\,p'\,x} & \text{if } x > 0. \end{cases} \tag{3.39}$$

Continuity of $\psi_E(x)$ and of its first derivative at $x = 0$ requires

$$A + B = A' + B' \qquad p\,\big(A - B\big) = p'\,\big(A' - B'\big) \tag{3.40}$$

These equations determine two of the four constants $A, B, A'$ and $B'$ in terms of the other. For example

$$\begin{aligned} A' &= \frac{1}{2}\,\big(1 + \frac{p}{p'}\big)\,A + \frac{1}{2}\,\big(1 - \frac{p}{p'}\big)\,B \\ B' &= \frac{1}{2}\,\big(1 - \frac{p}{p'}\big)\,A + \frac{1}{2}\,\big(1 + \frac{p}{p'}\big)\,B \end{aligned} \tag{3.41}$$

which we can write in a matrix form

$$\begin{pmatrix} A' \\ B' \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 + \frac{p}{p'} & 1 - \frac{p}{p'} \\ 1 - \frac{p}{p'} & 1 + \frac{p}{p'} \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix} \tag{3.42}$$

Equivalently,

$$\begin{aligned} B &= \frac{p - p'}{p + p'}\,A + \frac{2\,p'}{p + p'}\,B' \\ A' &= \frac{2\,p}{p + p'}\,A + \frac{p' - p}{p + p'}\,B' \end{aligned} \tag{3.43}$$

The general solution of the Schrödinger equation is therefore the linear superposition of the solution with $A = 1$ and $B' = 0$.

$$\psi^+_E(x) = \begin{cases} \mathrm{e}^{\frac{i}{\hbar}\,p\,x} + \frac{p-p'}{p+p'}\,\mathrm{e}^{-\frac{i}{\hbar}\,p\,x} & \text{if } x < 0, \\ \frac{2\,p}{p+p'}\,\mathrm{e}^{\frac{i}{\hbar}\,p'\,x} & \text{if } x > 0 \end{cases} \tag{3.44}$$

and of the solution with $A = 0$ and $B' = 1$

$$\psi^-_E(x) = \begin{cases} \frac{2\,p'}{p+p'}\,\mathrm{e}^{-\frac{i}{\hbar}\,p\,x} & \text{if } x < 0, \\ \frac{p'-p}{p+p'}\,\mathrm{e}^{\frac{i}{\hbar}\,p'\,x} + \,\mathrm{e}^{-i\,k'\,x} & \text{if } x > 0 \end{cases} \tag{3.45}$$

We conclude that the spectrum for $E > V_0$ is *continuous* and *doubly degenerate*.

The solution $\psi^+_E(x)$ is characterized by the absence, for $x > 0$, of the component with negative momentum. It is the quantum solution which corresponds to the classical trajectory in which the particle coming from $x = -\infty$ with momentum $p$ flies off to $x = +\infty$ with momentum $p'$. For $x < 0$ this solution is the superposition of an incoming wave-particle of momentum $p$ and a reflected wave-particle with momentum $-p$; for $x > 0$ it consists of the transmitted wave-particle of momentum $p'$. The existence for $x < 0$ of a reflected component is a quantum mechanical effect which has no analogue in the classical context.

The energy eigenstate $\psi^+_E(x)$ is relevant for the description of the physical process in which a particle with momentum $p$, coming from $x = -\infty$, hits the barrier.

In an analogous way, the energy eigenstate $\psi^-_E(x)$ is the quantum equivalent of the classical trajectory in which the the particle coming from $x = +\infty$ with momentum $-p'$ flies off to $x = -\infty$ with momentum $-p$.

## 3.6 Tunnel effect

Let us now consider a potential barrier of finite length

$$V(x) = \begin{cases} 0 & \text{if } x < 0, \\ V_0 & \text{if } 0 < x < L \\ 0 & \text{if } x > L \end{cases} \tag{3.46}$$

with $V_0 > 0$. Classical trajectories with $0 < E < V_0$ are restricted either to $x < 0$ or to $x > L$. They describe particles coming from $x = \pm\infty$ bouncing off at the potential barrier in $x = 0$ or $x = L$ and going back to $x = \mp\infty$. Classical trajectories with $E > V_0$ instead either start from $x = -\infty$ with momentum $p = \sqrt{2\,m\,E} > 0$ and fly off to $x = \infty$ with the same momentum, or come from $x = +\infty$ with negative momentum $-p$ and fly off to $x = -\infty$ with the same momentum.

The general solution of the Schrödinger equation for $E > 0$ is

$$\psi_E(x) = \begin{cases} A\,\mathrm{e}^{i\,k\,x} + B\,\mathrm{e}^{-i\,k\,x} & \text{if } x < 0 \\ A'\,\mathrm{e}^{i\,k'\,x} + B'\,\mathrm{e}^{-i\,k'\,x} & \text{if } 0 < x < L \\ A''\,\mathrm{e}^{i\,k\,x} + B''\,\mathrm{e}^{-i\,k\,x} & \text{if } x > L \end{cases} \tag{3.47}$$

where

$$k = \frac{p}{\hbar} = \frac{\sqrt{2\,m\,E}}{\hbar} \qquad k' = \frac{p'}{\hbar} = \frac{\sqrt{2\,m\,(E - V_0)}}{\hbar} \tag{3.48}$$

and $A, B, A', B', A'', B''$ are complex constants. According to the general discussion of section 3.4 we expect the energy spectrum to be continuous and doubly degenerate for any $E > 0$.

Continuity of the wave function and of its first derivative at $x = 0$ gives the same equations (3.42) that we have already solved in the previous subsection:

$$\begin{pmatrix} A' \\ B' \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + \frac{k}{k'} & 1 - \frac{k}{k'} \\ 1 - \frac{k}{k'} & 1 + \frac{k}{k'} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} \tag{3.49}$$

Continuity of $\psi_E(x)$ and of its first derivative at $x = L$ gives equations analogous to (3.40) relating $A'', B''$ to $A', B'$

$$\begin{aligned} & A'\,\mathrm{e}^{i\,k'\,L} + B'\,\mathrm{e}^{-i\,k'\,L} = A''\,\mathrm{e}^{i\,k\,L} + B''\,\mathrm{e}^{-i\,k\,L} \\ & k'\left(A'\,\mathrm{e}^{i\,k'\,L} - B'\,\mathrm{e}^{-i\,k'\,L}\right) = k\left(A''\,\mathrm{e}^{i\,k\,L} - B''\,\mathrm{e}^{-i\,k\,L}\right) \end{aligned} \tag{3.50}$$

These equations can be obtained by those in (3.40) by means of the replacements

$$p \leftrightarrow p'$$

$$A \to A'\,\mathrm{e}^{i\,k'\,L} \qquad B \to B'\,\mathrm{e}^{-i\,k'\,L} \tag{3.51}$$

$$A' \to A''\,\mathrm{e}^{i\,k\,L} \qquad B' \to B''\,\mathrm{e}^{-i\,k\,L} \tag{3.52}$$

We therefore obtain for the solution of (3.50)

$$\begin{aligned} \begin{pmatrix} A''\,\mathrm{e}^{i\,k\,L} \\ B''\,\mathrm{e}^{-i\,k\,L} \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 1 + \frac{k'}{k} & 1 - \frac{k'}{k} \\ 1 - \frac{k'}{k} & 1 + \frac{k'}{k} \end{pmatrix} \begin{pmatrix} A'\,\mathrm{e}^{i\,k'\,L} \\ B'\,\mathrm{e}^{-i\,k'\,L} \end{pmatrix} = \\ &= \frac{1}{2} \begin{pmatrix} 1 + \frac{k'}{k} & 1 - \frac{k'}{k} \\ 1 - \frac{k'}{k} & 1 + \frac{k'}{k} \end{pmatrix} \begin{pmatrix} \mathrm{e}^{i\,k'\,L} & 0 \\ 0 & \mathrm{e}^{-i\,k'\,L} \end{pmatrix} \begin{pmatrix} A' \\ B' \end{pmatrix} \end{aligned} \tag{3.53}$$

Plugging (3.49) into (3.53) we obtain

$$\begin{aligned} \begin{pmatrix} A''\,\mathrm{e}^{i\,k\,L} \\ B''\,\mathrm{e}^{-i\,k\,L} \end{pmatrix} &= \frac{1}{4} \begin{pmatrix} 1 + \frac{k'}{k} & 1 - \frac{k'}{k} \\ 1 - \frac{k'}{k} & 1 + \frac{k'}{k} \end{pmatrix} \begin{pmatrix} \mathrm{e}^{i\,k'\,L} & 0 \\ 0 & \mathrm{e}^{-i\,k'\,L} \end{pmatrix} \times \\ &\times \begin{pmatrix} 1 + \frac{k}{k'} & 1 - \frac{k}{k'} \\ 1 - \frac{k}{k'} & 1 + \frac{k}{k'} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} \end{aligned} \tag{3.54}$$

Hence

$$\begin{aligned}\begin{pmatrix}A''\\ B''\end{pmatrix} = \frac{1}{4}\begin{pmatrix}\mathrm{e}^{-i\,k\,L} & 0\\ 0 & \mathrm{e}^{i\,k\,L}\end{pmatrix}\begin{pmatrix}1+\frac{k'}{k} & 1-\frac{k'}{k}\\ 1-\frac{k'}{k} & 1+\frac{k'}{k}\end{pmatrix}\times\\ \times\begin{pmatrix}\mathrm{e}^{i\,k'\,L} & 0\\ 0 & \mathrm{e}^{-i\,k'\,L}\end{pmatrix}\begin{pmatrix}1+\frac{k}{k'} & 1-\frac{k}{k'}\\ 1-\frac{k}{k'} & 1+\frac{k}{k'}\end{pmatrix}\begin{pmatrix}A\\ B\end{pmatrix} =\\ \equiv \mathcal{T}\begin{pmatrix}A\\ B\end{pmatrix}\end{aligned} \tag{3.55}$$

where we introduced the matrix $\mathcal{T}$,

$$\begin{aligned}&\mathcal{T} = \frac{1}{4}\times\\ &\times\begin{pmatrix}\frac{(k+k')^2}{k\,k'}\,\mathrm{e}^{-i\,(k-k')\,L} - \frac{(k-k')^2}{k\,k'}\,\mathrm{e}^{-i\,(k+k')\,L} & -\frac{k^2-(k')^2}{k\,k'}\left(\mathrm{e}^{-i\,(k-k')\,L} - \mathrm{e}^{-i\,(k+k')\,L}\right)\\ -\frac{k^2-(k')^2}{k\,k'}\left(\mathrm{e}^{i\,(k-k')\,L} - \mathrm{e}^{i\,(k+k')\,L}\right) & \frac{(k+k')^2}{k\,k'}\,\mathrm{e}^{i\,(k-k')\,L} - \frac{(k-k')^2}{k\,k'}\,\mathrm{e}^{i\,(k+k')\,L}\end{pmatrix}\end{aligned} \tag{3.56}$$

which is called the *transmission matrix* of the potential barrier. Eqs. (3.55) and (3.49) completely determine the solution $\psi_E(x)$ once $A$ and $B$ are given. As anticipated, therefore, the energy spectrum is continuous and doubly degenerate.

Let us now pause for a moment to discuss some general properties of the transfer matrix. Observe that if $\psi_E(x)$ is a solution of the Schrödinger equation, the function

$$W(x) = \psi^*_E(x)\,\psi'_E(x) - (\psi^*_E(x))'\,\psi_E(x) \tag{3.57}$$

is in fact constant. Indeed

$$\begin{aligned}W'(x) &= \psi^*_E(x)\,\psi''_E(x) - (\psi^*_E(x))''\,\psi_E(x) =\\ &= \psi^*_E(x)\,\frac{2\,m}{\hbar^2}\,(V(x)-E)\,\psi_E(x) +\\ &-\frac{2\,m}{\hbar^2}\,(V(x)-E)\,\psi^*_E(x)\,\psi_E(x) = 0\end{aligned} \tag{3.58}$$

Let us apply this result to a solution of the Schrödinger equation for a potential barrier $V(x)$ which *vanishes* at $x \to \pm\infty$. The solution at spatial infinity has the form

$$\psi_E(x) = \begin{cases}A\,\mathrm{e}^{i\,k\,x} + B\,\mathrm{e}^{-i\,k\,x} & \text{if } x \to -\infty\\ \tilde{A}\,\mathrm{e}^{i\,k\,x} + \tilde{B}\,\mathrm{e}^{-i\,k\,x} & \text{if } x \to +\infty\end{cases} \tag{3.59}$$

where $k^2 = \frac{2\,m\,E}{\hbar^2}$. Let us compute $W(x)$ at $x \to -\infty$:

$$\begin{aligned} W(-\infty) = {} & i\,k\,(A^*\,\mathrm{e}^{-i\,k\,x} + B^*\,\mathrm{e}^{i\,k\,x})\,(A\,\mathrm{e}^{i\,k\,x} - B\,\mathrm{e}^{-i\,k\,x}) + \\ & + i\,k(A^*\,\mathrm{e}^{-i\,k\,x} - B^*\,\mathrm{e}^{i\,k\,x})\,(A\,\mathrm{e}^{i\,k\,x} + B\,\mathrm{e}^{-i\,k\,x}) = \\ & = 2\,i\,k\,(|A|^2 - |B|^2) \end{aligned} \tag{3.60}$$

We obtain, analogously, for $W(+\infty)$

$$W(+\infty) = 2\,i\,k\,(|\tilde{A}|^2 - |\tilde{B}|^2) \tag{3.61}$$

From (3.58) we conclude

$$|A|^2 - |B|^2 = |\tilde{A}|^2 - |\tilde{B}|^2 \tag{3.62}$$

The coefficients $\tilde{A}$ and $\tilde{B}$ are linear functions of $A$ and $B$, since the equation is linear:

$$\tilde{A} = \alpha\,A + \beta\,B \qquad \tilde{B} = \gamma\,A + \delta\,B \tag{3.63}$$

Moreover, if $\psi_E(x)$ is a solution, $\psi_E^*(x)$ is an energy eigenfunction with the same value of $E$, since the Schrödinger equation is real. Hence

$$\psi_E^*(x) = \begin{cases} B^*\,\mathrm{e}^{i\,k\,x} + A^*\,\mathrm{e}^{-i\,k\,x} & \text{if } x \to -\infty \\ \tilde{B}^*\,\mathrm{e}^{i\,k\,x} + \tilde{A}^*\,\mathrm{e}^{-i\,k\,x} & \text{if } x \to +\infty \end{cases} \tag{3.64}$$

The linear relation between $\tilde{B}^*$ and the coefficients $(B^*, A^*)$ should be the same as the one between $\tilde{A}$ and $(A, B)$

$$\tilde{B}^* = \alpha\,B^* + \beta\,A^* \quad \Rightarrow \tilde{B} = \beta^*\,A + \alpha^*\,B \tag{3.65}$$

We conclude that

$$\gamma = \beta^* \qquad \delta = \alpha^* \tag{3.66}$$

and

$$\begin{pmatrix} \tilde{A} \\ \tilde{B} \end{pmatrix} = \mathcal{T}\begin{pmatrix} A \\ B \end{pmatrix} \qquad \mathcal{T} = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} \tag{3.67}$$

Inserting these relations into Eq. (3.62) we obtain

$$\begin{aligned} |A|^2 - |B|^2 = {} & |\alpha\,A + \beta B|^2 - |\beta^*\,A + \alpha^*\,B|^2 = \\ & = (|\alpha|^2 - |\beta|^2)\,|A|^2 + (|\beta|^2 - |\alpha|^2)\,|B|^2 \end{aligned} \tag{3.68}$$

Therefore

$$|\alpha|^2 - |\beta|^2 = 1 \tag{3.69}$$

Consider now the solution with $\tilde{B} = 0$ and $A = 1$:

$$\psi_E^{(+)}(x) = \begin{cases} \mathrm{e}^{i\,k\,x} + B\,\mathrm{e}^{-i\,k\,x} & \text{if } x \to -\infty \\ \tilde{A}\,\mathrm{e}^{i\,k\,x} & \text{if } x \to +\infty \end{cases} \tag{3.70}$$

We can express the coefficients $B$ and $\tilde{A}$ in terms of the elements of the transfer matrix

$$\begin{aligned} &\tilde{A} = \alpha + \beta\,B \qquad 0 = \beta^* + \alpha^*\,B \ \Rightarrow \\ &B = -\frac{\beta^*}{\alpha^*} \qquad \tilde{A} = \frac{1}{\alpha^*} \end{aligned} \tag{3.71}$$

The solution $\psi_E^{(+)}(x)$ can be interpreted as describing a flux of particle of momentum $k$ coming from $x \to -\infty$, part of which, proportional to $|B|^2$, is reflected by the barrier and part, proportional to $|\tilde{A}|^2$, is transmitted through the barrier to $x \to +\infty$. We define therefore the *reflection* and *transmission* coefficients $R$ and $T$ of the potential barrier as

$$R = \Big| -\frac{\beta^*}{\alpha^*}\Big|^2 = \frac{|\beta|^2}{|\alpha|^2} \qquad T = \Big|\frac{1}{\alpha^*}\Big|^2 = \frac{1}{|\alpha|^2} \tag{3.72}$$

Note that, indeed,

$$R + T = \frac{|\beta|^2 + 1}{|\alpha|^2} = 1 \tag{3.73}$$

thanks to the relation (3.69). The collective intepretation of the Schrödinger wave function, which we will develop in Section (4.2), shows that $R$ and $T$ are precisely the reflected and transmitted fractions of a *flux* of incoming non-interacting particles, with momentum $p$, hitting the potential barrier from the negative spatial infinity.

Going back to the square barrier (3.46), we obtain from the transfer matrix (3.56)

$$\begin{aligned} -\frac{\beta^*}{\alpha^*} &= \frac{(k^2 - (k')^2)\,\big(\mathrm{e}^{-i\,k'\,L} - \mathrm{e}^{i\,k'\,L}\big)}{(k + k')^2\,\mathrm{e}^{-i\,k'\,L} - (k - k')^2\,\mathrm{e}^{i\,k'\,L}} \\ \frac{1}{\alpha^*} &= \frac{4\,k\,k'\,\mathrm{e}^{-i\,k\,L}}{(k + k')^2\,\mathrm{e}^{-i\,k'\,L} - (k - k')^2\,\mathrm{e}^{i\,k'\,L}} \end{aligned} \tag{3.74}$$

The reflection and transmission coefficients for the square potential barrier are therefore

$$\begin{aligned} R &= \frac{|k^2-(k')^2|^2\left|\mathrm{e}^{-i\,k'\,L}-\mathrm{e}^{i\,k'\,L}\right|^2}{\left|(k+k')^2\,\mathrm{e}^{-i\,k'\,L}-(k-k')^2\,\mathrm{e}^{i\,k'\,L}\right|^2} \\ T &= \frac{16\,k^2\,|k'|^2}{\left|(k+k')^2\,\mathrm{e}^{-i\,k'\,L}-(k-k')^2\,\mathrm{e}^{i\,k'\,L}\right|^2} \end{aligned} \tag{3.75}$$

Let us specialize this formula to the case

$$0 < E < V_0 \tag{3.76}$$

when $p'$ is *purely imaginary*

$$p' = i\,|p'| = i\,\sqrt{2\,m\,(V_0-E)} \tag{3.77}$$

The transmission coefficient becomes

$$\begin{aligned} T &= \frac{16\,k^2\,|k'|^2}{\left|(k+i\,|k'|)^2\,\mathrm{e}^{-|k'|\,L}-(k-i\,|k'|)^2\,\mathrm{e}^{|k'|\,L}\right|^2} = \\ &= \frac{4\,k^2\,(k')^2}{(k^2-(k')^2)^2\,\sinh^2(k'\,L)+4\,k^2\,(k')^2\,\cosh^2(k'\,L)} = \\ &= \frac{16\,k^2\,|k'|^2\,\mathrm{e}^{-2\,|k'|\,L}}{(k^2+|k'|^2)^2\left|1-\frac{(k+i\,|k'|)^2}{(k-i|k'|)^2}\,\mathrm{e}^{-2\,|k'|\,L}\right|^2} = \\ &= \frac{16\,k^2\,|k'|^2\,\mathrm{e}^{-2\,|k'|\,L}}{(k^2+|k'|^2)^2\left|1-\mathrm{e}^{-2\,|k'|\,L+4\,i\,\theta_E}\right|^2} = \\ &= \frac{16\,k^2\,|k'|^2\,\mathrm{e}^{-2\,|k'|\,L}}{(k^2+|k'|^2)^2}\,\frac{1}{1+\mathrm{e}^{-4\,|k'|\,L}-2\,\mathrm{e}^{-2\,|k'|\,L}\,\cos(4\,\theta_E)} \\ &= \frac{16\,E\,\big(V_0-E\big)\,\mathrm{e}^{-2\,|k'|\,L}}{V_0^2}\,\frac{1}{1+\mathrm{e}^{-4\,|k'|\,L}-2\,\mathrm{e}^{-2\,|k'|\,L}\,\cos(4\,\theta_E)} \end{aligned} \tag{3.78}$$

where

$$\tan\theta_E = \frac{|k'|}{k} \tag{3.79}$$

When

$$|k'|\,L \gg 1 \Rightarrow \mathrm{e}^{-2\,|k'|\,L} \ll 1 \tag{3.80}$$

this formula simplifes

$$T \approx \frac{16\,E\,(V_0 - E)\,\mathrm{e}^{-2\sqrt{\frac{2\,m\,L^2}{\hbar^2}(V_0-E)}}}{V_0^2} \tag{3.81}$$

$T$ is the probability that a particle of energy $E < V_0$ and positive momentum $p$ coming from the left of the potential barrier "tunnels" to the its right hand side. This is the so-called "tunnel effect" and, of course, has no classical analogue. Let us remark that this probability is proportional to the exponential

$$T \propto \mathrm{e}^{-2\,\frac{|p'|\,L}{\hbar}} = \mathrm{e}^{-\frac{2}{\hbar}\int_0^L |p'|\,dx} = \mathrm{e}^{-\frac{2}{\hbar}\int_0^L \sqrt{2\,m\,(V(x)-E)}\,dx} \tag{3.82}$$

This formula, although derived for the squared potential barrier (3.46), is in fact valid for a general undimensional potential barrier $V(x)$. The computation of the expression for the multiplicative factors in front of the exponential in the case of general potentail is less elementary and we refer to the bibliography for its derivation.

**Problem 3.** Show that the exponential factor which controls the probability of tunneling through the barrier

$$V(x) = \begin{cases} +\infty & \text{if } x < 0, \\ -V_0 & \text{if } 0 < x < L \\ \frac{Z\,e^2}{4\,\pi\,\epsilon_0\,x} & \text{if } x > L \end{cases} \tag{3.83}$$

in the limit

$$E \ll \frac{Z\,e^2}{4\,\pi\,\epsilon_0\,L} \tag{3.84}$$

is

$$T \propto \mathrm{e}^{-\frac{\pi\,Z\,e^2}{4\,\pi\,\epsilon_0\,\hbar}\sqrt{\frac{2\,m}{E}}} = \mathrm{e}^{-\frac{Z\,e^2}{2\,\epsilon_0\,\hbar\,v}} \tag{3.85}$$

Historically, the most famous application of this formula was due to Gamow (1929) who used it in the context of the alpha radioactive decay. $\alpha$ particles are helium nuclei; they have electric charge $+2\,e$ and mass $m$ which is about 4 times the proton mass

$$m \approx 6.6 \times 10^{-27}\ Kg \approx 7300\,m_{el} \tag{3.86}$$

$\alpha$ particles are emitted by heavy nuclei like uranium, thorium, actinium, radium, etc, with large atomic numbers $Z_A \gtrsim 50$. The potential felt by an $\alpha$ particle can be schematically represented by the function (3.83) with

$$Z = 2 \times (Z_A - 2) \approx 2\, Z_A \tag{3.87}$$

and $L$ equal to the nucleus radius:

$$L \approx 10^{-15}\, m \tag{3.88}$$

Then

$$\frac{Z\, e^2}{4\, \pi\, \epsilon_0\, L} \approx Z \times 5 \times 10^4 \times 27.2\, ev \approx Z \times 1\, Mev \tag{3.89}$$

Since the energy $E$ of the emitted $\alpha$ particle is in the range

$$E = 3\text{–}7\, Mev \tag{3.90}$$

the condition (3.84) for the validity of (3.85) applies. To estimate the decay life time of $\alpha$-decaying radioactive nuclei, we observe that the velocity of the $\alpha$ particle inside the nuclei is

$$v = \sqrt{\frac{2\, E}{m}} = \sqrt{\frac{E}{E_1^H}\, \frac{m_{el}}{m}\, \frac{e^2}{\hbar\, c}}\, c \approx (1.2\text{–}1.8) \times 10^7\, m/s \tag{3.91}$$

where $E_1^H = \frac{m_{el}\, e^4}{2\, (4\, \pi\, \epsilon_0)^2\, \hbar^2} = 13.6\, ev$ is the ionization energy of the electron and $m_{el} = 0.9 \times 10^{-30}\, kg$ is the electron mass. Since the particle is trapped in a potential box of size $L$, the frequency with which the particle hits the potential barrier is

$$\frac{v}{2\, L} \approx (6.0\text{–}9.1) \times 10^{21} sec^{-1} \sim 10^{22}\, sec^{-1} \tag{3.92}$$

The probability per unit of time for the particle to escape the box is therefore

$$\frac{dP}{dt} \equiv \frac{1}{\tau} \sim \frac{v}{2\, L}\, \mathrm{e}^{-\frac{Z_A\, e^2}{\epsilon_0\, \hbar\, v}} \tag{3.93}$$

where $\tau$ is the decay-time of the radioactive element. One obtains in this way the Gamow-Geiger-Nuttal law[13]

$$\log\frac{1}{\tau} = -\frac{e^2\sqrt{m}}{\sqrt{2}\,\epsilon_0\,\hbar}\,\frac{Z_A}{\sqrt{E}} + \log\frac{\sqrt{E}}{\sqrt{2\,m}\,L} \tag{3.94}$$

which gives the dependence of the decay-time on the energy $E$ of the emitted $\alpha$ particle and on the atomic number $Z_A$ of the radioactive nucleus.

To have an idea of the order of magnitude of the life-times predicted by Gamow formula (3.93), consider the example of $E = 5.0\,Mev$, $v = 1.55\times 10^7\,m/s$, $L = 10^{-15}\,m$:

$$\tau \sim 4.1\times 10^{-30}\,\mathrm{e}^{1.77\,Z_A}\,yrs \sim 4.1\times 10^{0.77\,Z_A-30}\,yrs \tag{3.95}$$

**Problem 4.** Find the reflection and transmission coefficients $R$ and $T$ for the potential barrier (3.46) for $E > V_0$.

**Problem 5.** Find the reflection and transmission coefficients $R$ and $T$ for the potential barrier (3.46) for $E < V_0$, in the limit in which $L \to 0$, $V_0 \to +\infty$ and $V_0\,L = \mathcal{V}$ constant. (This is the so-called "thin barrier" limit.)

**Solution.** We can consider the potential to be

$$V(x) = \mathcal{V}\,\delta(x) \tag{3.96}$$

Let us integrate the Schrödinger equation on a little interval $[-\epsilon, \epsilon]$ surrounding the origin:

$$\begin{aligned} 0 &= \int_{-\epsilon}^{\epsilon} dx\,\Big[-\frac{\hbar^2}{2\,m}\,\psi_E''(x) + (\mathcal{V}\,\delta(x) - E)\,\psi_E(x)\Big] = \\ &= -\frac{\hbar^2}{2\,m}\,(\psi'(\epsilon) - \psi'(-\epsilon) + \mathcal{V}\,\delta(x)\,\psi_E(0) \end{aligned} \tag{3.97}$$

---

[13] The Geiger-Nuttal law was first empirically derived by Geiger and Nuttal (1911), in the form

$$\log\frac{1}{\tau} = -\frac{a_1}{\sqrt{E}} + a_2,$$

with $a_1$ and $a_2$ constants. This has the same form as the Gamow formula (3.94), with the "constant" $a_2$ being replaced by a function which depends logarithmically on the energy. This dependence is negligible with respect to the strong power-like dependence of the term proportional to $a_1$. Moreover in the derivation we outlined, due to Gamow (1928), we were not very precise regarding the multiplicative factors in front of the tunneling exponential: this in principle might further affect the (weak) energy dependence of $a_2$.

Taking the limit for $\epsilon \to 0^+$ we obtain for the discontinuity of the first derivative at the origin

$$\Delta\psi'(0) = \frac{2\,m\,\mathcal{V}}{\hbar^2}\,\psi_E(0) \tag{3.98}$$

The wave function has the form

$$\psi_E(x) = \begin{cases} A_1\,\mathrm{e}^{i\,k\,x} + A_2\,\mathrm{e}^{-i\,k\,x} & \text{if } x < 0 \\ \tilde{A}_1\,\mathrm{e}^{i\,k\,x} + \tilde{A}_2\,\mathrm{e}^{-i\,k\,x} & \text{if } x > 0 \end{cases} \tag{3.99}$$

with $k^2 = \frac{2\,m\,E}{\hbar^2}$. Continuity of the wave function at $x = 0$ implies

$$\tilde{A}_1 + \tilde{A}_2 = A_1 + A_2 \tag{3.100}$$

while the discontinuity of the first derivative at $x = 0$ gives

$$\tilde{A}_1 - \tilde{A}_2 = A_1 - A_2 + \frac{2\,m\,\mathcal{V}}{i\,k\,\hbar^2}\,(A_1 + A_2) \tag{3.101}$$

Hence

$$\begin{aligned} \tilde{A}_1 &= (1 - \frac{i\,m\,\mathcal{V}}{k\,\hbar^2})\,A_1 - \frac{i\,m\,\mathcal{V}}{k\,\hbar^2}\,A_2 \\ \tilde{A}_2 &= \frac{i\,m\,\mathcal{V}}{k\,\hbar^2}\,A_1 + (1 + \frac{i\,m\,\mathcal{V}}{k\,\hbar^2})\,A_2 \end{aligned} \tag{3.102}$$

The transfer matrix is therefore

$$\mathcal{T} = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} = \begin{pmatrix} 1 - \frac{i\,m\,\mathcal{V}}{k\,\hbar^2} & -\frac{i\,m\,\mathcal{V}}{k\,\hbar^2} \\ \frac{i\,m\,\mathcal{V}}{k\,\hbar^2} & 1 + \frac{i\,m\,\mathcal{V}}{k\,\hbar^2} \end{pmatrix} \tag{3.103}$$

and the transmission and the reflection coefficients

$$\begin{aligned} T &= \Big|\frac{1}{\alpha^*}\Big|^2 = \Big|\frac{1}{1 + \frac{i\,m\,\mathcal{V}}{k\,\hbar^2}}\Big|^2 = \frac{1}{1 + \frac{m^2\,\mathcal{V}^2}{k^2\,\hbar^4}} \\ R &= \Big| -\frac{\beta^*}{\alpha^*}\Big|^2 = \Big| -\frac{\frac{i\,m\,\mathcal{V}}{k\,\hbar^2}}{1 + \frac{i\,m\,\mathcal{V}}{k\,\hbar^2}}\Big|^2 = \frac{\frac{m^2\,\mathcal{V}^2}{k^2\,\hbar^4}}{1 + \frac{m^2\,\mathcal{V}^2}{k^2\,\hbar^4}} \end{aligned} \tag{3.104}$$

**Problem 6.** Determine the discrete levels and the energy eigenfunctions of an unidimensional particle moving in the potential well

$$V(x) = \begin{cases} 0 & \text{if } |x| > \frac{a}{2}, \\ V_0 & \text{if } |x| \leq \frac{a}{2}. \end{cases} \tag{3.105}$$

for $V_0 < 0$.

**Problem 7.** An unidimensional particle moves on the real semi-line $x > 0$ subject to the potential

$$V(x) = \begin{cases} +\infty & \text{if } x < 0 \\ -V_0 & \text{if } 0 < x < L, \\ 0 & \text{if } x > L. \end{cases} \tag{3.106}$$

for $V_0 > 0$. Find the range of values of $V_0$ for which there is only one single bound state.

**Problem 8.** Determine the discrete levels of an unidimensional particle moving in the potential well (3.105), using the expression for the transmission matrix.

**Problem 9.** Determine the energy eigenfunctions for the potential barrier

$$V(x) = \begin{cases} 0 & \text{if } |x| > \frac{a}{2}, \\ V_0 & \text{if } |x| \leq \frac{a}{2}. \end{cases} \tag{3.107}$$

for $V_0 > 0$, which are also eigenstates of the spatial inversion operator $I : \psi(x) \to \psi(-x)$.

**Problem 10.**

A 1-dimensional particle moves on the half line $x > 0$ in the potential

$$V(x) = \begin{cases} +\infty & \text{if } x < 0, \\ 0 & \text{if } 0 < x < L. \\ V_0 & \text{if } L < x < L + a. \\ 0 & \text{if } L + a < x. \end{cases} \tag{3.108}$$

with $V_0 > 0$. Determine the energy eigenfunctions.

## 3.7 Periodic potentials

In this section we will study the properties of the energy spectrum of one-dimensional Hamiltonians with *periodic* potentials $V(x)$:

$$V(x + a) = V(x) \tag{3.109}$$

This kind of Hamiltonians can be used to model, for example, the behaviour of electrons moving in a lattice formed by the ions of a solid.

Let us introduce the linear operator $\hat{T}_a$ whose action on wave functions is defined to be

$$\hat{T}_a \, \psi(x) = \psi(x + a) \tag{3.110}$$

It is simple to show that $\hat{T}_a$ is *unitary*:

$$\begin{aligned} \langle \hat{T}_a \, \psi_1, \hat{T}_a \, \psi_2 \rangle &= \int_{-\infty}^{\infty} dx \, \psi_1^*(x + a) \, \psi_2(x + a) = \\ &= \int_{-\infty}^{\infty} dx' \, \psi_1^*(x') \, \psi_2(x') = \langle \psi_1, \psi_2 \rangle \end{aligned} \tag{3.111}$$

Let us $\psi_{t_a}$ be an eigenvector of $\hat{T}_a$ with eigenvalue $t_a$:

$$\hat{T}_a \, \psi_{t_a} = t_a \, \psi_{t_a} \tag{3.112}$$

From (3.111), it is immediately seen that eigenvalues $t_a$ of the unitary $\hat{T}_a$ are complex numbers of unit modulus

$$\begin{aligned} &\langle \hat{T}_a \; \psi_{t_a}, \hat{T}_a \, \psi_{t_a} \rangle = |t_a|^2 \, \langle \psi_{t_a}, \psi_{t_a} \rangle = \langle \psi_{t_a}, \psi_{t_a} \rangle \\ &\quad \Rightarrow \; |t_a|^2 = 1 \end{aligned} \tag{3.113}$$

We can therefore write

$$t_a = \mathrm{e}^{i \, \phi(a)} \tag{3.114}$$

where $\phi(a)$ is a real phase. Let us observe that

$$(\hat{T}_a)^n = \hat{T}_{n \, a} \tag{3.115}$$

for any relative integer $n$. Therefore

$$(\hat{T}_a)^n \, \psi_{t_a} = \mathrm{e}^{i \, n \, \phi(a)} \, \psi_{t_a} = \hat{T}_{n \, a} \, \psi_{t_a} \tag{3.116}$$

Hence

$$\mathrm{e}^{i \, n \, \phi(a)} = \mathrm{e}^{i \, \phi(n \, a)} \tag{3.117}$$

for any integer $n$ and real $a$. Eq. (3.117) implies that the generic eigenvalue of $\hat{T}_a$ has the form

$$\mathrm{e}^{i \, \phi(a)} = \mathrm{e}^{i \, \tilde{k} \, a} \tag{3.118}$$

where $\tilde{k}$ is a real number. In conclusion, eigevalues and eigenfunctions of $\hat{T}_a$

$$\hat{T}_a \, \psi_{\tilde{k}}(x) = \psi_{\tilde{k}}(x+a) = \mathrm{e}^{i \, \tilde{k} \, a} \, \psi_{\tilde{k}}(x) \tag{3.119}$$

are labelled by the real number $\tilde{k}$. $\hbar \, \tilde{k}$ is called the *quasi-momentum* of the wave-function $\psi_{\tilde{k}}(x)$. Let us observe that quasi-momenta[14] $\tilde{k}$ which differ by integer multiples of $\frac{2\,\pi}{a}$ correspond to the same eigenvalue of $\hat{T}_a$ :

$$\tilde{k} \sim \tilde{k} + \frac{2\,\pi}{a} \, n \qquad \text{for } n \text{ integer} \tag{3.120}$$

Hence, in order to label distinct $\hat{T}_a$ eigenvalues, the quasi-momentum $\tilde{k}$ can be restricted to take values in a interval of length $\frac{2\pi}{a}$: i.e., in the segment

$$\tilde{k} \in \, ] - \frac{\pi}{a}, \frac{\pi}{a}] \tag{3.121}$$

which is called the *fundamental Brillouin zone.*

Let $\psi_{\tilde{k}}(x)$ be an eigenfunction of $\hat{T}_a$ with quasi-momentum $\tilde{k}$. The function $\phi_{\tilde{k}}(x)$ defined by the relation

$$\psi_{\tilde{k}}(x) \equiv \mathrm{e}^{i \, \tilde{k} \, x} \, \phi_{\tilde{k}}(x) \tag{3.122}$$

is *periodic.* Indeed

$$\begin{aligned} \phi_{\tilde{k}}(x+a) &= \mathrm{e}^{-i \, \tilde{k} \, (x+a)} \, \psi_{\tilde{k}}(x+a) = \mathrm{e}^{-i \, \tilde{k} \, (x+a)} \, \mathrm{e}^{i \, \tilde{k} \, a} \, \psi_{\tilde{k}}(x) = \\ &= \mathrm{e}^{-i \, \tilde{k} \, x} \, \psi_{\tilde{k}}(x) = \phi_{\tilde{k}}(x) \end{aligned} \tag{3.123}$$

were we used the property (3.119). The eigenfunctions (3.122) are called *Bloch waves* of pseudo-momentum $\tilde{k}$.

Let us now consider the Hamiltonian of a one-dimensional system

$$\hat{H} = \frac{\hat{p}^2}{2\,m} + V(\hat{x}) \tag{3.124}$$

with $V(x)$ periodic of period $a$

$$V(x+a) = V(x) \tag{3.125}$$

---

[14] For simplicity, we will use the term quasi-momenta for both the wave-number $\tilde{k}$ and the momentum $\hbar \, \tilde{k}$.

It is easily seen that $\hat{H}$ commutes with $\hat{T}_a$

$$[\hat{H}, \hat{T}_a] = 0 \tag{3.126}$$

Therefore $\hat{H}$ and $\hat{T}_a$ admit a system of simultaneous eigenfunctions

$$\begin{aligned} \hat{H}\, \psi_{E_n(\tilde{k})}(x) &= E_n(\tilde{k})\, \psi_{E_n(\tilde{k})}(x) \\ \hat{T}_a, \psi_{E_n(\tilde{k})}(x) &= \mathrm{e}^{i\,\tilde{k}\,a}\, \psi_{E_n(\tilde{k})}(x) \end{aligned} \tag{3.127}$$

which write as

$$\psi_{E_n(\tilde{k})}(x+a) = \mathrm{e}^{i\,\tilde{k}\,x}\, \phi_{E_n(\tilde{k})}(x) \tag{3.128}$$

with $\phi_{E_n(\tilde{k})}(x)$ periodic. The energy eigenvalues $E_n(\tilde{k})$ of the simultaneous eigenfunctions depend on the continuous quasi-momentum $\tilde{k}$ and on a quantum number $n$, which labels distinct energy eigenfunctions with the same quasi-momentum $\tilde{k}$. The quantum number $n$ will be, in general, *discrete*. Hence, the energy spectrum $E_n(\tilde{k})$ will be, in general, made of discrete *energy bands*, labelled by the index $n$.

An intuitive way to understand why the quantum number $n$ is discrete is to observe that, once we fix $\tilde{k}$, the energy eigenfunctions are identified by the periodic functions $\phi_E(x)$ in Eq. (3.128). Therefore $\phi_E(x)$ is a solution of the *homogenous*, linear, second order differential equation which is obtained by plugging (3.128) into the Schrödinger equation. Searching for periodic solutions amounts to impose *two distinct* boundary conditions

$$\phi_{E_n(\tilde{k})}(0) = \phi_{E_n(\tilde{k})}(a) \qquad \phi'_{E_n(\tilde{k})}(0) = \phi'_{E_n(\tilde{k})}(a) \tag{3.129}$$

on such an equation. The discussion of Section (3.4) made clear, however, that two distinct boundary conditions for a linear, homogenous second order differential equations, select, for generic $E$, the null solution: only for discrete, special values of the energy $E$, non-trivial solutions of two distinct boundary conditions exist. In the following we will show more explicitly how this works, by analyzing in detail some specific potentials.

As an example, consider the periodic potential barrier of hight $V_0$ and thickness $L$

$$V(x) = \begin{cases} 0 & \text{for} \quad -\frac{a}{2} < x < \frac{a}{2} - L \\ V_0 & \text{for} \quad \frac{a}{2} - L < x < \frac{a}{2} \end{cases} \tag{3.130}$$

We know from (3.42) that the solution of the corresponding Schrödinger equation takes the form

$$\psi_E(x) = \begin{cases} A_1\,\mathrm{e}^{i\,k\,x} + A_2\,\mathrm{e}^{-i\,k\,x} & \text{if } \frac{a}{2} < x < \frac{a}{2} - L, \\ A_1'\,\mathrm{e}^{i\,k'\,x} + A_2'\,\mathrm{e}^{-i\,k'\,x} & \text{if } \frac{a}{2} - L < x < \frac{a}{2}, \\ A_1''\,\mathrm{e}^{i\,k\,x} + A_2''\,\mathrm{e}^{-i\,k\,x} & \text{if } \frac{a}{2} < x < a - L, \end{cases} \tag{3.131}$$

where

$$k^2 = \frac{2\,m\,E}{\hbar^2} \qquad k'^2 = \frac{2\,m\,(E - V_0)}{\hbar^2} \tag{3.132}$$

The coefficients of the plane waves to the left and to the right of the barrier are related by

$$\begin{pmatrix} A_1'' \\ A_2'' \end{pmatrix} = \mathcal{T} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \tag{3.133}$$

where $\mathcal{T}$ is the transmission matrix

$$\mathcal{T} = \begin{pmatrix} \alpha(k) & \beta(k) \\ \beta^*(k) & \alpha^*(k) \end{pmatrix} \tag{3.134}$$

with

$$|\alpha|^2 - |\beta|^2 = 1 \tag{3.135}$$

We calculated explicitly the functions $\alpha(k)$ and $\beta(k)$ for the periodic square potential (3.130) in subsection (3.6), but let us stress that the relations (3.134) and (3.134) are valid for any barrier of this kind. The quasi-periodicity condition

$$\psi_E(-\frac{a}{2} + x) = \mathrm{e}^{-i\,\tilde{k}\,a}\,\psi_E(x + \frac{a}{2}) \qquad 0 < x < \frac{a}{2} \tag{3.136}$$

leads to

$$\begin{aligned} A_1\,\mathrm{e}^{-i\,\frac{k\,a}{2}}\,\mathrm{e}^{i\,k\,x} + A_2\,\mathrm{e}^{+i\,\frac{k\,a}{2}}\,\mathrm{e}^{-i\,k\,x} &= \\ = \mathrm{e}^{-i\,\tilde{k}\,a}\left[A_1''\,\mathrm{e}^{i\,k\,\frac{a}{2}}\,\mathrm{e}^{i\,k\,x} + A_2''\,\mathrm{e}^{-i\,k\,\frac{a}{2}}\,\mathrm{e}^{-i\,k\,x}\right]& \end{aligned} \tag{3.137}$$

One obtains therefore that the wave function is quasi-periodic if and only if

$$\begin{aligned} A_1'' &= A_1\,\mathrm{e}^{-i\,(k-\tilde{k})\,a} \\ A_2'' &= A_2\,\mathrm{e}^{i\,(k+\tilde{k})\,a} \end{aligned} \tag{3.138}$$

or, using the matrix notation, if

$$\mathcal{T}\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{-i\,(k-\tilde{k})\,a} & 0 \\ 0 & \mathrm{e}^{i\,(k+\tilde{k})\,a} \end{pmatrix}\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \tag{3.139}$$

We conclude that the condition for quasi-periodicity is equivalent to the vanishing of the following determinant

$$\det\begin{pmatrix} \alpha - \mathrm{e}^{-i\,(k-\tilde{k})\,a} & \beta \\ \beta^* & \alpha^* - \mathrm{e}^{i\,(k+\tilde{k})\,a} \end{pmatrix} = 0 \tag{3.140}$$

We obtain in this way the equation

$$1 - \mathrm{e}^{i\,\tilde{k}\,a}\,(\mathrm{e}^{i\,k\,a}\,\alpha + \mathrm{e}^{-i\,k\,a}\,\alpha^*) + \mathrm{e}^{2\,i\,\tilde{k}\,a} = 0 \tag{3.141}$$

where we used the property (3.135) of the transmission matrix $\mathcal{T}$. Putting

$$\alpha(k) = |\alpha|(k)\,\mathrm{e}^{i\,\theta_\alpha(k)} = \alpha_1(k) + i\,\alpha_2(k) \tag{3.142}$$

where $\alpha_1$ and $\alpha_2$ are the real and imaginary parts of $\alpha$, Eq. (3.141) rewrites as follows

$$\cos\tilde{k}\,a = |\alpha|\,\cos(\theta_\alpha + k\,a) = \alpha_1\,\cos k\,a - \alpha_2\,\sin\,k\,a \tag{3.143}$$

This equation is valid for any periodic potential barrier: it gives $\cos\tilde{k}\,a$, i.e. the quasi-momentum[15], as a function of $k$ or, equivalently, of the energy $E$. Since $|\alpha| > 1$, the function on the right hand side of (3.143), as $k$ varies, will cross the roots $k_n^\pm$ of the equation

$$\cos(\theta_\alpha(k_n^\pm) + k_n^\pm\,a) = \pm\frac{1}{|\alpha(k_n^\pm)|} \qquad n = 1, 2, \ldots \tag{3.144}$$

The roots $k_n^\pm$ divide the $k$ line into a sequence of segments where the absolute value of the function on the right hand side of (3.143) is, alternatively, less or greater than 1: the formers correspond to energy bands while the latter are the energy *gaps* where no level exists.

Inverting the relation (3.143), one obtains $k$ as function of $\tilde{k}$: this is called the *dispersion relation*. For each allowed $k$ interval the function

---

[15] Remember that the quasi-momentum $\tilde{k}$ is restricted to the Brillouin zone, therefore $\cos\tilde{k}\,a$ determines it up to a sign.

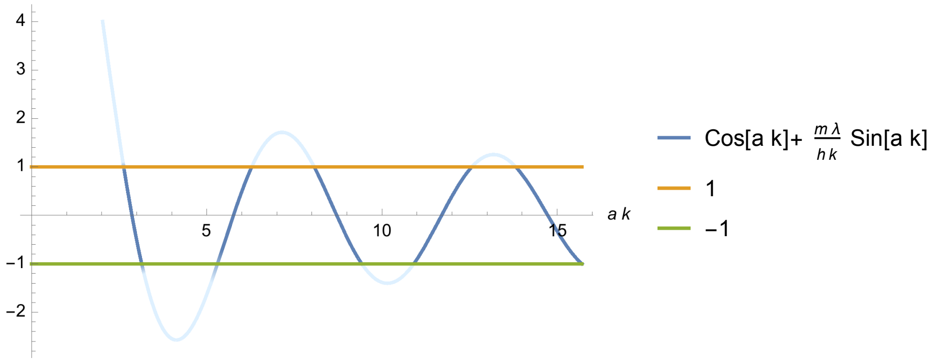


Figure 1: The r.h.s. of Eq.(3.143) as a function of $a\,k$, for the thin periodic barrier with $\frac{m\,\lambda\,a}{\hbar^2} = 10$. The darker portions of the graph are the regions where the absolute value of the function is less than 1 and solutions of the equation for $\tilde{k}$ exist.

$|\alpha|\,\cos(\theta_\alpha + k\,a)$ has a different inverse, and correspondingly for each energy band one obtains a different dispersion relation $E_n(\tilde{k})$.

Since the analytical discussion of Eq. (3.143) for the square periodic potential (3.130) is somewhat involved, let us consider the limit of a thin periodic barrier: in this limit the hight of the barrier $V_0 \to \infty$ as its thickness $L \to 0$ while the product $V_0\,L \equiv \lambda$ is kept fixed. The resulting periodic potential is therefore

$$V(x) = \sum_n \lambda\,\delta(x - \frac{a}{2} + n\,a) \tag{3.145}$$

As obtained in one of the exercises of a previous subsection, the transmission matrix $\mathcal{T}$ for this potential is

$$\alpha = 1 - i\,\frac{m\,\lambda}{\hbar^2\,k} \qquad \beta = i\,\frac{m\,\lambda}{\hbar^2\,k} \tag{3.146}$$

The equation (3.141) writes therefore

$$\cos\tilde{k}\,a = \frac{\cos\,(k - \bar{k})\,a}{\cos a\,\bar{k}} = \cos k\,a + \tan\bar{k}\,a\,\sin k\,a \tag{3.147}$$

where $\bar{k}$ is a *function* of $k$ defined by

$$\tan\bar{k}\,a \equiv \frac{m\,\lambda}{\hbar^2\,k} \qquad |\bar{k}\,a| < \frac{\pi}{2} \tag{3.148}$$

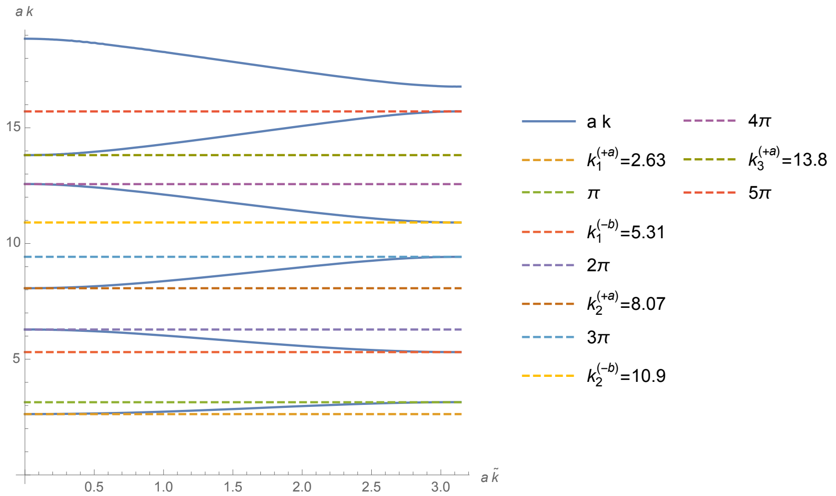


Figure 2: $a\,k$ as a function of $a\,\tilde{k}$ for $\frac{m\,\lambda\,a}{\hbar^2} = 10$. The horizontal dashed lines denote the boundaries of the $k$ bands.

Let us study the solutions of (3.147) as functions of the pseudo-momentum $\tilde{k}$, for $\lambda > 0$. To start with, note that Eq. (3.147) does not depend on the sign of $\tilde{k}$. We can therefore restrict $\tilde{k}$ to be in half of the Brillouin zone $\tilde{k} \in [0, \frac{\pi}{a}]$. The reflected values of $\tilde{k}$ will give the same energy levels, and they will corresponds to Bloch waves which "propagate" in the opposite direction of the real $x$ axis.

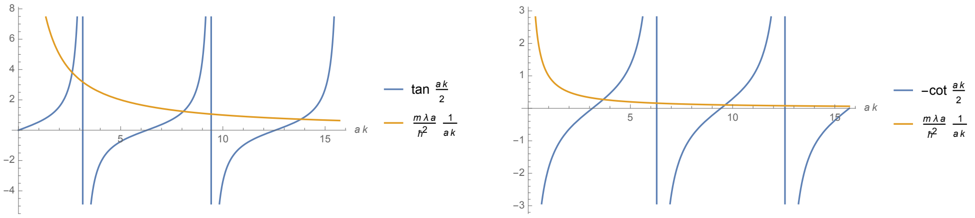
The r.h.s. of Eq. (3.147) as a function of $a\,k$, with $\frac{m\,\lambda\,a}{\hbar^2} = 10$, is graphed in Fig. 1. The darker portions of the graph denote the regions for which $\tilde{k}$ solving (3.147) exists. In Fig. 2 the inverse relations which give $a\,k$ as function of the quasi-momentum $a\,\tilde{k}$ are pictured.

Equation (3.144) which determines the locations of bands and gaps writes for the thin periodic barrier as

$$\cos a\,(k_n^{\pm} - \bar{k}(k_n^{\pm})) = \pm \cos a\,\bar{k}(k_n^{\pm}) \tag{3.149}$$

Let us first look at the $k_n^{+}$ roots. They are two types: those, which we will call of type $(a)$, such that

$$\begin{aligned} (a) \qquad & a\,(k_n^{+} - \bar{k}(k_n^{+})) = a\,\bar{k}(k_n^{+}) + 2\,n\,\pi \Leftrightarrow \\ & a\,k_n^{(+a)} = 2\,a\,\bar{k}(k_n^{(+a)}) + 2\,n\,\pi, \end{aligned} \tag{3.150}$$



(a) (b)

Figure 3: (a) The roots $a\,k_n^{(+a)}$, with $n = 1, 2, 3$, for $\frac{m\,\lambda\,a}{\hbar^2} = 10$. (b) The roots $a\,k_n^{(-b)}$, with $n = 1, 2$, for $\frac{m\,\lambda\,a}{\hbar^2} = 10$.

and those of type $(b)$, for which

$$\begin{aligned} (b) \quad & a\,(k_n^+ - \bar{k}(k_n^+)) = -a\,\bar{k}(k_n^+) + 2\,n\,\pi \Rightarrow \\ & a\,k_n^{(+b)} = 2\,n\,\pi \end{aligned} \tag{3.151}$$

From (3.148) we deduce that $k_n^{(a+)}$ are the solutions of the equation

$$\tan\frac{k_n^{(+a)}\,a}{2} = \frac{m\,\lambda}{\hbar^2\,k_n^{(a+)}} \tag{3.152}$$

These roots are represented in Fig.3(a) as intersections of the curves $\tan\frac{k\,a}{2}$ and $\frac{m\,\lambda}{\hbar^2\,k}$ for $\frac{m\,\lambda\,a}{\hbar^2} = 10$. It is easily seen from the graph that $k_n^{(+a)}$ satisfy

$$2\,(n-1)\,\pi < a\,k_n^{(+a)} < (2\,n-1)\,\pi \qquad n = 1, 2, \ldots \tag{3.153}$$

Turning to the $k_n^-$'s, we can analogously distinguish the $(a)$ type solutions

$$\begin{aligned} (a) \qquad & a\,(k_n^- - \bar{k}(k_n^-)) = \pi - a\,\bar{k}(k_n^-) + 2\,n\,\pi \Rightarrow \\ & a\,k_n^{-(a)} = (2\,n+1)\,\pi, \end{aligned} \tag{3.154}$$

and the solutions of type $(b)$:

$$\begin{aligned} (b) \qquad & a\,(k_n^- - \bar{k}(k_n^-)) = a\,\bar{k}(k_n^-) - \pi + 2\,n\,\pi \Rightarrow \\ & a\,k_n^{(-b)} = 2\,a\,\bar{k}(k_n^{(-b)}) + (2\,n-1)\,\pi \end{aligned} \tag{3.155}$$

Eq. (3.148) implies that $k_n^{(-b)}$ are the roots of the equation

$$-\cot\frac{k_n^{(-b)}\,a}{2} = \frac{m\,\lambda}{\hbar^2\,k_n^{(-b)}} \tag{3.156}$$

These roots are represented in Fig. 3 as intersections of the curves $-\cot\frac{k\,a}{2}$ and $\frac{m\,\lambda}{\hbar^2\,k}$ for $\frac{m\,\lambda\,a}{\hbar^2}=10$. They satisfy

$$(2\,n-1)\,\pi < a\,k_n^{(-b)} < 2\,n\,\pi \qquad n=1,2,\ldots \tag{3.157}$$

In conclusion, as we can see from Fig. 1, the allowed ranges for $k$ are

$$\begin{aligned} a\,k_n^{(+a)} &\le a\,k \le a\,k^{(-a)} = (2\,n+1)\,\pi \\ a\,k_n^{(-b)} &\le a\,k \le a\,k^{(+b)} = 2\,n\,\pi \end{aligned} \tag{3.158}$$

Note that the lower edges of bands of type $(a)$ correspond to $\tilde{k}=0$ while the higher edges correspond to $\tilde{k}=\frac{\pi}{a}$: i.e. $k$ *increases* as $\tilde{k}$ increases from 0 to $\frac{\pi}{a}$. For the bands of type (b), the reverse is true, the lower edges of bands correspond to $\tilde{k}=\frac{\pi}{a}$ while the higher edges correspond to $\tilde{k}=0$: $k$ *decreases* as $\tilde{k}$ increases from 0 to $\frac{\pi}{a}$. (See Fig. 2).

Summarizing, the energy bands corresponding to the allowed ranges of $k$, Eq. (3.158), are

$$\frac{\hbar^2\,k_n^2}{2\,m} \le E_n \le \frac{\hbar^2\,n^2\,\pi^2}{2\,m\,a^2} \qquad n=1,2,\ldots \tag{3.159}$$

where the set $\{k_n\}$, is the union of the sets $\{k_n^{(+a)}\}$ and $\{k_n^{(-b)}\}$. Recalling that $\epsilon_n=\frac{\hbar^2\,n^2\,\pi^2}{2\,m\,a^2}$ are precisely the energy levels for a particle in a box of length $a$, we see that the bands for the periodic potentials, are

$$E_n \in [\epsilon_n-\Delta_n,\epsilon_n] \qquad n=1,2,\ldots \tag{3.160}$$

with widths $\Delta_n$ equal to

$$\Delta_n = \frac{\hbar^2\,n^2\,\pi^2}{2\,m\,a^2} - \frac{\hbar^2\,k_n^2}{2\,m} \qquad n=1,2,\ldots \tag{3.161}$$

As a consistency check, let us note that when $\lambda\to+\infty$, the solutions of equations (3.152) and (3.156) approach the values

$$a\,k_n \to n\,\pi \tag{3.162}$$

Therefore, in this limit, the widths $\Delta_n$ vanish, in agreement with the fact that as $\lambda$ increases the periodic barrier becomes more and more impenetrable and the energy spectrum becomes discrete. Note also that when $k \to +\infty$, Eq. (3.147) becomes

$$\cos \tilde{k}\, a = \cos\, k\, a + O(\frac{1}{k}) \tag{3.163}$$

Therefore in the limit of large $k$ the allowed range for $a\, k$ in the $n$-th band covers more and more of the interval $[(n-1)\,\pi, n\,\pi]$, the energy gaps between the bands shrink to zero and the spectrum becomes continuous.

## 3.8 Harmonic Oscillator

The Hamiltonian on an unidimensional harmonic oscillator is

$$\hat{H}(x,p) = \frac{\hat{p}^2}{2\,m} + \frac{1}{2}\, m\, \omega^2\, \hat{x}^2 \tag{3.164}$$

The associated Schrödinger equation is

$$(-\frac{\hbar^2}{2\,m}\,\frac{d^2}{d\,x^2} + \frac{1}{2}\, m\, \omega^2\, \hat{x}^2)\psi''_E(x) = E\,\psi_E(x) \tag{3.165}$$

The discussion of the previous sections indicates that this system has a discrete, positive and non-degenerate energy spectrum. This spectrum can be derived by looking for the normalizable solutions of equation (3.165), which can be found explicitly in terms of elementary functions. The techniques involved are however somewhat involved (see bibliography for a description of these methods). Therefore in this section we will derive the eigenvalues and eigenvectors of the energy of this system by means of a simpler and more instructive approach, which does not use the Schrödinger representation.

Let us introduce the operators which are linear combinations of $\hat{x}$ and $\hat{p}$

$$a \equiv \frac{\lambda}{\hbar}\,(\hat{p} - i\, m\, \omega\, \hat{x}) \quad a^\dagger \equiv \frac{\lambda}{\hbar}\,(\hat{p} + i\, m\, \omega\, \hat{x}) \quad \lambda \equiv \sqrt{\frac{\hbar}{2\, m\, \omega}} \tag{3.166}$$

Note that:

- $a$ and $a^\dagger$ are not hermitian operators, therefore they certainly do not correspond to observables.
- $\lambda$ has the dimensions of a length: it is proportional to the amplitude of the classical harmonic oscillations of energy $E_0 = \frac{1}{2}\,\hbar\,\omega$.

- $a$ and $a^\dagger$ are dimensionless.

Recalling the canonical commutation relations, we obtain

$$\begin{aligned} & \big[a, a^\dagger\big] = \frac{\lambda^2}{\hbar^2}\big(i\,m\,\omega\,(-i\,\hbar) - i\,m\,\omega\,(i\,\hbar)\big) = \frac{2\,m\,\omega\,\lambda^2}{\hbar} = 1 \\ & \big[a, a\big] = \big[a^\dagger, a^\dagger\big] = 0 \end{aligned} \tag{3.167}$$

Consider the *hermitian* operator

$$\hat{N} \equiv a^\dagger\,a \tag{3.168}$$

Observe that this operator has only *positive* eigenvalues. Indeed, if $|n\rangle$ is an eigenvector of $\hat{N}$ of eigenvalue $n$, then

$$a^\dagger\,a\,|n\rangle = n\,|n\rangle \tag{3.169}$$

Taking the hermitian product of both side of this equation with the vector $|n\rangle$ we obtain

$$\langle n|\,a^\dagger\,a\,|n\rangle = n\,\langle n|n\rangle \tag{3.170}$$

or, equivalently,

$$\big|a\,|n\rangle\big|^2 = n\big||n\rangle\big|^2 \tag{3.171}$$

Therefore

$$n \geq 0 \tag{3.172}$$

i.e., $\hat{N}$ has a *non-negative* spectrum.

Consider now the commutators:

$$\big[\hat{N}, a\big] = -a \qquad \big[\hat{N}, a^\dagger\big] = a^\dagger \tag{3.173}$$

where we used (3.167). If $|n\rangle$ is an eigenvector of $\hat{N}$ of eigenvalue $n$, one has

$$\begin{aligned} \hat{N}\big(a^\dagger\,|n\rangle\big) &= +\big[\hat{N}, a^\dagger\big]\,|n\rangle + a^\dagger\,\hat{N}\,|n\rangle = a^\dagger\,|n\rangle + n\,a^\dagger\,|n\rangle = \\ &= (n+1)\,\big(a^\dagger\,|n\rangle\big) \\ \hat{N}\big(a\,|n\rangle\big) &= +\big[\hat{N}, a\big]\,|n\rangle + a\,\hat{N}\,|n\rangle = -a\,|n\rangle + n\,a\,|n\rangle = \\ &= (n+1)\,\big(a\,|n\rangle\big) \end{aligned} \tag{3.174}$$

Therefore if $n$ is an eigenvalue of $N$ then $n-1$ is another eigenvalue, *unless* $a\,|n\rangle = 0$. By iterating this reasoning we conclude that the existence of an eigenvalue $n$ ensures the existence of the eigenvalues $n-1, n-2, \ldots, n-k\cdots$ and this sequence stops if and only if there is a positive integer $k$ such that

$$a^k\,|n\rangle = 0 \tag{3.175}$$

In the same way one shows that the existence of an eigenvalue $n$ ensures the existence of the eigenvalues $n+1, n+2, \ldots, n+k\cdots$, corresponding to the eigenvectors

$$a^\dagger\,|n\rangle, \ldots, (a^\dagger)^k\,|n\rangle\ \ldots \tag{3.176}$$

and this sequence stops if and only if there is a positive integer $k$ such that

$$(a^\dagger)^k\,|n\rangle = 0 \tag{3.177}$$

Since we just showed that $\hat{N}$ has a non-negative spectrum, an integer $k$ for which (3.175) is true must exist. Therefore there must be a $\bar{n} \equiv n-k+1$ for which

$$a\,|\bar{n}\rangle = 0 \tag{3.178}$$

This implies

$$0 = \big|a\,|\bar{n}\rangle\big|^2 = \langle\bar{n}|\,a^\dagger\,a\,|\bar{n}\rangle = \langle\bar{n}|\,\hat{N}\,|\bar{n}\rangle = \bar{n}\,\langle\bar{n}|\bar{n}\rangle \tag{3.179}$$

Hence

$$\bar{n} = 0 \tag{3.180}$$

We conclude that there exists an eigenvector of $\hat{N}$ with zero eigenvalue; let us denote it by $|0\rangle$. This vector is annihilated by $a$

$$a\,|0\rangle = 0 \tag{3.181}$$

For what we said before, acting on $|0\rangle$ with $a^\dagger$ one obtains eigenvectors with *integer* eigenvalues $n = 1, 2, 3, \cdots$. We would like to know if this sequence stops, that is if there exist a *non-negative* integer $k_{max}$ such that

$$(a^\dagger)^{k_{max}+1}|0\rangle = 0 \tag{3.182}$$

If this were the case, there would exist an eigenvector $|k_{max}\rangle$ such that

$$a^\dagger\,|k_{max}\rangle = 0 \tag{3.183}$$

Repeating the steps we took above

$$\begin{aligned}0 = \left|a^\dagger\,|k_{max}\rangle\right|^2 = \langle k_{\max}|\,a\,a^\dagger\,|k_{max}\rangle = \langle k_{max}|\,\big(1+\hat{N}\big)\,|k_{max}\rangle = \\ = (k_{max}+1)\,\langle k_{max}|k_{max}\rangle\end{aligned} \tag{3.184}$$

i.e.

$$k_{max}+1 = 0, \tag{3.185}$$

in contradiction with the fact that $k_{max}$ is non-negative. We conclude that (3.182) is *never* satisfied and that the operator $\hat{N}$ has eigenvalues

$$n = 0, 1, 2, \dots \tag{3.186}$$

consisting of *all* natural numbers.

The relevance of the operator $\hat{N}$ for the problem of the harmonic oscillators becomes apparent when we express it in terms of $\hat{x}$ and $\hat{p}$

$$\begin{aligned}\hat{N} &= \frac{\lambda^2}{\hbar^2}\,\big(\hat{p}+i\,m\,\omega\,\hat{x}\big)\,\big(\hat{p}-i\,m\,\omega\,\hat{x}\big) = \\ &= \frac{\lambda^2}{\hbar^2}\,\big[\hat{p}^2+m^2\,\omega^2\,\hat{x}^2+i\,m\,\omega\big(\hat{x}\,\hat{p}-\hat{p}\,\hat{x}\big)\big] = \\ &= \frac{\lambda^2}{\hbar^2}\,\big[\hat{p}^2+m^2\,\omega^2\,\hat{x}^2-m\,\omega\,\hbar\big] = \\ &= \frac{2\,m\,\lambda^2}{\hbar^2}\,\big[\frac{\hat{p}^2}{2\,m}+\frac{1}{2}\,m\,\omega^2\,\hat{x}^2-\frac{1}{2}\,\hbar\,\omega\big]\end{aligned} \tag{3.187}$$

Hence the hamiltonian of the harmonic oscillator writes as

$$\hat{H} = \frac{\hat{p}^2}{2\,m}+\frac{m\,\omega^2}{2}\,\hat{x}^2 = \frac{\hbar^2}{2\,m\,\lambda^2}\,\hat{N}+\frac{1}{2}\,\hbar\,\omega = \hbar\,\omega\,\big(\hat{N}+\frac{1}{2}\big) \tag{3.188}$$

Therefore the eigenvectors $|n\rangle$ of $\hat{N}$ are also eigenvectors of the harmonic oscillator hamiltonian

$$\hat{H}|n\rangle = \hbar\,\omega\,\big(n+\frac{1}{2}\big)\,|n\rangle \tag{3.189}$$

Incidentally, since we know already (from the discussion of the previous section) that the spectrum of $\hat{H}$ is non-degenerate and discrete, this shows that $\hat{N}$ also has a non-degenerate spectrum and thus there is only one eigenvector $|n\rangle$ for any integer $n$ eigenvalue of $\hat{N}$.

Let us obtain the *normalized* eigenvectors $|n\rangle$ of $\hat{H}$ and $\hat{N}$ starting from the normalized ground state, i.e. the eigenstate with lowest energy

$$a\,|0\rangle = 0 \qquad \langle n|n\rangle = 1 \tag{3.190}$$

Acting with $a^\dagger$ on $|n\rangle$ we obtain an eigenvector with $\hat{N}$ with eigenvalue $n+1$:

$$a^\dagger\,|n\rangle = c_n\,|n+1\rangle \tag{3.191}$$

The norm of this vector is

$$\begin{aligned}\left|a^\dagger\,|n\rangle\right|^2 = |c_n|^2 = \langle n|\,a\,a^\dagger\,|n\rangle = \langle n|\,\big(\hat{N}+1\big)|n\rangle = (n+1)\,\langle n|n\rangle = \\ = n+1\end{aligned} \tag{3.192}$$

Therefore we can take

$$c_n = \sqrt{n+1} \tag{3.193}$$

and

$$a^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle \tag{3.194}$$

**Problem 10.** Show that

$$\begin{aligned} a\,|n\rangle &= \sqrt{n}\,|n-1\rangle \\ |n\rangle &= \frac{1}{\sqrt{n!}}\,(a^\dagger)^n\,|0\rangle \end{aligned} \tag{3.195}$$

**Problem 11.** Compute the averages of $\hat{x}^2, \hat{x}, \hat{p}^2, \hat{p}$ on the ground state $|0\rangle$ of the harmonic oscillator and show it is a state of minimal uncertainty. (*Hint*: write $\hat{x}$ and $\hat{p}$ in terms of $a$ and $a^\dagger$.)

# 4 Time evolution

## 4.1 The Schrödinger picture

So far we have discussed the quantum mechanical description of a system at a given fixed time. The problem of time evolution in quantum mechanics is the problem to determine the state of the system $v(t)$ at a time $t$ knowing the state of the system $v(t_0)$ at a given initial time $t_0$ and the forces which act on the system at any given time.

Given the linear structure of the space of states it is natural to assume that time evolution be determined by a *linear* and *unitary* operator

$$\begin{aligned} &v(t) = U(t;t_0)\, v_0 \qquad v(t_0) = v_0 \qquad U(t_0,t_0) = \mathbb{I} \\ &U(t;t_0)\, U^\dagger(t;t_0) = U^\dagger(t;t_0)\, U(t;t_0) = \mathbb{I} \end{aligned} \tag{4.196}$$

We have a consistency condition

$$U(t;t_1)\, U(t_1;t_0) = U(t;t_0) \tag{4.197}$$

If we are considering a system subject to forces which are *independent of time* then $U(t;t_0)$ depends only on the difference $t - t_0$

$$U(t;t_0) = U(t-t_0) \tag{4.198}$$

The consistency condition (4.197) becomes

$$U(t_1)\, U(t_2) = U(t_1 + t_2) \tag{4.199}$$

This implies that $U(t)$ has the form

$$U(t) = \mathrm{e}^{-i\hat{K}\, t} \tag{4.200}$$

where $\hat{K}$ is an hermitian operator independent of time

$$\hat{K} = \hat{K}^\dagger \tag{4.201}$$

Since the hermitian operator $\hat{K}$ encodes the dynamics of the system it is natural to postulate that it is proportional to the Hamiltonian of the system

$$\hat{K} = \frac{1}{\hbar}\hat{H} \qquad U(t) = \mathrm{e}^{-\frac{i}{\hbar}\hat{H}\, t} \tag{4.202}$$

In conclusion, the time evolution of a state is described by the first order linear differential equation

$$\begin{aligned} \frac{d\,v(t)}{dt} &= -\frac{i}{\hbar}\,\hat{H}\,v(t) \\ v(0) &= v_0 \end{aligned} \tag{4.203}$$

whose solution can be written, for time independent Hamiltonians, as follows

$$v(t) = \mathrm{e}^{-\frac{i}{\hbar}\,\hat{H}\,t}\,v_0 \tag{4.204}$$

Although (4.203) has been derived by assuming that the forces of the systems are time independent it maintains its validity even when the system is subject to time-dependent forces: in this case however $\hat{H}$ becomes explicitly time dependent and the solution of equation (4.203) is *not* given by (4.204).

Working for example in the Schrödinger representation, the time evolution equation becomes a differential equation of the form

$$i\,\hbar\,\frac{\partial\,\psi(\vec{x},t)}{\partial x} = \hat{H}(\vec{x}, -i\,\hbar\,\vec{\nabla}_x, t)\,\psi(\vec{x},t) \tag{4.205}$$

where $H(\vec{x},\vec{p},t)$ is the (possibly time dependent) Hamiltonian.

Going back to the case of forces independent of time, let us look for states which do not evolve in time: we are going to call these states *stationary*. Remembering the fact that vectors which differ by a multiplicative complex number describe the same state, we conclude that vectors corresponding to stationary states can be written as

$$v(t) = C(t)\,v_0 \tag{4.206}$$

where $C(t)$ is a time dependent complex *number*. Therefore

$$\frac{d\,v(t)}{dt} = \dot{C}(t)\,v_0 = -\frac{i}{\hbar}\,C(t)\,\hat{H}\,v_0 \tag{4.207}$$

and thus

$$\hat{H}\,v_0 = i\,\hbar\,\frac{\dot{C}(t)}{C(t)}\,v_0 \tag{4.208}$$

Since, by hypothesis, $\hat{H}$ is time independent, we deduce

$$i\,\hbar\,\frac{\dot{C}(t)}{C(t)} = E \tag{4.209}$$

where $E$ is a time independent constant. Then

$$C(t) = C(0)\, \mathrm{e}^{-\frac{i}{\hbar}E\,t} = \mathrm{e}^{-\frac{i}{\hbar}E\,t} \tag{4.210}$$

Summarizing, stationary states are described by vectors which evolve as

$$v(t) = \mathrm{e}^{-\frac{i}{\hbar}E\,t}\, v_0 \tag{4.211}$$

where $v_0$ is an eigenvector of $\hat{H}$ of eigenvalue $E$

$$\hat{H}\, v_0 = E\, v_0 \tag{4.212}$$

Conversely, any eigenvector of the Hamiltonian is a vector which evolves as in (4.211) and therefore describes a stationary state. In summary, in quantum mechanics stationary states are states of definite energy.

Let us now look for those observables $\mathcal{O}$ whose time average on any state do not depend on time. For such observables we must have

$$\langle v_0, \mathcal{O}\, v_0\rangle = \langle v(t), \mathcal{O}\, v(t)\rangle = \langle v_0, U^\dagger(t)\, \mathcal{O}\, U(t)\, v_0\rangle \qquad \forall v_0 \in \mathcal{H} \tag{4.213}$$

Therefore

$$U^\dagger(t)\, \mathcal{O}\, U(t) = \mathcal{O} \qquad \forall\, t \tag{4.214}$$

and thus

$$\left[\hat{H}, \mathcal{O}\right] = 0 \tag{4.215}$$

Hence, in quantum mechanics, observables which commute with the Hamiltonian have time independent averages on every state. Such observables are said to be *conserved* and they correspond to the integrals of motion of classical mechanics.

The knowledge of stationary states, i.e. of the eigenstates of the Hamiltonian, allows one to recover the time evolution of any state. Let us consider in fact a basis of vectors made of eigenvectors of the Hamiltonian $\{\psi_{E_n}\}$. Any state can be decomposed in this basis

$$v_0 = \sum_n \langle \psi_{E_n}, v_0\rangle\, \psi_{E_n} \tag{4.216}$$

The time evolution of $v_0$ is

$$v(t) = \mathrm{e}^{-\frac{i}{\hbar}\,t\,\hat{H}}\,v_0 = \sum_n \langle \psi_{E_n}, v_0\rangle\,\mathrm{e}^{-\frac{i}{\hbar}\,t\,\hat{H}}\,\psi_{E_n} = \\ = \sum_n \langle \psi_{E_n}, v_0\rangle\,\mathrm{e}^{-\frac{i}{\hbar}\,t\,E_n}\,\psi_{E_n} \tag{4.217}$$

In this sense, therefore, the solution of energy eigenstates problem in quantum mechanics gives the solution of the time evolution problem. This explains why, among all observables, the Hamiltonian has a prominent place. When the Hamiltonian has a continuous spectrum, Eq. (4.217) generalizes to:

$$v(t) = \int dE \langle \psi_E, v_0\rangle\,\mathrm{e}^{-\frac{i}{\hbar}\,t\,E}\,\psi_E \tag{4.218}$$

where the generalized eigenvectors $\psi_E$ are normalized with a Dirac delta function of the energy

$$\langle \psi_E, \psi_{E'}\rangle = \delta(E - E') \tag{4.219}$$

**Problem 1.** Let $E_+$ and$E_-$ the two energy eigenvalues of a 2-states system and $|+\rangle$ $|+\rangle$ the two corresponding energy eigenstates. Suppose that at the time $t = 0$ the system is in the state $\psi(0) = \frac{1}{\sqrt{2}}\big(|+\rangle + |-\rangle\big)$. Compute: (a) the probability that at the time $t > 0$ the system is in the state $|+\rangle$; (b) the probability that at the time $t > 0$ the system is in the state $\psi(0)$.

**Problem 2**

An harmonic oscillator of frequency $\omega$ is, at time $t = 0$, in the state

$$\psi_0 = |0\rangle + |1\rangle + |2\rangle \tag{4.220}$$

- a) Compute the average of $\hat{p}$ and $\hat{p}^2$ at the time $t$.
- b) Write the normalized wave function of the state at the time $t$ in the Schrödinger representation.
- c) What is the probability that at the time $t > 0$ the system is still in the state $\psi_0$?

## 4.2 Particle flux

We have explained that the scalar quantity

$$\rho(\vec{x},t) \equiv \big|\psi(\vec{x},t)\big|^2 \tag{4.221}$$

for a normalized wave function $\psi(\vec{x},t)$ at a given time $t$

$$\int d^3\vec{x}\,|\psi(\vec{x},t)|^2 = 1 \tag{4.222}$$

should be interpreted as the *probability density* to find the particle described by the wave function $\psi(\vec{x},t)$ in the position $\vec{x}$ at the time $t$.

Let us consider now the vector quantity

$$\vec{J}(\vec{x},t) = -\frac{i\hbar}{2\,m}\big[\psi^*(\vec{x},t)\,\vec{\nabla}_x\,\psi(\vec{x},t) - \vec{\nabla}_x\,\psi^*(\vec{x},t)\,\psi(\vec{x},t)\big] \tag{4.223}$$

where the time-dependent wave function satisfies the time-dependent non-relativistic Schrödinger equation

$$i\,\hbar\,\partial_t\,\psi(\vec{x},t) = -\frac{\hbar^2}{2\,m}\,\vec{\nabla}^2\,\psi(\vec{x},t) + V(\vec{x},t)\,\psi(\vec{x},t) \tag{4.224}$$

**Problem 3.** Show that for a plane wave $\psi(\vec{x},t) = \mathrm{e}^{-i\,\frac{p^2}{2\,m\,\hbar}\,t + \frac{i}{\hbar}\,\vec{p}\cdot\vec{x}}$ the vector $\vec{J}$ is

$$\vec{J}(\vec{x},t) = \frac{\vec{p}}{m} = \vec{v} \tag{4.225}$$

The physical interpretation of the vector $\vec{J}(\vec{x},t)$ is elucidated by the following equation, which is called the *continuity equation* and which relates $\vec{J}$ to the time derivative of $\rho$:

$$\frac{\partial\rho(\vec{x},t)}{\partial t} = -\vec{\nabla}\cdot\vec{J}(\vec{x},t) \tag{4.226}$$

The proof of this equation is as follows

$$\begin{aligned}
\partial_t\big|\psi(\vec{x},t)|^2 &= \partial_t\,\psi^*(\vec{x},t)\,\psi(\vec{x},t) + \psi^*(\vec{x},t)\,\partial_t\,\psi(\vec{x},t) = \\
&= \frac{i}{\hbar}\,\big[-\frac{\hbar^2}{2\,m}\,\vec{\nabla}^2\,\psi^*(\vec{x},t) + V(\vec{x},t)\,\psi^*(\vec{x},t)\big]\,\psi(\vec{x},t) + \\
&\quad -\frac{i}{\hbar}\,\psi^*(\vec{x},t)\,\big[-\frac{\hbar^2}{2\,m}\,\vec{\nabla}^2\,\psi(\vec{x},t) + V(\vec{x},t)\,\psi(\vec{x},t)\big] = \\
&= -\frac{i\,\hbar}{2\,m}\,\big[\vec{\nabla}^2\,\psi^*(\vec{x},t)\,\psi(\vec{x},t) - \psi^*(\vec{x},t)\,\vec{\nabla}^2\,\psi(\vec{x},t)\big] = \\
&= -\frac{i\,\hbar}{2\,m}\,\vec{\nabla}\cdot\big[\vec{\nabla}\,\psi^*(\vec{x},t)\,\psi(\vec{x},t) - \psi^*(\vec{x},t)\,\vec{\nabla}\,\psi(\vec{x},t)\big]
\end{aligned} \tag{4.227}$$

To better understand the physical content of the continuity equation let us specialize it to the unidimensional case:

$$\frac{\partial\rho(x,t)}{\partial t} = -\partial_x\, J_x(x,t) \tag{4.228}$$

Consider an interval $I = [a,b]$ on the real line. The probability to find the particle on the interval $I$ at the time $t$ is

$$P(t,I) = \int_a^b dx \rho(x,t) \tag{4.229}$$

The variation of this probability in an interval of time $\Delta t$ is

$$\begin{aligned}\Delta P(t,I) = \Delta t\, \partial_t \int_a^b dx \rho(x,t) = -\Delta t \int_a^b \partial_x\, J_x(x,t) = \\ = -\Delta t\,\big(J(b,t) - J(a,t)\big)\end{aligned} \tag{4.230}$$

The l.h.s. of this equation represents the variation in a time $\Delta t$ of the probability to find the particle in the interval $I$. On the other hand, the probability to find the particle *somewhere* on the real line is, at all times, equal to one. Therefore we conclude from the equation above that the variation of the probability to find the particle *outside* of the interval $I$ is precisely

$$\Delta t\,\big(J(b,t) - J(c,t)\big) \tag{4.231}$$

We can therefore interpret $\vec{J}(\vec{x},t)$ as a *probability flux*: $\Delta t\, J(b,x)$ is the probability that, in a time $\Delta t$, the particle exits the interval $[a,b]$ from the right side, while $-\Delta t\, J(a,x)$ is the probability that, in the same interval of time, the particles escapes from the interval through the opposite boundary at $x = a$.

So far, we have given to $\rho(\vec{x},t)$ the (standard) interpretation of the probability density of finding a *single* particle at the location $\vec{x}$. The continuity equation (4.226) suggests also a *collective* interpretation for the same quantity. Consider a large number $N$ of non-interacting particles all in the same quantum mechanical state $\psi(\vec{x},t)$: then

$$\Delta^3\, \rho(\vec{x},t) \tag{4.232}$$

is the fraction of these $N$ particles which are located in the volume $\Delta^3$ around the point $\vec{x}$; and the quantity

$$\vec{n}\cdot\, \vec{J}(\vec{x},t)\, \Delta A\, \Delta t \tag{4.233}$$

is the fraction of those $N$ particles which crosses, in a time $\Delta t$, the small surface of area $\Delta A$, centered around the point $\vec{x}$, whose normal unit vector is $\vec{n}$. Therefore, if we choose the normalization of $\psi(\vec{x},t)$ to be, rather than (4.222),

$$\int d^3\vec{x}\,\rho(\vec{x},t) = N \tag{4.234}$$

then $\rho(\vec{x},t)$ is directly the density at the point $\vec{x}$ of this fluid of non-interacting particles and $\vec{J}(\vec{x},t)$ is its flux.

One of the advantages of this interpretation is that, by considering the limit $N \to \infty$ of an infinite number of particles, one is able to give a direct physical meaning to non-normalizable wave functions like those corresponding to generalized eigenstates. For example, the wave function $\psi(\vec{x},t) = \mathrm{e}^{-i\,\frac{p^2}{2\,m\,\hbar}\,t + \frac{i}{\hbar}\,\vec{p}\cdot\vec{x}}$ is interpreted as describing a fluid of independent particles of uniform unitary density

$$\rho(\vec{x},t) = \left|\mathrm{e}^{-i\,\frac{p^2}{2\,m\,\hbar}\,t + \frac{i}{\hbar}\,\vec{p}\cdot\vec{x}}\right|^2 = 1 \tag{4.235}$$

moving with constant speed

$$\vec{J}(\vec{x},t) = \frac{\vec{p}}{m} \tag{4.236}$$

## 4.3 The Heisenberg picture

In the previous subsection we described time evolution in quantum mechanics in a framework in which the vectors representing states evolve with time. This is called the *Schrödinger picture* of time evolution. There is an alternative and equivalent formulation of time evolution which goes under the name of *Heisenberg picture*.

To understand this alternative formulation, consider an observable $\mathcal{O}$ and its matrix elements between two generic states $w$ and $v$. The matrix elements depend on time according to

$$\langle w(t), \mathcal{O}\,v(t)\rangle = \langle w_0, U^\dagger(t)\,\mathcal{O}\,U(t)\,v_0\rangle \tag{4.237}$$

Let us define time dependent observables

$$\mathcal{O}_H(t) \equiv U^\dagger(t)\,\mathcal{O}\,U(t) \tag{4.238}$$

Then the time dependent matrix element above rewrites as

$$\langle w(t), \mathcal{O}\, v(t)\rangle = \langle w_0, \mathcal{O}_H(t)\, v_0\rangle \tag{4.239}$$

We can therefore think of time evolution as time evolution of observables, with vectors being time independent.

The differential equation which describes the time evolution of observables in the Heisenberg picture is, from (4.238),

$$\frac{d\,\mathcal{O}_H(t)}{dt} = \frac{i}{\hbar}\, U^\dagger(t)\left[\hat{H}, \mathcal{O}\right] U(t) = \frac{i}{\hbar}\left[\hat{H}, \mathcal{O}_H(t)\right] \tag{4.240}$$

These are ordinary differential equations for operators, and should be compared to the classical equation of motions which are ordinary differential equations for functions. Consider for example a one-dimensional system described by the Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2\,m} + V(\hat{x}) \tag{4.241}$$

The equations of motion for $\hat{x}$ and $\hat{p}$ are, according to (4.240)

$$\begin{aligned}
\frac{d\,\hat{x}_H(t)}{dt} &= \frac{i}{\hbar}\left[\hat{H}, \hat{x}_H(t)\right] = \frac{i}{\hbar}\, U^\dagger(t)\left[\hat{H}, \hat{x}\right] U(t) = \\
&= \frac{1}{m}\, U^\dagger(t)\, \hat{p}\, U(t) = \frac{\hat{p}_H(t)}{m} \\
\frac{d\,\hat{p}_H(t)}{dt} &= \frac{i}{\hbar}\left[\hat{H}, \hat{p}_H(t)\right] = \frac{i}{\hbar}\, U^\dagger(t)\left[\hat{H}, \hat{p}\right] U(t) = \\
&= -\, U^\dagger(t)\, V'(\hat{x})\, U(t) = -V'(\hat{x}_H(t))
\end{aligned} \tag{4.242}$$

Compare this with the classical equations of motion for the same system which are

$$\begin{aligned}
\frac{d\,x(t)}{dt} &= \frac{p(t)}{m} \\
\frac{d\,p(t)}{dt} &= -V'(x(t))
\end{aligned} \tag{4.243}$$

We see that the Heisenberg picture for time evolution reproduces operatorial differential equations which are in form identical to their classical counterparts.

**Problem 4.** Determine the time evolution of the *free* wave packet

$$v_{\bar{p},\Delta_p}(x) = \frac{1}{(\pi)^{\frac{1}{4}}\,\Delta_p^{\frac{1}{2}}}\int dp\,\mathrm{e}^{-\frac{(p-\bar{p})^2}{2\,\Delta^2}}\,\frac{\mathrm{e}^{\frac{i}{\hbar}\,p\,x}}{\sqrt{2\,\pi\,\hbar}} \tag{4.244}$$

**Problem 5.** Compute the uncertainties of $\hat{x}$ and $\hat{p}$ on the *free* wave packet (4.244) for $t > 0$. (*Hint*: Use the Heisenberg picture of time evolution.)

From the solution of *Problem 5* we see that if a free particle of mass $m$ is at time $t = 0$ in a state of minimal uncertainty

$$\Delta x\,\Delta p \sim \hbar \tag{4.245}$$

the uncertainty on its position increases for great times $t \to \infty$ as

$$\Delta x(t) \approx \frac{\Delta p}{m}\,t \approx \frac{\hbar\,t}{m\,\Delta x} \tag{4.246}$$

**Problem 6.** Estimate the speed at which the following free wave packets, of minimal indeterminacy at $t = 0$, spread for large $t$:

- An electron with $\Delta x \approx 10^{-8}\,cm$.
- An electron with $\Delta x \approx 10^{-2}\,cm$.
- A grain of sand of $m = 10^{-12}g$, with $\Delta x = 10^{-5}cm$.

# 5 Symmetries

We have seen that changes of representations are associated to unitary transformations. Representations which are related by unitary transformations are physically equivalent. Alternatively we can say that *different* representations related by unitary transformations are *different descriptions* of the *same* physical system.

Let us elaborate this point: suppose we have system described by a linear space of states $\mathcal{H}$ on which the observables $\mathcal{O}$ are defined. Consider now a unitary linear transformation

$$U:\ \mathcal{H} \to \mathcal{H}' \tag{5.1}$$

which sends vector $v \in \mathcal{H}$ to $v' = U\,v$. Then the description of the system by means of the state $v \in \mathcal{H}$ and the observables $\mathcal{O}$ is physically equivalent to its description by means of $v' \in \mathcal{H}'$ and observables $\mathcal{O}'$ such that

$$\mathcal{O}' = U\,\mathcal{O}\,U^\dagger \qquad v' = U\,v \tag{5.2}$$

For example, the average of $\mathcal{O}$ on the state $v$ is the same as the average of $\mathcal{O}'$ on the state $v'$

$$\langle v, \mathcal{O}\,v\rangle = \langle v, U^\dagger\,\mathcal{O}'\,U\,v\rangle = \langle v', \mathcal{O}'\,v'\rangle \tag{5.3}$$

In the same way, any transition amplitude $v \to w$ is the same as the transition amplitude between $v' \to w'$.

$$\langle w, v\rangle = \langle U^\dagger\,w', U^\dagger\,v'\rangle = \langle w', U\,U^\dagger\,v'\rangle = \langle w', v'\rangle \tag{5.4}$$

## 5.1 Translations

There are special changes of descriptions of a physical system which have a *geometrical* significance: these are changes of description associated to changes of reference frames.

For example, we can describe a given physical system either by means of cartesian coordinates $x^i$ or by means of coordinates $(x')^i$ related to the first ones by a translation

$$(x')^i = x^i + a^i \tag{5.5}$$

where $a^i$ are the coordinates of a constant vector. Let $\mathcal{H}$ be the space of states, and $v \in \mathcal{H}$ and a $v' \in \mathcal{H}$ be the vectors describing a particular state of the system in the descriptions associated, respectively, to the $x^i$ and $(x')^i$ coordinate systems. We expect that the description of the same physical state in the system of coordinates $(x')^i$ be unitarily related to the description in the $x^i$ system by Eqs. (5.1) and (5.2):

$$v' = U(a)\, v \qquad (\hat{x}')^i = \hat{x}^i + a^i = U(a)\, \hat{x}^i\, U^\dagger(a) \tag{5.6}$$

where $U(a)$ is a unitary operator which depends on the vector $a^i$.

If we now consider a third system of cartesian coordinates

$$(x'')^i = (x')^i + b^i = x^i + (a^i + b^i) \tag{5.7}$$

Then

$$\begin{aligned}(\hat{x}'')^i &= U(a+b)\, \hat{x}^i\, U(a+b)^\dagger = U(b)\, (\hat{x}')^i\, U(b)^\dagger = \\ &= U(b)\, U(a)\, \hat{x}^i \left(U(b)\, U(a)\right)^\dagger \end{aligned} \tag{5.8}$$

We conclude that[16]:

$$U(a+b) = U(a)\, U(b) \tag{5.9}$$

The solution of this equation has the form

$$U(a) = \mathrm{e}^{i \sum_j a^j \hat{K}^j} \tag{5.10}$$

where $\hat{K}^j$ is *hermitian*. Let us then consider (5.6) for $a \to 0$:

$$\hat{x}^i + a^i = \hat{x}^i + i \sum_j a^j \left[\hat{K}^j, \hat{x}^i\right] + O(a^2) \tag{5.11}$$

Therefore $K^j$ must satisfy the equations

$$\left[\hat{K}^j, \hat{x}^i\right] = -i\, \delta^{ij} \tag{5.12}$$

---

[16] Strictly speaking (5.8) would be satisfied even if $U(a+b) = U(a)\, U(b)\, \mathrm{e}^{i\,\phi(a,b)}$ where $\phi(a,b)$ is a real phase. It can be shown, however that this case can be reduced, by a redefinition, to the situation in which $\phi(a,b) \equiv 0$.

Recalling the canonical commutation relation, we observe that a solution to this equation is [17]:

$$\hat{K}^i = \frac{1}{\hbar}\,\hat{p}^i \tag{5.13}$$

The conclusion of this discussion is that the unitary operator that implements the change of description associated to the change of coordinates system (5.5) is

$$U(a) = \mathrm{e}^{\frac{i}{\hbar}\,a^i\,\hat{p}^i} \tag{5.14}$$

According to (5.2) the same physical observables corresponds, in the two coordinate systems, to different operators, $\mathcal{O}$ and $\mathcal{O}'$, related to each other by

$$\mathcal{O}' = \mathrm{e}^{\frac{i}{\hbar}\,a^i\,\hat{p}^i}\,\mathcal{O}\,\mathrm{e}^{-\frac{i}{\hbar}\,a^i\,\hat{p}^i} \tag{5.15}$$

If the observable in question is the Hamiltonian, this writes

$$\hat{H}' = \mathrm{e}^{\frac{i}{\hbar}\,a^i\,\hat{p}^i}\,\hat{H}\,\mathrm{e}^{-\frac{i}{\hbar}\,a^i\,\hat{p}^i} \tag{5.16}$$

We will say that the system in question is *invariant* — or *symmetric* — under a translation by $a^i$ if

$$\hat{H}' = \hat{H} \tag{5.17}$$

We will also say that the system is *invariant* under translations if $\hat{H}' = \hat{H}$ for *any* $a^i$. In this case we can take $a^i \to 0$, and the condition (5.17) become equivalent to the vanishing of the commutator:

$$\big[\hat{p}^i, \hat{H}\big] = 0 \tag{5.18}$$

$\hat{p}^i$ is called the *generator* of the translational symmetry: the system is invariant by translations if and only if its Hamiltonian commutes with the generator $\hat{p}^i$.

---

[17] Again, there are other solutions, but one can show that they are all equivalent to this one.

## 5.2 Rotations

We can analogously discuss the unitary transformations associated to changes of description related to *rotations* of the reference frame.

Let $(x')^i$ and $x^i$ be two systems of cartesian coordinates related by a rotation,

$$(x')^i = R^i_j\, x^j \qquad R^\intercal R = R\, R^\intercal = \mathbb{I} \tag{5.19}$$

where $R$ is a $3 \times 3$ orthogonal real matrix whose elements are $R^i_j$.

To be concrete, let us start by considering a rotation around the $z$ axis

$$\begin{aligned} x' &= \cos\theta\, x - \sin\theta\, y \\ y' &= \sin\theta\, x + \cos\theta\, y \\ z' &= z \end{aligned} \tag{5.20}$$

where $\theta$ is the rotation angle. As for translations, we expect that the description associated to the system of coordinates $(x')^i$ be unitarily related to the description in the $x^i$ system:

$$\begin{aligned} v' &= U_z(\theta)\, v \\ \hat{x}' &= \cos\theta\, \hat{x} - \sin\theta\, \hat{y} = U_z(\theta)\, \hat{x}\, U_z^\dagger(\theta) \\ \hat{y}' &= \sin\theta\, \hat{x} + \cos\theta\, \hat{y} = U_z(\theta)\, \hat{y}\, U_z^\dagger(\theta) \\ \hat{z}' &= \hat{z} = U_z(\theta)\, \hat{z}\, U_z^\dagger(\theta) \end{aligned} \tag{5.21}$$

where $U_z(\theta)$ is a unitary operator implementing the rotation around the $z$ axis by an angle $\theta$, $v$ and $v'$ are the two vectors describing the same physical state in the two descriptions.

By considering a third system of cartesian coordinates $(x'')^i$, related to the system $(x')^i$ by a rotation around the $z$ axis of an angle $\phi$, we obtain

$$\begin{aligned} x'' &= \cos\phi\, x' - \sin\phi\, y' \\ y'' &= \sin\phi\, x' + \cos\phi\, y' \\ z'' &= z' \end{aligned} \tag{5.22}$$

Reasoning as in (5.8), we conclude that

$$U_z(\phi + \theta) = U_z(\phi)\, U_z(\theta) \tag{5.23}$$

Let us write the solution of this equation in the form

$$U_z(\theta) = \mathrm{e}^{\frac{i}{\hbar}\,\theta\, L_z} \tag{5.24}$$

where $L_z$ is *hermitian.* Taking the limit of (5.21) for $\theta \to 0$, we obtain

$$\begin{aligned}
\hat{x} - \theta\,\hat{y} + O(\theta^2) &= \hat{x} + \frac{i}{\hbar}\left[L_z, \hat{x}\right]\theta + O(\theta^2) \\
\hat{y} + \theta\,\hat{x} + O(\theta^2) &= \hat{y} + \frac{i}{\hbar}\left[L_z, \hat{y}\right]\theta + O(\theta^2) \\
\hat{z} &= \hat{z} + \frac{i}{\hbar}\left[L_z, \hat{z}\right]\theta + O(\theta^2)
\end{aligned} \tag{5.25}$$

Therefore $L_z$ must satisfy the equations

$$-\hat{y} = \frac{i}{\hbar}\left[L_z, \hat{x}\right] \qquad +\hat{x} = \frac{i}{\hbar}\left[L_z, \hat{y}\right] \qquad 0 = \frac{i}{\hbar}\left[L_z, \hat{z}\right] \tag{5.26}$$

By considering the transformation properties of the $p^i$'s instead than those of the $x^i$'s

$$\begin{aligned}
\hat{p}'_x &= \cos\theta\,\hat{p}_x - \sin\theta\,\hat{p}_y \\
\hat{p}'_y &= \sin\theta\,\hat{p}_x + \cos\theta\,\hat{p}_y \\
\hat{p}'_z &= \hat{p}_z
\end{aligned} \tag{5.27}$$

we arrive in a similar way to[18]

$$-\hat{p}_y = \tfrac{i}{\hbar}\left[L_z, \hat{p}_x\right] \qquad +\hat{p}_x = \tfrac{i}{\hbar}\left[L_z, \hat{p}_y\right] \qquad 0 = \tfrac{i}{\hbar}\left[L_z, \hat{p}_z\right] \tag{5.28}$$

*One* solution of both (5.26) and (5.28) is

$$L_z = \hat{x}\,\hat{p}_y - \hat{y}\,\hat{p}_x \tag{5.29}$$

Repeating this discussion for rotations around the $x$ and the $y$ axis, we arrive to the conclusions that those are implemented by unitary transformations $U_x(\theta)$ and $U_y(\theta)$ given by

$$\begin{aligned}
U_x(\theta) &= \mathrm{e}^{\frac{i}{\hbar}\,\theta\, L_x} \\
U_y(\theta) &= \mathrm{e}^{\frac{i}{\hbar}\,\theta\, L_y}
\end{aligned}$$

---

[18] We can derive similar commutation relations for all observables which are *vectors*, i.e. which transform under rotation as in (5.27).

where $L_x$ and $L_y$ satisfy

$$\begin{array}{lll} 0 = \frac{i}{\hbar}\left[L_x, \hat{x}\right] & -\hat{z} = \frac{i}{\hbar}\left[L_x, \hat{y}\right] & \hat{y} = \frac{i}{\hbar}\left[L_x, \hat{z}\right] \\ \hat{z} = \frac{i}{\hbar}\left[L_y, \hat{x}\right] & 0 = \frac{i}{\hbar}\left[L_y, \hat{y}\right] & -\hat{x} = \frac{i}{\hbar}\left[L_y, \hat{z}\right] \\ 0 = \frac{i}{\hbar}\left[L_x, \hat{p}_x\right] & -\hat{p}_z = \frac{i}{\hbar}\left[L_x, \hat{p}_y\right] & \hat{p}_y = \frac{i}{\hbar}\left[L_x, \hat{p}_z\right] \\ \hat{p}_z = \frac{i}{\hbar}\left[L_y, \hat{p}_x\right] & 0 = \frac{i}{\hbar}\left[L_y, \hat{p}_y\right] & -\hat{p}_x = \frac{i}{\hbar}\left[L_y, \hat{p}_z\right] \end{array} \tag{5.30}$$

These equations admit solutions analogous to (5.29),

$$L_x = \hat{y}\,\hat{p}_z - \hat{z}\,\hat{p}_y \qquad L_y = \hat{z}\,\hat{p}_x - \hat{x}\,\hat{p}_z \tag{5.31}$$

Eqs. (5.29) and (5.31) can be rewritten in a more compact way as

$$L_i = \sum_{j,k} \epsilon_{ijk}\,\hat{x}^j\,\hat{p}^j \tag{5.32}$$

We recognize the expression for the *angular momentum* of a non-relativistic 3-dimensional particle.

**Problem 1.** Show that the operators $L_i$ defined in (5.32) satisfy

$$\left[L_i, L_j\right] = i\,\hbar \sum_k \epsilon_{ijk}\,L_k \tag{5.33}$$

**Problem 2.** Consider the infinitesimal rotations $R_i(\theta)$ around the $x^i$-axis

$$R_i(\theta) = \mathbb{I}_3 + i\,\theta\,\mathcal{L}_i + O(\theta^2) \tag{5.34}$$

where $\mathcal{L}_i$ are $3 \times 3$ matrices. Show that

$$\left[\mathcal{L}_i, \mathcal{L}_j\right] = i \sum_k \epsilon_{ijk}\,\mathcal{L}_k \tag{5.35}$$

The commutation relations (5.35) are called the commutation relations of the rotation group. The fact that the operators $\frac{L_i}{\hbar}$ satisfy these same commutation relations is a consistency requirement that can be obtained in a manner completely analogous to (5.23) by considering the effect on the observables $x^i$ and $p^i$ of *two* infinitesimal rotations around *different* axis[19].

Let us now pause to observe the following. We explicitly verified ( in Problem 1) that the expressions (5.32) for $L_i$ do act correctly on the $\hat{x}^i$'s and the $\hat{p}^i$'s, as required by Eqs. (5.26), (5.28) and (5.30). On the other hand, since the latter equations involve commutators, the addition to the $L_i$'s of operators $S_i$ which *commute* with both the $\hat{x}^i$'s and the $\hat{p}^i$'s gives new solutions of the same equations. In other words the $J_i$ defined by

$$J_i = L_i + S_i \tag{5.36}$$

with

$$\left[S_i, \hat{x}^j\right] = \left[S_i, \hat{p}^j\right] = 0 \tag{5.37}$$

have the same commutators with the $\hat{x}^i$ and $\hat{p}^i$ as the $L_i$'s, and therefore they also satisfy Eqs. (5.26), (5.28) and (5.30).

The $S_i$, on the other hand, cannot be arbitrary. Consistency requires that $J_i$'s also satisfy the commutation relations of the rotation group (5.33)

$$[J_i, J_j] = i\,\hbar \sum_k \epsilon_{ijk}\, J_k \tag{5.38}$$

This implies as well that

$$\left[S_i, S_j\right] = i\,\hbar \sum_k \epsilon_{ijk}\, S_k \tag{5.39}$$

We said earlier that both the $\hat{x}^i$'s and the $\hat{p}^i$'s form a complete set of observables for non-relativistic *spinless* particle. In this case there is no non-trivial $S_i$ that we can consider in (5.36): the generators of rotations of spinless particles are given by the $L_i$'s. The latters are called the *orbital* angular momenta.

It turns out, however, that there exist particles for which neither the $\hat{x}^i$'s nor the $\hat{p}^i$ form a complete set of observables. The states of these particles are specified by introducing extra observables $S_i$, called *spin* variables, which behave as angular momenta. We have seen that, for consistency, such operators must satisfy both (5.37) and (5.39). Thus, for particles with spin, the operators

---

[19] We refer to the bibliography for an explicit proof of this statement.

$J_i$ implementing the changes of description associated to rotations are the sum of the orbital and the spin angular generators, as indicated in (5.36).

Let $\hat{H}$ and $\hat{H}'$ be the operators associated to the energy of the same physical system in two frames rotated, with respect to each other, by an angle $\theta$ around the $x^i$ axis

$$\hat{H}' = U_i(\theta)\hat{H}\,U_i(\theta) \tag{5.40}$$

We will say that the system in question is *invariant* — or *symmetric* — under a rotation of $\theta$ around the $x^i$ axis if

$$\hat{H}' = \hat{H} \tag{5.41}$$

We will also say that the system is *invariant* by rotations if $\hat{H}' = \hat{H}$ for *any* $U_i(\theta)$. In this case we can take $\theta \to 0$, and the condition (5.41) becomes equivalent to the vanishing of the commutator:

$$\big[J_i, \hat{H}\big] = 0 \tag{5.42}$$

We see that a system is invariant under rotations if and only if its Hamiltonian commutes with the generator $J_i$.

## 5.3 Representations of the angular momentum

We have seen in the previous subsection that the components of the angular momentum are hermitian operators which satisfy the commutation relations

$$\big[J_i, J_j\big] = i\,\hbar \sum_k \epsilon_{ijk}\, J_k \tag{5.43}$$

Therefore $J_i$ are not all simultaneously diagonalizable. Consider however the operator

$$\vec{J}^2 = \sum_i J_i^2 \tag{5.44}$$

**Problem 3.** Show that

$$\big[J_i, \vec{J}^2\big] = 0 \tag{5.45}$$

It follows that any one of the angular momentum components and $\vec{J}^2$ are simultaneously diagonalizable. Hence there exists a basis of simultaneous eigenvectors of, say, $J_3$ and $\vec{J}^2$. The set of compatible observables $\{J_3, \vec{J}^2\}$ is not necessarily complete: this means that the simultaneous eigenvectors of $J_3$ and $\vec{J}^2$ might be degenerate.

Let us therefore suppose to add to $J_3$ and $\vec{J}^2$ enough commuting observables, collectively denoted by $\{\mathcal{O}\}$, to form a complete compatible set of observables $\{J_3, \vec{J}^2, \mathcal{O}\}$:

$$\big[J_3, \mathcal{O}\big] = \big[\vec{J}^2, \mathcal{O}\big] = 0 \tag{5.46}$$

In this subsection we will make the *additional* assumption that the $\mathcal{O}$'s not only commute with $J_3$ and $\vec{J}^2$ but also commute with $J_1$ and $J_2$:

$$\big[J_i, \mathcal{O}\big] = 0, \qquad\qquad i = 1, 2, 3 \tag{5.47}$$

Of course Eq. (5.47) implies Eq. (5.46) but the reverse is not true. In the applications we have in mind, among the $\{\mathcal{O}\}$ there is the hamiltonian $\hat{H}$: in these cases Eq. (5.47) means that we are considering systems invariant under (all) rotations (See Eq. (5.42)).

Let us denote by $|\mu^2, m, \lambda\rangle$ the vectors of the *orthonormal* basis of simultaneous eigenvectors of $\{\vec{J}^2, J_3, \mathcal{O}\}$:

$$\begin{aligned}
&\vec{J}^2\,|\mu^2, m, \lambda\rangle = \hbar^2\,\mu^2\,|\mu^2, m, \lambda\rangle \\
&J_3\,|\mu^2, m, \lambda\rangle = \hbar\, m\,|\mu^2, m, \lambda\rangle \\
&\mathcal{O}\,|\mu^2, J_3, \lambda\rangle = \lambda\,|\mu^2, m, \lambda\rangle
\end{aligned} \tag{5.48}$$

We wrote the eigenvalues of $\vec{J}^2$ as $\hbar^2\,\mu^2$ and those of $J_3$ as $\hbar\, m$ so that $\mu^2$ and $m$ are dimensionless numbers.

Since all the $J_i$'s commute with both $\vec{J}^2$ and $\mathcal{O}$, when $J_i$ acts on $|\mu^2, J_3, \lambda\rangle$ it produces linear combinations of eigenvectors with the *same* value of $\lambda$ and $\mu^2$. Define the linear combinations of $J_1$ and $J_2$

$$J_+ \equiv J_1 + i\, J_2 \qquad J_- \equiv J_1 - i\, J_2 = \big(J_+\big)^\dagger \tag{5.49}$$

The commutation relations (5.43) rewrite in terms of $J_\pm$ and $J_3$ as follows

$$\begin{aligned}
&\big[J_3, J_\pm\big] = \pm\hbar J_\pm \\
&\big[J_+, J_-\big] = 2\,\hbar\, J_3
\end{aligned} \tag{5.50}$$

Moreover

$$\vec{J}^2 = J_+\, J_- + J_3^2 - \hbar\, J_3 = J_-\, J_+ + J_3^2 + \hbar\, J_3 \tag{5.51}$$

Therefore[20]:

$$\begin{aligned} J_+\, |\mu^2, m, \lambda\rangle &= C^{(+)}_{\mu,m,\lambda}\, |\mu^2, m+1, \lambda\rangle \\ J_-\, |\mu^2, m, \lambda\rangle &= C^{(-)}_{\mu,m,\lambda}\, |\mu^2, m-1, \lambda\rangle \end{aligned} \tag{5.52}$$

where $C^{(\pm)}_{\mu,m,\lambda}$ are numbers, possibly vanishing.

We see that $J_+$ ($J_-$) raises (lowers) by one the the eigenvalue of $J_3$, similarly to the creation and annihilation operators associated to the harmonic oscillator. We want to know if and when, by repeatedly acting with $J_\pm$ on a given eigenstate of $J_3$, we will eventually obtain a null vector.

Note that for any vector $v$

$$\langle v|\vec{J}^2 - J_3^3|v\rangle = \langle v|J_1^2 + J_2^2|v\rangle \geq 0 \tag{5.53}$$

Therefore

$$\langle \mu^2, m, \lambda|\vec{J}^2 - J_3^3|\mu^2, m, \lambda\rangle = \hbar^2\,(\mu^2 - m^2) \geq 0 \tag{5.54}$$

and hence

$$|m| \leq \mu \tag{5.55}$$

This implies that there exist both a $m_{max}$ and a $m_{min}$ such that

$$\begin{aligned} J_+|\mu^2, m_{max}, \lambda\rangle &= 0 \\ J_-|\mu^2, m_{min}, \lambda\rangle &= 0 \end{aligned} \tag{5.56}$$

On the other hand the norms of the vectors in (5.52) are, remembering (5.51),

$$\begin{aligned} \langle \mu^2, m, \lambda|J_-\, J_+|\mu^2, m, \lambda\rangle &= \left|C^{(+)}_{\mu^2,m,\lambda}\right|^2 = \hbar^2\left(\mu^2 - m^2 - m\right) \\ \langle \mu^2, m, \lambda|J_+\, J_-|\mu^2, m, \lambda\rangle &= \left|C^{(-)}_{\mu^2,m,\lambda}\right|^2 = \hbar^2\left(\mu^2 - m^2 + m\right) \end{aligned} \tag{5.57}$$

---

[20] Compare with Eqs. (3.174), where we did a similar computation for the harmonic oscillator.

and thus

$$\begin{aligned} C^{(+)}_{\mu^2,m_{max},\lambda} &= 0 \Leftrightarrow \mu^2 = m_{max}\,(m_{max}+1) \\ C^{(-)}_{\mu^2,m_{min},\lambda} &= 0 \Leftrightarrow \mu^2 = m_{min}\,(m_{min}-1) \end{aligned} \tag{5.58}$$

The solutions of the equation

$$m_{max}\,(m_{max}+1) = m_{min}\,(m_{min}-1) \tag{5.59}$$

are

$$m_{max} = -m_{min} \qquad m_{max} = m_{min} - 1 \tag{5.60}$$

The second solution is not acceptable, since $m_{max} \geq m_{min}$. The eigenvalues $m$ are hence

$$-m_{max}, -m_{max}+1, -m_{max}+2, \ldots, m_{max}-1, m_{max}+1 \tag{5.61}$$

This requires that

$$2\,m_{max}+1 = \text{positive integer} \Leftrightarrow m_{max} = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots \tag{5.62}$$

Usually one writes $m_{max} = J$ and the eigenvalues of $\vec{J}^2$ become

$$\mu^2 = J(J+1) \qquad m = -J, -J+1, \ldots, J-1, J \tag{5.63}$$

Since the value of $\mu^2 = J\,(J+1)$ is determined uniquely by the value of $J$, the eigenvectors of the basis $|\mu^2, m, \lambda\rangle$ are often denoted by $|J, m, \lambda\rangle$. From Eqs. (5.57) we obtain the action of $J_\pm$ on them

$$\begin{aligned} J_+\,|J,m,\lambda\rangle &= \hbar\,\sqrt{J\,(J+1) - m\,(m+1)}\,|J, m+1, \lambda\rangle \\ J_-\,|J,m,\lambda\rangle &= \hbar\,\sqrt{J\,(J+1) - m\,(m-1)}\,|J, m-1, \lambda\rangle \end{aligned} \tag{5.64}$$

Summarizing, the eigenspace with fixed values of $\vec{J}^2 = \hbar^2\,J\,(J+1)$ and $\lambda$ has dimension $2\,J+1$ with $J$ either positive integer or positive half-integer. These $2\,J+1$ eigenvectors, labelled by the values of $J_3 = \hbar\,m$ with $m = -J, -J+1, \ldots, J-1, J$, form a so-called *multiplet* or *irreducible representation* of the angular momentum.

It can be seen that if $\vec{J} = \vec{L}$, where $\vec{L}$ is the *orbital* angular momentum defined in (5.32), $J$ only takes *integer* eigenvalues. The spin variables $\vec{S}$, which are required to describe 3-dimensional non-relativistic particles with spin, give rise instead to representations with both integer and half-integer values for $J$.

**Problem 4.** A spin 1 system is in the state with $S_z = +1$. Find the probabilities that the measurement of $\hat{S}_x$ gives the values $S_x = -1, 0, 1$.

## 5.4 Composition of angular momenta

Let us consider a system composed by 2 particles. Let $\vec{J}_1$ and $\vec{J}_2$ be the angular momentum operators of the each particle. $\vec{J}_1$ and $\vec{J}_2$ commute with each other, since they are dynamical variables relative to different particles

$$\left[J_i^{(1)}, J_j^{(2)}\right] = 0 \tag{5.65}$$

Consider now the sum of $\vec{J}_1$ and $\vec{J}_2$,

$$\vec{J} = \vec{J}_1 + \vec{J}_2 \tag{5.66}$$

which is called the *total* angular momentum of the system. We have

$$\left[J_i, J_j\right] = i\,\hbar\,\epsilon_{ijk}\,J_k \tag{5.67}$$

Let $|J_1, m_1, \lambda_1; J_2, m_2, \lambda_2\rangle$ be the basis of the eigenvectors of the composite system. For *fixed* $\lambda_1$, $J_1$, $\lambda_2$ and $J_2$ these vectors span a subspace $\mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2}$ of dimension:

$$(2J_1 + 1) \times (2\,J_2 + 1) = \dim \mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2} \tag{5.68}$$

When the components $J_i$ of the total angular momentum act on vectors in $\mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2}$ they give vectors which are still in $\mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2}$: indeed, neither $\vec{J}_1$ nor $\vec{J}_2$ change the values of $\lambda_1$, $J_1$, $\lambda_2$ and $J_2$:

$$J_i :\ \mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2} \to \mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2} \tag{5.69}$$

Since the $J_i$'s satisfy the commutation rules of the angular momentum, $\vec{J}^2$ and $J_3$ can be simultaneously diagonalized on $\mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2}$. Let us denote by $|J, m, \alpha\rangle$ the basis of $\mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2}$ made of simultaneous eigenvectors of $\vec{J}^2$ and $J_3$. The index $\alpha$ denotes the *a priori* possible degeneration of the eigenspace with fixed $J$ and $m$. We would like to determine what are the possible values of $J$ and their possible degeneracies.

Let us consider the vector

$$|J_1, J_1, \lambda_1; J_2, J_2, \lambda_2\rangle \in \mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2} \tag{5.70}$$

This vector is an eigenvector of $J_3$ with eigenvalue

$$m = J_1 + J_2 \tag{5.71}$$

Therefore it must belong to a multiplet with $J \geq J_1 + J_2$. On the other hand, if $J > J_1 + J_2$ there should be a state in $\mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2}$ with $m > J_1 + J_2$. But the maximal value of $m$ on $\mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2}$ is $J_1 + J_2$ since $m = m_1 + m_2$ and all states have $m_1 \leq J_1$ and $m_2 \leq J_2$. Therefore the multiplet to which (5.70) belongs has

$$J = J_1 + J_2 \tag{5.72}$$

This is also the maximal value of $J$ which appears in $\mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2}$. Moreover there are no other multiplets with this same value of $J$. In fact, if there were such a multiplet, there would exist another state with $m = J_1 + J_2$ orthogonal to (5.70). But (5.70) is the *only* state in $\mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2}$ with $m = J_1 + J_2$. Hence the eigenvectors $|J_1 + J_2, m; \lambda_1, \lambda_2\rangle$, with $m = -J_1 - J_1, \ldots J_1 + J_2$ are non-degenerate.

Consider now the subspace of $\mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2}$ with

$$m = J_1 + J_2 - 1 \tag{5.73}$$

There are two linearly independent states with this value of $m$

$$|J_1, J_1 - 1, \lambda_1; J_2, J_2, \lambda_2\rangle \quad \text{and} \quad |J_1, J_1, \lambda_1; J_2 - 1, J_2, \lambda_2\rangle \tag{5.74}$$

One linear combination of these two states must be the state $|J_1 + J_2, J_1 + J_2 - 1\rangle$, which is in the same multiplet as $|J_1 + J_2, J_1 + J_2\rangle$. The orthogonal linear combination must be a state with $J \geq J_1 + J_2 - 1$. But $J$ of this state cannot be $J_1 + J_2$ since we already said that there is only *one* multiplet with $J = J_1 + J_2$. Hence the orthogonal linear combination is a state with

$$J = J_1 + J_2 - 1 \tag{5.75}$$

Again, there are no other multiplets with this value of $J$ in $\mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2}$. If there were, there would be more than two states with $m = J_1 + J_2 - 1$. Hence the eigenvectors $|J_1 + J_2 - 1, m\rangle$ are non-degenerate.

Continuing reasoning in this way we see that the values of $J$ which appear in $\mathcal{H}_{J_1,\lambda_1;J_2,\lambda_2}$ are

$$J = |J_1 - J_2|, |J_1 - J_2| + 1, \ldots, J_1 + J_2 - 1, J_1 + J_2 \tag{5.76}$$

and each one of those values appears only once: in other words the

$$|J, m; \lambda_1, \lambda_2\rangle \tag{5.77}$$

with $\lambda_1$ and $\lambda_2$ fixed, are non-degenerate.

**Problem 5.** Show that

$$\sum_{J=|J_1-J_2|}^{J=J_1+J_2} (2\,J+1) = (2J_1+1) \times (2\,J_2+1) \tag{5.78}$$

## 5.5 Angular momentum in the Schrödinger representation

Consider the expression for $\vec{L}^2$, the square of the orbital angular momentum:

$$\begin{aligned}\vec{L}^2 &= \epsilon_{ijk}\,\hat{x}^j\,\hat{p}^k\,\epsilon_{imn}\,\hat{x}^m\,\hat{p}^n = \big(\delta_{jm}\,\delta_{kn} - \delta_{jn}\,\delta_{km}\big)\,\hat{x}^j\,\hat{p}^k\,\hat{x}^m\,\hat{p}^n = \\ &= \hat{x}^j\,\hat{p}^k\,\hat{x}^j\,\hat{p}^k - \hat{x}^j\,\hat{p}^k\,\hat{x}^k\,\hat{p}^j = \\ &= \hat{\vec{x}}^2\,\hat{\vec{p}}^2 - i\,\hbar\,\hat{x}^j\,\hat{p}^j - \big(\hat{x}^k\,\hat{p}^k\big)^2 + 2\,i\,\hbar\,\hat{x}^j\,\hat{p}^j = \\ &= \hat{\vec{x}}^2\,\hat{\vec{p}}^2 - \big(\hat{\vec{x}}\cdot\hat{\vec{p}}\big)^2 + i\,\hbar\,\hat{\vec{x}}\cdot\hat{\vec{p}}\end{aligned} \tag{5.79}$$

In the Schrödinger representation this becomes

$$\begin{aligned}\vec{L}^2 &= -\hbar^2\,\vec{x}^2\,\vec{\nabla}^2 + \hbar^2\,\big(\vec{x}\cdot\vec{\nabla}\big)^2 + \hbar^2\,\vec{x}\cdot\,\vec{\nabla} = \\ &= \hbar^2\,\big(-\vec{x}^2\,\vec{\nabla}^2 + D^2 + D\big)\end{aligned} \tag{5.80}$$

where we introduced the linear differential operator

$$D \equiv \vec{x}\cdot\vec{\nabla} = x^i\,\partial_i \tag{5.81}$$

Notice that when acting on a wave function $\psi(\vec{x})$

$$D\,\psi(\vec{x}) = r\,\frac{\partial\psi(r,\theta,\phi)}{\partial r} \tag{5.82}$$

where $(r, \theta, \phi)$ are spherical coordinates.

### 5.5.1 Spherical harmonics

Let us consider homogenous polynomial functions $P_n(x)$ of the coordinates $x^i$ of degree $n$. For example

$$\begin{aligned} P_0(x) &= c_{0;1} \\ P_1(x) &= \sum_i c_{1;i}\, x^i \\ P_2(x) &= \sum_{i,j} c_{2;ij}\, x^i\, x^j \\ P_3(x) &= \sum_{i,j,k} c_{3;ijk}\, x^i\, x^j\, x^k \\ \dots & \quad \dots \end{aligned} \tag{5.83}$$

The coefficients $c_{n;i_1\dots i_n}$ are symmetric in the indices $i_i \dots i_n$. The polynomials of given degree $n$ form a linear space, since linear combinations of homogenous polynomial of given degree are polynomials of the same degree.

We have

$$D\, P_n(x) = n\, P_n(x) \tag{5.84}$$

Therefore from Eq. (5.80) we have

$$\vec{L}^2\, P_n(x) = -\hbar^2\, \vec{x}^2 \left(\vec{\nabla}^2\, P_n(x)\right) + \hbar^2\, n\,(n+1)\, P_n(x) \tag{5.85}$$

We see that if the homogenous polynomial $P_n(x)$ is *harmonic*, i.e. if it satisfies

$$\vec{\nabla}^2\, P_n(x) = 0 \tag{5.86}$$

then it is an eigenfunction of $\vec{L}^2$ with $L = n$

$$\vec{L}^2\, P_n(x) = \hbar^2\, n\,(n+1)\, P_n(x) \tag{5.87}$$

The harmonic condition (5.86) means

$$\sum_i c_{n;iii_3\dots i_n} = 0 \tag{5.88}$$

In other words among the homogeneous polynomials the harmonic ones are those defined by symmetric tensors $c_{n;i_1,\dots,i_n}$ which are *traceless*. Harmonic polynomials of a given degree form a linear space since linear combinations of harmonic polynomials are harmonic. Let us discuss how many linearly independent harmonic polynomials we have for each $n$.

For $n = 0$ we have, of course, just one linearly independent polynomial, the constant one

$$H_0(x) = 1 \tag{5.89}$$

For $n = 1$ every polynomial is harmonic. A basis of the space of harmonic polynomials of degree one, is the following:

$$H_{1,1}(x) = x_1 \qquad H_{1,2}(x) = x_2 \qquad H_{1,3}(x) = x_3 \tag{5.90}$$

For $n = 2$ among the 6 polynomials of degree 2 there are only 5 linearly independent harmonic combinations. A basis for the harmonic polynomials of degree 2 is, for example,

$$\begin{aligned} &H_{2,1}(x) = x_1\,x_2 \qquad H_{2,2}(x) = x_2\,x_3 \qquad H_{2,1}(x) = x_3\,x_1 \\ &H_{2,1}(x) = x_1^2 - x_3^2 \qquad H_{2,1}(x) = x_2^2 - x_3^3 \end{aligned} \tag{5.91}$$

Going to $n = 3$, of the 10 polynomials of degree 2 there are 7 linearly independent harmonic combinations. A basis for harmonic polynomials with $n = 3$ is for example

$$\begin{aligned} &H_{3,1}(x) = x_1\,x_2\,x_3 \qquad H_{3,2}(x) = x_1\left(x_2^2 - x_3^2\right) \\ &H_{3,3}(x) = x_2\left(x_1^2 - x_3^2\right) \qquad H_{3,4}(x) = x_3\left(x_1^2 - x_2^2\right) \\ &H_{3,5}(x) = x_1^3 - 3\,x_1\,x_2^2 \qquad H_{3,6}(x) = x_2^3 - 3\,x_3\,x_1^2 \\ &H_{3,7}(x) = x_3^3 - 3\,x_3\,x_1^2 \end{aligned} \tag{5.92}$$

**Problem 6.** Show that the harmonic homogenous polynomials of degree $n$ form a linear space of dimension $2\,n + 1$.

We see therefore that the harmonic homogenous polynomials of degree $n$ form a representation of the angular momentum with total angular momentum $l = n$. Therefore there exists a basis of harmonic polynomials which are also eigenfunctions of $L_3$. To find this basis, let us introduce the following combinations of coordinates

$$x_\pm = x_1 \pm i\,x_2 \tag{5.93}$$

**Problem 7.** Show that harmonic polynomials of $x_\pm, x_3$ satisfy the equation

$$\Big[\frac{\partial^2}{\partial x_3^2} + 4\,\frac{\partial}{\partial x_+}\,\frac{\partial}{\partial x_-}\Big]\,H(x_+, x_-, x_3) = 0 \tag{5.94}$$

**Problem 8.** Check that

$$\begin{aligned} L_3 &= \hbar\,\big(x_+\,\partial_+ - x_-\,\partial_-\big) \\ L_\pm &= \pm\,\hbar\,\big(2\,x_3\,\partial_\mp - x_\pm\,\partial_3\big) \end{aligned} \tag{5.95}$$

The action of $L_3$ on a generic monomial of degree $n$ is, therefore:

$$L_3\,\big(x_+^{n_+}\,x_-^{n_-}\,x_3^{n-n_+-n_-}\big) = \hbar\,(n_+ - n_-)\,\big(x_+^{n_+}\,x_-^{n_-}\,x_3^{n-n_+-n_-}\big) \tag{5.96}$$

Consider, to start, the three harmonic polynomials with $n = 1$. They form a multiplet of angular momentum $l = 1$. According to (5.96), the only polynomial which has $m = +1$ is proportional to

$$H_{1;+1}(x) = x_+ \tag{5.97}$$

Let us act on this with $L_-$. Recalling (5.64), we have

$$H_{1;0}(x) = \frac{1}{\sqrt{2}}\,L_-\,H_{1;+1}(x) = -\sqrt{2}\,x_3 \tag{5.98}$$

Acting on $H_{1;0}$ again with $L_-$ we obtain

$$H_{1;-1}(x) = \frac{1}{\sqrt{2}}\,L_-\,H_{1;0}(x) = \frac{1}{\sqrt{2}}\,\big(-\sqrt{2}\,x_-\big) = -x_- \tag{5.99}$$

Let us consider the $n = 2$ polynomials. The only one with $m = +2$ is

$$H_{2;+2}(x) = x_+^2 \tag{5.100}$$

Acting on this with $L_-$ we obtain a polynomial proportional to $H_{2;1}(x)$

$$H_{2;+1}(x) = \frac{1}{2}\,L_-\,H_{2;2}(x) = \frac{1}{2}\,\big(-4\,x_3\,x^+\big) = -2\,x_3\,x_+ \tag{5.101}$$

One obtains in a similar way $H_{2,0}(x)$

$$H_{2;0}(x) = \frac{1}{\sqrt{6}}\, L_-\, H_{2;+1}(x) = \frac{1}{\sqrt{6}}\left(4\, x_3 - 2\, x_-\, x_+\right) = \\ = \sqrt{\frac{2}{3}}\left(2\, x_3^2 - x_-\, x_+\right) \tag{5.102}$$

and analogously

$$H_{2;-1}(x) = 2\, x_-\, x_3 \qquad H_{2;-2} = x_-^2 \tag{5.103}$$

**Problem 9.** Find the basis $H_{n;m}(x)$, of harmonic polynomials of degree $n = 3$ and $n = 4$ which are eigenfunctions of $L_3$ with $m$.

Summarizing, for each $l$ non-negative integer number there are harmonic homogeneous polynomials $H_{l;m}(x)$ of degree $l$ which are simultaneous eigenfunctions of $\vec{L}^2$ and $L_3$:

$$\begin{aligned} \vec{L}^2\, H_{l;m}(x) &= \hbar^2\, l\,(l+1)\, H_{l;m}(x) \\ L_3\, H_{l;m}(x) &= \hbar\, m\; H_{l;m}(x) \end{aligned} \tag{5.104}$$

Since the harmonic polynomials $H_{l;m}(x)$ are homogenous of degree $l$ we can define the *spherical harmonic* functions

$$Y_{l;m}(\theta, \phi) \equiv \frac{H_{l;m}(x)}{r^l} \tag{5.105}$$

which do not depend, when expressed in spherical coordinates, on $r$. We have so far not specified the overall normalization of the harmonic polynomials. Let us do it now, by choosing it as follows

$$\int d\phi\, d\theta\, \sin\theta \left|Y_{l;m}(\theta\, \phi)\right|^2 = 1 \tag{5.106}$$

From the expressions above, we obtain for the spherical harmonics with $l = 0$ and $l = 1$

$$\begin{aligned} Y_{0;0}(\theta, \phi) &= \frac{1}{\sqrt{4\,\pi}} \\ Y_{1;0}(\theta, \phi) &= \sqrt{\frac{3}{4\,\pi}}\,\frac{x_3}{r} = \sqrt{\frac{3}{4\,\pi}}\,\cos\theta \\ Y_{1;\pm}(\theta, \phi) &= \mp\, i\,\sqrt{\frac{3}{8\,\pi}}\,\frac{x_1 \pm i\, x_2}{r} = \mp\, i\,\sqrt{\frac{3}{8\,\pi}}\,\sin\theta\, \mathrm{e}^{\pm i\,\phi} \end{aligned} \tag{5.107}$$

## 5.6 Schrödinger equation in central potential

Consider a 3-dimensional relativistic particle in a *central potential*, i.e. in a potential which depends only on the radial coordinate $r \equiv |\vec{x}|$

$$\hat{H} = \frac{\vec{p}^2}{2\,m} + V(r) \tag{5.108}$$

This system is invariant under rotations

$$\left[L_i, \hat{H}\right] = 0 \tag{5.109}$$

Therefore $\{\hat{H}, \vec{L}^2, L_3\}$ are a compatible set of commuting observables. We will see this set is also complete. Let $|E, l, m\rangle$ be the associated basis of simultaneous eigenvectors, and $\psi_{E,l,m}(\vec{x})$ the corresponding wave-functions in the Schrödinger representation. The Schrödinger equation for the energy eigenfunctions is

$$\left[\frac{-\hbar^2\,\vec{\nabla}^2}{2\,m} + V(r)\right]\psi_{E,l,m}(\vec{x}) = E\,\psi_{E,l,m}(\vec{x}) \tag{5.110}$$

Recalling the formula we derived in Eq. (5.80),

$$-\hbar^2\,\vec{x}^2\,\vec{\nabla}^2\,\psi_{E,l,m}(\vec{x}) = \left(\vec{L}^2 - \hbar^2\,D\,(D+1)\right)\psi_{E,l,r}(\vec{x}) \tag{5.111}$$

we can rewrite the Schrödinger equation as follows

$$\left[\frac{-\hbar^2}{2\,m\,r^2}\,r\,\partial_r\,(r\,\partial_r + 1) + \frac{1}{2\,m\,r^2}\,\vec{L}^2 + V(r)\right]\psi_{E,l,m} = E\,\psi_{E,l,m} \tag{5.112}$$

or, equivalently,

$$\frac{-\hbar^2}{2\,m}\,\partial_r^2\left(r\,\psi_{E,l,m}\right) + \left[\frac{\hbar^2\,l\,(l+1)}{2\,m\,r^2} + V(r)\right]\left(r\,\psi_{E,l,m}\right) = E\left(r\,\psi_{E,l,m}\right)$$

It is convenient to define a modified wavefunction

$$\chi_{E,l,m} \equiv r\,\psi_{E,l,m} \tag{5.113}$$

and an *effective potential*

$$V_{eff}(r) \equiv V(r) + \frac{\hbar^2\,l\,(l+1)}{2\,m\,r^2} \tag{5.114}$$

Then the Schrödinger equations rewrites as follows

$$\frac{-\hbar^2}{2\,m}\,\partial_r^2\,\chi_{E,l,m} + \left[\frac{\hbar^2\,l\,(l+1)}{2\,m\,r^2} + V(r)\right]\chi_{E,l,m} = E\,\chi_{E,l,m}$$

This should be solved together with the equations stating that the $\psi_{E,l,m}$'s form a multiplet of eigenfunctions with angular momentum $l$ and $m = -l, \ldots, l$:

$$\begin{aligned} \vec{L}^2\,\psi_{E,l,m}(x) &= \hbar^2\,l\,(l+1)\,\psi_{E,l,m}(x) \\ L_3\,\psi_{E,l,m}(x) &= \hbar\,m\;\psi_{E,l,m}(x) \end{aligned} \tag{5.115}$$

We already know that solutions of these latter equations are given by the spherical harmonics

$$\psi_{E,l,m} = R_{E,l}(r)\,Y_{l;m}(\theta,\phi) \tag{5.116}$$

where $R_{E,l}(r)$ is function which depends only on the radial coordinate $r$. Defining

$$\chi_{E,l}(r) \equiv r\,R_{E,l}(r) \tag{5.117}$$

we obtain an equation for the radial wavefunction $\chi_{E,l}(r)$

$$\frac{-\hbar^2}{2\,m}\,\chi''_{E,l}(r) + V_{eff}(r)\,\chi_{E,l}(r) = E\,\chi_{E,l}(r)$$

This equation has the form of a Schrödinger equation for a unidimensional problem of a particle of mass $m$ confined on the semi-axis $r > 0$ in the potential $V(r)$. From Eq. (5.117) $\chi_{E,l}(r)$ satisfies

$$\chi_{E,l}(0) = 0 \tag{5.118}$$

Note that the norm on the radial wavefunction which is implied by the familiar Schrödinger norm

$$\int d^3\vec{x}\,\psi^*_{E,l,m}(\vec{x})\,\psi_{E',l',m'}(\vec{x}) = \delta_{E,E'}\,\delta_{l,l'}\,\delta_{m,m'} \tag{5.119}$$

is

$$\int_0^\infty dr\,\chi^*_{E,l}(r)\,\chi_{E',l'}(r) = \delta_{E,E'}\,\delta_{l,l'} \tag{5.120}$$

**Problem 10.**

Consider a 3-dimensional particle of mass $m$ moving in a harmonic potential

$$V(r) = \frac{1}{2}\, m\, \omega^2\, r^2 \tag{5.121}$$

Show that:

- a) The energy levels are given by $E_n = \hbar\, (n + \frac{3}{2})\, \omega$ with $n = 0, 1, 2, \ldots$.
- b) The ground state is non-degenerate and has $l = 0$. Find the *normalized* radial wave function $\chi_{E_0,0}(r)$ with $l = 0$.
- c) The first excited level $E_1$ has degeneracy 3 and $l = 1$. Find the *normalized* radial wave function $\chi_{E_1,1}(r)$ with $l = 1$.

(*Hint*: Since one can write the 3-dimensional hamiltonian $\hat{H} = \sum_{i=1}^{3} \hat{H}_i$ where $\hat{H}_i = \frac{p_i^2}{2\, m} + \frac{1}{2}\, m\, \omega^2\, x_i^2$ is the hamiltonian of the 1-dimensional harmonic oscillator, the energy eigenfunctions of the 3-dimensional oscillator are

$$\psi_{n_1,n_2,n_3}(\vec{x}) = \psi_{n_1}(x_1)\, \psi_{n_2}(x_2)\, \psi_{n_3}(x_3) \tag{5.122}$$

where $\psi_{n_i}(x_i)$ are the eigenfunctions of the 1-dimensional harmonic oscillators.)

## 5.7 Hydrogenoids' energy levels and eigenfunctions

For the hydrogenoids

$$V(r) = -\frac{Z\, e^2}{4\, \pi\, \epsilon_0}\, \frac{1}{r} \tag{5.123}$$

where $Z$ is the atomic number. The effective potential (5.114) is therefore

$$V_{eff}(r) = -\frac{Z\, e^2}{4\, \pi\, \epsilon_0}\, \frac{1}{r} + \frac{\hbar^2\, l\, (l+1)}{2\, m\, r^2} \tag{5.124}$$

For $r \to 0$ the "centrifugal" term dominates when $l \neq 0$, while for $r \to +\infty$ the term which survives is the Coulombian potential

$$V_{eff}(r) \approx \begin{cases} \frac{\hbar^2\, l\, (l+1)}{2\, m\, r^2} & \text{if } r \to 0,\ \ l \neq 0 \\ -\frac{Z\, e^2}{4\, \pi\, \epsilon_0}\, \frac{1}{r} & \text{if } r \to +\infty. \end{cases} \tag{5.125}$$

The minimum of the potential, for $l \neq 0$, occurs for

$$V'_{eff}(r_0) = 0 = \frac{Z\,e^2}{4\,\pi\,\epsilon_0}\,\frac{1}{r_0^2} - \frac{\hbar^2\,l\,(l+1)}{m\,r_0^3} \Leftrightarrow r_0 = a_B\,\frac{l\,(l+1)}{Z} \tag{5.126}$$

where $a_B$ is the Bohr's radius

$$a_B = \frac{4\,\pi\,\epsilon_0\,\hbar^2}{m\,e^2} \tag{5.127}$$

The value of the potential at the minimum is

$$\begin{aligned} & l \neq 0 \\ & V(r_0) = -\frac{Z\,e^2}{8\,\pi\,\epsilon_0}\,\frac{1}{r_0} = -\frac{Z^2}{l\,(l+1)}\,\frac{1}{2}\,\frac{e^4\,m}{\hbar^2\,(4\,\pi\,\epsilon_0)^2} = \frac{Z^2}{l\,(l+1)}\,E_1^H \end{aligned}$$

where

$$E_1^H = -\frac{1}{2}\,\frac{e^4\,m}{\hbar^2\,(4\,\pi\,\epsilon_0)^2} \tag{5.128}$$

is the fundamental level of the hydrogen.

According to our general discussion about the features of the energy spectrum of unidimensional problems, we anticipate that:

- The energy spectrum is continuous for $E > 0$, and non-degenerate for $l$ and $m$ *fixed.*
- The discrete levels occur in the negative range of energies $\frac{Z^2}{l\,(l+1)}\,E_1^H < E < 0$ and is non-degenerate for $l$ and $m$ *fixed.*

Let us find the simplest eigenfunction with $l = 0$. The radial Schrödinger equation for $l = 0$ is

$$\frac{\chi''_{E,0}(r)}{\chi_{E,0}(r)} = -\frac{2\,m\,E}{\hbar^2} - \frac{2\,Z}{a_B\,r} \tag{5.129}$$

It is useful to introduce the adimensional radial coordinate

$$\rho \equiv \frac{Z\,r}{a_B} \tag{5.130}$$

in terms of which the equation above writes

$$\frac{\chi''_{E,0}(\rho)}{\chi_{E,0}(\rho)} = -\frac{2\,m\,E\,a_B^2}{Z^2\,\hbar^2} - \frac{2}{\rho} = -\frac{E}{Z^2\,|E_1^H|} - \frac{2}{\rho} \tag{5.131}$$

Let us seek for a solution of the form

$$\chi_{E,0}(\rho) = \rho\,\mathrm{e}^{-c\,\rho} \qquad c > 0 \tag{5.132}$$

which satisfies the boundary condition (5.118). We have

$$\frac{\chi''_{E,0}(\rho)}{\chi_{E,0}(\rho)} = c^2 - \frac{2\,c}{\rho} = -\frac{E}{Z^2\,|E_1^H|} - \frac{2}{\rho} \tag{5.133}$$

Hence we choose

$$c = 1 \qquad -\frac{E}{Z^2\,|E_1^H|} = 1 \tag{5.134}$$

We obtain in this way the normalized ground state wave function of the hydrogenoids

$$\chi_{E,0}(r) = 2\left(\frac{Z}{a_B}\right)^{\frac{3}{2}} r\,\mathrm{e}^{-\frac{Z\,r}{a_B}} \tag{5.135}$$

and the corresponding energy level

$$E_1 = -Z^2\,|E_1^H| = -\frac{1}{2}\,\frac{Z^2\,e^4\,m}{\hbar^2\,(4\,\pi\,\epsilon_0)^2} \tag{5.136}$$

**Problem 11.** Show that (5.135) has norm one.
**Problem 12.** Find the normalized radial wave function of the first excited state of the hydrogenoid atom with $l = 0$ by looking for solutions of (5.131) of the form

$$\chi_{E,0}(\rho) = C_2\,(\rho^2 + a_1\,\rho)\,\mathrm{e}^{-c\,\rho} \qquad c > 0 \tag{5.137}$$

Let us now consider the radial Schrödinger equation for generic angular momenta $l$:

$$\chi''_{E,l}(\rho) = \chi_{E,l}(\rho)\,\Big[\frac{l\,(l+1)}{\rho^2} - \frac{2}{\rho} - \epsilon\Big] \tag{5.138}$$

where we introduced the adimensional parameters $\rho$ and $\epsilon$

$$\rho \equiv \frac{Z\, r}{a_B} \qquad E = \epsilon\, Z^2\, \frac{\hbar^2}{2\, m\, a_B^2} = \epsilon\, \frac{Z^2\, e^4\, m}{(4\, \pi\, \epsilon_0)^2\, 2\, \hbar^2} \tag{5.139}$$

**Problem 13.** Find the normalized radial wave function of the lowest energy state of the hydrogenoid atom with $l = 1$ by looking for solutions of (5.138) of the form

$$\chi_{E,1}(\rho) = \rho^2\, \mathrm{e}^{-c\, \rho} \qquad c > 0 \tag{5.140}$$

The answers of Problems 11,12 and 13 suggest to look for general solutions of (5.138) of the form

$$\chi_{E,l}(\rho) = P_n(\rho)\, \mathrm{e}^{-c\, \rho} \tag{5.141}$$

where

$$P_n(\rho) = \rho^n + a_{n-1}\, \rho^{n-1} + a_{n-2}\, \rho^{n-2} + \cdots + a_2\, \rho^2 + a_1\, \rho \tag{5.142}$$

is a polynomial of degree $n$ vanishing for $\rho = 0$, as required by the boundary condition (5.118).

Plugging this ansätz inside the Schrödinger equation (5.138) we obtain the equation

$$(c^2 + \epsilon)\, P_n(\rho) + 2\, \Big[\frac{P_n(\rho)}{\rho} - c\, P_n'(\rho)\Big] + \Big[P_n''(\rho) - \frac{l\,(l+1)}{\rho^2}\Big] = 0 \tag{5.143}$$

The l.h.s. of this equation is a polynomial in $\rho$ of degree $n$. In order for this equation to be satisfied all the coefficients of this polynomial must vanish.

Monomials of degree $n$ are contained only in the first of the three terms in the l.h.s. of Eq. (5.143): thus the equation has solution only if

$$c^2 + \epsilon = 0 \tag{5.144}$$

When this condition holds Eq. (5.143) reduces to

$$2\, \Big[\frac{P_n(\rho)}{\rho} - c\, P_n'(\rho)\Big] + \Big[P_n''(\rho) - \frac{l\,(l+1)}{\rho^2}\Big] = 0 \tag{5.145}$$

The l.h.s.of Eq. (5.145) is a polynomial of degree $n - 1$. The terms of degree $n - 1$ come from the first square bracket

$$2\, \big[\rho^{n-1} - c\, n\, \rho^{n-1} + \cdots\big] + \cdots = 0 \tag{5.146}$$

where the dots denotes terms of degree lower than $n-1$. We deduce that

$$1 - c\, n = 0 \tag{5.147}$$

In conclusion

$$c = \frac{1}{n} \qquad \epsilon = -\frac{1}{n^2} \tag{5.148}$$

We obtain in this way the celebrated formula for the Bohr energy levels

$$E_n = -\frac{1}{n^2}\,\frac{Z^2\, e^4\, m}{(4\,\pi\,\epsilon_0)^2\, 2\,\hbar^2} \tag{5.149}$$

The corresponding radial wave functions have the following form

$$\chi_{E,l}(\rho) = P_n(\rho)\, \mathrm{e}^{-\frac{\rho}{n}} \tag{5.150}$$

This formula implies that the larger the $n$ the larger the size of the corresponding bound state. This agrees with Eq. (1.158) for the radii of the quantized orbits obtained in Section 1 when we discussed the Bohr-Sommerfeld quantization of the hydrogen atom.

When Eqs. (5.148) are satisfied, equation (5.145) writes

$$\begin{aligned} &2\,\big[\frac{a_{n-1}}{n}\,\rho^{n-2} + \frac{2\,a_{n-2}}{n}\,\rho^{n-3} + \cdots + \frac{(n-2)\,a_2}{n}\,\rho + \frac{(n-1)\,a_1}{n}\,\big] + \\ &+\big[\big(n\,(n-1) - l\,(l+1)\big)\,\rho^{n-2} + \big((n-1)\,(n-2) - l\,(l+1)\big)\,a_{n-1}\,\rho^{n-3} + \cdots \\ &\quad +\big(2\cdot 1 - l\,(l+1)\big)\,a_2\big] - \frac{2\,l\,(l+1)\,a_1}{\rho} = 0 \end{aligned} \tag{5.151}$$

This is equivalent to the following algebraic equations for the coefficients of the poynomial $P_n(\rho)$

$$\begin{aligned} a_{n-1} &= \frac{n}{2}\,\big[l\,(l+1) - n\,(n-1)\big] \\ a_{n-2} &= \frac{1}{2}\,\frac{n}{2}\,\big[l\,(l+1) - (n-1)\,(n-2)\big]\,a_{n-1} \\ \cdots &= \cdots \\ a_2 &= \frac{1}{n-2}\,\frac{n}{2}\,\big[l\,(l+1) - 3\cdot 2\big]\,a_3 \\ a_1 &= \frac{1}{n-1}\,\frac{n}{2}\,\big[l\,(l+1) - 1\cdot 2\big]\,a_2 \\ l\,(l+1)\,a_1 &= 0 \end{aligned} \tag{5.152}$$

Let us discuss the properties of the solutions of these equations.

For $l = 0$ the last of Eqs. (5.152) is satisfied for any $a_1$. The other equations determine, recursively, all the coefficients $a_{n-1}, a_{n-2}, \cdots a_1$

$$P_{n,0}(\rho) = \rho^n - \frac{n^2\,(n-1)}{2}\,\rho^{n-1} + \cdots + a_1\,\rho \tag{5.153}$$

Note that all the coefficients $a_1, a_2, \cdots a_{n-1}$ are non-vanishing.

Consider now $l = 1$. In this case, the last equation of (5.152) requires

$$a_1 = 0 \tag{5.154}$$

The next equation in (5.152)

$$a_1 = \frac{1}{n-1}\,\frac{n}{2}\,\big[l\,(l+1) - 1\cdot 2\big]\,a_2 = 0 \tag{5.155}$$

is, for $l = 1$, satisfied even if $a_2 \neq 0$. The rest of the equations determine, recursively, the non-vanishing coefficients $a_2, a_3, \cdots a_{n-1}$ :

$$P_{n,1}(\rho) = \rho^n + \frac{n}{2}\,(2 - n\,(n-1))\,\rho^{n-1} + \cdots + a_2\,\rho^2 \tag{5.156}$$

This state has energy $E_n$ as the state with $l = 0$.

Going to $l = 2$, we obtain analogously

$$a_1 = a_2 = 0 \tag{5.157}$$

The remaining equations determine, recursively, the non-zero coefficients $a_3, a_4, \cdots a_{n-1}$.

For higher values of $l$, as long as $l < n$, we obtain a polynomial $P_{n,l}(\rho)$

$$P_{n,l}(\rho) = \rho^n + \frac{n}{2}\,\big[l\,(l+1) - n\,(n-1)\big]\,\rho^{n-1} + \cdots + a_{l+1}\,\rho^{l+1} \tag{5.158}$$

with

$$a_1 = a_2 = a_3 = \cdots = a_l = 0 \qquad a_{l+1} \neq 0 \tag{5.159}$$

The highest possible value of the angular momentum, $l = n - 1$, corresponds to the polynomial

$$P_{n,n-1}(\rho) = \rho^n \tag{5.160}$$

We conclude that for any fixed degree $n$, there is (up to a normalization factor) a unique polynomial solution $P_{n,l}(\rho)$ for each $l = 0, 1, 2, \cdots, n-1$. The polynomials $P_{n,l}(\rho)$, for $l = 0, 1, \ldots, n-1$ are called Laguerre polynomials. The corresponding radial wave functions

$$\chi_{E_n,l}(\rho) = P_{n,l}(\rho)\, \mathrm{e}^{-\frac{\rho}{n}} \qquad\qquad 0 \leq l \leq n-1 \qquad n = 1, 2, \cdots \tag{5.161}$$

have energy $E_n$ independent of $l$. Let us list the first few Laguerre polynomials

$$\begin{aligned}
&P_{1,0}(\rho) = 2\,\rho \\
&P_{2,0}(\rho) = \frac{1}{2^{\frac{3}{2}}}\left(\rho^2 - 2\,\rho\right) \qquad P_{2,1}(\rho) = \frac{1}{2^{\frac{3}{2}} \cdot 3^{\frac{1}{2}}}\rho^2 \\
&P_{3,0}(\rho) = \frac{2^2}{3^{\frac{9}{2}}}\left(\rho^3 - 9\,\rho^2 + \frac{27}{2}\,\rho\right) \quad P_{3,1}(\rho) = \frac{2^{\frac{3}{2}}}{3^{\frac{9}{2}}}\left(\rho^3 - 6\,\rho^2\right) \\
&P_{3,2}(\rho) = \frac{2^{\frac{3}{2}}}{3^{\frac{9}{2}} \cdot 5^{\frac{1}{2}}}\rho^3
\end{aligned} \tag{5.162}$$

The coefficients in front of the polynomials $P_{n,l}(\rho)$ have been chosen in such a way that the radial wave functions are normalized as follows

$$\int_0^\infty d\rho\, \left|\chi_{E_n,l}(\rho)\right|^2 = 1 \tag{5.163}$$

Let us observe that for small distances from the origin, that is for $\rho \ll 1$ the radial wave function has the following behavior

$$\chi_{E_n,l}(\rho) = a_{l+1}\,\rho^{l+1} + O(\rho^{l+2}) \qquad \rho \ll 1 \tag{5.164}$$

One significant consequence of this is that for a given level $n$, electrons which are in states with higher angular momentum $l$ are more distant, on average, from the center of attraction than those with lower $l$.

# 6 Atoms

The Hamiltonian of an atom with $Z$ electrons in the non-relativistic approximation is

$$\begin{aligned}\hat{H} &= \sum_{a=1}^{Z}\Big(\frac{\vec{p}_a^{\,2}}{2\,m} - \frac{Z\,e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_a|}\Big) + \sum_{a>b}\frac{e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_a-\vec{x}_b|} = \\ &= \sum_{a=1}^{Z} H_a^{(0)} + \sum_{a>b}\frac{e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_a-\vec{x}_b|} = \\ &\equiv \hat{H}_0 + \hat{H}_{repulsion} \qquad (6.1)\end{aligned}$$

We know the spectrum and the eigenvectors of the first term,

$$\hat{H}_0 = \sum_{a=1}^{Z} H_a^{(0)} \qquad (6.2)$$

The energy eigenfunctions of $\hat{H}_0$ are simply given by the product of energy eigenfunctions of the hydrogenoid-like Hamiltonians $H_a^{(0)}$:

$$|(n_1,l_1,m_1);(n_1,l_1,m_1);\cdots;(n_Z,l_Z,m_Z)\rangle \qquad (6.3)$$

and the corresponding eigenvalues are

$$E_{n_1,\cdots,n_Z} = E_{n_1} + \cdots + E_{n_Z} \qquad (6.4)$$

Neglecting the repulsion term

$$\hat{H}_{repulsion} = \sum_{a>b}\frac{e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_a-\vec{x}_b|} \qquad (6.5)$$

gives a very crude approximation to atomic levels. The approximation in which $\hat{H}_{repulsion}$ is neglected is called the independent electrons approximation (I.E.A.). In the following we discuss, in the simplest cases, ways to improve on it.

## 6.1 Fundamental level of He

We will consider the Hamiltonian

$$\hat{H} = \frac{\vec{p}_1^{\,2}}{2\,m} - \frac{Z\,e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_1|} + \frac{\vec{p}_2^{\,2}}{2\,m} - \frac{Z\,e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_2|} + \frac{e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_1-\vec{x}_2|} \qquad (6.6)$$

For the helium $Z = 2$ but it is convenient to leave $Z$ as a parameter in the hamiltonian. When $Z \neq 2$ the hamiltonian (6.6) describes ions, e.g. $H^-$ ($Z = 1$), $Li^+$, ($Z = 3$) $Be^{++}$ ($Z = 4$) and so on. The fundamental state of (6.6) in the independent electrons approximation, in which the repulsion between the two electrons is neglected, is

$$\psi_0 = |(1,0,0);(1,0,0)\rangle \tag{6.7}$$

whose energy is

$$E_0^{IEA} = 2 \times Z^2\, E_1^H = -Z^2\, \frac{e^2}{4\,\pi\,\epsilon_0\, a_B} \tag{6.8}$$

where $E_1^H = -\frac{1}{2}\,\frac{e^2}{4\,\pi\,\epsilon_0\,a_B} \approx -13.6\,eV$ is the ground energy of hydrogen. The first ionization energy is the energy $E_1$ necessary to extract one of the electrons. Since, if we extract one of the electrons we obtain the hydrogenoid whose ground state energy is $Z^2\, E_1^H$, we have

$$E_0 + E_1 = Z^2\, E_1^H \tag{6.9}$$

or equivalently, the first ionization energy, in the independent electrons approximation is

$$E_1^{IEA} = Z^2\, |E_1^H| \tag{6.10}$$

For helium we obtain

$$E_1^{IEA} \approx 4 \times 13.6\, eV = 54.4\, eV \qquad \text{He} \tag{6.11}$$

much bigger than the experimental value

$$E_1^{(exp)} \approx 24.6\, eV \qquad \text{He} \tag{6.12}$$

A first improvement on this value is to add to $E_0$ the *average* of the repulsion term $\hat{H}_{repulsion}$ on the unperturbed state $\psi_0$. This approximation is known as the *perturbative* approximation, at first order. The average of the repulsion energy on $\psi_0$ is

$$\begin{aligned}
\Delta E_0^{pert} &= \langle(1,0,0);(1,0,0)|\, \frac{e^2}{4\,\pi\,\epsilon_0\, r_{12}}\, |(1,0,0);(1,0,0)\rangle = \\
&= \int d^3\vec{x}_1 \int d^3\vec{x}_2\, \psi_{100}^*(\vec{x}_1)\, \psi_{100}^*(\vec{x}_2)\, \frac{e^2}{4\,\pi\,\epsilon_0\, r_{12}}\, \psi_{100}(\vec{x}_1)\, \psi_{100}(\vec{x}_2) = \\
&= \int d^3\vec{x}_1 \int d^3\vec{x}_2\, |\psi_{100}(\vec{x}_1)|^2\, |\psi_{100}(\vec{x}_2)|^2\, \frac{e^2}{4\,\pi\,\epsilon_0\, r_{12}} = \\
&= \int d^3\vec{x}_1 \int d^3\vec{x}_2\, \rho(\vec{x}_1)\, \rho(\vec{x}_2)\, \frac{1}{4\,\pi\,\epsilon_0\, r_{12}}
\end{aligned} \tag{6.13}$$

where

$$r_{12} = |\vec{x}_1 - \vec{x}_2| \tag{6.14}$$

is the distance between the two electrons and[21]

$$\rho(\vec{x}) = e\,|\psi_{100}(\vec{x})|^2 = e\,\frac{Z^3}{\pi\,a_B^3}\,\mathrm{e}^{-\frac{2\,Z\,r}{a_B}} \qquad r \equiv |\vec{x}| \tag{6.15}$$

is the density of electric charge associated with the electronic wave function $\psi_{100}(\vec{x})$. $\Delta E_0^{pert}$ is twice the electrostatic energy of the electric charge density (6.15). Since the total electric charge of the distribution (6.15) is $e$ and its effective size of the order of $\frac{a_B}{Z}$ we expect, on dimensional grounds, that

$$\Delta E_0^{pert} \propto \frac{Z\,e^2}{4\,\pi\,\epsilon_0\,a_B} \tag{6.16}$$

Let us compute the exact numerical coefficient in front of this expression.

The electric field created by the electric charge density (6.15) is

$$\begin{aligned} E(r) &= \frac{1}{\epsilon_0\,r^2}\int_0^r dx\,x^2\rho(x) = \frac{e\,Z^3}{\pi\,\epsilon_0\,r^2\,a_B^3}\int_0^r dx\,x^2\,\mathrm{e}^{-\frac{2\,Z\,x}{a_B}} = \\ &= \frac{e}{8\,\pi\,\epsilon_0\,r^2}\int_0^{\frac{2\,Z\,r}{a_B}} d\rho\,\rho^2\,\mathrm{e}^{-\rho} = \\ &= \frac{e}{4\,\pi\,\epsilon_0\,r^2}\Big[1 - \mathrm{e}^{-\frac{2\,Z\,r}{a_B}}\big(1 + \frac{2\,Z\,r}{a_B}\,(1 + \frac{Z\,r}{a_B})\big)\Big] \end{aligned} \tag{6.17}$$

Twice the electrostatic energy of this field is

$$\begin{aligned} 4\,\pi\,\epsilon_0\int_0^\infty dr\,r^2\,E(r)^2 &= \frac{e^2}{16\,\pi\,\epsilon_0}\int_0^\infty dr\,\frac{1}{r^2}\,\times \\ &\qquad \times\Big[2 - \mathrm{e}^{-\frac{2\,Z\,r}{a_B}}\big(2 + \frac{2\,Z\,r}{a_B}\,(2 + \frac{2\,Z\,r}{a_B})\big)\Big]^2 = \\ &= \frac{e^2\,Z}{8\,\pi\,\epsilon_0\,a_B}\int_0^\infty d\rho\,\frac{1}{\rho^2}\Big[2 - \mathrm{e}^{-\rho}\big(2 + \rho\,(2+\rho)\big)\Big]^2 = \\ &= \frac{e^2\,Z}{8\,\pi\,\epsilon_0\,a_B}\times\frac{5}{4} \end{aligned} \tag{6.18}$$

---

[21] Recall that $\psi_{100}(\vec{x}) = Y_{00}(\theta,\phi)\,R_{10}(r) = \frac{R_{10}(r)}{2\,\sqrt{\pi}}$.

Thus

$$\Delta E_0^{pert} = \frac{e^2\,Z}{4\,\pi\,\epsilon_0\,a_B} \times \frac{5}{8} = \frac{5}{4}\,Z\,|E_1^H| \tag{6.19}$$

Hence a corrected estimate of the ground state energy is

$$E_0' = E_0 + \Delta E_0 = -\frac{e^2}{4\,\pi\,\epsilon_0\,a_B} \times \big(Z^2 - \frac{5}{8}\,Z\big) = -2\,|E_1^H|\,\big(Z^2 - \frac{5}{8}\,Z\big) \tag{6.20}$$

The first ionization energy $E_1$ becomes in this perturbative approximation

$$E_1^{pert} = Z^2\,|E_1^H| - \Delta E_0 = \big(Z^2 - \frac{5}{4}\,Z\big)\,|E_1^H| \tag{6.21}$$

For the Helium, this gives $E_1 \approx 13.6 \times \frac{3}{2}\,eV = 20.4\,eV$ in better agreement with the experimental value.

### 6.1.1 Variational methods

Another approach to the same problem is to replace the approximated wave function of the two-electron system

$$\psi_{100}(\vec{x}_1)\,\psi_{100}(\vec{x}_2) \tag{6.22}$$

with the same functions in which $Z$ is replaced by a parameter $Z'$ which will be determined by minimizing the total average energy. Thus one wants to compute the function of $Z'$

$$\begin{aligned} E(Z') \equiv \int d^3\vec{x}_1 \int d^3\vec{x}_2\,\psi^*_{100}(\vec{x}_2; Z')\,\psi^*_{100}(\vec{x}_2; Z') \times \\ \times \hat{H}_Z\left[\psi_{100}(\vec{x}_1; Z')\,\psi_{100}(\vec{x}_2; Z')\right] = \\ \equiv \langle \hat{H}_Z \rangle_{Z'} \end{aligned} \tag{6.23}$$

where we denoted by $\psi_{100}(\vec{x}; Z')$ the wave function of the ground state of the hydrogenoid with atomic number $Z'$, different than the $Z$ which appears in the hamiltonian $\hat{H}_Z$. We have

$$\begin{aligned}\hat{H}_Z = \frac{\vec{p}_1^2}{2\,m} - \frac{Z'\,e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_1|} + \frac{\vec{p}_2^2}{2\,m} - \frac{Z'\,e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_2|} + \\ + \frac{(Z'-Z)\,e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_1|} + \frac{(Z'-Z)\,e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_2|} + \frac{e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_1-\vec{x}_2|} = \\ = \hat{H}_{Z'} + \frac{(Z'-Z)\,e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_1|} + \frac{(Z'-Z)\,e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_2|} = \\ = \hat{H}^{(0)}_{Z'} + \hat{H}_{repulsion} + \frac{(Z'-Z)\,e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_1|} + \frac{(Z'-Z)\,e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_2|} \end{aligned} \tag{6.24}$$

The average of $\hat{H}^{(0)}_{Z'}$ on the ground state $\psi_{100}(\vec{x}_1;Z')\,\psi_{100}(\vec{x}_2;Z')$ is of course the same as (6.8) with $Z$ replaced by $Z'$:

$$\langle\hat{H}^{(0)}_{Z'}\rangle_{Z'} = -2\,(Z')^2\,|E_1^H| \tag{6.25}$$

The average of the repulsion term $\hat{H}_{repulsion}$ is analogously obtained from (6.19) by means of the replacement $Z \to Z'$

$$\langle\hat{H}_{repulsion}\rangle_{Z'} = \frac{5}{4}\,Z'\,|E_1^H| \tag{6.26}$$

The average of the last two terms in (6.24) both involve the same integral[22]

$$\begin{aligned}\int d^3\vec{x}_1 \int d^3\vec{x}_2\,|\psi_{100}(\vec{x}_1;Z')|^2\,|\psi_{100}(\vec{x}_2;Z')|^2\,\frac{e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_1|} = \\ = \int d^3\vec{x}_1\,|\psi_{100}(\vec{x}_1;Z')|^2\,\frac{e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_1|} = \\ = \frac{e^2\,(Z')^3}{\epsilon_0\,\pi\,a_B^3}\int_0^\infty dr\,r\,\mathrm{e}^{-\frac{2\,Z'\,r}{a_B}} = \frac{e^2\,Z'}{4\,\pi\,\epsilon_0\,a_B}\int_0^\infty d\rho\,\rho\,\mathrm{e}^{-\rho} = \\ = \frac{e^2\,Z'}{4\,\pi\,\epsilon_0\,a_B} = -2\,Z'\,E_1^H\end{aligned} \tag{6.27}$$

---

[22] It is possible to obtain this result without performing explicitly the integral, by exploiting the fact that $-\frac{e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_1|} = \partial_Z\,H^{(0)}_Z$ and the general result

$$\langle\psi^E_\lambda, \partial_\lambda\,H_\lambda\,\psi^E_\lambda\rangle = \partial_\lambda E(\lambda),$$

where $H_\lambda$ is an Hamiltonian dependent on a parameter $\lambda$, $\psi^E_\lambda$ is its ($\lambda$-dependent) eigenfunction with eigenvalue $E(\lambda)$. Applying this formula to the case at hand we obtain $\langle\frac{e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_1|}\rangle_Z = -\partial_Z\left(Z^2\,E_1^H\right) = -2\,Z\,E_1^H$, in agreement with the explicit computation (6.27).

Hence

$$E(Z') = -2\,(Z')^2\,|E_1^H| + \frac{5}{4}\,Z'\,|E_1^H| + 4\,Z'\,(Z' - Z)\,|E_1^H| \quad (6.28)$$

Minimizing this function with respect to $Z'$ with obtain

$$\frac{5}{4} + 4\,(Z' - Z) = 0 \Leftrightarrow Z' = Z - \frac{5}{16} \quad (6.29)$$

from which[23]

$$E(Z')\big|_{Z'=Z-\frac{5}{16}} = 2\,E_1^H\,(Z-\frac{5}{16})^2 \tag{6.30}$$

---

[23] The fact that the final result for the variational energy takes the form of a perfect square is not an accident. It is due to the fact that energy at the lowest order is $E_Z^{(0)} \propto Z^2$ and that the first order correction $\Delta E_0^{pert} \propto Z$. Let us explain why. Suppose $H_0(\lambda)$ is a $\lambda$-dependent hamiltonian whose eigenvector $\psi_\lambda^E$ has eigenvalue $E(\lambda)$. Consider the perturbed hamiltonian $\tilde{H} = H_0(\lambda) + \hat{V}$. As we have seen, the eigenvalue $\tilde{E}_\lambda$ of $\tilde{H}$ corresponding to $E(\lambda)$ is, to first order,

$$\tilde{E}(\lambda) = E(\lambda) + V(\lambda) + \cdots$$

where $V(\lambda) = \langle \psi_\lambda, \hat{V}\,\psi_\lambda\rangle \equiv \langle \hat{V}\rangle_\lambda$ is the average of the perturbation on the unperturbed eigenvector $\psi_\lambda^E$. A variational computation of the energy level $\tilde{E}(\lambda)$ which uses the wave functions $\psi_{\lambda'}^E$ with $\lambda' \neq \lambda$ gives

$$\langle \tilde{H}\rangle_{\lambda'} = \langle H_0(\lambda)\rangle_{\lambda'} + V(\lambda') = E(\lambda') + V(\lambda') + \int_{\lambda'}^{\lambda} d\eta \langle \partial_\eta H_0(\eta)\rangle_{\lambda'}$$

Assume that $H_0(\lambda)$ is *linear* in $\lambda$, so that $\partial_\lambda^2 H_0(\lambda) = 0$. Then

$$\langle \tilde{H}\rangle_{\lambda'} = E(\lambda') + V(\lambda') - (\lambda' - \lambda)\,E'(\lambda')$$

Extremizing with respect to $\lambda'$ we obtain

$$\lambda' - \lambda = \frac{V'(\lambda')}{E''(\lambda')}$$

The variational estimate for the energy level is therefore

$$\tilde{E}(\lambda) = E(\lambda') + V(\lambda') - \frac{V'(\lambda')\;E'(\lambda')}{E''(\lambda')}.$$

In general this expression contains terms of any order in $V$. However for $E(\lambda) = E_0\lambda^2$ and $V(\lambda) = V_0\,\lambda$ it gives

$$\tilde{E}(\lambda) = E(\lambda') = E(\lambda + \frac{V_0}{2\,E_0}) = E_0\,\big(\lambda + \frac{V_0}{2\,E_0}\big)^2$$

in agreement with the result (6.30). The variational estimate for the energy, expanded to second order in $V$, is in the generic case

$$\tilde{E}(\lambda) = E(\lambda) + V(\lambda) + \frac{(V'(\lambda))^2}{2\,E''(\lambda)} + \cdots$$

The first ionization energy in this variational approximation is therefore

$$E_1^{var} = \big(2\,(Z-\frac{5}{16})^2\ - Z^2\big)\,|E_1^H| = \big(Z^2 - \frac{5}{4}\,Z + \frac{25}{128}\big)\,|E_1^H| \tag{6.31}$$

For He, this gives

$$E_1^{var} \approx 23.1\,eV \qquad \text{He} \tag{6.32}$$

The agreement with the experimental value (6.12) is further improved with respect to the perturbative estimate (6.21)

## 6.2 Helium excited states: Pauli exclusion principle

Let us consider the excited states of helium in the independent electrons approximation. Among these states consider the ones in which only *one* of the two electron is excited. Given the level to which one of the two electrons is excited, there two independent vectors which are degenerate in the independent electrons approximation:

$$\psi^{(1)}_{n,l,m} = |(1,0,0);(n,l,m)\rangle \qquad \psi^{(2)}_{n,l,m} = |(n,l,m);(1,0,0)\rangle \quad n>1 \tag{6.33}$$

Consider the matrix elements of the Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{H}_{repulsion} = \hat{H}_0 + \frac{e^2}{4\,\pi\,\epsilon_0\,|\vec{x}_1 - \vec{x}_2|} \tag{6.34}$$

between the states (6.33)

$$\begin{aligned} H_{ab} = \langle\psi^{(a)}_{n,l,m}, \hat{H}\,\psi^{(b)}_{n,l,m}\rangle = E_0\,\delta_{ab} + \langle\psi^{(a)}_{n,l,m}, \hat{H}_{repulsion}\,\psi^{(b)}_{n,l,m}\rangle = \\ = E_0\,\delta_{ab} + H^{(rep)}_{ab} \qquad a,b = 1,2 \end{aligned} \tag{6.35}$$

where $E_0$ is the energy of the states $\psi^{(1)}_{n,l,m}$ in the independent electron approximation

$$E_0 = -4\,\big(1 + \frac{1}{n^2}\big)\,|E_1^H| \tag{6.36}$$

It is clear that, since $\hat{H}_{repulsion}$ is invariant under the exchange of the coordinates of the two electrons, we have the following relations between the matrix elements $H^{(rep)}_{ab}$

$$H^{(rep)}_{11} = H^{(rep)}_{22} \equiv J \qquad H^{(rep)}_{12} = H^{(rep)}_{21} \equiv K \tag{6.37}$$

$J$, which is called the *direct integral* is an obviously positive number

$$\begin{aligned} J &= \int d^3\vec{x}_1 \int d^3\vec{x}_2\, \frac{\rho_{100}(\vec{x}_1)\, \rho_{nlm}(\vec{x}_2)}{4\,\pi\,\epsilon_0\, r_{12}} \\ \rho_{n,l,m}(\vec{x}) &\equiv e\, |\psi_{nlm}(\vec{x})|^2 \end{aligned} \tag{6.38}$$

while $K$ is the so-called *exchange* integral

$$\begin{aligned} K &= \int d^3\vec{x}_1 \int d^3\vec{x}_2\, \frac{\rho_{ex}(\vec{x}_1)\, \rho^*_{ex}(\vec{x}_2)}{4\,\pi\,\epsilon_0\, r_{12}} \\ \rho_{ex}(\vec{x}_1) &\equiv e\, \psi^*_{100}(\vec{x})\, \psi_{nlm}(\vec{x}) \end{aligned} \tag{6.39}$$

This is twice the electrostatic energy of the complex valued charge density $\rho_{ex}(\vec{x})$: standard electrostatic arguments show that this integral is positive too, like $J$:

$$J > 0 \tag{6.40}$$

Therefore the $2 \times 2$ matrix $H_{ab}$

$$H = \begin{pmatrix} E_0 + J & K \\ K & E_0 + J \end{pmatrix} \tag{6.41}$$

has two eigenvectors

$$\psi^{(\pm)}_{nlm} = \frac{1}{\sqrt{2}} \left( \psi^{(1)}_{n,l,m} \pm \psi^{(2)}_{n,l,m} \right) \tag{6.42}$$

whose eigenvalues

$$E_\pm = E_0 + J \pm K \tag{6.43}$$

satisfy

$$E_+ > E_- \tag{6.44}$$

In other words, the state $\psi^{(-)}_{nlm}$ — which is *antisymmetric* under the exchange of $\vec{x}_1$ with $\vec{x}_2$ — has *lower energy* than state $\psi^{(+)}_{nlm}$, *symmetric* under the same exchange. Physically this is easily understood: the state described by the vector $\psi^{(-)}_{nlm}$ has a wave function which *vanishes* for $\vec{x}_1 = \vec{x}_2$. Hence such a wave function describes a state in which the two electrons are, on average, farther apart than in the state $\psi^{(+)}_{nlm}$. Thus the repulsion energy for this state is *lower* than the repulsion energy for the symmetric state.

Summarizing, the excited levels of helium we are discussing, which are doubly degenerate in the independent electrons approximation, are in fact, because of the electrostatic repulsion, *doublets* whose energies are separated by $2\,K$. The states in the doublets with higher (lower) energies have wave functions which are (anti)symmetric under the exchange of $\vec{x}_1$ and $\vec{x}_2$. The fundamental state manifests no such splitting and its wave function is symmetric under the exchange of $\vec{x}_1$ and $\vec{x}_2$.

Experimentally the existence of these doublets is observed. The splitting of the states $\psi^{(\pm)}_{200}$ is $0.8\,eV$, while that of the states $\psi^{(\pm)}_{210}$ is $0.25\,eV$. It was also observed, however, that there are no spectral lines corresponding to transitions from the antisymmetric to the symmetric states[24].

This is explained as a consequence of the *Pauli exclusion principle.* Each of the electrons carries spin $\frac{1}{2}$: therefore the system of two electrons has one state of total spin $S = 0$ (a spin *singlet*) and three states of total spin $S = 1$ (a spin *triplet*). The $S = 0$ state is *antisymmetric* under the exchange of the spins of the two electrons while the $S = 1$ state is *symmetric.*

Since in the approximation we are considering the Hamiltonian does not depend on spin variables each state should be, were not for the Pauli exclusion principle, *four-fold* degenerate. The exclusion principle, which stipulates that the *total* (orbital plus spin) wave function of the electron system should be antisymmetric, remove part of this degeneracy. Indeed, the states with orbital wave functions $\psi^{(+)}_{nlm}$ must have $S = 0$ while those with wave functions $\psi^{(-)}_{nlm}$ have spin $S = 1$. Since, to first approximation, electromagnetic transitions cannot change $S$, this explains why there are no lines corresponding to transitions from the levels $\psi^{(+)}_{nlm}$ to the levels $\psi^{(-)}_{nlm}$.

**Problem 1.** Compute the energy levels corresponding to the helium first excited configuration, $(1s\,2s)$.

---

[24] More precisely, spectral lines corresponding to the transition from antisymmetric to symmetric states are very weak.

*Solution.* The direct integral is

$$
\begin{aligned}
J &= e^2 \int d^3\vec{x}_2 \big|\psi_{200}(\vec{x}_2)\big|^2 \int \frac{d^3\vec{x}_1}{\pi} \, \frac{|\psi_{100}(\vec{x}_1)|^2}{4\,\pi\,\epsilon_0\,|\vec{x}_1 - \vec{x}_2|} = \\
&= \frac{e^2}{4\,\pi\,\epsilon_0} \int d^3\vec{x}_2 \big|\psi_{200}(\vec{x}_2)\big|^2 \int \frac{d^3\vec{x}_1}{\pi} \, \frac{|\psi_{100}(\vec{x}_1 + \vec{x}_2)|^2}{|\vec{x}_1|} = \\
&= \frac{e^2\,Z}{4\,\pi\,\epsilon_0\,a_B} \int d^3\vec{x}_2 \big|\psi_{200}(\vec{x}_2)\big|^2 \frac{\big[1 - \mathrm{e}^{-2\,\rho_2}\,\big(1 + \rho_2\big)\big]}{\rho_2} = \\
&= \frac{e^2\,Z}{4\,\pi\,\epsilon_0\,a_B} \frac{1}{2} \int_0^\infty d\rho_2\,\rho_2\,\big(1 - \frac{1}{2}\,\rho_2\big)^2\,\mathrm{e}^{-\rho_2}\,\big[1 - \mathrm{e}^{-2\,\rho_2}\,\big(1 + \rho_2\big)\big] = \\
&= \frac{17}{81}\,\frac{e^2\,Z}{4\,\pi\,\epsilon_0\,a_B}
\end{aligned}
\tag{6.45}
$$

where

$$\rho_2 = \frac{Z}{a_B}\,|\vec{x}_2| \tag{6.46}$$

For helium, $Z = 2$ and thus

$$J \approx 11.417\,eV \tag{6.47}$$

The exchange integral is

$$K = e^2 \int d^3\vec{x}_2\,\psi^*_{200}(\vec{x}_2)\,\psi_{100}(\vec{x}_2) \int d^3\vec{x}_1\,\frac{\psi^*_{100}(\vec{x}_1)\,\psi_{200}(\vec{x}_1)}{4\,\pi\,\epsilon_0\,|\vec{x}_1 - \vec{x}_2|} \tag{6.48}$$

Introducing the adimensional variables

$$\vec{\rho}_1 = \frac{Z\,\vec{x}_1}{a_B} \qquad \vec{\rho}_2 = \frac{Z\,\vec{x}_2}{a_B}, \tag{6.49}$$

we obtain for the exchange integral

$$\begin{aligned}
K &= \frac{e^2\,Z}{4\,\pi\,\epsilon_0\,a_B} \int \frac{d^3\vec{\rho}_2}{2\,\sqrt{2}\,\pi} \big(1 - \frac{1}{2}\,|\vec{\rho}_2|\big)\,\mathrm{e}^{-\frac{3}{2}|\vec{\rho}_2|} \int \frac{d^3\vec{\rho}_1}{2\,\sqrt{2}\,\pi}\, \frac{\big(1-\frac{1}{2}\,|\vec{\rho}_1|\big)\,\mathrm{e}^{-\frac{3}{2}|\vec{\rho}_1|}}{|\vec{\rho}_1 - \vec{\rho}_2|} = \\
&= \frac{e^2\,Z}{4\,\pi\,\epsilon_0\,a_B} \int_0^\infty d\rho_2\,\rho_2^2\,\sqrt{2}\,\big(1-\frac{1}{2}\,\rho_2\big)\,\mathrm{e}^{-\frac{3}{2}\,\rho_2} \times \\
&\times \int_0^\infty \frac{d\rho_1\,\rho_1^2\,\big(1-\frac{1}{2}\,\rho_1\big)\,\mathrm{e}^{-\frac{3}{2}\rho_1}}{\sqrt{2}} \int_{-1}^{1} \frac{dt}{\sqrt{\rho_1^2+\rho_2^2-2\,t\,\rho_1\,\rho_2}} = \\
&= \frac{e^2\,Z}{4\,\pi\,\epsilon_0\,a_B} \int_0^\infty d\rho_2\,\rho_2^2\,\sqrt{2}\,\big(1-\frac{1}{2}\,\rho_2\big)\,\mathrm{e}^{-\frac{3}{2}\,\rho_2} \times \\
&\times \int_0^\infty \frac{d\rho_1\,\rho_1^2\,\big(1-\frac{1}{2}\,\rho_1\big)\,\mathrm{e}^{-\frac{3}{2}\rho_1}}{\sqrt{2}}\, \frac{\rho_1+\rho_2-|\rho_1-\rho_2|}{\rho_1\,\rho_2} = \\
&= \frac{e^2\,Z}{4\,\pi\,\epsilon_0\,a_B} \int_0^\infty d\rho_2\,\rho_2^2\,\sqrt{2}\,\big(1-\frac{1}{2}\,\rho_2\big)\,\mathrm{e}^{-\frac{3}{2}\,\rho_2} \times \\
\times&\Big[\int_0^{\rho_2} \frac{d\rho_1\,\rho_1^2\,\big(1-\frac{1}{2}\,\rho_1\big)\,\mathrm{e}^{-\frac{3}{2}\rho_1}}{\sqrt{2}}\,\frac{2}{\rho_2} + \int_{\rho_2}^\infty \frac{d\rho_1\,\rho_1^2\,\big(1-\frac{1}{2}\,\rho_1\big)\,\mathrm{e}^{-\frac{3}{2}\rho_1}}{\sqrt{2}}\,\frac{2}{\rho_1}\Big] = \\
&= \frac{e^2\,Z}{4\,\pi\,\epsilon_0\,a_B} \int_0^\infty d\rho_2\,\rho_2^2\,\big(1-\frac{1}{2}\,\rho_2\big)\,\mathrm{e}^{-\frac{3}{2}\,\rho_2} \times \frac{4}{27}\,\mathrm{e}^{-\frac{3}{2}\,\rho_2}\,(2+3\,\rho_2) = \\
&= \frac{16}{729}\,\frac{e^2\,Z}{4\,\pi\,\epsilon_0\,a_B} \qquad (6.50)
\end{aligned}$$

For helium $Z=2$, and thus

$$K \approx 1.194\,eV \qquad (6.51)$$

The result for the energies of the spin singlet and spin triplet are

$$\begin{aligned}
E_+(S=0) &= -\frac{Z^2\,e^2}{4\,\pi\,\epsilon_0\,2\,a_B}\,\big(1+\frac{1}{4}\big) + J + K = \\
&= -\frac{e^2}{4\,\pi\,\epsilon_0\,a_B}\,\big(\frac{5}{8}\,Z^2 - \frac{17}{81}\,Z - \frac{16}{729}\,Z\big) = -\frac{e^2}{4\,\pi\,\epsilon_0\,a_B}\,\big(\frac{5}{8}\,Z^2 - \frac{169}{729}\,Z\big) = \\
&\approx -55.4\,eV \\
E_-(S=1) &= -\frac{Z^2\,e^2}{4\,\pi\,\epsilon_0\,2\,a_B}\,\big(1+\frac{1}{4}\big) + J - K = \\
&= -\frac{e^2}{4\,\pi\,\epsilon_0\,a_B}\,\big(\frac{5}{8}\,Z^2 - \frac{17}{81}\,Z + \frac{16}{729}\,Z\big) = -\frac{e^2}{4\,\pi\,\epsilon_0\,a_B}\,\big(\frac{5}{8}\,Z^2 - \frac{137}{729}\,Z\big) = \\
&\approx -57.8\,eV \qquad (6.52)
\end{aligned}$$

The splitting between the two levels is

$$\Delta E = 2\,K = 2.39\,eV \tag{6.53}$$

These results should be compared to the experimental results

$$E_{exp}(S=1) = -59.2\,eV \quad E_{exp}(S=0) = -58.4\,eV \quad \Delta E_{exp} = 0.8\,eV \tag{6.54}$$

In fact the agreement with the experimental data is not very good.

**Problem 2.** Apply the variational method to the computation of the energy levels corresponding to the helium first excited configuration, $(1s\,2s)$.

*Solution.* We approximate the wave functions of the degenerate states $|1s;2s\rangle$ and $|2s;1s\rangle$ with hydrogenoid wave functions with generic value of the atomic number $Z'$. We write the helium hamiltonian as in (6.24)

$$\hat{H}_Z = \hat{H}^{(0)}_{Z'} + \hat{H}_{repulsion} + \frac{(Z'-Z)\,e^2}{4\,\pi\,\epsilon_0}\left[\frac{1}{|\vec{x}_1|} + \frac{1}{|\vec{x}_2|}\right] \tag{6.55}$$

with $Z = 2$. The matrix elements of the hamiltonian between the degenerate states are

$$\begin{aligned} H_{11} = \langle 1s;2s|\hat{H}_Z|1s;2s\rangle &= \frac{e^2}{4\,\pi\,\epsilon_0\,a_B}\left(-\frac{5}{8}\,(Z')^2 + \frac{17}{81}\,Z'\right) + \\ &+ \frac{(Z'-Z)\,e^2}{4\,\pi\,\epsilon_0}\,\langle 1s;2s|\left(\frac{1}{|\vec{x}_1|} + \frac{1}{|\vec{x}_2|}\right)|1s;2s\rangle \end{aligned} \tag{6.56}$$

and

$$\begin{aligned} H_{21} = \langle 2s;1s|\hat{H}_Z|1s;2s\rangle &= \frac{e^2}{4\,\pi\,\epsilon_0\,a_B}\,\frac{16}{729}\,Z' + \\ &+ \frac{(Z'-Z)\,e^2}{4\,\pi\,\epsilon_0}\,\langle 2s;1s|\left(\frac{1}{|\vec{x}_1|} + \frac{1}{|\vec{x}_2|}\right)|1s;2s\rangle = \\ &= \frac{e^2}{4\,\pi\,\epsilon_0\,a_B}\,\frac{16}{729}\,Z' \end{aligned} \tag{6.57}$$

Recalling (6.27), we write

$$\langle 1s;2s|\frac{1}{|\vec{x}_1|}|1s;2s\rangle = \langle 1s|\frac{1}{|\vec{x}_1|}|1s\rangle = \frac{Z'}{a_B} \tag{6.58}$$

Analogously

$$\begin{aligned} &\langle 1s;2s|\frac{1}{|\vec{x}_2|}|1s;2s\rangle = \langle 2s|\frac{1}{|\vec{x}_2|}|2s\rangle = \\ &= \frac{Z'}{a_B}\int_0^\infty \frac{d\rho\,\rho}{2}\,\big(1-\frac{1}{2}\,\rho\big)\,\mathrm{e}^{-\rho} = \frac{Z'}{a_B}\,\frac{1}{4} \end{aligned} \tag{6.59}$$

Therefore

$$H_{11} = \frac{e^2}{4\,\pi\,\epsilon_0\,a_B}\,\Big(-\frac{5}{8}\,(Z')^2 + \frac{17}{81}\,Z' + \frac{5\,(Z'-Z)\,Z'}{4}\Big) \tag{6.60}$$

Hence

$$\begin{aligned} &E_\pm(Z') = H_{11} \pm H_{21} = \\ &= \frac{e^2}{4\,\pi\,\epsilon_0\,a_B}\,\Big(-\frac{5}{8}\,(Z')^2 + \frac{17}{81}\,Z' + \frac{5\,(Z'-Z)\,Z'}{4} \pm \frac{16}{729}\,Z'\Big) \end{aligned} \tag{6.61}$$

Let us minimize $E_\pm(Z')$ with respect to $Z'$

$$Z'_- = Z - \frac{548}{3645} \qquad Z'_+ = Z - \frac{676}{3645} \tag{6.62}$$

From this we obtain[25]

$$\begin{aligned} E_-^{var} &= E_(Z'_-) = -\frac{e^2}{4\,\pi\,\epsilon_0\,a_B}\,\frac{5}{8}\,(Z - \frac{548}{3645})^2 \\ E_+^{var} &= E_(Z'_+) = -\frac{e^2}{4\,\pi\,\epsilon_0\,a_B}\,\frac{5}{8}\,(Z - \frac{676}{3645})^2 \end{aligned} \tag{6.63}$$

For the helium $(Z=2)$ we obtain only a small improvement in the agreement with experimental data

$$\begin{aligned} Z'_- &= 1.85 \qquad E_-^{var} \approx -58.2\,eV \\ Z'_+ &= 1.81 \qquad E_+^{var} \approx -56\,eV \end{aligned} \tag{6.64}$$

---

[25] As we observed earlier, the fact the the final variational result is a perfect square is to be expected a priori. See footnote (23).

# 7 Molecules

## 7.1 Valence bond theory: hydrogen molecule

The (non-relativistic) $H_2$ molecule is the described by the Hamiltonian

$$\begin{aligned}\hat{H} = \sum_{a=1,2}\Big(\frac{\vec{P}_a^2}{2\,M}+\frac{\vec{p}_a^2}{2\,m}\Big)+\frac{e^2}{4\,\pi\,\epsilon_0}\,\frac{1}{|\vec{X}_1-\vec{X}_2|}+ \\ -\frac{e^2}{4\,\pi\,\epsilon_0}\sum_{a=1,2}\Big(\frac{1}{|\vec{x}_a-\vec{X}_1|}+\frac{1}{|\vec{x}_a-\vec{X}_2|}\Big)+ \\ +\frac{e^2}{4\,\pi\,\epsilon_0}\,\frac{1}{|\vec{x}_1-\vec{x}_2|}\end{aligned} \tag{7.1}$$

where $\vec{X}_{1,2}$ and $\vec{P}_{1,2}$ are the positions and momenta of the two protons of mass $M$ and $\vec{x}_{1,2}$ and $\vec{p}_{1,2}$ are the positions and momenta of the two electrons of mass $m$.

The first approximation we will make (the so-called *Born-Oppenheimer* approximation) is to consider the two protons, which are much heavier than the electrons, as being at *fixed* positions. We will study therefore the problem of two electrons moving in the potential created by two protons located at a given distance

$$R = |\vec{R}| = |\vec{X}_1-\vec{X}_2| \tag{7.2}$$

The electronic hamiltonian is

$$\begin{aligned}\hat{H}_R = \sum_{a=1,2}\frac{\vec{p}_a^2}{2\,m}-\frac{e^2}{4\,\pi\,\epsilon_0}\sum_{a=1,2}\Big(\frac{1}{|\vec{x}_a-\vec{X}_1|}+\frac{1}{|\vec{x}_a-\vec{X}_2|}\Big)+ \\ +\frac{e^2}{4\,\pi\,\epsilon_0}\,\frac{1}{|\vec{x}_1-\vec{x}_2|}=\end{aligned} \tag{7.3}$$

The eigenvalues $E(R)$ of this hamiltonian depend, parametrically, on the distance $R$ between the protons. The function

$$V_{eff}(R) = \frac{e^2}{4\,\pi\,\epsilon_0}\,\frac{1}{R}+E(R)-2\,E_1^H \tag{7.4}$$

represent the "effective" potential felt by the nuclei: the first term is the coulombian repulsion; the second — which is called the *electronic term* —

encodes the effect of the electronic cloud; the third term is minus the energy of the two separated atoms. We will interpret a minimum of $V_{eff}(R)$ for $R = R_0$ such that $V_{eff}(R_0) < 0$ as the existence of a bound state of the two hydrogen atoms of size $R$ and binding energy $|V_{eff}(R_0)|$.

The system (7.3) is still too complicated to be analyzed exactly. We will describe its eigenstates by using an approximation similar to the one we used to discuss the excited levels of helium[26]

When $R$ is large, the system reduces to two non-interacting hydrogen atoms, which are described by the Hamiltonian

$$\hat{H}^{(0)}_{R=\infty} = \sum_{a=1,2} \Big( \frac{\vec{p}_a^{\,2}}{2\,m} - \frac{e^2}{4\,\pi\,\epsilon_0}\,\frac{1}{|\vec{x}_a - \vec{X}_a|} \Big) \tag{7.5}$$

The eigenfunctions of $\hat{H}^{(0)}_{R=\infty}$ are products of hydrogen wave functions. As in the case of the helium atom, if one neglects the spins of the electrons, the ground state of $\hat{H}^{(0)}_{R=\infty}$ is doubly degenerate:

$$\begin{aligned} \Psi^{(1)} &= \psi_{100}(\vec{x}_1 - \vec{X}_1)\,\psi_{100}(\vec{x}_2 - \vec{X}_2) \equiv |\psi_{100}(1);\psi_{100}(2)\rangle \\ \Psi^{(2)} &= \psi_{100}(\vec{x}_1 - \vec{X}_2)\,\psi_{100}(\vec{x}_2 - \vec{X}_1) \equiv |\psi_{100}(2);\psi_{100}(1)\rangle \end{aligned} \tag{7.6}$$

In one configuration one electrons is attached to the first nucleus and the other to the second; in the other configuration the positions of the two electrons are exchanged. Both $\Psi^{(1)}$ and $\Psi^{(2)}$ are eigenstates of $\hat{H}^{(0)}_{R=\infty}$ with same eigenvalue

$$E_0 = 2\,E^H_1 \tag{7.7}$$

Notice, however, an important difference with respect to the analysis of the excited levels of the helium: the degenerate states $\Psi^{(1)}$ and $\Psi^{(2)}$ are *not* orthogonal:

$$\langle \Psi^{(1)}, \Psi^{(2)} \rangle = |S_R|^2 \tag{7.8}$$

where $S_R$ is the *overlap* integral

$$S_R \equiv \langle \psi_{100}(1), \psi_{100}(2) \rangle = \int d^3\vec{x}\,\psi^*_{100}(\vec{x} - \vec{X}_1)\,\psi_{100}(\vec{x} - \vec{X}_2) \tag{7.9}$$

---

[26] This approximation has been applied to the hydrogen molecules for the first by Heitler and London (1927). It is the basis of of *valence bond* theory.

**Problem 1.** Show that $S_R$ is real and

$$S_R = s(\rho) = \mathrm{e}^{-\rho}\left(1 + \rho + \frac{1}{3}\rho^2\right) \qquad \rho \equiv \frac{R}{a_B} \tag{7.10}$$

When $R$ is finite the interaction hamiltonian $\hat{H}_{int}$ between the two hydrogen atoms become non-negligeable

$$\begin{aligned} \hat{H}_{int} &= \hat{H}_R - \hat{H}^{(0)}_{R=\infty} = \\ &= \frac{e^2}{4\pi\epsilon_0}\left[\frac{1}{|\vec{x}_1 - \vec{x}_2|} - \frac{1}{|\vec{x}_1 - \vec{X}_2|} - \frac{1}{|\vec{x}_2 - \vec{X}_1|}\right] \end{aligned} \tag{7.11}$$

This hamiltonian is *symmetric* under the exchange of the coordinates of the two electrons

$$\vec{x}_1 \leftrightarrow \vec{x}_2 \tag{7.12}$$

Hence, the eigenstates of $\hat{H}_R$ have *definite* parity under this exchange. Therefore, to first approximation, the normalized eigenstates of $\hat{H}_R$ are the linear combinations of $\Psi^{(1)}$ and $\Psi^{(2)}$ with definite parities

$$\begin{aligned} \Psi^{(+)} &= N_+\big(|\psi_{100}(1);\psi_{100}(2)\rangle + |\psi_{100}(2);\psi_{100}(1)\rangle\big) \\ \Psi^{(-)} &= N_-\big(|\psi_{100}(1);\psi_{100}(2)\rangle - |\psi_{100}(2);\psi_{100}(1)\rangle\big) \end{aligned} \tag{7.13}$$

The normalization factors $N_\pm$ are obtained by the equations

$$\begin{aligned} \left|\Psi^{(\pm)}\right|^2 = |N_\pm|^2\left(1 + 1 \pm 2S_R^2\right) = 2|N_\pm|^2\left(1 \pm S_R^2\right) = 1 \Rightarrow \\ N_\pm = \frac{1}{2\left(1 \pm S_R^2\right)} \end{aligned} \tag{7.14}$$

The (approximate) eigenvalues of $\Psi^{(\pm)}$ are

$$E_\pm = \langle\Psi^{(\pm)}, \hat{H}_R\,\Psi^{(\pm)}\rangle = \frac{1}{2\left(1 \pm S_R^2\right)}\left[H_{11} + H_{22} \pm \big(H_{12} + H_{21}\big)\right] \tag{7.15}$$

where $H_{ab}$, with $a, b = 1, 2$, are the matrix elements of $\hat{H}_R$ between the states $\Psi^{(1)}$ and $\Psi^{(2)}$. Because of the symmetry of $\hat{H}_R$ under the exchange (7.12) we have

$$H_{11} = H_{22} \qquad H_{12} = H_{21} \tag{7.16}$$

Therefore

$$E_\pm = \langle \Psi^{(\pm)}, \hat{H}_R \, \Psi^{(\pm)} \rangle = \frac{1}{\left(1 \pm S_R^2\right)} \left[ H_{11} \pm H_{12} \right] \tag{7.17}$$

The matrix elements of $\hat{H}_R$ can be written as

$$H_{ab} = E_{ab}^{(0)} + H_{ab}^{(int)} \tag{7.18}$$

where

$$E_{ab}^{(0)} = 2\, E_1^H \begin{pmatrix} 1 & S_R^2 \\ S_R^2 & 1 \end{pmatrix} \tag{7.19}$$

are the matrix elements of the non-interacting $H^{(0)}_{R=\infty}$, and $H_{ab}^{(int)}$ are the matrix elements of $H^{int}$:

$$\begin{aligned}
H_{11}^{(int)} &= \frac{e^2}{4\,\pi\,\epsilon_0} \int d^3\vec{x}_1 \int d^3\,\vec{x}_2\, \rho_{100}(\vec{x}_1 - \vec{X}_1)\, \rho_{100}(\vec{x}_2 - \vec{X}_2) \,\times \\
&\qquad\qquad \times \Big[ \frac{1}{|\vec{x}_1 - \vec{x}_2|} - \frac{1}{|\vec{x}_1 - \vec{X}_2|} - \frac{1}{|\vec{x}_2 - \vec{X}_1|} \Big] = \\
&= \frac{e^2}{4\,\pi\,\epsilon_0} \int d^3\vec{x}_1 \int d^3\,\vec{x}_2\, \frac{\rho_{100}(\vec{x}_1 - \vec{X}_1)\, \rho_{100}(\vec{x}_2 - \vec{X}_2)}{|\vec{x}_1 - \vec{x}_2|} + \\
&\quad -2 \times \frac{e^2}{4\,\pi\,\epsilon_0} \int d^3\vec{x}\, \frac{\rho_{100}(\vec{x} - \vec{X}_1)}{|\vec{x} - \vec{X}_2|} = \\
&= \frac{e^2}{4\,\pi\,\epsilon_0\, a_B} \Big[ I_2(\rho) - 2\, I_1(\rho) \Big] \\
H_{12}^{(int)} &= \frac{e^2}{4\,\pi\,\epsilon_0} \int d^3\vec{x}_1 \int d^3\,\vec{x}_2\, \rho^*_{ex}(\vec{x}_1; \vec{X}_1, \vec{X}_2)\, \rho_{ex}(\vec{x}_2; \vec{X}_1, \vec{X}_2) \,\times \\
&\qquad\qquad \times \Big[ \frac{1}{|\vec{x}_1 - \vec{x}_2|} - \frac{1}{|\vec{x}_1 - \vec{X}_2|} - \frac{1}{|\vec{x}_2 - \vec{X}_1|} \Big] = \\
&= \frac{e^2}{4\,\pi\,\epsilon_0} \int d^3\vec{x}_1 \int d^3\,\vec{x}_2\, \frac{\rho_{ex}(\vec{x}_1; \vec{X}_1, \vec{X}_2)\, \rho^*_{ex}(\vec{x}_2; \vec{X}_1, \vec{X}_2)}{|\vec{x}_1 - \vec{x}_2|} + \\
&\quad - \frac{e^2}{4\,\pi\,\epsilon_0}\, S_R \int d^3\vec{x} \Big[ \frac{\rho^*_{ex}(\vec{x}; \vec{X}_1, \vec{X}_2)}{|\vec{x} - \vec{X}_2|} + \frac{\rho_{ex}(\vec{x}; \vec{X}_1, \vec{X}_2)}{|\vec{x} - \vec{X}_1|} \Big] = \\
&= \frac{e^2}{4\,\pi\,\epsilon_0\, a_B} \Big[ J_2(\rho) - 2\, J_1(\rho)\, s(\rho) \Big]
\end{aligned} \tag{7.20}$$

We have introduced

$$
\begin{aligned}
&\rho_{ex}(\vec{x};\vec{X}_1,\vec{X}_2) \equiv \psi^*_{100}(\vec{x}-\vec{X}_1)\,\psi_{100}(\vec{x}-\vec{X}_2)\\
&\vec{\rho} \equiv \frac{\vec{R}}{a_B} \qquad \rho \equiv \frac{R}{a_B}\\
&I_2(\rho) \equiv a_B \int d^3\vec{x}_1 \int d^3\,\vec{x}_2\, \frac{\rho_{100}(\vec{x}_1-\vec{X}_1)\,\rho_{100}(\vec{x}_2-\vec{X}_2)}{|\vec{x}_1-\vec{x}_2|} =\\
&\quad = \int \frac{d^3\vec{x}_1}{\pi} \int \frac{d^3\vec{x}_2}{\pi}\, \frac{\mathrm{e}^{-2\,|\vec{x}_1|}\,\mathrm{e}^{-2\,|\vec{x}_2-\vec{\rho}|}}{|\vec{x}_1-\vec{x}_2|} =\\
&\quad = \frac{1}{\rho}\left[1-\mathrm{e}^{-2\,\rho}\left(1+\frac{11}{8}\,\rho+\frac{3}{4}\,\rho^2+\frac{1}{6}\,\rho^3\right)\right]\\
&I_1(\rho) \equiv a_B \int d^3\vec{x}\, \frac{\rho_{100}(\vec{x}-\vec{X}_1)}{|\vec{x}-\vec{X}_2|} = \int \frac{d^3\vec{x}}{\pi}\,\frac{\mathrm{e}^{-2\,|\vec{x}-\vec{\rho}|}}{|\vec{x}|} =\\
&\quad = \frac{1}{\rho}\left(1-\mathrm{e}^{-2\,\rho}\,(1+\rho)\right)\\
&J_1(\rho) \equiv a_B \int d^3\vec{x}\, \frac{\rho^*_{ex}(\vec{x};\vec{X}_1,\vec{X}_2)}{|\vec{x}-\vec{X}_2|} = \int \frac{d^3\vec{x}}{\pi}\,\frac{\mathrm{e}^{-|\vec{x}|}\,\mathrm{e}^{-|\vec{x}-\vec{\rho}|}}{|\vec{x}|} =\\
&\quad = \mathrm{e}^{-\rho}\,(1+\rho)\\
&J_2(\rho) = a_B \int d^3\vec{x}_1 \int d^3\,\vec{x}_2\, \frac{\rho_{ex}(\vec{x}_1;\vec{X}_1,\vec{X}_2)\,\rho^*_{ex}(\vec{x}_2;\vec{X}_1,\vec{X}_2)}{|\vec{x}_1-\vec{x}_2|} =\\
&\quad = \int d^3\vec{x}_1 \int d^3\,\vec{x}_2\, \frac{\mathrm{e}^{-|\vec{x}_1|-|\vec{x}_1-\vec{\rho}|}\,\mathrm{e}^{-|\vec{x}_2|-|\vec{x}_2-\vec{\rho}|}}{|\vec{x}_1-\vec{x}_2|}
\end{aligned}
\tag{7.21}
$$

The "exchange" integral $J_2(\rho)$ is *very hard* to compute analytically. Heitler and London were only able to estimate it; Siugura (1927) was the first to evaluate it exactly.

Summarizing, the energy terms $E_\pm(R)$ write as follows

$$E_\pm(R) = 2\,E^H_1 + \frac{1}{\left(1\pm S^2_R\right)}\left[H^{int}_{11} \pm H^{int}_{12}\right] \tag{7.22}$$

and the effective potential (7.4) felt by the nuclei becomes

$$V^\pm_{eff}(R) = \frac{e^2}{4\,\pi\,\epsilon_0}\,\frac{1}{R} + \frac{1}{\left(1\pm S^2_R\right)}\left[H^{int}_{11} \pm H^{int}_{12}\right] \tag{7.23}$$

Contrary to the helium case, $H_{12}$ is *not* positive definite because of the Coulomb attraction between the electrons and the nuclei. It turns out that it is $\Psi^{(+)}$ the state with *lower* energy for all values of $R$. Because of the Pauli exclusion principle, $\Psi^{(+)}$ has total spin $S = 0$ while $\Psi^{(-)}$ has $S = 1$. Therefore in the ground state the spins of the electrons are anti-parallel and $S = 0$.

$E_+(R)$ is sufficiently negative to make $V^+_{eff}(R)$ have the right shape to describe a *binding* potential with a negative minimum around

$$R_{min} = 1.64\, a_B \qquad V^{(+)}_{eff}(R_{min}) = -0.23\, |E^H_1| \approx -3.15\, eV \tag{7.24}$$

in discrete agreement with the experimental values

$$R_{exp} = 1.400\, a_B \qquad V_{exp} = -0.349\, |E^H_1| \tag{7.25}$$

$V^-_{eff}(R)$ turns out to be, instead, non-binding.

The result (7.24) can be improved by replacing the hydrogen wave functions $\psi_{100}(\vec{x})$ with hydrogenoid wave functions with generic atomic number $Z$, and performing a variational computation to find the optimal $Z$. The variational computation gives

$$\begin{aligned} &R_{min} = 1.414\, a_B \qquad V^{(+)}_{eff}(R_{min}) = -0.278\, |E^H_1| \approx -3.78\, eV \\ &Z_{eff} = 1.166 \end{aligned} \tag{7.26}$$

in better agreement with the experimental results.

## 7.2 The hydrogen molecule-ion

$H_2^+$ is described in the Born-Oppenheimer approximation by

$$\hat{H}_R = \frac{\vec{p}^2}{2\, m} - \frac{e^2}{4\, \pi\, \epsilon_0} \left( \frac{1}{|\vec{x} - \vec{X}_1|} + \frac{1}{|\vec{x} - \vec{X}_2|} \right) \tag{7.27}$$

where $\vec{X}_{1,2}$ are the positions the two protons of mass $M$ and $\vec{x}$ and $\vec{p}$ is the position and momenta of the electron of mass $m$. The system (7.27) is exactly solvable. However the resulting energy eigenfunctions are quite involved. Here we present an approximation method to deduce the spectrum of $\hat{H}_R$ which is known as the *linear combination of atomic orbitals* (LCAO) approximation.

When

$$R = |\vec{X}_1 - \vec{X}_2| \tag{7.28}$$

is very large the eigenfunctions of $\hat{H}_R$ are approximated by the hydrogen atomic wavefunctions around either one of the nuclei. In the $R \to \infty$ the ground state is degenerate

$$\begin{aligned} \Psi^{(1)}(\vec{x}) &= \psi_{100}(\vec{x} - \vec{X}_1) \\ \Psi^{(2)}(\vec{x}) &= \psi_{100}(\vec{x} - \vec{X}_2) \end{aligned} \tag{7.29}$$

The symmetry of $\hat{H}_R$ under the exchange of the coordinates of the protons implies that its eigenstates have a definite symmetry under the same exchange. The approximate eigenstates are therefore

$$\psi_\pm = N_\pm \big( \Psi^{(1)}(\vec{x}) \pm \Psi^{(2)}(\vec{x}) \big) \tag{7.30}$$

where $N_\pm$ are given in terms of the overlap $S_R$

$$\begin{aligned} 1 &= |N_\pm|^2 \left(1 + 1 \pm 2\, S_R\right) = 2\, |N_\pm|^2 \left(1 \pm S_R\right) \\ S_R &= \int d^3\vec{x}\, \psi^*_{100}(\vec{x} - \vec{X}_1)\, \psi_{100}(\vec{x} - \vec{X}_2) = s(\rho) \\ \rho &= \frac{R}{a_B} \end{aligned} \tag{7.31}$$

The corresponding approximate eigenvalues are

$$E_\pm = \langle \psi_\pm, \hat{H}_R\, \psi_\pm \rangle = \frac{1}{\big(1 \pm S_R\big)} \big( H_{11} \pm \Re(H_{12}) \big) \tag{7.32}$$

where $H_{11}$ and $H_{12}$ are the matrix elements of $\hat{H}_R$ between $\Psi^{(1)}$ and $\Psi^{(2)}$

$$\begin{aligned} H_{11} &= E^H_1 - \frac{e^2}{4\,\pi\,\epsilon_0} \int d^3\vec{x}\, \frac{\rho_{100}(\vec{x} - \vec{X}_1)}{|\vec{x} - \vec{X}_2|} \equiv \\ &\equiv E^H_1 - \frac{e^2}{4\,\pi\,\epsilon_0\, a_B}\, I_1(\rho) \\ H_{12} &= E^H_1\, S_R - \frac{e^2}{4\,\pi\,\epsilon_0} \int d^3\vec{x}\, \frac{\rho_{ex}(\vec{x}; \vec{X}_1, \vec{X}_2)}{|\vec{x} - \vec{X}_2|} \equiv \\ &\equiv E^H_1\, S_R - \frac{e^2}{4\,\pi\,\epsilon_0}\, J_1(\rho) \end{aligned} \tag{7.33}$$

where $I_1(\rho)$ and $J_1(\rho)$ are the same integrals already introduced in Eq. (7.21)

$$\begin{aligned}
&\rho_{ex}(\vec{x};\vec{X}_1,\vec{X}_2) \equiv \psi^*_{100}(\vec{x}-\vec{X}_1)\,\psi_{100}(\vec{x}-\vec{X}_2)\\
&I_1(\rho) \equiv a_B \int d^3\vec{x}\,\frac{\rho_{100}(\vec{x}-\vec{X}_1)}{|\vec{x}-\vec{X}_2|} = \int \frac{d^3\vec{x}}{\pi}\,\frac{\mathrm{e}^{-2\,|\vec{x}-\vec{\rho}|}}{|\vec{x}|}\\
&J_1(\rho) \equiv a_B \int d^3\vec{x}\,\frac{\rho^*_{ex}(\vec{x};\vec{X}_1,\vec{X}_2)}{|\vec{x}-\vec{X}_2|} = \int \frac{d^3\vec{x}}{\pi}\,\frac{\mathrm{e}^{-|\vec{x}|}\,\mathrm{e}^{-|\vec{x}-\vec{\rho}|}}{|\vec{x}|} =\\
&\qquad = \mathrm{e}^{-\rho}\,(1+\rho)
\end{aligned} \tag{7.34}$$

**Problem 2.** Show that

$$\begin{aligned}
&I_1(\rho) = \frac{1}{\rho}\left(1-\mathrm{e}^{-2\,\rho}\,(1+\rho)\right) \equiv \frac{1}{a_B}\,f(\rho) \qquad \rho = \frac{R}{a_B}\\
&J_1(\rho) = \mathrm{e}^{-\rho}\,(1+\rho) \equiv \frac{1}{a_B}\,g(\rho)
\end{aligned} \tag{7.35}$$

The effective potential felt by the two nuclei of the hydrogen ion is therefore

$$V_\pm(R) \equiv \frac{e^2}{4\,\pi\,\epsilon_0\,R} + E_\pm - E_1^H = \frac{e^2}{4\,\pi\,\epsilon_0\,a_B}\left[\frac{1}{\rho} - \frac{I_1(\rho)\pm J_1(\rho)}{\left(1\pm s(\rho)\right)}\right] \tag{7.36}$$

The lowest energy state is the *symmetric* state, and it is a binding state. It turns out that the "resonance" term $J_1(\rho)$ is the crucial term to obtain a binding potential whose negative minimum is

$$R_{min} = 2.5\,a_B \qquad V_+(R_{min}) = -0.13\,|E_1^H| \tag{7.37}$$

in discrete, although not perfect, agreement with the experimental values

$$R_{exp} = 2.00\,a_B \qquad\qquad V_{exp} = -0.205\,|E_1^H| \tag{7.38}$$

$V_-(R)$ turns out to be, instead, non-binding.

An improvement on the result the result (7.37) is obtained by using for $\psi_{100}(\vec{x})$ the hydrogenoid wave function with *generic* $Z$

$$\psi_{100}(\vec{x}) = \frac{Z^{\frac{3}{2}}}{\sqrt{\pi}\,a_B^{\frac{3}{2}}}\,\mathrm{e}^{-\frac{Z}{a}\,|\vec{x}|}$$

One obtains in this way a potential $V(R,Z)$ which depends on both $R$ and $Z$:

$$
\begin{aligned}
V(R,Z) &= \frac{e^2}{4\,\pi\,\epsilon_0\,a_B}\Big[\frac{Z}{\rho} - Z\,\frac{I_1(\rho)+(2-Z)\,J_1(\rho)+(1-Z)}{1+s(\rho)} - \frac{Z^2-1}{2}\Big] \\
\rho &\equiv \frac{Z\,R}{a_B}
\end{aligned}
\tag{7.39}
$$

Extremizing this potential with respect to both $R$ and $Z$ one obtains

$$
\begin{aligned}
&R_{min} = \frac{\rho_{min}}{Z_{min}}\,a_B = 2.003\,a_B \qquad Z_{min} = 1.2380 \\
&V_+(R_{min}, Z_{min}) = -0.1730\,|E_1^H|
\end{aligned}
\tag{7.40}
$$

in better agreement with the experimental results.

## 7.3 Huckel theory for conjugated hydrocarbons

Butiadene is a chain $CH_2 = CH - CH = CH_2$ in which each one of the carbon atoms provides an electron in a $p$-orbital (with level $n = 2$). The four $p$-electrons are delocalized and are shared by the the molecule. The approximation we are going to discuss can be applied to more general linear or ring shaped conjugated systems, involving many carbon atoms each one providing a $p$ electron.

The atoms of the chain are supposed to lie on a plane, which will will take as the $(x,y)$ plane. The $p$-orbitals are those aligned in the direction normal to this plane, i.e. are those with $L_z = 0$. The space of states accessible to each electron is the linear span of the $p$ wave functions, which we will denote by

$$
\psi_i = \psi_{210}(\vec{x} - \vec{X}_i) \qquad i = 1,2,3,4 \tag{7.41}
$$

where $\vec{X}_i$ are the positions of the carbon atoms. The matrix elements

$$
H_{ij} = \langle \psi_i, H\,\psi_j \rangle \tag{7.42}
$$

of the hamiltonian between these states is supposed to have the form

$$
H = \begin{pmatrix} \alpha & \beta & 0 & 0 \\ \beta & \alpha & \beta & 0 \\ 0 & \beta & \alpha & \beta \\ 0 & 0 & \beta & \alpha \end{pmatrix} \tag{7.43}
$$

The overlaps

$$S_{ij} = \langle \psi_i, \psi_j \rangle = \delta_{ij} \tag{7.44}$$

are taken to be zero for $i \neq j$. A generic *normalized* state in the 4-dimensional linear span of the $\psi_i$'s is

$$\psi = \sum_i c_i \, \psi_i \tag{7.45}$$

with

$$\sum_i |c_i|^2 = 1 \tag{7.46}$$

The energy on such a state is

$$E(c) = \langle \, \psi, H \, \psi \rangle = \sum_{i,j} H_{ij} \, c_i^* \, c_j \tag{7.47}$$

We want to find the extremum of $E(c)$ on the space of vectors (7.45) with $c_i$ obeying the normalization condition (7.46). To this end, consider the following function of $c^i$'s and $\epsilon$

$$W(c, \epsilon) \equiv E(c) - \epsilon \, \Big( \sum_i |c_i|^2 - 1 \Big) \tag{7.48}$$

The extrema equations for $W(c, \epsilon)$ are

$$\begin{aligned} \frac{\partial \, W(c)}{\partial c_i^*} &= H_{ij} \, c_j - \epsilon \, c_i = 0 \\ \frac{\partial \, W(c)}{\partial \epsilon} &= \sum_i |c_i|^2 - 1 = 0 \end{aligned} \tag{7.49}$$

Hence the solutions of this equation give precisely the minima of $E(c)$ which satisfy the constraint (7.46).

Eqs. (7.49) mean that the $(c_1, c_2, \ldots, c_N)$ is a normalized eigenvectors of the matrix $H_{ij}$ with eigenvalue $\epsilon$.

The equation which determines the eigenvalues of $H_{ij}$ is

$$\det \begin{pmatrix} \alpha - \epsilon & \beta & 0 & 0 \\ \beta & \alpha - \epsilon & \beta & 0 \\ 0 & \beta & \alpha - \epsilon & \beta \\ 0 & 0 & \beta & \alpha - \epsilon \end{pmatrix} = 0 \tag{7.50}$$

or equivalently

$$\det \begin{pmatrix} x & 1 & 0 & 0 \\ 1 & x & 1 & 0 \\ 0 & 1 & x & 1 \\ 0 & 0 & 1 & x \end{pmatrix} = 0 \tag{7.51}$$

where

$$x \equiv \frac{\alpha - \epsilon}{\beta} \Leftrightarrow \epsilon = \alpha - \beta\, x \tag{7.52}$$

The equation (7.51) has four roots

$$\begin{aligned} & x^4 - 3\,x^2 + 1 = 0 \Rightarrow x = \pm\, x_{1,2} \\ & x_1 = \frac{1+\sqrt{5}}{2} \approx 1.618 \qquad x_2 = \frac{\sqrt{5}-1}{2} \approx 0.618 \end{aligned} \tag{7.53}$$

The four eigenvalues of the energy are

$$\epsilon_{\pm}^{(1)} = \alpha \mp \beta\, x_1 \qquad \epsilon_{\pm}^{(2)} = \alpha \mp \beta\, x_2 \tag{7.54}$$

The eigenvector associated to eigenvalue $\epsilon = \alpha - \beta\, x$ is

$$\begin{aligned} & c_2 = -x\, c_1 \\ & c_3 = -c_1 - x\, c_2 = -c_1 + x^2\, c_1 = (-1 + x^2)\, c_1 \\ & c_4 = -c_2 - x\, c_3 = x\, c_1 + x\, c_1 - x^3\, c_1 = (2\,x - x^3)\, c_1 \end{aligned} \tag{7.55}$$

These eigenvectors describe electrons which are delocalized among the four carbons: they are called *molecular orbitals*. They are indicated by $\pi$ since they come from atomic $p$ orbitals.

$\beta$ is negative. Therefore two lowest levels are

$$\epsilon_{-}^{(1)} = \alpha + \beta\, x_1 \qquad \epsilon_{-}^{(2)} = \alpha + \beta\, x_2 \tag{7.56}$$

The levels $\epsilon_{+}^{(1,2)}$ are non-binding.

In the independent electrons approximation, these two levels are doubly occupied and give the $1\pi^2 2\pi^2$ molecular orbital configuration. The total energy is

$$E_0 = 4\,\alpha + 2\,\beta\,\big(x_1 + x_2\big) \approx 4\,\alpha + 4.472\,\beta \tag{7.57}$$

**Problem 3.** Apply Huckel theory to the ethylene molecule $CH_2 = CH_2$, in which two $p$ electrons are shared by the molecule. Show that the single electrons levels are

$$\epsilon_\pm = \alpha \pm \beta \tag{7.58}$$

and find the corresponding eigenvectors.

The ground state of ethylene has configuration $1\pi^2$ whose energy is (in the independent electrons approximation)

$$E_0 = 2\,\alpha + 2\,\beta \tag{7.59}$$

This shows the increase in stability that the shared delocalized electrons provide to butiadene: if butiadene had two localized double bonds, the electrons ground state energy would be twice that of ethylene, i.e. $4\,(\alpha + \beta)$. The difference

$$E_0^{butiadene} - 2\,E_0^{ethylene} = (4.472 - 4)\beta = 0.472\,\beta \tag{7.60}$$

is called the *delocalization* energy of butiadiene.

**Problem 4.** Apply Huckel theory to the benzene molecule in which 6 $p$ electrons are shared by a planar molecule which for *closed* (ring) chain. Compute the electron levels and the delocalization energy with respect to ethylene.

## 7.4 Particle on a ring: a simplified model for benzene

Consider a particle of mass $M$ constrained to move around a circular ring of radius $R$. The energy of this system is just the kinetic energy

$$H = \frac{1}{2}\,m\,v^2 = \frac{1}{2}\,M\,R^2\,\dot{\phi}^2 = \frac{p_\phi^2}{2\,M\,R^2} = \frac{p_\phi^2}{2\,I} \tag{7.61}$$

where $\phi(t)$ is the angular coordinate on the circle, $\dot{\phi}(t)$ the time derivative,

$$I = M\,R^2 \tag{7.62}$$

is the moment of inertia of the system and

$$p_\phi = M\,R^2\,\dot{\phi} = R\,m\,v = L_z \tag{7.63}$$

is the canonical momentum conjugate to $\phi$, which coincides with the component of the angular momentum along the direction perpendicular to the plane of the circle.

At quantum level we know that $p = M\,v$ is represented in the Schrödinger representation by

$$\hat{p} = -i\,\hbar\frac{\partial}{\partial x} \tag{7.64}$$

where $x = \phi\,R$ is the linear coordinate along the trajectory of the particle. Hence

$$p_\phi \to \hat{p}_\phi = -i\,\hbar\,\frac{\partial}{\partial\phi} \tag{7.65}$$

which is indeed the operator which represents $L_z$ in the Schrödinger representation. In conclusion, the quantum Hamiltonian is

$$\hat{H} = \frac{L_z^2}{2\,I} \tag{7.66}$$

and the Schrödinger equation is

$$-\frac{\hbar^2}{2\,I}\frac{\partial^2}{\partial\phi^2}\,\psi(\phi) = E\,\psi(\phi) \tag{7.67}$$

where $\psi(\phi)$ is the wave function in the Schrödinger representation.

Since $L_z$ and $\hat{H}$ commute, there is a simultaneous basis of common eigenvectors:

$$\begin{aligned} &\psi_m(\phi) = \frac{\mathrm{e}^{i\,m\,\phi}}{\sqrt{2\,\pi}} \\ &L_z\,\psi_m(\phi) = \hbar\,m\,\psi_m(\phi) \\ &H\,\psi_m(\phi) = E_m\,\psi_m(\phi) \qquad E_m = \frac{\hbar^2\,m^2}{2\,I} \end{aligned} \tag{7.68}$$

Therefore $\hat{H}$ has a discrete spectrum labelled by $m = 0, 1, 2, \ldots$: the ground state has $m = 0$ has is non-degenerate, while all levels with $E_m > 0$ are two-fold degenerates, and correspond to two angular momentum eigenstates $\psi_{\pm m}(\phi)$.

If we think of the system of the 6 carbon electrons of benzene as moving freely on the ring, the fundamental level will be occupied by 2 electrons with opposite spin, while the doubly degenerate level with $m = 1$ will be occupied by 4 electrons. This is the $1\pi^2 2\,\pi^4$ configuration. The first excited state will be associated to the excitation of an electron to the level $m = 2$. The difference of energy with respect to the ground state is

$$\Delta E = E_2 - E_1 = \frac{3\,\hbar^2}{2\,I} \tag{7.69}$$

**Problem 5.** Knowning that the radius $R$ of benzene is $R \approx 1.39\,10^{-8} cm$ find the prediction of this ring model for wave-length of the emission radiation associated to the transition from the first excited state to the fundamental level. Compare with the experimental value $\lambda_{exp} = 2.68\,10^{-5}\,cm$

## 7.5 Cyclic conjugated hydrocarbons: an unidimensional model for crystals

We want to discuss the Huckel model for a molecule which form a planar close chain with $N$ carbon atoms providing $N$ $p$ electrons. We will consider the limit for $N$ large to describe a one-dimensional crystal. The hamiltonian for the N-electrons Huckel model describing a closed chain is the $N \times N$ matrix

$$H = \begin{pmatrix} \alpha & \beta & 0 & \dots & \beta \\ \beta & \alpha & \beta & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \beta & \alpha & \beta \\ \beta & 0 & \dots & \beta & \alpha \end{pmatrix} \tag{7.70}$$

It is clear that this system is symmetric under the shift of the location of the $N$ carbon atoms

$$T : X_1 \to X_2 \to X_3 \to \dots \to X_{N-1} \to X_N \to X_1 \tag{7.71}$$

The matrix which implements this rotation of the ring is

$$T = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 \end{pmatrix} \tag{7.72}$$

**Problem 6.** Show that $T$ and $H$ commute.

Since $T$ and $H$ commute, they have a basis of *simultaneous* eigenvectors. Let us study, therefore, $T$ eigenvectors:

$$T\,v = \xi\,v \tag{7.73}$$

$T$ is a unitary operator (not an hermitian one). Therefore

$$|\xi| = 1 \tag{7.74}$$

Moreover from

$$T^N = \mathbb{I}_N \tag{7.75}$$

it follows that

$$\xi^N = 1 \tag{7.76}$$

Hence the $T$ has $N$ complex eigenvalues which are the $N$ complex $N$-th roots of the unity

$$\xi_k = \mathrm{e}^{\frac{2\,\pi\,k\,i}{N}}, \qquad k = 0, 1, 2, \dots, N-1 \tag{7.77}$$

Since $\xi_k$ are all different, $T$ is non degenerate. The $T$ eigenvectors

$$v_k = (c_1^{(k)}, \dots, c_N^{(k)}) \tag{7.78}$$

corresponding to the eigenvalues $\xi_k$, satisfy

$$\begin{pmatrix} -\xi_k & 1 & 0 & \dots & 0 \\ 0 & -\xi_k & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & -\xi_k & 1 \\ 1 & 0 & \dots & 0 & -\xi_k \end{pmatrix} \begin{pmatrix} c_1^{(k)} \\ c_2^{(k)} \\ \dots \\ c_{N-1}^{(k)} \\ c_{N-1}^{(k)} \end{pmatrix} = 0 \tag{7.79}$$

Therefore

$$\begin{aligned} c_2^{(k)} &= \xi_k\, c_1^{(k)} \\ c_3^{(k)} &= \xi_k\, c_2^{(k)} = \xi_k^2\; c_1^{(k)} \\ \ldots &= \ldots \\ c_N^{(k)} &= \xi_k\, c_{N-1}^{(k)} = \xi_k^{N-1}\, c_1^{(k)} \end{aligned} \tag{7.80}$$

The normalized eigenvector $v_k$ is therefore

$$v_k = \frac{1}{\sqrt{N}} \left(1, \xi_k, \xi_k^2, \ldots, \xi_k^{N-1}\right) \tag{7.81}$$

Since $v_k$ are not degenerate eigenvectors of $T$ they are also eigenvectors of $H$. To compute the $H$ eigenvalues it is convenient to note that[27]

$$H = \alpha\, \mathbb{I}_N + \beta \left(T + T^{-1}\right) \tag{7.82}$$

Hence

$$H\, v_k = \left[\alpha + \beta \left(\xi_k + \frac{1}{\xi_k}\right)\right] v_k \tag{7.83}$$

The energy eigenvalues are therefore

$$\epsilon_k = \alpha + 2\,\beta\, \cos\frac{2\,\pi\,k}{N} \tag{7.84}$$

Let us remark that

$$\epsilon_{N-k} = \epsilon_k \tag{7.85}$$

Therefore:

- when $N$ odd, there are $\frac{N-1}{2}$ two-fold degenerate eigenvalues $\epsilon_k$ with $k = 1, 2, \ldots, \frac{N-1}{2}$, and *one* single non-degenerate eigenvalue for $k = 0$

$$\epsilon_0 = \alpha + 2\,\beta \tag{7.86}$$

- when for $N$ even, there are $\frac{N}{2} - 1$ two-fold degenerates eigenvalues $\epsilon_k$ for $k = 1, 2, \ldots, \frac{N}{2} - 1$ and *two* non-degenerate eigenvalues with $k = 0$ and $k = \frac{N}{2}$

$$\begin{aligned} \epsilon_0 &= \alpha + 2\,\beta \\ \epsilon_{\frac{N}{2}} &= \alpha - 2\,\beta \end{aligned} \tag{7.87}$$

---

[27] This relation, in particular provide another way to prove the $T$ and $H$ commute.

In the independent electron approximation, assuming as usual $\beta < 0$, the ground state is the state in which the $N$ electrons fill the molecular orbitals starting from the lowest $\epsilon_0 = \alpha + 2\,\beta$. When

$$N = 2 + 4\,k \tag{7.88}$$

all the binding levels are filled, giving rise to particular stable molecules, a little bit like the atoms of noble gases. Planar conjugated ings which satisfy (7.88) are said to obey Huckel's rule for *aromaticity*. For example, benzene corresponds to $N = 6$, $k = 1$.

**Problem 7.** Find the energy difference between the first excited state and the ground state of benzene in Huckel's model.

For other values of $N$ the last level to be filled ($k = \frac{N}{4}$ or $k = \frac{N \pm 1}{4}$) has less than 4 electrons. Molecules with $N = 4\,k$ are called *anti-aromatic*.

When $N \to \infty$ the discrete levels (7.84) become a continuous *band* covering the interval $[\alpha + 2\,\beta, \alpha - 2\,\beta]$ and $\frac{k}{N}$ becomes a continuous index, the *pseudo-momentum* taking value in a finite interval.