Notes
on
Quantum Mechanics
with Examples of Solved Problems

Ennio Gozzi*

Department of Physics, Miramare Section, University of Trieste

Strada Costiera 11, P.O. Box 586, Trieste, Italy

* gozzi@ts.infn.it

“ La meccanica quantistica?
Non è altro che materia
che dorme e sogna ”

Fabia Gardinazzi
Poetessa
(1959–2010)

TABLE OF CONTENTS

INTRODUCTION

These notes on Quantum Mechanics (QM) are not original but are based on material I took and adapted from many books: Landau, Messiah, Sakurai, Cohen-Tannoudji, etc. and in large part from the handwritten notes of my colleague GianCarlo Ghirardi who taught this course for many many years before retiring. Differently from many other books, these notes contain 34 problems with their detailed worked-out solutions. I will present in class at least 10 extra problems with solutions of which I will provide copy of my handwritten notes. Several other problems are left as homeworks and I advise you to do them because they may end up being part of the final written exam. All these problems are not original but taken from various sources: Landau, Lim, D'Emilio and Picasso, et al. and some, very original ones, were generously provided to me by GianCarlo Ghirardi.

My personal advice to the students of this course is to study first these notes (or even better the original handwritten ones of Ghirardi) and only later on, if needed, to turn to the books quoted above. This advice is given only in order to speed up the preparation of the exam and avoid getting lost in the details and notation of many different books, notation that I have tried to make uniform in these notes.

A second advice is to work out many more of the 50 problems presented and solved in these notes. The solution of problems is what a physicist should learn to do in every course and later on in his professional life. It is the only manner to really master the theoretical aspects presented in class or learned from the book.

A third very important advice (but also very *personal*) is to waste no time, at

this stage of your preparation, in wondering about the very "*strange*" behaviour of quantum mechanics. This "strange" behaviour ranges from interference effects, to the tunnel effect, to the Schrödinger cat paradox, to the non-locality intrinsic in QM and many other things. This "strange" behaviour is a consequence of a set of *postulates* which are given at the beginning and that we will spell out in details. Once these postulates are accepted, the "strange" behaviour of QM comes as a consequence. So somehow I like to look at quantum mechanics as a *mathematical* theory with its own postulates and consequences (theorems) and one should not wonder too much if the consequences are rather "strange". After all, the *experiments* confirm this "strange" behaviour. Remember that in physics only experiments are the ultimate test of a theory and QM is the best tested theory around. The "strange" behavior is only due to the fact that we, as human beings, are macroscopic objects and so we can develop an intuition only about macroscopic phenomena and not about microscopic ones like the things that happen, for example, in an atom and that are described in a strange way by QM. I would like sometime to be transformed into an electron for just a day, and I would like that not only because in that manner I would loose weight, but also because for one day I would experience a different world and develop the intuition to "understand" QM. Remember anyhow that the role of theoretical physics is not to "*understand*" phenomena but to develop the equations which describe these physical phenomena and that can *predict* new ones. In this respect the strange theory which is QM has been an extremely successful one. So as an electron I would only be able to develop an intuition for these strange phenomena but I would not be able to explain them.

As I said I like to see QM as a mathematical theory with its own postulates and theorems and corollaries of theorems and so on. In a mathematical theory the *only* things that we can do are the following:

1. to check that the postulates are *not in contradiction* with each other (and sometimes contradictions are subtle);

2. to check if the postulates are *redundant*, so that we may be able to find a smaller set which is sufficient to explain all of QM phenomena;

3. to check if we can *replace* one postulate with a new one and still obtain a *consistent* theory.This is somehow analog to what has been done in non-Euclidean geometries where the fifth postulate of Euclid has been replaced by a new one and all the others left the same. The new theory was consistent, had no contradiction but it described a new world (a curved one instead of the flat one described by the Euclidean geometry).This same thing could happen in QM where, with some a new postulate replacing an old one, we may obtain a consistent theory but which describes a different world;

4. to check if one or two of the postulates could be unified in a new postulate with the same physical consequences we have tested so far and the potential to explain new phenomena. This is the road followed, for example, by those who modified the Schrödinger equation (for a review see ref [1]). With the postulate of a new and different time evolution, they could replace both the postulate of the Schrödinger time evolution and the one of collapse of the wave function (at least for measurement of the position operator);

5. to check if the postulates of QM could be proved as theorems of a more basic theory than QM.

These are the things that we, as theoreticians, should do instead of wondering about the "strange" behaviour of QM. As experimentalists instead we should look for new phenomena which QM could (or could not) explain.

A deep unsolved problem that I *personally* feel is still open is the transition from the *microscopic* to the *macroscopic* world and the interaction between the two. The chair on which you are seated is made of hundred of trillions of billions of atoms and so it should be possible in principle to describe it using QM because atoms behave according to QM, but instead we are unable to describe the chair using QM. So we use an "*effective*" theory which is classical mechanics (CM). I call it "effective" because somehow it manages to describes hundred of trillions of billions of atoms in a single stroke. People usually think that the transition from QM to CM is achieved by sending the Planck constant to zero. This is not correct, not only because the Planck constant has dimension but also because in this "semiclassical" limit we do not get exactly CM but CM plus some "phases" which spoils the classical behaviour. The same happens when, instead of the semiclassical limit, we take the limit of large N (N being the number of atoms). So the problem of how to get CM from QM is still open. Also the problem of how a macroscopic object (subject to the rules of CM) interacts with a microscopic one (subject to the rules of QM) is still open and it is at the basis of the so called "measurement problem" and of all the modern mesoscopic physics phenomena discovered over the last 20 years. Attempts to correctly achieve the transition from QM to CM have been done by many people. The two I know better are summarized in ref.[2] and [3].

The solution of the problem of the transition and interaction between QM and CM may shed some light not only on several of the "strange" features of QM and on mesoscopic physics phenomena but also on CM itself. As we saw before, CM is only an "effective" theory and not a "*fundamental*" one like QM. It could (in principle) be obtained from QM by using some procedure and limit (but not the ones mentioned above) like for example the one of considering as basic variables blocks of phase space of dimension $\gg \hbar$ and not points in phase-space. This may overthrow concepts like

the one of "*degrees of freedom*" and similar ones which are considered well-established and not questionable concepts of CM. In this procedure it will for sure emerge that CM is not only an "effective theory" but also an "*approximate*" one . This may lead to the conclusion that CM is *not* the correct theory to describe *all* macroscopic phenomena. Actually there are some macroscopic phenomena that we are unable to explain and which were discovered over the last 30 years. They are those astronomical phenomena which need the introduction of theoretical concepts like "*dark matter*" and "*dark energy*". May it be that instead of all these "dark things" of which we have not found any experimental evidence, the way out is that CM may have to be modified (or improved in his "approximate" status) at those large scales? These improvements can come from a more rigorous and less approximate derivation of CM from QM. All this gives you an idea of the problems that we may be able to solve starting from a better understanding of QM.

In this book you will not find any of these more advanced topics or others like quantum relativistic equations, modifications to the Schrödinger equation, entanglement, non-locality, quantum information, teleportation etc. which are left for more advanced courses taught by my colleagues at the University of Trieste.

These notes are in this LaTeX and Web form thanks to the hard work of Alessandro Candolini and the financial support of the "Consorzio per lo sviluppo della fisica" of Trieste University. I wish to warmly thank both Alessandro and the Consorzio expecially his president prof. Ghirardi and its director prof. Mardirossian. I wish also to thank in advance those colleagues and students who, reading these notes, will indicate to me any misprints and conceptual or computational errors which for sure are present in these notes.

These notes are dedicated to the memory of my dear friends Fabia and Paolo.

Ennio Gozzi

Trieste, September , 2012

Bibliography

[1] A. Bassi and G. C. Ghirardi, Phys. Rept. **379** (2003) 257.

[2] G. C. Ghirardi, R. Grassi and F. Benatti, Found. Of Phys. **25** (1995) 5.
A. Bassi, G. C. Ghirardi, Phys. Lett. A **275**, (2000) 373.

[3] A. A. Abrikosov (jr), E. Gozzi, D. Mauro, Ann. of Phys. **317** (2005) 24.

Chapter 1

SCHRÖDINGER EQUATION. WRONSKIAN THEOREM.

1.1 The time-dependent Schrödinger equation: general properties

The time-dependent Schrödinger equation for the wave function $\psi(x,t)$ of a quantum particle of mass m moving in one dimension is

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2}+V(x)\psi(x,t)=i\hbar\frac{\partial\psi(x,t)}{\partial t}\,, \tag{1.1}$$

where $V(x)$ is the potential acting on the particle.

The main properties of this equation are

- if $\psi(x,t)$ is a solution, then also all the functions of the form

$$\psi_\varphi(x,t)=\mathrm{e}^{i\varphi}\,\psi(x,t) \tag{1.2}$$

  are solutions of the same equation, provided that the phase φ does not depend on x,t. This follows from the fact that Eqn. (1.1) is linear in ψ. The phase φ can anyhow depend on such external parameters as a constant electric or magnetic field.

- The set of all solutions of the form (1.2) is referred to as "ray". We shall take into account only one element for each ray, but it is important that, once you have chosen the phase, you do not change the phase when you sum two states.

Example 1.1. Let us suppose to choose the phase α for the state ψ_1 and the phase β for another state ψ_2, thus the sum of the two states is

$$\psi = \psi_1 \, \mathrm{e}^{i\alpha} + \psi_2 \, \mathrm{e}^{i\beta} \ .$$

If now we change the phase of ψ_1 for an amount γ and the phase of ψ_2 for an amount δ, the sum would become

$$\tilde{\psi} = \psi_1 \, \mathrm{e}^{i\alpha} \, \mathrm{e}^{i\gamma} + \psi_2 \, \mathrm{e}^{i\beta} \, \mathrm{e}^{i\delta}$$

Clearly, ψ and $\tilde{\psi}$ do not belong to the same ray, since in general

$$\tilde{\psi} \neq \psi \, \mathrm{e}^{i\Delta} \ ,$$

where Δ is a phase. Thus, the sum of two states is a state while the sum of two rays may not be a ray,

- Thus, you must sum states and not rays.

1.2 Solution of the time-dependent Schrödinger equation

We now turn to the problem of how to solve the time-dependent Schrödinger equation (1.1). Let us seek a solution of the form *

$$\Psi(x,t) = \psi(x)\varphi(t) \ . \tag{1.3}$$

Inserting (1.3) in Eqn. (1.1) and dividing by $\Psi(x,t)$ we obtain

$$\frac{1}{\psi(x)} \left[-\frac{\hbar^2}{2m} \frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} + V(x)\psi(x) \right] = i\hbar \frac{1}{\varphi(t)} \frac{\mathrm{d}\varphi(t)}{\mathrm{d}t} \ . \tag{1.4}$$

* A general solution is just a linear combinations of them: $\tilde{\Psi}(x,t) = \sum_i C_i \psi_i(x)\varphi_i(t)$.

The left-hand side of Eqn. (1.4) depends only on x while the right-hand side depends only on t. The only way out is that both sides be equal to a constant, say E:

$$i\hbar \frac{1}{\varphi(t)} \frac{\mathrm{d}\varphi(t)}{\mathrm{d}t} = E \ , \tag{1.5}$$

$$\frac{1}{\psi(x)} \left[-\frac{\hbar^2}{2m} \frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} + V(x)\psi(x) \right] = E \ . \tag{1.6}$$

The solution of Eqn. (1.5) is (up to a constant)

$$\varphi(t) = \mathrm{e}^{-iEt/\hbar} \ . \tag{1.7}$$

and Eqn. (1.6) can be written as:

$$-\frac{\hbar^2}{2m} \frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} + V(x)\psi(x) = E\psi(x) \ . \tag{1.8}$$

Eqn. (1.8) is the so-called stationary (or time-independent) Schrödinger equation.

Few remarks:

- In the above derivation, special care must be taken in dealing with the zeros of $\psi(x)$ and $\varphi(t)$, however it is possible to prove that they do not lead to problems.

- Total, and not partial, derivatives have been used in writing Eqn. (1.4) since ψ and φ depend only on x and t respectively.

1.2.1 Properties of the solution

1. Since $\psi(x)$ is related to the probability density $\rho(x)$ of finding the particle in x via the relation [Born interpretation of the $\psi(x)$]

$$\rho(x) = |\psi(x)|^2 \ , \tag{1.9}$$

we must require that ψ be normalized:

$$\int_{-\infty}^{+\infty} \rho(x)\,\mathrm{d}x = 1 = \int_{-\infty}^{+\infty} |\psi(x)|^2\,\mathrm{d}x \ ,$$

thus $\psi(x)$ must be square-integrable: $\psi \in L^2$.

2. Moreover $\psi(x)$ must be finite for all x, otherwise if $\psi(x)$ were infinite at some point x_0 then also the probability density would become infinite at that point x_0.

3. $\psi(x)$ must be single-valued because from it we will derive observable results;

4. $\psi(x)$ and its derivative will be chosen to be continuos at least away from singularities of the potential. This choice is "compatible" with the physical interpretation of the wave-function and its derivative. "Compatible" means that the continuity of the wave-function shall imply the continuity of the probability density which is a reasonable physical condition. Analogously for the derivative of the wave-function:its continuity is "compatible" with the continuity of the probability current defined below. A more mathematical argument is the following based on the fact that the derivability of a function implies the continuity of the same. In the Schroedinger equation we have the second derivative of the wave-function and, if this does exists, it implies that the first derivative exists and is a continuos function. This last in turn, as it is a derivative, it implies that the wave-function itself is continuos. So this is the reason why the wave-function and its derivative must be continuos, at least for regular potential.

Before concluding this section, we should recall the continuity equation which in the one-dimensional case reads

$$\frac{\partial \rho(x,t)}{\partial t} = -\frac{\partial J(x,t)}{\partial x} \ ,$$

where the *probability density* $\rho(x,t)$ and the *probability current* $J(x,t)$ are given respectively by

$$\begin{aligned} \rho(x,t) &= |\Psi(x,t)|^2 = \Psi(x,t)\Psi^*(x,t) \ , \\ J(x,t) &= \frac{i\hbar}{2m}\left[-\Psi^*(x,t)\frac{\partial \Psi(x,t)}{\partial x} + \Psi(x,t)\frac{\partial \Psi^*(x,t)}{\partial x}\right] \ . \end{aligned}$$

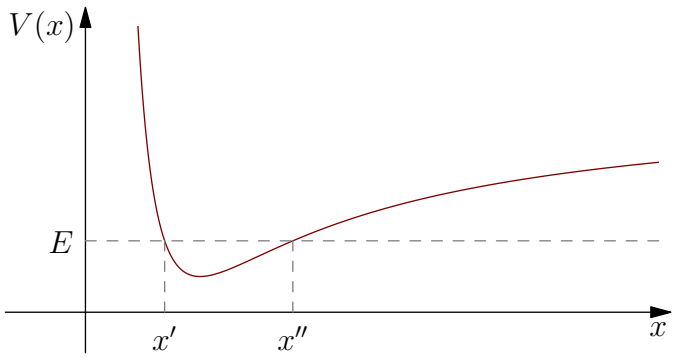


FIG. 1.1. The potential $V(x)$ used in the qualitative analysis of the stationary Schrödinger equation and the energy E chosen for the analysis.

1.3 Discrete energy levels

We would like to understand in this section if the Schrödinger equation (1.8) can have *discrete* energy solutions like several experimental evidences seem to indicate for atoms and other systems.

1.3.1 Qualitative analysis

The stationary Schrödinger equation (1.8) can be rewritten as

$$\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} = \frac{2m}{\hbar^2}\left[V(x) - E\right]\psi(x) \; . \tag{1.10}$$

Since it is a second-order linear differential equation, the values of $\psi(x)$ and its first-order derivative at some initial point $x = a$ must be provided in order to solve the equation, *i.e.*:

$$\psi(a) = \psi_a \; , \quad \left.\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\right|_{x=a} = \psi'_a \; .$$

Let us consider a potential $V(x)$ like the one plotted in Fig. 1.1. Between x' and x'' we have $V(x) - E < 0$. If we choose $x' < a < x''$ and we suppose $\psi(a) > 0$, then it follows from Eqn. (1.10) that

$$\left.\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2}\right|_{x=a} < 0 \; ,$$

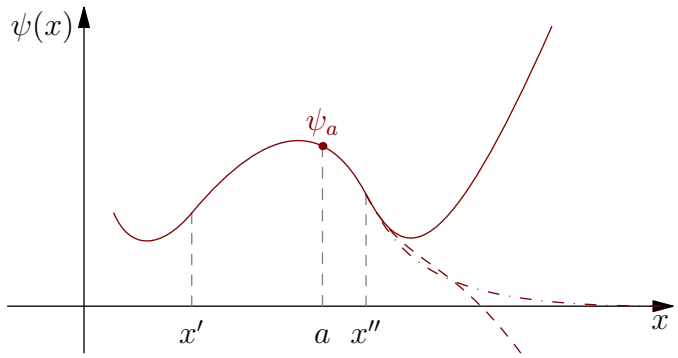


FIG. 1.2. Three attempts to find an acceptable solution to the time-independent Schrödinger equation.

and $\psi(x)$ remains concave downwards up to x'' (see Fig.1.2). After the point x'' the wave function goes into a region where $V(x) - E > 0$, thus it becomes concave upwards. In this case, the wave function diverges towards $+\infty$, as shown in Fig. 1.2 (solid line).

Another possibility is to decrease the slope of the wave function after x''. In this way ψ intersects the x-axis and it becomes negative. Beyond the point of intersection we have

$$\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} < 0 \,,$$

thus the wave function becomes concave downwards and again it diverges, this time towards $-\infty$, see Fig. 1.2 (dashed line).

A third possibility occurs when the slope of the wave function after x'' is such that ψ decreases very slowly so that it remains concave upwards, it does not intersect the x-axis and tends to zero asymptotically. This solution has a behavior at least at $+\infty$ that allows it to be normalizable. See Fig. 1.2.

1.3.2 Quantitative analysis and Wronskian theorem

Let us write the stationary Schrödinger equation

$$\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} = \frac{2m}{\hbar^2}[V(x) - E]\psi(x)$$

in the following way

$$\frac{\mathrm{d}^2 z(x)}{\mathrm{d}x^2} + [\tilde{E} - U(x)]z(x) = 0 \,, \tag{1.11}$$

where

$$\tilde{E} \equiv \frac{2m}{\hbar^2}E \,, \quad U(x) \equiv \frac{2m}{\hbar^2}V(x) \,, \quad z(x) \equiv \psi(x) \,.$$

Let us take $U(x)$ continuous and bounded from below.

Given two arbitrary solutions $z_1(x)$ and $z_2(x)$ of Eqn. (1.11) corresponding to two different energies, we can define the following object, which is called the *Wronskian*:

$$W(z_1, z_2) \equiv z_1 z_2' - z_2 z_1' \,, \tag{1.12}$$

where $z_1' = \mathrm{d}z_1/\,\mathrm{d}x$ and $z_2' = \mathrm{d}z_2/\,\mathrm{d}x$. If, at a certain point $x = a$, the Wronskian (1.12) is zero, then

$$z_1(a)z_2'(a) - z_2(a)z_1'(a) = 0 \,,$$

that is,

$$\frac{z_2'(a)}{z_2(a)} = \frac{z_1'(a)}{z_1(a)} \,,$$

which means that the logarithmic derivatives* of the two solutions are the same at that point.

Theorem 1.1 (Wronskian theorem)**.** If $z_1(x)$ and $z_2(x)$ are solutions of the following equations

$$z_1''(x) + F_1(x)z_1(x) = 0 \,, \tag{1.13a}$$

$$z_2''(x) + F_2(x)z_2(x) = 0 \,, \tag{1.13b}$$

* The logarithmic derivative of $z(x)$ is $\frac{\mathrm{d}}{\mathrm{d}x}\log z(x)$.

where $F_1(x)$ and $F_2(x)$ are generic functions, then it is possible to prove that the variation of the associated Wronskian $W(z_1, z_2)$ in some interval (a, b) is given by

$$W(z_1, z_2)\Big|_a^b = \int_a^b [F_1(x) - F_2(x)]\, z_1(x) z_2(x)\, \mathrm{d}x \; . \tag{1.14}$$

Proof. If we multiply Eqn. (1.13a) by z_2 and Eqn. (1.13b) by z_1 and take the difference of the two, we obtain

$$\underbrace{z_2 z_1'' - z_1 z_2''}_{-W'(z_1,z_2)} + (F_1 - F_2) z_1 z_2 = 0 \; . \tag{1.15}$$

The first term is, modulo a sign, nothing but the derivative of the Wronskian $W(z_1, z_2)$ with respect to x, as you can see by direct differentiation of Eqn. (1.12):

$$W'(z_1, z_2) = z_1 z_2'' - z_2 z_1'' \; .$$

Integrating Eqn. (1.15) in x between a and b yields just Eqn. (1.14). ∎

Corollary 1.1.1. If Eqs. (1.13) correspond to the stationary Schrödinger equation (1.11) for two different energies $\tilde{E}_1$ and $\tilde{E}_2$ respectively, then $F_1(x) = \tilde{E}_1 - U(x)$, $F_2(x) = \tilde{E}_2 - U(x)$ and Eqn. (1.14) becomes

$$W(z_1, z_2)\Big|_a^b = (\tilde{E}_1 - \tilde{E}_2) \int_a^b z_1(x) z_2(x)\, \mathrm{d}x \; . \tag{1.16}$$

Corollary 1.1.2. If the two solutions z_1 and z_2 are associated with the same energy $\tilde{E}_1 = \tilde{E}_2$, then we get from Eqn. (1.16)

$$W(z_1, z_2)\Big|_{x=a} = W(z_1, z_2)\Big|_{x=b} \; ,$$

so

$$W(z_1, z_2) = \text{const} \; ,$$

i.e., the Wronskian is independent of x since the interval (a, b) is arbitrary.

Corollary 1.1.3. If $z(x, \tilde{E})$ is a solution of Eqn. (1.11) and if its logarithmic derivative, *i.e.*:

$$F(x, \tilde{E}) = \frac{z'(x, \tilde{E})}{z(x, \tilde{E})}$$

at the point $x = a$ has a fixed value f_a which is independent of $\tilde{E}$, *i.e.*:

$$F(a, \tilde{E}) = f_a \ , \quad \text{for all } \tilde{E} \ ,$$

then $F(x, \tilde{E})$ is a monotonic function of $\tilde{E}$, increasing if $x < a$ and decreasing if $x > a$. This is because the derivative of F with respect to $\tilde{E}$ is given by (see proof below)

$$\frac{\partial F(x, \tilde{E})}{\partial \tilde{E}} = -\frac{1}{z^2(x, \tilde{E})} \int_a^x z^2(\xi, \tilde{E}) \, \mathrm{d}\xi \ . \tag{1.17}$$

and from Eqn. (1.17) we see that the derivative is negative if $x > a$ and positive if $x < a$.

Proof. Let us consider the following boundary conditions:

$$z(a, \tilde{E}) = z_a \ , \quad z'(a, \tilde{E}) = z'_a \ ,$$

and replace $\tilde{E}$ with $\tilde{E} + \delta\tilde{E}$. There will be two nearby solutions z and $z + \delta z$ which correspond to the values $\tilde{E}$ and $\tilde{E} + \delta\tilde{E}$ and satisfy the same boundary conditions. According to Corollary 1.1.1, at first order in δ we obtain

$$W(z, z + \delta z)\Big|_a^b = -\delta\tilde{E} \int_a^b z^2(x) \, \mathrm{d}x \ . \tag{1.18}$$

At the point $x = a$, the Wronskian is zero: $W(z, z+\delta z)|_{x=a} = 0$. This is so because the two solutions satisfy the same boundary conditions, since z_a and z'_a are independent of $\tilde{E}$. For all other values of x in the interval (a, b) we have

$$W(z, z + \delta z) = z\delta z' - z'\delta z = z^2\delta\left(\frac{z'}{z}\right) = z^2\delta F \ .$$

Thus, Eqn. (1.18) becomes

$$z^2\delta F\Big|_{x=b} = -\delta\tilde{E} \int_a^b z^2(x) \, \mathrm{d}x \ ,$$

that is,

$$\frac{\delta F}{\delta\tilde{E}}\Big|_{x=b} = -\frac{1}{z^2(b)} \int_a^b z^2(x) \, \mathrm{d}x \ .$$

from which it follows that F is a monotonic function of $\tilde{E}$, increasing or decreasing for $b < a$ or $b > a$ respectively. ∎

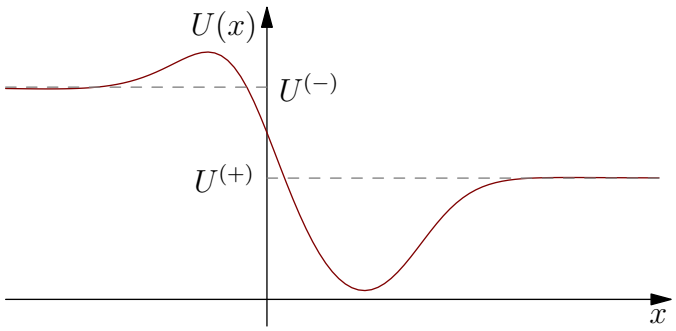


FIG. 1.3. Potential $U(x)$ and its asymptotic behavior at $\pm\infty$.

Let us now take a potential $U(x)$ such as the one plotted in Fig. 1.3, with $U^{(-)} > U^{(+)}$.

First case. We consider $\tilde{E} > U^{(-)}$. In this case, $\tilde{E} - U(x)$ is positive at the two extrema and constant, Eqn. (1.11) becomes at $+\infty$

$$z'' + \underbrace{[\tilde{E} - U^{(+)}]}_{k} z = 0 \,,$$

that is,

$$z'' = -kz \,, \qquad k \equiv \tilde{E} - U^{(+)} > 0 \,,$$

and the asymptotic behavior of z at $+\infty$ is

$$z \sim \mathrm{e}^{i\sqrt{k}x} \quad \text{for } x \to +\infty \,. \tag{1.19a}$$

At $-\infty$, Eqn. (1.11) becomes

$$z'' + \underbrace{[\tilde{E} - U^{(-)}]}_{\tilde{k}} z = 0 \,,$$

that is,

$$z'' = -\tilde{k}z \,, \qquad \tilde{k} \equiv \tilde{E} - U^{(-)} > 0 \,,$$

and the solution is

$$z \sim \mathrm{e}^{i\sqrt{\tilde{k}}x} \quad \text{for } x \to -\infty \,. \tag{1.19b}$$

Solutions (1.19) are bounded and oscillating and they can be fitted continuously with each other. So there is no constraint on the energy. Every value of energy is allowed and the spectrum is continuous.

Remark. Oscillating solutions are allowed, even if — strictly speaking — they are not truly normalizable. They are normalizable in a generalized sense (theory of distributions). In fact, the scalar product of two plane waves e^{ikx} and $\mathrm{e}^{ik'x}$ is

$$\int_{-\infty}^{+\infty} \mathrm{e}^{ikx}\, \mathrm{e}^{-ik'x}\, \mathrm{d}x \sim \delta(k-k') \; .$$

On the RHS we do not have the usual Kronecker delta but the Dirac delta (a distribution) of which we will give details later on. The space of square-integrable functions L^2 must be enlarged in order to include also those functions which are square-integrable in the distributional sense. If we call this space $\tilde{L}^2$, it is possible to prove that L^2 is dense in $\tilde{L}^2$, that is to say, each element in $\tilde{L}^2$ has a "nearby" element in L^2.

Second case. Let us now consider $U^{(-)} > \tilde{E} > U^{(+)}$. In this case, $\tilde{E}-U$ is negative for $x \to -\infty$ so Eqn. (1.11) becomes

$$z'' = -kz \; , \qquad k \equiv \tilde{E} - U^{(-)} < 0 \; ,$$

and the asymptotic behavior of z at $-\infty$ is given by

$$z \sim \mathrm{e}^{\sqrt{|k|}x} \quad \text{for } x \to -\infty \; , \tag{1.20a}$$

which tends to zero for $x \to -\infty$ and therefore can be accepted. At $+\infty$, in Eqn. (1.11) we have instead

$$z'' = -k'z \; , \qquad k' \equiv \tilde{E} - U^{(+)} > 0 \; ,$$

and the solution is

$$z \sim \mathrm{e}^{i\sqrt{k'}x} \quad \text{for } x \to +\infty \; , \tag{1.20b}$$

which is an oscillating function. Again it is possible to match the two solutions (1.20) and therefore also in this case there are no constrains on the energy and the spectrum is continuous.

Third case. We consider $U^{(-)} > U^{(+)} > \tilde{E}$. Eqn. (1.11) becomes at both extrema

$$z'' = -k^{\pm} z \ , \qquad k^{\pm} \equiv \tilde{E} - U^{\pm} < 0 \ ,$$

and the asymptotic solutions are

$$z \sim \mathrm{e}^{\mp\sqrt{|k^{\pm}|}x} \quad \text{for } x \to \pm\infty \ , \tag{1.21}$$

In this case, the solutions are bounded and damped at both extrema. Now we have to fit them in order to guarantee the continuity of $z(x)$ and of its first-order derivative. Let us consider the logarithmic derivatives of the two solutions, and indicate them with $F_+(x, \tilde{E})$ and $F_-(x, \tilde{E})$. They will be equal to each other at a certain point $x = a$. This is so because both z and z' satisfy certain boundary conditions at the point $x = a$, and since z_- and z_+ converge to the same solution z, both z_- and z_+ must satisfy the same boundary conditions in a. Now F_+ is valid for $x > a$ and according to Corollary 1.1.3 it is a monotonic decreasing function of $\tilde{E}$, while F_- is valid for $x < a$ and is a monotonic increasing function of $\tilde{E}$. Now F_+ and F_- must be equal in $x = a$ and this can happen only for some particular values $\tilde{E}_n$ of $\tilde{E}$. This is the proof that in the present case only certain values of $\tilde{E}$ are allowed. These values are fixed and isolated and the spectrum is discrete.

1.4 The Dirac delta

In this section we will provide some introductory details on the Dirac delta that we mentioned before. Let us recall the main properties of the Dirac's δ-function.

A *function* is a map between, for example, the space of real numbers $\mathbb{R}$ and $\mathbb{R}$, or between the space of complex numbers $\mathbb{C}$ and $\mathbb{C}$, or between $\mathbb{C}$ and $\mathbb{R}$. A *functional* is a map between a space of functions and the space of real or complex numbers:

$$\mathcal{F}[f(x)] \to \mathbb{R} \ .$$

Example 1.2 (functionals).

- The *action*

$$S = \int_{t_0}^{t_1} \mathcal{L}(q(t), \dot{q}(t), t)\, \mathrm{d}t$$

associates to a function $q(t)$ the real number S obtained by evaluating the integral above.

- The *Dirac's delta function* $\delta(x)$ is defined via the relation

$$\int_{-\infty}^{+\infty} \delta(x - x_0) f(x)\, \mathrm{d}x = f(x_0)\ , \tag{1.22}$$

that is, the Dirac delta associates to the function $f(x)$ the number $f(x_0)$.

Eqn. (1.22) suggests that it is possible to "represent" the Dirac delta distribution as the limit of a suitable sequence of functions. For example, if we choose $x_0 = 0$ we have that the action of the Dirac delta in Eqn. (1.22), namely

$$\int_{-\infty}^{+\infty} \delta(x) f(x)\, \mathrm{d}x = f(0)\ ,$$

is the same as the one achieved by the following operation:

$$\delta(x) \sim \lim_{l\to\infty} \frac{l}{\sqrt{\pi}}\, \mathrm{e}^{-l^2x^2}\ ,$$

which must be understood as performed under an integral:

$$\int_{-\infty}^{+\infty} \delta(x) f(x)\, \mathrm{d}x = \lim_{l\to\infty} \int_{-\infty}^{+\infty} \frac{l}{\sqrt{\pi}}\, \mathrm{e}^{-l^2x^2}\, f(x)\, \mathrm{d}x\ .$$

(Prove the latter equation as an homework.)

Note that the set of all functions $f_l(x) = l/\sqrt{\pi}\exp(-l^2x^2)$ as we send l to ∞ is infinite. These functions are Gaussians and the area below them is equal to one. (See Fig. 1.4.)

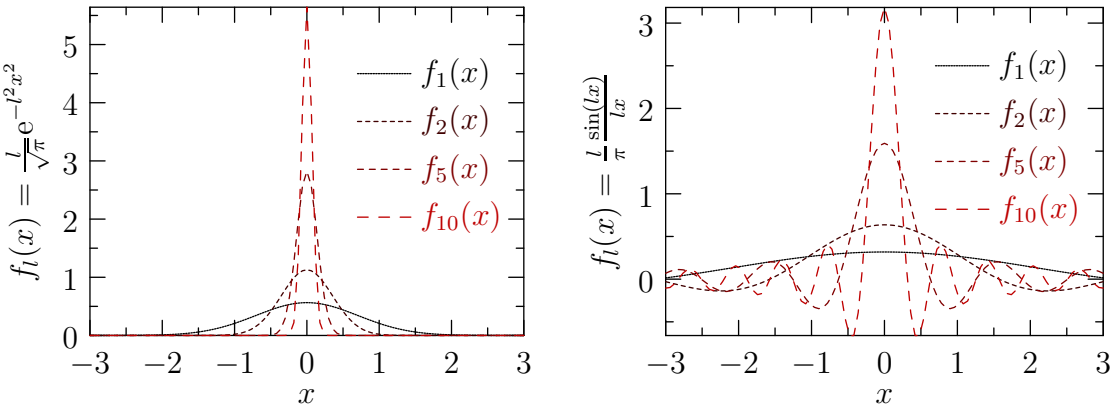


FIG. 1.4. Representations of the Dirac delta $\delta(x)$.

Other representations of the Dirac delta are given by

$$\frac{1}{l\pi}\frac{\sin^2 lx}{x^2}\,, \qquad \frac{1}{l\pi}\frac{1}{x^2+\frac{1}{l^2}}\,, \qquad \frac{1}{\pi}\frac{\sin lx}{x}\,.$$

(See Fig. 1.4.)

The Dirac delta $\delta(x-y)$ is "somehow" the generalization to the continuous case of the Kronecker delta δ^i_j, the latter being defined for discrete indexes i, j as

$$\delta^i_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

By definition, Kronecker delta acts on a vector f^i as

$$f^i = \sum_j \delta^i_j f^j\,.$$

Now, if the indexes i, j become continuous, the various f^j become functions f^x or $f(x)$ and the sum becomes an integral over x, that is

$$f(y) = \int \delta(x-y) f(x)\,\mathrm{d}x\,.$$

1.5 Harmonic oscillator and Hermite polynomials

Find the spectrum and the wave functions of the stationary Schrödinger equation for an harmonic oscillator with potential $V(x) = \frac{1}{2}kx^2$.

The stationary Schrödinger equation is

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} + \frac{1}{2}kx^2\psi(x) = E\psi(x) \; .$$

First, let us change variables :

$$\xi \equiv \alpha x \; , \qquad \alpha \equiv \left(\frac{mk}{\hbar^2}\right)^{1/4} \; .$$

The Schrödinger equation in the new variable ξ reads

$$\frac{\mathrm{d}^2\psi(x(\xi))}{\mathrm{d}\xi^2} + (\lambda - \xi^2)\psi = 0 \; , \tag{1.23}$$

where

$$\lambda \equiv \frac{2E}{\hbar}\left(\frac{m}{k}\right)^{1/2} = \frac{2E}{\hbar\omega_c} \; , \quad \omega_c = \left(\frac{k}{m}\right)^{1/2} \; . \tag{1.24}$$

It is easy to prove that for $\xi \to +\infty$ the function

$$\psi(\xi) = \xi^n \, \mathrm{e}^{-\xi^2/2}$$

satisfies Eqn. (1.23) with $\lambda = 0$. To prove this fact, let us calculate the first and second order derivatives of $\psi(\xi)$:

$$\begin{aligned}
\frac{\mathrm{d}\psi(\xi)}{\mathrm{d}\xi} &= n\xi^{n-1}\,\mathrm{e}^{-\xi^2/2} - \xi^{n+1}\,\mathrm{e}^{-\xi^2/2} \; , \\
\frac{\mathrm{d}^2\psi(\xi)}{\mathrm{d}\xi^2} &= [n(n-1)\xi^{n-2} - n\xi^n - (n+1)\xi^n + \xi^{n+2}]\,\mathrm{e}^{-\xi^2/2} \; .
\end{aligned}$$

For $\xi \to \infty$, the second-order derivative behaves like

$$\frac{\mathrm{d}^2\psi(\xi)}{\mathrm{d}\xi^2} \sim \xi^{n+2}\,\mathrm{e}^{-\xi^2/2} \; ,$$

and for $\lambda = 0$ Eqn. (1.23) is asymptotically fullfilled, as you can verify by direct substition:

$$\frac{\mathrm{d}^2\psi(\xi)}{\mathrm{d}\xi^2} - \xi^2\psi(\xi) \overset{\xi\to\infty}{\sim} \xi^{n+2}\,\mathrm{e}^{-\xi^2/2} - \xi^2\xi^n\,\mathrm{e}^{-\xi^2/2} = 0 \; .$$

Now, we try to find a complete (*i.e.*, not only asymptotic) solution of Eqn. (1.23) of the form

$$\psi(\xi) = H(\xi)\,\mathrm{e}^{-\xi^2/2}\,, \tag{1.25}$$

where $H(\xi)$ is a polynomial in ξ. Inserting Eqn. (1.25) into Eqn. (1.23) yields the following equation for the polynomial $H(\xi)$:

$$\frac{\mathrm{d}^2H(\xi)}{\mathrm{d}\xi^2} - 2\xi\frac{\mathrm{d}H(\xi)}{\mathrm{d}\xi} + (\lambda-1)H(\xi) = 0\,. \tag{1.26}$$

Now, let us make the following *ansatz* for $H(\xi)$:

$$H(\xi) = \xi^s\left[a_0 + a_1\xi + a_2\xi^2 + \ldots\right]\,, \tag{1.27}$$

with $a_0 \neq 0$ and $s \geq 0$. By substituting into Eqn. (1.26) and by equating to zero all the coefficients of the various powers of ξ, we get

$$\begin{aligned}
s(s-1)a_0 &= 0\,,\\
(s+1)sa_1 &= 0\,,\\
(s+2)(s+1)a_2 - (2s+1-\lambda)a_0 &= 0\,,\\
(s+3)(s+2)a_3 - (2s+3-\lambda)a_1 &= 0\,,\\
&\vdots\\
(s+\nu+2)(s+\nu+1)a_{\nu+2} - (2s+2\nu+1-\lambda)a_\nu &= 0\,,
\end{aligned} \tag{1.28}$$

where ν is an integer number. Since $a_0 \neq 0$, we get from the first equation that

$$s = 0 \quad \text{or} \quad s = 1\,.$$

The second equation gives

$$s = 0 \quad \text{or} \quad a_1 = 0\,,$$

or both of them. The third equation gives a_2 as a function of a_0 and the fourth equation gives a_3 as a function of a_1. In general, from $\nu = 3$, the ν–th equation gives $a_{\nu-1}$ as a function of $a_{\nu-3}$.

According to problems presented at the end of this chapter, the wave functions $\psi(x)$ for one dimensional system cannot be degenerate for bound states and moreover — since the potential is even $V(x) = V(-x)$ — they will be either even or odd, that is,

$$\psi(x) = \pm\psi(-x) \ .$$

Now, let us consider separately the two cases $s = 0$ and $s = 1$ that solve the first equation in (1.28)

$s = 0$: The first equation in (1.28) is satisfied and also the second is fulfilled without having to choose $a_1 = 0$. However, in this way the parity would not be defined, since we would have both the zeroth power with a_0 and the first power with a_1. Instead, if we choose $a_1 = 0$, the second equation in (1.28) is still satisfied and we get the *even* polynomials.

$s = 1$: In this case, we get the *odd* polynomials since

$$\xi^s[a_0 + a_1\xi + \ldots] = \xi[a_0 + a_1\xi + \ldots] \ ,$$

and a_1 must be zero in order to satisfy the second equation in (1.28).

In (1.28) there will be a finite or an infinite number of terms depending on how we choose s and λ. If the series did not stop, then the ratio of the coefficients $a_{\nu+2}/a_\nu$ would become for large ν

$$\frac{a_{\nu+2}}{a_\nu} \overset{\nu\to\infty}{\to} \frac{2s + 2\nu + 1 - \lambda}{(s + \nu + 2)(s + \nu + 1)} \sim \frac{2}{\nu} \ , \tag{1.29}$$

but these coefficients are just the same of the expansion of

$$\xi^n \, \mathrm{e}^{2\xi^2} \ , \tag{1.30}$$

this means that $\psi(\xi) = H_n(\xi) \, \mathrm{e}^{-\xi^2/2}$ would become $\xi^n \, \mathrm{e}^{2\xi^2 - \xi^2/2} = \xi^n \, \mathrm{e}^{3/2\xi^2}$ which is not normalizable.

Let us check that the coefficients of the series expansion of (1.30) goes as in Eqn. (1.29). The ν–th term of the series expansion of (1.30) is

$$\xi^n(\xi^2 2)^m \frac{1}{m!}$$

and the $(\nu-1)$–th is

$$\xi^n(\xi^2 2)^{m-1} \frac{1}{(m-1)!} \ ,$$

thus the coefficients are $2^\nu/\nu!$ and $2^{\nu-1}/(\nu-1)!$ respectively and their ratio is $2/\nu$ as in Eqn. (1.29).

Thus, in order to have normalizable wave functions the series (1.28) must stop, that is we must have polynomials with a finite number of terms. From Eqs. (1.28) we see that this happens if

$$\lambda = 2s + 2\nu + 1 \ .$$

In fact, in this way $a_{\nu+2}$ is zero and the following coefficients obtained from $a_{\nu+2}$ will be zero too. Now, let us call $s + \nu = n$, which is an integer number. The above relation becomes

$$\lambda = 2n + 1 \ .$$

By replacing λ with E via Eqn. (1.24) we get

$$\boxed{E_n = \left(n + \frac{1}{2}\right)\hbar\omega_c \ .} \tag{1.31}$$

These are the discrete energy levels. The lowest energy occurs for $n = 0$ and it is given by

$$E_0 = \frac{1}{2}\hbar\omega_c \ ,$$

which is called the *zero-point* energy.

Since ν is finite, H_n will be polynomials and not series. In particular,

$$H_0(\xi) = 1 \ ,$$

$$H_1(\xi) = 2\xi \ ,$$

$$H_2(\xi) = 4\xi^2 - 2 \ , \ldots$$

These polynomials satisfy the equation

$$H_n''(\xi) - 2\xi H_n'(\xi) + 2nH_n(\xi) = 0 \ .$$

and they are referred to as "Hermite polynomials".

We want to stress the fact that also in the case of the harmonic oscillator the discreteness of the energy spectrum comes from the normalization requirement of the wave functions.

1.6 Schrödinger equation for multidimensional problems

Until now, we have dealt exclusively with one-dimensional problems. Now, we turn to the two- and three-dimensional ones. We will focus on those problems where the potential is separable, which in the three-dimensional case and using Cartesian coordinates means

$$V(x,y,z) = V_1(x) + V_2(y) + V_3(z) \ . \tag{1.32}$$

It may happen that a potential is not separable in Cartesian coordinates, but it is separable in another set of coordinates, for example in cylindrical or spherical coordinates. This means that using for example spherical coordinates, we have

$$V(r,\vartheta,\varphi) = V_1(r) + V_2(\vartheta) + V_3(\varphi) \ .$$

In which set of coordinates the potential is separable depends on the symmetries of the system. For the time being, we will consider potentials which are separable in Cartesian coordinates. In this case, the stationary Schrödinger equation reads

$$\left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right) + V_1(x) + V_2(y) + V_3(z)\right]\psi(x,y,z) = E\psi(x,y,z) \ .$$

We look for solutions of the form*

$$\psi(x,y,z) = \psi_1(x)\psi_2(y)\psi_3(z) \ .$$

* We use this ansatz because a general solutions is a linear combination of them: $\tilde{\psi}(x,y,z) = C^{i,j,k}\psi_i(x)\psi_j(y)\psi_k(z)$.

Inserting this into the Schrödinger equation we obtain

$$
\begin{aligned}
&-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_1(x)}{\mathrm{d}x^2}\psi_2(y)\psi_3(z) + V_1(x)\psi_1(x)\psi_2(y)\psi_3(z) + \\
&-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_2(y)}{\mathrm{d}y^2}\psi_1(x)\psi_3(z) + V_2(y)\psi_1(x)\psi_2(y)\psi_3(z) + \\
&-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_3(z)}{\mathrm{d}z^2}\psi_1(x)\psi_2(y) + V_3(z)\psi_1(x)\psi_2(y)\psi_3(z) = E\psi_1(x)\psi_2(y)\psi_3(z)\ .
\end{aligned}
$$

by dividing both sides by $\psi_1(x)\psi_2(y)\psi_3(z)$ we get

$$
\begin{aligned}
&\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_1(x)}{\mathrm{d}x^2} + V_1(x)\psi_1(x)\right]\frac{1}{\psi_1(x)} + \\
&\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_2(y)}{\mathrm{d}y^2} + V_2(y)\psi_2(y)\right]\frac{1}{\psi_2(y)} + \\
&\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_3(z)}{\mathrm{d}z^2} + V_3(z)\psi_3(z)\right]\frac{1}{\psi_3(z)} = E\ .
\end{aligned}
$$

Note that each of the three terms on the left-hand side is a function of a different variable: the first is a function of x, the second of y and the third of z. In order to have the sum equal to a constant, namely E, each term must be a constant:

$$
\begin{aligned}
&\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_1(x)}{\mathrm{d}x^2} + V_1(x)\psi_1(x)\right]\frac{1}{\psi_1(x)} = E_1\ , \\
&\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_2(y)}{\mathrm{d}y^2} + V_2(y)\psi_2(y)\right]\frac{1}{\psi_2(y)} = E_2\ , \\
&\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_3(z)}{\mathrm{d}z^2} + V_3(z)\psi_3(z)\right]\frac{1}{\psi_3(z)} = E_3\ ,
\end{aligned}
$$

with $E_1 + E_2 + E_3 = E$. The initial problem of solving the three-dimensional Schrödinger equation is thus reduced to solving the following three Schrödinger equations in one dimension:

$$
\begin{aligned}
&-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_1(x)}{\mathrm{d}x^2} + V_1(x)\psi_1(x) = E_1\psi_1(x)\ , \\
&-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_2(y)}{\mathrm{d}y^2} + V_2(y)\psi_2(y) = E_2\psi_2(y)\ , \\
&-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_3(z)}{\mathrm{d}z^2} + V_3(z)\psi_3(z) = E_3\psi_3(z)\ .
\end{aligned}
$$

The three solutions are associated with energy eigenvalues E_1, E_2 and E_3 such that

$$E_1 + E_2 + E_3 = E \ .$$

It may happen that by adding three different eigenvalues E_1', E_2' and E_3' we obtain the same energy E:

$$E_1' + E_2' + E_3' = E \ .$$

The three wave functions corresponding to the eigenvalues E_1', E_2' and E_3' — which we shall denote with $\psi_1'(x)$, $\psi_2'(y)$ and $\psi_3'(z)$ — in general will be different from those associated with E_1, E_2 and E_3 — say, $\psi_1(x)$, $\psi_2(y)$ and $\psi_3(z)$. This means that the global wave functions

$$\psi_E(x,y,z) = \psi_1(x)\psi_2(y)\psi_3(z)$$

and

$$\psi_E'(x,y,z) = \psi_1'(x)\psi_2'(y)\psi_3'(z)$$

have the same energy E but they are two different functions. In this case, we speak of *degeneracy*, *i.e.*, there are different wave functions associated to the same energy.

Now, we want to give an example of a potential which is separable in spherical coordinates but not in Cartesian coordinates. We define the spherical coordinates in the usual way:

$$x = r\sin\vartheta\cos\varphi \ , \qquad 0 < r < \infty \ , \tag{1.33a}$$

$$y = r\sin\vartheta\sin\varphi \ , \qquad 0 \leq \vartheta \leq \pi \ , \tag{1.33b}$$

$$z = r\cos\vartheta \ , \qquad 0 \leq \varphi < 2\pi \ . \tag{1.33c}$$

Let us consider the following potential

$$V(r,\vartheta,\varphi) = kr^2 + l^2\tan\vartheta + s^2\sin^2\varphi \ . \tag{1.34}$$

Of course, it is separable in spherical coordinates, since it can be written in the form

$$V(r,\vartheta,\varphi) = V_1(r) + V_2(\vartheta) + V_3(\varphi) \ ,$$

where

$$V_1(r) = kr^2 \ , \quad V_2(\vartheta) = l^2 \tan\vartheta \ , \quad V_3(\varphi) = s^2 \sin^2\varphi \ .$$

By inverting* Eqs. (1.33) and inserting into Eqn. (1.34) we get

$$\tilde{V}(x,y,z) = k(x^2+y^2+z^2) + l^2 \frac{\sqrt{x^2+y^2}}{z} + s^2 \frac{y^2}{x^2+y^2} \ ,$$

which is not separable as a sum $\tilde{V}_1(x) + \tilde{V}_2(y) + \tilde{V}_3(x)$.

1.7 Central potentials

With the word "central potentials" we mean those potentials which, once they are written in polar coordinates, depend only on r, that is

$$V(x,y,z) = U(r) \ .$$

It is not difficult to prove that the Laplacian in spherical coordinates reads

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{r^2 \sin\vartheta}\frac{\partial}{\partial \vartheta}\left(\sin\vartheta \frac{\partial}{\partial \vartheta}\right) + \frac{1}{r^2 \sin^2\vartheta}\frac{\partial^2}{\partial \varphi^2} \ .$$

It should be noted that the transformation from the Cartesian coordinates to the spherical coordinates is not one-to-one along the z-axis (in fact, ϑ can be both 0 or π and φ can take any value). The Laplacian reflects this fact via the presence of singularities in $\vartheta = 0$, $\vartheta = \pi$ and $r = 0$. Thus, the solution of the stationary Schrödinger equation can have spurious singularities in those points due to the singularities of the Laplacian in spherical coordinates.

Using the expression for the Laplacian in spherical coordinates, the stationary Schrödinger equation for a central potential $U(r)$ becomes

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial}{\partial\vartheta}\right) + \frac{1}{r^2}\frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial\varphi^2}\right]\tilde{\psi}(r,\vartheta,\varphi) + U(r)\tilde{\psi}(r,\vartheta,\varphi) = E\tilde{\psi}(r,\vartheta,\varphi) \ . \qquad (1.35)$$

* In the points where it is possible.

We seek solutions of the form

$$\tilde{\psi}(r,\vartheta,\varphi) = R(r)Y(\vartheta,\varphi) \; .$$

Inserting this into Eqn. (1.35) and multiplying by $r^2/(RY)$, we get:

$$\frac{1}{R(r)}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R(r)}{\mathrm{d}r}\right) + \frac{2m}{\hbar^2}[E-U(r)]r^2 = \\ -\left[\frac{1}{Y\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial Y}{\partial\vartheta}\right) + \frac{1}{Y\sin^2\vartheta}\frac{\partial^2 Y}{\partial\varphi^2}\right] \; .$$

Note that the left-hand side depends only on r while the right-hand side is a function only of ϑ and φ. For the two sides to be equal to each other it must be that they both equal a constant, which we indicate with λ. This leads to:

$$\frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) + \frac{2m}{\hbar^2}[E-U(r)]r^2 = \lambda \; , \tag{1.36a}$$

$$\frac{1}{Y\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial Y}{\partial\vartheta}\right) + \frac{1}{Y\sin^2\vartheta}\frac{\partial^2 Y}{\partial\varphi^2} = -\lambda \; . \tag{1.36b}$$

Eqn. (1.36b) is basically related to the quantization of angular momentum and it will be treated in chapter 5. Instead, the solution of Eqn. (1.36a) can be already worked out at this point. Let us rewrite $R(r) = \chi(r)/r$, where $\chi(r)$ is a new function. Inserting this "parametrization" of $R(r)$ into Eqn. (1.36a) we get

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\chi(r)}{\mathrm{d}r^2} + \left[U(r) + \frac{\lambda\hbar^2}{2mr^2}\right]\chi(r) = E\chi(r) \; . \tag{1.37}$$

This is nothing but the Schrödinger equation for a particle moving in a "effective" potential given by the sum of the original potential $U(r)$ plus the "centrifugal" term $\frac{\lambda\hbar^2}{2mr^2}$. We shall see in Chapter 9 that the centrifugal potential is linked to the angular momentum and it is non-zero only when the angular momentum does not vanish. In fact, it is possible to prove that this term represents the rotational kinetic energy.

In the case of the hydrogen atom, $U(r)$ is given by

$$U(r) = -\frac{e^2}{r} \; ,$$

or, if we consider one electron spinning around a nucleus of charge Ze, it is

$$U(r) = -\frac{Ze^2}{r} \, .$$

The radial equation (1.37) becomes

$$-\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) - \frac{Ze^2}{r}R + \frac{\lambda\hbar^2}{2mr^2}R = ER \, ,$$

where m is not the mass of electron, but its reduced mass

$$m = \frac{m_e m_N}{m_e + m_N} \, ,$$

where m_e is the actual mass of the electron and m_N is the mass of the nucleus. Since $m_e \ll m_N$, we have $m \approx m_e$.

Let us introduce the following new variables

$$\rho \equiv \alpha r \, , \quad \alpha^2 \equiv \frac{8m\,|E|}{\hbar^2} \, , \quad \delta = \frac{Ze^2}{\hbar}\left(\frac{m}{2\,|E|}\right)^{1/2} \, .$$

The radial equation above is then transformed into the following one:

$$\frac{1}{\rho^2}\frac{\mathrm{d}}{\mathrm{d}\rho}\left(\rho^2\frac{\mathrm{d}R(r(\rho))}{\mathrm{d}\rho}\right) + \left[\frac{\delta}{\rho} - \frac{1}{4} - \frac{\lambda}{\rho^2}\right]R = 0 \, .$$

This is a differential equation which is well-known, and we shall solve it in chapter 9 after having obtained the value of λ in chapter 5 via the quantization of the angular momentum.

1.8 Problems and Solutions

Problem 1.1. Show that the following wave function

$$\psi(x,t) = A\exp\left[-\frac{\sqrt{Cm}}{2\hbar}x^2 - \frac{i}{2}\sqrt{\frac{C}{m}}t\right] \, ,$$

is a solution of the Schrödinger equation (1.1) for an harmonic oscillator with potential

$$V(x) = \frac{1}{2}Cx^2 \, .$$

Calculate the probability density associated with this wave function and compare the result with the *classical* probability of finding the particle in a interval $\mathrm{d}x$.

SOLUTION. First, we must evaluate the derivatives of ψ:

$$\frac{\partial\psi(x,t)}{\partial t} = -\frac{i}{2}\sqrt{\frac{C}{m}}\psi(x,t) \ ,$$
$$\frac{\partial\psi(x,t)}{\partial x} = -\frac{\sqrt{Cm}}{\hbar}x\psi(x,t) \ ,$$
$$\frac{\partial^2\psi(x,t)}{\partial x^2} = -\frac{\sqrt{Cm}}{\hbar}\psi(x,t) + \frac{Cm}{\hbar^2}x^2\psi(x,t) \ .$$

Replacing these expressions in the Schrödinger equation (1.1) we get

$$\frac{\hbar^2}{2m}\frac{\sqrt{Cm}}{\hbar}\psi(x,t) - \frac{\hbar^2}{2m}\frac{Cm}{\hbar^2}x^2\psi(x,t) + \frac{1}{2}Cx^2\psi(x,t) = i\hbar\left(-\frac{i}{2}\right)\sqrt{\frac{C}{m}}\psi(x,t) \ ,$$

from which it follows

$$\frac{\hbar}{2}\sqrt{\frac{C}{m}}\psi(x,t) - \frac{1}{2}Cx^2\psi(x,t) + \frac{1}{2}Cx^2\psi(x,t) = \frac{\hbar}{2}\sqrt{\frac{C}{m}}\psi(x,t) \ ,$$

which is identically satisfied.

Accordingly to the Born interpretation of the wave function, the quantum probability density $\rho(x,t)$ of finding the particle in x at the time t is given by

$$\rho(x,t) = |\psi(x,t)|^2 = A^2\exp\left[-\frac{\sqrt{Cm}}{\hbar}x^2\right] \ ,$$

which does not depend on t; this is a common feature of all the wave functions corresponding to fixed values of the energy. $\rho(x)$ is plotted in Fig. 1.5(b).

Let us now evaluate the *classical* probability $P(x)\,\mathrm{d}x$ of finding the particle in a small interval $\mathrm{d}x$ centered at x. This probability will be proportional to the time taken by the particle to cross $\mathrm{d}x$, thus it is inversely proportional to the velocity of the particle in x:

$$P(x) = \frac{B^2}{v(x)} \ ,$$

where B is a constant. At a given energy E we have

$$E = \frac{1}{2}mv^2 + \frac{1}{2}Cx^2 \ ,$$

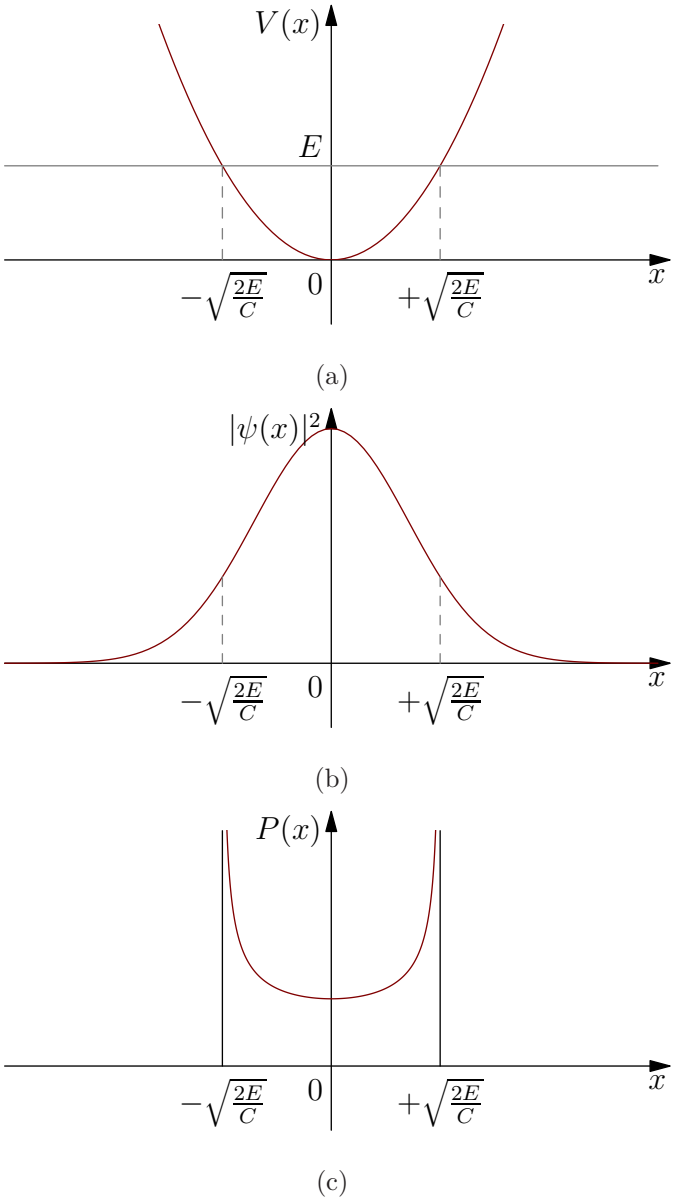


FIG. 1.5. Comparison between (c) classical and (b) quantum probability densities for an harmonic oscillator whose potential is shown in (a).

thus

$$v(x) = \sqrt{\frac{2}{m}}\sqrt{E - \frac{1}{2}Cx^2}\,,$$

and

$$P(x) = \frac{B^2}{\sqrt{\frac{2}{m}}\sqrt{E - \frac{1}{2}Cx^2}}.$$

$P(x)$ is plotted in Fig. 1.5(c).

Note that the probability of finding the particle at the origin $x = 0$ is minimum in the classical case, while it is maximum in the quantum case. The classical particle has its maximum velocity at $x = 0$ and therefore goes rapidly through the origin and the probability of finding the particle there is minimum. In the quantum case instead the situation is just the opposite and the particle is more likely to be found at the origin. Note however that there is a non-null quantum probability of finding the particle also outside the extrema of the oscillation range. The difference between the classical and quantum case decreases at higher energies. At low energies, quantum effects becomes dominant.

$\star \quad \star \quad \star$

Problem 1.2. Show that in a one-dimensional problem the energy spectrum of bound states is always non-degenerate. By degeneracy we mean that there are two or more linearly independent solutions associated to the same energy value.

SOLUTION. Let us prove this fact *ab absurdo* and suppose the opposite is true. Let $\psi_1(x)$ and $\psi_2(x)$ be two linearly independent solutions corresponding to the same energy E:

$$\psi_1''(x) + \frac{2m}{\hbar^2}[E - V(x)]\psi_1(x) = 0\,, \quad \psi_2''(x) + \frac{2m}{\hbar^2}[E - V(x)]\psi_2(x) = 0\,.$$

Dividing the former by ψ_1 and the latter by ψ_2 (and staying away from their zeros) we get

$$\frac{\psi_1''(x)}{\psi_1(x)} = \frac{2m}{\hbar^2}[-E + V(x)]\,, \quad \frac{\psi_2''(x)}{\psi_2(x)} = \frac{2m}{\hbar^2}[-E + V(x)]\,,$$

which implies that

$$\frac{\psi_1''(x)}{\psi_1(x)} = \frac{\psi_2''(x)}{\psi_2(x)} = \frac{2m}{\hbar^2}[-E + V(x)]\,,$$

and from the first equality it follows that

$$\psi_1''\psi_2 - \psi_2''\psi_1 = (\psi_1'\psi_2)' - (\psi_2'\psi_1)' = 0 \, .$$

By integrating the previous equation we get

$$\psi_1'\psi_2 - \psi_2'\psi_1 = \text{const.}$$

This equation must hold for all x and since the wave function of a bound state must be zero at ∞, the constant appearing on the right-hand side of this equation must also be zero, which implies

$$\frac{\psi_1'}{\psi_1} = \frac{\psi_2'}{\psi_2} \, .$$

By integrating, we get

$$\log \psi_1 = \log \psi_2 + \log C \, ,$$

which leads to

$$\psi_1 = C\psi_2 \, .$$

Therefore, the two states are linearly dependent. This in contradiction with the hypothesis and it implies that there is no degeneracy.

$\star \quad \star \quad \star$

Problem 1.3. Show that the first derivatives of the *stationary* wave functions are continuous even if the potential $V(x)$ has a *finite* discontinuity.

SOLUTION. Suppose the discontinuity is at the point x_0. Let us consider an interval $(x_0 - \varepsilon, x_0 + \varepsilon)$ and replace $V(x)$ in such interval with a potential $V_1(x)$ linearly interpolating between $V(x_0 - \varepsilon)$ and $V(x_0 + \varepsilon)$. In this way, we obtain a continuous potential whose solutions satisfy the stationary Schrödinger equation

$$\frac{\mathrm{d}^2\psi_1(x)}{\mathrm{d}x^2} + \frac{2m}{\hbar^2}[E - V_1(x)]\psi_1(x) = 0 \, .$$

By integrating this equation in the interval $(x_0-\varepsilon, x_0+\varepsilon)$ we get

$$\int_{x_0-\varepsilon}^{x_0+\varepsilon} \frac{\mathrm{d}^2\psi_1(x)}{\mathrm{d}x^2}\,\mathrm{d}x = \frac{2m}{\hbar^2}\int_{x_0-\varepsilon}^{x_0+\varepsilon}[V_1(x)-E]\psi_1(x)\,\mathrm{d}x\ ,$$

from which it follows

$$\left.\frac{\mathrm{d}\psi_1(x)}{\mathrm{d}x}\right|_{x=x_0+\varepsilon} - \left.\frac{\mathrm{d}\psi_1(x)}{\mathrm{d}x}\right|_{x=x_0-\varepsilon} = \frac{2m}{\hbar^2}\int_{x_0-\varepsilon}^{x_0+\varepsilon}[V_1(x)-E]\psi_1(x)\,\mathrm{d}x\ . \tag{1.38}$$

In the limit $\varepsilon \to 0$, we have $V_1(x) \to V(x)$ and $\psi_1(x) \to \psi(x)$. Since the integrand on the right-hand side is finite, it remains finite also in the limit $\varepsilon \to 0$. Let us recall now the mean-value theorem, which states that

$$\int_a^b f(x)\,\mathrm{d}x = \bar{f}[b-a]\ ,$$

where $\bar{f}$ is the mean value of the function $f(x)$ over the interval (a,b). If $\bar{f}$ remains finite and $(b-a)\to 0$ then also the integral of $f(x)$ must vanish. This is exactly what happens in the case of Eqn. (1.38), since the integrand is finite. Thus,

$$\lim_{\varepsilon\to 0}\left\{\left.\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\right|_{x=x_0+\varepsilon} - \left.\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\right|_{x=x_0-\varepsilon}\right\} = 0\ ,$$

that is,

$$\left.\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\right|_{x=x_0-0} = \left.\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\right|_{x=x_0+0}\ ,$$

and therefore the derivatives are continuous.

$\star \quad \star \quad \star$

Remark. If the potential $V(x)$ of the previous problem had an ∞ discontinuity, then it would have been impossible to apply the mean value theorem. In such cases, the derivative ψ' could have a finite gap. We will see an example in the next problems.

Problem 1.4. The potential energy of a system is

$$U(x) = \tilde{U}(x) + \alpha\delta(x-x_0)\ ,$$

where $\tilde{U}(x)$ is a bounded function. Which is the behavior of the solution of the Schrödinger equation $\psi(x)$ and its derivative near the point x_0?

SOLUTION. The Schrödinger equation reads

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2}+[\tilde{U}(x)+\alpha\delta(x-x_0)]\psi(x)=E\psi(x)\ . \tag{1.39}$$

$\psi(x)$ must be continuous because of its probabilistic meaning. Let us investigate if also the derivative remain continuous. By integrating Eqn. (1.39) between $x_0-\varepsilon$ and $x_0+\varepsilon$ with $\varepsilon>0$ and by sending $\varepsilon\to 0$ we obtain

$$-\frac{\hbar^2}{2m}\int_{x_0-\varepsilon}^{x_0+\varepsilon}\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2}\,\mathrm{d}x=-\int_{x_0-\varepsilon}^{x_0+\varepsilon}[\tilde{U}(x)+\alpha\delta(x-x_0)-E]\psi(x)\,\mathrm{d}x\ ,$$

from which it follows

$$-\frac{\hbar^2}{2m}\left[\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\bigg|_{x=x_0+\varepsilon}-\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\bigg|_{x=x_0-\varepsilon}\right]=-\alpha\psi(x_0)-\int_{x_0-\varepsilon}^{x_0+\varepsilon}[\tilde{U}(x)-E]\psi(x)\,\mathrm{d}x\ .$$

If $\psi(x)$ is continuous and $\tilde{U}(x)$ is bounded we can use the mean-value theorem, so in the limit $\varepsilon\to 0$ the integral on the right-hand side of the previous equation vanishes and we get

$$\lim_{\varepsilon\to 0}-\frac{\hbar^2}{2m}\left[\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\bigg|_{x=x_0+\varepsilon}-\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\bigg|_{x=x_0-\varepsilon}\right]=-\alpha\psi(x_0)\ ,$$

that is,

$$\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\bigg|_{x=x_0+0}-\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\bigg|_{x=x_0-0}=\frac{2m\alpha}{\hbar^2}\psi(x_0)\ . \tag{1.40}$$

This means that the derivative is discontinuous at the point x_0. Remember the result proved in Problem (1.3): for potentials with a *finite* gap the derivative of $\psi(x)$ must be continuous. In this case however the Dirac delta has in x_0 an *infinite* jump, so the result of Problem (1.3) cannot hold and in fact the derivative is discontinuous.

⋆ ⋆ ⋆

Problem 1.5. Find the energy levels and the wave functions of the bound states of a particle in a potential $U(x)=-\alpha\delta(x)$, with $\alpha>0$.

SOLUTION. The Schrödinger equation is

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2}-\alpha\delta(x)\psi(x)=E\psi(x)\ .$$

For $x \neq 0$, this is nothing but the usual equation of a free particle:

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} = E\psi(x) \ .$$

The solutions of the equation above are easily found to be exponentials, however here we consider bound states so we must require that the solutions go to zero at $\pm\infty$, thus the only allowed solutions take the form

$$\psi(x) = \begin{cases} A\,\mathrm{e}^{-kx} & \text{if } x > 0 \\ \\ B\,\mathrm{e}^{kx} & \text{if } x < 0 \end{cases} ,$$

where k is related to E. The relation between k and E can be found by inserting the solution $\psi(x)$ into the Schrödinger equation of the free particle; we get

$$-\frac{\hbar^2}{2m}A(k^2)\,\mathrm{e}^{-kx} = EA\,\mathrm{e}^{-kx} \ ,$$

from which it follows

$$E = -\frac{\hbar^2 k^2}{2m} \ .$$

E is negative (bound state) and k has not yet been determined. Now, let us determine A and B. By using Eqn. (1.40), the jump in the derivative of $\psi(x)$ is given by

$$\left.\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\right|_{x=x_0+0} - \left.\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\right|_{x=x_0-0} = \frac{2m(-\alpha)}{\hbar^2}\psi(x_0) \ .$$

[Note that in problem 1.4 we have $\alpha\delta(x - x_0)$ while here we have $-\alpha\delta(x)$.] Hence,

$$A(-k)\,\mathrm{e}^{-k\cdot 0} - kB\,\mathrm{e}^{k\cdot 0} = -\frac{2m\alpha}{\hbar^2}A \ .$$

In the right-side of the latter equation we could have used B as well. In fact, because of the continuity condition, we would get $A = B$. Therefore,

$$A(-k) - kA = -\frac{2m\alpha}{\hbar^2}A \ ,$$

that is,

$$k = \frac{m\alpha}{\hbar^2} \ .$$

This implies that there is only one value of k which corresponds to a bound state of our problem, its energy is given by

$$E = -\frac{\hbar^2}{2m}k^2 = -\frac{\hbar^2}{2m}\frac{m^2\alpha^2}{\hbar^4} = -\frac{m\alpha^2}{2\hbar^2} \ .$$

A is found by requiring $\psi(x)$ to be normalized: $\int |\psi(x)|^2 \, \mathrm{d}x = 1$. The normalized wave function corresponding to the only bound state is given by $\psi(x) = \sqrt{k}\, \mathrm{e}^{-k|x|}$. (Prove the latter result as an homework.)

$\star \quad \star \quad \star$

Problem 1.6. Prove that under an even potential $V(x) = V(-x)$ the eigenfunctions solution of the stationary Schrödinger equation for bound states are even or odd, but in any case they always have well-defined parity.

SOLUTION. Let us start with

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\right]\psi(x) = E\psi(x) \ .$$

Now, we prove that also $\psi(-x)$ is a solution of the same equation with the same value of energy E. Replacing x with $-x$ in the equation above yields

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(-x)\right]\psi(-x) = E\psi(-x) \ ,$$

and exploiting the fact that $V(x) = V(-x)$ we get

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\right]\psi(-x) = E\psi(-x) \ ,$$

This proves that both $\psi(x)$ and $\psi(-x)$ are solutions of the Schrödinger equation with the same energy E.

Now, suppose $\psi(x)$ and $\psi(-x)$ are bound states belonging to the discrete spectrum. From Problem (1.2) we already know that in this case there cannot be degeneracy, hence $\psi(x)$ and $\psi(-x)$ must be linearly dependent:

$$\psi(x) = C\psi(-x) \ .$$

Normalizing

$$\int_{-\infty}^{+\infty} |\psi(x)|^2 \, \mathrm{d}x = 1 \, ,$$

we get

$$|C|^2 \int_{-\infty}^{+\infty} |\psi(-x)|^2 \, \mathrm{d}x = 1 \, ,$$

and thus $|C|^2 = 1$. From this it follows that (apart from a phase) $C = \pm 1$ and therefore $\psi(x) = \pm\psi(-x)$, that is the wave functions are even or odd. This result holds for the discrete spectrum. For the continuous part of the spectrum there can be degeneracy and so it is not always true that $\psi(x) = C\psi(-x)$ and the energy eigenfunctions may have parity that it is not well-defined.

$\star \quad \star \quad \star$

Problem 1.7. A two-dimensional harmonic oscillator has Hamiltonian given by

$$H = \frac{p_x^2}{2M} + \frac{1}{2}kx^2 + \frac{p_y^2}{2M} + \frac{1}{2}ky^2 \, . \tag{1.41}$$

Study the spectrum and calculate the degeneracy.

Solution. The two-dimensional stationary Schrödinger equation reads in this case

$$\left[-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial x^2} - \frac{\hbar^2}{2M}\frac{\partial^2}{\partial y^2} + \frac{1}{2}kx^2 + \frac{1}{2}ky^2 \right] \psi(x,y) = E\psi(x,y) \, . \tag{1.42}$$

Since the potential is separable, we can look for solutions of the form

$$\psi(x,y) = \psi_1(x)\psi_2(y) \, .$$

So Eqn. (1.42) becomes equivalent to the following two equations:

$$\left(-\frac{\hbar^2}{2M}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{1}{2}kx^2 \right) \psi_1(x) = E_1\psi_1(x) \, ,$$
$$\left(-\frac{\hbar^2}{2M}\frac{\mathrm{d}^2}{\mathrm{d}y^2} + \frac{1}{2}ky^2 \right) \psi_2(y) = E_2\psi_2(y) \, ,$$

with $E_1 + E_2 = E$. E_1 and E_2 are given by the solution of the one-dimensional harmonic oscillator, that is,

$$E_1 = \left(n + \frac{1}{2}\right)\hbar\omega_c \ , \quad E_2 = \left(m + \frac{1}{2}\right)\hbar\omega_c \ ,$$

where $\omega_c = (k/M)^{1/2}$ and n, m are (non-negative) integer numbers labeling the eigenvalues. The total energy is

$$E = E_1 + E_2 = (n + m + 1)\hbar\omega_c = (N + 1)\hbar\omega_c \ ,$$

where $N = n + m$ is any non-negative integer number. The corresponding wave functions are

$$\psi_N(x, y) = \psi_{1,n}(x)\psi_{2,m}(y) \ ,$$

where we have denoted with $\psi_{1,n}(x)$ the eigenfunctions corresponding to the eigenvalue $E_{1,n} = (n + 1/2)\hbar\omega_c$ and with $\psi_{2,m}(y)$ the eigenfunctions with eigenvalue $E_{2,m} = (m + 1/2)\hbar\omega_c$, $N = n + m$. It is clear that there are many different values of n and m such that $n + m = N$ with N fixed and whose corresponding wave functions have the same energy. For example, we can choose $n' = n - 1$ and $m' = m + 1$ and clearly we get the same N:

$$N = n + m = n' + m' \ ,$$

and in the same way we can choose $n'' = n - 2$ and $m'' = m + 2$. Associated to these different n, m there are different total wave functions $\psi(x, y)$:

$$\psi'_N(x, y) = \psi_{1,n-1}(x)\psi_{2,m+1}(y) \ ,$$
$$\psi''_N(x, y) = \psi_{1,n-2}(x)\psi_{2,m+2}(y) \ .$$

Therefore, there is degeneracy. Consider $N = 0$. In this case, the only allowed combination of n, m is

$$n = 0 \ , \quad m = 0 \ ,$$

since N must be non-negative, so there is only one solution in this case, namely

$$\psi_0(x, y) = \psi_{1,0}(x)\psi_{2,0}(y) \ .$$

The energy of this level is $E_0 = \hbar\omega_c$ and there is not degeneracy. Let us consider now $N = 1$. In this case, there are two allowed combinations of n, m, namely

$$n = 1\,, \quad m = 0 \qquad \text{and} \qquad n = 0\,, \quad m = 1\,.$$

The two corresponding total wave functions are

$$\psi_1(x,y) = \psi_{1,1}(x)\psi_{2,0}(y)\,,$$
$$\psi_1'(x,y) = \psi_{1,0}(x)\psi_{2,1}(y)\,.$$

$\star \quad \star \quad \star$

Problem 1.8. Calculate the spectrum and the eigenfunctions of a two dimensional infinite well potential of sides a and b:

$$V(x,y) = \begin{cases} 0 & \text{if } 0 < x < a\,, \quad 0 < y < b\,; \\ +\infty & \text{otherwise}\,. \end{cases}$$

Solution. The Schrödinger equation in the range $0 < x < a$ and $0 < y < b$ is

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

and since we have infinite wells, so that the particle cannot escape, the boundary conditions are

$$\psi(0,y) = \psi(a,y) = 0\,, \quad \psi(x,0) = \psi(x,b) = 0\,.$$

The potential is separable, hence we seek solutions of the form

$$\psi(x,y) = \psi_1(x)\psi_2(y)\,.$$

Following the general procedure outlined in this section, it is easy to prove that the functions

$\psi_1(x)$ and $\psi_2(x)$ must satisfy the following equations:

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_1(x)}{\mathrm{d}x^2} = E_1\psi_1(x)\ , \tag{1.43a}$$

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_2(y)}{\mathrm{d}y^2} = E_2\psi_2(y)\ , \tag{1.43b}$$

$$\tag{1.43c}$$

with $E_1 + E_2 = E$. Using the following boundary conditions

$$\psi_1(0) = \psi_1(a) = 0\ , \quad \psi_2(0) = \psi_2(b) = 0\ ,$$

the boundary conditions for $\psi(x, y)$ are automatically satisfied. Therefore, the solutions of Eqs. (1.43) are those of two one-dimensional infinite wells, one in the x-direction and the other in the y-direction:

$$\left\{\begin{array}{rcl} \psi_{1,n}(x) & = & \sqrt{\frac{2}{a}}\sin\frac{n\pi x}{a} \\ E_1^{(n)} & = & \frac{\hbar^2}{2m}\left(\frac{n\pi}{a}\right)^2 \end{array}\right. , \quad \left\{\begin{array}{rcl} \psi_{2,m}(y) & = & \sqrt{\frac{2}{b}}\sin\frac{m\pi y}{b} \\ E_2^{(m)} & = & \frac{\hbar^2}{2m}\left(\frac{m\pi}{b}\right)^2 \end{array}\right. ,$$

where n, m are integer numbers, $n, m > 0$. Thus, the total eigenfunctions of the two-dimensional infinite well potential are given by

$$\psi_{n,m}(x, y) = \sqrt{\frac{2}{a}}\sqrt{\frac{2}{b}}\sin\frac{n\pi x}{a}\sin\frac{m\pi y}{b}\ ,$$

and the corresponding energies are

$$E_{n,m} = \frac{\hbar^2}{2m}\left[\left(\frac{\pi}{a}\right)^2 n^2 + \left(\frac{\pi}{b}\right)^2 m^2\right]\ .$$

$$\star \quad \star \quad \star$$

Chapter 2

AN INTRODUCTION TO HILBERT SPACES FOR PHYSICISTS

2.1 Basic definitions

Definition 2.1 (Hilbert space)**.** An Hilbert space is a *linear space* endowed with a *scalar product* and such that it is *complete* and *separable*.

Let us now clarify the meaning of the various terms.

2.1.1 Linear space

A set S is a "linear space" over a numerical field $\mathbb{N}$ if

1. for all pair of elements ψ, φ of S it is possible to assign one and only one well-defined element χ in S, called the "sum" of ψ and φ and denoted with $\chi = \psi + \varphi$;

2. for all pair $a \in \mathbb{N}$ and $\psi \in S$ it is possible to assign one and only one element in S called "product" of ψ by a and denoted with $a\psi$.

The above operations must satisfy the following properties:

$$\begin{aligned} \psi + \varphi &= \varphi + \psi \,, & a(b\varphi) &= (ab)\varphi \,, \\ (\varphi + \psi) + \chi &= \varphi + (\psi + \chi) \,, & 1 \cdot \varphi &= \varphi \,, \\ a(\varphi + \psi) &= a\varphi + a\psi \,, & 0 \cdot \varphi &= 0 \cdot \psi = \omega \,, \\ (a + b)\psi &= a\psi + b\psi \,, \end{aligned}$$

where $\omega \in S$ is referred to as the "null" element of S. ω has the following properties:

$$\forall \varphi \in S\ , \quad \varphi + \omega = \varphi\ ,$$

and it is unique. Moreover,

$$a\omega = \omega\ .$$

Definition 2.2 (linear independence)**.** The elements $\varphi_1, \varphi_2, \ldots, \varphi_n \in S$ are said to be "linearly independent" if the equation

$$a_1\varphi_1 + a_2\varphi_2 + \ldots + a_n\varphi_n = \omega$$

implies $a_1 = a_2 = \ldots = a_n = 0$.

Definition 2.3 (dimension)**.** A linear space S is said to have "finite dimension" n if it is always possible to find n linearly independent elements in S but it is impossible to find $n + 1$ linearly independent elements. If there are arbitrarily many linearly independent vectors the linear space is said to be infinite-dimensional.

2.1.2 Scalar product

Definition 2.4. A "scalar product" is a mapping $S \times S$ into $\mathbb{C}$, denoted with (φ, ψ), satisfying the following properties:

1. $(\varphi, a\psi_1 + b\psi_2) = a(\varphi, \psi_1) + b(\varphi, \psi_2)$;

2. $(\varphi, \psi) = (\psi, \varphi)^*$;

3. $(\varphi, \varphi) \geq 0$ and $(\varphi, \varphi) = 0$ if and only if $\varphi = \omega$.

From these properties, one can easily prove that

$$(a\varphi_1 + b\varphi_2, \psi) = a^*(\varphi_1, \psi) + b^*(\varphi_2, \psi)\ .$$

The norm of ψ is defined as

$$\|\psi\| \equiv \sqrt{(\psi, \psi)} \, .$$

The distance between two elements $\varphi, \psi \in S$ is given by

$$d_{\varphi,\psi} \equiv \|\varphi - \psi\| \ ,$$

and it is possible to prove that

1. $\|\varphi\| \geq 0$;

2. $\|a\varphi\| = |a| \, \|\varphi\|$;

3. $\|\varphi + \psi\| \leq \|\varphi\| + \|\psi\|$ (triangle inequality).

Theorem 2.1 (Schwarz inequality)**.**

$$|(\varphi, \psi)| \leq \|\varphi\| \, \|\psi\| \ .$$

Definition 2.5 (Orthogonality)**.** Two vectors ψ and φ in an Hilbert space are said to be "orthogonal" if $(\psi, \varphi) = 0$.

2.1.3 Sequences and convergence

In this section, we collect some basic definitions.

Definition 2.6. A sequence $\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots$ of elements of S is said to **converge** to an element φ if the numerical sequence $\|\varphi_1 - \varphi\|, \ldots, \|\varphi_n - \varphi\|, \ldots$ converges to zero.

Definition 2.7. A set is called **closed** if it contains all its limit points.

Definition 2.8. A set U is said to be **dense** in S if U together with all its limit points is equal to S.

Theorem 2.2 (Cauchy convergence criterion)**.** A sequence $\varphi_1, \varphi_2, \ldots, \varphi_n \ldots$ satisfies the Cauchy criterion if for each $\varepsilon > 0$, there exists an integer N such that $\|\varphi_m - \varphi_n\| \leq \varepsilon$ for all $n, m \geq N$. This is a *necessary* condition for the convergence of a sequence .

Definition 2.9 (completeness)**.** A set S is said to be **complete** if all Cauchy sequences in S do converge to an element of S.

Definition 2.10 (separability)**.** A set S is said to be **separable** if for any $\varphi \in S$ and $\varepsilon > 0$ there exists a φ_n, of a countable sequence, such that $\|\varphi - \varphi_n\| < \varepsilon$. (Here, n depends on both ε and φ.)

For example, the real line $\mathbb{R}$ is a separable set since the rational numbers are dense in $\mathbb{R}$ and there are countably many rational numbers.

Example 2.1. Consider the space $\mathcal{C}_0(0,1)$ of all continuous functions $f(t)$, $0 \leq t \leq 1$, with the norm given by

$$\|f\|^2 = \int_0^1 |f(t)|^2 \, \mathrm{d}t \; .$$

We shall prove that this space is not complete.

Let us consider the sequence of functions $f_n(t)$ defined as

$$f_n(t) = \begin{cases} 0 & \text{if } 0 \leq t \leq \frac{1}{2} - \frac{1}{n} \\ 1 & \text{if } \frac{1}{2} + \frac{1}{n} \leq t \leq 1 \\ \text{linear in } t & \text{if } \frac{1}{2} - \frac{1}{n} < t < \frac{1}{2} + \frac{1}{n} \end{cases} .$$

It is easy to prove that the **discontinuous** function

$$f(t) = \begin{cases} 0 & \text{if } 0 \leq t \leq \frac{1}{2} \\ 1 & \text{if } \frac{1}{2} < t \leq 1 \end{cases}$$

satisfies the relation

$$\lim_{n \to \infty} \int_0^1 |f_n(t) - f(t)|^2 \, \mathrm{d}t = 0 \; .$$

Thus, the sequence $f_n(t)$ being convergent is also a Cauchy sequence, but it does not converge to a continuous functions, that is, to an element in $\mathcal{C}_0(0,1)$, therefore $\mathcal{C}_0(0,1)$ is not complete and cannot be an Hilbert space.

Before considering specific examples of Hilbert spaces, we need some other definition.

2.1.4 Orthonormal systems

Definition 2.11. A sequence $\{\varphi_n\} \in S$ is said to be **orthonormal** if

$$(\varphi_m, \varphi_n) = \delta_{m,n} \ ,$$

for any two elements φ_n and φ_m in the sequence.

Definition 2.12. An orthonormal set $\{\varphi_n\} \in S$ is said to be a **complete orthonormal system** if there exists no $\varphi \neq \omega$ orthogonal to every element of the sequence, *i.e.*,

$$(\varphi, \varphi_k) = 0 \quad \forall k \quad \Rightarrow \quad \varphi = \omega \ .$$

For example, in the three-dimensional Euclidean space the set of unit vectors along the x- and y-directions is not a complete orthonormal set, even if it is an orthonormal one.

Definition 2.13. Let U be a set of elements of an Hilbert space. The set of all the linear combinations of the vectors in U is again a linear space, called the **linear variety** spanned by U. If we add all its limit points, we obtain a closed linear variety that contains U, which we shall denote with $[U]$: this is the closed linear variety spanned by U.

Theorem 2.3. In an n-dimensional Hilbert space, every orthonormal set has a number of elements less than or equal to n. It is a complete orthonormal system only if it has n elements.

Theorem 2.4. In a infinite-dimensional Hilbert space, every orthonormal set has a finite or countably infinite number of elements. If it is a complete orthonormal system then it must have countably infinite many elements.

It should be emphasized that the converse is not generally true: it may happen that a orthonormal set with an infinite number of elements is *not* a complete orthonormal system. In fact , it is sufficient to remove one element from a complete set in order to get a non-complete one which has infinite elements.

Let $\{\varphi_k\}$ be an orthonormal set. Given an element φ, we define the coefficients

$$a_k = (\varphi_k, \varphi) \ ,$$

and build the series

$$\sum_k a_k \varphi_k = \chi \ .$$

This series is always convergent in an Hilbert space, moreover $\chi - \varphi$ is orthogonal to each φ_k and if the set is complete this means that $\chi - \varphi = \omega$.

Theorem 2.5. Let $\{\varphi_n\}$ be an orthonormal set. It is complete if and only if one of the following three conditions is satisfied:

1. for any $\varphi \in S$,

$$\varphi = \sum_k a_k \varphi_k \ ,$$

   where $a_k = (\varphi_k, \varphi)$;

2. the closed linear variety spanned by $\{\varphi_k\}$ is equal to S;

3. for any pair of elements $\varphi, \psi \in S$,

$$(\varphi, \psi) = \sum_k (\varphi, \varphi_k)(\varphi_k, \psi) \ .$$

2.2 Examples of Hilbert spaces

2.2.1 The space $l^2(\infty)$

$l^2(\infty)$ is the linear space having as elements all the sequences of complex numbers $\{a_k\}$ such that $\sum_{k=1}^{n} |a_k|^2$ converges as $n \to +\infty$.

The basic operations in $l^2(\infty)$ are defined as follows. Let $a = \{a_k\}$, $b = \{b_k\}$ and $c = \{c_k\}$ in $l^2(\infty)$ and $\alpha, \beta \in \mathbb{C}$ (i.e. they belong to the complex numbers).

1. We define
$$\alpha a = \{\alpha a_k\} , \qquad a + b = \{a_k + b_k\} .$$

2. The scalar product is defined as
$$(a, b) = \sum_k a_k^* b_k .$$
Note that it satisfies all the properties to be actually a scalar product over $l^2(\infty)$ since
$$(a, b) = (b, a)^* ,$$
$$(a, \alpha b + \beta c) = \alpha(a, b) + \beta(a, c) ,$$
and (a, a) is a positive real number and it is zero only if $a = (0, 0, 0, \ldots)$.

It is possible to prove that $l^2(\infty)$ is complete and separable.

Theorem 2.6. Every infinite-dimensional Hilbert space is isomorphic to $l^2(\infty)$.

2.2.2 The space $\mathcal{L}^2$

$\mathcal{L}^2$ is the space of the square-integrable complex-valued functions ψ of k real variables $(q_1, \ldots, q_k)$, *i.e.*,
$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} |\psi(q_1, \ldots, q_k)|^2 \, \mathrm{d}q_1 \cdots \mathrm{d}q_k < \infty .$$

The basic operations in $\mathcal{L}^2$ are defined as follows. Given any two functions f and g in $\mathcal{L}^2$ and a complex number α we define αf to be the function $\alpha f(q_1, \ldots, q_k)$ and $f + g$ to be the function $f(q_1, \ldots, q_k) + g(q_1, \ldots, q_k)$. It is possible to prove that αf and $f + g$ defined in such way actually belong to $\mathcal{L}^2$. The case for αf is trivial, let us consider the case of $f + g$.

We have to prove that $\int \cdots \int |f + g|^2 \, \mathrm{d}q_1 \cdots \mathrm{d}q_k < \infty$. For simplicity, we shall denote the integration element as $\mathrm{d}q = \mathrm{d}q_1 \cdots \mathrm{d}q_k$. We have:

$$\int |f + g|^2 \, \mathrm{d}q = \int |f|^2 \, \mathrm{d}q + \int |g|^2 \, \mathrm{d}q + \int 2\Re f^* g \, \mathrm{d}q \; .$$

Since

$$2\Re f^* g \leq 2 \, |f| \, |g| \; ,$$

it follows

$$|f + g|^2 \leq |f|^2 + |g|^2 + 2 \, |f| \, |g| = (|f| + |g|)^2 \; .$$

Moreover, as

$$(|f| - |g|)^2 \geq 0 \; ,$$

it also true that

$$|f|^2 + |g|^2 \geq 2 \, |f| \, |g| \; .$$

Hence,

$$|f + g|^2 \leq |f|^2 + |g|^2 + |f|^2 + |g|^2 = 2 \, |f|^2 + 2 \, |g|^2 \; ,$$

and since both $\int |f|^2 \, \mathrm{d}q$ and $\int |g|^2 \, \mathrm{d}q$ are $< \infty$ also $\int |f + g|^2 \, \mathrm{d}q < \infty$.

The $\mathcal{L}^2$ scalar product is defined as

$$(f, g) = \int f^*(q_1, \ldots, q_k) g(q_1, \ldots, q_k) \, \mathrm{d}q_1 \ldots \mathrm{d}q_k$$

The properties of the scalar product are all satisfied except the one which says that $(f, f) = 0$ implies $f = 0$. In fact, $(f, f) = 0$ means $\int |f|^2 \, \mathrm{d}q = 0$ and this does not imply that $f = 0$, it is sufficient that $|f| \neq 0$ over a set of zero measure $(\mathrm{d}q)$.

Therefore, we consider equivalent two functions which differ on a subset having zero measure. The null element of our Hilbert space is given by the (equivalent class of) the functions that are zero almost everywhere.

Theorem 2.7. $\mathcal{L}^2$ is infinite-dimensional.

Proof. Let $O_1, \ldots, O_n$ be n non-overlapping subsets having Lebesgue measure greater than zero, but finite. Let $f_1(q_1, \ldots, q_k)$ take value 1 in O_1 and zero elsewhere. Since $\int |f_1|^2 \, \mathrm{d}q$ is equal to the measure of O_1, which for hypotesis is finite, f_1 belongs to $\mathcal{L}^2$. In the same way we can define the functions $f_2, \ldots, f_n$ and these n functions are linearly independent. Since the construction holds for an arbitrary large n, this means that for all n it is possible to specify n linearly independent functions in $\mathcal{L}^2$, thus $\mathcal{L}^2$ is infinite dimensional. ■

von Neumann in his book "*Mathematical Foundations of Quantum Mechanics*" has shown that $\mathcal{L}^2$ is complete and separable.

2.3 Operators in Hilbert spaces

We start with a more general setting: we consider operators between generic spaces, that is, mappings from a set R into another set R'. We shall denote the mapping with $\hat{T}$ and write

$$\varphi' = \hat{T}\varphi \, ,$$

where $\varphi \in R$ and $\varphi' \in R'$. The set R on which $\hat{T}$ acts is called the **domain** of $\hat{T}$ and the set R' obtained from R via $\hat{T}$ is the **range**. Sometimes, it may be useful to denote the domain of an operator $\hat{T}$ using the notation D_T instead of R, and the range using the notation R_T instead of R'. Hereafter, we shall deal with operators over linear spaces. First of all, let us recall some basic definitions.

Sum of two operators. The **sum** of two operators $\hat{T}_1$ and $\hat{T}_2$ is the operator $\hat{T}_1 + \hat{T}_2$ defined as the operator $\hat{T}_3$ which, once applied to φ, gives the state $\hat{T}_1\varphi + \hat{T}_2\varphi$.

The latter is nothing but the state obtained as the sum of the two vectors $\hat{T}_1\varphi$ and $\hat{T}_2\varphi$: the sum of these two vectors is a well-defined vector, since we are in a linear space. It is easy to see that the domain of $\hat{T}_3$ is $D_{T_3} = D_{T_1} \cap D_{T_2}$.

Inverse of an operator. If the operator $\hat{T}$ maps different vectors into different vectors,

$$\varphi' = \hat{T}\varphi \ ,$$

it is possible to define the inverse $\hat{T}^{-1}$ of $\hat{T}$ as

$$\varphi = \hat{T}^{-1}\varphi' \ .$$

Note that $D_{T^{-1}} = R_T$ and $R_{T^{-1}} = D_T$.

Additive operators. The additive operators are those for which the following property holds:

$$\hat{T}(\varphi + \psi) = \hat{T}\varphi + \hat{T}\psi \ ,$$

for all $\varphi, \psi \in D_T$.

Homogeneous operators. if for all $a \in \mathbb{C}$ and for all $\psi \in D_T$ we have

$$\hat{T}(a\psi) = a\hat{T}\psi \ ,$$

then the operator is said to be **homogeneous**. If instead

$$\hat{T}(a\psi) = a^*\hat{T}\psi \ ,$$

then the operator is said to be **anti-homogeneous**.

Linear operators. The **linear** operators are those which are both additive and homogeneous. Those which are instead additive and anti-homogeneous are called **anti-linear**.

Theorem 2.8. The domain and the range of a linear (or anti-linear) operator are linear spaces.

Theorem 2.9. Necessary and sufficient condition for a linear (or anti-linear) operator $\hat{T}$ to be invertible is that $\hat{T}\varphi = \omega$ implies $\varphi = \omega$.

We introduce the notion of continuity for an operator.

Definition 2.14. An operator $\hat{T}$ is said to be **continuous** at $\psi_0 \in S$ if for each $\varepsilon > 0$ there exists a δ_ε such that for all $\tilde{\psi}$ satisfying

$$\|\tilde{\psi} - \psi_0\| \leq \delta_\varepsilon$$

it follows that $\|\hat{T}\tilde{\psi} - \hat{T}\psi_0\| < \varepsilon$.

Related to the notion of continuity is the notion of boundeness.

Definition 2.15. An operator $\hat{T}$ is said to be **bounded** in a subset $U \subset S$ if there exists a real number R such that for all $\varphi \in U$

$$\|\hat{T}\varphi\| < R\,\|\varphi\| \ .$$

Definition 2.16. Two operators, indicated by $\hat{T}$ and $\hat{T}^\dagger$, are said to be **adjoint** of each other if they have the same domain and in this domain

$$(\hat{T}\varphi, \psi) = (\varphi, \hat{T}^\dagger\psi) \ , \tag{2.1}$$

for all φ, ψ.

Eqn. (2.1) implies also that (prove the following as an homework):

$$(\varphi, \hat{T}\psi) = (\hat{T}^\dagger\varphi, \psi) \ .$$

You can prove also that

$$(\hat{T}^\dagger)^\dagger = \hat{T} \ , \quad (a\hat{T})^\dagger = a^*\hat{T}^\dagger \ , \quad (\hat{T}_1\hat{T}_2)^\dagger = \hat{T}_2^\dagger\hat{T}_1^\dagger \ .$$

(Do this last one as an homework.)

Definition 2.17. If $\hat{T} = \hat{T}^\dagger$, the operator is said to be **Hermitian**.

Definition 2.18. An operator $\hat{U}$ is said to be **unitary** if

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \mathbb{1} \ ,$$

As a consequence of the previous definition, it follows

$$\hat{U}^\dagger = \hat{U}^{-1} \ .$$

Thus, unitary operators are invertible.

Theorem 2.10. An operator is unitary if and only if it is defined everywhere and norm-preserving.

Let us see only one part of the proof, namely, if an operator is unitary then it preserves the norms:

$$(\hat{U}\varphi, \hat{U}\psi) = (\varphi, \hat{U}^\dagger\hat{U}\psi) = (\varphi, \psi) \ .$$

Theorem 2.11. If $\hat{T}$ is Hermitian and $\hat{U}$ is unitary, then also $\hat{U}\hat{T}\hat{U}^{-1} \equiv \hat{T}'$ is Hermitian.

2.3.1 Projection operators

Let $\mathscr{N}$ be any subset of some Hilbert space $\mathscr{H}$ and consider the set of all the elements of $\mathscr{H}$ which are orthogonal to all the elements of $\mathscr{N}$. This is a closed linear subset which we shall denote with $\mathscr{H} - \mathscr{N}$.

Theorem 2.12. Let $\mathscr{N}$ be a closed linear subset of an Hilbert space $\mathscr{H}$. Then each $\varphi \in \mathscr{H}$ can be decomposed in one and only one way as the sum $\varphi = \psi + \chi$ of two vectors $\psi \in \mathscr{N}$ and $\chi \in \mathscr{H} - \mathscr{N}$. ψ is called the **projection** of φ in $\mathscr{N}$.

Definition 2.19. Let $\mathscr{N}$ be a closed linear subset in an Hilbert space $\mathscr{H}$. The operator $\hat{P}_\mathscr{N}$ (defined everywhere in $\mathscr{H}$) which assigns to each $\varphi \in \mathscr{H}$ its projection

in $\mathscr{N}$ (as defined by the previous theorem) is called the **projection operator** of $\mathscr{N}$. Using the notation employed in the statement of the previous theorem,

$$\hat{P}_{\mathscr{N}}\varphi = \psi \ .$$

Theorem 2.13. It is possible to prove that $\hat{P}_{\mathscr{N}}$ is linear, Hermitian and idempotent (that is, $\hat{P}^2_{\mathscr{N}} = \hat{P}_{\mathscr{N}}$).

Theorem 2.14. An operator $\hat{P}$ defined everywhere is a projection operator if and only if it is linear, Hermitian and idempotent. The closed linear subset $\mathscr{N}$ on which $\hat{P}$ is the projector is the one spanned by all the vectors ψ_N such that $\hat{P}\psi_N = \psi_N$.

Theorem 2.15. Let $\hat{P}$ be a projector. It is possible to prove the following identities:

$$\|\hat{P}\varphi\|^2 = (\hat{P}\varphi, \varphi) \ , \quad \|\hat{P}\varphi\| \leq \|\varphi\| \ , \quad \|\hat{P}\varphi - \hat{P}\psi\| = \|\hat{P}(\varphi - \psi)\| \leq \|\varphi - \psi\| \ ,$$

and moreover

$$\hat{P}_{\mathscr{H}-\mathscr{N}} = \mathbb{1} - \hat{P}_{\mathscr{N}} \ .$$

Homework 2.1. Let $\hat{P}$ and $\hat{Q}$ be projector operators on the two linear varietys $\mathscr{M}$ and $\mathscr{N}$, respectively. Prove that $\hat{P}\hat{Q}$ is a projector operator if and only if $\hat{P}$ and $\hat{Q}$ commutes among themselves and show that in that case $\hat{P}\hat{Q}$ project on the variety $\mathscr{M} \cap \mathscr{N}$.

Homework 2.2. If $\hat{P}_1$, $\hat{P}_2$, $\hat{P}_3$, ..., $\hat{P}_N$ are projector operators, prove that $\hat{P}_1 + \hat{P}_2 + \hat{P}_3 + \ldots + \hat{P}_N$ is a projector operator if and only if $\hat{P}_i\hat{P}_j = \delta_{i,j}\hat{P}_i$.

2.4 The eigenvalue problem

A key role in quantum mechanics is played by the eigenvalue problem, that means to solve the equation

$$\hat{H}\psi_r = h_r\psi_r \ , \tag{2.2}$$

where $\hat{H}$ is a given linear* operator, h_r is a real or complex number known as **eigenvalue** and ψ_r is any non-zero† element belonging to the Hilbert space and which is referred to as **eigenvector** or **eigenstate** corresponding to h_r. The name **eigenfunction** is also used having in mind a concrete realization of the Hilbert space as a space of functions, which is often the case in quantum mechanics. The set of all the eigenvalues h_r is the so-called **spectrum** of $\hat{H}$.

Theorem 2.16.
- The *discrete* spectrum of a linear, *Hermitian* operator is made of a finite or countably infinite set of *real* numbers.
- Eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proof. First, we prove the fact that the eigenvalues of a (linear) Hermitian operator are real. Since $\hat{H}$ is Hermitian,

$$(\hat{H}\psi_r, \psi_r) = (\psi_r, \hat{H}\psi_r) \ .$$

Since ψ_r is eigenvector of $\hat{H}$ with eigenvalue h_r, the previous equation reads

$$(h_r\psi_r, \psi_r) = (\psi_r, h_r\psi_r) \ ,$$

that is,

$$h_r^*(\psi_r, \psi_r) = h_r(\psi_r, \psi_r) \ .$$

Therefore, $h_r^* = h_r$, which means that h_r is real.

* The eigenvalue problem can be stated without problem also for non-Hermitian operators, however in the following we are mainly concerned with Hermitian ones, since this is the case relevant for quantum mechanics.

† Of course, Eqn. (2.2) would hold trivially for any real or complex number h_r if ψ_r was the null vector of the Hilbert space. Moreover, if ψ_r is any solution of Eqn. (2.2) it is easily shown that $\alpha\psi_r$ is a solution too, for any real or complex number α. So, it is sufficient to consider normalized solutions: $\|\psi_r\| = 1$. However, for the continuous spectrum (see later) normalization must be interpreted in the sense of distributions.

Let us consider two different eigenvectors ψ_r and ψ_s corresponding to different eigenvalues h_r and h_s, respectively: we have

$$\begin{aligned}(\hat{H}\psi_r, \psi_s) &= (\psi_r, \hat{H}\psi_s) \ , \\ (h_r\psi_r, \psi_s) &= (\psi_r, h_s\psi_s) \ , \\ h_r^*(\psi_r, \psi_s) &= h_s(\psi_r, \psi_s) \ . \end{aligned} \tag{2.3}$$

From Eqn. (2.3) and using the fact that h_r is real, we get

$$(h_r - h_s)(\psi_r, \psi_s) = 0 \ .$$

Since $h_r \neq h_s$, we obtain

$$(\psi_r, \psi_s) = 0 \ ,$$

that is, ψ_r and ψ_s are orthogonal and thus also linear independent. ■

In the case of continuous spectrum, the Hilbert space must be enlarged to include "improper" eigenstates which formally satisfy Eqn. (2.2) but that strictly speaking do not belong to the Hilbert space and that are normalized in the generalized sense of the Dirac-delta that we mentioned in the previous chapter. We have already encountered such situation discussing plane wave function solutions of the free particle Hamiltonian. The reader should realize at this point that the stationary Schrödinger equation is nothing but an eigenvalue problem for the Hamiltonian differential operator in the Hilbert space $\mathcal{L}^2$ of the square-integrable wave functions.

We now introduce the concept of **degeneracy** of an eigenvalue. When two or more *linearly independent* eigenvectors $\psi_r, \tilde{\psi}_r, \ldots$ correspond to the same eigenvalue h_r,

$$\hat{H}\psi_r = h_r\psi_r \qquad \text{and} \qquad \hat{H}\tilde{\psi}_r = h_r\tilde{\psi}_r \ ,$$

we say that h_r is a *degenerate* eigenvalue. Any linear combination of $\psi_r, \tilde{\psi}_r, \ldots$ is an

eigenvector with the same eigenvalue h_r:

$$\begin{aligned}\hat{H}(\alpha\psi_r + \beta\tilde{\psi}_r) &= \alpha\hat{H}\psi_r + \beta\hat{H}\tilde{\psi}_r \\ &= \alpha h_r\psi_r + \beta h_r\tilde{\psi}_r \\ &= h_r(\alpha\psi_r + \beta\tilde{\psi}_r) \ .\end{aligned}$$

That is, for any given eigenvalue h_r all linear independent eigenvectors corresponding to h_r form a (closed) linear eigenvariety. For any eigenvariety one can extract two or more (depending on the dimension of the eigenvariety) mutually orthogonal eigenvectors that span the whole variety.

The set of all discrete and continuous eigenvectors, ψ_k, ψ_λ, is said to be **complete** if for any $\varphi \in \mathscr{H}$

$$\varphi = \sum_k (\psi_k, \varphi)\psi_k + \int \mathrm{d}\lambda(\psi_\lambda, \varphi)\psi_\lambda \ , \tag{2.4}$$

and for any pair $\varphi, \chi \in \mathscr{H}$

$$(\varphi, \chi) = \sum_k (\varphi, \psi_k)(\psi_k, \chi) + \int \mathrm{d}\lambda(\varphi, \psi_\lambda)(\psi_\lambda, \chi) \ . \tag{2.5}$$

(Here, $\mathscr{H}$ denotes the Hilbert space.)

Hermitian operators whose spectrum is complete in the sense specified above are called **hypermaximal** operator. In the following, we shall assume that all the Hermitian operators are hypermaximal.

Remark. There are eigenvalues and eigenvectors also for non-Hermitian operators.

Theorem 2.17. If ψ_i is eigenvector of the operator $\hat{A}$ with eigenvalue a_i:

$$\hat{A}\psi_i = a_i\psi_i \ ,$$

then

$$\hat{A}^n\psi_i = a_i^n\psi_i \ ,$$

for every positive integer n.

Theorem 2.18. If $\hat{A}$ is a linear Hermitian operator, the equation $\hat{A}^n\psi = 0$ is equivalent to $\hat{A}\psi = 0$.

Theorem 2.19. Let

$$F(\hat{H}) = \hat{H}^n + a_1\hat{H}^{n-1} + \ldots + a_n = 0$$

be the lowest-degree equation satisfied by the linear Hermitian operator $\hat{H}$. Then

- the roots of $F(x) = 0$ are never degenerate;
- the roots of $F(x) = 0$ are the only eigenvalues of $\hat{H}$;
- the set of eigenvectors of $\hat{H}$ is complete.

The equation $F(\hat{H}) = 0$ is called the **secular equation**.

2.5 Functions of one operator

We will specify here what we mean by a function $F(\hat{H})$ of an operator $\hat{H}$.

Let us suppose that $\hat{H}$ is a hyper-maximal Hermitian operator acting on an Hilbert space $\mathscr{H}$. We define $F(\hat{H})$ in terms of functions of the eigenvalues of $\hat{H}$, namely, $F(\hat{H})$ is the operator which has the same eigenvectors ψ_r of $\hat{H}$ and $F(h_r)$ are the corresponding eigenvalues, *i.e.*,

$$\hat{H}\psi_r = h_r\psi_r \qquad \text{and} \qquad F(\hat{H})\psi_r = F(h_r)\psi_r \ .$$

It is not even necessary for F to be defined over all x, it is sufficient that F is defined for the set of points $x = h_r$.

If $\hat{H}$ is hypermaximal, $F(\hat{H})$ is well-defined for all vectors ψ: from Eqn. (2.4), *i.e.*,

$$\psi = \sum_r \alpha_r\psi_r + \int \alpha_\lambda\psi_\lambda \,\mathrm{d}\lambda \ ,$$

it follows

$$F(\hat{H})\psi = \sum_r \alpha_r F(h_r)\psi_r + \int \alpha_\lambda F(h_\lambda)\psi_\lambda \, \mathrm{d}\lambda \, .$$

For this state to be well-defined, it must belong to the Hilbert space. For example, if the Hilbert space is $\mathscr{H} = \mathcal{L}^2$, one must have

$$\sum |F(h_r)\alpha_r|^2 + \int |\alpha_\lambda F(h_\lambda)|^2 \, \mathrm{d}\lambda < \infty \, .$$

The adjoint of $F(\hat{H})$ is indicated by $F^\dagger(\hat{H})$ and it is such that

$$(\varphi, F(\hat{H})\psi) = (F^\dagger(\hat{H})\varphi, \psi) \, .$$

It is easy to prove that

$$F^\dagger(\hat{H})\psi_r = F^*(h_r)\psi_r \, .$$

2.6 Commutativity and compatibility

If ψ is eigenvector of both $\hat{A}$ and $\hat{B}$, then $[\hat{A}, \hat{B}]\psi = 0$. In fact, from

$$\hat{A}\psi = a\psi \qquad \text{and} \qquad \hat{B}\psi = b\psi$$

one gets

$$\begin{aligned} [\hat{A}, \hat{B}]\psi = \hat{A}\hat{B}\psi - \hat{B}\hat{A}\psi &= \hat{A}(b\psi) - \hat{B}(a\psi) \\ = b\hat{A}\psi - a\hat{B}\psi \quad &= ba\psi - ab\psi = 0 \, . \end{aligned}$$

Theorem 2.20. Two hyper-maximal Hermitian operators share a complete set of eigenvectors if and only if they commute.

Proof. Let us suppose that there is no degeneracy. Suppose $[\hat{A}, \hat{B}] = 0$. Let $\{\varphi_{b_n}\}$ be a complete set of eigenstates of $\hat{B}$ and ψ_a be an eigenvector of $\hat{A}$ corresponding to the eigenvalue a, *i.e.*,

$$\hat{A}\psi_a = a\psi_a \, .$$

Let us expand ψ_a on the eigenvectors φ_{b_n}:*

$$\psi_a = \sum_{b_n} c_a^{b_n} \varphi_{b_n} \ . \tag{2.6}$$

Since

$$\hat{A}\psi_a = a\psi_a \quad \Rightarrow \quad \left(\hat{A} - a\right)\psi_a = 0 \ ,$$

it follows from Eqn. (2.6) that

$$\sum_{b_n} (\hat{A} - a) c_a^{b_n} \varphi_{b_n} = 0 \ . \tag{2.7}$$

Now, we shall prove that the state

$$\tilde{\psi} \equiv (\hat{A} - a) c_a^{b_n} \varphi_{b_n}$$

is an eigenvector of $\hat{B}$ with eigenvalue b_n:

$$\begin{aligned} \hat{B}\tilde{\psi} = \hat{B}\left[(\hat{A} - a) c_a^{b_n} \varphi_{b_n}\right] &= (\hat{A} - a)\hat{B} c_a^{b_n} \varphi_{b_n} = (\hat{A} - a) c_a^{b_n} b_n \varphi_{b_n} \\ &= b_n \left[(\hat{A} - a) c_a^{b_n} \varphi_{b_n}\right] = b_n \tilde{\psi} \end{aligned}$$

It is possible to prove that eigenvectors belonging to different eigenvalues are linearly independent (prove this fact as homework) thus Eqn. (2.7) (which is a linear combination of eigenvectors of $\hat{B}$) leads to

$$(\hat{A} - a) c_a^{b_n} \varphi_{b_n} = 0$$

and hence

$$\hat{A}\left(c_a^{b_n} \varphi_{b_n}\right) = a\left(c_a^{b_n} \varphi_{b_n}\right) \ ,$$

that is, $c_a^{b_n} \varphi_{b_n}$ are eigenvectors of $\hat{A}$ with eigenvalue a, but we already know they are also eigenvectors of $\hat{B}$ with eigenvalue b_n, therefore the various $c_a^{b_n} \varphi_{b_n}$ are simultaneously eigenvectors of both $\hat{A}$ and $\hat{B}$. Eqn. (2.6) is the expansion of one eigenstate of

* For simplicity, we are assuming that $\hat{B}$ does not have a continuous spectrum, otherwise Eqn. (2.6) must be replaced by the more general Eqn. (2.4).

$\hat{A}$ on the eigenstates of $\hat{B}$. Since the set of eigenstates of $\hat{A}$ is complete, any $\psi \in \mathscr{H}$ can be written as

$$\psi = \sum_{a_n} c_\psi^{a_n} \psi_{a_n}$$

but the eigenstates ψ_{a_n} can be expanded in terms of the eigenstates of $\hat{B}$, therefore we can conclude that ψ can be expanded in common eigenstates of $\hat{A}$ and $\hat{B}$.

For a proof of the converse, let $\{\varphi_n\}$ be the complete set of common eigenstates of $\hat{A}$ and $\hat{B}$. Clearly,

$$[\hat{A}, \hat{B}]\varphi_n = 0 \ , \tag{2.8}$$

because we have

$$\begin{aligned}[\hat{A}, \hat{B}]\varphi_n &= (\hat{A}\hat{B} - \hat{B}\hat{A})\varphi_n = \hat{A}(b_n\varphi_n) - \hat{B}(a_n\varphi_n) \\ &= b_n a_n \varphi_n - a_n b_n \varphi_n = 0 \ .\end{aligned}$$

Now, let ψ be any vector in the Hilbert space. We can expand ψ over the basis of φ_n

$$\psi = \sum_n c_n \varphi_n \ ,$$

From this formula and using the previous result (2.8) we find immediately $[\hat{A}, \hat{B}]\psi = 0$, thus $\hat{A}$ and $\hat{B}$ commute because ψ is any element of the Hilbert space. ■

2.7 Complete set of commuting operators

The main result of the last section, namely the theorem on simultaneous diagonalizability of two commuting Hermitian operators, has been obtained assuming no degeneracy.

Now, let us suppose there are two linearly independent eigenstates φ_r and $\tilde{\varphi}_r$ corresponding to the same (degenerate) eigenvalue a_r of $\hat{A}$. Clearly, the eigenvalue a_r is not sufficient to uniquely characterize its eigenvectors. Let $\hat{B}$ be an hypermaximal Hermitian operator commuting with $\hat{A}$. In the eigenvariety associated with the eigenvalue a_r, we look for eigenvectors of $\hat{B}$. Such eigenstates are linear combinations of

φ_r and $\tilde{\varphi}_r$, and we shall denote them with ψ_r and $\tilde{\psi}_r$. Suppose that they correspond to two different eigenvalues b_r and $\tilde{b}_r$ of $\hat{B}$, respectively. Then, we can say that the two eigenstates ψ_r and $\tilde{\psi}_r$ are uniquely specified by a pair of different eigenvalues:

$$\psi_r \to (a_r, b_r) \,, \qquad \tilde{\psi}_r \to (a_r, \tilde{b}_r) \,.$$

If this is actually the case, the degeneracy is removed. Otherwise, we can iterate the procedure and look for a third operator $\hat{C}$ which commutes both with $\hat{A}$ and $\hat{B}$. We look for linear combinations of ψ_r and $\tilde{\psi}_r$, say χ_r and $\tilde{\chi}_r$, that are also eigenstates of $\hat{C}$. In general, we repeat the procedure until all degeneracies have been lifted. If the eigenvalues of $\hat{C}$ associated to χ_r and $\tilde{\chi}_r$ are different, say, $c_r, \tilde{c}_r$, then the triples (a_r, b_r, c_r) and $(a_r, b_r, \tilde{c}_r)$ are different and characterize uniquely χ_r and $\tilde{\chi}_r$. If this is the case, $\hat{A}$, $\hat{B}$ and $\hat{C}$ are said to form a **complete** set of commuting operators.

Definition 2.20. A set of mutually commuting hypermaximal operators is said to be **complete** when specifying the eigenvalues with respect to all the operators, we can determine uniquely a common eigenvector.

Theorem 2.21. Let $\{\hat{A}, \hat{B}, \hat{C}, \ldots\}$ be a complete set of commuting hypermaximal operators. Any other hyper-maximal operator $\hat{F}$ which commutes with all the operators in this set must be a function of $\hat{A}, \hat{B}, \hat{C}, \ldots$

Definition 2.21. A subspace S in the domain of an operator $\hat{A}$ is called **invariant** if by applying $\hat{A}$ to any of the elements in S we get a new state which belongs to S too. Therefore, the image of an invariant subspace S is again within S.

Definition 2.22. A set of operators is said to be **irreducible** if they have no common invariant subspace.

Theorem 2.22. If a set of operators $\hat{A}, \hat{B} \ldots$ is irreducible, then any operator which commutes with all of them is a multiple of the identity.

2.8 Dirac notation

We summarize briefly how to translate the notation used so far in the Dirac notation.

- $|\psi\rangle$ = ket. They represent vectors in the Hilbert space.
- $\langle\psi|$ = bra. They represent the space of *linear* mappings $\psi(|\varphi\rangle)$ that associate to every state $|\varphi\rangle$ a corresponding complex number given by the scalar product (ψ, φ).
- scalar product: $(\psi, \varphi) \to \langle\psi|\varphi\rangle$;
- eigenvalue problem:

$$\hat{H}\psi_r = h_r\psi_r \quad \to \quad \hat{H}\,|\psi_r\rangle = h_r\,|\psi_r\rangle \ ;$$

- superposition of states:

$$|\psi\rangle = \alpha\,|\psi_1\rangle + \beta\,|\psi_2\rangle \ , \qquad |\psi\rangle = \int_{\xi_1}^{\xi_2} f(\xi)\,|\xi\rangle\,\mathrm{d}\xi \ ;$$

- **Theorem**. The object $\hat{O} \equiv |u\rangle\,\langle v|$ acts as a linear operator.

  Proof. Let us apply $\hat{O}$ to a state $|\psi\rangle$:

$$\hat{O}\,|\psi\rangle = |u\rangle\,\langle v|\psi\rangle \ .$$

  Now, $\langle v|\psi\rangle$ is nothing but a number that multiplies the state $|u\rangle$, thus $\hat{O}\,|\psi\rangle$ has produced a state. Let us now prove how it acts on a linear superposition of states:

$$\begin{aligned}\hat{O}\,[\alpha\,|\psi\rangle + \beta\,|\varphi\rangle] &= |u\rangle\,\langle v|\,[\alpha\,|\psi\rangle + \beta\,|\varphi\rangle] \\ &= \alpha\,|u\rangle\,\langle v|\psi\rangle + \beta\,|u\rangle\,\langle v|\varphi\rangle \\ &= \alpha\hat{O}\,|\psi\rangle + \beta\hat{O}\,|\varphi\rangle \ ,\end{aligned}$$

  so this proves that $\hat{O} = |u\rangle\,\langle v|$ is a linear operator. ■

- **Dual space**: Given a ket $|\psi\rangle$, the corresponding bra $\langle\psi|$ is defined as that linear map for which

$$\langle\psi|\psi\rangle = \|\psi\|^2 \ .$$

- **Operators**: Let us suppose that acting with an operator $\hat{A}$ on a state $|\varphi\rangle$ we get the state $|\varphi'\rangle \equiv \hat{A}\,|\varphi\rangle$. The scalar product of this state with another state $|\psi\rangle$ is

$$\langle\psi|\varphi'\rangle = \langle\psi|\left(\hat{A}\,|\varphi\rangle\right) \ ;$$

  to indicate this scalar product we introduce the notation $\langle\psi|\hat{A}\varphi\rangle$.

- **Adjoint**: The adjoint $\hat{A}^\dagger$ of an operator $\hat{A}$ is defined by

$$\langle\hat{A}^\dagger\psi|\varphi\rangle = \langle\psi|\hat{A}\varphi\rangle \ ,$$

  where we have used the notation previously introduced. From this relation, it is clear that $\hat{A}$ acts on the space of kets, while $\hat{A}^\dagger$ acts on the space of bras, *i.e.*, if

$$|\psi'\rangle = \hat{A}\,|\psi\rangle \ ,$$

  then

$$\langle\psi|\,\hat{A}^\dagger = \langle\psi'| \ .$$

2.9 Spectral decomposition

Theorem 2.23. For a hypermaximal Hermitian operator $\hat{H}$ having spectrum h_n and corresponding eigenstates $|h_n\rangle$ the following **spectral decomposition** formula holds:

$$\hat{H} = \sum_{n=1}^{\infty} h_n\,|h_n\rangle\,\langle h_n| \ . \tag{2.9}$$

In other words, $\hat{H}$ is the sum of the projector operators $|h_n\rangle\,\langle h_n|$ multiplied by the corresponding eigenvalues h_n.

For operators having also a continuous spectrum $|\xi_h\rangle$, Eqn. (2.9) must be generalized as

$$\hat{H} = \sum_{n=1}^{\infty} h_n \, |h_n\rangle \, \langle h_n| + \int_{\xi_{h_1}}^{\xi_{h_2}} \xi_h \, |\xi_h\rangle \, \langle \xi_h| \, \mathrm{d}\xi_h \ . \tag{2.10}$$

Proof. We restrict ourselves to the case (2.9) and we assume no degeneracy.

Since $\hat{H}$ is hyper-maximal, its eigenstates form a complete basis for the Hilbert space. Therefore, any state $|\psi\rangle$ in the Hilbert space can be expanded in the following way

$$|\psi\rangle = \sum_{n}^{\infty} c_n \, |h_n\rangle \ . \tag{2.11}$$

The operator identity (2.9) must be understood in the sense that if we apply the operator $\hat{H}$ to some state $|\psi\rangle$ the result $\hat{H} \, |\psi\rangle$ is just the same state we would obtain applying the right-hand side of Eqn. (2.9), *i.e.*:

$$\hat{H} \, |\psi\rangle = \left(\sum_{n} h_n \, |h_n\rangle \, \langle h_n| \right) |\psi\rangle \ . \tag{2.12}$$

Let us prove this formula. By inserting the expansion formula (2.11) in the left-hand side of Eqn. (2.12) we get

$$\hat{H} \, |\psi\rangle = \hat{H} \left(\sum_{n} c_n \, |h_n\rangle \right) = \sum_{n} c_n \hat{H} \, |h_n\rangle = \sum_{n} c_n h_n \, |h_n\rangle \ ,$$

while the right-hand side of Eqn. (2.12) is nothing but

$$\left(\sum_{n} h_n \, |h_n\rangle \, \langle h_n| \right) |\psi\rangle = \sum_{n} h_n \, |h_n\rangle \, \langle h_n|\psi\rangle \ .$$

Our goal would be met if we were able to prove that the expansion coefficients in Eqn. (2.11) are actually given by $c_n = \langle h_n|\psi\rangle$. This result on c_n can be easily derived directly by means of Eqn. (2.11):

$$\langle h_i|\psi\rangle = \sum_{n} c_n \underbrace{\langle h_i|h_n\rangle}_{\delta_{n,i}} = c_i \ ,$$

hence $c_n = \langle h_n|\psi\rangle$. This completes the proof. ∎

Here is a summary of some important identities encountered in this chapter. First of all, remember the important Eqn. (2.4), *i.e.*:

$$\varphi = \sum_k (\psi_k, \varphi)\psi_k + \int \mathrm{d}\lambda (\psi_\lambda, \varphi)\psi_\lambda \ .$$

We can write it using the Dirac notation as follows:

$$|\varphi\rangle = \sum_k \langle\psi_k|\varphi\rangle \, |\psi_k\rangle + \int \mathrm{d}\lambda \, \langle\psi_\lambda|\varphi\rangle \, |\psi_\lambda\rangle \ ,$$

or

$$|\varphi\rangle = \sum_k |\psi_k\rangle \, \langle\psi_k|\varphi\rangle + \int \mathrm{d}\lambda \, |\psi_\lambda\rangle \, \langle\psi_\lambda|\varphi\rangle \ .$$

This is possible because $\langle\psi_k|\varphi\rangle$ is a number. Since the above expression must hold for every state $|\varphi\rangle$ in the Hilbert space, it is usually written as an operator identity by dropping $|\varphi\rangle$ in every term:

$$\mathbb{1} = \sum_k |\psi_k\rangle \, \langle\psi_k| + \int \mathrm{d}\lambda \, |\psi_\lambda\rangle \, \langle\psi_\lambda| \ , \tag{2.13}$$

where $\mathbb{1}$ is the identity operator. Eqn. (2.13) is referred to as **decomposition or resolution of the identity**. It means that applying the left-hand side to a state $|\varphi\rangle$ we get the same result as if we apply the right-hand side to the same state.

Let us now turn to the **spectral decomposition**, Eqn. (2.10):

$$\hat{H} = \sum_{n=1}^{\infty} h_n \, |h_n\rangle \, \langle h_n| + \int_{\xi_{h_1}}^{\xi_{h_2}} \mathrm{d}\xi_h \, |\xi_h\rangle \, \langle\xi_h| \, \xi_h \ . \tag{2.14}$$

It is then easy to generalize this to a generic function $F(\hat{H})$ of the operator $\hat{H}$ as :

$$F(\hat{H}) = \sum_{n=1}^{\infty} F(h_n) \, |h_n\rangle \, \langle h_n| + \int_{\xi_{h_1}}^{\xi_{h_2}} F(\xi_h) \, |\xi_h\rangle \, \langle\xi_h| \, \mathrm{d}\xi_h \ . \tag{2.15}$$

2.10 Tensor Product

The tensor product, usually indicated with the symbol $\otimes$, is an operation which can be applied to any couple of vector-spaces even infinite-dimensional ones like the

Hilbert spaces. Before giving more mathematical details let us illustrate the case with a physical example. Let us consider a system made of two particles **1** and **2** and take for each particle a particular wave function: $\Psi(r_1)$ and $\widetilde{\Psi}(r_2)$. The product $\Psi(r_1)\widetilde{\Psi}(r_2)$ represents a wave function of the combined system. Of course it is not the most general wave function for the combined system. This will be a $\Phi(r_1, r_2)$ which is not in general factorized into two parts like the wave function above. It will be in general a linear combination of factorized wave functions . In fact let us look at $\Phi(r_1, r_2)$ as a wave function of particle **1**, then it can be expanded on a basis of wave functions of particle **1** as:

$$\Phi(r_1, r_2) = \sum C_i \Psi_i(r_1). \tag{2.16}$$

Clearly the coefficients C_i must be functions of r_2, i.e.:

$$\Phi(r_1, r_2) = \sum C_i(r_2)\Psi_i(r_1)$$

and we can interpret the $C_i(r_2)$ as wave functions of r_2 multiplied by some coefficients a_i like:

$$C_i(r_2) = a_i\widetilde{\Psi}_i(r_2)$$

and the relation (2.16) becomes:

$$\Phi(r_1, r_2) = \sum_i a_i\widetilde{\Psi}_i(r_2)\Psi_i(r_1). \tag{2.17}$$

So we can say that the space of wave functions of the combined system is formed by the linear combinations of products of wave functions of particle **1** and of wave functions of particle **2**. If the Hilbert space of particle **1** is indicated with $\mathscr{H}^{(1)}$ and the one of particle **2** with $\mathscr{H}^{(2)}$, the Hilbert space of the combined system of the two particles ,whose wave functions we have built in the example above, is defined as the *tensor product* of the two separate Hilbert spaces and indicated as:

$$\mathscr{H} \equiv \mathscr{H}^{(1)} \otimes \mathscr{H}^{(2)} \tag{2.18}$$

The states in Eqn. (2.17) can be indicated also in the following way:

$$\Phi(r_1, r_2) = \sum a_i \widetilde{\Psi}_i(r_2) \otimes \Psi_i(r_1)$$

The observables associated to particle **1** will be indicated with the symbol $\widehat{O}_{(1)}$ and those associated to particle **2** as $\widehat{O}_{(2)}$. If $\widehat{O}_{(1)}$ acts on $\Psi(r_1)$ as follows:

$$\widehat{O}_{(1)}\Psi(r_1) = \Psi'(r_1)$$

then on the tensor product it will act as follows:

$$\widehat{O}_{(1)}(\Psi(r_1)\widetilde{\Psi}(r_2)) = \Psi'(r_1)\widetilde{\Psi}(r_2). \tag{2.19}$$

Similarly an observable $\widehat{O}_{(2)}$ acting on particle **2** as :

$$\widehat{O}_{(2)}\Psi(r_2) = \Psi'(r_2)$$

will act on the tensor product space as

$$\widehat{O}_{(2)}(\Psi(r_1)\widetilde{\Psi}(r_2)) = \Psi(r_1)\widetilde{\Psi}'(r_2) \tag{2.20}$$

An equivalent manner to indicate this procedure is the following: an operator $\widehat{O}_{(1)}$ acting on the Hilbert space $\mathscr{H}_1$ get extended, in the tensor product space $\mathscr{H} \equiv \mathscr{H}_1 \otimes \mathscr{H}_2$, into the following operator:

$$\widehat{O}_{(1)} \Longrightarrow \widehat{O}_{(1)} \otimes \mathbb{1}$$

where $\mathbb{1}$ is the identity operator acting on the Hilbert space of particle **2**. The notation above summarize the operation indicated in Eqn. (2.19) where the state $\widetilde{\Psi}(r_2)$ is left intact.

The same procedure we follow for an operator $\widehat{O}_{(2)}$ acting on the states of particle **2**. In the tensor product space this operator get extended as follows:

$$\widehat{O}_{(2)} \Longrightarrow \mathbb{1} \otimes \widehat{O}_{(2)}$$

In general the operators acting in the Hilbert space tensor product of the two, will have the form

$$\widehat{O}(r_1, r_2) \equiv \sum \alpha_i \widehat{A}_i(r_1) \otimes \widehat{B}_i(r_2)$$

and will act on a generic state belonging to the tensor product space:

$$\widetilde{\Psi}(r_1, r_2) = \sum C_j \Phi_j(r_1) \otimes \Psi_j(r_2)$$

in the following manner:

$$\widehat{O}(r_1, r_2)\widetilde{\Psi}(r_1, r_2) = \sum_{i,j} \alpha_i C_j \widehat{A}_i(r_1)\Phi_j(r_1) \otimes \widehat{B}_i(r_2)\Psi_j(r_2) \tag{2.21}$$

One last remark: as $\hat{r}_1$ and $\hat{r}_2$ refer to different particle, they commute. The same happens for the operator $\widehat{A}_i(r_1)$ and $\widehat{B}_i(r_2)$

2.11 Problems and Solutions

Problem 2.1. Prove that

$$(|u\rangle \langle v|)^\dagger = |v\rangle \langle u| \ .$$

SOLUTION. Let $\hat{O} \equiv |u\rangle \langle v|$. We know that

$$|\tilde{\psi}\rangle = \hat{O} |\psi\rangle \quad \Rightarrow \quad \langle\psi| \hat{O}^\dagger = \langle\tilde{\psi}| \ .$$

Let us now investigate if it possible from

$$(|u\rangle \langle v|)^\dagger |\psi\rangle \equiv |\tilde{\psi}\rangle \ , \tag{2.22}$$

to prove that:

$$(|v\rangle \langle u|) |\psi\rangle = |\tilde{\psi}\rangle \ . \tag{2.23}$$

We know that Eqn. (2.22) is equivalent to

$$\begin{aligned}\langle\tilde{\psi}| &= \langle\psi|\left(|u\rangle\langle v|\right)\\ &= \langle\psi|u\rangle\langle v| \ ,\end{aligned}$$

i.e., $\langle\tilde{\psi}| = a\langle v|$ with $a = \langle\psi|u\rangle$. Now, taking the dual we get (see the proof afterwards)

$$|\tilde{\psi}\rangle = a^*|v\rangle \ , \tag{2.24}$$

i.e.,

$$|\tilde{\psi}\rangle = (\langle\psi|u\rangle)^*|v\rangle = \langle u|\psi\rangle|v\rangle \ .$$

$\langle u|\psi\rangle$ is a number, and we can equivalently put it in the following form:

$$|\tilde{\psi}\rangle = |v\rangle\left(\langle u|\psi\rangle\right) = \left(|v\rangle\langle u|\right)|\psi\rangle \ ,$$

which is exactly Eqn. (2.23).

Let us now prove Eqn. (2.24). Remember the definition of the dual: consider

$$\langle\tilde{\psi}|\tilde{\psi}\rangle = \|\tilde{\psi}\|^2 \ ,$$

if

$$\langle\tilde{\psi}| = \langle v|\,a \ ,$$

then we get

$$\langle\tilde{\psi}|\tilde{\psi}\rangle = \|\,|v\rangle\,\|^2\,\|a\|^2 \ .$$

This holds if $|\tilde{\psi}\rangle = a^*|v\rangle$, because in this way we have

$$\langle\tilde{\psi}|\tilde{\psi}\rangle = aa^*\langle v|v\rangle = \|a\|^2\,\|\,|v\rangle\,\|^2 \ .$$

$\star \quad \star \quad \star$

Problem 2.2. Prove that the operator

$$\hat{P}_\psi = |\psi\rangle\langle\psi|$$

with $\langle\psi|\psi\rangle = 1$ is a projection operator.

SOLUTION. It is necessary to prove the following facts.

1. $\hat{P}_\psi$ is linear (we have already done this before)
2. $\hat{P}_\psi$ is Hermitian, *i.e.*, $\hat{P}_\psi^\dagger = \hat{P}_\psi$. Prove this as homework. You have to construct the adjoint $\hat{P}_\psi^\dagger$.
3. $\hat{P}_\psi$ is idempotent, *i.e.*, $\hat{P}_\psi^2 = \hat{P}_\psi$. We have

$$\hat{P}_\psi^2 = (|\psi\rangle\langle\psi|)(|\psi\rangle\langle\psi|) = |\psi\rangle \underbrace{\langle\psi|\psi\rangle}_{1} \langle\psi| = |\psi\rangle\langle\psi| = \hat{P}_\psi \ .$$

$\star \quad \star \quad \star$

Problem 2.3. Prove that, in the Hilbert space $\mathcal{L}^2$ of the square-integrable functions on x, the operator $i\partial/\partial x$ is Hermitian.

SOLUTION. $\mathcal{L}^2$ is endowed with the scalar product defined for any two square-integrable functions $f(x)$ and $g(x)$ of one variable x as

$$\langle f|g\rangle = \int_{-\infty}^{+\infty} f^*(x)g(x)\,\mathrm{d}x \ .$$

An operator $\hat{O}$ is Hermitian if $\hat{O}^\dagger = \hat{O}$ where $\hat{O}^\dagger$ is defined via the relation

$$\langle \hat{O}^\dagger f|g\rangle = \langle f|\hat{O}g\rangle \ ,$$

therefore for an Hermitian operator $\hat{O}$ we have

$$\langle \hat{O}f|g\rangle = \langle f|\hat{O}g\rangle \ .$$

Let us consider $\hat{O} = i\partial/\partial x$. The above relation becomes

$$\int_{-\infty}^{+\infty} \left(i\frac{\partial f}{\partial x}\right)^* g(x)\,\mathrm{d}x \stackrel{?}{=} \int_{-\infty}^{+\infty} f^*(x) i\frac{\partial g}{\partial x}\,\mathrm{d}x \ .$$

We prove that this equality actually holds by integrating by parts the integral on the left-hand side. We have

$$-i\int_{-\infty}^{+\infty} \frac{\partial f^*}{\partial x} g(x)\,\mathrm{d}x = -i\underbrace{f^*(x)g(x)|_{-\infty}^{+\infty}}_{0} + \int_{-\infty}^{+\infty} f^*(x) i\frac{\partial g}{\partial x}\,\mathrm{d}x$$

The surface term vanishes since both f^* and g must go to zero at $\pm\infty$ for f, g to belong to $\mathcal{L}^2$ (otherwise, they would not be square-integrable). This proves that the equality above holds, thus $i\partial/\partial x$ is Hermitian.

$\star \quad \star \quad \star$

Chapter 3

POSTULATES OF QUANTUM MECHANICS

First let us recall the postulates of classical mechanics.

Postulate 1. The *initial state* of a system is given by a point (q_0, p_0) in phase space. The states of the system at any later time are also points in phase space.

Postulate 2. The *time evolution* is given by Hamilton's equations of motion:

$$\dot{q} = \{q, H\}_{P.B.} , \qquad \dot{p} = \{p, H\}_{P.B.} .$$

where $\{\cdot, \cdot\}_{P.B.}$ denotes the Poisson brackets.

Postulate 3. The *observables* (*e.g.*, energy, position, angular momentum, etc.) are real functions $F(q, p)$ defined on the phase space. The measured quantities are the values of F on the points of the trajectory.

In quantum mechanics the above three postulates are replaced by the following ones:

Postulate 1. The *initial state* of a quantum system is an element $|\psi\rangle$ of an Hilbert space. The states obtained by time evolution are also elements of the Hilbert space.

Postulate 2. The *time evolution* of a state of a quantum system (in absence of any measurement process) is given by the time-dependent Schrödinger equation:

$$i\hbar \frac{\partial |\psi\rangle}{\partial t} = \hat{H} |\psi\rangle \ ,$$

where $\hat{H}$ is an operator, called Hamiltonian or Schrödinger operator, which we shall soon define.

Postulate 3. The *observables* of a quantum system are hyper-maximal Hermitian operators defined on the Hilbert space of the states of the system. The outcomes of a measurement of an observable are the eigenvalues of the corresponding Hermitian operator (which, since the operator is Hermitian, are real).

In quantum mechanics however beside these three postulates it is necessary to introduce also some additional ones.

Postulate 4. A system is prepared by measuring a *complete set of commuting operators*.

As we have already seen, a complete set of commuting Hermitian operators determines uniquely a complete set of common eigenvectors, each one completely characterized by its eigenvalues. For instance, let us suppose that the complete set is $(\hat{A}, \hat{B}, \hat{C}, \ldots, \hat{M})$. The common eigenstates are obtained by diagonalizing simultaneously $\hat{A}, \hat{B}, \hat{C}, \ldots, \hat{M}$, that is, we have something like

$$\begin{cases} \hat{A}\,|\,\rangle = \ a_1 \ \ |\,\rangle \\ \hat{B}\,|\,\rangle = \ b_3 \ \ |\,\rangle \\ \hat{C}\,|\,\rangle = \ c_5 \ \ |\,\rangle \\ \qquad \vdots \\ \hat{M}\,|\,\rangle = m_{15}\ |\,\rangle \end{cases}$$

where the numbers $1, 3, 5, \ldots$ are just to say that we have more than one eigenvalue and we select one particular eigenvalue among all of them. Since the set is complete the correspondence between the eigenstate $|\,\rangle$ and the set of numbers $(a_1, b_3, c_5, \ldots, m_{15})$ is one-to-one, we can use the set of eigenvalues to label the state $|\,\rangle$, in particular we shall use the notation:

$$|a_1, b_3, c_5, \ldots, m_{15}\rangle$$

Measuring the observables $\hat{A}, \ldots, \hat{M}$ we obtain, according to the postulate 3, a string of numbers like $(a_i, b_j, c_k, \ldots, m_l)$, this string is in one-to-one correspondence with an

eigenstate

$$|\,\rangle \leftrightarrow (a_i, b_j, c_k, \ldots, m_l)$$

thus obtaining this string of eigenvalues from a measure is equivalent to obtaining a state. This is the way the initial state is actually prepared, that is, by measuring a *complete* set of commuting operators.

Postulate 5 (probability). The next question is the following one: since in general an observable $\hat{O}$ has many different eigenvalues $o_1, o_2, \ldots, o_n, \ldots$ what is the outcome of a measurement of the observable $\hat{O}$? Actually, in quantum mechanics one must speak about the *probability* of getting o_i in a measurement of $\hat{O}$. Let us suppose the system has been prepared in the state $|\psi_0\rangle$ at the time $t = 0$ and let this state evolve until the time t is reached. At the time t the state of the system is $|\psi\rangle$. Since $\hat{O}$ is an observable (that is, an hypermaximal Hermitian operator) the set of all its eigenstates $\{|o_i\rangle\}$ is a complete set of orthonormal states on which one can expand $|\psi\rangle$:

$$|\psi\rangle = \sum_i c_i \,|o_i\rangle \ .$$

Let us suppose that both $|\psi\rangle$ and the states $|o_i\rangle$ are *normalized* and that the eigenvalues o_i are not degenerate. Then, the probability of finding o_i in a measurement of $\hat{O}$ is

$$P_{o_i} = |c_i|^2 \ . \tag{3.1}$$

This procedure is possible if $\hat{O}$ is a complete set of commuting operators by itself. Otherwise, there is degeneracy . In this case we must look for Hermitian operators different than $\hat{O}$ which commute each other and with $\hat{O}$ and that form a complete set of commuting operators. The expansion formula for $|\psi\rangle$ becomes in this case

$$|\psi\rangle = \sum_{i,r} c_{i,r} \,|o_i, r\rangle \ ,$$

where the index r labels the degeneracy. In this case the probability of finding o_i in a measurement of $\hat{O}$ becomes

$$P_{o_i} = \sum_r |c_{i,r}|^2 \ . \tag{3.2}$$

Postulate 6 (Measurement postulate). Let us suppose we prepare a system in the state $|\psi\rangle$ and perform a measurement of an observable $\hat{O}$. Let us also suppose that we get the eigenvalue o_i. If immediately after we perform another measurement of $\hat{O}$ we get o_i again with probability 1. This means that immediately after the first measurement the original state $|\psi\rangle$ has become $|o_i\rangle$:

$$|\psi\rangle \xrightarrow{\text{measurement of } \hat{O}} |o_i\rangle \xrightarrow{\text{measurement of } \hat{O}} |o_i\rangle \ .$$

In other words, the state $|\tilde{\psi}\rangle$ in which the wave function has collapsed, that is,

$$|\psi\rangle \xrightarrow{\text{measurement of } \hat{O}} |\tilde{\psi}\rangle$$

has no components in the directions other than that of $|o_i\rangle$.

If, instead of performing the second measurement immediately after the first one, in which we got o_i, we wait until the time t is reached then the probability of finding o_i again is in fact different from 1, but it is the same we would obtain by evolving the state $|o_i\rangle$ from the first measurement to the time t using the Schrödinger equation:

$$|\psi\rangle \xrightarrow{\text{measurement of } \hat{O}} |o_i\rangle \xrightarrow{\text{measurement of } \hat{O}} |o_{i,t}\rangle \ .$$

It should be noted this postulate implies there exist in quantum mechanics two different ways in which states transform: the (reversible) time evolution given by the time-dependent Schrödinger equation and the (irreversible) wave packet reduction associated to any measurements process.

Let us now put forward a technical remark regarding the measurement postulate in the case in which the measured observable $\hat{O}$ does not form a complete set of commuting Hermitian observables by itself and degeneracy occurs. Let o_i be the outcome of a measurement of $\hat{O}$. A whole set of eigenstates $|o_i, r\rangle$ are now associated to the eigenvalue o_i, being r the index which labels the degeneracy. The question is: in which of the various $|o_i, r\rangle$ should the state collapse immediately after the measurement? We know the various $|o_i, r\rangle$ span an eigenvariety. The state in which $|\psi\rangle$ collapses is the "projection" of $|\psi\rangle$ on this eigenvariety, and we shall denote this state with $|\tilde{\psi}\rangle_{o_i}$. By denoting with $\hat{P}_i$ the projection operator on the eigenvariety spanned by $|o_i, r\rangle$, the measurement process can be represented schematically in the following way:

$$|\psi\rangle \xrightarrow{\substack{\text{measurement of } \hat{O} \\ \text{with outcome } o_i}} |\tilde{\psi}\rangle_{o_i} = \frac{\hat{P}_i |\psi\rangle}{\| \hat{P}_i |\psi\rangle \|} .$$

3.1 Dirac's correspondence rules

These rules, to some extent, allow us to build most of the observables in quantum mechanics starting from their classical analogous.

Postulate 7. 1. In classical mechanics, observables are functions of q and p, for example $O_1(q,p)$ and $O_2(q,p)$. The corresponding quantum observables are obtained* from the classical ones by replacing p, q with the associated quantum position and momentum operators $\hat{q}, \hat{p}$ (defined in the following)

$$O_1(q,p) \to \hat{O}_1(\hat{q},\hat{p}) ,$$
$$O_2(q,p) \to \hat{O}_2(\hat{q},\hat{p}) .$$

2. Further, the Poisson brackets go into the commutator divided by $i\hbar$:

$$\{O_1(q,p), O_2(q,p)\}_{\text{P.B.}} \to \frac{1}{i\hbar}[\hat{O}_1(\hat{q},\hat{p}), \hat{O}_2(\hat{q},\hat{p})] .$$

* Excluding certain "ambiguities" in the ordering, see remark 2 on the next page.

For example,

$$\{q,p\}_{\text{P.B.}} = 1 \quad \rightarrow \quad \frac{1}{i\hbar}[\hat{q},\hat{p}] = 1 \ .$$

So, we have

$$[\hat{q},\hat{p}] = i\hbar \ . \tag{3.3}$$

These are called Heisenberg commutation rules.

Now, let us discuss the limits of this correspondence principle.

1. There exist quantum observables having no classical analogous, such as spin, isospin, and so. In this case the correspondence principle does not work.

2. Since $\hat{q}$ and $\hat{p}$ do not commute at the quantum level, ambiguities may arise due to the fact that more than one quantum Hermitian operator can be associated to the same classical observable by changing the ordering in which the various $\hat{q}$ and $\hat{p}$ terms are considered. To clarify this point, let us consider for example the observable $O_1 = pq$. We can write this expression equivalently at the classical level as $O_1 = qp$ or $O_1 = (qp + pq)/2$. According to the correspondence rule, the associated quantum observables are

$$\begin{aligned} qp &\rightarrow \hat{q}\hat{p} = \hat{O}_1 \\ \frac{qp+pq}{2} &\rightarrow \frac{\hat{q}\hat{p}+\hat{p}\hat{q}}{2} = \tilde{\hat{O}}_1 \\ pq &\rightarrow \hat{p}\hat{q} = \tilde{\tilde{\hat{O}}}_1 \end{aligned}$$

It is easy to prove that $\hat{O}_1 \neq \tilde{\hat{O}}_1 \neq \tilde{\tilde{\hat{O}}}_1$. For example,

$$\hat{O}_1 = \hat{q}\hat{p} = \hat{p}\hat{q} + [\hat{q},\hat{p}] = \tilde{\tilde{\hat{O}}}_1 + i\hbar \ .$$

This means that, in constructing the quantum observables using the correspondence principle starting from the classical observables, we must be aware of the

fact that there exist quantum observables with no classical analogous and, for those having classical analogous, we must prescribe the *ordering* of $\hat{q}$ and $\hat{p}$. The latter ambiguity is referred to as *ordering ambiguity*.

3.2 More on the postulates of quantum mechanics

There is a further postulate called "*spin-statistics*" which we will present in chapter 8. In this section we shall comment on the postulates of quantum mechanics introduced in the previous sections and we shall try to understand them more deeply.

3.2.1 Expectation values

We have seen from postulate 3 that the outcomes of a measurement are probabilistic and not deterministic: one can only calculate the *probability* that by measuring an observable $\hat{O}$ the outcome will be an eigenvalue o_i. Accordingly to the postulate 3 the probability is given by the following rule: if the system is in the state $|\psi\rangle$, first of all you have to expand $|\psi\rangle$ on the basis of the eigenstates of $\hat{O}$ (resolving degeneracy if any)

$$|\psi\rangle = \sum_i c_i \, |o_i\rangle \ , \quad c_i = \langle o_i|\psi\rangle \ ,$$

then the probability of finding o_i performing a measurement of $\hat{O}$ on the state $|\psi\rangle$ is

$$P_i = |c_i|^2 \ .$$

(Do not confuse this P_i with the projector operator $\hat{P}_i$ introduced in the last section.) Thus, one can only speak about "mean value" $\overline{\hat{O}}$ of an observable $\hat{O}$ on a quantum state $|\psi\rangle$:

$$\overline{\hat{O}} = \sum_i P_i o_i = \sum_i |c_i|^2 \, o_i \ .$$

It is not difficult to prove that the mean value can also be obtained using the following expression :*

$$\overline{\hat{O}} = \langle\psi|\hat{O}|\psi\rangle \ . \tag{3.4}$$

In fact, as the $|o_i\rangle$ are a complete set, we can write

$$\sum_i |o_i\rangle \langle o_i| = 1 \ ,$$

and using this we get:

$$\overline{\hat{O}} = \langle\psi|\hat{O}|\psi\rangle = \langle\psi| \hat{O} \left(\sum_i |o_i\rangle \langle o_i| \right) |\psi\rangle = \sum_i \langle\psi|\hat{O}|o_i\rangle \langle o_i|\psi\rangle = \sum_i \langle\psi|o_i|o_i\rangle \langle o_i|\psi\rangle$$
$$= \sum_i o_i \langle\psi|o_i\rangle \langle o_i|\psi\rangle = \sum_i o_i \, |\langle o_i|\psi\rangle|^2 = \sum_i o_i \, |c_i|^2 \ .$$

3.2.2 Compatibility of observables and measurements.

Let us consider two commuting hyper-maximal Hermitian operators $\hat{A}$ and $\hat{B}$:

$$[\hat{A}, \hat{B}] = 0 \ .$$

In literature it is usually found the statement that these two observables can be measured simultaneously, and are said to be mutually compatible. This means the following thing: let us suppose we measure $\hat{A}$ and find the eigenvalue a_1. Immediately after, say at the time $\varepsilon/2$ where ε is a strictly positive real number as small as we want, we measure $\hat{B}$ and we find say b_2. At a time $\varepsilon/2$ after the measurement of $\hat{B}$ we perform a measurement of $\hat{A}$ again and we find a_1 with probability 1 because of postulate 3. The reason why we consider "small" times ε is that we do not want the state to change "too much" during the time evolution.

To understand how this concept of "compatibility" of two observables is linked to the postulates of quantum mechanics, we consider now the case in which the

* The formula is also correct in the case of degeneracy and when observables have continuous spectrum.

eigenvalue a_1 above shows degeneracy and thus there exists a whole eigenvariety of eigenstates associated to a_1. After the measurement of $\hat{A}$, the state $|\psi\rangle$ of the system immediately before the measurement of $\hat{B}$ collapses into $|\tilde{\psi}\rangle_{a_1}$, i.e. the projection of $|\psi\rangle$ on the eigenvariety associated to a_1. For the postulates of quantum mechanics, if after a "small" time $\varepsilon/2$ we measured $\hat{A}$ again we would find again a_1. After a further identical time interval let us perform a measurement of $\hat{B}$. Since $\hat{A}$ and $\hat{B}$ commutes, from the theorems on theory of Hilbert spaces we know $\hat{A}$ and $\hat{B}$ share at least one complete set of eigenstates. This means there must be at least one eigenstate of $\hat{B}$ in the eigenvariety associated to a_1. Let us suppose that there are two such eigenstates. If we find b_2 in the measurement of $\hat{B}$, $|\tilde{\psi}\rangle_{a_1}$ will collapse in the state $|a_1, b_2\rangle$. It is now clear that if, after a time $\varepsilon/2$, we measure $\hat{A}$ again, since the system is in the state $|a_1, b_2\rangle$ we will obtain a_1 again.

Therefore, we can say that a measurement of $\hat{B}$ has not changed the result of a measurement of $\hat{A}$: we get a_1 at the time $t = 0$ and at the time ε we get a_1 again. It is said also that a measurement of $\hat{B}$ does not disturb $\hat{A}$. That is, we can measure simultaneously $\hat{A}$ and $\hat{B}$. This is the meaning of *compatibility* of two observables.

3.2.3 Incompatible observables

We know that two non-commuting observables cannot share a complete set of eigenstates. They can have a set of eigenstates in common, but this set cannot be complete.

For example, $\hat{q}$ and $\hat{p}$ do not commute: $[\hat{q}, \hat{p}] = i\hbar$, and it is easy to show that these two observables have no eigenstate in common. In fact, if there were one such common eigenstate, say $|v\rangle$, then we would get

$$\begin{aligned}[\hat{q}, \hat{p}]\,|v\rangle &= \hat{q}\hat{p}\,|v\rangle - \hat{p}\hat{q}\,|v\rangle = \hat{q}p_v\,|v\rangle - \hat{p}q_v\,|v\rangle \\ &= p_v\hat{q}\,|v\rangle - q_v\hat{p}\,|v\rangle = p_vq_v\,|v\rangle - q_vp_v\,|v\rangle = 0\ ,\end{aligned}$$

while from $[\hat{q}, \hat{p}] = i\hbar$ we would get $[\hat{q}, \hat{p}]\,|v\rangle = i\hbar\,|v\rangle \neq 0$.

Two non-commuting operators $\hat{A}$ and $\hat{B}$, $[\hat{A}, \hat{B}] = \hat{C} \neq 0$, can have in general one eigenstate in common but it is easy to prove that this eigenstate must be an eigenstate of $\hat{C}$ with eigenvalue zero. In the case of $\hat{q}$ and $\hat{p}$, $i\hbar$ does not admit zero as eigenvalue.

Let us return to our discussion about measurements of incompatible observables. Let us expand the state $|q_0\rangle$ (in which our system has collapsed after a measurement of $\hat{q}$) on the basis of the eigenstates of $\hat{p}$:

$$|q_0\rangle = \sum_p C(q_0, p) \, |p\rangle \ . \tag{3.5}$$

Here the summation notation has been used, however the reader should keep in mind that $\hat{p}$ has a continuous spectrum and the summation must be replaced actually by an integration over the continuous variable p. Since $\hat{q}$ and $\hat{p}$ have no common eigenstates, $C(q_0, p)$ are always different from 1 for any p. After a time $\varepsilon/2$ we perform a measurement of $\hat{p}$ and let us suppose to find p_0. After a second amount of time $\varepsilon/2$ we measure again $\hat{q}$. The outcome is not q_0 with probability 1 as in case of the previous section when the two measured observables were compatible. What we have to do is to expand $|p_0\rangle$ on the eigenstates of the position operator

$$|p_0\rangle = \sum_q \tilde{C}(q, p_0) \, |q\rangle$$

where $\tilde{C}(q, p_0)$ are different from 1 for any q. So in general there is a non-zero probability $P_q = |\tilde{C}(q, p_0)|^2$ of finding a value q, and thus the outcome of the second measurement of $\hat{q}$ is not necessarily q_0.

The conclusion is that in the case of two non-commuting observables, the measurement of one observable "disturbs" the second observable. Consequently, the measurement postulate must be restated in the following way: after a "small" time ε the system is still in q_0 if and only if no measurement process has been done in between of a quantum observable which does not commute with the first.

3.3 Heisenberg uncertainty principle

Besides the probabilistic aspects of measurements, there are also correlations among them if they satisfy proper commutation relations. We will analyze here the most famous case. Let $\hat{A}$ and $\hat{B}$ be two non-commuting hyper-maximal Hermitian operators such that their commutator satisfy the relation

$$[\hat{A}, \hat{B}] = i\hat{C} \ , \tag{3.6}$$

where $\hat{C}$ is an Hermitian operator. Moreover we require that the states obtained by applying these operators to the original states belong to the same original Hilbert space.

Given the average value of the observable $\hat{A}$ with respect to the state $|\psi\rangle$

$$\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle$$

and the mean square

$$\Delta \hat{A} = \sqrt{\langle [\hat{A} - \langle \hat{A} \rangle]^2 \rangle}$$

it is possible to prove that

$$\Delta \hat{A} \Delta \hat{B} \geq \frac{1}{2} | \langle \hat{C} \rangle | \ . \tag{3.7}$$

In particular, in the case in which $\hat{A} = \hat{q}$ and $\hat{B} = \hat{p}$ we have $\hat{C} = \hbar$ and the previous relation becomes

$$\Delta \hat{q} \Delta \hat{p} \geq \frac{1}{2} \hbar \ .$$

This is Heisenberg uncertainty principle.

We now give a proof of Eqn. (3.7). We start by evaluating $(\Delta \hat{A})^2$. Let us diagonalize $\hat{A}$:

$$\hat{A} \, |a_k\rangle = a_k \, |a_k\rangle \ .$$

Being $\hat{A}$ hyper-maximal, we can expand any $|\psi\rangle$ on the basis of the eigenstates of $\hat{A}$ according to the usual expansion formula:

$$|\psi\rangle = \sum_k c_k \, |a_k\rangle \ , \qquad c_k = \langle a_k | \psi \rangle \ .$$

By inserting the latter in the expression for $(\Delta\hat{A})^2$ we get

$$\begin{aligned}(\Delta\hat{A})^2 &= \langle\psi|[\hat{A}-\langle\hat{A}\rangle]^2|\psi\rangle \\ &= \sum_{j,k} c_k^* c_j \langle a_k|[\hat{A}-\langle\hat{A}\rangle]^2|a_j\rangle \\ &= \sum_{j,k} c_k^* c_j \left(a_k-\langle\hat{A}\rangle\right)\left(a_j-\langle\hat{A}\rangle\right)\underbrace{\langle a_k|a_j\rangle}_{\delta_k^j} \\ &= \sum_k |c_k|^2 \left(a_k-\langle\hat{A}\rangle\right)^2 .\end{aligned}$$

Since $|c_k|^2$ is nothing but the probability of finding a_k in a measurement of $\hat{A}$, the last formula is nothing but the usual formula of the standard deviation.

By multiplying $\left(\Delta\hat{A}\right)^2$ by $\left(\Delta\hat{B}\right)^2$ we get

$$\begin{aligned}(\Delta\hat{A})^2(\Delta\hat{B})^2 &= \langle\psi|[\hat{A}-\langle\hat{A}\rangle]^2|\psi\rangle\,\langle\psi|[\hat{B}-\langle\hat{B}\rangle]^2|\psi\rangle \\ &= \|(\hat{A}-\langle\hat{A}\rangle)\,|\psi\rangle\,\|^2\,\|(\hat{B}-\langle\hat{B}\rangle)\,|\psi\rangle\,\|^2 \qquad (3.8)\end{aligned}$$

since both $\hat{A}$ and $\hat{B}$ are Hermitian and since $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$ are real. Let us introduce the new operators

$$\tilde{A}\equiv\hat{A}-\langle\hat{A}\rangle\ , \qquad \tilde{B}\equiv\hat{B}-\langle\hat{B}\rangle\ .$$

Then Eqn. (3.8) can be rewritten as

$$(\Delta\hat{A})^2(\Delta\hat{B})^2 = \|\tilde{A}\,|\psi\rangle\,\|^2\,\|\tilde{B}\,|\psi\rangle\,\|^2\ .$$

Using the Schwarz inequality on the vectors $\tilde{A}\,|\psi\rangle$ and $\tilde{B}\,|\psi\rangle$ we get

$$\left\|\left(\hat{A}-\langle\hat{A}\rangle\right)|\psi\rangle\right\|\left\|\left(\hat{B}-\langle\hat{B}\rangle\right)|\psi\rangle\right\| \geq \left|\left\langle\psi\,\middle|\,\left(\hat{A}-\langle\hat{A}\rangle\right)\left(\hat{B}-\langle\hat{B}\rangle\right)\,\middle|\,\psi\right\rangle\right| ,$$

from which it follows

$$\Delta\hat{A}\Delta\hat{B} \geq \left|\left\langle\psi\,\middle|\,\left(\hat{A}-\langle\hat{A}\rangle\right)\left(\hat{B}-\langle\hat{B}\rangle\right)\,\middle|\,\psi\right\rangle\right| , \qquad (3.9)$$

The modulus of a complex number is always bigger than the modulus of its imaginary part, thus on the right-hand side of Eqn. (3.9) we can take the imaginary part and

perform the following steps:

$$\begin{aligned}
\Delta\hat{A}\Delta\hat{B} &\geq \left|\frac{1}{2i}\left[\langle\psi|(\hat{A}-\langle\hat{A}\rangle)(\hat{B}-\langle\hat{B}\rangle)|\psi\rangle - \langle\psi|(\hat{A}-\langle\hat{A}\rangle)(\hat{B}-\langle\hat{B}\rangle)|\psi\rangle^*\right]\right| \\
&= \left|\frac{1}{2i}\left[\langle\psi|(\hat{A}-\langle\hat{A}\rangle)(\hat{B}-\langle\hat{B}\rangle)|\psi\rangle - \langle\psi|(\hat{B}-\langle\hat{B}\rangle)(\hat{A}-\langle\hat{A}\rangle)|\psi\rangle\right]\right| \\
&= \left|\frac{1}{2i}\langle\psi|[(\hat{A}-\langle\hat{A}\rangle),(\hat{B}-\langle\hat{B}\rangle)]|\psi\rangle\right| \\
&= \frac{1}{2}\,|\,\langle\hat{C}\rangle\,|\ .
\end{aligned}$$

$\Delta\hat{A}\Delta\hat{B} = 0$ only if $\langle\psi|[\hat{A},\hat{B}]|\psi\rangle = 0$, that is when $|\,\langle\hat{C}\rangle\,| = 0$.

3.4 Position and momentum representations

Let us start from the commutation relation

$$[\hat{q},\hat{p}] = i\hbar\ , \tag{3.10}$$

and ask ourselves if there exists a complete set of eigenstates common to both operators:

$$\begin{cases} \hat{q}\,|q,p\rangle &= q\,|q,p\rangle \\ \hat{p}\,|q,p\rangle &= p\,|q,p\rangle \end{cases}$$

If there were such $|q,p\rangle$, it would be possible to expand any state $|\psi\rangle$ as

$$|\psi\rangle = \int \mathrm{d}q\,\mathrm{d}p\ C(q,p)\,|q,p\rangle\ ,$$

Now, applying the commutatator $[\hat{q},\hat{p}]$ to $|\psi\rangle$ using the expansion formula above would yield

$$[\hat{q},\hat{p}]\,|\psi\rangle = \int qp\,C(q,p)\,|q,p\rangle\,\mathrm{d}q\,\mathrm{d}p - \int pq\,C(q,p)\,|q,p\rangle\,\mathrm{d}q\,\mathrm{d}p = 0\ ,$$

while we know by Eqn. (3.10) that

$$[\hat{q},\hat{p}]\,|\psi\rangle = i\hbar\,|\psi\rangle \neq 0\ .$$

Thus, as a consequence of the commutation relation (3.10) there cannot be such common eigenstate $|q, p\rangle$.

Another question: is it possible to represent $\hat{q}$ and $\hat{p}$ operators by means of finite-dimensional matrices? If it were possible then, by taking the trace of Eqn. (3.10), we would get :

$$\mathrm{Tr}[\hat{q}, \hat{p}] = \mathrm{Tr}(\hat{q}\hat{p}) - \mathrm{Tr}(\hat{p}\hat{q}) = \mathrm{Tr}(\hat{q}\hat{p}) - \mathrm{Tr}(\hat{q}\hat{p}) = 0 \ ,$$

where the two main properties of the trace, namely linearity and cycliccity, were used. But using the right-hand side of Eqn. (3.10) we get directly

$$\mathrm{Tr}[\hat{q}, \hat{p}] = \mathrm{Tr}\, i\hbar = i\hbar N \ ,$$

being N the dimension of the matrix. For the commutation relation (3.10) to hold, it is thus not possible to represent $\hat{q}$ and $\hat{p}$ by means of finite-dimensional matrices.

Now, since $\hat{q}$ is Hermitian under the usual scalar product (that is, the $\mathcal{L}^2$ scalar product $\langle\psi|\varphi\rangle = \int \psi^*(q)\varphi(q)\,\mathrm{d}q$, being $\psi(q)$ and $\varphi(q)$ two square-integrable complex function of q) we can diagonalize it and get a complete orthonormal basis

$$\hat{q}\,|q\rangle = q\,|q\rangle \ .$$

We remark that here "complete" means

$$\int |q\rangle\,\langle q|\,\mathrm{d}q = \mathbb{1} \ ,$$

where $\mathbb{1}$ denotes here the identity operator. The above relation is nothing but the usual resolution of the identity formula for the Hermitian operator $\hat{q}$.

Let us consider a state $|\psi\rangle$ and let us expand it on the basis of the eigenstates of $|q\rangle$:

$$|\psi\rangle = \int |q\rangle\,\langle q|\psi\rangle\,\mathrm{d}q \ .$$

The expansion coefficients are $\langle q|\psi\rangle$. The latter can be interpreted as a function of q

$$\langle q|\psi\rangle = \psi(q) \ .$$

which is the wave function associated to the state $|\psi\rangle$ in the position representation, also known as q representation.

The operator $\hat{q}$ acts on a state $|\psi\rangle$ in the following way:

$$\hat{q}\,|\psi\rangle = \hat{q}\int |q\rangle \underbrace{\langle q|\psi\rangle}_{\psi(q)}\,\mathrm{d}q = \int \hat{q}\,|q\rangle\,\psi(q)\,\mathrm{d}q = \int q\psi(q)\,|q\rangle\,\mathrm{d}q\ .$$

therefore

$$\langle q'|\hat{q}|\psi\rangle = \int q\psi(q)\,\langle q'|q\rangle\,\mathrm{d}q\ .$$

The left-hand side above is simply

$$\langle q'|\hat{q}|\psi\rangle = q'\,\langle q'|\psi\rangle = q'\psi(q')\ ,$$

and this should be equal to $\int q\psi(q)\,\langle q'|q\rangle\,\mathrm{d}q$. This fact tells us that $\langle q'|q\rangle$ must behave like a Dirac's delta function $\delta(q-q')$, so that

$$\int q\psi(q)\,\underbrace{\langle q'|q\rangle}_{\delta(q-q')}\,\mathrm{d}q = \int q\psi(q)\delta(q-q')\,\mathrm{d}q = q'\psi(q')\ .$$

Hence, the eigenstates $|q\rangle$ belong to a generalized Hilbert space and are normalized in the distributional sense of the Dirac's delta function.

Let us turn to the momentum operator $\hat{p}$. It is often said that commutation relation (3.10) implies $\hat{p}$ can be realized as*

$$\hat{p} = -i\hbar\frac{\mathrm{d}}{\mathrm{d}q}\ .$$

In fact, if we apply the commutation relation (3.10) to $\psi(q)$ and we use this representation for $\hat{p}$ as a differential operator, Eqn. (3.10) is actually fulfilled. Let us check this fact:

$$[\hat{q},\hat{p}]\psi(q) = -i\hbar\left(q\frac{\mathrm{d}\psi(q)}{\mathrm{d}q} - \frac{\mathrm{d}q\psi(q)}{\mathrm{d}q}\right) = i\hbar\psi(q)\ .$$

* Total derivative was used having in mind the one-dimensional case; in multi-dimensional case partial derivatives must be used in writing the representation above.

Let us now work out the matrix elements of $\hat{p}$ in the position representation. First, the matrix elements of the commutator $[\hat{q}, \hat{p}]$ are evaluated between any two position eigenstates $|q\rangle$ and $\langle q'|$:

$$\langle q'|[\hat{q}, \hat{p}]|q\rangle = \langle q'|\hat{q}\hat{p}|q\rangle - \langle q'|\hat{p}\hat{q}|q\rangle = (q' - q)\,\langle q'|\hat{p}|q\rangle \ .$$

Next, Eqn. (3.10) is used to get

$$\langle q'|[\hat{q}, \hat{p}]|q\rangle = \langle q'|i\hbar|q\rangle = i\hbar\,\langle q'|q\rangle = i\hbar\delta(q - q') \ .$$

By matching the two previous expressions we are led to

$$(q' - q)\,\langle q'|\hat{p}|q\rangle = i\hbar\delta(q - q') \ . \tag{3.11}$$

In other words, $\langle q'|\hat{p}|q\rangle$ is something that, multiplied by $q' - q$, behaves under integration with respect to q' as a delta function $\delta(q - q')$, that is

$$\int (q' - q)\,\langle q'|\hat{p}|q\rangle\,\mathrm{d}q' = i\hbar \int \delta(q - q')\,\mathrm{d}q' \ .$$

The latter equation is fulfilled if

$$\langle q'|\hat{p}|q\rangle = -i\hbar\frac{\mathrm{d}}{\mathrm{d}q'}\delta(q - q') \ .$$

In fact, integrating by parts yields

$$\begin{aligned}\int (q' - q)\left(-i\hbar\frac{\mathrm{d}}{\mathrm{d}q'}\right)\delta(q - q')\,\mathrm{d}q' \;&=\; i\hbar\int\left(\frac{\mathrm{d}}{\mathrm{d}q'}(q' - q)\right)\delta(q - q')\,\mathrm{d}q' + \\ &\quad (q' - q)(-i\hbar)\delta(q - q')\Big|_{-\infty}^{+\infty} \\ &=\; \int i\hbar\delta(q - q')\,\mathrm{d}q' \ .\end{aligned}$$

Strictly speaking instead of $\hat{p} = -i\hbar\,\mathrm{d}/\,\mathrm{d}q$ one should more properly say that the representation of $\hat{p}$ in the basis of position eigenstates $|q\rangle$ is given by the matrix elements

$$\langle q'|\hat{p}|q\rangle = -i\hbar\frac{\mathrm{d}}{\mathrm{d}q'}\delta(q - q') \ .$$

From this equation it is easy to find how $\hat{p}$ acts on a state $|\psi\rangle$ in the position representation:

$$\begin{aligned}\langle q'|\hat{p}|\psi\rangle &= \langle q'|\,\hat{p}\int |q\rangle\,\langle q|\psi\rangle\,\mathrm{d}q \\ &= \int(-i\hbar)\left(\frac{\mathrm{d}}{\mathrm{d}q'}\delta(q-q')\right)\psi(q)\,\mathrm{d}q \\ &= -i\hbar\frac{\mathrm{d}}{\mathrm{d}q'}\int\delta(q-q')\psi(q)\,\mathrm{d}q \\ &= -i\hbar\frac{\mathrm{d}\psi(q')}{\mathrm{d}q'}\ .\end{aligned}$$

This clarifies the meaning of the statement that $\hat{p}$ acts on $\psi(q)$ as $-i\hbar\,\mathrm{d}/\,\mathrm{d}q$.

Momentum representation

Instead of diagonalizing $\hat{q}$, we now turn to the problem of diagonalizing $\hat{p}$, which we shall prove in an exercise to be Hermitian and thus to have a complete basis of eigenstates:

$$\hat{p}\,|p'\rangle = p'\,|p'\rangle\ , \quad \int |p'\rangle\,\langle p'|\,\mathrm{d}p' = \mathbb{1}\ ,$$

where $\mathbb{1}$ denotes as usual the identity operator.

We are interested in how the momentum eigenstates $|p\rangle$ are represented in the basis of position eigenstates $|q\rangle$, that is, what are the matrix elements $\langle q|p\rangle$. We have

$$\langle q'|\hat{p}|p'\rangle = p'\,\langle q'|p'\rangle\ ,$$

and by inserting the resolution of the identity formula $\int |q\rangle\,\langle q|\,\mathrm{d}q = 1$ (completeness of position eigenstates) in the left-hand side above we get

$$\langle q'|\hat{p}|p'\rangle = \int\langle q'|\hat{p}|q\rangle\,\langle q|p'\rangle\,\mathrm{d}q = p'\,\langle q'|p'\rangle\ .$$

By using the representation of $\hat{p}$ derived previously, the equation above can be rewritten as

$$-i\hbar\int\frac{\mathrm{d}}{\mathrm{d}q'}\delta(q-q')\,\langle q|p'\rangle\,\mathrm{d}q = p'\,\langle q'|p'\rangle\ .$$

The derivative with respect to q' does not affect the integration with respect to q and thus it can be brought outside the integration,

$$-i\hbar \int \frac{\mathrm{d}}{\mathrm{d}q'}\delta(q-q') \langle q|p'\rangle \,\mathrm{d}q = -i\hbar \frac{\mathrm{d}}{\mathrm{d}q'} \int \delta(q-q') \langle q|p'\rangle \,\mathrm{d}q = -i\hbar \frac{\mathrm{d}}{\mathrm{d}q'} \langle q'|p'\rangle \ ,$$

yielding the following differential equation for the matrix elements $\langle q'|p'\rangle$:

$$-i\hbar \frac{\mathrm{d}}{\mathrm{d}q'} \langle q'|p'\rangle = p' \langle q'|p'\rangle$$

whose solution is

$$\langle q'|p'\rangle = C(p')\, \mathrm{e}^{\frac{i}{\hbar}q'p'} \ ,$$

where $C(p')$ is a function to be determined. We do that by requiring

$$\langle q'|q\rangle = \delta(q-q') \ .$$

Because of the completeness of momentum eigenstates we have

$$\langle q'|q\rangle = \int \langle q'|p\rangle \langle p|q\rangle \,\mathrm{d}p = \int |C(p)|^2 \, \mathrm{e}^{\frac{i}{\hbar}(q'-q)p} \,\mathrm{d}p$$

and this should be equal to $\delta(q'-q)$; by recalling the integral representation of the Dirac's delta

$$\delta(q-q') = \frac{1}{2\pi} \int \mathrm{e}^{ip(q'-q)} \,\mathrm{d}p \ .$$

we get

$$|C(p)| = \frac{1}{\sqrt{2\pi\hbar}} \ ,$$

and thus

$$\langle q|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \, \mathrm{e}^{\frac{i}{\hbar}qp} \ .$$

Homework 3.1. Prove that $\langle p'|p''\rangle = \delta(p'-p'')$.

Let us now check that

$$\int |p\rangle \langle p| \,\mathrm{d}p = \mathbb{1} \ . \tag{3.12}$$

Let us sandwich the above relation between $\langle q'|$ and $|q\rangle$ and we get:

$$\int \mathrm{d}p \, \langle q'|p\rangle \, \langle p|q\rangle = \langle q'|\mathbb{1}|q\rangle = \delta(q-q') \; .$$

The left-hand side can be rewritten as

$$\begin{aligned} \int \mathrm{d}p \frac{1}{\sqrt{2\pi\hbar}} \, \mathrm{e}^{\frac{i}{\hbar}q'p} \, \frac{1}{\sqrt{2\pi\hbar}} \, \mathrm{e}^{-\frac{i}{\hbar}qp} &= \frac{1}{2\pi\hbar} \int \mathrm{d}p \, \mathrm{e}^{\frac{i}{\hbar}p(q'-q)} \\ &= \frac{1}{2\pi} \int \mathrm{d}\left(\frac{p}{\hbar}\right) \mathrm{e}^{i\left(\frac{p}{\hbar}\right)(q-q')} \\ &= \delta(q-q') \; , \end{aligned}$$

and this proves Eqn. (3.12).

Fourier transform

It is possible to prove that the position and the momentum representations of a given abstract state $|\psi\rangle$ are one the Fourier transformed of the other respectively.

By using the completeness of the momentum eigenstates we can expand $|\psi\rangle$ in the usual way

$$|\psi\rangle = \int |p\rangle \, \langle p|\psi\rangle \, \mathrm{d}p \; ,$$

where $\langle p|\psi\rangle$ can be viewed as a function of p, the so-called momentum-space wave function

$$\tilde{\psi}(p) \equiv \langle p|\psi\rangle \; .$$

The position-space wave function had already been defined as

$$\psi(q) = \langle q|\psi\rangle \; .$$

By using the resolution of the identity and by applying the formula for the matrix elements $\langle q|p\rangle$ found in the previous section, we get

$$\begin{aligned} \psi(q) = \langle q|\psi\rangle &= \int \langle q|p\rangle \, \langle p|\psi\rangle \, \mathrm{d}p \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int \mathrm{e}^{\frac{i}{\hbar}qp} \, \tilde{\psi}(p) \, \mathrm{d}p \; . \end{aligned}$$

Setting $\hbar = 1$ yields

$$\psi(q) = \frac{1}{\sqrt{2\pi}} \int \mathrm{e}^{iqp}\, \tilde{\psi}(p)\, \mathrm{d}p \ ,$$

which is exactly the Fourier transform. Of course, if $\hbar$ is not set to one, the correct formula is the one given before.

Let us now look for the momentum representation of the position operator $\hat{q}$. We note that in the momentum representation $\hat{p}$ is represented as a multiplicative operator, *i.e.*:

$$\hat{p}\,|p\rangle = p\,|p\rangle \ .$$

Let us consider the commutation relation (3.10) as usual. The matrix element of the commutator between two momentum eigenstates $\langle p|$ and $|p'\rangle$ is

$$\langle p|[\hat{q},\hat{p}]|p'\rangle = \langle p|\hat{q}\hat{p}|p'\rangle - \langle p|\hat{p}\hat{q}|p'\rangle = (p'-p)\,\langle p|\hat{q}|p'\rangle \ .$$

Because of the commutator relation, this is also equal to

$$\langle p|[\hat{q},\hat{p}]|p'\rangle = \langle p|i\hbar|p'\rangle = i\hbar\,\langle p|p'\rangle = i\hbar\delta(p'-p) \ ,$$

thus

$$\langle p|\hat{q}|p'\rangle = i\hbar\frac{\mathrm{d}}{\mathrm{d}p}\delta(p'-p) \ ,$$

which means that the operator $\hat{q}$ acts as $+i\hbar\frac{\mathrm{d}}{\mathrm{d}p}$ in the momentum representation.

3.5 Time-dependent Schrödinger equation: formal solution and time-evolution operator

The time-dependent Schrödinger equation is (formally)

$$i\hbar\frac{\mathrm{d}\,|\psi(t)\rangle}{\mathrm{d}t} = \hat{H}\,|\psi(t)\rangle \ . \tag{3.13}$$

The formal solution is given by*

$$|\psi(t)\rangle = \mathrm{e}^{-\frac{i}{\hbar}\hat{H}t}\,|\psi(0)\rangle \ . \tag{3.14}$$

The operator-valued exponential function is understood here in the sense we have used to define a function of an operator, namely, by its action on the eigenstates of the original operator:

$$\mathrm{e}^{-\frac{i}{\hbar}\hat{H}t}\,|E_n\rangle = \underbrace{\mathrm{e}^{-\frac{i}{\hbar}E_n t}}_{\text{This is the exponential of a number}}|E_n\rangle$$

Using this in (3.14) we immediately get (3.13). Another definition of the exponential of an operator is the following one:

$$\mathrm{e}^{-\frac{i}{\hbar}\hat{H}t} = \sum_{n=0}^{\infty}\left(-\frac{i\hat{H}t}{\hbar}\right)^n\frac{1}{n!}.$$

Note that this is not a Taylor expandion but a converging series for all value of t. Using this expression we can now prove that the solution (3.14) formally satisfies the time-dependent Schrödinger equation. In fact, inserting the (3.14) into the left-hand side of Eqn. (3.13) yields

$$\begin{aligned}
i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{n=0}^{\infty}\left(-\frac{i\hat{H}t}{\hbar}\right)^n\frac{1}{n!}\right]|\psi(0)\rangle &= i\hbar\sum_{n=1}^{\infty}\frac{1}{n!}\left(\frac{-i\hat{H}t}{\hbar}\right)^{n-1} n\left(\frac{-i\hat{H}}{\hbar}\right)|\psi(0)\rangle \\
&= i\hbar\sum_{n=1}^{\infty}\frac{1}{(n-1)!}\left(\frac{-i\hat{H}t}{\hbar}\right)^{n-1}\left(\frac{-i\hat{H}}{\hbar}\right)|\psi(0)\rangle \\
&= \sum_{n=1}^{\infty}\frac{i}{(n-1)!}\left(\frac{t}{\hbar}\right)^{n-1}(-i\hat{H})^n\,|\psi(0)\rangle \\
&= \sum_{n=0}^{\infty}\frac{i}{n!}\left(\frac{t}{\hbar}\right)^{n}(-i\hat{H})^{n+1}\,|\psi(0)\rangle \ ,
\end{aligned} \tag{3.15}$$

* This holds for time-independent Hamiltonian operators. Hamiltonian operators which depend *explicitly* on time may occur, for example, when dealing with time-varying electric or magnetic fields. We shall return on this when we shall discuss the interaction picture and the time-dependent perturbation theory.

while introducing the solution into the right-hand side of the equation yields

$$\hat{H}\sum_{n=0}^{\infty}\left(-\frac{i\hat{H}t}{\hbar}\right)^n\frac{1}{n!}\left|\psi(0)\right\rangle=\sum_{n=0}^{\infty}\left(\frac{t}{\hbar}\right)^n\frac{i}{n!}(-i\hat{H})^{n+1}\left|\psi(0)\right\rangle\ ,$$

which is equal to (3.15).

Now we turn to the problem of how the formal solution (3.14) can be practically used to find $|\psi(t)\rangle$ once $|\psi(0)\rangle$ is known. Since the Hamiltonian $\hat{H}$ is an Hermitian operator, its eigenstates $|E_n\rangle$,

$$\hat{H}\left|E_n\right\rangle=E_n\left|E_n\right\rangle\ ,$$

form a complete orthornormal set and can be used to expand any given $|\psi(0)\rangle$ according to the formula

$$\left|\psi(0)\right\rangle=\sum_n c_n\left|E_n\right\rangle\ ,\quad c_n=\langle E_n|\psi(0)\rangle\ .$$

Now, we apply Eqn. (3.14):

$$\begin{aligned}\left|\psi(t)\right\rangle&=\mathrm{e}^{-\frac{i}{\hbar}\hat{H}t}\sum_n c_n\left|E_n\right\rangle\\&=\sum_n c_n\,\mathrm{e}^{-\frac{i}{\hbar}\hat{H}t}\left|E_n\right\rangle\\&=\sum_n c_n\,\mathrm{e}^{-\frac{i}{\hbar}E_nt}\left|E_n\right\rangle\\&=\sum_n c_n(t)\left|E_n\right\rangle\end{aligned}$$

where

$$c_n(t)=c_n\,\mathrm{e}^{-\frac{i}{\hbar}E_nt}=\langle E_n|\psi(0)\rangle\,\mathrm{e}^{-\frac{i}{\hbar}E_nt}\ .$$

So, the probability of finding E_n by performing an energy measurement at the time t is given by

$$P_{E_n}(t)=\left|c_n(t)\right|^2\ .$$

However, it is easily seen that

$$\left|c_n(t)\right|^2=c_n^*(t)c_n(t)=c_n^*\underbrace{\mathrm{e}^{+\frac{i}{\hbar}E_nt}\,\mathrm{e}^{-\frac{i}{\hbar}E_nt}}_{=1}c_n=\left|c_n\right|^2\ .$$

This means the probability above does not change with time.

Another important concept is the so-called time-evolution operator $\hat{U}(t)$ defined by

$$\hat{U}(t) = \mathrm{e}^{-\frac{i}{\hbar}\hat{H}t} \ .$$

Since $\hat{H}$ is Hermitian, $\hat{U}(t)$ is unitary:

$$\hat{U}^{\dagger}(t) = \mathrm{e}^{\frac{i}{\hbar}\hat{H}t} = \hat{U}^{-1}(t) \ .$$

In fact,

$$\hat{U}^{\dagger}(t)\hat{U}(t) = \mathrm{e}^{\frac{i}{\hbar}\hat{H}t}\,\mathrm{e}^{-\frac{i}{\hbar}\hat{H}t} = \mathbb{1} \ ,$$

being $\mathbb{1}$ the identity operator.

3.6 Density Matrix

One of the postulate of Quantum Mechanics tells us that to prepare a system we need a *complete* set of commuting observables. The string of eigenvalues, each associated to a different observables, will uniquely characterize the state of the system. What about if we do not have a complete set of observables or if we have many many particles and we cannot perform measurements on all of them because they are too many. In both cases we will have some probabilities $p_1, p_2, \cdots p_m \cdots$ that the particle is in one the normalized states $\langle 1|\,, \langle 2| \cdots \langle m| \cdots$. If we now ask which is the average value of an observable $\widehat{O}$, as a result we will get:

$$\langle \widehat{O} \rangle = \sum_i p_i \langle i|\, \widehat{O}\, |i\rangle \tag{3.16}$$

This statistical mixture can be obtained by introducing what is called the *density operator* or matrix, defined as:

$$\widehat{\rho} \equiv \sum_i p_i\, |i\rangle\, \langle i| \tag{3.17}$$

where the $\langle i|$ are normalized states but not necessarily orthogonal to each other, and the p_i are probabilities that sum up to one, so:

$$p_i \geq 0 \ , \quad \sum_i p_i = 1$$

The average value defined in Eqn.(3.16) can then be written as

$$\langle \widehat{O} \rangle = Tr(\widehat{\rho}\, \widehat{O}) \tag{3.18}$$

This is easy to prove, in fact from

$$Tr(\widehat{\rho}\, \widehat{O}) = \sum_i p_i Tr(|i\rangle \langle i|\, \widehat{O}) \tag{3.19}$$

in order to get (3.18) we have just to show that:

$$Tr(|i\rangle \langle i|\, \widehat{O}) = \langle i|\, \widehat{O}\, |i\rangle$$

This can be proved easily considering that $|i\rangle \langle i| \equiv \widehat{P}_i$ is a projection operator, so $\widehat{P}_i^2 = \widehat{P}_i$. If we then use the cyclic property of the trace we have:

$$Tr\, |i\rangle \langle i|\, \widehat{O} = Tr \widehat{P}_i^2 \widehat{O} = Tr \widehat{P}_i \widehat{O} \widehat{P}_i$$

The last expression can be written as

$$Tr\, |i\rangle \langle i|\, \widehat{O}\, |i\rangle \langle i| = \langle i|\, \widehat{O}\, |i\rangle\, Tr\, |i\rangle \langle i| = \langle i|\, \widehat{O}\, |i\rangle\, Tr \widehat{P}_i.$$

We know that

$$Tr \widehat{P}_i = Tr\, |i\rangle \langle i| = \langle i|i\rangle = 1$$

The last step comes from the fact that the trace operation is the sum of the diagonal elements of an operator, so:

$$Tr\, |i\rangle \langle i| = \sum_j \langle j|i\rangle \langle i|j\rangle = \langle i|i\rangle \langle i|i\rangle = \langle i|i\rangle$$

where we have choosen a basis of orthogonal vectors $\langle j|$ of which $\langle i|$ is just one of them. We could choose this basis because the trace in invariant under a unitary change of basis:

$$Tr\widehat{O} = TrU\widehat{O}U^{\dagger} = TrU^{\dagger}U\widehat{O} = TrU^{-1}U\widehat{O} = Tr\widehat{O}$$

Another property of the density operator can be obtained if we choose in Eqn.(3.18) $\widehat{O} = \mathbb{1}$. We get:

$$Tr\widehat{\rho} = 1$$

The density matrix will help us not only in getting average values but also in getting probability distributions. For example if $\widehat{P}_D$ is the projector operator upon the subspace spanned by the eigenvectors of $\widehat{O}$ located in a certain domain D of the spectrum of $\widehat{O}$, then the probability $\mathcal{P}_D$ that the measurement of $\widehat{O}$ belongs to D is

$$\mathcal{P}_D = \sum_m p_m \langle m| \widehat{P}_D |m\rangle = Tr(\widehat{\rho}\widehat{P}_D) \tag{3.20}$$

In particular, using Eqn.(3.20), the probability of finding the system in the quantum state represented by the Hilbert state element $|\chi\rangle$ is given by

$$\mathcal{P}_\chi = Tr(\widehat{\rho}|\chi\rangle\langle\chi|) = \langle\chi|\widehat{\rho}|\chi\rangle$$

So, given $\widehat{\rho}$, we can calculate all *average* values and *statistical* distributions of measurements of observables.

In general one needs to introduce a density matrix $\widehat{\rho}$ when the system cannot be specified by a single state $|\Psi\rangle$. If this instead this were the case, then $\widehat{\rho}$ would be:

$$\widehat{\rho} = |\Psi\rangle\langle\Psi|$$

and, if Ψ was normalized, we would have

$$\widehat{\rho}^2 = \widehat{\rho} \tag{3.21}$$

These kinds of density matrices are called *pure*. In this case we could use the state itself $\langle\Psi|$ instead of the density matrix.

In general anyhow the $\widehat{\rho}$ is not pure and does not have the form above but the following one:

$$\widehat{\rho} = \sum p_i \left|i\right\rangle \left\langle i\right| \tag{3.22}$$

and as a consequence

$$\widehat{\rho}^2 \neq \widehat{\rho}.$$

In this case we say that the density matrix is a *mixed* one because it gives a statistical mixture which cannot be specified by a single state.

We can say that the density matrix is a unifying formalism to treat both systems that can be specified by a single state and those which are statistical mixture of many states.

Several properties of the density matrix are of relevance and we will list them here.

From the fact that the p_i are positive probabilities which sum up to one: $p_i \geq 0$ and $\sum_i p_i = 1$, we can easily derive that $p_i^2 \leq 1$ and from this we can prove that:

$$\widehat{\rho}^2 \leq 1$$

Other properties of $\widehat{\rho}$ are the following: first it is hermitian (because the p_i are real)

$$\widehat{\rho}^\dagger = \widehat{\rho},$$

second, because the probabilities sum up to one and the states are normalized, its trace is one:

$$Tr\widehat{\rho} = 1$$

and third it is a positive semidefinite operator:

$$\left\langle\Psi\right| \widehat{\rho} \left|\Psi\right\rangle \geq 0 \tag{3.23}$$

This last property can be easily proved using Eqn. (3.22). In fact using that expression we have $\langle\Psi|\,\widehat{\rho}\,|\Psi\rangle = \sum_i p_i| < \psi|i > |^2$ and from the fact that the $p_i \geq 0$ and $| < \psi|i > |^2 \geq 0$ we get Eqn.(3.23).

Let us now attack the problem of whether a physical system is specified by one and only one density matrix or not. For the states we know that there is not a one to one correspondence: any state up to a global phase specify the same physical system. This ambiguity on the states is not transferred to the density matrix, in fact :

$$\widehat{\rho} = \sum_i p_i \,|i\rangle\,\langle i| \longrightarrow \sum_i p_i e^{i\alpha_i}\,|i\rangle\,\langle i|\,e^{-i\alpha_i} = \sum_i p_i \,|i\rangle\,\langle i|\,.$$

Anyhow the $\widehat{\rho}$ has a lot of other freedoms. It is possible, for example, to prove that the set of states $|i\rangle$ and $|m\rangle$ generate the same density matrix if (and only if) they are transformed into each other by a unitary matrix $U_{i,m}$:

$$|i\rangle = \sum_m U_{i,m}\,|m\rangle$$

The number of states $|i\rangle$ and $|m\rangle$ must be the same and normalized. We will leave the proof of this theorem as an **homework**. Basically it will turn out that:

$$\widehat{\rho} = \sum_i p_i \,|i\rangle\,\langle i| = \sum_m q_m \,|m\rangle\,\langle m|$$

where the q_m are defined by:

$$\sqrt{p_i}\,|i\rangle = \sum_m U_{im}\sqrt{q_m}\,|m\rangle$$

3.6.1 Time Evolution of the Density Matrix

As the density matrix is built out of kets and bras, the time evolution is obtained by making the time-evolution of these components.

Suppose

$$\widehat{\rho}(0) = \sum p_i \left|i\right\rangle \left\langle i\right|$$

then

$$\widehat{\rho}(t) = \sum p_i \left|i,t\right\rangle \left\langle i,t\right| = \sum_i p_i \widehat{U}(t) \left|i\right\rangle \left\langle i\right| \widehat{U}^\dagger(t) = \widehat{U}(t)\widehat{\rho}(0)\widehat{U}^\dagger(t) \tag{3.24}$$

We have seen previously in this chapter the expression of the operator of time evolution $\widehat{U}(t)$ which is:

$$\widehat{U}(t) = e^{-i\frac{\widehat{H}}{\hbar}t}$$

and so Eqn.(3.24) becomes:

$$\widehat{\rho}(t) = e^{-i\frac{\widehat{H}}{\hbar}t}\widehat{\rho}(0)e^{i\frac{\widehat{H}}{\hbar}t}.$$

Taking the derivative with respect to t of the expression above we get the equation of motion of $\widehat{\rho}$:

$$\frac{\partial\widehat{\rho}}{\partial t} = \frac{i}{\hbar}[\widehat{\rho},\widehat{H}]$$

which is called von Neumann equation and is the quantum analogue of the Liouville equation.

3.7 Problems and Solutions

Problem 3.1. By using the derivation of the uncertainty relation presented in this chapter , look for the states of minimal uncertainty which satisfy

$$\Delta\hat{q}\Delta\hat{p} = \frac{1}{2}\hbar\,. \tag{3.25}$$

SOLUTION. The equality holds when the following two conditions, encountered in the various steps of the Heisenberg uncertainty theorem, are both satisfied:

1. the Schwarz inequality becomes an equality. This is true if and only if the two vectors in the inequality are proportional to each other:

$$[\hat{A} - \langle\hat{A}\rangle]\,|\psi\rangle = k[\hat{B} - \langle B\rangle]\,|\psi\rangle \ ,$$

   with some constant k.

2. the complex number in the right-hand side of Eqn. (3.9) is purely imaginary:

$$\langle\psi|(\hat{A} - \langle\hat{A}\rangle)(\hat{B} - \langle\hat{B}\rangle)|\psi\rangle = i\gamma \ ,$$

   with γ a real constant.

Thus,

$$k^* \underbrace{\langle\psi|(\hat{B} - \langle\hat{B}\rangle)(\hat{B} - \langle\hat{B}\rangle)|\psi\rangle}_{\|(\hat{B}-\langle\hat{B}\rangle)|\psi\rangle\|^2} = i\gamma \ .$$

This means k must be purely imaginary: $k = i\varepsilon$, with ε real. This is due to the fact that the expectation value on the left-hand side above is real: in fact, it is nothing but the square of the norm $\|(\hat{B} - \langle\hat{B}\rangle)\,|\psi\rangle\,\|$. So we have

$$\left(\hat{A} - \langle\hat{A}\rangle\right)|\psi\rangle = i\varepsilon\left(\hat{B} - \langle\hat{B}\rangle\right)|\psi\rangle \ .$$

Now, we choose $\hat{A} = \hat{q}$ and $\hat{B} = \hat{p} = -i\hbar\frac{\mathrm{d}}{\mathrm{d}q}$ and we image that $|\psi\rangle$ is already realized on q as a wave function $\psi(q)$ in the $\mathcal{L}^2$ space. The previous equation becomes

$$[\hat{q} - \langle\hat{q}\rangle]\,\psi(q) = i\varepsilon\left[-i\hbar\frac{\mathrm{d}}{\mathrm{d}q} - \langle\hat{p}\rangle\right]\psi(q) \ ,$$

that is,

$$\varepsilon\hbar\frac{\mathrm{d}\psi(q)}{\mathrm{d}q} + [\langle\hat{q}\rangle - i\varepsilon\,\langle\hat{p}\rangle]\,\psi(q) - q\psi(q) = 0 \ ,$$

from which it follows

$$\frac{\mathrm{d}\psi}{\psi} = \frac{q + i\varepsilon\,\langle\hat{p}\rangle - \langle\hat{q}\rangle}{\varepsilon\hbar}\,\mathrm{d}q \ ,$$

which can be integrated to yield

$$\log\psi = \frac{1}{\varepsilon\hbar}\left[\frac{q^2}{2} + (i\varepsilon\langle\hat{p}\rangle - \langle\hat{q}\rangle)q\right] + \text{const.}$$

Finally, by taking the exponential of both sides we get

$$\begin{aligned}\psi(q) &= A\exp\left[\frac{q^2}{2\varepsilon\hbar} - \frac{\langle\hat{q}\rangle}{\varepsilon\hbar}q + \frac{i\langle\hat{p}\rangle}{\hbar}q\right] \\ &= A\exp\left[\frac{(q-\langle\hat{q}\rangle)^2}{2\varepsilon\hbar} + i\frac{\langle\hat{p}\rangle}{\hbar}q - \frac{\langle\hat{q}\rangle^2}{2\varepsilon\hbar}\right] .\end{aligned}$$

A can be found by normalizing the wave function. Of course, ε should be less than zero in order ψ to be normalizable. The result is

$$\psi(q) = \sqrt[4]{\frac{2}{\hbar|\varepsilon\pi|}}\exp\left[\frac{-(q-\langle\hat{q}\rangle)^2}{2|\varepsilon|\hbar} + i\frac{\langle\hat{p}\rangle}{\hbar}q\right] . \tag{3.26}$$

$\star \quad \star \quad \star$

Problem 3.2. Determine the momentum probability distribution of the energy eigenstates of an harmonic oscillator whose mass and frequency are m and ω respectively.

SOLUTION. The Hamiltonian is

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 , \tag{3.27}$$

and the energy eigenfunctions in the position representation are

$$\psi_n(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{2}}\frac{1}{2^{n/2}\sqrt{n!}}\,\mathrm{e}^{-\frac{m\omega x^2}{2\hbar}}\,H_n\left(x\sqrt{\frac{m\omega}{\hbar}}\right) , \tag{3.28}$$

where

$$H_n(\xi) = (-1)^n\,\mathrm{e}^{\xi^2}\,\frac{\mathrm{d}^n\,\mathrm{e}^{-\xi^2}}{\mathrm{d}\xi^n}$$

are the Hermite polynomials.

Of course, we could Fourier transform $\psi_n(x)$ to obtain the momentum-space wave function $\tilde{\psi}(p)$ whose square modulus is nothing but the desired probability density function. However, it seems easier to write $\hat{H}$ in the momentum representation:

$$\hat{H}_p = \frac{p^2}{2m} - \frac{\hbar^2}{2} m\omega^2 \frac{\mathrm{d}^2}{\mathrm{d}p^2} \, .$$

The eigenvalue problem for the operator $\hat{H}_p$ reads

$$\left(-\frac{\hbar^2}{2} m\omega^2 \frac{\mathrm{d}^2}{\mathrm{d}p^2} + \frac{p^2}{2m} \right) \tilde{\psi}_n(p) = E_n \tilde{\psi}_n(p) \, ,$$

which can be written as

$$\frac{\mathrm{d}^2 \tilde{\psi}_n(p)}{\mathrm{d}p^2} + \frac{2}{\hbar^2 m\omega^2} \left(E_n - \frac{p^2}{2m} \right) \tilde{\psi}_n(p) = 0 \, . \tag{3.29}$$

The position-space stationary Schrödinger equation had the form

$$\frac{\mathrm{d}^2 \psi_n(x)}{\mathrm{d}x^2} + \frac{2m}{\hbar^2} \left(E_n - \frac{1}{2} m\omega^2 x^2 \right) \psi_n(x) = 0 \, . \tag{3.30}$$

By performing the replacement

$$\frac{p}{\sqrt{m\omega\hbar}} \quad \text{with} \quad x\sqrt{\frac{m\omega}{\hbar}}$$

Eqn. (3.29) formally becomes equivalent to Eqn. (3.30) whose solution are Eqn. (3.28). By the same substitution into Eqn. (3.28), and properly normalizing, we get:

$$\tilde{\psi}_n(p) = \frac{1}{\sqrt{2^n n!}\sqrt{\pi m\omega\hbar}} \, \mathrm{e}^{-\frac{p^2}{2m\omega\hbar}} \, H_n\left(\frac{p}{\sqrt{m\omega\hbar}} \right) \, .$$

$\star \quad \star \quad \star$

Problem 3.3. Let

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{q})$$

be the Hamiltonian of a one-dimensional quantum system having discrete spectrum with eigenstates $|n\rangle$ and eigenvalues E_n. By using the commutation relation

Eqn. (3.10) between position coordinate $\hat{q}$ and momentum $\hat{p}$, prove the following relation:

$$\sum_n (E_n - E_0) \left|\langle n|\hat{q}|0\rangle\right|^2 = \frac{\hbar^2}{2m} ,$$

where $|0\rangle$ and E_0 are the ground state and its corresponding eigenvalue respectively.

SOLUTION. By direct evaluation we have

$$[\hat{H}, \hat{q}] = \frac{1}{2m}[\hat{p}^2, \hat{q}] = -\frac{i\hbar}{m}\hat{p} ,$$

hence

$$\Big[[\hat{H}, \hat{q}], \hat{q}\Big] = -\frac{i\hbar}{m}[\hat{p}, \hat{q}] = -\frac{\hbar^2}{m} ,$$

from which it follows

$$\Big\langle n \Big| \Big[[\hat{H}, \hat{q}], \hat{q}\Big] \Big| n \Big\rangle = -\frac{\hbar^2}{m} . \tag{3.31}$$

The left-hand side in the latter expression is also equal to

$$\begin{aligned}
\Big\langle n \Big| \Big[[\hat{H}, \hat{q}], \hat{q}\Big] \Big| n \Big\rangle &= \langle n|\hat{H}\hat{q}^2 - 2\hat{q}\hat{H}\hat{q} + \hat{q}^2\hat{H}|n\rangle \\
&= 2E_n \langle n|\hat{q}^2|n\rangle - 2\langle n|\hat{q}\hat{H}\hat{q}|n\rangle \\
&= \sum_m 2E_n \langle n|\hat{q}|m\rangle \langle m|\hat{q}|n\rangle - \sum_{m,l} 2\langle n|\hat{q}|m\rangle \langle m|\hat{H}|l\rangle \langle l|\hat{q}|n\rangle \\
&= \sum_m 2E_n \left|\langle n|\hat{q}|m\rangle\right|^2 - \sum_{m,l} 2\langle n|\hat{q}|m\rangle E_m \delta_{l,m} \langle l|\hat{q}|n\rangle \\
&= \sum_m \Big\{2E_n \left|\langle n|\hat{q}|m\rangle\right|^2 - 2E_m \left|\langle n|\hat{q}|m\rangle\right|^2\Big\} \\
&= 2\sum_m (E_n - E_m) \left|\langle n|\hat{q}|m\rangle\right|^2 .
\end{aligned}$$

By taking $n = 0$ and comparing with the right-hand side of Eqn. (3.31) we get

$$\sum_m (E_m - E_0) \left|\langle m|\hat{q}|0\rangle\right|^2 = \frac{\hbar^2}{2m} .$$

$\star \quad \star \quad \star$

Problem 3.4. Let $\hat{f}$ be an Hermitian operator having a discrete spectrum with N different eigenvalues. Prove that $\hat{f}^N$ can be expressed as a linear combination of the operators $1, \hat{f}, \ldots, \hat{f}^{N-1}$.

Solution. We define the following operator:

$$\begin{aligned} \hat{G} &\equiv \prod_{i=1}^{N} \left(\hat{f} - f_i\right) \\ &= (\hat{f} - f_1)(\hat{f} - f_2)\cdots(\hat{f} - f_N) \, . \end{aligned}$$

Let $\hat{f}$ be hypermaximal, so that its eigenstates $|f_i\rangle$ form a complete orthonormal basis and any state $|\psi\rangle$ can be expanded as

$$|\psi\rangle = \sum_k a_k \, |f_k\rangle \ , \quad a_k = \langle f_k|\psi\rangle \ .$$

Now, let us apply $\hat{G}$ to $|\psi\rangle$:

$$\hat{G}\,|\psi\rangle = \Big[(\hat{f} - f_1)(\hat{f} - f_2)\cdots(\hat{f} - f_N)\Big] \left(\sum_{k=1}^{N} a_k \, |f_k\rangle\right) \ .$$

We note that

$$\begin{aligned} \hat{G}\,|\psi\rangle &= \Big[(\hat{f} - f_1)(\hat{f} - f_2)\cdots(\hat{f} - f_N)\Big] \left(a_N \, |f_N\rangle + \sum_{k=1}^{N-1} a_k \, |f_k\rangle\right) \\ &= \Big[(\hat{f} - f_1)(\hat{f} - f_2)\cdots(\hat{f} - f_N)\Big] \, a_N \, |f_N\rangle + \\ &\quad \Big[(\hat{f} - f_1)(\hat{f} - f_2)\cdots(\hat{f} - f_N)\Big] \left(\sum_{k=1}^{N-1} a_k \, |f_k\rangle\right) \\ &= \Big[(\hat{f} - f_1)(\hat{f} - f_2)\cdots(\hat{f} - f_N)\Big] \left(\sum_{k=1}^{N-1} a_k \, |f_k\rangle\right) \ . \end{aligned}$$

This is so because

$$(\hat{f} - f_N) a_N \, |f_N\rangle = a_N \hat{f} \, |f_N\rangle - a_N f_N \, |f_N\rangle = 0.$$

By considering the next terms in the sum, we get by induction

$$\hat{G}\,|\psi\rangle = 0\ ,$$

that is, $\hat{G} = 0$ ($\hat{G}$ is the null operator). Thus,

$$\hat{G} = 0 = \hat{f}^N - \sum_{i=1}^{N} f_i \hat{f}^{N-1} + \frac{1}{2}\sum_{i,j} f_i f_j \hat{f}^{N-2} + \cdots + (-1)^N \prod_{i=1}^{N} f_i\ ,$$

and as a consequence

$$\hat{f}^N = \sum_{i=1}^{N} f_i \hat{f}^{N-1} + \frac{1}{2}\sum_{i,k} f_i f_k \hat{f}^{N-2} + \cdots + (-1)^N \prod_{i=1}^{N} f_i\ .$$

$\star \quad \star \quad \star$

Problem 3.5. A particle of mass m moves in a infinite potential well of width a. At the time $t = 0$ the particle was prepared in the state given by the wave function

$$\psi(x, t = 0) = A \sin^3 \frac{\pi x}{a}\ . \tag{3.32}$$

The questions are:

1. Find the wave function $\psi(x, t)$ at any later time $t > 0$.
2. Find the time T at which the particle turns back to its initial state.

SOLUTION. We recall the energy levels of the one-dimensional infinite potential well are

$$E_n = \frac{\hbar^2 \pi^2 (n+1)^2}{2ma^2}\ , \quad n = 0, 1, 2, \ldots$$

The corresponding normalized wave eigenfunctions are

$$\psi_n(x) = \langle x | E_n \rangle = \sqrt{\frac{2}{a}} \sin \frac{\pi(n+1)x}{a}\ .$$

As we have just seen, the state $|\psi(t)\rangle$ at any later time is found by applying the unitary time-evolution operator $\hat{U}(t) = \exp(-i\hat{H}t/\hbar)$ to the initial state $|\psi(0)\rangle$. For this procedure to work properly we have to determine the expansion coefficients c_n of $|\psi(0)\rangle$ on the Hamiltonian eigenstates $|E_n\rangle$ whose wave function $\langle x|E_n\rangle$ is given above:

$$|\psi(0)\rangle = \sum_n c_n |E_n\rangle \ , \quad c_n = \langle E_n|\psi(0)\rangle \ .$$

Inserting the identity $\int |x\rangle \langle x| \, \mathrm{d}x = 1$ yields

$$c_n = \langle E_n|\psi(0)\rangle = \int \underbrace{\langle E_n|x\rangle}_{\psi_n^*(x)} \underbrace{\langle x|\psi(0)\rangle}_{\psi(x,t=0)} \mathrm{d}x \ .$$

The integral involved in the calculation of c_n must in principle be solved for each n:

$$c_n = \sqrt{\frac{2}{a}} A \int_0^a \sin\frac{\pi(n+1)x}{a} \sin^3\frac{\pi x}{a} \, \mathrm{d}x \ .$$

There is however a faster way to get the correct result, that is, we can manipulate directly the initial wave function Eqn. (3.32) in order to identify the components. We can write

$$\begin{aligned}
\psi(x, t=0) &= A \sin^3\frac{\pi x}{a} \\
&= A \sin\frac{\pi x}{a}\left[1 - \cos^2\frac{\pi x}{a}\right] \\
&= A \sin\frac{\pi x}{a} - A\frac{1}{2}\left[1 + \cos\frac{2\pi x}{a}\right]\sin\frac{\pi x}{a} \\
&= \frac{A}{2}\sin\frac{\pi x}{a} - \frac{A}{2}\sin\frac{\pi x}{a}\cos\frac{2\pi x}{a} \\
&= \frac{A}{2}\sin\frac{\pi x}{a} - \frac{A}{4}\left[\sin\frac{3\pi x}{a} - \sin\frac{\pi x}{a}\right] \\
&= \frac{3}{4}A\sin\frac{\pi x}{a} - \frac{1}{4}A\sin\frac{3\pi x}{a}
\end{aligned}$$

The trigonometric identity

$$\cos p \sin q = \frac{1}{2}\left[\sin(p+q) - \sin(p-q)\right]$$

has been used in the last steps above. In this way we get the initial state as a sum of $\langle x|E_0\rangle$

and $\langle x|E_2\rangle$. So, the time evolution is given by

$$\begin{aligned}\psi(x,t) &= \mathrm{e}^{-\frac{i}{\hbar}\hat{H}t}\,\psi(x,t=0)\\ &= \mathrm{e}^{-\frac{i}{\hbar}\hat{H}t}\sum_{n=0}^{\infty} c_n\psi_n(x) \quad c_n = \int_0^a \psi_n^*(x)\psi(x,t=0)\,\mathrm{d}x\ ,\\ &= \sum_{n=0}^{\infty} c_n\,\mathrm{e}^{-\frac{i}{\hbar}E_n t}\,\psi_n(x)\\ &= c_0\,\mathrm{e}^{-\frac{i}{\hbar}E_0 t}\,\psi_0(x) + c_2\,\mathrm{e}^{-\frac{i}{\hbar}E_2 t}\,\psi_2(x)\ ;\end{aligned}$$

where, $c_0 = \frac{3}{4}A\sqrt{\frac{a}{2}}$ and $c_2 = -\frac{1}{4}A\sqrt{\frac{a}{2}}$. The final result is

$$\psi(x,t) = \frac{A}{4}\sqrt{\frac{2}{a}}\left[3\sin\frac{\pi x}{a}\exp\left(-i\frac{\hbar\pi^2 t}{2ma^2}\right) - \sin\frac{3\pi x}{a}\exp\left(-9i\frac{\hbar\pi^2 t}{2ma^2}\right)\right]\ ,$$

which can also be written as

$$\psi(x,t) = \frac{A}{4}\sqrt{\frac{2}{a}}\exp\left(-i\frac{\hbar\pi^2 t}{2ma^2}\right)\left[3\sin\frac{\pi x}{a} - \exp\left(-i\frac{4\hbar\pi^2 t}{ma^2}\right)\sin\frac{3\pi x}{a}\right]\ .$$

For the wave function $\psi(x,T)$ at a later time $T > 0$ to be equal to $\psi(x,t=0)$ the following condition must hold:

$$\exp\left(-i\frac{\hbar\pi^2 T}{2ma^2}\right) = \exp\left(-4i\frac{\hbar\pi^2 T}{ma^2}\right) = 1\ ,$$

that is

$$\frac{\hbar\pi^2 T}{2ma^2} = n(2\pi)\ , \quad 4\frac{\hbar\pi^2 T}{ma^2} = 8n(2\pi)\ ,$$

for some integer n. By taking $n = 1$ (we find the smallest period T)

$$\frac{\hbar\pi^2 4T_{\min}}{ma^2} = 16\pi\ ,$$

from which it follows

$$T_{\min} = \frac{16\pi ma^2}{4\hbar\pi^2} = \frac{4ma^2}{\hbar\pi}\ .$$

$\star \quad \star \quad \star$

Problem 3.6. The state of a free particle at initial time $t = 0$ is given by the following Gaussian wave function:

$$\psi(x, t = 0) = A \exp\left(-\frac{x^2}{2a^2} + \frac{imv_0 x}{\hbar}\right) .$$

Calculate the wave function $\psi(x, t)$ at the time t and also the expectation value and the mean square deviation of the position operator $\hat{x}$, that is, $\langle\psi(t)|\hat{x}|\psi(t)\rangle$ and $\langle\psi(t)|\Delta\hat{x}^2|\psi(t)\rangle$ respectively.

Solution. We need to recall the following Gaussian integrals:

$$\int_0^\infty \exp\left(-ax^2\right) \mathrm{d}x = \frac{1}{2}\sqrt{\frac{\pi}{a}} ,$$

$$\int_{-\infty}^{+\infty} \exp\left(-ax^2 + bx\right) \mathrm{d}x = \sqrt{\frac{\pi}{a}} exp - \frac{b^2}{4a} ,$$

$$\int_0^\infty \exp\left(-ax^2\right) x^{2k} \mathrm{d}x = \frac{1 \times 3 \times \cdots \times (2k-1)}{2^{k+1}} \sqrt{\frac{\pi}{a^{2k+1}}} ,$$

$$\int_0^\infty \exp\left(-ax^2\right) x^{2k+1} \mathrm{d}x = \frac{k!}{2a^{k+1}} ,$$

$$\int_{-\infty}^{+\infty} \mathrm{d}x_1 \cdots \int_{-\infty}^{+\infty} \mathrm{d}x_N \exp\left(-\sum a_{ij} x_j x_i\right) = \pi^{\frac{N}{2}} \left[\det A_{ij}\right]^{-1/2} .$$

The Hamiltonian of a free particle is

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

its eigenstates are

$$\hat{H} \left|\pm p\right\rangle = \frac{p^2}{2m} \left|\pm p\right\rangle ,$$

thus $\hat{H}$ does not form a complete set since we still have the degeneracy $|\pm p\rangle$. For this reason, we shall use instead the eigenstates of $\hat{p}$:

$$\hat{p} \left|p\right\rangle = p \left|p\right\rangle .$$

We expand the initial state $|\psi(0)\rangle$ on the basis of the momentum eigenstates:

$$\left|\psi(0)\right\rangle = \int C(p) \left|p\right\rangle \mathrm{d}p , \quad C(p) = \langle p|\psi(0)\rangle .$$

It should be noted that we know how the momentum eigenstates do evolve in time under the free Hamiltonian. $C(p')$ is given by

$$\begin{aligned}\langle p'|\psi(0)\rangle &= \int \langle p'|x\rangle \langle x|\psi(0)\rangle \,\mathrm{d}x \\ &= \int \frac{1}{\sqrt{2\pi\hbar}} \mathrm{e}^{-\frac{ip'x}{\hbar}} \psi(x,0)\,\mathrm{d}x \\ &= \frac{A}{\sqrt{2\pi\hbar}} \int \exp\left[-\frac{x^2}{2a^2} + i\frac{mv_0x}{\hbar} - i\frac{p'x}{\hbar}\right]\mathrm{d}x\,.\end{aligned}$$

The latter is a Gaussian integral, which can be evaluated exactly to yield:

$$C(p) = aA\exp\left[-\frac{1}{2\hbar}(mv_0a - pa)^2\right]\,.$$

By applying the time-evolution operator, we get

$$\begin{aligned}\psi(x,t) &= \mathrm{e}^{-\frac{i}{\hbar}\hat{H}t}\,\psi(x,t=0) \\ &= \mathrm{e}^{-\frac{i}{\hbar}\hat{H}t}\int C(p)\,\underbrace{\varphi_p(x)}_{\langle x|p\rangle}\,\mathrm{d}p \\ &= \frac{aA}{\sqrt{2\pi\hbar}}\int \exp\left[-\frac{ip^2t}{2m\hbar} + i\frac{px}{\hbar} - \frac{a^2}{2\hbar^2}(mv_0-p)^2\right]dp \\ &= A\left[1+\frac{i\hbar t}{ma^2}\right]^{-\frac{1}{2}}\exp\Big\{\left[-ma^2\hbar^2(x-v_0t) + \hbar^3x^2t + ia^4m^2v_0\hbar(2x-v_0t)\right] \\ &\qquad \times\left[2m(a^4\hbar^2 + \frac{t^2\hbar^2}{m^2})\right]^{-1}\Big\}\end{aligned}$$

A is determined by normalizing the wave function: $\int |\psi(x,t)|^2\,\mathrm{d}x = 1$. The reason for which we require $\psi(x,t)$ to be normalized is that we are interested in finding mean values and mean square deviations, thus the wave function must be normalized. We have

$$|\psi(x,t)|^2 = \frac{A^2}{\sqrt{1+\frac{\hbar^2t^2}{m^2a^4}}}\exp\left[-\frac{(x-v_0t)^2}{a^2\left(1+\frac{t^2\hbar^2}{m^2a^4}\right)}\right]\,.$$

By imposing $\int |\psi(x,t)|^2\,\mathrm{d}x = 1$ one finds

$$A^2 = (\pi a^2)^{-\frac{1}{2}}\,.$$

The expectation value of the position operator $\hat{x}$ on the evolved state $|\psi(t)\rangle$ is

$$\langle\psi(t)|\hat{x}|\psi(t)\rangle = \int x\,|\psi(x,t)|^2\,\mathrm{d}x\,.$$

Again, this is nothing but a Gaussian integral which can be calculate explicitly. The result is $v_0 t$. It should be noted at this point that the quantum expectation value obey the classical laws of motion. This is a more general result which we shall encounter again later: the so-called "Ehrenfest theorem".

The mean square deviation of the position operator $\hat{x}$ on the evolved state $|\psi(t)\rangle$ can be calculated in the same manner using the formulas for the Gaussian integrals and the result is

$$\frac{a^2}{2}\left(1+\frac{t^2\hbar^2}{m^2a^4}\right) .$$

Note that at $t = 0$ the mean square deviation is $a/\sqrt{2}$ while for any other $t > 0$ the mean square deviation is greater than the initial value. This spreading of the wave function is a typical quantum effect.

$\star \quad \star \quad \star$

Chapter 4

HARMONIC OSCILLATOR IN OPERATORIAL FORM.

4.1 Harmonic Oscillator in Operatorial Form.

The Hamiltonian operator of the one-dimensional quantum harmonic oscillator is given by

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + \frac{1}{2}m\omega^2\hat{q}^2 \ . \tag{4.1}$$

Let us recall that the position and momentum operators satisfy the commutation relation

$$[\hat{q}, \hat{p}] = i\hbar \ .$$

To simplify the notations, let us introduce the following operators:

$$\widehat{\tilde{H}} \equiv \frac{\hat{H}}{\hbar\omega}$$

and

$$\hat{Q} \equiv \left(\frac{m\omega}{\hbar}\right)^{\frac{1}{2}} \hat{q} \ ,$$
$$\hat{P} \equiv \left(\frac{1}{m\hbar\omega}\right)^{\frac{1}{2}} \hat{p} \ ,$$

so that

$$\widehat{\tilde{H}} = \frac{1}{2}\left(\hat{P}^2 + \hat{Q}^2\right) \ ,$$

with

$$[\hat{Q}, \hat{P}] = i \ .$$

Let us also define the following other operators:

$$\hat{a} \equiv \frac{1}{2}\sqrt{2}\left(\hat{Q} + i\hat{P}\right) ,$$
$$\hat{a}^\dagger \equiv \frac{1}{2}\sqrt{2}\left(\hat{Q} - i\hat{P}\right) .$$

A direct computation shows that :

$$[\hat{a}, \hat{a}^\dagger] = 1$$

and

$$\widehat{\tilde{H}} = \frac{1}{2}\left(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\right) .$$

By introducing the operator

$$\hat{N} \equiv \hat{a}^\dagger\hat{a} ,$$

we can rewrite the operator $\widehat{\tilde{H}}$ as

$$\hat{\tilde{H}} = \hat{N} + \frac{1}{2} . \tag{4.2}$$

It is easy to work out the commutation rules of $\hat{a}$ and $\hat{a}^\dagger$ with $\hat{N}$:

$$\hat{N}\hat{a} = \hat{a}\left(\hat{N} - 1\right) , \tag{4.3}$$
$$\hat{N}\hat{a}^\dagger = \hat{a}^\dagger\left(\hat{N} + 1\right) . \tag{4.4}$$

We are now in a position to prove the following theorem.

THEOREM:

If

$$\hat{N}\,|\nu\rangle = \nu\,|\nu\rangle ,$$

i.e., if $|\nu\rangle$ is an eigenvector of $\hat{N}$ with corresponding eigenvalue ν, then:

a) $\nu \geq 0$;

b) if $\nu = 0$, then $\hat{a}\,|\nu\rangle = 0$;

c) if $\nu \neq 0$, then $\hat{a}\,|\nu\rangle$ is a non-zero vector whose norm is given by $\nu\,\langle\nu|\nu\rangle$ and which

is an eigenvector of $\hat{N}$ with eigenvalue $\nu - 1$; $\hat{a}^\dagger |\nu\rangle$ is a state which is always different from zero (even if $\nu = 0$) and its norm is given by

$$(\nu + 1) \langle \nu | \nu \rangle \ .$$

Morewover it is an eigenvector of $\hat{N}$ with eigenvalue $\nu + 1$.

proof:

By hypothesis,

$$\hat{N} |\nu\rangle = \nu |\nu\rangle \ ,$$

with $\langle \nu | \nu \rangle > 0$. So,

$$\langle \nu | \hat{N} | \nu \rangle = \langle \nu | \hat{a}^\dagger \hat{a} | \nu \rangle = \| \hat{a} |\nu\rangle \|^2 = \nu \langle \nu | \nu \rangle \ . \tag{4.5}$$

and

$$\left\| \hat{a}^\dagger |\nu\rangle \right\|^2 = \langle \nu | \hat{a} \hat{a}^\dagger | \nu \rangle = \langle \nu | \hat{N} + 1 | \nu \rangle = (\nu + 1) \langle \nu | \nu \rangle \ . \tag{4.6}$$

Now, by definition all the vectors in the Hilbert space have non-negative norms and the only vector having zero norm is the null vector. From Eqn. (4.5) we get that, since $\| \hat{a} |\nu\rangle \|^2$ is the norm of a state, it is a positive number as well as $\langle \nu | \nu \rangle$, so we have $\nu \geq 0$. This proves proposition **a)**.

If $\nu = 0$, then $\| \hat{a} |\nu\rangle \| = 0$, and consequently $\hat{a} |\nu\rangle = 0$, which proves proposition **b)**. If $\nu \neq 0$, then $\hat{a} |\nu\rangle$ is a vector with non-zero norm given in (4.5). Using Eqn. (4.3) we get

$$\hat{N} \hat{a} |\nu\rangle = \hat{a} \left(\hat{N} - 1 \right) |\nu\rangle = (\nu - 1) \hat{a} |\nu\rangle \ .$$

This proves part of point **c)**. From (4.6) we get the norm of $\hat{a}^\dagger |\nu\rangle$. Since $\nu \geq 0$ the minimum value of this norm is $1 \cdot \langle \nu | \nu \rangle$ when $\nu = 0$, and even this is a norm different from zero. Now let us take into account Eqn. (4.4):

$$\hat{N} \hat{a}^\dagger |\nu\rangle = \hat{a}^\dagger \left(\hat{N} + 1 \right) |\nu\rangle = (\nu + 1) \hat{a}^\dagger |\nu\rangle \ .$$

This relation tells us that $\hat{a}^\dagger |\nu\rangle$ is eigenvector of $\hat{N}$ with eigenvalue $\nu + 1$. This completes the proof of the theorem. Via repeated applications of the operator $\hat{a}$ on

the eigenvector $|\nu\rangle$ we get that the states :

$$\hat{a}\,|\nu\rangle \ , \quad \hat{a}^2\,|\nu\rangle \ , \quad \ldots , \quad \hat{a}^p\,|\nu\rangle \ ,$$

are eigenvectors of $\hat{N}$ with eigenvalues

$$\nu - 1\,, \quad \nu - 2\,, \quad \ldots\ , \quad \nu - p\,.$$

From proposition **a)** of the previous theorem, these eigenvalues must be positive or zero, so there must exists a certain positive integer p for which we have $\nu - p = 0$, *i.e.*, $\nu = p$. Therefore, we get the important result that the generic eigenvalue ν is zero or a positive *integer* because p is integer. If $\nu - p$ were not zero but the smallest non-integer positive number , then we would always be able to construct the state $\hat{a}\,|\nu - p\rangle$ which would be an eigenvector of $\hat{N}$ with eigenvalue $(\nu - p - 1)$, whose norm would be proportional to $\nu - p - 1$ and negative, but this cannot happen in a Hilbert space. This proves that the eigenvalues ν are given by all the (non-negative) integers.

From this and Eqn.(4.2) we get that, calling by n a positive integer, the energy levels of the original Hamiltonian are:

$$E_n = (n + \frac{1}{2})\hbar\omega$$

Now, $\hat{N}$ is a Hermitian operator, so its eigenvectors $|0\rangle$, $|1\rangle$, ..., $|n\rangle$, ..., make up a basis for the Hilbert space. There are no degeneracies and they make up a complete set. From the relation

$$\langle\nu|\hat{a}^\dagger\hat{a}|\nu\rangle = \|\hat{a}\,|\nu\rangle\|^2 = \nu\,\langle\nu|\nu\rangle \ ,$$

we get that $\hat{a}\,|\nu\rangle$ is normalized if $\nu = 1$ and $\langle\nu|\nu\rangle = 1$. Then, from

$$\langle\nu|\hat{a}\hat{a}^\dagger|\nu\rangle = \langle\nu|\hat{N} + 1|\nu\rangle = (\nu + 1)\,\langle\nu|\nu\rangle \ ,$$

we get that

$$\left\|\hat{a}^\dagger\,|1\rangle\right\|^2 = 2\,\langle 1|1\rangle = 2$$

so

$$\left\|\hat{a}^\dagger\,|1\rangle\right\| = \sqrt{2}\;.$$

But we already know that $\hat{a}^\dagger\,|1\rangle \propto |2\rangle$ and if we suppose that all the states are normalized, *i.e.*, $\langle 2|2\rangle = 1$, then the correct proportionality factor is such that:

$$\hat{a}^\dagger\,|1\rangle = \sqrt{2}\,|2\rangle\;,$$

and in general we shall have

$$\hat{a}^\dagger\,|n\rangle = (n+1)^{1/2}\,|n+1\rangle\;.$$

Analogously, we can determine the normalization coefficient in the relation $\hat{a}\,|n\rangle \propto |n-1\rangle$. If $|n\rangle$ is normalized then

$$\langle n|\hat{a}^\dagger\hat{a}|n\rangle = \|\hat{a}\,|n\rangle\|^2 = n\,\langle n|n\rangle = n\;,$$

and so

$$\|\hat{a}\,|n\rangle\| = \sqrt{n}\;, \qquad \hat{a}\,|0\rangle = 0\;.$$

We can conclude that the vectors

$$|n\rangle = (n!)^{-\frac{1}{2}}\;\hat{a}^{\dagger\,n}\,|0\rangle\;,$$

are eigenstates of $\hat{N}$:

$$\hat{N}\,|n\rangle = n\,|n\rangle\;,$$

and they make an orthonormal basis:

$$\langle n|n'\rangle = \delta_{n,n'}\;.$$

To check the orthogonality, let us consider for example $\langle 1|2\rangle$:

$$\begin{aligned}
\langle 1|2\rangle &= \frac{1}{\sqrt{2}}\langle 0|\hat{a}\hat{a}^\dagger\hat{a}^\dagger|0\rangle \\
&= \frac{1}{\sqrt{2}}\langle 0|\left(\hat{a}^\dagger\hat{a}+1\right)\hat{a}^\dagger|0\rangle \\
&= \frac{1}{\sqrt{2}}\langle 0|\hat{a}^\dagger\hat{a}\hat{a}^\dagger|0\rangle + \frac{1}{\sqrt{2}}\langle 0|\hat{a}^\dagger|0\rangle \\
&= \frac{1}{\sqrt{2}}\langle 0|\hat{a}^\dagger\hat{a}\hat{a}^\dagger|0\rangle \\
&= \frac{1}{\sqrt{2}}\langle 0|\hat{a}^\dagger\left(1-\hat{a}^\dagger\hat{a}\right)|0\rangle \\
&= \frac{1}{\sqrt{2}}\langle 0|\hat{a}^\dagger|0\rangle - \frac{1}{\sqrt{2}}\langle 0|\hat{a}^\dagger\hat{a}^\dagger\hat{a}|0\rangle = 0\ .
\end{aligned}$$

The same can be easily proved for all states labelled by different integers.

The fact the states are labelled by positive integer agrees with the result we got in chapter 1) where we solved explicitly the Schroedinger equation associated to the harmonic oscillator and did not use operator methods like we did here. The wave-functions we obtained in chapter 1) are nothing else and the x-representation of the states $|n\rangle$,i.e:

$$\phi_n^{osc.}(x) = \langle x|n\rangle$$

Chapter 5

ANGULAR MOMENTUM QUANTIZATION

In classical mechanics the angular momentum of a particle is defined as

$$\boldsymbol{L} = \boldsymbol{x} \times \boldsymbol{p} \ ,$$

where the symbol "$\times$" indicates the vector product. In components, we get

$$L_x = yp_z - zp_y \ , \quad L_y = zp_x - xp_z \ , \quad L_z = xp_y - yp_x \ .$$

By using the correspondence rules, the three components of the angular momentum operator in quantum mechanics are:

$$\begin{aligned}
\hat{L}_x &= \hat{y}\hat{p}_z - \hat{z}\hat{p}_y = -i\hbar \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \right) \ , \\
\hat{L}_y &= \hat{z}\hat{p}_x - \hat{x}\hat{p}_z = -i\hbar \left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z} \right) \ , \\
\hat{L}_z &= \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = -i\hbar \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} \right) \ .
\end{aligned}$$

We know that observables in quantum mechanics must be Hermitian operators. Are $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$ actually Hermitian? It is easy to check that in fact they are.

By using the fundamental commutation rules among position and momentum observables, namely,

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{i,j} \ ,$$

one can immediately calculate the corresponding commutation rules among the vari-

ous components of the angular momentum operator, for example

$$\begin{aligned}
[\hat{L}_x, \hat{L}_y] &= [\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z] \\
&= [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z] - [\hat{z}\hat{p}_y, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z] \\
&= \hat{y}[\hat{p}_z, \hat{z}]\hat{p}_x - [\hat{z}\hat{p}_y, (-\hat{x}\hat{p}_z)] \\
&= -i\hbar\hat{y}\hat{p}_x + \hat{x}[\hat{z}, \hat{p}_z]\hat{p}_y \\
&= i\hbar\left(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x\right) \\
&= i\hbar\hat{L}_z \ .
\end{aligned}$$

The other commutators are:

$$[\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y \ , \qquad [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x \ .$$

There is a compact notation to write these formulas, namely,

$$[\hat{L}_i, \hat{L}_j] = i\hbar\varepsilon_{ijk}\hat{L}_k \ , \tag{5.1}$$

where ε_{ijk} is the totally-antisymmetric tensor whose value is

$$\begin{cases}
\ \ 1 & \text{if} \quad \text{even permutation of : } x, y, z \\
\ \ 0 & \text{if} \quad \text{two indices are repeated} \\
-1 & \quad\ \ \text{otherwise}
\end{cases}$$

Note that this result is the one obtained, via the correspondence principle, from the classical case where the Poisson brackets between the components of the angular momentum satisfy the relation

$$\{L_i, L_j\} = \varepsilon_{ijk}L_k \ . \tag{5.2}$$

Homework 5.1. A) Check Eqn. (5.2). **B)** Find the commutator between $\hat{L}_i$ and $\hat{q}_j$, where the latter are the position variables.

5.1 Spectrum of the angular momentum operators

In this section we look for the spectrum of the angular momentum operators. We shall see that the spectrum of the angular momentum is *discrete* in contrast with the classical case. All the results of this section actually relies only on the commutation rules (5.1) and thus they apply to every triplet of operators whose algebra satisfies Eqn. (5.1). To emphasize this fact, we shall replace the notation $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$ with the more general $\hat{J}_i$ where i stands for x, y or z, or more generally $i = 1, 2, 3$. The commutation relations read

$$[\hat{J}_i, \hat{J}_j] = i\hbar\varepsilon_{ijk}\hat{J}_k \ . \tag{5.3}$$

The angular momentum is not the only operator satisfying this algebra: also spin, isotopic spin, etc. do obey the same commutation rules.

It should be noted at this point that the three operators $\hat{J}_i$ do not commute with each other. Therefore, we cannot diagonalize two components simultaneously. But it is easy to check that the following two operators do commute: $\hat{J}_z$ and $\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$. In fact,

$$\begin{aligned}
[\hat{J}_z, \hat{J}^2] &= [\hat{J}_z, \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2] \\
&= [\hat{J}_z, \hat{J}_x^2] + [\hat{J}_z, \hat{J}_y^2] + \underbrace{[J_z, \hat{J}_z^2]}_{=0} \\
&= [\hat{J}_z, \hat{J}_x]\hat{J}_x + \hat{J}_x[\hat{J}_z, \hat{J}_x] + [\hat{J}_z, \hat{J}_y]\hat{J}_y + \hat{J}_y[\hat{J}_z, \hat{J}_y] \\
&= i\hbar\left(\hat{J}_y\hat{J}_x + \hat{J}_x\hat{J}_y - \hat{J}_x\hat{J}_y - \hat{J}_y\hat{J}_x\right) \\
&= 0 \ .
\end{aligned}$$

In the same way it is also possible to prove that

$$[\hat{J}_x, \hat{J}^2] = 0 \ , \qquad [\hat{J}_y, \hat{J}^2] = 0 \ .$$

Since $\hat{J}_z$ and $\hat{J}^2$ are hermitian and commute, they can be diagonalized simultaneously

and their common eigenstates can be chosen to be orthogonal:

$$\begin{cases} \hat{J}^2 \left|\lambda, m\right\rangle &= \hbar^2 \lambda \left|\lambda, m\right\rangle \\ \hat{J}_z \left|\lambda, m\right\rangle &= \hbar\ m \left|\lambda, m\right\rangle \end{cases} \tag{5.4}$$

The eigenvalues $\hbar^2\lambda$ and $\hbar m$ have been written in this form in order for λ and m to be dimensionless, since $\hat{J}^2$ and $\hat{J}_z$ have dimensions of $\hbar^2$ and $\hbar$ respectively. $|\lambda, m\rangle$ denotes the (normalized) common eigenstate. For the time being, no restriction is required for λ and m; of course, they must be *real* numbers since $\hat{J}^2$ and $\hat{J}_z$ are Hermitian and thus their eigenvalues are real numbers.

The next step is to introduce the two auxiliary operators

$$\hat{J}_{(+)} \equiv \hat{J}_x + i\hat{J}_y \ , \quad \hat{J}_{(-)} \equiv \hat{J}_x - i\hat{J}_y \ .$$

These two operators are one the Hermitian conjugate of the other:

$$\hat{J}^\dagger_{(+)} = \hat{J}_{(-)} \ , \quad \hat{J}^\dagger_{(-)} = \hat{J}_{(+)} \ .$$

Moreover, it is easy to prove that the following relations hold:

$$[\hat{J}_z, \hat{J}_{(+)}] = \hbar \hat{J}_{(+)} \ , \tag{5.5a}$$

$$[\hat{J}_z, \hat{J}_{(-)}] = -\hbar \hat{J}_{(-)} \ , \tag{5.5b}$$

$$[\hat{J}_{(+)}, \hat{J}_{(-)}] = 2\hbar \hat{J}_z \ , \tag{5.5c}$$

as well as the following ones

$$\hat{J}_{(+)}\hat{J}_{(-)} = \hat{J}^\dagger_{(-)}\hat{J}_{(-)} = \hat{J}^2 - \hat{J}_z(\hat{J}_z - \hbar) \ , \tag{5.6a}$$

$$\hat{J}_{(-)}\hat{J}_{(+)} = \hat{J}^\dagger_{(+)}\hat{J}_{(+)} = \hat{J}^2 - \hat{J}_z(\hat{J}_z + \hbar) \ . \tag{5.6b}$$

Homework 5.2. Check Eqs. (5.5)–(5.6).

Let us now apply Eqn. (5.6a) on a *given* eigenstate $|\lambda, m\rangle$ satisfying the eigenvalue condition (5.4):

$$\hat{J}_{(+)}\hat{J}_{(-)} \left|\lambda, m\right\rangle = (\hat{J}^2 - \hat{J}_z^2 + \hbar\hat{J}_z) \left|\lambda, m\right\rangle = \hbar^2[\lambda - m(m-1)] \left|\lambda, m\right\rangle \ . \tag{5.7a}$$

Analogously, applying Eqn. (5.6b) on $|\lambda, m\rangle$ yields:

$$\hat{J}_{(-)}\hat{J}_{(+)}\,|\lambda, m\rangle = (\hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z)\,|\lambda, m\rangle = \hbar^2[\lambda - m(m+1)]\,|\lambda, m\rangle \ . \tag{5.7b}$$

Applying instead Eqs. (5.5) on $|\lambda, m\rangle$ yields:

$$\hat{J}_z\hat{J}_{(+)}\,|\lambda, m\rangle = (\hat{J}_{(+)}\hat{J}_z + \hbar\hat{J}_{(+)})\,|\lambda, m\rangle = \hbar(m+1)\hat{J}_{(+)}\,|\lambda, m\rangle \ , \tag{5.8a}$$

$$\hat{J}_z\hat{J}_{(-)}\,|\lambda, m\rangle = (\hat{J}_{(-)}\hat{J}_z - \hbar\hat{J}_{(-)})\,|\lambda, m\rangle = \hbar(m-1)\hat{J}_{(-)}\,|\lambda, m\rangle \ . \tag{5.8b}$$

Together, these relations tell us that if $|\lambda, m\rangle$ is eigenstate of $\hat{J}_z$ with eigenvalue m, then $\hat{J}_{(+)}\,|\lambda, m\rangle$ and $\hat{J}_{(-)}\,|\lambda, m\rangle$ are eigenstates of $\hat{J}_z$ too with eigenvalues $\hbar(m+1)$ and $\hbar(m-1)$ respectively. However, in deriving this conclusion the following fact should be stressed. Let us consider the norms of $\hat{J}_{(+)}\,|\lambda, m\rangle$ and $\hat{J}_{(-)}\,|\lambda, m\rangle$; these can be calculated directly by taking the scalar product of Eqs. (5.7) above with $\langle\lambda, m|$. Assuming as usual that $|\lambda, m\rangle$ is correctly normalized, the result is

$$\left\|\hat{J}_{(+)}\,|\lambda, m\rangle\right\|^2 = \hbar^2[\lambda - m(m+1)] \ , \tag{5.9a}$$

$$\left\|\hat{J}_{(-)}\,|\lambda, m\rangle\right\|^2 = \hbar^2[\lambda - m(m-1)] \ . \tag{5.9b}$$

Homework 5.3. Prove Eqs. (5.9).

Clearly, the terms on the right-hand sides of Eqs. (5.9) above must be non-negative real numbers. In fact, by definition the norm of a vector is a non-negative real number, moreover here we are considering the *square* of a norm. This imposes a constraint on the values that λ and m can take. To explain this point, let us generalize Eqs. (5.7)-(5.9) to the case in which $\hat{J}_{(+)}$ or $\hat{J}_{(-)}$ are applied multiple times on $|\lambda, m\rangle$. By induction, one can prove that the following relations hold:

$$\hat{J}_z\left(\hat{J}^p_{(+)}\,|\lambda, m\rangle\right) = \hbar(m+p)\left(\hat{J}^p_{(+)}\,|\lambda, m\rangle\right) \ , \tag{5.10a}$$

$$\hat{J}_z\left(\hat{J}^q_{(-)}\,|\lambda, m\rangle\right) = \hbar(m-q)\left(\hat{J}^q_{(-)}\,|\lambda, m\rangle\right) \ , \tag{5.10b}$$

where $p, q = 1, 2, \ldots$ are the number of times the operators $\hat{J}_{(+)}$ and $\hat{J}_{(-)}$ respectively have been applied on $|\lambda, m\rangle$. Each time $\hat{J}_{(+)}$ (respectively, $\hat{J}_{(-)}$) is applied on $|\lambda, m\rangle$ the corresponding eigenvalue of $\hat{J}_z$ increases or decreases respectively of one unit.

Let us now fix a particular value of m, say m_0. By applying many times the operators $\hat{J}_{(+)}$ and $\hat{J}_{(-)}$ on $|\lambda, m_0\rangle$ one can generate the whole series of eigenstates of the type $|\lambda, m\rangle$:

$$\ldots, \hat{J}^2_{(-)} |\lambda, m_0\rangle , \hat{J}_{(-)} |\lambda, m_0\rangle , |\lambda, m_0\rangle , \hat{J}_{(+)} |\lambda, m_0\rangle , \hat{J}^2_{(+)} |\lambda, m_0\rangle , \ldots$$

Of course, in general neither $\hat{J}^p_{(+)} |\lambda, m_0\rangle$ nor $\hat{J}^q_{(-)} |\lambda, m_0\rangle$ is guaranteed to be correctly normalized even assuming that $|\lambda, m_0\rangle$ is normalized. In fact, we have seen that the contrary happens in the case $p = 1$, see Eqs. (5.9). The associated *normalized* eigenstates are indicated as:

$$\ldots, |\lambda, m_0 - 2\rangle , |\lambda, m_0 - 1\rangle , |\lambda, m_0\rangle , |\lambda, m_0 + 1\rangle , |\lambda, m_0 + 2\rangle , \ldots$$

and the corresponding eigenvalues for $\hat{J}_z$ are given by

$$\ldots, \hbar(m_0 - 2), \hbar(m_0 - 1), \hbar m_0, \hbar(m_0 + 1), \hbar(m_0 + 2), \ldots$$

The chain must stop in both directions. Let us now explain why. Eqs. (5.7) can be generalized as follows:

$$\begin{aligned}
\hat{J}_{(+)}\hat{J}_{(-)} \left(\hat{J}^q_{(-)} |\lambda, m_0\rangle \right) &= \hbar^2[\lambda - (m_0 - q)(m_0 - q - 1)] \left(\hat{J}^q_{(-)} |\lambda, m_0\rangle \right) , \\
\hat{J}_{(-)}\hat{J}_{(+)} \left(\hat{J}^p_{(+)} |\lambda, m_0\rangle \right) &= \hbar^2[\lambda - (m_0 + p)(m_0 + p + 1)] \left(\hat{J}^p_{(+)} |\lambda, m_0\rangle \right) ,
\end{aligned}$$

and Eqs. (5.9) become

$$\begin{aligned}
\left\| \hat{J}_{(+)} |\lambda, m_0 + p\rangle \right\|^2 &= \langle \lambda, m_0 + p | \hat{J}_{(-)}\hat{J}_{(+)} |\lambda, m_0 + p\rangle \\
&= \hbar^2[\lambda - (m_0 + p)(m_0 + p + 1)] , \qquad (5.11a) \\
\left\| \hat{J}_{(-)} |\lambda, m_0 - q\rangle \right\|^2 &= \langle \lambda, m - q | \hat{J}_{(+)}\hat{J}_{(-)} |\lambda, m_0 - q\rangle \\
&= \hbar^2[\lambda - (m_0 - q)(m_0 - q - 1)] , \qquad (5.11b)
\end{aligned}$$

or, stated in another way,

$$\left\| \hat{J}^{p+1}_{(+)} \left| \lambda, m_0 \right\rangle \right\|^2 = \hbar^2 [\lambda - (m_0 + p)(m_0 + p + 1)] \left\| \hat{J}^{p}_{(+)} \left| \lambda, m_0 \right\rangle \right\|^2 ,$$
$$\left\| \hat{J}^{q+1}_{(-)} \left| \lambda, m_0 \right\rangle \right\|^2 = \hbar^2 [\lambda - (m_0 - q)(m_0 - q - 1)] \left\| \hat{J}^{q}_{(-)} \left| \lambda, m_0 \right\rangle \right\|^2 .$$

Homework 5.4. Verify Eqs. (5.10)–(5.11).

Since the (square of) the norm of a vector in a Hilbert space must be (by definition) a non-negative real quantity, both quantities appearing on the right-hand side of Eqs. (5.11)

$$\lambda - (m_0 + p)(m_0 + p + 1) \qquad \text{and} \qquad \lambda - (m_0 - q)(m_0 - q - 1) \tag{5.12}$$

must be non-negative. Remember that p and q are integer because they indicates the number of times we apply the $\hat{J}_{(+)}$ and $\hat{J}_{(-)}$. Moreover, both p and q can be made arbitrary large since we can apply $\hat{J}_{(+)}$ or $\hat{J}_{(-)}$ on $|\lambda, m_0\rangle$ as many times as we want, and therefore for any given λ and m_0 it is always possible to choose two positive integer numbers p and q large enough to let these quantities in (5.12) to become negative. The only way out is to admit the existence of two (non-negative) integer numbers q_0 and p_0 such that the chain stops, that is,

$$\hat{J}^{p_0}_{(+)} \left| \lambda, m_0 \right\rangle \neq 0 , \qquad \hat{J}^{p_0+1}_{(+)} \left| \lambda, m_0 \right\rangle = 0 ,$$
$$\hat{J}^{q_0}_{(-)} \left| \lambda, m_0 \right\rangle \neq 0 , \qquad \hat{J}^{q_0+1}_{(-)} \left| \lambda, m_0 \right\rangle = 0 .$$

This implies that

$$\hat{J}_{(-)}\hat{J}_{(+)} \left(\hat{J}^{p_0}_{(+)} \left| \lambda, m_0 \right\rangle \right) = \hat{J}_{(-)} \left(\hat{J}^{p_0+1}_{(+)} \left| \lambda, m_0 \right\rangle \right) = 0 ,$$
$$\hat{J}_{(+)}\hat{J}_{(-)} \left(\hat{J}^{q_0}_{(-)} \left| \lambda, m_0 \right\rangle \right) = \hat{J}_{(+)} \left(\hat{J}^{q_0+1}_{(-)} \left| \lambda, m_0 \right\rangle \right) = 0 ,$$

from which it follows that at some point the expression in (5.12) must become zero, i.e.:

$$\lambda = (m_0 + p_0)(m_0 + p_0 + 1) , \tag{5.13a}$$

$$\lambda = (m_0 - q_0)(m_0 - q_0 - 1) . \tag{5.13b}$$

As the left-hand sides are equal, we get that:

$$(m_0 + p_0)(m_0 + p_0 + 1) = (m_0 - q_0)(m_0 - q_0 - 1) \ .$$

Performing the calculation, this relation leads to

$$m_0^2 + m_0 p_0 + m_0 p_0 + p_0^2 + m_0 + p_0 = m_0^2 + q_0^2 - 2m_0 q_0 - m_0 + q_0 \ ,$$

and thus

$$m_0 = \frac{q_0 - p_0}{2} \ ,$$

where q_0 and p_0 are by construction two non-negative integer numbers. From here we get that m_0 must be either integer or half-integer. Now, if you define

$$j \equiv \frac{q_0 + p_0}{2} \ ,$$

we get from one of the Eqs. (5.13) that

$$\lambda = j(j+1) \ .$$

This means the eigenvalues of $\hat{J}^2$ are of the form $j(j+1)\hbar^2$, where $j = (q_0 + p_0)/2$ is an integer or half-integer: $j = 0, 1/2, 1, 3/2, \ldots$ Note also that for a given j the smallest eigenvalue of $\hat{J}_z$ is given by $m_{\min}\hbar$, where

$$m_{\min} = m_0 - q_0 = \left(\frac{q_0 - p_0}{2}\right) - q_0 = -j \ ,$$

and analogously the largest eigenvalue of $\hat{J}_z$ is $m_{\max}\hbar$, where

$$m_{\max} = m_0 + p_0 = \left(\frac{q_0 - p_0}{2}\right) + p_0 = j \ .$$

Therefore, the eigenvalues of $\hat{J}_z$ are of the form $m\hbar$ where

$$m = \underbrace{-j, -j+1, -j+2, \ldots, j-2, j-1, j}_{(2j+1) \text{ terms.}} \ .$$

The fact that m and j can take only this discrete set of values is known as **angular momentum quantization**.

From now on we shall denote the eigenstates of $\hat{J}^2$ and $\hat{J}_z$ with $|j,m\rangle$:

$$\hat{J}^2 |j,m\rangle = \hbar^2 j(j+1) |j,m\rangle \ ,$$
$$\hat{J}_z |j,m\rangle = \hbar m |j,m\rangle \ .$$

Once j is fixed (do remember that j must be integer or half-integer) the set of all $2j+1$ eigenstates $|j,m\rangle$, where $m = -j, -j+1, \ldots, j-1, j$, makes up what is called a **multiplet** with fixed j. We can move within the multiplet leaving unchanged the eigenvalue of $\hat{J}^2$ and changing only the eigenvalue of $\hat{J}_z$ one unit at a time (in units of $\hbar$) by using the operators $\hat{J}_{(+)}$ and $\hat{J}_{(-)}$. For example, we can start from $|j,j\rangle$ (with the highest eigenvalue $m = j$ of $\hat{J}_z$) and move downwards by applying recursively $\hat{J}_{(-)}$ until the eigenstate $|j,-j\rangle$ is reached. Basically , we can construct the general state $|j,m\rangle$ as

$$|j,m\rangle = \tilde{c}^{(-)}_{j,m} \hat{J}^{j-m}_{(-)} |j,j\rangle \ .$$

The coefficient $\tilde{c}^{(-)}_{j,m}$ accounts for the normalization. We can also start from $|j,-j\rangle$ and move upwards through successive applications of $\hat{J}_{(+)}$:

$$|j,m\rangle = \tilde{c}^{(+)}_{j,m} \hat{J}^{j+m}_{(+)} |j,-j\rangle \ .$$

Let us evaluate the coefficients $\tilde{c}^{(-)}_{j,m}$ and $\tilde{c}^{(+)}_{j,m}$. The simplest case is

$$\hat{J}_{(-)} |j,m\rangle = c^{(-)}_{j,m} |j,m-1\rangle \ , \quad \hat{J}_{(+)} |j,m\rangle = c^{(+)}_{j,m} |j,m+1\rangle \ .$$

We have already calculated the norm of $\hat{J}_{(+)} |j,m\rangle$ and $\hat{J}_{(-)} |j,m\rangle$. By using Eqs. (5.9), and making the choice of having real and positive coefficients, we find

$$c^{(-)}_{j,m} = \hbar\sqrt{(j+m)(j-m+1)} \ ,$$
$$c^{(+)}_{j,m} = \hbar\sqrt{(j-m)(j+m+1)} \ .$$

where we have made the choice for the coefficients $c^{(\pm)}_{j,m}$ to be real and positive. This choice will be crucial to determine the relative phases among states of the same

multiplets. Thus,

$$\hat{J}_{(-)}\,|j,m\rangle = \hbar\sqrt{(j+m)(j-m+1)}\,|j,m-1\rangle \ , \tag{5.14a}$$

$$\hat{J}_{(+)}\,|j,m\rangle = \hbar\sqrt{(j-m)(j+m+1)}\,|j,m+1\rangle \ . \tag{5.14b}$$

By iterating the same procedure, we get the general formulas

$$|j,m\rangle = \frac{1}{\hbar^{j-m}}\frac{1}{\sqrt{(2j)!}}\sqrt{\frac{(j+m)!}{(j-m)!}}\hat{J}_{(-)}^{j-m}\,|j,j\rangle \ , \tag{5.15a}$$

$$|j,m\rangle = \frac{1}{\hbar^{j+m}}\frac{1}{\sqrt{(2j)!}}\sqrt{\frac{(j-m)!}{(j+m)!}}\hat{J}_{(+)}^{j+m}\,|j,-j\rangle \ . \tag{5.15b}$$

Homework 5.5. Prove Eqs. (5.15).

From Eqs. (5.14) we can derive the matrix elements

$$\langle j,m'|\hat{J}_{(-)}|j,m\rangle = \hbar\sqrt{(j+m)(j-m+1)}\delta_{m',m-1} \ , \tag{5.16a}$$

$$\langle j,m'|\hat{J}_{(+)}|j,m\rangle = \hbar\sqrt{(j-m)(j+m+1)}\delta_{m',m+1} \ . \tag{5.16b}$$

So the matrix elements of $\hat{J}_{(-)}$ and $\hat{J}_{(+)}$ are different from zero only for the elements next to the main diagonal. From the matrix elements of $\hat{J}_{(-)}$ and $\hat{J}_{(+)}$ it is possible to calculate also the matrix elements of $\hat{J}_x$ and $\hat{J}_y$ since

$$\hat{J}_x = \frac{1}{2}(\hat{J}_{(+)} + \hat{J}_{(-)}) \ , \quad \hat{J}_y = \frac{1}{2i}(\hat{J}_{(+)} - \hat{J}_{(-)}) \ .$$

The matrix elements of $\hat{J}_z$ can be easily calculated, because $\hat{J}_z$ is diagonal in the representation of its eigenstates:

$$\langle j,m'|\hat{J}_z|j,m\rangle = \hbar m\delta_{m,m'} \ .$$

The explicit matrix representations of $\hat{J}_x$, $\hat{J}_y$ and $\hat{J}_z$ are for $j = 1/2$:

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

for $j = 1$:

$$\hat{J}_x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} , \quad \hat{J}_y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} , \quad \hat{J}_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} ;$$

for $j = 3/2$:

$$\hat{J}_x = \frac{\hbar}{2} \begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & 2 & 0 \\ 0 & 2 & 0 & \sqrt{3} \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix} , \quad \hat{J}_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i\sqrt{3} & 0 & 0 \\ i\sqrt{3} & 0 & -2i & 0 \\ 0 & 2i & 0 & -i\sqrt{3} \\ 0 & 0 & i\sqrt{3} & 0 \end{pmatrix} ,$$

and

$$\hat{J}_z = \frac{\hbar}{2} \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix} .$$

5.2 Spherical Harmonics

We now come back to the specific case of *orbital* angular momentum. We shall employ the notation $\hat{L}$ instead of $\hat{J}$ when referring to the orbital angular momentum, and j will be replaced by l. In the last section we have said that j is integer or half-integer. As we shall discuss in this section, l and therefore m must be *integer* numbers for the orbital angular momentum.

The first step is to look for a representation of the abstract eigenstates $|l, m\rangle$ of $\hat{L}^2$ and $\hat{L}_z$. This is somewhat analogous to the case in which the eigenstates of the Hamiltonian $\hat{H}$, say $|E\rangle$, were represented on the basis of the position eigenstates: $\langle x|E\rangle = \psi_E(x)$. In that case, $\psi_E(x)$ were the eigenfunctions of the Schrödinger operator in the position representation, and were found to be square integrable functions of the spatial coordinate x. Here we shall do the same thing with the eigenstates

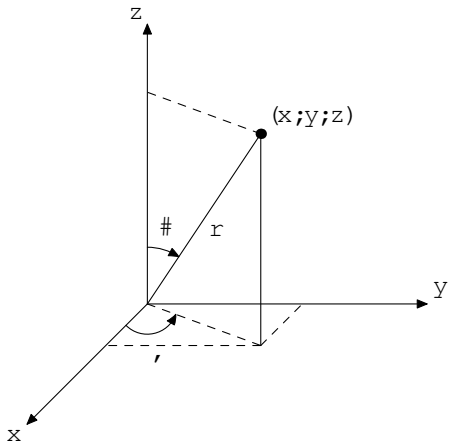


FIG. 5.1. Polar coordinates.

$|l, m\rangle$ but the basis is not that of the Cartesian coordinates $|x, y, z\rangle$ but that of the polar coordinates $|r, \vartheta, \varphi\rangle$:

$$\langle r, \vartheta, \varphi | l, m \rangle = Y_{l,m}(\vartheta, \varphi) \ .$$

As we shall see, $Y_{l,m}$ are functions not depending explicitly on r and are usually indicated with the symbol $Y_{l,m}$; they are referred to as **spherical harmonics**. The next goal is to calculate these functions.

First of all, we recall the usual relation between Cartesian and polar coordinates (see Fig. 5.1)

$$x = r \cos\varphi \sin\vartheta \ , \tag{5.17a}$$

$$y = r \sin\varphi \sin\vartheta \ , \tag{5.17b}$$

$$z = r \cos\vartheta \ . \tag{5.17c}$$

The three components of the angular momentum can also be rewritten in terms of polar coordinates. For example, the z-component of the angular momentum reads

$$\hat{L}_z = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \ ;$$

x and y can be replaced directly by using Eqs. (5.17) while the derivatives $\partial/\partial x$ and $\partial/\partial y$ can be rewritten in polar coordinates by using the chain rule of differentiation:

$$\frac{\partial}{\partial x} = \frac{\partial \vartheta}{\partial x}\frac{\partial}{\partial \vartheta} + \frac{\partial \varphi}{\partial x}\frac{\partial}{\partial \varphi} + \frac{\partial r}{\partial x}\frac{\partial}{\partial r} ,$$

and analogously for the other derivatives. In this way one finds

$$\begin{aligned}
\hat{L}_x &= i\hbar\left(\sin\varphi\frac{\partial}{\partial\vartheta} + \cot\vartheta\cos\varphi\frac{\partial}{\partial\varphi}\right) , \\
\hat{L}_y &= i\hbar\left(-\cos\varphi\frac{\partial}{\partial\vartheta} + \cot\vartheta\sin\varphi\frac{\partial}{\partial\varphi}\right) , \\
\hat{L}_z &= -i\hbar\frac{\partial}{\partial\varphi} .
\end{aligned}$$

Homework 5.6. Derive the expressions above for the components of the orbital angular momentum in polar coordinates.

In the representation of $|r, \vartheta, \varphi\rangle$ the eigenvalue problem $\hat{L}_z\,|l, m\rangle = m\hbar\,|l, m\rangle$ becomes

$$\underbrace{-i\hbar\frac{\partial}{\partial\varphi}}_{\hat{L}_z} Y_{l,m}(\vartheta, \varphi) = m\hbar Y_{l,m}(\vartheta, \varphi) . \tag{5.18}$$

Homework 5.7. Try to derive Eqn. (5.18). Bring in mind that the resolution of the identity formula in polar coordinates reads

$$1 = \int |\vartheta, \varphi\rangle\,\langle\vartheta, \varphi|\sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\varphi .$$

The latter relation follows from the form of the integration volume

$$\int \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int r^2\,\mathrm{d}r \int \sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\varphi .$$

The solution of the eigenvalue problem (5.18) is given by

$$Y_{l,m}(\vartheta, \varphi) = \frac{\mathrm{e}^{im\varphi}}{\sqrt{2\pi}}\Theta_l^m(\vartheta) , \tag{5.19}$$

where $\Theta_l^m(\vartheta)$ is a function to be determined which depends only on ϑ while the whole dependence on φ is included in the factor $\mathrm{e}^{im\varphi}$. We have put a factor $1/\sqrt{2\pi}$ as we did for the normalization of plane wave functions.

In order to determine the function $\Theta_l^m(\vartheta)$ we need to express the operators $\hat{L}_{(+)}$ and $\hat{L}_{(-)}$ in polar coordinates:

$$\hat{L}_{(\pm)} = \hat{L}_x \pm i\hat{L}_y = \hbar\, \mathrm{e}^{\pm i\varphi} \left(\pm \frac{\partial}{\partial \vartheta} + i \cot\vartheta \frac{\partial}{\partial \varphi} \right) . \tag{5.20}$$

We recall from the previous section that

$$\hat{L}_{(+)} \, |l, l\rangle = 0 \, ,$$

which now becomes

$$\hat{L}_{(+)} Y_{l,l}(\vartheta, \varphi) = 0 \, ,$$

and using Eqn. (5.19)

$$\left(\frac{\mathrm{d}}{\mathrm{d}\vartheta} - l \frac{\cos\vartheta}{\sin\vartheta} \right) \Theta_l^l(\vartheta) = 0 \, ,$$

we get

$$\frac{\mathrm{d}}{\mathrm{d}\vartheta} \left(\frac{1}{\sin^l \vartheta} \Theta_l^l(\vartheta) \right) = 0 \, .$$

Its solution is given by

$$\Theta_l^l(\vartheta) = k \sin^l \vartheta \, .$$

The value of k is fixed by requiring $\Theta_l^l(\vartheta)$ to be normalized to 1:

$$\int_0^\pi \left| \Theta_l^l(\vartheta) \right|^2 \sin\vartheta \, \mathrm{d}\vartheta = 1 \, .$$

The result is

$$\Theta_l^l(\vartheta) = (-1)^l \sqrt{\frac{(2l+1)!}{2}} \frac{1}{2^l l!} \sin^l \vartheta \, .$$

By repeatedly using Eqn. (5.20) we can obtain a general expression for $\Theta_l^m(\vartheta)$:

$$\Theta_l^m(\vartheta) = \frac{1}{\hbar^{l-m}} \frac{1}{\sqrt{2l!}} \sqrt{\frac{(l+m)!}{(l-m)!}} \hat{L}_{(-)}^{l-m} \Theta_l^l(\vartheta) \, .$$

The final result is

$$Y_{l,m}(\vartheta, \varphi) = (-1)^l \sqrt{\frac{2l+1}{4\pi}} \sqrt{\frac{(l+m)!}{(l-m)!}} \, \mathrm{e}^{im\varphi} \sin^{-m}\vartheta \frac{\mathrm{d}^{l-m}}{\mathrm{d}(\cos\vartheta)^{l-m}} (\sin\vartheta)^{2l} \, . \tag{5.21}$$

Remark: sometimes $Y_{l,m}(\vartheta,\varphi)$ are denoted as $Y_l^m(\vartheta,\varphi)$. The spherical harmonics can be written in a more compact way in terms of associated Legendre functions $P_l^m(\cos\vartheta)$.

The abstract states $|l,m\rangle$ form a complete orthonormal basis as they are eigenstates of the Hermitian operators $\hat{L}^2$ and $\hat{L}_z$

$$\hat{L}^2\,|l,m\rangle = \hbar^2 l(l+1)\,|l,m\rangle$$
$$\hat{L}_z\,|l,m\rangle = \hbar m\,|l,m\rangle\,.$$

Thus also their representation (spherical harmonics) makes up a basis on which we can expand every function of ϑ and φ, *i.e.*:

$$\psi(\vartheta,\varphi) = \sum_{l,m} c_{l,m} Y_{l,m}(\vartheta,\varphi)\,.$$

For a given l we know that

$$\langle l,m|l,m'\rangle = \delta_{m,m'}\,,$$

and this relation turns into

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \mathrm{d}\vartheta\, \sin\vartheta\, Y^*_{l,m}(\vartheta,\varphi) Y_{l,m'}(\vartheta,\varphi) = \delta_{m,m'}$$

or more generally

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \mathrm{d}\vartheta\, \sin\vartheta\, Y^*_{l,m}(\vartheta,\varphi) Y_{l',m'}(\vartheta,\varphi) = \delta_{m,m'}\delta_{l,l'}\,.$$

Let us now consider in some detail the case of a plane wave $\exp(i\,\boldsymbol{k}\cdot\boldsymbol{r})$ in three dimensions and find its expansion on spherical harmonics. As $\boldsymbol{k}\cdot\boldsymbol{r}$ is a scalar product, it is invariant under rotations and there is no loss of generality if we restrict ourselves to the case in which the vector $\boldsymbol{k}$ lies along the z-axis:

$$\mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}} = \mathrm{e}^{ikr\cos\vartheta}\,.$$

Expanding on spherical harmonics yields an expression of the form

$$\mathrm{e}^{ikr\cos\vartheta} = \sum_{l,m} c_{l,m}(r) Y_{l,m}(\vartheta,\varphi)\,.$$

In general, the expansion coefficients $c_{l,m}$ depend on r. In the case of plane waves, the left-hand side in the previous equation does not depend on φ, that implies that the expansion can be rewritten in terms of $\Theta_{l,m}(\vartheta)$ only, that is,

$$\mathrm{e}^{ikr\cos\vartheta} = \sum_{l,m} \tilde{c}_{l,m}(r)\Theta_{l,m}(\vartheta) \ .$$

The coefficients of the expansion depend on l only and can be written in terms of the so-called spherical Bessel functions J_l:

$$\tilde{c}_{l,m}(r) \propto J_l(kr) \ .$$

As $\exp(ikr\cos\vartheta)$ depends on ϑ via $\cos\vartheta$, the $\Theta_{l,m}(\vartheta)$ which depend on ϑ via $\cos\vartheta$ are only those having $m = 0$, therefore

$$\mathrm{e}^{ikr\cos\vartheta} = \sum_{l} c_l(kr)\Theta_{l,0}(\vartheta) \ , \quad c_l(kr) \propto J_l(kr) \ .$$

This is the spherical harmonics expansion of the plane wave.

5.3 Addition of angular momenta

Let us now consider two systems with given angular momenta and we want to find the angular momentum of the composite system. The results of this section are not restricted to the case of orbital angular momentum $\hat{L}_i$ but to any operator $\hat{J}_i$ satisfying the algebra (5.3).

It should be stressed that here we are not dealing with the problem of having to sum two numbers, instead we have to sum two operators. For the i-components this reads

$$\hat{J}_i^{(T)} = \hat{J}_i^{(1)} + \hat{J}_i^{(2)} \ .$$

Hereafter T denotes the total system $(1 + 2)$ and the labels (1) and (2) indicates the two systems.

Being an operator, $\hat{J}^{(T)}$ has its own spectrum of eigenvalues. What we shall prove in this section is that the possible eigenvalues of $(\hat{J}^{(T)})^2$ are given by one among the set of values

$$|j^{(1)} - j^{(2)}|, \ldots, (j^{(1)} + j^{(2)}) ,$$

where $j^{(1)}$ and $j^{(2)}$ are the eigenvalues of $(\hat{J}^{(1)})^2$ and $(\hat{J}^{(2)})^2$ respectively, that is

$$(\hat{J}^{(1)})^2 \,|\,\rangle = j^{(1)}(j^{(1)}+1)\hbar^2 \,|\,\rangle \ .$$

and an analogous relation holds for $(\hat{J}^{(2)})^2$ and $(\hat{J}^{(T)})^2$.

(1) and (2) refer to two different systems, so the angular momentum components of the first system do commute with each component of the second system:

$$[\hat{J}_i^{(1)}, \hat{J}_j^{(2)}] = 0 \ .$$

From this relation it is easy to prove that the following four operators commute with each other:

$$(\hat{J}^{(1)})^2, \quad (\hat{J}^{(2)})^2, \quad \hat{J}_z^{(1)}, \quad \hat{J}_z^{(2)} \ , \tag{5.22}$$

and the same is true also for the following set of four operators

$$(\hat{J}^{(1)})^2, \quad (\hat{J}^{(2)})^2, \quad (\hat{J}^{(T)})^2, \quad \hat{J}_z^{(T)} \ . \tag{5.23}$$

Homework 5.8. Prove that the operators (5.23) commute with each other.

Let us diagonalize simultaneously the set (5.22). The eigenvalue problem reads

$$\begin{aligned}
(\hat{J}^{(1)})^2 \,|j^{(1)}, j^{(2)}, m^{(1)}, m^{(2)}\rangle &= \hbar^2 j^{(1)}(j^{(1)}+1) \,|j^{(1)}, j^{(2)}, m^{(1)}, m^{(2)}\rangle \ , \\
(\hat{J}^{(2)})^2 \,|j^{(1)}, j^{(2)}, m^{(1)}, m^{(2)}\rangle &= \hbar^2 j^{(2)}(j^{(2)}+1) \,|j^{(1)}, j^{(2)}, m^{(1)}, m^{(2)}\rangle \ , \\
\hat{J}_z^{(1)} \,|j^{(1)}, j^{(2)}, m^{(1)}, m^{(2)}\rangle &= \hbar m^{(1)} \qquad\quad |j^{(1)}, j^{(2)}, m^{(1)}, m^{(2)}\rangle \ , \\
\hat{J}_z^{(2)} \,|j^{(1)}, j^{(2)}, m^{(1)}, m^{(2)}\rangle &= \hbar m^{(2)} \qquad\quad |j^{(1)}, j^{(2)}, m^{(1)}, m^{(2)}\rangle \ .
\end{aligned}$$

If instead we diagonalize the second set of operators, namely the set (5.23), we get

$$\begin{aligned}
(\hat{J}^{(1)})^2 \,|j^{(1)}, j^{(2)}, J^{(T)}, M^{(T)}\rangle &= \hbar^2 j^{(1)}(j^{(1)}+1) \quad |j^{(1)}, j^{(2)}, J^{(T)}, M^{(T)}\rangle \ , \\
(\hat{J}^{(2)})^2 \,|j^{(1)}, j^{(2)}, J^{(T)}, M^{(T)}\rangle &= \hbar^2 j^{(2)}(j^{(2)}+1) \quad |j^{(1)}, j^{(2)}, J^{(T)}, M^{(T)}\rangle \ , \\
(\hat{J}^{(T)})^2 \,|j^{(1)}, j^{(2)}, J^{(T)}, M^{(T)}\rangle &= \hbar^2 J^{(T)}(J^{(T)}+1) \,|j^{(1)}, j^{(2)}, J^{(T)}, M^{(T)}\rangle \ , \\
\hat{J}_z^{(T)} \,|j^{(1)}, j^{(2)}, J^{(T)}, M^{(T)}\rangle &= \hbar M^{(T)} \qquad\qquad\quad |j^{(1)}, j^{(2)}, J^{(T)}, M^{(T)}\rangle
\end{aligned}$$

The space of the states $|j^{(1)}, j^{(2)}, m^{(1)}, m^{(2)}\rangle$ has dimension

$$(2j^{(2)}+1)(2j^{(1)}+1)$$

once $j^{(1)}$ and $j^{(2)}$ have been fixed. Now, it is easy to prove (homework) that $(\hat{J}^{(1)})^2$ and $(\hat{J}^{(2)})^2$ commute with $\hat{J}_i^{(T)}$ (the components of the total angular momentum operator $\hat{J}^{(T)}$). Therefore we can look for the eigenstates of $(\hat{J}^{(T)})^2$ and $\hat{J}_z^{(T)}$ among the eigenstates of $(\hat{J}^{(1)})^2$ and $(\hat{J}^{(2)})^2$ at fixed $j^{(1)}$ and $j^{(2)}$. This is because $(\hat{J}^{(T)})^2$ and $\hat{J}_z^{(T)}$ are made up of $\hat{J}_i^{(T)}$.

In order to simplify the notation we will denote the common eigenstates of $(\hat{J}^{(1)})^2$ and $(\hat{J}^{(2)})^2$ at fixed $j^{(1)}$ and $j^{(2)}$ with $|m^{(1)}, m^{(2)}\rangle$, i.e. the states $|j^{(1)}, j^{(2)}, m^{(1)}, m^{(2)}\rangle$ will be denoted by $|m^{(1)}, m^{(2)}\rangle$ once $j^{(1)}$ and $j^{(2)}$ have been fixed. Similarly , the states $|j^{(1)}, j^{(2)}, J^{(T)}, M^{(T)}\rangle$ will be indicated by $|J^{(T)}, M^{(T)}\rangle$.

Let us try to write $|J^{(T)}, M^{(T)}\rangle$ in the basis of $|m^{(1)}, m^{(2)}\rangle$. In this way we shall obtain the possible values of $J^{(T)}$ and $M^{(T)}$ once $j^{(1)}$ and $j^{(2)}$ are fixed. The solution of this problem is based on the following two observations, namely

1. Every vector $|m^{(1)}, m^{(2)}\rangle$ is an eigenstate of $\hat{J}_z^{(T)}$ with corresponding eigenvalue

$$M^{(T)} = m^{(1)} + m^{(2)} \ ,$$

that is

$$\hat{J}_z^{(T)} \,|m^{(1)}, m^{(2)}\rangle = \hbar(m^{(1)} + m^{(2)}) \,|m^{(1)}, m^{(2)}\rangle \ .$$

This is because

$$\hat{J}_z^{(T)} = \hat{J}_z^{(1)} + \hat{J}_z^{(2)} \ .$$

2. For every value of $J^{(T)}$ there are various $M^{(T)}$ and therefore there is a certain number, say $N(J^{(T)})$, of vectors in the space of $|m^{(1)}, m^{(2)}\rangle$. Every set contains all the eigenvalues $M^{(T)}$ from $-J^{(T)}$ to $J^{(T)}$.

   We shall instead denote with $n(M^{(T)})$ the number of states $|m^{(1)}, m^{(2)}\rangle$ which have the same $M^{(T)}$.

It is easy to prove that the relation among $n(M^{(T)})$ and $N(J^{(T)})$ is

$$n(M^{(T)}) = \sum_{J^{(T)} \geq |M^{(T)}|} N(J^{(T)}) \ , \tag{5.24}$$

because if $J^{(T)}$ is larger than $M^{(T)}$ then surely it has some $(M')^{(T)}$ which is equal to $M^{(T)}$, $(M')^{(T)}$ being in the range from $-J^{(T)}$ to $J^{(T)}$.

From Eqn. (5.24) it follows that

$$N(J^{(T)}) = n(J^{(T)}) - n(J^{(T)} + 1) \ . \tag{5.25}$$

In fact, $n(J^{(T)})$ and $n(J^{(T)} + 1)$ differ only for one term, namely that of $J^{(T)}$.

From Eqn. (5.25) one sees that in order to determine $N(J^{(T)})$ it is sufficient to determine $n(M^{(T)})$. Now, $n(M^{(T)})$ is the number of pairs $(m^{(1)}, m^{(2)})$ such that

$$M^{(T)} = m^{(1)} + m^{(2)} \ .$$

To find this number one can employ a diagram like the one plotted in Fig. 5.2 for the special case $j^{(1)} = 7/2$ and $j^{(2)} = 2$. $n(M^{(T)})$ is the number of pairs $(m^{(1)}, m^{(2)})$ in that plot that are situated along the line of equation $M^{(T)} = m^{(1)} + m^{(2)}$. If $j^{(1)} > j^{(2)}$ we get from Fig. 5.2

$$n(M^{(T)}) = \begin{cases} 0 & \text{if} \quad |M^{(T)}| > j^{(1)} + j^{(2)} \\ j^{(1)} + j^{(2)} + 1 - M^{(T)} & \text{if} \quad j^{(1)} + j^{(2)} \geq |M^{(T)}| \geq |j^{(1)} - j^{(2)}| \\ 2j^{(2)} + 1 & \text{if} \quad |j^{(1)} - j^{(2)}| \geq M^{(T)} \geq 0 \end{cases}$$

If we put these numbers into Eqn. (5.25) we obtain

$$N(J^{(T)}) = 1$$

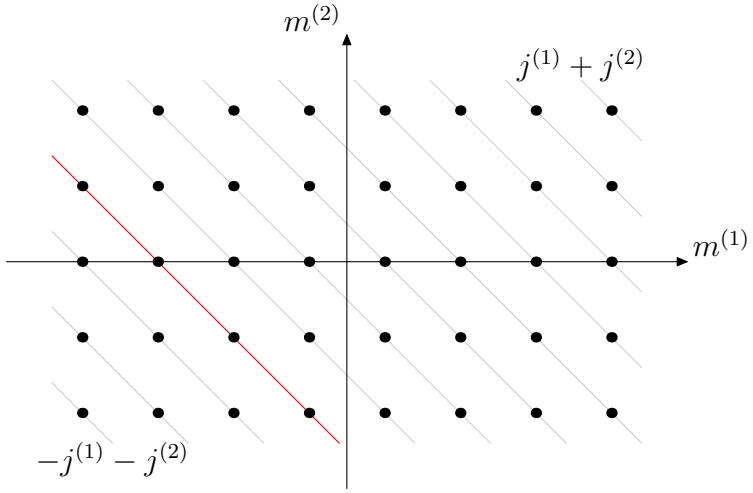


FIG. 5.2. Addition of angular momenta. The bullets display all the pairs $(m^{(1)}, m^{(2)})$ for $j^{(1)} = 7/2$ and $j^{(2)} = 2$. Remember that the only allowed values of $m^{(1)}$ and $m^{(2)}$ given $j^{(1)}$ and $j^{(2)}$ are $m^{(1,2)} = -j^{(1,2)}, -j^{(1,2)} + 1, \ldots, j^{(1,2)} - 1, j^{(1,2)}$ thus $m^{(1)} = -7/2, -5/2, \ldots, 5/2, 7/2$ and $m^{(2)} = -2, -1, 0, 1, 2$. The lines correspond to $M^{(T)} = m^{(1)} + m^{(2)}$. For example, the red line in figure is the line of the pairs $(m^{(1)}, m^{(2)})$ such that $M^{(T)} = -j^{(1)} - j^{(2)} + 3$. To count the number of pairs $n(M^{(T)})$ you have simply to count how many pairs are situated along the line corresponding to $M^{(T)}$.

for

$$J^{(T)} = (j^{(1)} + j^{(2)}), (j^{(1)} + j^{(2)} - 1), \ldots, |j^{(1)} - j^{(2)}| \ ,$$

otherwise for all the other eigenvalues we get $N(J^{(T)}) = 0$.

Remark. These are the eigenstates at $j^{(1)}$ and $j^{(2)}$ fixed.

From the previous considerations we obtain the following :

Theorem 5.1. In the $(2j^{(1)} + 1)(2j^{(2)} + 1)$-dimensional space spanned by

$$|j^{(1)}, j^{(2)}, m^{(1)}, m^{(2)}\rangle$$

with $j^{(1)}$ and $j^{(2)}$ fixed, the possible values of the total angular momentum $J^{(T)}$ are

$$J^{(T)} = (j^{(1)} + j^{(2)}), (j^{(1)} + j^{(2)} - 1), \ldots, |j^{(1)} - j^{(2)}| \ .$$

and to each of the possible values of $J^{(T)}$ it is associated one and only one multiplet of eigenstates $|J^{(T)}, M^{(T)}\rangle$ of the total angular momentum.

5.4 Clebsh-Gordon Coefficients

In the last section dealing with the sum of angular momenta, we have made use of two basis: the $|j_1, j_2, J, M\rangle$ and $|j_1, j_2, m_1, m_2\rangle$. We did not insert any other label into these kets because we made the hypothesis that $\hat{j^2}_1, \hat{j^2}_2, \hat{J}^2, \hat{M}$ form a complete set of observables and the same for the operators associated to the other ket. If this were not so, we would have had to introduce a further label, which we will indicate with the symbol "α", to indicate the eigenvalues of a set of operators to be added to the previous ones to make them a complete set . The states would be indicated as: $|\alpha, j_1, j_2, J, M\rangle$, and $|\alpha, j_1, j_2, m_1, m_2\rangle$.

In order to define the states $|\alpha, j_1, j_2, J, M\rangle$ unambiguously we will assume them to have norm one and we will adopt an appropriate convention for its phase. In the Hilbert space $\mathscr{H}(\alpha, j_1, j_2)$ the states $|\alpha, j_1, j_2, J, M\rangle$ and $|\alpha, j_1, j_2, m_1, m_2\rangle$ form two orthonormal basis and we can pass from one to the other by a unitary transformation:

$$|\alpha, j_1, j_2, J, M\rangle = \sum_{m_1, m_2} |\alpha, j_1, j_2, m_1, m_2\rangle \langle \alpha, j_1, j_2, m_1, m_2 | \alpha, j_1, j_2, m_1, m_2\rangle .$$

It is possible to prove that the coefficients of the transformation, i.e:

$$\langle \alpha, j_1, j_2, m_1, m_2 | \alpha, j_1, j_2, m_1, m_2\rangle$$

are independent of α. We know in fact that in the Hilbert space $\mathscr{H}(\alpha, j_1, j_2)$ the operator $\hat{J}^2$ and $\hat{J}_z$ are represented by matrices which are independent of α . In fact indicating $\langle \alpha, j_1, j_2, J, M|$ with the shortened symbol $\langle \alpha, J, M|$, we know that:

$$\langle \alpha, J, M| \hat{J}_z |\alpha', J', M'\rangle = M \delta_{\alpha,\alpha'} \delta_{J,J'} \delta_{M,M'}$$

and

$$\langle \alpha, J, M| \hat{J}_\pm |\alpha', J', M'\rangle = \sqrt{J(J \pm 1) - MM'} \delta_{\alpha,\alpha'} \delta_{J,J'} \delta_{M,M'\pm 1}.$$

The same will happen for the matrices representing $\hat{j}_1$ and $\hat{j}_2$. As a consequence also the components of the eigenvectors of $\hat{J}^2$ and $\hat{J}_z$, i.e $\langle \alpha, j_1, j_2, m_1, m_2 | \alpha, j_1.j_2, J, M\rangle$

, will be independent of "α". So we can say that these transition coefficients have a pure geometrical interpretation depending only on the angular momenta and their orientation and not upon the physical nature of the dynamical variables labelled by α. These coefficients are called *Clebsh-Gordon* (C.G) coefficients and will be written without the label α, as $\langle j_1, j_2, m_1, m_2 | j_1.j_2, J, M\rangle$. Taking account of this fact the expansion of the state $|\alpha, j_1, j_2, J, M\rangle$ can then be written as::

$$|\alpha, j_1, j_2, J, M\rangle = \sum_{m_1,m_2} |\alpha, j_1, j_2, m_1, m_2\rangle \langle j_1, j_2, m_1, m_2 | j_1.j_2, J, M\rangle .$$

In order to completely determine the C.G. coefficients, we should fix the phases of the vectors $\langle \alpha, j_1, j_2, J, M|$. For a fixed J the relative phases of the 2J associated states have been fixed via the choice of taking real and positive the coefficients of Eqn.(5-14a,b). Once this is done the $\langle \alpha, j_1, j_2, J, M|$ have only the freedom of a phase which can change with J. We fix this phase by requiring that, whatever* is J, the components of $\langle \alpha, j_1, j_2, J, J|$ along $\langle \alpha, j_1, j_2, J, J - j_1|$ be real and positive i.e:

$$\langle j_1, j_2, J, J - j_1 | JJ\rangle \geq 0.$$

Many properties of the C.G coefficients follow directly from their definition and the requirements above for the phases. To get all of the C.G coefficients we have just to use the recursion relations (5-14-a,b) and the orthonormality relation between the states. It is then easy to find several symmetry properties of the C.G coefficients which greatly facilitates their calculation and tabulation as was done by Racah. For more details the reader should study sect.27 of chapter XIII and appendix C of Messiah.

5.5 Problems and Solutions

* J can go from $|j_1 - j_2|$ to $j_1 + j_2$

Problem 5.1. Let us consider two weakly interacting systems whose states are specified by the quantum numbers

$$(l^{(1)}, m^{(1)}) \ , \quad (l^{(2)}, m^{(2)})$$

of the angular momentum and of its projection along the z-axis. Calculate the possible values of the total angular momentum $\hat{L}^{(T)}$ and the mean values of $\hat{L}_i^{(T)}$ and $(\hat{L}^{(T)})^2$ on the corresponding eigenstates.

SOLUTION. Note that here not only $l^{(1)}$ and $l^{(2)}$ but also $m^{(1)}$ and $m^{(2)}$ are fixed. It is possible to prove that

$$\max\left\{|l^{(1)} - l^{(2)}|\,, |m^{(1)} + m^{(2)}|\right\} \leq L^{(T)} \leq l^{(1)} + l^{(2)} \ . \tag{5.26}$$

The upper limit comes obviously from the theory of addition of angular momenta developed in the previous section. In fact, we have seen that

$$|l^{(1)} - l^{(2)}| \leq L^{(T)} \leq l^{(1)} + l^{(2)} \ .$$

The lower limit however is different. Why? We know that

$$-L^{(T)} \leq M^{(T)} \leq L^{(T)}$$

and $M^{(T)} = m^{(1)} + m^{(2)}$. Thus in our present case if $|m^{(1)} + m^{(2)}|$ is higher than $|l^{(1)} - l^{(2)}|$, we must stop at $|m^{(1)} + m^{(2)}|$.

Of course, since 1 and 2 are two different systems, the angular momentum components of the first system do commute with each component of the second one:

$$[\hat{L}_i^{(1)}, \hat{L}_j^{(2)}] = 0 \ .$$

It is also possible to prove that

$$\bar{\hat{L}}_x^{(1)} = \bar{\hat{L}}_x^{(2)} = 0 = \bar{\hat{L}}_y^{(1)} = \bar{\hat{L}}_y^{(2)} \ .$$

We have

$$\begin{aligned}
\bar{\hat{L}}_x^{(T)} &= \bar{\hat{L}}_x^{(1)} + \bar{\hat{L}}_x^{(2)} = 0 \,, \\
\bar{\hat{L}}_y^{(T)} &= \bar{\hat{L}}_y^{(1)} + \bar{\hat{L}}_y^{(2)} = 0 \,, \\
\bar{\hat{L}}_z^{(T)} &= \bar{\hat{L}}_z^{(1)} + \bar{\hat{L}}_z^{(2)} = (m^{(1)} + m^{(2)})\hbar. \,.
\end{aligned}$$

(Prove this as **homework**.)

Let us now consider $\overline{(\hat{L}^{(T)})^2}$:

$$\overline{(\hat{L}^{(T)})^2} = \overline{\left(\hat{\boldsymbol{L}}^{(1)} + \hat{\boldsymbol{L}}^{(2)}\right)^2} = \overline{(\hat{L}^{(1)})^2} + \overline{(\hat{L}^{(2)})^2} + 2\overline{\hat{\boldsymbol{L}}^{(1)} \cdot \hat{\boldsymbol{L}}^{(2)}} \,.$$

We know that

$$\overline{(\hat{L}^{(1)})^2} = \hbar^2 l^{(1)}(l^{(1)}+1) \,, \quad \overline{(\hat{L}^{(2)})^2} = \hbar^2 l^{(2)}(l^{(2)}+1) \,.$$

We have only to calculate

$$2\overline{\hat{\boldsymbol{L}}^{(1)} \cdot \hat{\boldsymbol{L}}^{(2)}} = 2\overline{\hat{L}_x^{(1)}\hat{L}_x^{(2)}} + 2\overline{\hat{L}_y^{(1)}\hat{L}_y^{(2)}} + 2\overline{\hat{L}_z^{(1)}\hat{L}_z^{(2)}} \,.$$

Remember that the states on which we are calculating the mean values are

$$|l^{(1)}, m^{(1)}, l^{(2)}, m^{(2)}\rangle = |l^{(1)}, m^{(1)}\rangle \otimes |l^{(2)}, m^{(2)}\rangle \ ,$$

where the symbol $\otimes$ stay for tensor product and an operative definition has been given in chapter 3, thus for example:

$$\begin{aligned}
\overline{\hat{L}_x^{(1)}\hat{L}_x^{(2)}} &= \left(\langle l^{(1)}, m^{(1)}| \otimes \langle l^{(2)}, m^{(2)}|\right) \hat{L}_x^{(1)}\hat{L}_x^{(2)} \left(|l^{(2)}, m^{(2)}\rangle \otimes |l^{(1)}, m^{(1)}\rangle\right) \\
&= \langle l^{(1)}, m^{(1)}|\hat{L}_x^{(1)}|l^{(1)}, m^{(1)}\rangle \, \langle l^{(2)}, m^{(2)}|\hat{L}_x^{(2)}l^{(2)}, m^{(2)}\rangle \\
&= \overline{\hat{L}_x^{(1)}} \ \overline{\hat{L}_x^{(2)}} = 0 \,.
\end{aligned}$$

In the same way

$$\overline{\hat{L}_y^{(1)}\hat{L}_y^{(2)}} = \overline{\hat{L}_y^{(1)}} \ \overline{\hat{L}_y^{(2)}} = 0 \,, \quad \overline{\hat{L}_z^{(1)}\hat{L}_z^{(2)}} = \overline{\hat{L}_z^{(1)}} \ \overline{\hat{L}_z^{(2)}} = \hbar^2 m^{(1)} m^{(2)} \,.$$

Thus,

$$\overline{(\hat{L}^{(T)})^2} = \hbar^2 \left[l^{(1)}(l^{(1)}+1) + l^{(2)}(l^{(2)}+1) + 2m^{(1)}m^{(2)}\right] \,. \tag{5.27}$$

$\star \quad \star \quad \star$

Problem 5.2. In the same setting of the previous problem, calculate the probabilities of the various values of the angular momentum, assuming $m^{(1)} = l^{(1)}$ and $m^{(2)} = l^{(2)} - 1$.

SOLUTION. By looking at Eqn.(5.26), the range of possible values of $\hat{L}^{(T)}$ is:

$$\max\left\{|l^{(1)} - l^{(2)}|, |m^{(1)} + m^{(2)}|\right\} \leq L^{(T)} \leq l^{(1)} + l^{(2)} .$$

In the special case in which $m^{(1)} = l^{(1)}$ and $m^{(2)} = l^{(2)} - 1$,

$$\max\left\{|l^{(1)} - l^{(2)}|, |m^{(1)} + m^{(2)}|\right\} = \max\left\{|l^{(1)} - l^{(2)}|, |l^{(1)} + l^{(2)} - 1|\right\}$$

and so the range becomes

$$l^{(1)} + l^{(2)} - 1 \leq L^{(T)} \leq l^{(1)} + l^{(2)} .$$

This relation tells that there are only two possible values of $L^{(T)}$, namely

$$l^{(1)} + l^{(2)} , \quad l^{(1)} + l^{(2)} - 1 .$$

If we denote by $P(L^{(T)})$ the probability of finding $L^{(T)}$ in a measurement, then

$$P(l^{(1)} + l^{(2)}) + P(l^{(1)} + l^{(2)} - 1) = 1 .$$

and the mean value $\overline{(\hat{L}^{(T)})^2}$ can be written (in units of $\hbar$)

$$\begin{aligned} \overline{(\hat{L}^{(T)})^2} &= \sum_{L^{(T)}} L^{(T)}(L^{(T)} + 1)P(L^{(T)}) \\ &= (l^{(1)} + l^{(2)})(l^{(1)} + l^{(2)} + 1)P(l^{(1)} + l^{(2)}) + \\ &\quad (l^{(1)} + l^{(2)} - 1)(l^{(1)} + l^{(2)}) \underbrace{\left[1 - P(l^{(1)} + l^{(2)})\right]}_{=P(l^{(1)}+l^{(2)}-1)} . \end{aligned} \tag{5.28}$$

From Eqn. (5.27) with $m^{(1)} = l^{(1)}$ and $m^{(2)} = l^{(2)} - 1$ we get

$$\overline{(\hat{L}^{(T)})^2} = l^{(1)}(l^{(1)} + 1) + l^{(2)}(l^{(2)} + 1) + 2l^{(1)}(l^{(2)} - 1) ,$$

comparison between this expression and Eqn. (5.28) yields

$$P(l^{(1)}+l^{(2)}) = \frac{l^{(2)}}{l^{(1)}+l^{(2)}} \,, \qquad P(l^{(1)}+l^{(2)}-1) = \frac{l^{(1)}}{l^{(1)}+l^{(2)}} \,.$$

$\star \quad \star \quad \star$

Problem 5.3. Prove that, starting from the commutation relation

$$[\hat{L}_i, \hat{O}] = 0 \,,$$

between the angular momentum operators $\hat{L}_i$ and an Hermitian operator $\hat{O}$, it is possible to deduce that the matrix elements

$$\langle L, M|\hat{O}|L, M'\rangle$$

are different from zero only if $M = M'$ and they are independent of M.

SOLUTION. Let us start from the commutation relation

$$\hat{L}_z\hat{O} - \hat{O}\hat{L}_z = 0 \,.$$

By applying it to the eigenstate $|M\rangle$ we get

$$\hat{L}_z\hat{O}\,|M\rangle = \hat{O}\hat{L}_z\,|M\rangle = M\hat{O}\,|M\rangle \;\,,$$

and this tells us that $\hat{O}\,|M\rangle$ is eigenstate of $\hat{L}_z$ with eigenvalue M. Thus,

$$\langle M'|\hat{O}M\rangle \propto \langle M'|M\rangle = \delta_{M',M} \,,$$

which is the first part of what we have to prove.

Let us now prove the second part. Since every $\hat{L}_i$ commutes with $\hat{O}$, $\hat{L}_{(+)}$ commutes with $\hat{O}$ too:

$$\hat{L}_{(+)}\hat{O} = \hat{O}\hat{L}_{(+)} \,,$$

and by taking the matrix elements of $\hat{L}_{(+)}\hat{O}$ between $\langle M+1|$ and $|M\rangle$ we get

$$\langle M+1|\hat{L}_{(+)}\hat{O}|M\rangle = \langle M+1|\hat{O}\hat{L}_{(+)}|M\rangle \ .$$

We know that $\hat{L}_{(+)}$ acts on $|M\rangle$ in the following way:

$$\hat{L}_{(+)}\,|M\rangle = \sqrt{(L-M)(L+M+1)}\,|M+1\rangle \ .$$

It is also easy to prove that

$$\langle M+1|\,\hat{L}_{(+)} = \sqrt{(L-M)(L+M+1)}\,\langle M| \ .$$

Hence,

$$\sqrt{(L-M)(L+M+1)}\,\langle M|\hat{O}|M\rangle = \langle M+1|\hat{O}|M+1\rangle\,\sqrt{(L-M)(L+M+1)}$$

which implies

$$\langle M|\hat{O}|M\rangle = \langle M+1|\hat{O}|M+1\rangle \ ,$$

This proves that the expectation values of $\hat{O}$ are independent of M.

$\star \quad \star \quad \star$

Problem 5.4. Find the eigenstates of the total angular momentum of two systems whose angular momenta are $L^{(1)} = 2$ and $L^{(2)} = 1$. How many states are there? How many of them have $M^{(T)} = 0$?

Solution. For the rules of addition of angular momenta, $L^{(T)}$ can take only the following values:

$$|L^{(1)} - L^{(2)}|, \ldots, L^{(1)} + L^{(2)} \ .$$

In our case, $L^{(1)} = 2$, $L^{(2)} = 1$ thus $|L^{(1)} - L^{(2)}| = 1$ and $L^{(1)} + L^{(2)} = 3$, therefore

$$L^{(T)} = \begin{cases} 3 & \text{or} \\ 2 & \text{or} \\ 1 & \end{cases}$$

Associated to each value of $L^{(T)}$ are the corresponding values of $M^{(T)}$, for example for $L^{(T)} = 3$ the possible values of $M^{(T)}$ are

$$M^{(T)} = -3, -2, -1, 0, 1, 2, 3 \ .$$

Therefore, the states $|L^{(T)}, M^{(T)}\rangle$ with $L^{(T)} = 3$ are

$$|3,3\rangle \, , |3,2\rangle \ \, , |3,1\rangle \, , |3,0\rangle \, , |3,-1\rangle \, , |3,-2\rangle \, , |3,-3\rangle \ \, .$$

There are seven states $|L^{(T)}, M^{(T)}\rangle$ with $L^{(T)} = 3$. For $L^{(T)} = 2$ the possible states are

$$|2,2\rangle \, , |2,-1\rangle \, , |2,0\rangle \, , |2,-1\rangle \, , |2,-2\rangle \ \, ,$$

there are five such states. Finally, there are three states with $L^{(T)} = 1$, namely

$$|1,1\rangle \, , |1,0\rangle \, , |1,-1\rangle \ \, .$$

There is a total of $7 + 5 + 3 = 15$ states, three of which have $M^{(T)} = 0$: $|3,0\rangle$, $|2,0\rangle$ and $|1,0\rangle$.

$\star \quad \star \quad \star$

Chapter 6

SYMMETRIES IN QUANTUM MECHANICS

6.1 Rotations, symmetries, unitary transformations and degeneracy

In quantum mechanics the basic quantities are the probabilities:

$$P = \langle\psi|\psi\rangle \ .$$

If now we change the states $|\psi\rangle$ via a unitary transformation $\hat{U}$:

$$\hat{U}\,|\psi\rangle = |\psi'\rangle \ ,$$

it is easy to prove that the probabilities are left unchanged. In fact,

$$P' = \langle\psi'|\psi'\rangle = \langle\psi|\hat{U}^\dagger\hat{U}|\psi\rangle = \langle\psi|\hat{U}^{-1}\hat{U}|\psi\rangle = \langle\psi|\psi\rangle = P \ .$$

Therefore, the probabilities do not change under unitary transformations of states.

What happens to the mean values of an observable? Let us consider the expectation value of an observable $\hat{O}$ on a state $|\psi\rangle$, it is given by

$$\bar{O} = \langle\psi|\hat{O}|\psi\rangle \ .$$

Now, let $|\psi\rangle$ be transformed according to an unitary transformation $\hat{U}$: $|\psi'\rangle = \hat{U}\,|\psi\rangle$. In order the mean value to be the same, $\hat{O}$ must be transformed according to

$$\hat{O}' = \hat{U}\hat{O}\hat{U}^\dagger \ .$$

In fact, in this way one gets

$$\langle\psi'|\hat{O}'|\psi'\rangle = \langle\psi|\hat{U}^\dagger\hat{U}\hat{O}\hat{U}^\dagger\hat{U}|\psi\rangle = \langle\psi|\hat{O}|\psi\rangle \ .$$

So by performing a unitary transformation we have to change both the states and the operators.

Somehow unitary transformations play in quantum mechanics the role that in classical mechanics is played by canonical transformations.

Let us now turn back to rotations. In classical mechanics, angular momentum is the generator of the rotations, in the sense that any observable O transforms under a rotation of an angle $\Delta\varphi$ around the z-axis according to the following rule:

$$O' = O + \Delta\varphi \{L_z, O\}_{\text{P.B.}} \ , \tag{6.1}$$

as long as $\Delta\varphi$ can be considered small enough.$(P.B)$ stays for Poisson Brackets. If we now look at infinitesimal time translations, these are generated by $\hat{H}$ and accordingly we have

$$O' = O + \Delta t \{H, O\}_{\text{P.B.}} \ .$$

By using the correspondence rule, we may infer the transformation in quantum mechanics:

$$\hat{O}' = \hat{O} + \Delta t \frac{[\hat{H}, \hat{O}]}{i\hbar} \ .$$

We also know that for finite transformations the time-evolution operator is given by

$$U_t = \mathrm{e}^{-\frac{i}{\hbar}\hat{H}t} \ . \tag{6.2}$$

In the same way, for the infinitesimal rotation given by Eqn. (6.1) the quantum analogous becomes:

$$\hat{O}' = \hat{O} + \Delta\varphi \frac{[\hat{L}_z, \hat{O}]}{i\hbar} \ . \tag{6.3}$$

For a *finite* rotation we should get the analogous of Eqn. (6.2), namely

$$\hat{U}_\varphi = \mathrm{e}^{-\frac{i}{\hbar}\hat{L}_z\varphi} \ .$$

This means that under finite rotations the states should transform according to

$$|\psi'\rangle = \hat{U}_\varphi |\psi\rangle \ ,$$

while observables $\hat{O}$ transform in the following way:

$$\hat{O}' = \hat{U}_\varphi \hat{O} \hat{U}_\varphi^\dagger = \mathrm{e}^{-\frac{i}{\hbar}\hat{L}_z\varphi}\, \hat{O}\, \mathrm{e}^{\frac{i}{\hbar}\hat{L}_z\varphi} \ .$$

Let us now see if from this, for small angles, we get (6.3). Let us start from

$$\hat{O}' = \mathrm{e}^{-\frac{i}{\hbar}\Delta\varphi\hat{L}_z}\, \hat{O}\, \mathrm{e}^{\frac{i}{\hbar}\Delta\varphi\hat{L}_z} \ ,$$

by expanding the exponentials we get

$$\begin{aligned} \hat{O}' &= \left[1 - \frac{i}{\hbar}\Delta\varphi\hat{L}_z + \ldots\right] \hat{O} \left[1 + \frac{i}{\hbar}\Delta\varphi\hat{L}_z + \ldots\right] \\ &= \hat{O} + \frac{i}{\hbar}\Delta\varphi[\hat{O}, \hat{L}_z] + \mathcal{O}(\Delta\varphi^2) \\ &= \hat{O} + \frac{\Delta\varphi}{i\hbar}[\hat{L}_z, \hat{O}] + \mathcal{O}(\Delta\varphi^2) \end{aligned}$$

which is nothing but Eqn. (6.3).

If $[\hat{L}_z, \hat{O}] = 0$, then $\hat{O}' = \hat{O}$, *i.e.*, the observable $\hat{O}$ is invariant under rotations.

Let us now try to derive explictely that $\hat{U}_\varphi = \mathrm{e}^{-i\varphi\hat{L}_z/\hbar}$ generates a *finite* rotation around the z-axis.

For an infinitesimal rotation we have:

$$|\psi'\rangle = \hat{U}_{\Delta\varphi}\, |\psi\rangle = \mathrm{e}^{-\frac{i}{\hbar}\Delta\varphi\hat{L}_z}\, |\psi\rangle \ .$$

Taking the derivative of $|\psi'\rangle$ with respect to φ, we get:

$$\frac{\mathrm{d}\,|\psi'\rangle}{\mathrm{d}\varphi} = -\frac{i}{\hbar}\hat{L}_z\, |\psi'\rangle \ .$$

If we perform explicitly a traslation of an infinitesimal angle ε on φ then $\psi(r, \vartheta, \varphi)$ goes into $\psi'(r, \vartheta, \varphi + \varepsilon)$, hence

$$\begin{aligned} \psi'(r, \vartheta, \varphi + \varepsilon) &= \psi(r, \vartheta, \varphi) + \varepsilon \frac{\mathrm{d}}{\mathrm{d}\varphi}\psi(r, \vartheta, \varphi) \\ &= \psi(r, \vartheta, \varphi) - \frac{i\varepsilon}{\hbar}\hat{L}_z\psi(r, \vartheta, \varphi) + \ldots \end{aligned}$$

and since ε is "small" we can write

$$|\psi'\rangle = \mathrm{e}^{-\frac{i}{\hbar}\varepsilon \hat{L}_z} |\psi\rangle \ .$$

To get a finite rotation, we have to multiply many infinitesimal rotations:

$$\mathrm{e}^{-\frac{i}{\hbar}\varepsilon_1 \hat{L}_z} \mathrm{e}^{-\frac{i}{\hbar}\varepsilon_2 \hat{L}_z} \ldots \tag{6.4}$$

By applying the Baker-Hausdorff formula

$$\mathrm{e}^{\hat{A}} \mathrm{e}^{\hat{B}} = \mathrm{e}^{\mathcal{H}(\hat{A},\hat{B})} \ ,$$

where

$$\mathcal{H}(\hat{A},\hat{B}) = \hat{A} + \hat{B} + \frac{1}{2}[\hat{A},\hat{B}] + \ldots$$

and the dots denote commutators of commutators, we get from Eqn. (6.4)

$$\mathrm{e}^{-\frac{i}{\hbar}(\varepsilon_1+\varepsilon_2+\ldots)\hat{L}_z} \ , \tag{6.5}$$

because the various $\hat{L}_z$ commute among themselves. Eqn. (6.5) means that even for finite rotations the generator is given by $\hat{U}(\varphi) = \mathrm{e}^{-i\varphi \hat{L}_z/\hbar}$ that is what we wanted to prove.

So far we have dealt with rotations around the z-axis. What's about generic rotations? It is well-known from analytic mechanics that any rotation can be written as a rotation of some angle α around a direction specified by a unit vector $\boldsymbol{n}$. Thus, three angles are needed in order to specify a generic rotation: the two that specify $\boldsymbol{n}$ and the rotation angle α. The unitary operators associated to such rotations are given by

$$\hat{U}(\alpha, \boldsymbol{n}) = \mathrm{e}^{-\frac{i}{\hbar}\alpha \boldsymbol{n}\cdot\hat{\boldsymbol{L}}} \ , \tag{6.6}$$

where $\boldsymbol{n} \cdot \hat{\boldsymbol{L}} = n_x \hat{L}_x + n_y \hat{L}_y + n_z \hat{L}_z$. Eqn. (6.6) reduces to the previous formula for the case of rotations about the z-axis.

Remark: The statement we made at the beginning of this section that mean values are left unchanged under unitary transformations, is always true except in the case

of the time evolution generated by $\mathrm{e}^{-i\hat{H}t/\hbar}$. Under this transformation, mean values change in time. We will see later that in this case states and observables must be treated on a different ground (Schrödinger and Heisenberg picture).

6.2 Symmetries, generators and degeneracy

In classical mechanics, a symmetry is a transformation which leaves the equation of motion invariant. It is easy to prove that at the Hamiltonian level this implies the existence of a quantity $Q(q,p)$ having null Poisson brackets with the Hamiltonian:

$$\{Q(q,p),H\}_{\text{P.B.}}=0\ .$$

If $Q(q,p)$ does not depend explicitly on time, then

$$\frac{\mathrm{d}Q(q,p)}{\mathrm{d}t}=\frac{\partial Q}{\partial q}\frac{\partial q}{\partial t}+\frac{\partial Q}{\partial p}\frac{\partial p}{\partial t}=\frac{\partial Q}{\partial q}\frac{\partial H}{\partial p}-\frac{\partial Q}{\partial p}\frac{\partial H}{\partial q}=\{Q,H\}_{\text{P.B.}}\ .$$

If $\{Q,H\}_{\text{P.B.}}=0$ then $\mathrm{d}Q/\,\mathrm{d}t=0$, *i.e.*, Q is a conserved quantity (Noether theorem).

By using the corrispondence principle,

$$\{\cdot,\cdot\}_{\text{P.B.}}\to\frac{1}{i\hbar}[\cdot,\cdot]\ ,$$

we get

$$\{Q,H\}_{\text{P.B.}}\to\frac{1}{i\hbar}[\hat{Q},\hat{H}]\ ,$$

and from $\{Q,H\}_{\text{P.B.}}=0$ we obtain $[\hat{Q},\hat{H}]=0$. This means that at the *quantum* level a symmetry is signaled by the presence of a charge $\hat{Q}$ which commutes with the Hamiltonian.

$\hat{Q}$ is referred to as "*charge*" or also as "*generator*" of a symmetry. The first name comes from the analogy with electromagnetism, where gauge invariance, namely,

$$A^\mu\to A^\mu+\partial^\mu\varepsilon(x)\ ,$$

leads to the conservation of charge. The second name comes from the fact that the transformation

$$\begin{aligned}\delta q&=\{q,Q\}_{\text{P.B.}}\varepsilon\ ,\\ \delta p&=\{p,Q\}_{\text{P.B.}}\varepsilon\ ,\end{aligned}\tag{6.7}$$

generates a symmetry of the equation of motion. Here, ε is an infinitesimal parameter, When ε does not depends on time, we speak of "global" symmetry, otherwise we speak of "local" or gauge transformation.

Homework 6.1. Prove that if $\{Q, H\}_{\text{P.B.}} = 0$ then the equation of motion are left unchanged in form, *i.e.*, Q generates a symmetry.

From the correspondence principle, it may seem that if a charge is conserved at the *classical* level then it is conserved also at *quantum* level. Actually, this is not always the case, at least when considering systems having infinitely many degrees of freedom, such as fields. In such cases quantum effects may break the symmetry. The two most studied examples of such effects are the spontaneous symmetry breaking and the anomalies, the former being related for example to the Higgs mechanism while the latter is responsible for the $\pi^0 \to 2\gamma$ decay and it also implies that there must be as many quarks as leptons.

Let us now turn to investigate what are the first consequences of the presence of a symmetry in quantum mechanics. The first consequence is *degeneracy*. Let us suppose we know the eigenstates of the Hamiltonian $\hat{H}$:

$$\hat{H}\,|E_n\rangle = E_n\,|E_n\rangle \ .$$

For simplicity, we are assuming the eigenstates to be discrete. We have learned that if $\hat{H}$ has a symmetry, then there must be a charge $\hat{Q}$ which commutes with it:

$$[\hat{Q}, \hat{H}] = 0 \ .$$

Now, it is easy to prove that the state

$$|\tilde{E}_n\rangle \equiv \hat{Q}\,|E_n\rangle$$

is again an eigenstate of $\hat{H}$ with the same eigenvalue of $|E_n\rangle$, namely E_n, thus we

have degeneracy. The proof goes as follows. From $[\hat{Q}, \hat{H}] = 0$ we get

$$\begin{aligned}
[\hat{Q}, \hat{H}] \, |E_n\rangle &= 0 \\
\hat{Q}\hat{H} \, |E_n\rangle - \hat{H}\hat{Q} \, |E_n\rangle &= 0 \\
\hat{Q} E_n \, |E_n\rangle - \hat{H} \, |\tilde{E}_n\rangle &= 0 \\
E_n \hat{Q} \, |E_n\rangle - \hat{H} \, |\tilde{E}_n\rangle &= 0 \\
E_n \, |\tilde{E}_n\rangle - \hat{H} \, |\tilde{E}_n\rangle &= 0 \ ,
\end{aligned}$$

from which it follows

$$\hat{H} \, |\tilde{E}_n\rangle = E_n \, |\tilde{E}_n\rangle \ .$$

If now we consider the unitary operator $\hat{U}$ associated to $\hat{Q}$, *i.e.*,

$$\hat{U}(\alpha) = \mathrm{e}^{-\frac{i}{\hbar}\hat{Q}\alpha} \ ,$$

where α is the parameter of the symmetry, we can construct a whole family of states

$$\hat{U}(\alpha) \, |E_n\rangle = |\tilde{E}_n, \alpha\rangle \ ,$$

and they are all degenerate with $|E_n\rangle$, *i.e.*:

$$\begin{aligned}
\hat{H} \, |\tilde{E}_n, \alpha\rangle = \hat{H} \, \mathrm{e}^{-\frac{i}{\hbar}\hat{Q}\alpha} \, |E_n\rangle = \mathrm{e}^{-\frac{i}{\hbar}\hat{Q}\alpha} \, \hat{H} \, |E_n\rangle = \mathrm{e}^{-\frac{i}{\hbar}\hat{Q}\alpha} \, E_n \, |E_n\rangle \\
= E_n \, \mathrm{e}^{-\frac{i}{\hbar}\hat{Q}\alpha} \, |E_n\rangle = E_n \, |\tilde{E}_n, \alpha\rangle \ .
\end{aligned}$$

Remark: not every degeneracy is due to symmetries. There are also accidental degeneracies.

6.3 Brief introduction to Lie groups.

The set of unitary operators $\hat{U}(\alpha, \boldsymbol{n}) = \mathrm{e}^{-\frac{i}{\hbar}\alpha \boldsymbol{n}\cdot\hat{\boldsymbol{L}}}$ which implement the generic rotations on the Hilbert space form a structure which in mathematics is called a *group*. A group $\mathcal{G}$ is a set of elements $g \in \mathcal{G}$ on which an internal operation is defined (let us indicated it with $*$) such that:

$$g * g' = g'' \in \mathcal{G}$$

Moreover in $\mathcal{G}$ there must exist an *identity* element, "e", such that:

$$e * g = g * e = g$$

and finally, given any element g there must exist in $\mathcal{G}$ the *inverse* of g, indicated by g^{-1} such that:

$$g * g^{-1} = g^{-1} * g = e.$$

Some groups are a set on which the concept of *continuity* can be defined and this, together with few other properties, make what are called the *Lie groups*. For these it is always possible to find what are called the *generators* $\widehat{O}_i$ of the group. They are such that every element U of the group can be written as :

$$U = \mathrm{e}^{i \sum_i \alpha_i \widehat{O}_i}$$

where the α_i are a set of continuos parameters which are as many as the group generators $\widehat{O}_i$. The group structure of U induces the following algebra on the generators:

$$[\widehat{O}_i, \widehat{O}_j] = C^k_{i,j} \widehat{O}_k$$

where $[\cdot, \cdot]$ is the commutator and the $C^k_{i,j}$ are constants called *structure constants*. The above commutation relation defines what is called a *Lie-algebra* among the generators. For the rotation group the analogue of this algebra are the commutators among the angular momentum operators. Given an *abstract* Lie algebra we can build the Lie-group but then we have to decide on which space it acts. The explicit realization of the algebra and the group elements on this space is called a "*representation*" of the group and of the algebra. For example for the rotation group and algebra we have seen in the previous chapter that the spaces on which it can act are vector spaces of dimension $2J + 1$ with J integer or half-integer. So we can say that each value of J labels a representation.

This is just a brief introduction to the concept of groups. More will be studied in advanced courses next year.

6.4 Scalar and Tensor Operators

From now on we shall concentrate con the rotation group.

There are observables which are invariant under rotation like for example:

$$\hat{R}^2 \equiv \hat{x}^2 + \hat{y}^2 + \hat{z}^2$$

These operators are called *scalars* and we will indicate them with $\hat{S}$. From the invariance of $\hat{S}$ under $U = \mathrm{e}^{-\frac{i}{\hbar}\alpha \boldsymbol{n}\cdot\hat{\boldsymbol{J}}}$, i.e:

$$\hat{S} = U\hat{S}U^\dagger$$

we get that $\hat{S}$ commutes with the generators $\hat{\boldsymbol{J}}$:

$$[\hat{S}, \hat{\boldsymbol{J}}] = 0. \tag{6.8}$$

In an exercise in chapter 5) we proved that, from the commutation relation (6.8), we can derive that:

$$\langle J, M|\, \hat{S}\, |J', M'\rangle = \delta_{J,J'}\delta_{M,M'}S^{(J)} \tag{6.9}$$

where $S^{(J)}$ is a quantity which depends on J but not on M. Eqn.(6.9) tells us also that the state $\hat{S}\,|J', M'\rangle$ is still a state with quantum numbers (J', M'), i.e $\hat{S}$ does not change the quantum numbers of the state to which it is applied. We can generalize the concept of scalar operator to *vector* and *tensor* operators. They are defined by the manner they transform under rotation. A vector operator $\hat{\boldsymbol{V}}$ is defined as:

$$U\hat{V}_i U^\dagger = \sum_{J=1}^{3} R_{ij}\hat{V}_J$$

where R_{ij} are the same three by three rotation matrix we would obtain if we apply a rotation to (x, y, z). If we take two vector operators $\hat{\boldsymbol{V}}$ and $\hat{\boldsymbol{W}}$, we can build the following object with 9 components:

$$\hat{V}_i\hat{W}_j \tag{6.10}$$

plus all their linear combinations. Their transformation properties would be

$$(\hat{V}_i\hat{W}_j)' = R_{ik}R_{js}(\hat{V}_k\hat{W}_s).$$

This is a particular *tensor*. We see that it is an object with multi-index and each of them transform linearly, in our case with the matrix R. Among the tensors we will pay particular attention to the *irreducible* ones. These are defined as follows: let us suppose the space of our tensors has dimension N. If there is a subset of this space of dimension $M < N$ such that any rotation brings every element of this subset into elements of the same subset, then we say that the tensor space is reducible, in the opposite case we say that it is irreducible. When it is reducible we can represents our tensor in the subspace M. For example the tensor we have built in Eqn.(6.10) is not irreducible. Its 9-dimensional space contains irreducible subspaces of dimension 1,3 and 5. They are respectively the space of the scalar product: $\sum_i \hat{V}_i\hat{W}_i$ that under rotation is left unchanged and it is a 1-dimensional space, the 3-dimensional space associated to the the vector product:$\hat{\boldsymbol{V}} \wedge \hat{\boldsymbol{W}} = \sum \epsilon_{ijk}V_jW_k$ which is a vector and so under rotation goes into a vector, and the 5-dimensional space associated to $J = 2$. The dimensions of all these spaces is always $2J + 1$ with J integer or half-integer. In general, it is easy to prove that, if the $(2J + 1)$ operators $\hat{T}^{(J)}_M$ with $-J \leq M \leq J$ are the standard components of an *irreducible* J-th order tensor operator $\hat{T}^{(J)}$, then they transform under rotation according to the law:

$$U\hat{T}^{(J)}_M U^{-1} = \sum_{M'} \hat{T}^{(J)}_{M'} R^{(J)}_{M'M} \tag{6.11}$$

where $R^{(J)}_{M'M}$ are the matrices of rotation for the states in the J representation. This means that these matrix elements are the same as those which enter the transformation:

$$U\,|J, M\rangle = \sum_{M'} |J, M'\rangle\, R^{(J)}_{M'M} \tag{6.12}$$

Using Eqn.(6.12) and (6.11) it is easy to prove that the operators $\hat{T}^{(J)}_M$ have the

following commutation relations with the generators $\hat{J}_{\pm}, \hat{J}_z$

$$[\hat{J}_{\pm}, \hat{T}^{(J)}_M] = \sqrt{J(J+1) - M(M \pm 1)}\hat{T}^{(J)}_{M\pm 1} \tag{6.13}$$

and

$$[\hat{J}_z, \hat{T}^{(J)}_M] = M\hat{T}^{(J)}_M \tag{6.14}$$

6.5 Wigner-Eckart Theorem

What we shall do in this section is to give the general form of the matrix elements of irreducible tensor operators. Somehow this is the analogue of what we did in Eqn. (6.9) for scalar operators. For convenience of notation in this section we will indicate the irreducible tensors with the notation: $\hat{T}^{(K)}_q$ which means we are in the (K) representation and take the q element. Let us now indicate with

$$\langle J', K, M', q | JM \rangle \tag{6.15}$$

the Clebsh-Gordon coefficients which in the full notation we would write as:

$$\langle J', K, M', q | J', K, J, M \rangle \tag{6.16}$$

where J' and K are the analogue of $j^{(1)}$ and $j^{(2)}$ used in the previous chapter and M' and q are the analogue of $m^{(1)}$ and $m^{(2)}$. Let us also make the assumption that $\hat{J}^2, J_z$ are not a complete set of operators and that, in order to specify uniquely the states, we need some extra quantum numbers which we will indicate with α. It is then possible to prove the following relation:

$$\langle \alpha, J, M | \hat{T}^{(K)}_q | \alpha', J'.M' \rangle = \frac{1}{\sqrt{(2J+1)}} \langle \alpha, J | |T^{(K)}| | \alpha', J' \rangle \cdot \langle J', K, M', q | J, M \rangle \tag{6.17}$$

The quantity that we have simbolically indicated with $\langle \alpha, J | |T^{(K)}| | \alpha', J' \rangle$ is called *reduced matrix elements* and it is a number which depends only on $(J, J', \alpha, \alpha', K)$ but *not* on (M, M', q). We will not calculate it explicitly except for vector operators. The other term on the RHS of Eqn.(6.17) is a Clebsh Gordon coefficient.

Eqn. (6.17) is called *Wigner-Eckart theorem* and for a detailed derivation we refer the reader to capt.XIII of Messiah. We will prove it later on for vector operators only. What instead we will do now is to derive some consequences of this theorem. One of these result is a *selection rule* which tells us that the matrix elements of $\hat{T}_q^{(K)}$ are different from zero only if:

$$q = M - M' \tag{6.18}$$

and

$$|J - J'| \leq K \leq (J + J') \tag{6.19}$$

This is so because on the R.H.S of Eqn. (6.17) we have a Clebsh-Gordon coefficient which is different from zero only if the relations (6.18) and (6.19) are satisfied. Let us apply all of this to a *vector-operator* $\hat{\boldsymbol{V}}$. It is called vector operator because, as we explained in previous section, under the rotation generators $\hat{\boldsymbol{J}}$ it transform like a vector which means it has the following commutation relations with $\hat{\boldsymbol{J}}$:

$$[\hat{J}_x, \hat{V}_x] = 0, [\hat{J}_x, \hat{V}_y] = i\hbar\hat{V}_z, [\hat{J}_x, \hat{V}_z] = -i\hbar\hat{V}_y \tag{6.20}$$

and cyclically with the other components of $\hat{\boldsymbol{J}}$. We can also build the $\hat{V}_\pm$ and $\hat{J}_\pm$ and derive from (6.20) the following set of commutation relations:

$$[\hat{J}_x, \hat{V}_\pm] = \mp\hbar\hat{V}_z, [\hat{J}_y, \hat{V}_\pm] = -i\hbar\hat{V}_z, [J_z, V_\pm] = \pm\hbar\hat{V}_\pm \tag{6.21}$$

and

$$[\hat{J}_+, \hat{V}_+] = 0, [\hat{J}_+, \hat{V}_-] = 2\hbar\hat{V}_z, [\hat{J}_-, \hat{V}_+] = -2\hbar\hat{V}_z, [\hat{J}_-, \hat{V}_-] = 0 \tag{6.22}$$

Let us now calculate the matrix elements of $\hat{V}_z$ between $\langle J, M|$ and $|J', M'\rangle$, i.e.:

$$\langle J'M'| \hat{V}_z |J, M\rangle \tag{6.23}$$

We will see that a lot of these elements will be zero. For example it is easy to prove that if $M \neq M'$ the matrix elements above are zero. The proof goes as follows: we know from the commutation relations above that $[\hat{V}_z, \hat{J}_z] = 0$ and if we sandwich it

among the states we get: $\langle J', M'| [\hat{V}_z, \hat{J}_z] |J, M\rangle = 0$. Applying $\hat{J}_z$ on the states on the left and on the right we get:

$$\langle J', M'| [\hat{V}_z, \hat{J}_z] |J, M\rangle = \hbar(M - M') \langle J'M'| \hat{V}_z |J, M\rangle = 0$$

Now if we choose $M \neq M'$ we get that the relation above is satisfied only if:

$$\langle J'M'| \hat{V}_z |J, M\rangle = 0.$$

Another easy thing to prove is that the matrix elements: $\langle J, M| V_\pm |J, M'\rangle$ are different from zero if and only if $M' = M \pm 1$. This goes as follows. Let us sandwich the commutation relation $[\hat{J}_z, \hat{V}_\pm] = \pm\hbar\hat{V}_\pm$ among states of different M, and we get

$$\langle J, M| \hat{V}_\pm |J, M'\rangle = \pm\frac{1}{\hbar} \langle JM| [\hat{J}_z, \hat{V}_\pm] |J, M'\rangle$$

If we apply, in the term on the RHS, the operator $\hat{J}_z$ on the bra and the ket, we get that the equality above becomes:

$$\pm(M \mp 1 - M') \langle J, M| \hat{V}_\pm |J, M'\rangle = 0$$

As a consequence, if $M \neq M' \pm 1$, we get that it must be:

$$\langle J, M| \hat{V}_\pm |J, M'\rangle = 0$$

We will now prove the *important* result that the matrix elements of $\hat{\boldsymbol{V}}$ are proportional to the matrix elements of $\hat{\boldsymbol{J}}$. Let us remember from (6.22) that $\hat{J}_+$ and $\hat{V}_+$ commutes so their matrix elements between any state are zero. Let us in particular sandwich it between $\langle J, M + 2|$ and $|J, M\rangle$ and we get:

$$\langle J, M + 2| \hat{J}_+\hat{V}_+ |J, M\rangle = \langle J, M + 2| \hat{V}_+\hat{J}_+ |J, M\rangle .$$

Let us now insert a completeness $\sum |J', M'\rangle \langle J'M'| = 1$ between the operators $\hat{J}_+$ and $\hat{V}_+$ above. Because of the manner the operator $\hat{J}_+$ acts on the states, the only

elements different from zero which survive ,after we have inserted the completeness, produce the following relation:

$$\langle J, M+2|\,\hat{J}_+\,|J, M+1\rangle\,\langle J, M+1|\,\hat{V}_+\,|J, M\rangle = \langle J, M+2|\,\hat{V}_+\,|J, M+1\rangle\,\langle J, M+1|\,\hat{J}_+\,|J, M\rangle$$

which is equivalent to :

$$\frac{\langle J, M+1|\,\hat{V}_+\,|J, M\rangle}{\langle J, M+1|\,\hat{J}_+\,|J, M\rangle} = \frac{\langle J, M+2|\,\hat{V}_+\,|J, M+1\rangle}{\langle J, M+2|\,\hat{J}_+\,|J, M+1\rangle}$$

This holds for all $-J \leq M \leq J$ so we can say that from the relation above we get:

$$\langle J, M+1|\,\hat{V}_+\,|J, M\rangle = \alpha(J)\,\langle J, M+1|\,\hat{J}_+\,|JM\rangle \tag{6.24}$$

where $\alpha(J)$ is a proportionality constant independent of M but which could, in principle, depends on J. As the RHS above is different from zero only if the bra is $\langle J, M+1|$, we can generalize Eqn.(6.24) as follows:

$$\langle J, M'|\,\hat{V}_+\,|J, M\rangle = \alpha(J)\,\langle J, M'|\,\hat{J}_+\,|JM\rangle \tag{6.25}$$

It is easy to prove that this proportionality can be extended to all the components of $\hat{\boldsymbol{V}}$ with the same proportinality factor $\alpha(J)$. Let us prove this for the z-component. We start from the commutation relation:

$$[\hat{J}_-, \hat{V}_+] = -2\hbar\hat{V}_z \tag{6.26}$$

and let us sandwich both sides between $\langle J, M|$ and $|J, M\rangle$. Acting with the $\hat{J}_-$ which is inside (6.26) on the bra and the ket on which we have sandwiched the above commutator, we obtain for the LHS of (6.26):

$$\hbar\sqrt{J(J+1) - M(M+1)}\,\langle J, M+1|\,\hat{V}_+\,|J, M\rangle - \hbar\sqrt{J(J+1) - M(M-1)}\,\langle J.M|\,\hat{V}_+\,|J, M-1\rangle$$

This last expression, once we use (6.25) in it, becomes:

$$\alpha\hbar\{\sqrt{J(J+1) - M(M+1)}\,\langle J, M+1|\,\hat{J}_+\,|J, M\rangle - \sqrt{J(J+1) - M(M-1)}\,\langle J, M|\,\hat{J}_+\,|J, M-1\rangle\}$$

This expression can be easily worked out and we get that it is equal to $-2\hbar\alpha M$ which can be written as $2\alpha(J)\langle J,M|\hat{J}_z|J,M\rangle$. So, going now back to (6.26), we can conclude that:

$$\langle JM|\hat{V}_z|JM\rangle = \alpha(J)\langle J,M|\hat{J}_z|J,M\rangle$$

which is what we wanted to prove. In a similar manner we can prove the same proportionality for all the other components of $\hat{\boldsymbol{V}}$. This is basically the Wigner-Eckart theorem for vector operators.

Chapter 7

SPIN

In the quantization of angular momentum, or better, of the operators $\hat{J}_i$ satisfying the same algebra of angular momentum operators, namely

$$[\hat{J}_i, \hat{J}_j] = i\hbar\varepsilon_{ijk}\hat{J}_k \ ,$$

we proved that we can get only integer or *half-integer* values. Now, we are interested in trying to check if the half-integer cases can be associated to some physical phenomena. Of course, half-integer values cannot be related to *orbital* angular momentum. In fact, suppose we perform a rotation about z on some wave function $\psi(\varphi)$. We know the unitary operator corresponding to this rotation is given by

$$\hat{U}(\alpha) = \mathrm{e}^{-\frac{i}{\hbar}\hat{L}_z\alpha} \ ,$$

being α the angle of rotation. Now, suppose we select $\psi(\varphi)$ as an eigenstate of $\hat{L}_z$ with eigenvalue $\hbar/2$, *i.e.*,

$$\hat{L}_z\psi_{\frac{1}{2}}(\varphi) = \frac{1}{2}\hbar\psi_{\frac{1}{2}}(\varphi) \ ,$$

and let us performe a rotation on $\psi_{\frac{1}{2}}$ of an angle of 2π:

$$\begin{aligned} \underbrace{\hat{U}(2\pi)\psi_{\frac{1}{2}}(\varphi)}_{\psi_{\frac{1}{2}}(\varphi+2\pi)} &= \mathrm{e}^{-\frac{i}{\hbar}\hat{L}_z 2\pi}\,\psi_{\frac{1}{2}}(\varphi) \\ &= \mathrm{e}^{-\frac{i}{\hbar}2\pi\frac{1}{2}\hbar}\,\psi_{\frac{1}{2}}(\varphi) \\ &= -\psi_{\frac{1}{2}}(\varphi) \ . \end{aligned}$$

We expect to come back to the same state instead this does not happen. There are two possible ways of overcome this difficulty: either $\psi_{\frac{1}{2}}(\varphi)$ cannot be an eigenstate

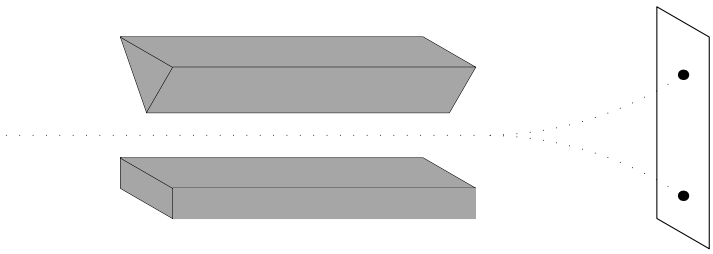


FIG. 7.1. Illustration of the Stern-Gerlach experiment.

of $\hat{L}_z$ or $\hat{L}_z$ is not the operator of rotation. Both are absurd conclusions. Thus there must be a different interpretation of the half-integer eigenvalues.

In 1922 Stern and Gerlach performed a well-known experiment in which a beam of atoms having only one valence electron was sent through a strongly inhomogeneus magnetic field like in Fig. 7.1. What did Stern and Gerlach find? The beam splitted into two beams!

Such results had been already explained within the framework of Bohr model by Uhlenbeck and Goudsmit by assuming that the electron carries an intrinsic angular momentum, called spin, whose components can take only the values $\pm\hbar/2$ and to which it corresponds an intrinsic magnetic moment $\pm\frac{\hbar}{2}\frac{e}{m_0 c}$.

A quantity which is used to measure magnetic moments is the so-called "Bohr magneton" μ_0, defined by

$$\mu_0 \equiv \frac{e\hbar}{2m_0 c} \,. \tag{7.1}$$

When a charged particle (with charge e) rotates with (orbital) angular momentum $m\hbar$ then its magnetic moment due to the orbital angular momentum is $m\mu_0 = \mu$. We will return on this issue when we will talk about the Zeeman effect.

If we measure the angular momentum in units of $\hbar$ and the magnetic moment in units of the Bohr magneton then the ratio

$$\frac{\mu/\mu_0}{L_z/\hbar} = g \,,$$

is called "gyromagnetic ratio" and is equal to 1. On the contrary, for the spin it

comes out that this ratio is 2. This is called "*anomalous gyromagnetic ratio*" and it is explained by the Dirac relativistic equation.

The interaction Hamiltonian between a magnetic moment and a magnetic field is

$$H_I = \boldsymbol{\mu} \cdot \boldsymbol{B} \ . \tag{7.2}$$

Thus, if the magnetic moment is generated by a rotation of a charge, the Hamiltonian is

$$H_I = \boldsymbol{\mu} \cdot \boldsymbol{B} \ ,$$

while in the case in which the magnetic moment is due to spin we get

$$H_I = 2\boldsymbol{\mu} \cdot \boldsymbol{B} \ .$$

In describing the spin we used at the beginning the expression "intrinsic angular momentum" but this is misleading. The electron seems to be a point-like particle at the present energy (7 TeV) and so the picture of something rotating on itself is not appropriate (at least at the present energy). What is more appropriate to say is that the electrons and many other particles behave as small magnets endowed with a magnetic moment which at the quantum level becomes a set of three operators with the same algebra of the one of the angular momentum. But this does not imply that this quantum number is a sort of "angular momentum".

Other quantum observables that you will encounter in particle physics, like the isotopic spin etc., have the same algebra as angular momentum but have nothing to do with it.

7.1 Quantum description of particles carrying spin

Since spin is not a space-time property of a particle, how can we realize the wave function of a particle with spin?

We have seen that the algebra associated to spin is

$$[\hat{s}_i, \hat{s}_j] = i\hbar\varepsilon_{ijk}\hat{s}_k \ , \tag{7.3}$$

s_z can only take the values $\pm\hbar/2$ the number of states in the multiplet is $2l+1=2$. This means the matrices $\hat{s}_i$ are two-by-two matrices. Two-by-two matrices satisfying the algebra (7.3) have been worked out in section 5.1 and take the form

$$\hat{s}_i = \frac{\hbar}{2}\hat{\sigma}_i \ , \tag{7.4}$$

where $\hat{\sigma}_i$ are the so-called Pauli matrices whose form is

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \qquad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ , \qquad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ .$$

Their form can be obtained via simple arguments as follows. $\hat{s}_z$ must have eigenvalues $\pm\hbar/2$, therefore $\hat{\sigma}_z$ must have eigenvalues ± 1 as it is clear from Eqn. (7.4). Since z is an arbitrary direction, and we could choose the x and y directions instead of z, then also $\hat{\sigma}_x$ and $\hat{\sigma}_y$ must have eigenvalues ± 1 . If this is so, their squares have only eigenvalues equal to 1 and thus they must be the identity matrix:

$$\hat{\sigma}_x^2 = \hat{\sigma}_y^2 = \hat{\sigma}_z^2 = \mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \ . \tag{7.5}$$

Now, consider the commutation rule

$$\hat{\sigma}_y\hat{\sigma}_z - \hat{\sigma}_z\hat{\sigma}_y = 2i\hat{\sigma}_x \ .$$

We multiply this expression by $\hat{\sigma}_z$ from the right and we make use of Eqn. (7.5):

$$\hat{\sigma}_y - \hat{\sigma}_z\hat{\sigma}_y\hat{\sigma}_z = 2i\hat{\sigma}_x\hat{\sigma}_z \ .$$

If we multiply instead from the left we get

$$\hat{\sigma}_z\hat{\sigma}_y\hat{\sigma}_z - \hat{\sigma}_y = 2i\hat{\sigma}_z\hat{\sigma}_x \ .$$

By summing the two previous expressions, it follows

$$2i\left(\hat{\sigma}_x\hat{\sigma}_z + \hat{\sigma}_z\hat{\sigma}_x\right) = 0 \ ,$$

i.e.,

$$\{\hat{\sigma}_x, \hat{\sigma}_z\} = 0 \ ,$$

where $\{\cdot, \cdot\}$ denotes the anti-commutator. In the same way it is possible to prove that in general:

$$\{\hat{\sigma}_i, \hat{\sigma}_j\} = \delta_{i,j} \mathbb{1} \ .$$

Using this formula and the commuation rules

$$[\hat{\sigma}_i, \hat{\sigma}_j] = 2i\varepsilon_{ijk}\hat{\sigma}_k \ ,$$

it is not difficult to prove that

$$\hat{\sigma}_r \hat{\sigma}_s = \delta_{r,s}\mathbb{1} + i\varepsilon_{rsk}\hat{\sigma}_k \ ,$$

where, in the last term on the R.H.S, summation over k is understood.

Now, we choose the representation in which $\hat{\sigma}_z$ is diagonal and we keep in mind that its eigenvalues are ± 1. Thus,

$$\hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \Rightarrow \qquad \hat{s}_z = \begin{pmatrix} \frac{\hbar}{2} & 0 \\ 0 & -\frac{\hbar}{2} \end{pmatrix} \ .$$

In order to obtain $\hat{\sigma}_x$ in this representation, we write $\hat{\sigma}_x$ in the form

$$\hat{\sigma}_x = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \ .$$

From the anti-commutation rules we must have:

$$\hat{\sigma}_x \hat{\sigma}_z = -\hat{\sigma}_z \hat{\sigma}_x \ ,$$

so, using $\hat{\sigma}_z = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, we get $a = d = 0$ and finally, from $\hat{\sigma}_x^2 = \mathbb{1}$, we obtain $bc = 1$. In fact,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} a & -b \\ c & -d \end{pmatrix} = - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -a & -b \\ c & d \end{pmatrix} \ ,$$

from which it follows $a = d = 0$, and from $\hat{\sigma}_x^2 = \mathbb{1}$ we get

$$\begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} = \begin{pmatrix} bc & 0 \\ 0 & bc \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \Rightarrow \qquad bc = 1 \; .$$

Since these matrices must be Hermitian, $c^* = b$ and choosing $b = 1$ we get:

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \; .$$

In the same way it is possible to prove that

$$\hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \; .$$

7.2 Pauli equation

If for a particle also spin has to be taken into account, then the Hilbert space of the particle becomes the tensor product of two Hilbert spaces, namely, the Hilbert space of wave functions $\psi(r)$ and the bidimensional Hilbert space of the spin degrees of freedom $\mathcal{H}_{(2)}^{\text{spin}}$:

$$\mathcal{H} = \mathcal{H}(r) \otimes \mathcal{H}_{(2)}^{\text{spin}} \; .$$

Thus,

$$\psi = \begin{pmatrix} \psi^+(x, y, z, t) \\ \psi^-(x, y, z, t) \end{pmatrix} \; .$$

Sometimes, another notation is used to denote this state, namely, $\psi(x, y, z, t; w)$, where w is the spin quantum number. Note that ψ can be written as

$$\psi = \begin{pmatrix} \psi^+ \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \psi^- \end{pmatrix} \; ,$$

where

$$\hat{\sigma}_z \begin{pmatrix} \psi^+ \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \psi^+ \\ 0 \end{pmatrix} = \begin{pmatrix} \psi^+ \\ 0 \end{pmatrix} \; ,$$

i.e., $\binom{\psi^+}{0}$ is eigenstate of $\hat{\sigma}_z$ with eigenvalue $+1$. In the same way one proves that $\binom{0}{\psi^-}$ is eigenstate of $\hat{\sigma}_z$ with eigenvalue -1:

$$\hat{\sigma}_z \begin{pmatrix} 0 \\ \psi^- \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ \psi^- \end{pmatrix} = - \begin{pmatrix} 0 \\ \psi^- \end{pmatrix} .$$

So, $\psi^+(x,y,z)$ is a state of the electron such that we have a probability 1 of finding $+\hbar/2$ in a measurement of $\hat{s}_z$. Therefore, $|\psi^+(x,y,z)|^2$ is the probability density of finding the particle in (x,y,z) with component $+\hbar/2$ of spin along z, while $|\psi^-(x,y,z)|^2$ is the probability density to localize the particle at (x,y,z) and with $-\hbar/2$ component of the spin along z.

The Hamiltonian of a particle with charge $-e$ and intrinsic magnetic moment μ within an electromagnetic field generated by the scalar potential $V(x,y,z)$ and the vector potential $\boldsymbol{A}(x,y,z)$ is given by

$$H = \underbrace{\frac{1}{2m}\left(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}\right)^2 - eV}_{H_0} + \boldsymbol{B}\cdot\boldsymbol{\mu}_s \, , \tag{7.6a}$$

where $\boldsymbol{\mu}_s = \mu_0\boldsymbol{\sigma}$ and so we have already taken account of the anomalous gyromagnetic ratio. H_0 is found by applying the minimal coupling prescription. H can be rewritten in terms of Pauli matrices as follows:

$$H = H_0 + \mu_0 \boldsymbol{B}\cdot\boldsymbol{\sigma} \, , \tag{7.6b}$$

where μ_0 is the Bohr magneton.

The Schrödinger equation for system with spin is called Pauli equation:

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi \, , \tag{7.7}$$

where $\psi = \binom{\psi_+}{\psi_-}$, with the normalization condition:

$$\int_V \left(|\psi_+|^2 + |\psi_-|^2\right) \mathrm{d}V = 1 \, .$$

Eqn. (7.7), rewritten in components, reads

$$\hat{H}_0 \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} + \mu_0 \boldsymbol{B} \cdot \boldsymbol{\sigma} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} .$$

Moreover we have

$$\boldsymbol{B} \cdot \boldsymbol{\sigma} = B_x \hat{\sigma}_x + B_y \hat{\sigma}_y + B_z \hat{\sigma}_z = \begin{pmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{pmatrix} ,$$

and by inserting this in the Pauli equation we get

$$\begin{aligned} \hat{H}_0 \psi_+ + \mu_0 B_z \psi_+ + \mu_0 (B_x - iB_y)\psi_- &= i\hbar \frac{\partial \psi_+}{\partial t} , \\ \hat{H}_0 \psi_- + \mu_0 (B_x + iB_y)\psi_+ - \mu_0 B_z \psi_- &= i\hbar \frac{\partial \psi_-}{\partial t} . \end{aligned}$$

Let us assume that the magnetic field is uniform and let's try to separate the spatial part of the wave function from its spin part. We write

$$\psi = \phi^0(x, y, z)\tilde{\varphi}_s , \qquad \tilde{\varphi}_s = \begin{pmatrix} a \\ b \end{pmatrix} ,$$

where $\phi^0(x, y, z)$ are solutions of the Schrödinger equation associated to $\hat{H}_0$. If we solve the stationary Pauli equation using the ansatz

$$u_n(x, y, z; w) = u_n^0(x, y, z)\tilde{\varphi}_s ,$$

where u_n^0 are eigenfunctions of H_0, the stationary Pauli equation splits into

$$\hat{H}_0 u_n^0(x, y, z) = E_n^0 u_n^0(x, y, z) , \tag{7.8a}$$

$$\mu_0 \boldsymbol{B} \cdot \boldsymbol{\sigma} \tilde{\varphi}_s = E'_s \tilde{\varphi}_s , \tag{7.8b}$$

and the total energy is

$$E_n = E_n^0 + E'_s .$$

By inserting the expression for $\tilde{\varphi}_s = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$ Eqn. (7.8b) becomes

$$\mu_0 B_z a + \mu_0 (B_x - iB_y) b = E'_s a , \tag{7.9a}$$

$$\mu_0 (B_x + iB_y) a - \mu_0 B_z b = E'_s b . \tag{7.9b}$$

The solutions a, b will be different from zero if the following determinant vanishes:

$$\det \begin{vmatrix} \mu_0 B_z - E'_s & \mu_0 (B_x - iB_y) \\ \mu_0 (B_x + iB_y) & -\mu_0 B_z - E'_s \end{vmatrix} = 0 \ ,$$

from which it follows

$$E_s = \pm \mu_0 |B| \ , \qquad |B| = \sqrt{B_x^2 + B_y^2 + B_z^2} \ .$$

By substituting these values into Eqs. (7.9) we find, for $E_s = +\mu_0 B$, the eigenfunctions:

$$\tilde{\varphi}_1 = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \ ,$$

where

$$a_1 = \sqrt{\frac{|B| + B_z}{2\,|B|}}\, \mathrm{e}^{i\alpha} \ , \qquad b_1 = \sqrt{\frac{|B| - B_z}{2\,|B|}}\, \mathrm{e}^{i\left(\alpha + \arctan \frac{B_y}{B_x}\right)} \ ,$$

and α is an arbitary real phase, while for $E_s = -\mu_0 B$ we get

$$\tilde{\varphi}_2 = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \ ,$$

where

$$a_2 = \sqrt{\frac{|B| - B_z}{2\,|B|}}\, \mathrm{e}^{i\alpha} \ , \qquad b_2 = -\sqrt{\frac{|B| + B_z}{2\,|B|}}\, \mathrm{e}^{i\left(\alpha + \arctan \frac{B_y}{B_x}\right)} \ .$$

In particular, if $\boldsymbol{B} = (0, 0, |B|)$, *i.e.*the magnetic field is along z, the solutions are

$$\tilde{\varphi}_1 \sim \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ , \qquad \tilde{\varphi}_2 \sim \begin{pmatrix} 0 \\ -1 \end{pmatrix} \ .$$

The corresponding energies will be different: $E'_n = E_n \pm \mu_0 |B|$.

The total angular momentum of the system, if besides spin we have also the orbital angular momentum, is one of the two following values: $j = l \pm \frac{1}{2}$. More properly, j characterizes the total *magnetic* moment, since spin is not actually an angular momentum as we explained at the beginning.

Spin appears not only in the Stern-Gerlach experiment, but also it plays a role in the explanation of the *anomalous* Zeeman effect, which will be presented after having discussed the Hydrogen atom.

7.3 *Singlet and triplet*

Consider two spin-1/2 particles. According to the rules of the sum of operators like angular momentum, we know that the sum can take only the values 1 and 0. Let us now look for the wave functions Ψ_{s,s_z} of the total spin and its z-component.

The form of the wave functions $\psi_{(1,0)}$ and $\psi_{(0,0)}$ can be sought according to one of the two following arguments:

1. Directly from the equation for the eigenfunctions of $\hat{s}^2$ (square of total spin);

2. or by using the shift operators $\hat{s}_\pm$.

We use method 1. As homework solve the exercise using the method 2.

The z-component for $s_{\rm tot} = 1$ can be one of the following three values:

$$s_z = -1, 0, 1 \ ,$$

while in the case $s_{\rm tot} = 0$ the only possibility is $s_z = 0$. Thus we have four states:

$$\psi_{s,s_z} = \psi_{1,1}, \psi_{1,0}, \psi_{1,-1}, \psi_{0,0} \ .$$

Let us now construct the explicit form of these wave functions. The single-particle wave functions are

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}_{(1)} \qquad \text{and} \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{(1)} \ ,$$

and

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}_{(2)} \qquad \text{and} \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{(2)} \ .$$

The wave-function $\psi_{1,1}$ must then be of the form

$$\psi_{1,1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{(1)} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{(2)} \ ,$$

where with $\otimes$ we indicate the tensor product. The total spin operator is

$$\hat{S}_z^{(\mathrm{tot})} = \hat{S}_z^{(1)} + \hat{S}_z^{(2)}$$

the operator $\hat{S}_z^{(1)}$ will act only on the first part of the tensor product above while $\hat{S}_z^{(2)}$ will act only on the second part, and therefore, by putting $\hbar = 1$, we get:

$$\begin{aligned}
\hat{S}_z^{(\mathrm{tot})}\psi_{1,1} &= \left(\hat{S}_z^{(1)} + \hat{S}_z^{(2)}\right)\left[\begin{pmatrix}1\\0\end{pmatrix}_{(1)} \otimes \begin{pmatrix}1\\0\end{pmatrix}_{(2)}\right] \\
&= \left[\frac{1}{2}\begin{pmatrix}1\\0\end{pmatrix}_{(1)} \otimes \begin{pmatrix}1\\0\end{pmatrix}_{(2)} + \frac{1}{2}\begin{pmatrix}1\\0\end{pmatrix}_{(1)} \otimes \begin{pmatrix}1\\0\end{pmatrix}_{(2)}\right] \\
&= 1\left[\begin{pmatrix}1\\0\end{pmatrix}_{(1)} \otimes \begin{pmatrix}1\\0\end{pmatrix}_{(2)}\right] \\
&= 1\psi_{1,1} \, .
\end{aligned}$$

Since s_{tot} can only be zero or one, a wave function with $s_z = 1$, like the one above, can only belong to $s_{\mathrm{tot}} = 1$.

In the same way on can easily construct the eigenstate $\psi_{1,-1}$, which is

$$\psi_{1,-1} = \begin{pmatrix}0\\1\end{pmatrix}_{(1)} \otimes \begin{pmatrix}0\\1\end{pmatrix}_{(2)} \, .$$

Let us now turn to the problem of constructing the state $\psi_{0,0}$. As it has a total $s_z = 0$, by following the previous rules it is easy to see that such a state will be of the form:

$$\psi_{0,0} = \alpha \begin{pmatrix}1\\0\end{pmatrix}_{(1)} \otimes \begin{pmatrix}0\\1\end{pmatrix}_{(2)} + \beta \begin{pmatrix}0\\1\end{pmatrix}_{(1)} \otimes \begin{pmatrix}1\\0\end{pmatrix}_{(2)} \, . \tag{7.10}$$

The coefficients α and β will be determined imposing that $\psi_{0,0}$ satisfies also $s^{(\mathrm{tot})} = 0$, *i.e.*,

$$\left(\hat{S}^{(\mathrm{tot})}\right)^2 \psi_{0,0} = s(s+1)\psi_{0,0} = 0 \, ,$$

from which it follows

$$\langle\psi_{0,0}|\left(\hat{S}^{(\mathrm{tot})}\right)^2|\psi_{0,0}\rangle = 0\ ,$$

and thus

$$\left\|\hat{S}_x^{(\mathrm{tot})}\psi_{0,0}\right\|^2 + \left\|\hat{S}_y^{(\mathrm{tot})}\psi_{0,0}\right\|^2 + \left\|\hat{S}_z^{(\mathrm{tot})}\psi_{0,0}\right\|^2 = 0\ .$$

Since the three terms on the left-hand side are norms, the only way the previous expression will be fulfilled is to require that each of the three terms be individually equal to zero. For instance, let us evaluate

$$\hat{S}_x\psi_{0,0} = \frac{1}{2}\left[\hat{\sigma}_x^{(1)} + \hat{\sigma}_x^{(2)}\right]\left\{\alpha\begin{pmatrix}1\\0\end{pmatrix}_{(1)}\otimes\begin{pmatrix}0\\1\end{pmatrix}_{(2)} + \beta\begin{pmatrix}0\\1\end{pmatrix}_{(1)}\otimes\begin{pmatrix}1\\0\end{pmatrix}_{(2)}\right\}\ .$$

By using $\hat{\sigma}_x = \left(\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right)$ whose action is

$$\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix}\ , \qquad \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}\ ,$$

we get

$$\hat{S}_x\psi_{0,0} = \frac{1}{2}(\alpha+\beta)\left\{\begin{pmatrix}1\\0\end{pmatrix}_{(1)}\otimes\begin{pmatrix}1\\0\end{pmatrix}_{(2)} + \begin{pmatrix}0\\1\end{pmatrix}_{(1)}\otimes\begin{pmatrix}0\\1\end{pmatrix}_{(2)}\right\}\ .$$

For this to be zero we must have $\alpha = -\beta$. So, the *normalized* state $\psi_{0,0}$ is given by

$$\psi_{0,0} = \frac{1}{\sqrt{2}}\left\{\begin{pmatrix}1\\0\end{pmatrix}_{(1)}\otimes\begin{pmatrix}0\\1\end{pmatrix}_{(2)} - \begin{pmatrix}0\\1\end{pmatrix}_{(1)}\otimes\begin{pmatrix}1\\0\end{pmatrix}_{(2)}\right\}\ .$$

Finally, let us construct $\psi_{1,0}$. Also $\psi_{1,0}$ will be of the form (7.10) but it must be orthogonal to $\psi_{0,0}$ since, having different s, they belong to different eigenvalues of $(\hat{S}^{(\mathrm{tot})})^2$ and we know that eigenstates belonging to different eigenvalues of an Hermitian operator are orthogonal. The only normalized state of the form (7.10) and orthogonal to $\psi_{0,0}$ is

$$\psi_{1,0} = \frac{1}{\sqrt{2}}\left\{\begin{pmatrix}1\\0\end{pmatrix}_{(1)}\otimes\begin{pmatrix}0\\1\end{pmatrix}_{(2)} + \begin{pmatrix}0\\1\end{pmatrix}_{(1)}\otimes\begin{pmatrix}1\\0\end{pmatrix}_{(2)}\right\}\ .$$

$\psi_{0,0}$ is referred to as *singlet* and it is an anti-symmetric state with respect to an interchange of the two particles, while $\psi_{1,0}$ is referred to as *triplet* and it is symmetric under the interchange of the two particles. They are called this way because $s^{(tot)} = 1$ has three states while $s^{(tot)} = 0$ has just one.

7.4 Explanation of the Stern-Gerlach experiment

An *uniform* magnetic field $\boldsymbol{B}$ produces a torque $\boldsymbol{\tau}$ on a magnetic dipole $\boldsymbol{\mu}_q$, given by

$$\boldsymbol{\tau} = \boldsymbol{\mu_q} \times \boldsymbol{B} \ ,$$

but it does not produce any force. In order to get a force on a magnetic dipole a *non-uniform* magnetic field is needed. This is analogous to the case of an electric dipole: if the electric field is uniform, both charges q and $-q$ feel exactly the same force (but opposite in direction) and so the net force acting on the dipole is zero. If instead the field is non-uniform one of the two charges , being located in a different position than the other, will experience a force different than the one acting on the other and the net force on the dipole will be non-vanishing. In the case of an electric dipole the force is

$$F = d_z \frac{\partial E}{\partial z} \ ,$$

where $d_z = e\Delta z$ is the electric dipole moment along z. (For simplicity, we are assuming the electric field to vary only in the z-direction.)

A similar force also occurs in the case of a magnetic dipole in a *non-uniform* magnetic field:

$$F = \mu_z \frac{\partial B}{\partial z} \ .$$

In the Stern-Gerlach experiment, electrons have a magnetic moment due to spin, and this magnetic moment can only take two values $\pm\mu_0$ because, taking account of the anomalous gyromagnetic ratio, we have:

$$\hat{\mu}_z = \mu_0 \hat{\sigma}_z \ .$$

The force acting on the electrons will then be

$$F_z = +\mu_0 \frac{\partial B}{\partial z} \qquad \text{or} \qquad F_z = -\mu_0 \frac{\partial B}{\partial z} \ ,$$

according to the components of the spin we are considering. As the forces have opposite directions, they will split the beam into two components. This is the explanation of the Stern-Gerlach experiment.

Chapter 8

IDENTICAL PARTICLES

8.1 General considerations

In this chapter we will deal with systems having two or more *identical* particles. In classical mechanics two particles are said to be identical if the Hamiltonian is invariant under the interchange of the corresponding coordinates of the two particles in the phase space, *i.e.*,

$$H(\varphi^{(1)}, \varphi^{(2)}) = H(\varphi^{(2)}, \varphi^{(1)}) \ .$$

Since in classical mechanics we can speak about trajectories (in phase space), two identical particles can be distinguished when the initial points of their trajectories are different. In fact, the trajectories (in the phase space) of two particles whose starting points are different will never cross, so there is no ambiguity at all in identifying the trajectory of one particle and the one of the other .

Things are different in quantum mechanics. Indeed, in quantum mechanics we are not dealing with points and trajectories in phase space but with states in an Hilbert space, and even if we prepare the initial states of the two particles at $t = 0$ localized one very far away from the other, time-evolution will spread the spatial distribution of such states and at some time in the future there will be some overlapping of the two wave functions. In the region of overlapping we will not be able to say if the wave function belongs to particle 1 or 2. So, the problem of identical particles must be studied in a more complete way in quantum mechanics. We will see in this chapter that such considerations will lead us to new interesting results.

Following the general rules of quantum mechanics, let us prepare the system of two particles by identifying a complete set of commuting operators. Let us denote the operators of this set in the following way:

$$\hat{O}_1(\xi_1,\xi_2)\,,\quad \hat{O}_2(\xi_1,\xi_2)\,,\quad \hat{O}_3(\xi_1,\xi_2)\,,\ \ldots\ ,\hat{O}_n(\xi_1,\xi_2)\,, \tag{8.1}$$

where ξ_1 and ξ_2 label variables corresponding to the particle 1 and 2, respectively.

Two particles are called identical if the set (8.1) is invariant under the exchange of ξ_1 and ξ_2, *i.e.*,

$$\begin{aligned}
\hat{O}_1(\xi_1,\xi_2) &= \hat{O}_1(\xi_2,\xi_1)\,,\\
\hat{O}_2(\xi_1,\xi_2) &= \hat{O}_2(\xi_2,\xi_1)\,,\\
\hat{O}_3(\xi_1,\xi_2) &= \hat{O}_3(\xi_2,\xi_1)\,,\\
&\vdots\\
\hat{O}_n(\xi_1,\xi_2) &= \hat{O}_n(\xi_2,\xi_1)\,.
\end{aligned} \tag{8.2}$$

As a consequence, if the state of the two particles is prepared in the state $\psi(\xi_1,\xi_2)$ then the state prepared using the set of operators with ξ_1 and ξ_2 interchanged will be $\psi(\xi_2,\xi_1)$ and these two wave functions must be equivalent, *i.e.*, they must be equal up to a phase:

$$\psi(\xi_1,\xi_2) = \mathrm{e}^{i\alpha}\,\psi(\xi_2,\xi_1)\,. \tag{8.3}$$

If we interchange again the two particles we get

$$\psi(\xi_1,\xi_2) = \mathrm{e}^{i\alpha}\,\mathrm{e}^{i\alpha}\,\psi(\xi_1,\xi_2)\,,$$

and thus $\mathrm{e}^{2i\alpha}=1$, so $\mathrm{e}^{i\alpha}=\pm 1$. From this it follows

$$\psi(\xi_1,\xi_2) = \pm\psi(\xi_2,\xi_1)\,, \tag{8.4}$$

which means the wave functions corresponding to two identical particles must be either *symmetric* or *anti-symmetric* under the exchange of the two particles.

Let us introduce an operator $\hat{P}$ which exchanges the two particle, i.e which acts in the following way on the wave function of two identical particles:

$$\hat{P}\psi(\xi_1,\xi_2) = \psi(\xi_2,\xi_1)\,.$$

It is clear that

$$\hat{P}^2 = I \ ,$$

thus the eigenvalues of $\hat{P}$ are ± 1. It is also easy to prove that $\hat{P} = \hat{P}^\dagger = \hat{P}^{-1}$ (work it out as an homework). Let $\hat{O}$ be any observable, then the action of $\hat{P}$ on $\hat{O}$ will be

$$\hat{P}^\dagger \hat{O} \hat{P} = \hat{P}^{-1} \hat{O} \hat{P} \ .$$

Clearly,

$$\hat{P}^\dagger \hat{O}(\xi_1, \xi_2) \hat{P} = \hat{O}(\xi_2, \xi_1) \ ,$$

and, since for identical particles $\hat{O}(\xi_1, \xi_2) = \hat{O}(\xi_2, \xi_1)$, we get

$$\hat{P}^\dagger \hat{O}(\xi_1, \xi_2) \hat{P} = \hat{O}(\xi_1, \xi_2) \ ,$$

which is equivalent (after having multiplied by $\hat{P}$ on the left) to:

$$\hat{O}(\xi_1, \xi_2) \hat{P} = \hat{P} \hat{O}(\xi_1, \xi_2) \ ,$$

i.e., $\hat{P}$ commutes with the whole complete set of observables. Therefore, we have two possibilities: either $\hat{P}$ is a function of $\hat{O}_1, \hat{O}_2, \ldots, \hat{O}_n$ or the set is not complete. For identical particles, the latter is what happens: for identical particles, $\hat{P}$ must be added to the initial set of operators in order to get a complete set.

When preparing the initial state, being $\hat{P}$ now part of the complete set of observables, it must be diagonalized together with the other observables of the set. Since its eigenvalues are ± 1, we immediately get symmetric or anti-symmetric states. But which states do we have to select?

The answer to the previous question is provided by quantum field theory through a well-known *theorem* called **spin-statistics** theorem. Since we have not yet studied quantum field theory, we take the content of that theorem as a new *postulate* for QM:

Spin-statistics postulate: Identical particles of half-integer spin must have anti-symmetric wave functions (fermions). Identical particles of integer spin must have symmetric wave functions (bosons).

The basic ingredients of quantum field theory that enable us to prove such result are:

1. Lorentz invariance;

2. locality of fields;

3. existence of a ground state for $\hat{H}$.

What about a system made by one boson and one fermion? The spin-statistics postulate does not apply to this case since the two particles are no longer identical (one of them is a boson, the other is a fermion).

8.2 Pauli exclusion principle

As a consequence of the spin-statistics postulate, two spin one-half particles never happen to be in the same state. In fact, if we denote with ψ_{Q_1} and ψ_{Q_2} the single-particle states of the two particles, where Q_1 and Q_2 are the quantum numbers corresponding to the single-particle wave functions of the first and the second particle respectively, the anti-symmetric wave function of the total system of the two particles is

$$\tilde{\psi}(\xi_1, \xi_2) = \frac{1}{\sqrt{2}} \left[\psi_{Q_1}(\xi_1)\psi_{Q_2}(\xi_2) - \psi_{Q_1}(\xi_2)\psi_{Q_2}(\xi_1)\right] \ . \tag{8.5}$$

If the states of the two particles were the same, *i.e.*, if $Q_1 = Q_2$ (same quantum numbers), we would get

$$\tilde{\psi}(\xi_1, \xi_2) = \frac{1}{\sqrt{2}} \left[\psi_{Q_1}(\xi_1)\psi_{Q_1}(\xi_2) - \psi_{Q_1}(\xi_2)\psi_{Q_1}(\xi_1)\right] = 0 \ ,$$

therefore the probability to find the two particles in the same state vanishes. This result is known as **Pauli exclusion principle**, and it explains various things concerning atomic systems. For example, for each atomic level we can have no more than two electrons, one having $s_z = \frac{1}{2}$ and the other with $s_z = -\frac{1}{2}$, where s_z is the spin z-component. If we had a third electron, this would have to occupy the next atomic level. It is as if the third electron felt a repulsive force from the core of the previous two electrons. The Pauli's exclusion principle is responsible for much of the periodic table of elements. Another consequence of the Pauli exclusion principle is the so-called exchange energy, discussed in the next section.

8.3 Exchange interaction

Even if no spin-dependent potential were included into the Schrödinger equation of two identical fermions, the effect of spin would appear in the energy of some levels via the Pauli's exclusion principle. Indeed, such effect is measurable.

Let us consider two spin-half particles interacting through a central potential $U(|\boldsymbol{r}_1 - \boldsymbol{r}_2|)$, where $\boldsymbol{r}_1$ denotes the vector position of the first particle and $\boldsymbol{r}_2$ that of the second one. The *total* wave-function (in both spin and spatial variables) must be anti-symmetric, because of the spin-statistics postulate. For $s^{(\mathrm{tot})} = 0$ (singlet) we already know, from the previous chapter, that the spin part of the wave-function is anti-symmetric so the spatial (*i.e.*, orbital) wave-function must be symmetric:

$$\phi_{s=0} = \frac{1}{\sqrt{2}} \left[\phi_1(r_1)\phi_2(r_2) + \phi_1(r_2)\phi_2(r_1)\right] \ , \tag{8.6}$$

while for $s = 1$ the spin part is symmetric and so the orbital part must be anti-symmetric:

$$\phi_{s=1} = \frac{1}{\sqrt{2}} \left[\phi_1(r_1)\phi_2(r_2) - \phi_1(r_2)\phi_2(r_1)\right] \ . \tag{8.7}$$

Now, let us calculate the mean value of the potential energy of these two states:

$$\begin{aligned}\int [\phi^*_{s=0} U(r_1 - r_2)\phi_{s=0}]\, \mathrm{d}^3 r_1\, \mathrm{d}^3 r_2 &= A + J\ ,\\ \int [\phi^*_{s=1} U(r_1 - r_2)\phi_{s=1}]\, \mathrm{d}^3 r_1\, \mathrm{d}^3 r_2 &= A - J\ ,\end{aligned}$$

where

$$A = \int U \left|\phi_1(r_1)\right|^2 \left|\phi_2(r_2)\right|^2 \mathrm{d}^3 r_1\, \mathrm{d}^3 r_2\ ,$$

and

$$J = \int U \phi_1(r_1)\phi_1^*(r_2)\phi_2(r_2)\phi_2^*(r_1)\, \mathrm{d}^3 r_1\, \mathrm{d}^3 r_2\ .$$

The latter integral is referred to as *exchange energy*.

The previous result shows that the singlet and triplet states have different (mean) potential energies. This effect is only due to the spin-statistics postulate, and it is an effect which has been measured.

8.4 Slater's determinant

If we are dealing with N identical spin one-half particles instead of only two, how can we construct the most general anti-symmetric wave-function? The answer is the following:

$$\psi(\xi_1, \xi_2, \ldots, \xi_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{Q_1}(\xi_1) & \psi_{Q_1}(\xi_2) & \cdots & \psi_{Q_1}(\xi_N) \\ \psi_{Q_2}(\xi_1) & \psi_{Q_2}(\xi_2) & \cdots & \psi_{Q_2}(\xi_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{Q_N}(\xi_1) & \psi_{Q_N}(\xi_2) & \cdots & \psi_{Q_N}(\xi_N) \end{vmatrix}\ , \tag{8.8}$$

where we must find at least N different combinations $Q_1, Q_2, \ldots . Q_N$ of single-particle quantum numbers in order to have N different single-particle wave-functions. If we cannot find such N different quantum numbers, $\psi(\xi_1, \xi_2, \ldots, \xi_N)$ is obviously zero. The wave-function above is known as "*Slater*" determinant.

8.5 Entangled states

Let us suppose we have two particles 1 and 2 and the corresponding Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$. The states of the system of the two particles belong to the tensor product of the two one-particle Hilbert spaces:

$$\mathcal{H}^{(\text{tot})} = \mathcal{H}_1 \otimes \mathcal{H}_2 \ .$$

Some wave-functions belonging to this Hilbert space will be of the form:

$$\tilde{\psi}^{(\text{tot})}(1,2) = \psi_1 \otimes \psi_2$$

but not all of them. In fact in the Hilbert space $\mathcal{H}^{(\text{tot})}$ we can also take linear combinations of such wave-functions such as:

$$\tilde{\tilde{\psi}}^{(\text{tot})}(1,2) = \alpha\psi_1 \otimes \psi_2 + \beta\tilde{\psi}_1 \otimes \tilde{\psi}_2 \ .$$

In general this is not be the tensor product of two states belonging to the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, *i.e.*, in general

$$\tilde{\tilde{\psi}}^{(\text{tot})}(1,2) = \alpha\psi_1 \otimes \psi_2 + \beta\tilde{\psi}_1 \otimes \tilde{\psi}_2 \neq \Phi(1) \otimes \Phi(2) \ .$$

If it is not the product , then the state of the two particles is referred to as an *entangled* state.

Looking at the state above, we see that we are not able to answer the question in which state is one of the two particles: particle 1 could be in ψ_1 or in $\tilde{\psi}_1$ and analogously for particle 2. The entangled states are one of the "*most peculiar feature*" of Quantum Mechanics, as Schrödinger pointed out in 1935. They are at the origin of the many *non-local* features of Quantum Mechanics.

Chapter 9

HYDROGEN ATOM

We have already encountered the hydrogen atom studying central potentials in section 1.7. In that occasion, we have approached the Schrödinger equation by using the method of the separation of variables. The result was that the wave function in spherical coordinates could be written in the form

$$\tilde{\psi}(r,\vartheta,\varphi) = R(r)Y(\vartheta,\varphi)\ . \tag{9.1}$$

The radial function $R(r)$ and the angular function $Y(\vartheta,\varphi)$ satisfy the differential equations

$$\frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) + \frac{2m}{\hbar^2}\left[E - U(r)\right]r^2 = \lambda\ , \tag{9.2a}$$

$$\frac{1}{Y\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial Y}{\partial\vartheta}\right) + \frac{1}{Y\sin^2\vartheta}\frac{\partial^2 Y}{\partial\varphi^2} = -\lambda\ , \tag{9.2b}$$

where λ is some arbitrary constant. The latter equation above now looks familiar, in fact by writing the angular momentum operator $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$ in spherical coordinates we get

$$\hat{L}^2 = -\hbar^2\left[\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial}{\partial\vartheta}\right) + \frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial\varphi^2}\right]\ , \tag{9.3}$$

and Eqn. (9.2b) can be rewritten as

$$\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial Y(\vartheta,\varphi)}{\partial\vartheta}\right) + \frac{1}{\sin^2\vartheta}\frac{\partial^2 Y(\vartheta,\varphi)}{\partial\varphi^2} = -\lambda Y(\vartheta,\varphi)\ ,$$

i.e., as

$$\hat{L}^2 Y(\vartheta,\varphi) = \lambda\hbar^2 Y(\vartheta,\varphi)\ .$$

This means that $Y(\vartheta,\varphi)$ are nothing more that the eigenfunctions of $\hat{L}^2$ and we already know from the general theory of angular momentum that their eigenvalues are of the form $l(l+1)$, therefore

$$\lambda = l(l+1) \ ,$$

and the functions $Y(\vartheta,\varphi)$ are the spherical harmonics $Y_{lm}(\vartheta,\varphi)$.

Inserting the value above for λ, Eqn. (9.2a) becomes

$$\frac{1}{R(r)}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R(r)}{\mathrm{d}r}\right) + \frac{2m}{\hbar^2}\left[E - U(r)\right]r^2 = l(l+1) \ .$$

If we parametrize $R(r)$ as

$$R(r) = \frac{\chi(r)}{r} \ ,$$

the $\chi(r)$ satisfies the equation:

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\chi(r)}{\mathrm{d}r^2} + \left[U(r) + \frac{l(l+1)\hbar^2}{2mr^2}\right]\chi(r) = E\chi(r) \ . \tag{9.4}$$

Note that this looks like a one-dimensional Schrödinger equation in the coordinate r for a potential given by $U(r)$ plus a centrifugal term $\frac{l(l+1)\hbar^2}{2mr^2}$.

The original Hamiltonian of our system, having a potential which depends only on r, is invariant under rotations and thus commutes with both $\hat{L}^2$ and $\hat{L}_z$. It is therefore possible to diagonalize simultaneously each of these three operators, $\hat{H}$, $\hat{L}^2$, $\hat{L}_z$:

$$\begin{cases} \hat{H}\psi &= E\psi \ , \\ \hat{L}^2\psi &= [l(l+1)]\hbar^2\psi \ , \\ \hat{L}_z\psi &= m\hbar\psi \ . \end{cases} \tag{9.5}$$

ψ can then be labelled by the three numbers: (E,l,m). The last two equations in (9.5) are those having as solution the spherical harmonics $Y_{lm}(\vartheta,\varphi)$.

Let us now turn back to the radial equation, *i.e.*, Eqn. (9.4). The wave function normalization condition reads

$$\int_0^{+\infty}\mathrm{d}r r^2\int_0^{\pi}\mathrm{d}\vartheta\sin\vartheta\int_0^{2\pi}\mathrm{d}\varphi\left|\tilde{\psi}(r,\vartheta,\varphi)\right|^2 < \infty \ .$$

which, once the wave function is written in the factorized form

$$\tilde{\psi}(r,\vartheta,\varphi) = R(r)Y(\vartheta,\varphi) \ ,$$

and taking into account the normalization of the spherical harmonics, yields

$$\int_0^\infty \mathrm{d}r r^2 \left|R(r)\right|^2 < \infty \ .$$

Let us remember that we wrote $R(r) = \frac{\chi(r)}{r}$, so from the above relation we get

$$\int_0^\infty \mathrm{d}r \left(\chi(r)\right)^2 < \infty \ ,$$

and we also gather that $\chi(r)$ must go to zero as $r \to 0$ at least as fast as r:

$$\chi(r) \overset{r\to 0}{\to} \mathcal{O}(r) \ .$$

Using $\chi(r)$ we can say that solving our problem is like solving the one-dimensional problem of a particle which feels the potential

$$U_{\text{eff}}(r) = U(r) + \frac{l(l+1)\hbar^2}{2mr^2} \ ,$$

for $r > 0$ while for $r < 0$ the potential is that of an infinite potential well, because we are requiring $\chi(r)$ to be zero at $r = 0$.

We now choose the radial equation to be the one of the electron in a hydrogen atoms which feels the Coulomb potential given by

$$U(r) = \frac{-e^2}{r} \ , \tag{9.6}$$

or in general to be the one of an electron spinning around a nucleus of charge Ze:

$$U(r) = \frac{-Ze^2}{r} \ . \tag{9.7}$$

The radial equation in $R(r)$ becomes

$$-\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) - \frac{Ze^2R}{r} + \frac{l(l+1)\hbar^2}{2mr^2}R = ER \ . \tag{9.8}$$

m is not the mass of the electron but its reduced mass:

$$m = \frac{m_e m_N}{m_e + m_N} \approx m_e \ .$$

Now we shall introduce the following dimensionless expressions ρ :

$$\rho \equiv \alpha r \ , \qquad \text{where} \qquad \alpha^2 \equiv \frac{8m\,|E|}{\hbar^2} \ ,$$

and let us denote with $\tilde{\lambda}$ the following quantity:

$$\tilde{\lambda} \equiv \frac{2mZe^2}{\alpha\hbar^2} \equiv \frac{Ze^2}{\hbar}\left(\frac{m}{2\,|E|}\right)^{1/2} \ .$$

The radial equation becomes

$$\frac{1}{\rho^2}\frac{\mathrm{d}}{\mathrm{d}\rho}\left(\rho^2\frac{\mathrm{d}R}{\mathrm{d}\rho}\right) + \left[\frac{\tilde{\lambda}}{\rho} - \frac{1}{4} - \frac{l(l+1)}{\rho^2}\right] R = 0 \ . \tag{9.9}$$

It is easy to prove that for sufficiently large ρ, an $R(\rho)$ of the form $R(\rho) \sim \rho^n\,\mathrm{e}^{\pm\frac{1}{2}\rho}$ satisfies the above equation. In fact, the kinetic term splits into three parts:

$$\left[A\rho^{n-1} + B\rho^n + C\rho^{n-2}\right]\mathrm{e}^{\pm\frac{1}{2}\rho} \tag{9.10}$$

and the same happens for the potential term, thus it is always possible to determine λ and l in such a way to make the kinetic and potential part equal and opposite to each other. Since we are assuming $R(\rho) \to 0$ for $r \to \infty$ in order to have properly normalized wave functions,

$$\int_0^\infty \mathrm{d}r r^2 R^2(r) < \infty \ ,$$

the solution with the plus sign in (9.10) cannot be taken. Therefore, we will seek for an exact solution of the form

$$R(\rho) = F(\rho)\,\mathrm{e}^{-\frac{1}{2}\rho} \ ,$$

where $F(\rho)$ is a polynomial of finite order in ρ.

The equation for $R(\rho)$ becomes one for $F(\rho)$:

$$F''(\rho) + \left(\frac{2}{\rho} - 1\right) F'(\rho) + \left[\frac{\tilde{\lambda} - 1}{\rho} - \frac{l(l+1)}{\rho^2}\right] F(\rho) = 0 \; . \tag{9.11}$$

Let us choose $F(\rho)$ of the form

$$F(\rho) = \rho^s \underbrace{\left[a_0 + a_1\rho + a_2\rho^2 + \ldots\right]}_{L(\rho)} ,$$

with $a_0 \neq 0$ and $s \geq 0$. The condition $s \geq 0$ is needed for $R(\rho)$ to be finite for $\rho = 0$. The equation for the expression $L(\rho)$ becomes

$$\rho^2 L'' + \rho\left[2(s+1) - \rho\right] L' + \Big[\rho(\tilde{\lambda} - s - 1) + s(s+1) - l(l+1)\Big] L = 0 \; . \tag{9.12}$$

If $\rho = 0$, Eqn. (9.12) turns out to be simply:

$$\left[s(s+1) - l(l+1)\right] L = 0 \; . \tag{9.13}$$

L was given by

$$L(\rho) = a_0 + a_1\rho + a_2\rho^2 + \ldots \; .$$

With $a_0 \neq 0$ we get $L(0) = a_0 \neq 0$, and as a consequence Eqn. (9.13) implies $s(s+1) - l(l+1) = 0$ which has the following two solutions:

$$s = l \; , \qquad s = -(l+1) \; .$$

However, the latter, namely $s = -(l+1)$, is not acceptable since $R(\rho)$ would diverge for $\rho \to 0$:

$$R(\rho) = \rho^{-(l+1)} \left[a_0 + a_1\rho + a_2\rho^2 + \ldots\right] \mathrm{e}^{-\frac{1}{2}\rho} \overset{\rho\to 0}{\to} \infty \; .$$

So, the only acceptable solution is $s = l$ and the equation for $L(\rho)$, [Eqn. (9.12)], becomes

$$\rho^2 L'' + \rho\left[2(l+1) - \rho\right] L' + \rho(\tilde{\lambda} - l - 1) L = 0 \; .$$

By inserting the expression for $L(\rho) = a_0 + a_1\rho + a_2\rho^2 + \ldots$ one gets the following recursive relation among the coefficients of $L(\rho)$:

$$a_{\nu+1} = \frac{\nu + l + 1 - \tilde{\lambda}}{(\nu+1)(\nu+2l+2)} a_\nu \ . \tag{9.14}$$

If the series did not stop, its asymptotic behavior would be

$$\frac{a_{\nu+1}}{a_\nu} \sim \frac{1}{\nu+1} \sim \frac{1}{\nu} \ ,$$

which is exactly the kind of behavior which we would have got from a solution of the form $\rho^n \, \mathrm{e}^\rho$, as it can be easily verified by expanding the exponential:

$$\rho^n \, \mathrm{e}^\rho = \rho^n \sum_{\nu=0}^{\infty} \rho^\nu \frac{1}{\nu!} \ ;$$

in fact in this case the coefficients would be

$$a_{\nu+1} = \frac{1}{(\nu+1)!} \ , \quad \text{and} \quad a_\nu = \frac{1}{\nu!} \ .$$

So their ratio is

$$\frac{a_{\nu+1}}{a_\nu} \sim \frac{1}{\nu+1} \sim \frac{1}{\nu} \ ,$$

i.e., of the type found above. But we have already said that a behavior $\rho^n \, \mathrm{e}^\rho$ would result in a radial solution $R(\rho)$ which diverges for $\rho \to \infty$. Thus, the only way out is that the series expansion of $L(\rho)$ terminates. This must happen in order for the wave function not to diverge.

By using the recursive equation (9.14) one has that $a_{\nu+1} = 0$ (and the same for all successive coefficients) when

$$\tilde{\lambda} = \nu + l + 1 \ ,$$

From now on we shall indicate ν with n' since it is an integer and the series terminate at the $(n'+1)^{\text{th}}$ power. Thus the above relation reads

$$\tilde{\lambda} = n' + l + 1 \ .$$

n' is referred to as *radial quantum number* for it is the highest power of the series representation of $L(\rho)$ and thus of the radial wave function. The equation above tells us that also $\tilde{\lambda}$ is an integer number, and we shall denote it with n:

$$n = n' + l + 1 \; .$$

n is called *total* quantum number.

Let us now recall the definition of $\tilde{\lambda}$, namely

$$\tilde{\lambda} = \frac{Ze^2}{\hbar} \left(\frac{m}{2\,|E|} \right)^{1/2} \; ,$$

since $\tilde{\lambda} = n$ we get

$$n = \frac{Ze^2}{\hbar} \left(\frac{m}{2\,|E|} \right)^{1/2} \; ,$$

and by inverting we get the following expression for the energy levels

$$E_n = -\,|E_n| = -\frac{mZ^2e^4}{2\hbar^2}\frac{1}{n^2} \; , \qquad n \geq 1 \; . \tag{9.15}$$

The rule here is that free electrons are at zero energy and bound states have negative energies. (We shall explain later the reason for $n \geq 1$.)

Notice that we have obtained the behavior $E_n \sim 1/n^2$ which is the same behavior experimentalists had discovered. The constant

$$R_H \equiv \frac{mZ^2e^4}{2\hbar^2}$$

is called *Rydberg's constant*, after the name of the experimentalist who first observed the $1/n^2$ behavior. Quantum mechanics has been capable of deriving this constant in terms of more basic quantities.

Let us now look at the wave functions. The radial solution $R(r)$ with energy eigenvalue E_n carries both a label n and a label l since l enters Eqn. (9.9). So we shall denote the R with $R_{nl}(r)$. However, energy levels do not depend on l and this means there is degeneracy. The total wave functions contain also the spherical harmonics $Y_{lm}(\vartheta, \varphi)$ and so there is a dependence also on the quantum number m:

$$\tilde{\psi}_{nlm}(r, \vartheta, \varphi) = R_{nl}(r) Y_{lm}(\vartheta, \varphi) \; ,$$

Besides l, the energy levels do not depend on m, *i.e.*, we get a further degeneracy.

The polynomials $L(\rho)$, solutions of Eqn. (9.12), are called *Laguerre polynomials*. Their general expression is:

$$L_{n+l}^{2l+1}(\rho) = \sum_{k=0}^{n-l-1} (-1)^{k+2l+1} \frac{[(n+l)!]^2 \rho^k}{(n-l-1-k)!(2l+1+k)!k!} \ .$$

In terms of these polynomials the radial wave function $R_{ln}(\rho)$ can be written as

$$R_{ln}(\rho) = \mathrm{e}^{-\frac{1}{2}\rho}\, \rho^l L_{n+l}^{2l+1}(\rho) c_{n,l} \ ,$$

where $c_{n,l}$ are proper normalization coefficients.

Let us now turn back to the expression for the energy levels and let us define the quantity

$$a_0 \equiv \frac{\hbar^2}{me^2} \ .$$

The energy levels can then be written as:

$$E_n = -\frac{Z^2 e^2}{2a_0 n^2} \ .$$

Let us remember that $n = n' + l + 1$ and that n' is a non-negative integer (the maximum degree of the polynomial $L(\rho)$). Since l can be equal to zero at its lowest value, it follows that $n \geq 1$. In particular, the first levels are

$$E_1 = -\frac{Z^2 e^2}{2a_0} \ , \quad E_2 = -\frac{Z^2 e^2}{8a_0} \ , \quad E_3 = -\frac{Z^2 e^2}{18a_0} \ \ldots$$

Notice that their separation becomes smaller and smaller as they get closer to zero. If we take the hydrogen atom ($Z = 1$) and we consider the first energy level $n = 1$, $l = 0$, $m = 0$, we can then calculate the probability density and see where it reaches its maximum:

$$\max \langle \tilde{\psi}_{n=1,l=0,m=0} | \tilde{\psi}_{n=1,l=0,m=0} \rangle \ .$$

It is in $r = a_0$. So a_0 has a specific physical meaning: it is the position at which it is highest the probability of finding the electron and it is called Bohr radius.

9.1 Degeneracy of the energy levels

As we said, the energy levels depend only on n, while the eigenfunctions $\tilde{\psi}$ depend also on l and m, thus there is degeneracy. Let us determine this degeneracy. Remember that $n = n' + l + 1$, where n' is positive. Once n is fixed, we can change both n' and l in a way that does not affect the values of n. In particular, since $n' \geq 0$, l can vary from 0 to $n-1$, *i.e.*, these are all the values l can take once n is fixed. Moreover, for every value of l there are correspondingly $2l+1$ allowed values of m. Therefore, the total number of states sharing the same value of n is

$$\sum_{l=0}^{n-1}(2l+1) = \frac{2n(n-1)}{2} + n = n^2 \ .$$

This is the total degeneracy.

The degeneracy in m is due to rotational invariance along z (or about any generic axis) and it is thus a common feature of all central potentials $U(r)$. On the contrary, degeneracy on l is typical of the Coulomb potential $1/r$. If you consider a valence electron in an atom with many electrons, the potential energy is only approximately of the Coulomb form due to the mutual repulsion forces among the electrons, and the degeneracy in l is no longer present.

This degeneracy in l is due to an extra symmetry present in the $1/r$ potential. It is a symmetry whose associated conserved charge is the Runge-Lenz vector. This vector is the one joining the focus of an ellipse with its perihelion. Using this extra conserved quantity Pauli gave a solution of the Hydrogen atom entirely in terms of operators in the same period during which Schrödinger was working out its solution.

9.2 Form of the eigenfunctions

Let us now discuss the structure of the energy levels as usually given in chemistry and how to translate it in the language of physicists.

The K shell is defined as the one with $n = 1$, and therefore $l = 0$ and $m = 0$. So,

there is only one eigenstate (it is referred to as $1s$ state, where 1 stays for $n = 1$ and s for $l = 0$). Its eigenfunction is given by

$$u_{n=1,l=0,m=0} = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0}\right)^{3/2} \exp\left(-Z\frac{r}{a_0}\right) .$$

There can be two electrons with this radial wave function, one with spin up and the other with spin down.

The L shell is defined as the one having $n = 2$, so l can assume two different values, namely $l = 0$ and $l = 1$. Therefore we have two kind of states: the state $2s$ ($n = 2$ and $l = 0$) whose eigenfunction is

$$u_{n=2,l=0,m=0} = \frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a_0}\right)^{3/2} \left(2 - \frac{Z}{a_0}r\right) \exp\left(-\frac{Zr}{2a_0}\right) ,$$

with two electrons with opposite spin, and the states $2p$ ($n = 2$, $l = 1$), with $m = -1, 0, 1$ and correspondingly we have three different eigenfunctions

$$\begin{aligned}
u_{n=2,l=1,m=-1} &= \frac{1}{8\sqrt{\pi}} \left(\frac{Z}{a_0}\right)^{3/2} \frac{Z}{a_0} r \exp\left(-\frac{Zr}{2a_0}\right) \sin\vartheta \exp\left(-i\varphi\right) , \\
u_{n=2,l=1,m=0} &= \frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a_0}\right)^{3/2} \frac{Z}{a_0} r \exp\left(-\frac{Zr}{2a_0}\right) \cos\vartheta , \\
u_{n=2,l=1,m=1} &= \frac{1}{8\sqrt{\pi}} \left(\frac{Z}{a_0}\right)^{3/2} \frac{Z}{a_0} r \exp\left(-\frac{Zr}{2a_0}\right) \sin\vartheta \exp\left(i\varphi\right) ,
\end{aligned}$$

Thus the L shell can carry eight electrons.

9.3 Zeeman effect

Let us now see how to remove the degeneracies:

- the degeneracy in l is removed by perturbing the original Coulomb potential in such a way that we move from $U(r) = 1/r$ to $\tilde{U}(r) = 1/r + F(r)$ where $F(r)$ is any polynomial in r or $1/r$.

- the degeneracy in m is removed by switching on a magnetic field. For example, the p levels (with $l = 1$) split themselves into three levels with different m and different energies. This is called **Zeeman effect**.

Let us try to explain the Zeeman effect. We know that if there is a magnetic field described through a vector potential $\boldsymbol{A}(x, y, z)$ the Hamiltonian becomes

$$H = \frac{1}{2m_0}\left(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}\right)^2 .$$

This is the so-called "minimal coupling scheme". At the quantum level we get

$$\hat{H} = \frac{1}{2m_0}\left[-i\hbar\frac{\partial}{\partial \boldsymbol{q}} - \frac{e}{c}\boldsymbol{A}(\hat{q})\right]^2 = \frac{1}{2m_0}\left[\hat{\boldsymbol{p}} - \frac{e}{c}\boldsymbol{A}(\hat{q})\right]^2 .$$

Let us recall that $\hat{p}$, being an operator, does not commute with A and thus the above expression becomes

$$\left[\hat{\boldsymbol{p}} - \frac{e}{c}\boldsymbol{A}(\hat{q})\right]^2 = \hat{p}^2 - \frac{e}{c}\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{A}} - \frac{e}{c}\hat{\boldsymbol{A}}\cdot\hat{\boldsymbol{p}} + \frac{e^2}{c^2}\hat{A}^2 .$$

Moreover, since $\hat{p} = -i\hbar\frac{\partial}{\partial q}$, we have

$$\boldsymbol{p}\cdot\boldsymbol{A} - \boldsymbol{A}\cdot\boldsymbol{p} = [\boldsymbol{p}, \boldsymbol{A}] = -i\hbar\boldsymbol{\nabla}\cdot\boldsymbol{A} .$$

Thus,

$$\hat{H} = \frac{\hat{p}^2}{2m_0} - \frac{e}{m_0 c}\hat{\boldsymbol{A}}\cdot\hat{\boldsymbol{p}} + i\hbar\frac{e}{2m_0 c}\boldsymbol{\nabla}\cdot\boldsymbol{A} + \frac{e^2}{2m_0 c^2}\boldsymbol{A}^2 .$$

Let us suppose we have a constant magnetic field along z:

$$\boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A} = |B|\,\hat{z} .$$

A possible choice of the vector potential $\boldsymbol{A}$ that produces such magnetic field is the following one:

$$A_x = -\frac{1}{2}By , \qquad A_y = \frac{1}{2}Bx , \qquad A_z = 0 . \tag{9.16}$$

Notice that, with this choice, $\boldsymbol{\nabla}\cdot\boldsymbol{A} = 0$ and the Hamiltonian in particular becomes

$$\hat{H} = \frac{\hat{p}^2}{2m_0} + \frac{e}{m_0 c}\hat{\boldsymbol{A}}\cdot\hat{\boldsymbol{p}} + \frac{e^2}{2m_0 c^2}\boldsymbol{A}^2 .$$

Now, by inserting Eqs. (9.16) into the right-hand side above we get

$$\hat{H} = \frac{\hat{p}^2}{2m_0} + \frac{eB}{2m_0 c}\left(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x\right) + \frac{e^2}{2m_0 c^2}\boldsymbol{A}^2 .$$

The latter term is order of $\frac{1}{c^2}$ and we shall neglect it. The Hamiltonian becomes

$$\hat{H} = \frac{\hat{p}^2}{2m_0} + \frac{e}{2m_0 c} B \cdot \hat{L}_z \ .$$

where $\hat{L}_z$ is the z component of the angular momentum.

We can also add a central potential (like in the case of the Hydrogen atom):

$$\hat{H} = \underbrace{\frac{\hat{p}^2}{2m_0} + U(r)}_{\hat{H}_r} + \underbrace{\frac{e}{2m_0 c} B \cdot \hat{L}_z}_{\hat{H}_B} \ . \tag{9.17}$$

This Hamiltonian does commute with both $\hat{L}_z$ and $\hat{L}^2$. Moreover, it commutes also with $\hat{H}_r$ alone,

$$\hat{H}_r = \frac{\hat{p}^2}{2m_0} + U(r) \ ,$$

which represents the Hamiltonian of the system when the magnetic field is turned off. So, we are allowed to diagonalize simultaneously the four operators $\hat{H} = \hat{H}_r + \hat{H}_B$, $\hat{H}_r$, $\hat{L}_z$ and $\hat{L}^2$. We get

$$\hat{H}_r \psi_{n,l,m} = E_{n,l} \psi_{n,l,m} \ , \tag{9.18}$$

(in the particular case of the Coulomb potential $E_{n,l}$ does not depend on l). We diagonalize in this base $\hat{H}_B$:

$$\begin{aligned} \hat{H}_B \psi_{n,l,m} &= \tilde{E}_{n,l,m} \psi_{n,l,m} \\ \frac{eB}{2m_0 c} \hat{L}_z \psi_{n,l,m} &= \frac{eB}{2m_0 c} \hbar m \psi_{n,l,m} \ , \end{aligned} \tag{9.19}$$

that is,

$$\tilde{E}_{n,l,m} = \frac{eB\hbar}{2m_0 c} m = m \mu_0 B \ , \qquad \mu_0 = \frac{e\hbar}{2m_0 c} \ , \tag{9.20}$$

where μ_0 is the Bohr's magneton. By combining Eqs. (9.18)–(9.20) we get

$$\hat{H} \psi_{n,l,m} = \left(E_{n,l} + m \mu_0 B \right) \psi_{n,l,m} \ ,$$

thus the new energy levels are

$$E'_{n,l,m} = E_{n,l} + m \mu_0 B \ .$$

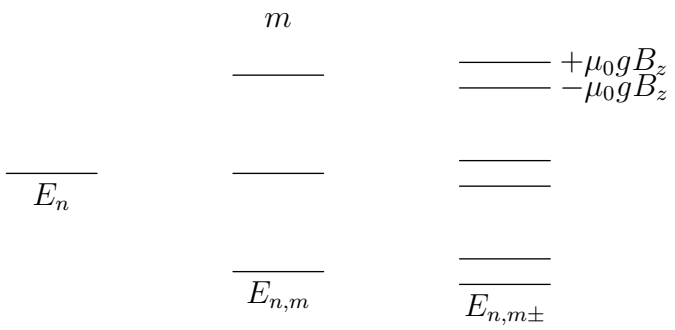


FIG. 9.1.

and they change with m. This effect is known as Zeeman effect.

Indeed, since the electron carries spin (and hence an intrinsic magnetic moment) if we switch on a magnetic field the Hamiltonian will be no longer given by Eqn. (9.17), but by

$$\hat{H} = \hat{H}_r + m\mu_0 B_z + 2\mu_s B_z \ , \tag{9.21}$$

where $\mu_s = \mu_0 s_z$ and $s_z = \pm\frac{1}{2}\hbar$. The factor two in front of the last term in Eqn. (9.21) is the gyromagnetic ratio of the electron. By diagonalizing Eqn. (9.21) one finds that the energy levels split according to the scheme in Fig. 9.1. This effect is called anomalous Zeeman effect.

Homework 9.1. Diagonalize Eqn. (9.21).

9.4 Landau Levels Problem

Look for the eigenfunctions and the energy levels of a charged particle in a constant magnetic field along z. Choose the following vector potential:

$$A_x = 0 \ , \qquad A_y = Bx \ , \qquad A_z = 0 \ . \tag{9.22}$$

Let's start by checking that Eqn. (9.22) gives the correct (*i.e.*, constant) magnetic field:

$$\begin{aligned} B_x &= \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} = 0 \ , \\ B_y &= \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} = 0 \ , \\ B_z &= \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} = B \ . \end{aligned}$$

The Hamiltonian is built according to the minimal coupling rule:

$$\hat{H} = \frac{1}{2m}\left[\hat{p}_x^2 + \left(\hat{p}_y - \frac{e}{c}B\hat{x}\right)^2 + \hat{p}_z^2\right] \ ,$$

The three operators $\hat{H}$, $\hat{p}_y$ and $\hat{p}_z$ do commute among themselves since $\hat{H}$ does not contain $\hat{y}$ and $\hat{z}$. On the contrary, $\hat{H}$ does not commute with $\hat{p}_x$ since $\hat{x}$ appears in the expression of $\hat{H}$.

So, we are allowed to diagonalize simultaneously $\hat{H}$, $\hat{p}_y$ and $\hat{p}_z$. Let us begin with $\hat{p}_y$ and $\hat{p}_z$: we have

$$\hat{p}_y\psi(x,y,z) = p_y\psi(x,y,z) \ ,$$

i.e.,

$$-i\hbar\frac{\partial}{\partial y}\psi(x,y,z) = p_y\psi(x,y,z) \ ,$$

whose solution is

$$\psi_{p_y}(x,y,z) = \mathrm{e}^{ip_y y/\hbar}\,\tilde{\psi}(x,z) \ , \tag{9.23}$$

with eventually a normalization factor $1/\sqrt{2\pi\hbar}$ in front of the right-hand side and $\tilde{\psi}(x,z)$ a function to be determined.

Remark. p_y is a number in Eqn. (9.23), in fact it is the eigenvalue of $\hat{p}_y$. For this reason we have labelled the wave function with p_y.

Now, we use Eqn. (9.23) in diagonalizing $\hat{p}_z$:

$$\hat{p}_z\psi_{p_y}(x,y,z) = p_z\psi_{p_y}(x,y,z) \ ,$$

i.e.,

$$-i\hbar\frac{\partial}{\partial z}\left[\mathrm{e}^{ip_y y/\hbar}\,\tilde{\psi}(x,z)\right] = p_z\,\mathrm{e}^{ip_y y/\hbar}\,\tilde{\psi}(x,z)\ ,$$

which yields

$$-i\hbar\frac{\partial}{\partial z}\tilde{\psi}(x,z) = p_z\tilde{\psi}(x,z)\ ,$$

whose solution is

$$\tilde{\psi}_{p_z}(x,z) = \mathrm{e}^{ip_z z/\hbar}\,\psi(x)\ ,$$

with $\psi(x)$ a function to be determined plus eventually a normalization constant $1/\sqrt{2\pi\hbar}$. Thus,

$$\psi_{p_y,p_z}(x,y,z) = \mathrm{e}^{ip_y y/\hbar}\,\mathrm{e}^{ip_z z/\hbar}\,\psi(x)\ .$$

Now, we use this expression to diagonalize $\hat{H}$:

$$\hat{H}\psi_{E,p_y,p_z}(x,y,z) = E\psi_{E,p_y,p_z}(x,y,z)\ ,$$

which is nothing but

$$\frac{1}{2m}\left[\hat{p}_x^2 + \left(\hat{p}_y - \frac{e}{c}B\hat{x}\right)^2 + \hat{p}_z^2\right]\mathrm{e}^{ip_y y/\hbar}\,\mathrm{e}^{ip_z z/\hbar}\,\psi(x) = E\,\mathrm{e}^{ip_y y/\hbar}\,\mathrm{e}^{ip_z z/\hbar}\,\psi(x)\ .$$

This equation becomes the following one:

$$\frac{1}{2m}\left[\hat{p}_x^2 + \left(p_y - \frac{e}{c}B\hat{x}\right)^2 + p_z^2\right]\psi_E(x) = E\psi_E(x)\ . \tag{9.24}$$

Note that in the equation above p_y and p_z are no longer operators, but numbers, while we are still using $\hat{p}_x$ as operator. We have labelled $\psi(x)$ with the corresponding energy eigenvalue E since the solutions of Eqn. (9.24) depend on E. Let us rewrite Eqn. (9.24) in the following way:

$$-\frac{\hbar^2}{2m}\psi_E''(x) + \frac{1}{2m}\left(p_y - \frac{e}{c}Bx\right)^2\psi_E(x) = \left(E - \frac{p_z^2}{2m}\right)\psi_E(x)\ .$$

Let us introduce the following quantity

$$E_T \equiv E - \frac{p_z^2}{2m}\ ,$$

where T means transverse. E_T is the total energy in the xy-plane. Now, we change variables and we replace x with

$$x' \equiv p_y - \frac{eB}{c}x \ .$$

The derivatives also change according to the chain rule:

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = \left|-\frac{eB}{c}\right|^2 \frac{\mathrm{d}^2\psi}{\mathrm{d}(x')^2} \ .$$

Eqn. (9.24) thus becomes

$$-\frac{\hbar^2}{2m}\left|-\frac{eB}{c}\right|^2 \psi_E''(x') + \frac{1}{2m}(x')^2\psi_E(x') = E_T\psi_E(x') \ .$$

By dividing it by $|-eB/c|^2$ we obtain

$$-\frac{\hbar^2}{2m}\psi_E''(x') + \frac{1}{2m}\left|-\frac{c}{eB}\right|^2 (x')^2\psi_E(x') = \left|-\frac{c}{eB}\right|^2 E_T\psi_E(x') \ . \tag{9.25}$$

Let us now recall the form of the Hamiltonian of an harmonic oscillator:

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}m\omega^2x^2 \ .$$

If we compare this with the left-hand side of Eqn. (9.25) we recognize that Eqn. (9.25) is formally the equation of an harmonic oscillator if we identify

$$m\omega^2 = \frac{1}{m}\left|\frac{c}{eB}\right|^2 \ ,$$

hence

$$\omega = \frac{c}{eBm} \ , \tag{9.26}$$

where we have chosen the positive solution for obvious reasons. So, Eqn. (9.25) is the equation of a quantum harmonic oscillator with frequency given by Eqn. (9.26). The right-hand side of Eqn. (9.25) tells us that the "energy" of this harmonic oscillator is $\left|\frac{c}{eB}\right|^2 E_T$, and knowing the expression for the energy levels of the harmonic oscillator we get

$$\left|\frac{c}{eB}\right|^2 E_T = \left(n + \frac{1}{2}\right)\hbar\omega \ ,$$

where ω is given by Eqn. (9.26). By substituting the expression for E_T, we get

$$\frac{c^2}{e^2B^2}\left(E - \frac{p_z^2}{2m}\right) = \left(n + \frac{1}{2}\right)\hbar\omega \ ,$$

and therefore

$$E_n = \frac{e\hbar B}{mc}\left(n + \frac{1}{2}\right) + \frac{p_z^2}{2m} \ , \tag{9.27}$$

where we labelled the energy eigenvalues with the integer index $n = 0, 1, \ldots$ since n appears on the right-hand side, however they are not discretized energy levels since p_z is continuous. The corresponding energy eigenfunctions are

$$\psi_{E,p_y,p_z}(x', y, z) = \frac{1}{2\pi\hbar}\exp\left[\frac{i}{\hbar}\left(p_y y + p_z z\right)\right]\psi_n(x') \ , \tag{9.28}$$

where ψ_n are the energy eigenfunctions of the one-dimensional quantum harmonic oscillator. Note that while the energy eigenfunctions do depend on p_y, the energy eigenvalues do not. This means there is degeneracy, and in particular an ∞-degeneracy since p_y is a continuous variable. These states are called **Landau levels**. They are localized in x (in fact, in x are Gaussian functions multiplied by Hermite polynomials) but delocalized in y and z where they are simply plane waves.

9.5 Review Problems and Solutions

Problem 9.1. Let's consider a quantum-mechanical system whose Hamiltonian has two eigenvalues $E_+ = \hbar\omega$ and $E_- = -\hbar\omega$ with corresponding (normalized) eigenstates $|+\rangle$ and $|-\rangle$. The system is prepared at the time $t = 0$ in the state

$$|\psi_0\rangle = \frac{1}{2}|+\rangle + \sqrt{\frac{3}{4}}|-\rangle \ . \tag{9.29}$$

Consider also an observable $\hat{O}$ with two eigenvalues a and b and corresponding eigenstates

$$|a\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle + |-\rangle\right) \ , \qquad |b\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle - |-\rangle\right) \ , \tag{9.30}$$

respectively. Consider the two following situations.

1. Let the system evolve from $t = 0$ to $t = 2\tau$ and calculate the probability of finding the value a if a measurement of $\hat{O}$ is performed on the system at $t = 2\tau$.
2. Alternatively, evolve the system until the time $t = \tau$ is reached, then perform a measurement of $\hat{O}$ without reading the outcome (such kind of measurements are called non-selective ones), then let the system evolve again until time 2τ and calculate the probability to obtain a if a measurement of $\hat{O}$ is performed at the time $t = 2\tau$.

Compare the two cases.

SOLUTION. We work in the basis in which the Hamiltonian $\hat{H}$ is diagonal, namely

$$\hat{H} = \begin{pmatrix} \hbar\omega & 0 \\ 0 & -\hbar\omega \end{pmatrix} ,$$

or in an equivalent manner

$$\hat{H} = \hbar\omega \left|+\right\rangle \left\langle+\right| - \hbar\omega \left|-\right\rangle \left\langle-\right| \ .$$

The time-evolution of the state $|\psi_0\rangle$ from $t = 0$ to $t = 2\tau$ is given by:

$$\begin{aligned} |\psi(2\tau)\rangle &= \mathrm{e}^{-\frac{i}{\hbar}\hat{H}2\tau} |\psi_0\rangle \\ &= \mathrm{e}^{-\frac{i}{\hbar}\hat{H}2\tau} \left[\frac{1}{2}|+\rangle + \sqrt{\frac{3}{4}}|-\rangle\right] \\ &= \frac{1}{2}|+\rangle \mathrm{e}^{-2i\omega\tau} + \sqrt{\frac{3}{4}} \mathrm{e}^{2i\omega\tau} |-\rangle \ . \end{aligned}$$

Now we look for the probability of finding a if a measurement of $\hat{O}$ is performed on the state $|\psi(2\tau)\rangle$. According to the basic rules of quantum mechanics, such probability is given

by

$$\begin{aligned} P(\hat{O}=a|2\tau) &= |\langle a|\psi(2\tau)\rangle|^2 \\ &= \left|\frac{1}{\sqrt{2}}\left(\langle +|+\langle -|\right)\left(\frac{1}{2}|+\rangle\, \mathrm{e}^{-2i\omega\tau}+\sqrt{\frac{3}{4}}|-\rangle\, \mathrm{e}^{2i\omega\tau}\right)\right|^2 \\ &= \frac{1}{2}\left[1+\sqrt{\frac{3}{4}}\cos\left(4\omega\tau\right)\right] \ . \end{aligned} \tag{9.31}$$

The notation $P(\hat{O}=a|2\tau)$ denotes the probability of finding the value a for the observable $\hat{O}$ at the time $t=2\tau$. We will use a similar notation also for the other probabilities which we shall calculate in the following.

Now, let us consider the second case. In this case, the system evolves until time $t=\tau$, then at $t=\tau$ a measurement of $\hat{O}$ is performed on the system but we do not read the resulting outcome. Since we do not know what the result of this measurement is, we need to calculate the probabilities of finding both results. The time-evolution of $|\psi_0\rangle$ from $t=0$ to $t=\tau$ is

$$|\psi(\tau)\rangle = \frac{1}{2}|+\rangle\, \mathrm{e}^{-i\omega\tau}+\sqrt{\frac{3}{4}}\,\mathrm{e}^{i\omega\tau}|-\rangle \ .$$

Now, the probability of finding a in a measurement of $\hat{O}$ at $t=\tau$ is

$$P(\hat{O}=a|\tau) = |\langle a|\psi(\tau)\rangle|^2 = \frac{1}{2}\left(1+\sqrt{\frac{3}{4}}\cos 2\omega\tau\right) \ .$$

But we can also have b as possible outcome of the measurement, which implies that after the measurement the state $|\psi(\tau)\rangle$ would have collapsed into the eigenstate $|b\rangle$, that is

$$|b\rangle = \frac{1}{\sqrt{2}}\left[|+\rangle - |-\rangle\right] \ .$$

The probability of finding b is

$$\begin{aligned} P(\hat{O}=b|\tau) &= |\langle b|\psi(\tau)\rangle|^2 \\ &= \left|\frac{1}{\sqrt{2}}\left(|+\rangle - |-\rangle\right)\left(\frac{1}{2}|+\rangle\, \mathrm{e}^{-i\omega\tau}+\sqrt{\frac{3}{4}}|-\rangle\, \mathrm{e}^{i\omega\tau}\right)\right|^2 \\ &= \frac{1}{2}\left(1-\sqrt{\frac{3}{4}}\cos 2\omega\tau\right) \ . \end{aligned}$$

Now, suppose we have performed a measurement of $\hat{O}$ and we have obtained a, then the state would become

$$|a\rangle = \frac{1}{\sqrt{2}} \left(|+\rangle + |-\rangle\right) \ ,$$

and this will then evolve from $t = \tau$ to $t = 2\tau$. At $t = 2\tau$ the state will be:

$$|a, 2\tau\rangle = \mathrm{e}^{-i\hat{H}\tau/\hbar} |a\rangle = \frac{1}{\sqrt{2}} \left(\mathrm{e}^{-i\omega\tau} |+\rangle + \mathrm{e}^{i\omega\tau} |-\rangle\right) \ .$$

If instead b had been obtained at $t = \tau$, the time evolution from $t = \tau$ to $t = 2\tau$ will lead at $t = 2\tau$ to the state

$$|b, 2\tau\rangle = \mathrm{e}^{-i\hat{H}\tau/\hbar} |b\rangle = \frac{1}{\sqrt{2}} \left(\mathrm{e}^{-i\omega\tau} |+\rangle - \mathrm{e}^{i\omega\tau} |-\rangle\right) \ .$$

Now, we have all the informations needed to calculate the probability of finding a at the time $t = 2\tau$. We shall denote this probability with $\tilde{P}(\hat{O} = a|2\tau)$. Considering the whole process, this composite probability is given by the sum of products of simpler probabilities and it has the following expression:

$$\tilde{P}(\hat{O} = a|2\tau) = P(\hat{O} = a|\tau) \left|\langle a|a, 2\tau\rangle\right|^2 + P(\hat{O} = b|\tau) \left|\langle a|b, 2\tau\rangle\right|^2 \ .$$

Since

$$|\langle a|b, 2\tau\rangle|^2 = \left| \frac{1}{\sqrt{2}} \left(\langle +| + \langle -|\right) \frac{1}{\sqrt{2}} \left(\mathrm{e}^{-i\omega\tau} |+\rangle - \mathrm{e}^{i\omega\tau} |-\rangle\right) \right|^2 = \sin^2 \omega\tau \ ,$$

and

$$|\langle a|a, 2\tau\rangle|^2 = \left| \frac{1}{\sqrt{2}} \left(\langle +| + \langle -|\right) \frac{1}{\sqrt{2}} \left(\mathrm{e}^{-i\omega\tau} |+\rangle + \mathrm{e}^{i\omega\tau} |-\rangle\right) \right|^2 = \cos^2 \omega\tau \ ,$$

we get

$$\tilde{P}(\hat{O} = a|2\tau) = \frac{1}{2} + \frac{1}{2}\sqrt{\frac{3}{4}} \cos^2 2\omega\tau \ .$$

Note that this result is different from that of the previous case given by Eqn. (9.31). This means that a measurement done at $t = \tau$ can affect the future results even if we do not read its outcomes.

$\star \quad \star \quad \star$

Problem 9.2. A particle which is rotating with uniform velocity in the xy-plane around a fixed point is prepared in the state

$$\psi(\varphi) = \frac{2}{\sqrt{2\pi}} \cos^2 \varphi \ .$$

1. Find the possible outcomes of a measurement of $\hat{M}_z$ and the corresponding probabilities.
2. Calculate the mean value of $\hat{M}_z$.
3. Find the possible outcomes of a measurement of the square of $\hat{M}_z$, *i.e.*, of $\hat{M}_z^2$, and the corresponding probabilities.
4. Calculate the mean value of $\hat{M}_z^2$.
5. Calculate the mean square deviation.

SOLUTION. First of all, we diagonalize $\hat{M}_z$ in order to get its spectrum and so the possible outcomes of a measurement of $\hat{M}_z$. We will use $\hbar = 1$. We have

$$\hat{M}_z = -i \frac{\mathrm{d}}{\mathrm{d}\varphi} \ ,$$

and thus the eigenvalue problem for $\hat{M}_z$ reads

$$-i \frac{\mathrm{d}}{\mathrm{d}\varphi} \psi_m(\varphi) = m \psi_m(\varphi) \ ,$$

whose solutions are

$$\psi_m(\varphi) = A \, \mathrm{e}^{im\varphi} \ ,$$

and in order ψ_m to be single-valued we must have $\psi_m(0) = \psi_m(2\pi)$ and thus m must be integer:

$$m = \pm 0, \pm 1, \pm 2 \ldots$$

A is found by imposing that ψ_m is properly normalized:

$$\|\psi_m\|^2 = A^2 \int_0^{2\pi} \mathrm{e}^{im\varphi} \, \mathrm{e}^{-im\varphi} \, \mathrm{d}\varphi = A^2 \int_0^{2\pi} \mathrm{d}\varphi = 2\pi A^2 = 1 \ ,$$

therefore $A = 1/\sqrt{2\pi}$. The eigenfunctions are thus

$$\begin{aligned}
\psi_0 &= \frac{1}{\sqrt{2\pi}} \,, \\
\psi_1 &= \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{i\varphi} \,, \\
\psi_{-1} &= \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{-i\varphi} \,, \\
\psi_2 &= \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{i2\varphi} \,, \\
\psi_{-2} &= \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{-i2\varphi} \;\; \ldots
\end{aligned} \tag{9.32}$$

The initial state is

$$\psi(\varphi) = \frac{2}{\sqrt{2\pi}} \cos^2 \varphi \,.$$

Since we have to compute probabilities we have to check that this state is actually normalized. Indeed, it is not normalized:

$$\|\psi\|^2 = \int_0^{2\pi} \psi^*(\varphi)\psi(\varphi) \, \mathrm{d}\varphi = \frac{3}{2} \,,$$

hence the normalized initial states is given by

$$\psi(\varphi) = \sqrt{\frac{2}{3}} \frac{2}{\sqrt{2\pi}} \cos^2 \varphi = \sqrt{\frac{4}{3\pi}} \cos^2 \varphi \,.$$

In order to obtain the possible outcomes of a measurement of $\hat{M}_z$ and their corresponding probabilities on the state $\psi(\varphi)$ we need to expand the state $\psi(\varphi)$ on the basis of the eigenfunctions of $\hat{M}_z$ given in Eqn. (9.32), *i.e.*:

$$\psi(\varphi) = \sum_{m=\pm 0}^{\pm\infty} c_m \psi_m(\varphi) \,,$$

where $c_m = \langle \psi_m(\varphi) | \psi(\varphi) \rangle$. In this case, there is an easier way to get the result without having to calculate all these scalar products. In fact, it is sufficient to rewrite the expression for the initial state in the following way:

$$\begin{aligned}
\psi(\varphi) &= \sqrt{\frac{2}{3}} \frac{2}{\sqrt{2\pi}} \cos^2 \varphi \\
&= \sqrt{\frac{2}{3}} \frac{2}{\sqrt{2\pi}} \left(\frac{1 + \cos 2\varphi}{2} \right) \\
&= \sqrt{\frac{2}{3}} \underbrace{\frac{1}{\sqrt{2\pi}}}_{\psi_0} + \frac{1}{\sqrt{6}} \underbrace{\frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{2i\varphi}}_{\psi_2} + \frac{1}{\sqrt{6}} \underbrace{\frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{-2i\varphi}}_{\psi_{-2}} \,.
\end{aligned}$$

Since our state contains only ψ_0, ψ_2 and ψ_{-2} the only possible outcomes of a measurement of $\hat{M}_z$ are $m = 0$, $m = \pm 2$. The corresponding probabilities, *i.e.*, $|c_m|^2$, can be read directly from the above expression of $\psi(\varphi)$ and in particular they are given by

$$P_0 = \left|\sqrt{\frac{2}{3}}\right|^2 = \frac{2}{3}\,, \qquad P_2 = \left|\frac{\sqrt{1}}{\sqrt{6}}\right|^2 = \frac{1}{6}\,, \qquad P_{-2} = \left|\frac{\sqrt{1}}{\sqrt{6}}\right|^2 = \frac{1}{6}\,.$$

The sum of these probabilities is

$$P_0 + P_2 + P_{-2} = 1\,,$$

as it should be.

The mean value of $\hat{M}_z$ is

$$\langle \hat{M}_z \rangle = \sum_m m P_m = 0 \times \frac{2}{3} + 2 \times \frac{1}{6} + (-2) \times \frac{1}{6} = 0\,.$$

The possible outcomes of a measurement of $\hat{M}_z^2$ are its eigenvalues and since for $\hat{M}_z$ they were $0, 2, -2$, for $\hat{M}_z^2$ we will have $0, 4, 4$:

$$\begin{aligned}
\hat{M}_z^2 \varphi_0 &= 0\,, \\
\hat{M}_z^2 \varphi_2 &= 4\varphi_2\,, \\
\hat{M}_z^2 \varphi_{-2} &= \hat{M}_z \hat{M}_z \varphi_{-2} = -2\hat{M}_z \varphi_{-2} = +4\varphi_{-2}\,.
\end{aligned}$$

Thus, the possible outcomes are only 0 and 4. The corresponding probabilities will be given by

$$P_0 = \left|\sqrt{\frac{2}{3}}\right|^2 = \frac{2}{3}\,, \qquad P_4 = P_2 + P_{-2} = \left|\frac{1}{\sqrt{6}}\right|^2 + \left|\frac{1}{\sqrt{6}}\right|^2 = \frac{1}{3}\,.$$

The mean value of $\hat{M}_z^2$ is

$$\langle \hat{M}_z^2 \rangle = \sum_m m^2 P_m = 0 + 4 \times \frac{1}{3} = \frac{4}{3}\,.$$

The mean square deviation is

$$\Delta M_z = \sqrt{\langle \hat{M}_z^2 \rangle - \langle \hat{M}_z \rangle^2}\,,$$

and since $\langle M_z \rangle = 0$ we get

$$\Delta M_z = \sqrt{\frac{4}{3}} = \frac{2}{\sqrt{3}}\,.$$

$\star \quad \star \quad \star$

Problem 9.3. The state of a quantum particle is given by

$$\psi(x) = C \exp\left[\frac{i}{\hbar} p_0 x\right] \phi(x) \ ,$$

with $\phi(x)$ real. Calculate the probability that a measurement of the momentum gives $3p_0$ as outcome.

SOLUTION. First of all, $\psi(x)$ needs to be normalized:

$$\|\psi\|^2 = \int_{-\infty}^{+\infty} \psi^*(x)\psi(x)\,\mathrm{d}x = 1 = \int_{-\infty}^{+\infty} |C|^2\,|\phi|^2\,\mathrm{d}x \ ,$$

hence

$$|C|^2 = \left\{\int_{-\infty}^{+\infty} |\phi(x)|^2\,\mathrm{d}x\right\}^{-1} \ .$$

Next we expand $\psi(x)$ on the basis of the eigenstates of the momentum operator $\hat{p}$, which are given by plane waves $\frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{ipx/\hbar}$:

$$\psi(x) = \int_{-\infty}^{+\infty} \tilde{c}(p) \frac{\mathrm{e}^{ipx/\hbar}}{\sqrt{2\pi\hbar}}\,\mathrm{d}p \ ,$$

and $|\tilde{c}(p)|^2 = P(p)$ is the probability density of finding p as outcome of a measurement of $\hat{p}$. Multiplying by $\mathrm{e}^{-ip'x/\hbar}/\sqrt{2\pi\hbar}$ and integrating with respect to x we get

$$\begin{aligned}\int_{-\infty}^{+\infty} \psi(x) \frac{\mathrm{e}^{-ip'x/\hbar}}{\sqrt{2\pi\hbar}}\,\mathrm{d}x &= \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} \tilde{c}(p)\,\mathrm{e}^{i(p-p')x/\hbar}\,\mathrm{d}p\,\mathrm{d}x \\ &= \int_{-\infty}^{+\infty} \tilde{c}(p)\delta\left(p-p'\right)\mathrm{d}p \\ &= \tilde{c}(p') \ .\end{aligned}$$

In the former manipulations, we have made use of the relation

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{i(k-k')x}\,\mathrm{d}x = \delta(k-k') \ .$$

We have found therefore that

$$\begin{aligned}\tilde{c}(p') &= \int_{-\infty}^{+\infty} \psi(x) \frac{\mathrm{e}^{-i\frac{p'x}{\hbar}}}{\sqrt{2\pi\hbar}}\,\mathrm{d}x \\ &= \int_{-\infty}^{+\infty} C\phi(x) \frac{\mathrm{e}^{i\frac{p_0 x}{\hbar}}}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{-i\frac{p'x}{\hbar}}\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int C\phi(x)\,\mathrm{e}^{i\frac{(p_0-p')x}{\hbar}}\,\mathrm{d}x\ .\end{aligned}$$

Now, if we want the probability of finding $3p_0$ in measuring $\hat{p}$ we have simply to put $p' = 3p_0$ in the above relation:

$$\tilde{c}(3p_0) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} C\phi(x)\,\mathrm{e}^{-i2\frac{p_0 x}{\hbar}}\,\mathrm{d}x\ ,$$

thus (since $\phi(x)$ is real)

$$\begin{aligned}P(3p_0) = |\tilde{c}(3p_0)|^2 &= \frac{1}{2\pi\hbar}\,|C|^2 \int_{-\infty}^{+\infty} \phi(x)\,\mathrm{e}^{-2i\frac{p_0 x}{\hbar}}\,\mathrm{d}x \int_{-\infty}^{+\infty} \phi^\star(x')\,\mathrm{e}^{+2i\frac{p_0 x'}{\hbar}}\,\mathrm{d}x' \\ &= \frac{1}{2\pi\hbar}\,|C|^2 \int \phi(x)\phi^\star(x')\,\mathrm{e}^{i2\frac{p_0(x'-x)}{\hbar}}\,\mathrm{d}x\,\mathrm{d}x'\ ,\end{aligned}$$

where

$$|C|^2 = \frac{1}{\int_{-\infty}^{+\infty} |\phi(x)|^2\,\mathrm{d}x}\ .$$

$\star \quad \star \quad \star$

Problem 9.4 (spin-orbit). In atoms with many electrons there might be interaction between the intrinsic magnetic moment of the electrons and the magnetic field created by their motion. This effect is called spin-orbit interaction, and in particular it might become important for the electrons in the most external shells. The interaction Hamiltonian is

$$\hat{H} = g\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{s}}\ ,$$

where $\boldsymbol{L}$ is the angular momentum operator, $\boldsymbol{s}$ is the spin of the electron and g is a coupling which can have dimensions and it is not the gyromagnetic ratio.

Suppose that an electron is in the p state (*i.e.*, $l = 1$) find the eigenvalues of $\hat{H}$ and their degeneracies.

SOLUTION. Let us define the operator

$$\hat{\boldsymbol{J}} = \hat{\boldsymbol{L}} + \hat{\boldsymbol{s}} \, ,$$

from which it follows

$$\hat{\boldsymbol{J}}^2 = \hat{\boldsymbol{L}}^2 + \hat{\boldsymbol{s}}^2 + \hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{s}} + \hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{L}} = \hat{\boldsymbol{L}}^2 + \hat{\boldsymbol{s}}^2 + 2\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{s}} \, .$$

We get this expression because $\hat{\boldsymbol{L}}$ and $\hat{\boldsymbol{s}}$ commute with each other because they refer to different quantities. From the expression above we obtain

$$\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{s}} = \frac{\hat{\boldsymbol{J}}^2 - \hat{\boldsymbol{L}}^2 - \hat{\boldsymbol{s}}^2}{2} \, ,$$

and thus

$$\hat{H} = \frac{g}{2} \left[\hat{\boldsymbol{J}}^2 - \hat{\boldsymbol{L}}^2 - \hat{\boldsymbol{s}}^2 \right] \, .$$

Let us now show that $\hat{H}$ commutes with $\hat{\boldsymbol{J}}^2$, $\hat{J}_z$, $\hat{\boldsymbol{L}}^2$, $\hat{\boldsymbol{s}}^2$. Let us begin with $\hat{\boldsymbol{J}}^2$. We need to prove that $\hat{J}^2$ commutes with $\hat{L}^2$ and $\hat{s}^2$. Remember that

$$\hat{J}^2 = \hat{L}^2 + \hat{s}^2 + 2\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{s}} \, .$$

Clearly, $\hat{L}^2$ commutes with $\hat{L}^2$ itself, moreover it surely commutes with $\hat{s}^2$ since they are different degrees of freedom. Finally, $\hat{L}^2$ commutes with $2\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{s}}$ because the latter is a scalar under rotations which are generated by the $\hat{L}_i$. So, $\hat{L}^2$ commutes with $\hat{J}^2$. The same is true for $\hat{s}^2$. In fact $\hat{s}^2$ commutes with $\hat{L}^2$ because they refer to different degrees of freedom, of course it commutes with $\hat{s}^2$ and finally it commutes also with $2\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{s}}$. The last follows from the fact that we may choose $\hat{\boldsymbol{L}}$ along z and thus $\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{s}} \sim \hat{s}_z$ and $\hat{s}_z$ commutes with $\hat{s}^2$. Now, let us consider $\hat{J}_z$. Recall that $\hat{J}_z = \hat{L}_z + \hat{s}_z$ and $\hat{L}_z$ and $\hat{s}_s$ commute with $\hat{L}^2$ and $\hat{s}^2$ and also with $\hat{\boldsymbol{L}} \cdot \hat{\boldsymbol{s}}$ because again we can choose $\hat{\boldsymbol{L}}$ along z and $\hat{L}_z \hat{s}_z$ commutes with $\hat{L}_z$

and $\hat{s}_z$. Therefore, since $\hat{J}^2$, $\hat{J}_z$, $\hat{L}^2$, $\hat{s}^2$ commutes with each other, we can diagonalize them simultaneously:

$$\begin{cases} \hat{J}^2 |j,m,l,s\rangle = \hbar^2 j(j+1) |j,m,l,s\rangle \\ \hat{J}_z |j,m,l,s\rangle = \hbar m |j,m,l,s\rangle \\ \hat{L}^2 |j,m,l,s\rangle = \hbar^2 l(l+1) |j,m,l,s\rangle \\ \hat{s}^2 |j,m,l,s\rangle = \hbar^2 s(s+1) |j,m,l,s\rangle \end{cases} . \tag{9.33}$$

We know that $l = 1$ and $s = 1/2$, so

$$|l-s| = \frac{1}{2} \leq j \leq \frac{3}{2} = l+s \ ,$$

that is, we can have $j = 1/2$ or $j = 3/2$. Now, we diagonalize $\hat{H}$ in the basis of Eqn. (9.33) because all these operators commute with $\hat{H}$:

$$\begin{aligned} \hat{H} |j,m,l,s\rangle &= \frac{g}{2} \left[\hat{J}^2 - \hat{L}^2 - \hat{s}^2 \right] |j,m,l,s\rangle \\ &= \hbar^2 \frac{g}{2} \left[j(j+1) - l(l+1) - s(s+1) \right] |j,m,l,s\rangle \ . \end{aligned}$$

Inserting $l = 1$ and $s = 1/2$ we get

$$\begin{aligned} \hat{H} |j,m,1,1/2\rangle &= \frac{\hbar^2 g}{2} \left[j(j+1) - 2 - \frac{3}{4} \right] |j,m,1,1/2\rangle \\ &= \frac{\hbar^2 g}{2} \left[j(j+1) - \frac{11}{4} \right] |j,m,1,1/2\rangle \ , \end{aligned}$$

therefore the energy is

$$E_{j,1,1/2} = \frac{g\hbar^2}{2} \left[j(j+1) - \frac{11}{4} \right] \ ,$$

that is, it depends on j, l and s in general but not on m and thus m signals a degeneracy. If we choose $j = 3/2$, we get:

$$E_{3/2,1,1/2} = \frac{g\hbar^2}{2} \left[\frac{3}{2}\frac{5}{2} - \frac{11}{4} \right] = \frac{g\hbar^2}{2} \ .$$

There are $2j+1 = 2\frac{3}{2}+1 = 4$ eigenstates associated with this energy and in particular they have the following form:

$$\left.\begin{array}{l} |3/2,3/2,1,1/2\rangle \\ |3/2,1/2,1,1/2\rangle \\ |3/2,-1/2,1,1/2\rangle \\ |3/2,-3/2,1,1/2\rangle \end{array}\right\} \quad \Rightarrow \quad E_{3/2,1,1/2} \ .$$

The other value of the energy is obtained by taking $j = 1/2$ and it is given by

$$E_{1/2,1,1/2} = \frac{g\hbar^2}{2}\left[\frac{1}{2}\frac{3}{2} - \frac{11}{4}\right] = -g\hbar^2 \,,$$

the degeneracy is $2j + 1 = 2\frac{1}{2} + 1 = 2$ and the associated eigenstates are

$$\left.\begin{array}{l} |1/2, 1/2, 1, 1/2\rangle \\ |1/2, -1/2, 1, 1/2\rangle \end{array}\right\} \quad \Rightarrow \quad E_{1/2,1,1/2} \,.$$

$$\star \quad \star \quad \star$$

Problem 9.5. At the time $t = 0$ an harmonic oscillator is prepared in the state

$$|\psi_0\rangle = \left(2\hat{a}^\dagger + 1\right)|0\rangle \ ,$$

where $\hat{a}$ and $\hat{a}^\dagger$ are defined in the following way

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} + i\frac{1}{\sqrt{2m\omega\hbar}}\hat{p} \,,$$
$$\hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} - i\frac{1}{\sqrt{2m\omega\hbar}}\hat{p} \,.$$

Calculate the mean values of position and momentum operators as functions of time.

SOLUTION. Let us recall that the Hamiltonian has the form

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \ ,$$

and our initial state is

$$|\psi_0\rangle = 2\,|1\rangle + |0\rangle \ ,$$

where $|1\rangle = \hat{a}^\dagger\,|0\rangle$. Its norm is

$$\langle\psi_0|\psi_0\rangle = \left(\langle 0| + \langle 1|\,2\right)\left(2\,|1\rangle + |0\rangle\right) = 5 \,,$$

thus the normalized state is

$$|\psi_0\rangle = \frac{1}{\sqrt{5}}\left[2\,|1\rangle + |0\rangle\right] \ .$$

The evolution in time gives

$$\begin{aligned} |\psi\rangle_t &= \mathrm{e}^{-i\hat{H}t/\hbar}\,\frac{1}{\sqrt{5}}\left[2\,|1\rangle + |0\rangle\right] \\ &= \frac{1}{\sqrt{5}}2\,\mathrm{e}^{-iE_1t/\hbar}\,|1\rangle + \frac{1}{\sqrt{5}}\,\mathrm{e}^{-iE_0t/\hbar}\,|0\rangle \ . \end{aligned}$$

Since

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

the state at the time t can be written as

$$|\psi\rangle_t = \frac{2}{\sqrt{5}}\,\mathrm{e}^{-i\frac{3}{2}t\omega}\,|1\rangle + \frac{1}{\sqrt{5}}\,\mathrm{e}^{-i\omega t/2}\,|0\rangle \ .$$

The mean values of $\hat{x}$ and $\hat{p}$ can be computed by inverting the expressions of $\hat{a}$ and $\hat{a}^\dagger$ in terms of $\hat{x}$ and $\hat{p}$, *i.e.*,

$$\begin{aligned} \hat{x} &= \sqrt{\frac{\hbar}{2m\omega}}\left(\hat{a}^\dagger + \hat{a}\right) \ , \\ \hat{p} &= i\sqrt{\frac{\hbar m\omega}{2}}\left(\hat{a}^\dagger - \hat{a}\right) \ , \end{aligned}$$

thus

$$\begin{aligned} \overline{X(t)} = \langle\psi(t)|\hat{x}|\psi(t)\rangle = \left[\frac{2}{\sqrt{5}}\,\mathrm{e}^{i\frac{3}{2}t\omega}\,\langle 1| + \frac{1}{\sqrt{5}}\,\mathrm{e}^{i\frac{1}{2}t\omega}\,\langle 0|\right]\left[\sqrt{\frac{\hbar}{2m\omega}}\left(\hat{a} + \hat{a}^\dagger\right)\right] \\ \times\left[\frac{2}{\sqrt{5}}\,\mathrm{e}^{-i\frac{3}{2}t\omega}\,|1\rangle + \frac{1}{\sqrt{5}}\,\mathrm{e}^{-i\frac{1}{2}t\omega}\,|0\rangle\right] \ , \end{aligned}$$

and since

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle \ , \qquad \hat{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle \ ,$$

the previous expression becomes

$$\begin{aligned} \overline{X(t)} &= \sqrt{\frac{\hbar}{2m\omega}}\left[\frac{2}{\sqrt{5}}\,\mathrm{e}^{i\frac{3}{2}t\omega}\,\langle 1| + \frac{1}{\sqrt{5}}\,\mathrm{e}^{i\frac{1}{2}t\omega}\,\langle 0|\right] \\ &\quad\times\left[\frac{2}{\sqrt{5}}\,\mathrm{e}^{-i\frac{3}{2}t\omega}\,|0\rangle + \frac{2\sqrt{2}}{\sqrt{5}}\,\mathrm{e}^{-i\frac{3}{2}t\omega}\,|2\rangle + \frac{1}{\sqrt{5}}\,\mathrm{e}^{-i\frac{1}{2}t\omega}\,|1\rangle\right] \\ &= \sqrt{\frac{\hbar}{2m\omega}}\left[\frac{2}{5}\,\mathrm{e}^{i\frac{2}{2}t\omega} + \frac{2}{5}\,\mathrm{e}^{-i\frac{2}{2}t\omega}\right] \\ &= \frac{4}{5}\sqrt{\frac{\hbar}{2m\omega}}\cos\omega t \ . \end{aligned}$$

In the same way one can find

$$\overline{P(t)} = \langle\psi(t)|\hat{p}|\psi(t)\rangle = \frac{2}{5}\sqrt{2\hbar m\omega}\,\sin\omega t\ .$$

Now, we consider the mean value of the energy:

$$\begin{aligned}
\langle\psi(t)|\hat{H}|\psi(t)\rangle &= \left[\frac{2}{\sqrt{5}}\,\mathrm{e}^{i\frac{3}{2}t\omega}\,\langle 1| + \frac{1}{\sqrt{5}}\,\mathrm{e}^{i\frac{1}{2}t\omega}\,\langle 0|\right]\left[\hbar\omega\left(\hat{N}+\frac{1}{2}\right)\right] \\
&\quad\times\left[\frac{2}{\sqrt{5}}\,\mathrm{e}^{-i\frac{3}{2}t\omega}\,|1\rangle + \frac{1}{\sqrt{5}}\,\mathrm{e}^{-i\frac{1}{2}t\omega}\,|0\rangle\right] \\
&= \left[\frac{2}{\sqrt{5}}\,\mathrm{e}^{i\frac{3}{2}t\omega}\,\langle 1| + \frac{1}{\sqrt{5}}\,\mathrm{e}^{i\frac{1}{2}t\omega}\,\langle 0|\right] \\
&\quad\times\left[\left(1+\frac{1}{2}\right)\frac{2}{\sqrt{5}}\,\mathrm{e}^{-i\frac{3}{2}t\omega}\,|1\rangle + \frac{1}{2}\frac{1}{\sqrt{5}}\,\mathrm{e}^{-i\frac{1}{2}t\omega}\,|0\rangle\right] \\
&= \hbar\omega\left[\frac{2}{\sqrt{5}}\frac{3}{2}\frac{2}{\sqrt{5}} + \frac{1}{\sqrt{5}}\frac{1}{\sqrt{5}}\frac{1}{2}\right] \\
&= \frac{\hbar\omega}{5}\left[6+\frac{1}{2}\right] = \frac{13}{10}\hbar\omega\ .
\end{aligned}$$

$\star \quad \star \quad \star$

Problem 9.6. Consider two particles of spin one-half. Show that the following operator associated with the two particles

$$\hat{S}_{12} = \left[3\frac{\left(\hat{\boldsymbol{\sigma}}_{(1)}\cdot\hat{\boldsymbol{r}}\right)\left(\hat{\boldsymbol{\sigma}}_{(2)}\cdot\hat{\boldsymbol{r}}\right)}{r^2} - \hat{\boldsymbol{\sigma}}_{(1)}\cdot\hat{\boldsymbol{\sigma}}_{(2)}\right]\ , \tag{9.34}$$

can be written as

$$\hat{S}_{12} = 2\left[3\frac{\left(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{r}}\right)^2}{r^2} - \left(\hat{\boldsymbol{s}}\right)^2\right]\ ,$$

where

$$\hat{\boldsymbol{s}} = \frac{1}{2}\left(\hat{\boldsymbol{\sigma}}_{(1)} + \hat{\boldsymbol{\sigma}}_{2}\right)\ ,$$

and the labels 1 and 2 refer to particles 1 or 2.

SOLUTION. Let us begin with the relation

$$\begin{aligned}(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{r}})^2 &= \frac{1}{4}\left[\hat{\boldsymbol{\sigma}}_{(1)}\cdot\hat{\boldsymbol{r}}+\hat{\boldsymbol{\sigma}}_{(2)}\cdot\hat{\boldsymbol{r}}\right]^2\\ &= \frac{1}{4}\left[\left(\hat{\boldsymbol{\sigma}}_{(1)}\cdot\hat{\boldsymbol{r}}\right)^2+\left(\hat{\boldsymbol{\sigma}}_{(2)}\cdot\hat{\boldsymbol{r}}\right)^2+2\left(\hat{\boldsymbol{\sigma}}_{(1)}\cdot\hat{\boldsymbol{r}}\right)\left(\hat{\boldsymbol{\sigma}}_{(2)}\cdot\hat{\boldsymbol{r}}\right)\right]\ ,\end{aligned}$$

where

$$(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{r}})^2=\hat{\sigma}_x^2x^2+\hat{\sigma}_y^2y^2+\hat{\sigma}_z^2z^2\ .$$

Since $\hat{\sigma}_x^2=\hat{\sigma}_y^2=\hat{\sigma}_z^2=\mathbb{1}$, we have that

$$(\hat{\boldsymbol{\sigma}}\cdot\hat{\boldsymbol{r}})^2=r^2\mathbb{1}\ .$$

Using this relation we get

$$\begin{aligned}(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{r}})^2 &= \frac{1}{4}\left[2r^2\mathbb{1}+2\left(\hat{\boldsymbol{\sigma}}_{(1)}\cdot\hat{\boldsymbol{r}}\right)\left(\hat{\boldsymbol{\sigma}}_{(2)}\cdot\hat{\boldsymbol{r}}\right)\right]\\ &= \frac{1}{2}\left[r^2\mathbb{1}+\left(\hat{\boldsymbol{\sigma}}_{(1)}\cdot\hat{\boldsymbol{r}}\right)\left(\hat{\boldsymbol{\sigma}}_{(2)}\cdot\hat{\boldsymbol{r}}\right)\right]\ ,\end{aligned}$$

or equivalently

$$2\left(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{r}}\right)^2-r^2\mathbb{1}=\left(\hat{\boldsymbol{\sigma}}_{(1)}\cdot\hat{\boldsymbol{r}}\right)\left(\hat{\boldsymbol{\sigma}}_{(2)}\cdot\hat{\boldsymbol{r}}\right)\ .$$

Using the L.H.S of the above expression in (9.34), we get

$$\begin{aligned}\hat{S}_{12} &= \left[3\frac{2\left(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{r}}\right)^2-r^2\mathbb{1}}{r^2}-\hat{\boldsymbol{\sigma}}_{(1)}\cdot\hat{\boldsymbol{\sigma}}_{(2)}\right]\\ &= 6\frac{\left(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{r}}\right)^2}{r^2}-3\mathbb{1}-\hat{\boldsymbol{\sigma}}_{(1)}\cdot\hat{\boldsymbol{\sigma}}_{(2)}\ .\end{aligned}$$

Let us now consider the last term on the right-hand side. Remember that

$$\hat{\boldsymbol{s}}=\frac{1}{2}\left(\hat{\boldsymbol{\sigma}}_{(1)}+\hat{\boldsymbol{\sigma}}_{(2)}\right)\ ,$$

thus

$$\begin{aligned}(\hat{\boldsymbol{s}})^2 &= \frac{1}{4}\left[\hat{\boldsymbol{\sigma}}_{(1)}^2+\hat{\boldsymbol{\sigma}}_{(2)}^2+2\hat{\boldsymbol{\sigma}}_{(1)}\cdot\hat{\boldsymbol{\sigma}}_{(2)}\right]\\ &= \frac{1}{4}\left[(3+3)\mathbb{1}+2\hat{\boldsymbol{\sigma}}_{(1)}\cdot\hat{\boldsymbol{\sigma}}_{(2)}\right]\\ &= \frac{3}{2}\mathbb{1}+\frac{1}{2}\hat{\boldsymbol{\sigma}}_{(1)}\cdot\hat{\boldsymbol{\sigma}}_{(2)}\ ,\end{aligned}$$

where we have used the fact that

$$\hat{\boldsymbol{\sigma}}^2_{(1)} = \hat{\sigma}^2_{(1),x} + \hat{\sigma}^2_{(1),y} + \hat{\sigma}^2_{(1),z} = \mathbb{1} + \mathbb{1} + \mathbb{1} = 3\mathbb{1} \ .$$

Therefore,

$$\hat{\boldsymbol{\sigma}}_{(1)} \cdot \hat{\boldsymbol{\sigma}}_{(2)} = 2\,(\hat{\boldsymbol{s}})^2 - 3\mathbb{1} \ ,$$

and so for $\hat{S}_{12}$ we get

$$\hat{S}_{12} = 6\frac{(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{r}})^2}{r^2} - 3\mathbb{1} + 3\mathbb{1} - 2\,(\hat{\boldsymbol{s}})^2 = 6\frac{(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{r}})^2}{r^2} - 2\,(\hat{\boldsymbol{s}})^2 \ ,$$

which is the expression we wanted to derive.

$\star \quad \star \quad \star$

Problem 9.7. Three Hermitian operators $\hat{A}$, $\hat{B}$ and $\hat{C}$ satisfy the following commutation rules:

$$[\hat{A}, \hat{C}] = 0 = [\hat{B}, \hat{C}] \ , \qquad [\hat{A}, \hat{B}] \neq 0 \ .$$

Prove that one of the eigenvalues of $\hat{C}$ must necessarily be degenerate.

SOLUTION. Let ψ_c be an eigenstate of $\hat{C}$ with eigenvalue c. We apply $[\hat{A}, \hat{C}] = 0$ to ψ_c:

$$[\hat{A}, \hat{C}]\psi_c = 0 \ ,$$

thus

$$\hat{A}\hat{C}\psi_c - \hat{C}\hat{A}\psi_c = c\left(\hat{A}\psi_c\right) - \hat{C}\left(\hat{A}\psi_c\right) = 0 \ ,$$

which means that $\hat{A}\psi_c$ is another eigenstate of $\hat{C}$ with the same eigenvalue c of ψ_c. If we assume that c is not degenerate then $\hat{A}\psi_c$ must be proportional to ψ_c:

$$\hat{A}\psi_c = a\psi_c \ .$$

Notice that in this way ψ_c turns out to be eigenstate of $\hat{A}$ with eigenvalue a. The same argument can be applied to $[\hat{B}, \hat{C}]\psi_c = 0$, yielding

$$\hat{B}\psi_c = b\psi_c \ ,$$

where b is the proportionality constant but also the eigenvalue.

Now, we use the fact that $[\hat{A}, \hat{B}] \neq 0$. We apply this commutator to the state ψ_c and we get

$$[\hat{A}, \hat{B}]\psi_c = \hat{A}\hat{B}\psi_c - \hat{B}\hat{A}\psi_c = ab\psi_c - ab\psi_c = 0 \, .$$

Since $\{\psi_c\}$ form a complete set, this result holds for arbitrary states ψ. This contradicts the fact that $[\hat{A}, \hat{B}]\psi \neq 0$. So we have to give up the hypothesis that ψ_c is not degenerate. It has to be given up at least for one of the eigenstates of $\hat{C}$.

$\star \quad \star \quad \star$

Problem 9.8. Consider a particle moving on a circle with radius 1 and angle φ. Its angular momentum is $\hat{P}_\varphi = -i\hbar\frac{\mathrm{d}}{\mathrm{d}\varphi}$. Since the commutator is $[\hat{\varphi}, \hat{P}_\varphi] = i\hbar$, do we have

$$\Delta\varphi\Delta P_\varphi \geq \frac{\hbar}{2} \, , \tag{9.35}$$

or not? Actually the range of φ is 2π, so the square variation must be smaller than 2π, *i.e.*: $\Delta\varphi \leq 2\pi$, and if Eqn. (9.35) were true we expect that it would be impossible to build a packet having ΔP_φ arbitrary small, since such request would imply $\Delta\varphi = \infty$. Comment on this.

Solution. By definition,

$$\left(\Delta\hat{A}_\psi\right)^2 = \langle\psi|(\hat{A} - \langle\hat{A}\rangle)^2|\psi\rangle \ ,$$

where the label ψ on $\Delta\hat{A}_\psi$ is to remember that the square deviation is calculated with respect to the state ψ. If ψ is an eigenstate of $\hat{A}$,

$$\hat{A}\psi_a = a\psi_a \ ,$$

it is easy to show that $\Delta\hat{A}_{\psi_a} = 0$, in fact we have (assuming ψ_a is normalized) :

$$\begin{aligned}
\langle\psi_a|(\hat{A} - \langle\hat{A}\rangle)^2|\psi_a\rangle &= \langle\psi_a|(\hat{A} - \langle\psi_a|\hat{A}|\psi_a\rangle)^2|\psi_a\rangle \\
&= \langle\psi_a|[\hat{A}^2 + (\langle\psi_a|\hat{A}\psi_a\rangle)^2 - 2\hat{A}\,\langle\psi_a|\hat{A}|\psi_a\rangle]|\psi_a\rangle \\
&= \langle\psi_a|\hat{A}^2 + a^2 - 2\hat{A}a|\psi_a\rangle \\
&= \langle\psi_a|a^2 + a^2 - 2a^2|\psi_a\rangle = 0\ .
\end{aligned}$$

So for example the states $\psi = \frac{\exp i\frac{m\varphi}{\hbar}}{\sqrt{2\pi\hbar}}$ have $\Delta\hat{P}_\varphi = 0$. Consider the operator

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}\varphi^2}\ , \qquad 0 \le \varphi \le 2\pi\ ,$$

and let us study its self-adjointness:

$$\langle\psi_1|\hat{H}\psi_2\rangle = \langle\hat{H}\psi_1|\psi_2\rangle - \frac{\hbar^2}{2m}\left[\psi_1^*\psi_2' - (\psi_1^*)'\psi_2\right]_0^{2\pi}\ . \tag{9.36}$$

$\hat{H}$ is Hermitian if the surface term is zero. There is a compact way to rewrite the surface term by introducing the following notation:

$$\Psi_{(1)} \equiv \begin{pmatrix}\psi_1\\ \psi_1'\end{pmatrix}\ , \qquad \Psi_{(2)} \equiv \begin{pmatrix}\psi_2\\ \psi_2'\end{pmatrix}\ .$$

So,

$$\psi_1^*\psi_2' - (\psi_1^*)'\psi_2 = \Psi_{(1)}^\dagger \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\Psi_{(2)}\ ,$$

and as a consequence the surface term in Eqn. (9.36) can be written as

$$-\frac{\hbar^2}{2m}\left[\Psi_{(1)}^\dagger(2\pi)\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\Psi_{(2)}(2\pi) - \Psi_{(1)}^\dagger(0)\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\Psi_{(2)}(0)\right]\ .$$

In order for this term to vanish, let us suppose that

$$\Psi_{1,2}(2\pi) = \hat{A}\Psi_{1,2}(0)\ , \tag{9.37}$$

where $\hat{A}$ is a 2×2 matrix, because $\Psi_{1,2}$ are 2- vectors. By inserting (9.37) in the surface term we get that it is zero if:

$$\hat{A}^\dagger\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\hat{A} = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\ . \tag{9.38}$$

The matrices which satisfy this relation are

$$\hat{A} = \mathrm{e}^{[i\alpha + \hat{\sigma}_2 \beta]} , \qquad \hat{\sigma}_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} , \tag{9.39}$$

where α and β are real numbers.

So, $\hat{H}$ is Hermitian on a space of wave functions which are labelled by two parameters α and β. They are called self-adjoint extensions of $\hat{H}$ because it is a space larger than the space of periodic functions.

If we take $\beta = 0$, the eigenfunctions of $\hat{H}$ are

$$\psi_n^{(\alpha)}(\varphi) = \frac{1}{\sqrt{2\pi}} \exp\left\{ i \left(\frac{\alpha}{2\pi} + n \right) \varphi \right\} , \tag{9.40}$$

and their corresponding eigenvalues are

$$E_n = \frac{\hbar^2}{2m} \left(\frac{\alpha}{2\pi} + n \right)^2 .$$

$\psi_n^{(\alpha)}(\varphi)$ satisfy Eqn. (9.37) with $\hat{A}$ given by Eqn. (9.39). In fact, Eqn. (9.37) and Eqn. (9.39) with $\beta = 0$ are

$$\Psi_1(2\pi) = \mathrm{e}^{i\alpha} \Psi_1(0) ,$$
$$\Psi_2(2\pi) = \mathrm{e}^{i\alpha} \Psi_2(0) .$$

Of course, these functions are not periodic, but only the eigenfunctions of $\hat{L}_z$ must be periodic.

So, we can say that the domain of self-adjointness of $\hat{H}$ is $C^2(0, 2\pi)$ with

$$\psi(2\pi) = \mathrm{e}^{i\alpha} \psi(0) . \tag{9.41}$$

Let us now apply the operator $\hat{\varphi}$ to $\psi(\varphi)$:

$$\hat{\varphi}\psi(\varphi) = \tilde{\psi}(\varphi) .$$

This $\tilde{\psi}$ does no longer fulfill Eqn. (9.41). In fact, we have

$$\tilde{\psi}(2\pi) = 2\pi\psi(2\pi) ,$$
$$\tilde{\psi}(0) = 0 \cdot \psi(0) = 0 ,$$

and so it is not true that $\tilde{\psi}(2\pi) = \mathrm{e}^{i\alpha}\,\tilde{\psi}(0)$. This means that the operator $\hat{\varphi}$ pushes ψ outside the domain of self-adjointness of $\hat{H}$, which is our Hilbert space. A crucial ingredient in the proof of the Heisenberg uncertainty principle is that $\tilde{\psi}(\varphi)$ must belong to the same Hilbert space we started from. As a consequence of this, what we can say is that the uncertainty principle does not hold anymore for the operators $\hat{\varphi}$ and $\hat{P}_\varphi$. This is the reason of the contradiction we have found. Since the uncertainty principle is not valid, we can construct a state with a perfectly well-defined value of $\hat{P}_\varphi$, $\psi(\varphi) = \mathrm{e}^{ip_\varphi\varphi/\hbar}/\sqrt{2\pi}$, even if $\Delta\varphi < 2\pi$. Note that all this goes through also in the case of periodic eigenfunctions.

$\star \quad \star \quad \star$

Problem 9.9. Prove that the expectation value of the force on the stationary states of the Schrödinger equation is always zero.

SOLUTION. What we have to prove is that

$$\langle\psi_E|\hat{F}|\psi_E\rangle = 0\,, \tag{9.42}$$

with

$$\hat{H}\psi_E = E\psi_E\,.$$

Using $\hat{F} = \dot{\hat{p}}$ and the correspondence rule

$$\dot{p} = \{p, H\}_{\text{P.B.}} \quad \Rightarrow \quad \dot{\hat{p}} = \frac{1}{i\hbar}[\hat{p}, \hat{H}]\,,$$

Eqn. (9.42) becomes

$$\begin{aligned}\left\langle \psi_E \left| \frac{1}{i\hbar}[\hat{p}, \hat{H}] \right| \psi_E \right\rangle &= \frac{1}{i\hbar}\left\langle \psi_E \left| \left(\hat{p}\hat{H} - \hat{H}\hat{p}\right) \right| \psi_E \right\rangle \\ &= \frac{E}{i\hbar}\left\{\langle\psi_E|\hat{p}|\psi_E\rangle - \langle\psi_E|\hat{p}|\psi_E\rangle\right\} \\ &= 0\,.\end{aligned}$$

$\star \quad \star \quad \star$

Chapter 10

TIME-INDEPENDENT PERTURBATION THEORY

There is a small number of potentials which are exactly solvable in quantum mechanics. Because of this fact, people have developed techniques to solve other potentials at least in an *approximate* way.

These techniques basically belong to two separate classes: perturbative techniques and non-perturbative ones. Here, we will deal with the first ones.

Let us suppose that the Hamiltonian $\hat{H}$ of our system can be splitted into two parts, say $\hat{H}_0$ and $g\hat{V}$, as follows:

$$\hat{H} = \hat{H}_0 + g\hat{V} \ , \tag{10.1}$$

where we are assuming to know exactly the eigenvalues and the eigenstates of $\hat{H}_0$, which is not necessarily the free Hamiltonian. For the time being, let us assume that the spectrum of $\hat{H}_0$ is a purely discrete one:

$$\hat{H}_0 \, |E_n^{(0)}\rangle = E_n^{(0)} \, |E_n^{(0)}\rangle \ . \tag{10.2}$$

Furthermore, let us assume that

$$\frac{\langle g\hat{V}\rangle}{\langle \hat{H}_0\rangle} \ll 1 \ , \tag{10.3}$$

where $\langle\cdot\rangle$ denotes the mean value with respect to arbitrary states.

Condition Eqn. (10.3) is sometimes expressed in a qualitative fashion by saying that g (the coupling) is somehow "small" and thus the perturbation $g\hat{V}$ with respect to $\hat{H}_0$ is small.

A more precise way of saying that the perturbation is small is the following one:

$$\frac{\langle E_n^{(0)}|g\hat{V}|E_n^{(0)}\rangle}{E_{n+1}-E_n} \ll 1 \ ,$$

that is to say, the shift on the energy levels $E_n^{(0)}$ due to the perturbation $g\hat{V}$ is smaller than the difference between the energy of the n-th level and the next level.

10.1 Non-degenerate case

In this section we deal with the case in which the energy levels of the unperturbed Hamiltonian $\hat{H}_0$ are non degenerate. The cases in which there is degeneracy will be covered in a later section.

Our task is to find an approximate expression of the energy levels E_n and their corresponding eigenstates $|E_n\rangle$ of the complete Hamiltonian $\hat{H}$:

$$\hat{H}\,|E_n\rangle = E_n\,|E_n\rangle \ . \tag{10.4}$$

Obviously, we want to require that, if we switch off the perturbation, then the energy levels get into those of the unperturbed Hamiltonian $\hat{H}_0$, *i.e.*,

$$\lim_{g\to 0} E_n = E_n^{(0)} \ , \qquad \lim_{g\to 0} |E_n\rangle = |E_n^{(0)}\rangle \ .$$

We *assume* that the exact values of E_n and $|E_n\rangle$ can be written as a series in g, since g is small. That is to say,

$$E_n = E_n^{(0)} + g\varepsilon_1 + g^2\varepsilon_2 + \ldots + g^m\varepsilon_m + \ldots \ , \tag{10.5}$$

$$|E_n\rangle = |E_n^{(0)}\rangle + g\,|1\rangle + g^2\,|2\rangle + \ldots + g^m\,|m\rangle + \ldots \ , \tag{10.6}$$

where the quantities ε_1, ε_2, ... and the states $|1\rangle$, $|2\rangle$, ... are to be determined.

We impose

$$\langle E_n|E_n^{(0)}\rangle = 1 \ . \tag{10.7}$$

We will see in a while that it is always possible to meet such requirement even if the eigenstates $|E_n^{(0)}\rangle$ are normalized, *i.e.*,

$$\langle E_n^{(0)}|E_n^{(0)}\rangle = 1 \ . \tag{10.8}$$

In fact, taking into account Eqn. (10.6), the latter requirement is fulfilled if

$$\langle E_n^{(0)}|1\rangle = 0 \ , \quad \langle E_n^{(0)}|2\rangle = 0 \ , \quad \ldots \ , \quad \langle E_n^{(0)}|m\rangle = 0 \ , \ldots \tag{10.9}$$

Using Eqs. (10.5)–(10.6) into Eqn. (10.4) we get

$$\begin{aligned} &\left[\hat{H}_0 + g\hat{V}\right]\left(|E_n^{(0)}\rangle + g\,|1\rangle + g^2\,|2\rangle + \ldots\right) \\ &\qquad = \left[E_n^{(0)} + g\varepsilon_1 + g^2\varepsilon_2 + \ldots\right]\left(|E_n^{(0)}\rangle + g\,|1\rangle + g^2\,|2\rangle + \ldots\right) \ , \end{aligned}$$

By equating terms with the same power of g on the left and right hand side in the previous expression we get the system of equations:

$$\begin{aligned} \left(\hat{H}_0 - E_n^{(0)}\right)|E_n^{(0)}\rangle &= 0 \ , \\ \left(\hat{H}_0 - E_n^{(0)}\right)|1\rangle + \left(\hat{V} - \varepsilon_1\right)|E_n^{(0)}\rangle &= 0 \ , \\ \left(\hat{H}_0 - E_n^{(0)}\right)|2\rangle + \left(\hat{V} - \varepsilon_1\right)|1\rangle - \varepsilon_2\,|E_n^{(0)}\rangle &= 0 \ , \ldots \end{aligned} \tag{10.10}$$

The reason for equating terms sharing the same power of g is that g is a small but otherwise arbitrary parameter.

Now we will see how it is possible to choose the states $|m\rangle$ in such a way to satisfy Eqn. (10.9) and thus Eqn. (10.7) to be fulfilled. The state $|E_n\rangle$ solution of

$$\hat{H}\,|E_n\rangle = E_n\,|E_n\rangle \ , \tag{10.11}$$

is defined up to a constant k:

$$|E_n\rangle' \equiv k\,|E_n\rangle \ ,$$

in fact $|E_n\rangle'$ is also a solution of Eqn. (10.11). We define the value of k in such a way that taking the limit $g \to 0$ we recover the results of the unperturbed Hamiltonian, *i.e.*,

$$\lim_{g\to 0} k\,|E_n\rangle = |E_n^{(0)}\rangle \ .$$

But this is not sufficient, in addition we need also to define the components of $|E_n\rangle$. The only state we know is $|E_n^{(0)}\rangle$ so we require

$$\langle E_n|E_n^{(0)}\rangle = 1 \,,$$

this is not in contraddiction with the assumption

$$\langle E_n^{(0)}|E_n^{(0)}\rangle = 1 \,,$$

since we have already seen in Eqn. (10.9) that it is sufficient to ensure that the states $|1\rangle, |2\rangle, \ldots, |m\rangle, \ldots$ are orthogonal to $|E_n^{(0)}\rangle$.

Is it possible to require such thing? Yes it is, in fact from Eqs. (10.10) we notice that we can add to each state $|1\rangle$, $|2\rangle$, ..., another state proportional to $|E_n^{(0)}\rangle$ and we get again Eqs. (10.10). Let us define

$$|1'\rangle \equiv |1\rangle + k\,|E_n^{(0)}\rangle \;, \tag{10.12}$$

and insert $|1'\rangle$ in the second equation of Eqs. (10.10):

$$\left(\hat{H}_0 - E_n^{(0)}\right)|1'\rangle + \left(\hat{V} - \varepsilon_1\right)|E_n^{(0)}\rangle = 0 \,.$$

Inserting the expression of $|1'\rangle$ yields

$$\left(\hat{H}_0 - E_n^{(0)}\right)|1\rangle + k\underbrace{\left(\hat{H}_0 - E_n^{(0)}\right)|E_n^{(0)}\rangle}_{0} + \left(\hat{V} - \varepsilon_1\right)|E_n^{(0)}\rangle = 0 \,,$$

hence

$$\left(\hat{H}_0 - E_n^{(0)}\right)|1\rangle + \left(\hat{V} - \varepsilon_1\right)|E_n^{(0)}\rangle = 0 \,.$$

Therefore, we can say that solutions of Eqn. (10.10) are defined up to a constant proportional to $|E_n^{(0)}\rangle$. The same reasoning applies to $|2\rangle$ and to the other states.

We can use this freedom to impose

$$\langle E_n^{(0)}|m'\rangle = 0 \quad \Leftrightarrow \quad \langle E_n^{(0)}|m\rangle + k\underbrace{\langle E_n^{(0)}|E_n^{(0)}\rangle}_{1} = 0 \,,$$

i.e., it is sufficient to choose

$$k = -\langle E_n^{(0)}|m\rangle \ . \tag{10.13}$$

This choice of k produces a state $|m'\rangle$ orthogonal to $\langle E_n^{(0)}|$. Of course, k changes with $|m\rangle$ so it would be better to put a label on k:

$$k_m = -\langle E_n^{(0)}|m\rangle \ .$$

Now, let us come back to Eqn. (10.10) and let us try to determine $\varepsilon_1, \varepsilon_2, \ldots$ and the states $|1\rangle, |2\rangle, \ldots$

10.1.1 First order corrections

Let us suppose we know only the state $|E_n^{(0)}\rangle$ and its energy $E_n^{(0)}$.

Consider the second equation in Eqs. (10.10) and braket with $\langle E_n^{(0)}|$:

$$\langle E_n^{(0)}|\hat{H}_0 - E_n^{(0)}|1\rangle + \langle E_n^{(0)}|\hat{V} - \varepsilon_1|E_n^{(0)}\rangle = 0 \ ,$$

so

$$\underbrace{\langle E_n^{(0)}|E_n^{(0)} - E_n^{(0)}|1\rangle}_{0} + \langle E_n^{(0)}|\hat{V}|E_n^{(0)}\rangle - \varepsilon_1 \underbrace{\langle E_n^{(0)}|E_n^{(0)}\rangle}_{1} = 0 \ ,$$

and thus

$$\varepsilon_1 = \langle E_n^{(0)}|\hat{V}|E_n^{(0)}\rangle \ . \tag{10.14}$$

This is the first-order perturbative correction to the energy and it is a function only of known things, namely, $\hat{V}$ and $|E_n^{(0)}\rangle$.

10.1.2 Second order corrections

Consider the third equation in Eqs. (10.10) and bracket it with respect to $\langle E_n^{(0)}|$:

$$\langle E_n^{(0)}|\hat{H}_0 - E_n^{(0)}|2\rangle + \langle E_n^{(0)}|\hat{V} - \varepsilon_1|1\rangle - \varepsilon_2 \underbrace{\langle E_n^{(0)}|E_n^{(0)}\rangle}_{1} = 0 \ ,$$

so

$$\langle E_n^{(0)}|\underbrace{E_n^{(0)} - E_n^{(0)}}_{0}|2\rangle + \langle E_n^{(0)}|\hat{V}|1\rangle - \varepsilon_1 \underbrace{\langle E_n^{(0)}|1\rangle}_{0} - \varepsilon_2 = 0 \ ,$$

from which it follows that

$$\varepsilon_2 = \langle E_n^{(0)}|\hat{V}|1\rangle \ . \tag{10.15}$$

In the same way (do this as an homework) it is possible to prove that

$$\varepsilon_i = \langle E_n^{(0)}|\hat{V}|i-1\rangle \ . \tag{10.16}$$

A key point in the derivation above is the fact that the states $|m\rangle$ can be chosen orthogonal to $\langle E_n^{(0)}|$.

From Eqs. (10.15)–(10.16) we notice that ε_2 can be determined once we know $|1\rangle$ and, more generally, in order to calculate the correction ε_m we should know the state $|m-1\rangle$ and so forth.

10.1.3 Correction to the states

To build the state $|1\rangle$ we should determine its components along a given orthogonal basis. As a basis we consider the one made by the eigenstates of $\hat{H}_0$, namely, $|E_1^{(0)}\rangle, |E_2^{(0)}\rangle, \ldots, |E_i^{(0)}\rangle, \ldots$. Thus, we should build

$$\langle E_i^{(0)}|1\rangle \ ,$$

with $i \neq n$ since, from Eqn. (10.9), we know that $\langle E_n^{(0)}|1\rangle = 0$.

Leu us start from Eqn. (10.10) and project along $\langle E_i^{(0)}|$. From the first equation we get

$$\langle E_i^{(0)}|\hat{H}_0 - E_n^{(0)}|E_n^{(0)}\rangle = 0 \ ,$$

thus

$$\langle E_i^{(0)}|E_i^{(0)} - E_n^{(0)}|E_n^{(0)}\rangle = \left(E_i^{(0)} - E_n^{(0)}\right) \underbrace{\langle E_i^{(0)}|E_n^{(0)}\rangle}_{0} = 0 \ ,$$

thus the first equation is trivially fulfilled; It happens that $\langle E_i^{(0)}|E_n^{(0)}\rangle = 0$ because the basis of the eigenstates of $\hat{H}_0$ is an orthonormal one.

Now, we consider the second equation in Eqn. (10.10) and bracket it with $\langle E_i^0|$

$$\langle E_i^{(0)}|\hat{H}_0 - E_n^{(0)}|1\rangle + \langle E_i^{(0)}|\hat{V} - \varepsilon_1|E_n^{(0)}\rangle = 0 \ ,$$

thus

$$\langle E_i^{(0)}|1\rangle \left(E_i^{(0)} - E_n^{(0)}\right) = -\langle E_i^{(0)}|\hat{V} - \varepsilon_1|E_n^{(0)}\rangle \ ,$$

therefore

$$\langle E_i^{(0)}|1\rangle = \frac{\langle E_i^{(0)}|(\hat{V} - \varepsilon_1)|E_n^{(0)}\rangle}{E_n^{(0)} - E_i^{(0)}} \ .$$

This equation gives the components of the state $|1\rangle$ on the basis of $|E_i^{(0)}\rangle$ and it requires only things which are known , namely, $\langle E_i^{(0)}|$, $\langle E_n^{(0)}|$, $\hat{V}$ and ε_1, which has been already determined through Eqn. (10.14) without having to know anything about $|1\rangle$.

Now we start from the third equation in Eqs. (10.10) and we bracket with $\langle E_i^{(0)}|$:

$$\langle E_i^{(0)}|\hat{H}_0 - E_n^{(0)}|2\rangle + \langle E_i^{(0)}|\hat{V} - \varepsilon_1|1\rangle - \varepsilon_2 \underbrace{\langle E_i^{(0)}|E_n^{(0)}\rangle}_{0} = 0 \ ,$$

thus

$$\left(E_i^{(0)} - E_n^{(0)}\right) \langle E_i^{(0)}|2\rangle = -\langle E_i^{(0)}|\hat{V} - \varepsilon_1|1\rangle \ ,$$

and therefore

$$\langle E_i^{(0)}|2\rangle = \frac{\langle E_i^{(0)}|(\hat{V} - \varepsilon_1)|1\rangle}{E_n^{(0)} - E_i^{(0)}} \ . \tag{10.17}$$

Eqn. (10.17) gives the components of $|2\rangle$ in the basis of the eigenstates $|E_i^{(0)}\rangle$ in terms of given quantities on the right-hand side, namely, once again $\varepsilon_1, \hat{V}, |E_i^{(0)}\rangle$ and $|1\rangle$ which has been already determined in the previous section without needing to know $|2\rangle$.

The procedure seems to be a recursive one. In fact, the general formula for the state $|l\rangle$ is

$$\begin{aligned}\langle E_i^{(0)}|l\rangle = \frac{1}{E_n^{(0)} - E_i^{(0)}} \Big[& \langle E_i^{(0)}|\hat{V} - \varepsilon_1|l-1\rangle - \varepsilon_2 \langle E_i^{(0)}|l-2\rangle \\ & - \ldots - \varepsilon_{l-1} \langle E_i^{(0)}|1\rangle \Big] \ ,\end{aligned} \tag{10.18}$$

with $i \neq n$.

Homework 10.1. Prove Eqn. (10.18).

From Eqn. (10.18) it follows that $|l\rangle$ is completely determined from quantities which are already known, namely, $\varepsilon_1, \ldots, \varepsilon_{l-1}$ and $|1\rangle, \ldots, |l-1\rangle$. Thus, it is a recursive procedure.

At this point, it should be emphasized that the string of states $|1\rangle$, $|2\rangle$, ..., $|l\rangle$, ... depends on which state $|E_n^{(0)}\rangle$ we are considering. Therefore, it would have been better to write Eqn. (10.6) in the form

$$|E_n\rangle = |E_n^{(0)}\rangle + g\,|1\rangle_n + g^2\,|2\rangle_n + \ldots g^m\,|m\rangle_n + \ldots \ ,$$

where the label $|\cdot\rangle_n$ on the states $|1\rangle$, $|2\rangle$, ... denotes that these states are different if we use different $|E_n^{(0)}\rangle$.

From Eqn. (10.18) it follows that the states $|l\rangle$ depend on which energy level $E_n^{(0)}$ and which state $|E_n^{(0)}\rangle$ we are dealing with. In fact, $E_n^{(0)}$ appear on the right-hand side of Eqn. (10.18).

Eqn. (10.18) can be written in a more elegant way by introducing suitable projectors.

10.1.4 Projection operators in perturbation theory

Let us introduce the operator

$$\hat{Q}_0 \equiv \mathbb{1} - |E_n^{(0)}\rangle\,\langle E_n^{(0)}| = \sum_{i\neq n} |E_i^{(0)}\rangle\,\langle E_i^{(0)}| \ . \tag{10.19}$$

The last result follows from the completeness of the eigenstates of $\hat{H}_0$:

$$\sum_j |E_j^{(0)}\rangle\,\langle E_j^{(0)}| = \mathbb{1} \ .$$

We can also introduce the following operator:

$$\hat{\tilde{Q}}_0 \equiv \hat{Q}_0 \frac{1}{E_n^{(0)} - \hat{H}_0} \hat{Q}_0 \ . \tag{10.20}$$

It is not difficult to prove, using Eqn. (10.19), that

$$\hat{\tilde{Q}}_0 = \sum_{i\neq n} \frac{|E_i^{(0)}\rangle\,\langle E_i^{(0)}|}{E_n^{(0)} - E_i^{(0)}} \ . \tag{10.21}$$

Using $\hat{\tilde{Q}}_0$ we can write Eqn. (10.18) as

$$|l\rangle = \hat{\tilde{Q}}_0 \left[\left(\hat{V} - \varepsilon_1 \right) |l-1\rangle - \varepsilon_2 |l-2\rangle - \ldots - \varepsilon_{l-1} |1\rangle \right] . \tag{10.22}$$

Eqn. (10.22) tells us that from $|E_n^{(0)}\rangle$ and ε_1 we get $|1\rangle$, from $|1\rangle$ we get ε_2, then from $|1\rangle$ and ε_2 we get $|2\rangle$, from $|2\rangle$ we get ε_3, and so.

Furthermore, notice that

$$\hat{\tilde{Q}}_0 |E_n^{(0)}\rangle = 0 , \tag{10.23}$$

with $n \neq i$. This follows from Eqn. (10.21) e from the fact that $\langle E_n^{(0)} | E_i^{(0)} \rangle = \delta_{i,n}$.

Notice that also $\hat{\tilde{Q}}_0$ should carry a label n since the right-hand side of Eqn. (10.21) contains $E_n^{(0)}$, thus there are as many $\hat{\tilde{Q}}_0$ and strings of states $|1\rangle$, $|2\rangle$, ..., as $E_n^{(0)}$.

Homework 10.2. Prove Eqn. (10.21).

Homework 10.3. Prove Eqn. (10.22).

10.1.5 Summary of formulas for first-order perturbation theory

Let us recall Eqn.(10.14):

$$\varepsilon_1 = \langle E_n^{(0)} | \hat{V} | E_n^{(0)} \rangle ,$$

and also Eqn. (10.5)

$$\begin{aligned} E_n &= E_n^{(0)} + \varepsilon_1 g + \mathcal{O}(g^2) \\ &= \langle E_n^{(0)} | \hat{H}_0 | E_n^{(0)} \rangle + g \langle E_n^{(0)} | \hat{V} | E_n^{(0)} \rangle + \mathcal{O}(g^2) \\ &= \langle E_n^{(0)} | (\hat{H}_0 + g\hat{V}) | E_n^{(0)} \rangle + \mathcal{O}(g^2) \\ E_n &= \langle E_n^{(0)} | \hat{H} | E_n^{(0)} \rangle + \mathcal{O}(g^2) . \end{aligned}$$

Therefore, the energy at the first order in g is nothing but $\hat{H}$ (the complete Hamiltonian) evaluated on the unperturbed state $|E_n^{(0)}\rangle$.

From the formula in Eqs. (10.6) we get

$$|E_n\rangle = |E_n^{(0)}\rangle + g |1\rangle + \mathcal{O}(g^2) ,$$

where

$$|1\rangle = \hat{\tilde{Q}}_0 \left(\hat{V} - \varepsilon_1\right) |E_n^{(0)}\rangle \ ,$$

and since $\hat{\tilde{Q}}_0 \, |E_n^{(0)}\rangle = 0$ we get

$$|1\rangle = \hat{\tilde{Q}}_0 \hat{V} \, |E_n^{(0)}\rangle \ ,$$

thus

$$\begin{aligned} |E_n\rangle &= |E_n^{(0)}\rangle + g \, |1\rangle + \mathcal{O}(g^2) \\ &= \left(1 + g\hat{\tilde{Q}}\hat{V}\right) |E_n^{(0)}\rangle + \mathcal{O}(g^2) \ . \end{aligned}$$

It is also possible to check that

$$\langle E_n | E_n \rangle = 1 + \mathcal{O}(g^2) \ , \qquad \langle E_n | E_n^{(0)} \rangle = 1 \ .$$

Another remark: Eqn. (10.18) can be simplified as follows:

$$\langle E_i^{(0)} | 1 \rangle = \frac{\langle E_i^{(0)} | \hat{V} | E_n^{(0)} \rangle - \varepsilon_1 \, \langle E_i^{(0)} | E_n^{(0)} \rangle}{E_n^{(0)} - E_i^{(0)}} \ ,$$

and from $\langle E_i^{(0)} | E_n^{(0)} \rangle = 0$ for $n \neq i$ we obtain

$$\langle E_i^{(0)} | 1 \rangle = \frac{\langle E_i^{(0)} | \hat{V} | E_n^{(0)} \rangle}{E_n^{(0)} - E_i^{(0)}} \ ,$$

i.e., the components of $|1\rangle$ on $\langle E_i^{(0)}|$ are given (modulo the denominator $E_n^{(0)} - E_i^{(0)}$) from the off-diagonal matrix elements of $\hat{V}$.

10.2 Perturbation theory: degenerate case

We will now treat the case in which the states are degenerate. For completeness, let us rewrite some of the previous equations. We are interested in solving

$$\hat{H} = \hat{H}^{(0)} + g\hat{V} \ , \tag{10.24}$$

and $E_1^{(0)}, E_2^{(0)}, E_3^{(0)}, \ldots, E_m^{(0)}, \ldots$ are the energy levels of the unperturbed Hamiltonian $\hat{H}^{(0)}$ which are exactly known.

Let's now suppose, differently from the previous section, that these energy levels are degenerate:

$$\hat{H}^{(0)} \left| E_n^{(0)}, \alpha \right\rangle = E_n^{(0)} \left| E_n^{(0)}, \alpha \right\rangle \ , \tag{10.25}$$

where α is the set of quantum numbers which are necessary to label the degeneracy, or if you wish, the quantum numbers associated with a set of operators which together with $\hat{H}^{(0)}$ form a complete set.

Like for the non-degenerate case, we require that

$$\lim_{g \to 0} E_n = E_n^{(0)} \ , \tag{10.26}$$

where E_n is the eigenvalue of $\hat{H}$.

Furthermore, we assume an expansion to be of the form

$$\begin{cases} E_n = E_n^{(0)} + g\varepsilon_1 + g^2\varepsilon_2 + \ldots + g^m\varepsilon_m + \ldots \\ |E_n\rangle = |E_n^{(0)}\rangle + g\,|1\rangle + g^2\,|2\rangle + \ldots + g^m\,|m\rangle + \ldots \end{cases} , \tag{10.27}$$

Inserting Eqs. (10.27) into the stationary Schrödinger equation

$$\hat{H}\,|E_n\rangle = E_n\,|E_n\rangle \ ,$$

we get, as before, the string of equations

$$\begin{cases} \left(\hat{H}^{(0)} - E_n^{(0)}\right) |E_n^{(0)}\rangle = 0 \ , \\ \left[\hat{H}^{(0)} - E_n^{(0)}\right] |1\rangle + \left(\hat{V} - \varepsilon_1\right) |E_n^{(0)}\rangle = 0 \ , \\ \left[\hat{H}^{(0)} - E_n^{(0)}\right] |2\rangle + \left(\hat{V} - \varepsilon_1\right) |1\rangle - \varepsilon_2\,|E_n^{(0)}\rangle = 0 \\ \vdots \end{cases} \tag{10.28}$$

Again, we impose

$$\langle E_n | E_n^{(0)} \rangle = 1 \ ,$$

and we get from the second equation of (10.27) that

$$\langle 1 | E_n^{(0)} \rangle = \langle 2 | E_n^{(0)} \rangle = \ldots = \langle m | E_n^{(0)} \rangle = \ldots = 0 \ .$$

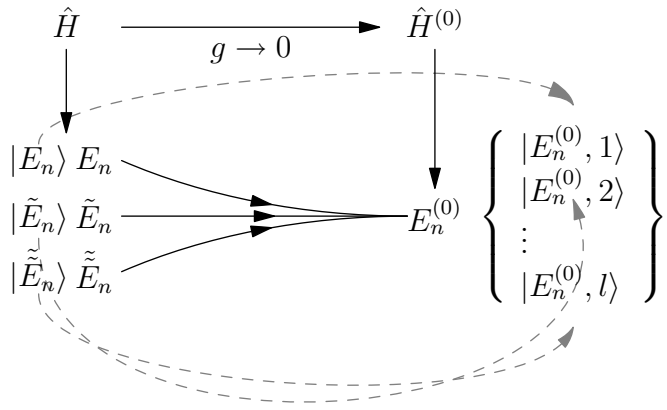


FIG. 10.1. Schematic description

Let us suppose that the eigenvalue $E_n^{(0)}$ of the unperturbed Hamiltonian $\hat{H}^{(0)}$ is l times degenerate, *i.e.*,

$$|E_n^{(0)}, \alpha\rangle \tag{10.29}$$

where α takes l different values. These vectors span a space which we shall denote with $\xi_n^{(0)}$. Let us denote with $\hat{P}^{(0)}$ the projector onto this space.

We have already required, see Eqn. (10.26), that the eigenvalue E_n of $\hat{H}$ goes to $E_n^{(0)}$ of $\hat{H}^{(0)}$ as $g \to 0$, but it may happen that there is more than one eigenvalue of $\hat{H}$ which go to $E_n^{(0)}$, see Fig. 10.1. In other words, there may be more than one eigenvalue E_n, $\tilde{E}_n$, $\tilde{\tilde{E}}_n$ of $\hat{H}$ which go to $E_n^{(0)}$ for $g \to 0$ and the corresponding states $|E_n\rangle$, $|\tilde{E}_n\rangle$, $|\tilde{\tilde{E}}_n\rangle$ go into different vectors of the l-dimensional space of eigenstates associated with $E_n^{(0)}$.

It may also happen that these states

$$|E_n\rangle \ , \qquad |\tilde{E}_n\rangle \ , \qquad |\tilde{\tilde{E}}_n\rangle$$

are degenerate too, with degeneracy $l_n, \tilde{l}_n, \tilde{\tilde{l}}_n$ respectively. We require (and this is true if the perturbation does not make $\hat{H}$ too different from $\hat{H}^{(0)}$) that the sum of the degeracy of $|E_n\rangle$, $|\tilde{E}_n\rangle$, $|\tilde{\tilde{E}}_n\rangle$, namely $l_n + \tilde{l}_n + \tilde{\tilde{l}}_n$, is equal to the degeneracy of $E_n^{(0)}$,

namely l, i.e:

$$l_n + \tilde{l}_n + \tilde{\tilde{l}}_n = l \ .$$

Let us denote with ξ_1, ξ_2, ξ_3 the subspaces of E_n with degeneracies l_n, $\tilde{l}_n$, $\tilde{\tilde{l}}_n$ respectively, so :

$$\lim_{g\to 0}(\xi_1 + \xi_2 + \xi_3) = \xi_n^{(0)} \ ,$$

where $\xi_n^{(0)}$ is the subspace spanned by the eigenstates $|E_n^{(0)}, \alpha\rangle$. The projector $\hat{P}$ onto $\xi_1 + \xi_2 + \xi_3$ for $g \to 0$ will go into the projector $\hat{P}^{(0)}$ (which was defined previously):

$$\lim_{g\to 0} \hat{P}_{\xi_1} + \hat{P}_{\xi_2} + \hat{P}_{\xi_3} = \hat{P}^{(0)} \ .$$

Let us now move back to the equation of the unperturbed Hamiltonian:

$$\left(\hat{H}^{(0)} - E^{(0)}\right) |E^{(0)}, \beta\rangle = 0 \ ,$$

where we have removed, for the moment, the label n of the energy level. $|E^{(0)}, \beta\rangle$ is a linear combination of the degenerate states $|E_n^{(0)}, \alpha\rangle$ of Eqn. (10.29). The only requirement is that this state belongs to the subspace $\xi^{(0)}$, *i.e.*,

$$\hat{P}^{(0)} |E^{(0)}, \beta\rangle = |E^{(0)}, \beta\rangle \ .$$

Now, consider Eqs. (10.28). The second equation is

$$\left(\hat{H}^{(0)} - E^{(0)}\right) |1\rangle + \left(\hat{V} - \varepsilon_1\right) |E_n^{(0)}\rangle = 0 \ ,$$

where $|E_n^{(0)}\rangle$ is one generic state which belongs to $\xi^{(0)}$. By projecting this equation onto $\xi^{(0)}$ we get

$$\hat{P}^{(0)} \left(\hat{H}^{(0)} - E^{(0)}\right) |1\rangle + \hat{P}^{(0)} \left(\hat{V} - \varepsilon_1\right) |E_n^{(0)}\rangle = 0 \ ,$$

i.e.,

$$\hat{P}^{(0)}\hat{H}^{(0)} |1\rangle - E^{(0)} \underbrace{\hat{P}^{(0)} |1\rangle}_{0} + \hat{P}^{(0)}\hat{V} |E_n^{(0)}\rangle - \varepsilon_1 \underbrace{\hat{P}^{(0)} |E_n^{(0)}\rangle}_{|E_n^{(0)}\rangle} = 0 \ ,$$

and hence

$$\hat{P}^{(0)}\hat{H}^{(0)}\,|1\rangle + \hat{P}^{(0)}\hat{V}\,|E_n^{(0)}\rangle - \varepsilon_1\,|E_n^{(0)}\rangle = 0\;. \tag{10.30}$$

$\hat{P}^{(0)}$ is the projector onto the subspace labelled by the states $|E_n^{(0)},\alpha_i\rangle$, *i.e.*:

$$\hat{P}^{(0)} = \sum_i |E_n^{(0)},\alpha_i\rangle\,\langle E_n^{(0)},\alpha_i|\;.$$

If now we put this expession into Eqn. (10.30) we get

$$\sum_i |E_n^{(0)},\alpha_i\rangle\,\langle E_n^{(0)},\alpha_i|\hat{H}^{(0)}|1\rangle + \hat{P}^{(0)}\hat{V}\hat{P}^{(0)}\,|E_n^{(0)}\rangle = \varepsilon_1\,|E_n^{(0)}\rangle\;,$$

so

$$E_n^{(0)}\sum_i |E_n^{(0)},\alpha_i\rangle\,\langle E_n^{(0)},\alpha_i|1\rangle + \hat{P}^{(0)}\hat{V}\hat{P}^{(0)}\,|E_n^{(0)}\rangle = \varepsilon_1\,|E_n^{(0)}\rangle\;,$$

therefore we end up in

$$\left(\hat{P}^{(0)}\hat{V}\hat{P}^{(0)}\right)|E_n^{(0)}\rangle = \varepsilon_1\,|E_n^{(0)}\rangle\;. \tag{10.31}$$

The latter is an eigenvalue equation. It is basically the diagonalization of the operator $\hat{V}$ in the space $\xi^{(0)}$, since it is there that we have projected $\hat{V}$ via the projector $\hat{P}^{(0)}$. Moreover, the eigenstates are of the type $|E_n^{(0)}\rangle$, *i.e.*, linear combinations (to be determined) of the states $|E_n^{(0)},\alpha_i\rangle$.

Now, remember that

$$\hat{P}^{(0)} = \sum_{\alpha_i} |E^{(0)},\alpha_i\rangle\,\langle E^{(0)},\alpha_i|\;,$$

and that we can always choose $|E^{(0)},\alpha_i\rangle$ to be mutually orthogonal. Eqn. (10.31) then becomes

$$\sum_{\alpha_i}\sum_{\alpha_j} \underbrace{|E^{(0)},\alpha_i\rangle\,\langle E^{(0)},\alpha_i|}_{\hat{P}^{(0)}}\hat{V}\,|E^{(0)},\alpha_j\rangle\,\langle E^{(0)},\alpha_j|E_n^{(0)}\rangle = \varepsilon_1\,|E_n^{(0)}\rangle\;.$$

If we bracket from the left with $\langle E^{(0)},\beta_k|$ we get

$$\sum_j \langle E^{(0)},\beta_k|\hat{V}|E^{(0)},\alpha_j\rangle\,\langle E^{(0)},\alpha_j|E_n^{(0)}\rangle = \varepsilon_1\,\langle E^{(0)},\beta_k|E_n^{(0)}\rangle\;,$$

which can be written as

$$\sum_j V_{\beta_k,\alpha_j}\psi^{(0)}_{\alpha_j} = \varepsilon_1\psi^{(0)}_{\beta_k} \ , \tag{10.32}$$

where

$$V_{\beta_k,\alpha_j} = \langle E^{(0)}, \beta_k|\hat{V}|E^{(0)}, \alpha_j\rangle \ ,$$

and

$$\psi^{(0)}_{\alpha_j} = \langle E^{(0)}, \alpha_j|E^{(0)}_n\rangle \ .$$

The ε_1 are determined by Eqn. (10.31) or equivalently by Eqn. (10.32).

From Eqn. (10.32), we see that ε_1 are eigenvalues of the matrix V_{β_k,α_j}. In the non-degenerate case, we had only one eigenvalue:

$$\varepsilon_1 = \langle E^{(0)}_n|\hat{V}|E^{(0)}_n\rangle \ .$$

Now instead we have many of them, since we have to diagonalize a matrix V_{β_k,α_j} of dimension $l \times l$. If the various ε_1 are all different one from the other, we will have exactly l of them and the corresponding energy levels will be

$$\begin{cases} E_1 = E_0 + g\varepsilon^1_{(1)} \\ E_2 = E_0 + g\varepsilon^2_{(1)} \\ \quad \vdots \\ E_l = E_0 + g\varepsilon^l_{(1)} \ , \end{cases} \tag{10.33}$$

where the index (1) denotes that we are dealing with the first-order perturbative correction, and we have l different corrections since Eqn. (10.32) has l different eigenvalues (as we have assumed). From Eqn. (10.33), if $\varepsilon^j_{(1)}$ are all different, one says that the perturbation has completely removed the degeneracy. In fact, one goes from one value for the energy $E^{(0)}$ to l different values $E_1, E_2, \ldots, E_l$.

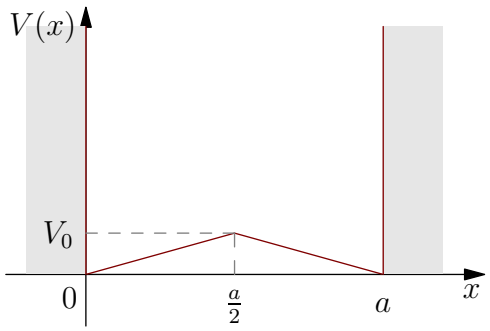


FIG. 10.2. Infinite well with a perturbation

10.3 Problems and Solutions

Problem 10.1. For a particle in an infinite-well potential of width a, *i.e.*, $0 < x < a$, find the first-order perturbative shift of the energy levels under the action of a piecewise linear perturbation potential as plotted in the figure (10.2).

SOLUTION. It is not difficult to show that the analytic form of this potential is

$$V(x) = \frac{V_0}{a}\left(a - |2x - a|\right) \ , \tag{10.34}$$

where V_0/a will be the "small" parameter. The exactly solvable system, *i.e.*, $\hat{H}_0$ in Eqn. (10.1), is the infinite potential well between 0 and a. Its energy levels and eigenfunctions are

$$E_n^{(0)} = \frac{\hbar^2\pi^2\,(n+1)^2}{2ma^2} \ , \tag{10.35a}$$

$$\psi_n^{(0)}(x) = \sqrt{\frac{2}{a}}\sin\frac{\pi\,(n+1)\,x}{a} \ , \tag{10.35b}$$

where n =0,1,2,... By using Eqn. (10.14), which gives the energy correction at first order in perturbation theory

$$\varepsilon_n^{(1)} = \langle\psi_n^{(0)}|\hat{V}|\psi_n^{(0)}\rangle \ ,$$

and using Eqs. (10.34)–(10.35) we get

$$\begin{aligned}
\varepsilon_n^{(1)} &= \int_0^a \mathrm{d}x \left\{ \sqrt{\frac{2}{a}} \sin \frac{\pi(n+1)x}{a} \left[\frac{V_0}{a} \left(a - |2x-a| \right) \right] \sqrt{\frac{2}{a}} \sin \frac{\pi\,(n+1)\,x}{a} \right\} \\
&= \frac{2}{a}\frac{V_0}{a} \int_0^a \sin^2 \left(\frac{\pi\,(n+1)\,x}{a} \right) \{ a - |2x-a| \}\, \mathrm{d}x \\
&= \frac{2V_0}{a^2} \int_0^a a \sin^2 \left(\frac{\pi\,(n+1)\,x}{a} \right) \mathrm{d}x - \frac{2V_0}{a^2} \int_0^a |2x-a| \sin^2 \left(\frac{\pi\,(n+1)\,x}{a} \right) \mathrm{d}x \\
&= \frac{2V_0}{a} \int_0^a \sin^2 \frac{\pi\,(n+1)\,x}{a}\, \mathrm{d}x - \frac{2V_0}{a^2} \int_0^{a/2} (a-2x) \sin^2 \left(\frac{\pi\,(n+1)\,x}{a} \right) \mathrm{d}x \\
&\qquad - \frac{2V_0}{a^2} \int_{a/2}^a (2x-a) \sin^2 \left(\frac{\pi\,(n+1)\,x}{a} \right) \mathrm{d}x \,.
\end{aligned}$$

The first integral above is evaluated by means of the duplication formula

$$\frac{2V_0}{a} \int_0^a \frac{1 - \cos \frac{2\pi(n+1)x}{a}}{2}\, \mathrm{d}x \,,$$

for the second one we use the identity

$$\int x \sin^2 x = \int x \left(\frac{1-\cos 2x}{2} \right) = \int \frac{x}{2} - \int \frac{x}{2} \cos 2x \,,$$

where the latter is evaluated by parts. The result is

$$\varepsilon_n^{(1)} = V_0 \left\{ \frac{1}{2} + \frac{1+(-1)^n}{\pi^2\,(n+1)^2} \right\} \,.$$

$\star \quad \star \quad \star$

FURTHER PROBLEMS WILL BE PROVIDED IN HANDWRITTEN FORM BY THE TEACHER

Chapter 11

SCHRÖDINGER, HEISENBERG AND INTERACTION PICTURES

11.1 Schrödinger and Heisenberg pictures

We have seen that the mean values of an observable $\hat{O}$ do not change under an unitary transformation $\hat{U}$ which changes the basis of the system:

$$\hat{O} \to \hat{O}' = \hat{U}\hat{O}\hat{U}^\dagger \ , \tag{11.1}$$

where states change as

$$\begin{array}{rcccl} |\psi\rangle & \to & |\psi'\rangle & = & \hat{U}\,|\psi\rangle \\ \langle\psi| & \to & \langle\psi'| & = & \langle\psi|\,\hat{U}^\dagger \ . \end{array} \tag{11.2}$$

In this way, we have

$$\langle\psi|\hat{O}|\psi\rangle = \langle\psi|\hat{U}^\dagger\hat{U}\hat{O}\hat{U}^\dagger\hat{U}|\psi\rangle = \langle\psi'|\hat{O}'|\psi'\rangle \ . \tag{11.3}$$

The usual picture of time evolution which we have used so far is an exception to this rule. It is the one in which the states evolve in time as:

$$|\psi(t)\rangle = \mathrm{e}^{-i\hat{H}t/\hbar}\,|\psi(0)\rangle \ ,$$

according to the Schrödinger equation and the observables instead do not change in time. This should be considered as a further *postulate* of quantum mechanics. If for the observable $\hat{x}$ we want to compute the mean value at time t we do the following:

$$\langle\psi(t)|\hat{x}|\psi(t)\rangle = \overline{\hat{x}(t)} \ . \tag{11.4}$$

That means we take $\hat{x}$ as given and use the states at time t. This is called *Schrödinger picture*.

As we said in the previous chapters, we will not apply the rules (11.1) to the unitary transformation $\hat{U}(t)$ representing time evolution. If we had applied rules (11.1)–(11.2) we would have got the analogous of Eqn. (11.3) and thus

$$\overline{\hat{x}(t)} = \overline{\hat{x}(0)} \ .$$

This does not make any sense, since it would tell us that the mean position does not change in time, as if the system had not evolved in time.

We shall denote quantities in the Schrödinger picture with the label S, for example, $|\psi_S(t)\rangle$ means that this is a state in the Schrödinger picture where the states, but not the observables, evolve in time.

It may seem unnatural to let the states evolve in time while letting the observables unchanged. Let us ask ourselves if it is possible to use a different picture in which the observables evolve in time but in such a way that we get the same mean value of the Schrödinger picture, namely Eqn. (11.4), which have been confirmed by experiments.

Such different picture does in fact exist, and it is called *Heisenberg picture*. The states in the Heisenberg picture will be denoted by $|\psi_H(t)\rangle$. Such states are obtained from those in the Schrödinger picture via a unitary transformation of this type:

$$|\psi_H(t)\rangle = \hat{S}^\dagger \, |\psi_S(t)\rangle \ , \tag{11.5}$$

where

$$\hat{S}^\dagger = \mathrm{e}^{i\hat{H}t/\hbar} \ .$$

Recall that

$$|\psi_S(t)\rangle = \mathrm{e}^{-i\frac{Ht}{\hbar}} \, |\psi_S(0)\rangle \ ,$$

by inserting this expression into Eqn. (11.5) we obtain

$$|\psi_H(t)\rangle = \hat{S}^\dagger \, |\psi_S(t)\rangle = \mathrm{e}^{i\hat{H}t/\hbar} \, |\psi_S(t)\rangle = \mathrm{e}^{i\hat{H}t/\hbar} \, \mathrm{e}^{-i\hat{H}t/\hbar} \, |\psi_S(0)\rangle = |\psi_S(0)\rangle \ . \tag{11.6}$$

Therefore, the states in the Heisenberg picture at any time t are equal to the states in the Schrödinger picture at the initial time $t = 0$.

Notice that on the left-hand side of Eqn. (11.6) " t "appears but it does not on the right-hand side, which means that in fact also on the left-hand side there is no *effective* dependence on time. Thus, also if we change t on the left nothing changes:

$$|\psi_H(t')\rangle = |\psi_H(t)\rangle = |\psi_H(0)\rangle = |\psi_S(0)\rangle \ .$$

As $\hat{S}^\dagger$ is a unitary operator we have to apply the general rules for transforming, besides the states, also the observables as outlined at the beginning of this section. Therefore, the change from an observable $\hat{O}_S$ in the Schrödinger picture to the same one $\hat{O}_H$ in the Heisenberg picture is

$$\hat{O}_H = \hat{S}^\dagger \hat{O}_S \hat{S} = \mathrm{e}^{i\hat{H}t/\hbar}\, \hat{O}_S\, \mathrm{e}^{-i\hat{H}t/\hbar} \ . \tag{11.7}$$

In this way, it is clear that, even if $\hat{O}_S$ does not depend explicitly on time, $\hat{O}_H$ is time-dependent, and for this reason we shall denote it with $\hat{O}_H(t)$.

11.1.1 Time evolution of $\hat{O}_H(t)$

Let's try to understand how $\hat{O}_H(t)$ evolves in time. By taking the time-derivative, we obtain:

$$\begin{aligned}
\frac{\mathrm{d}\hat{O}_H(t)}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\mathrm{e}^{i\hat{H}t/\hbar}\,\hat{O}_S\,\mathrm{e}^{-i\hat{H}t/\hbar}\right] = \frac{\mathrm{d}}{\mathrm{d}t}\left[\mathrm{e}^{i\hat{H}t/\hbar}\right]\hat{O}_S\,\mathrm{e}^{-i\hat{H}t/\hbar} + \mathrm{e}^{i\hat{H}t/\hbar}\,\hat{O}_S\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathrm{e}^{-i\hat{H}t/\hbar}\right] \\
&= \frac{i}{\hbar}\hat{H}\underbrace{\mathrm{e}^{i\hat{H}t/\hbar}\,\hat{O}_S\,\mathrm{e}^{-i\hat{H}t/\hbar}}_{\hat{O}_H(t)} + \mathrm{e}^{i\hat{H}t/\hbar}\,\hat{O}_S\left(-\frac{i}{\hbar}\hat{H}\right)\mathrm{e}^{-i\hat{H}t/\hbar} \\
&= \frac{i}{\hbar}\hat{H}\hat{O}_H(t) + \underbrace{\mathrm{e}^{i\hat{H}t/\hbar}\,\hat{O}_S\,\mathrm{e}^{-i\hat{H}t/\hbar}}_{\hat{O}_H(t)}\left(-\frac{i}{\hbar}\hat{H}\right) \\
&= \frac{i}{\hbar}\hat{H}\hat{O}_H(t) - \frac{i}{\hbar}\hat{O}_H(t)\hat{H} = \frac{i}{\hbar}[\hat{H}, \hat{O}_H(t)] \ ,
\end{aligned}$$

hence

$$\frac{\mathrm{d}\hat{O}_H(t)}{\mathrm{d}t} = \frac{i}{\hbar}[\hat{H}, \hat{O}_H(t)] \ . \tag{11.8}$$

Notice that this is exactly the same equation we would get by using the correspondence principle to the classical equations of motion

$$\frac{\mathrm{d}O}{\mathrm{d}t} = \{O, H\}_{\text{P.B.}} = -\{H, O\}_{\text{P.B.}} \,, \tag{11.9}$$

The corrispondence rules say that

$$\{\cdot, \cdot\}_{\text{P.B.}} \longrightarrow \frac{1}{i\hbar}[\cdot, \cdot] \,,$$

so by applying these to Eqn. (11.9) we exactly get Eqn. (11.8).

Among all operators, only one is the same in both Schrödinger and Heisenberg pictures, and this is the Hamiltonian operator $\hat{H}$. In fact,

$$\hat{H}_H(t) = \mathrm{e}^{i\hat{H}t/\hbar}\, \hat{H}_S\, \mathrm{e}^{-i\hat{H}t/\hbar} = \mathrm{e}^{i\hat{H}t/\hbar}\, \mathrm{e}^{-i\hat{H}t/\hbar}\, \hat{H}_S = \hat{H}_S \,.$$

These two different pictures reflect the different approaches Schrödinger and Heisenberg used to develop Quantum Mechanics. Heisenberg was using all the time the analogies with classical mechanics and so he made heavy use of the correspondence principle, while Schrödinger had postulated his equation and was calculating everything using that. Only afterwards they realized their approaches were giving the same physical results.

Now, we shall prove that the expectation value of the observables are the same in the Schrödinger and Heisenberg pictures, like Eqn. (11.4). The expectation value of an observable $\hat{O}_H$ in the Heisenberg picture is

$$\begin{aligned} \langle \psi_H(t)|\hat{O}_H(t)|\psi_H(t)\rangle &= \underbrace{\langle \psi_S(t)|\, \mathrm{e}^{-i\hat{H}t/\hbar}}_{\langle \psi_H(t)|}\, \underbrace{\mathrm{e}^{i\hat{H}t/\hbar}\, \hat{O}_S\, \mathrm{e}^{-i\hat{H}t/\hbar}}_{\hat{O}_H(t)}\, \underbrace{\mathrm{e}^{i\hat{H}t/\hbar}\, |\psi_S(t)\rangle}_{|\psi_H(t)\rangle} \\ &= \langle \psi_S(t)|\hat{O}_S|\psi_S(t)\rangle \,. \end{aligned}$$

Expectation values are the quantities that experimentalists measure in their experiments, so it is very important that they do not depend on the particular picture we use to describe time-evolution.

Since the expectation values are those measured by experimentalists, it is also important to understand how these expectation values evolve in time. This is investigated in the following section.

11.1.2 Time evolution of the average values.

Let us consider an observable $\hat{O}$ and let us calculate its expectation value working in the Schrödinger picture:

$$\overline{\hat{O}}(t) = \langle\psi(t)|\hat{O}|\psi(t)\rangle \ .$$

Taking the time derivative we get

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\overline{\hat{O}}(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\langle\psi(t)|\hat{O}|\psi(t)\rangle\right] \\ &= \frac{\mathrm{d}\,\langle\psi(t)|}{\mathrm{d}t}\hat{O}\,|\psi(t)\rangle + \langle\psi(t)|\,\hat{O}\frac{\mathrm{d}\,|\psi(t)\rangle}{\mathrm{d}t} \ . \end{aligned} \tag{11.10}$$

Remember that, using the time-dependent Schrödinger equation,

$$i\hbar\frac{\mathrm{d}\,|\psi(t)\rangle}{\mathrm{d}t} = \hat{H}\,|\psi(t)\rangle \ ,$$

which can also be written as:

$$\frac{\mathrm{d}\,|\psi(t)\rangle}{\mathrm{d}t} = -\frac{i}{\hbar}\hat{H}\,|\psi(t)\rangle \ , \tag{11.11}$$

and whose dual is

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\langle\psi(t)| = \langle\psi(t)|\,\hat{H}\frac{i}{\hbar} \ . \tag{11.12}$$

So we get from Eqn. (11.10) that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\overline{\hat{O}}(t) &= \frac{i}{\hbar}\,\langle\psi(t)|\hat{H}\hat{O}|\psi(t)\rangle - \frac{i}{\hbar}\,\langle\psi(t)|\hat{O}\hat{H}|\psi(t)\rangle \\ &= \frac{i}{\hbar}\,\langle\psi(t)|[\hat{H},\hat{O}]|\psi(t)\rangle \\ &= \frac{i}{\hbar}\overline{[\hat{H},\hat{O}]} \end{aligned} \tag{11.13}$$

This is the equation of motion of the expectation values. Notice that, in general, on the right-hand side of Eqn. (11.13) it is not possible to factorize the mean value on each term of the commutator, *i.e.*,

$$\overline{[\hat{H},\hat{O}]} \neq [\overline{\hat{H}},\overline{\hat{O}}] \; . \tag{11.14}$$

Homework 11.1. Explain why in general Eqn. (11.14) does hold.

11.1.3 Constants of motion

Let us suppose that $\hat{O}$ commutes with $\hat{H}$:

$$[\hat{O},\hat{H}] = 0 \; .$$

Then, the right-hand side of Eqn. (11.13) is zero and so

$$\frac{\mathrm{d}\overline{\hat{O}}(t)}{\mathrm{d}t} = 0 \; , \tag{11.15}$$

which means that the mean value does not change in time. In this case, $\hat{O}$ is said to be a constant of motion.

If $\hat{O}$ is a constant of motion, then it is possible to prove that its eigenvalues do not depend on time. Let us start from Eqn. (11.15) and in particular we consider the expectation value with respect to an eigenstate $|\psi_n\rangle$ of $\hat{O}$:

$$\hat{O}\,|\psi_n\rangle = o_n\,|\psi_n\rangle \; .$$

The time derivative of the mean value of $\hat{O}$ with respect to the eigenstate $|\psi_n\rangle$ is easily evaluated to be

$$\begin{aligned}\left.\frac{\mathrm{d}\overline{\hat{O}}}{\mathrm{d}t}\right|_{\psi_n} &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\langle\psi_n(t)|\hat{O}|\psi_n(t)\rangle\right] \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\langle\psi_n(0)|\,\mathrm{e}^{i\hat{H}t/\hbar}\,\hat{O}\,\mathrm{e}^{-i\hat{H}t/\hbar}\,|\psi_n(0)\rangle\right] \; ,\end{aligned}$$

which must be zero according to Eqn. (11.15). $\hat{O}$ commutes with $\hat{H}$, thus it is possible to interchange $\hat{O}$ with $\mathrm{e}^{-i\hat{H}t/\hbar}$ in the previous expression to obtain

$$\begin{aligned}\left.\frac{\mathrm{d}\overline{\hat{O}}(t)}{\mathrm{d}t}\right|_{\psi_n} &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\langle\psi_n(0)|\,\mathrm{e}^{i\hat{H}t/\hbar}\,\mathrm{e}^{-i\hat{H}t/\hbar}\,\hat{O}|\psi_n(0)\rangle\right] = \frac{\mathrm{d}}{\mathrm{d}t}\left[\langle\psi_n(0)|\hat{O}|\psi_n(0)\rangle\right] \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\langle\psi_n(0)|o_n|\psi_n(0)\rangle\right] = \frac{\mathrm{d}}{\mathrm{d}t}\left[o_n\,\langle\psi_n(0)|\psi_n(0)\rangle\right] \\ &= \frac{\mathrm{d}o_n}{\mathrm{d}t} = 0\ ,\end{aligned}$$

and this proves also the result that the eigenvalues of observable which are constants of motion do not depend on time.

11.1.4 Evolution equation of $\hat{q}$ and $\hat{p}$ and Ehrenfest equation

Let us consider the position and momentum operators in the Heisenberg picture: $\hat{q}_H$ and $\hat{p}_H$. Their time evolution is obtained by use of Eqn. (11.8), *i.e.*,

$$\frac{\mathrm{d}\hat{q}_H}{\mathrm{d}t} = \frac{i}{\hbar}[\hat{H},\hat{q}_H]\ , \qquad \frac{\mathrm{d}\hat{p}_H}{\mathrm{d}t} = \frac{i}{\hbar}[\hat{H},\hat{p}_H]\ , \tag{11.16}$$

These are the *Heisenberg equations*, and are the quantum analogous of the Hamilton equations.

Let us now consider the expectation values of Eqn. (11.16):

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\langle\hat{q}_H\rangle = \frac{i}{\hbar}\,\langle[\hat{H},\hat{q}_H]\rangle\ , \qquad \frac{\mathrm{d}}{\mathrm{d}t}\,\langle\hat{p}_H\rangle = \frac{i}{\hbar}\,\langle[\hat{H},\hat{p}_H]\rangle\ . \tag{11.17}$$

These are called *Ehrenfest equations*.

Some remarks are needed here. Let us suppose that

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

then Eqs. (11.17) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\langle\hat{q}_H\rangle = \left\langle\frac{\hat{p}_H}{m}\right\rangle\ , \qquad \frac{\mathrm{d}}{\mathrm{d}t}\,\langle\hat{p}_H\rangle = -\left\langle\frac{\partial V}{\partial q}(\hat{q})\right\rangle\ . \tag{11.18}$$

It is important to *stress* that these equations do not tell us that the average values obey classical mechanics. In fact in the second of equations. (11.18) we are not allowed to replace $\left\langle \frac{\partial V}{\partial q}(\hat{q}) \right\rangle$ with $\frac{\partial V(\langle \hat{q} \rangle)}{\partial \langle \hat{q} \rangle}$. If it were possible, then Eqs. (11.18) would be identical, in form, to the classical equations of motion:

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{\hat{q}}_H = \frac{\overline{\hat{p}}_H}{m} \ , \qquad \frac{\mathrm{d}}{\mathrm{d}t}\overline{\hat{p}}_H = -\frac{\partial V(\overline{\hat{q}})}{\partial \overline{\hat{q}}} \ . \tag{11.19}$$

But it is not so, since

$$\left\langle \frac{\partial V}{\partial q} \right\rangle \neq \frac{\partial V(\overline{\hat{q}})}{\partial \overline{\hat{q}}} \ . \tag{11.20}$$

Why? It is sufficient to consider, for example, a potential such as

$$V(\hat{q}) = \hat{q}^3 \ ,$$

to show that

$$\left\langle \frac{\partial V}{\partial q} \right\rangle = \left\langle 3\hat{q}^2 \right\rangle \ , \qquad \frac{\partial V(\overline{\hat{q}})}{\partial \overline{\hat{q}}} = 3\overline{\hat{q}}^2 \ . \tag{11.21}$$

These two expressions are not the same, in fact their difference is given by

$$3\overline{\hat{q}^2} - 3\overline{\hat{q}}^2 = 3\left[\overline{\hat{q}^2} - \overline{\hat{q}}^2\right] = 3\left(\Delta q\right)^2 \ ,$$

i.e., it is equal to the mean square displacement.

Homework 11.2. For an arbitrary potential $V(\hat{q})$, calculate the difference

$$\left\langle \frac{\partial V}{\partial \hat{q}} \right\rangle - \frac{\partial V(\langle \hat{q} \rangle)}{\partial \langle \hat{q} \rangle} \ ,$$

the result should be expressed in terms of 1) higher-order derivatives of V, 2) the mean square displacement and 3) higher-order moments (generalization of the mean square displacement).

11.2 Interaction picture

This is an intermediate picture, in the sense that the time-dependence of the states is only partially moved on to the observables.

Let us split the Hamiltonian in the Schrödinger picture into two parts: the free part $\hat{H}_S^{(0)}$ (more generally, $\hat{H}_S^{(0)}$ is the exactly-solvable part of the Hamiltonian) and the interacting part $\hat{V}_S$ (*i.e.*, the perturbative part):

$$\hat{H}_S = \hat{H}_S^{(0)} + \hat{V}_S \ . \tag{11.22}$$

Next, we introduce the following unitary operator:

$$\hat{S}_0^\dagger = \mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar} \ .$$

By acting with this operator, we define a new picture (the *interaction picture*, thus the label I) according to the following rules:

$$\begin{aligned} |\psi_I(t)\rangle &= \hat{S}_0^\dagger\,|\psi_S(t)\rangle = \mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,|\psi_S(t)\rangle \\ \hat{O}_I(t) &= \hat{S}_0^\dagger \hat{O}_S \hat{S}_0 \quad = \mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{O}_S\,\mathrm{e}^{-i\hat{H}_S^{(0)}t/\hbar} \ . \end{aligned}$$

First of all, notice that the mean values do not change in the new picture, as it should be:

$$\begin{aligned} \langle\psi_I|\hat{O}_I|\psi_I\rangle &= \langle\psi_S|\hat{S}_0\hat{S}_0^\dagger\hat{O}_S\hat{S}_0\hat{S}_0^\dagger|\psi_S\rangle \\ &= \langle\psi_S|\hat{O}_S|\psi_S\rangle \ . \end{aligned}$$

Notice that $\hat{O}_I(t)$ depends on time, while $\hat{O}_S$ does not, and also $|\psi_I\rangle$ depends on time, but in a way different from that of $|\psi_S\rangle$. Let us see how.

Let us derive the evolution equation for $|\psi_I(t)\rangle$. Our starting point is

$$|\psi_I(t)\rangle = \mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,|\psi_S(t)\rangle \ .$$

Taking the time-derivative of both sides yields

$$\begin{aligned} \frac{\mathrm{d}\,|\psi_I(t)\rangle}{\mathrm{d}t} &= \frac{i}{\hbar}\hat{H}_S^{(0)}\underbrace{\mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,|\psi_S(t)\rangle}_{|\psi_I(t)\rangle} + \mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,\frac{\mathrm{d}\,|\psi_S(t)\rangle}{\mathrm{d}t} \\ &= \frac{i}{\hbar}\hat{H}_S^{(0)}\,|\psi_I(t)\rangle - \frac{i}{\hbar}\,\mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\left[\hat{H}_S^{(0)} + \hat{V}_S\right]|\psi_S(t)\rangle = -\frac{i}{\hbar}\,\mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{V}_S\,|\psi_S(t)\rangle \\ &= -\frac{i}{\hbar}\underbrace{\mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{V}_S\,\mathrm{e}^{-i\hat{H}_S^{(0)}t/\hbar}}_{\hat{V}_I(t)}\underbrace{\mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,|\psi_S(t)\rangle}_{|\psi_I(t)\rangle} \ , \end{aligned}$$

thus

$$i\hbar\frac{\mathrm{d}\,|\psi_I(t)\rangle}{\mathrm{d}t} = \hat{V}_I\,|\psi_I(t)\rangle\ . \tag{11.23}$$

This is the analogous of the Schrödinger equation but in the *interaction picture*. From this equation we discover why this picture is called *interaction picture*, namely because the role of the Hamiltonian is played here by the interaction $\hat{V}_I$. This picture is particularly suitable for doing time-dependent perturbation theory, as we will see.

11.2.1 Evolution of the operators in the interaction picture

Our starting point is

$$\hat{O}_I(t) = \mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{O}_S\,\mathrm{e}^{-i\hat{H}_S^{(0)}t/\hbar}\ .$$

Taking the time-derivative of both sides, we get

$$\begin{aligned}\frac{\mathrm{d}\hat{O}_I(t)}{\mathrm{d}t} &= \frac{i}{\hbar}\hat{H}_S^{(0)}\underbrace{\mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{O}_S\,\mathrm{e}^{-i\hat{H}_S^{(0)}t/\hbar}}_{\hat{O}_I(t)} - \frac{i}{\hbar}\underbrace{\mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{O}_S\,\mathrm{e}^{-i\hat{H}_S^{(0)}t/\hbar}}_{\hat{O}_I(t)}\hat{H}_S^{(0)}\\ &= \frac{i}{\hbar}\hat{H}_S^{(0)}\hat{O}_I - \frac{i}{\hbar}\hat{O}_I\hat{H}_S^{(0)}\\ &= \frac{i}{\hbar}[\hat{H}_S^{(0)},\hat{O}_I]\end{aligned} \tag{11.24}$$

In the interaction picture, $\hat{H}_S^{(0)}$ becomes

$$\hat{H}_I^{(0)} = \mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{H}_S^{(0)}\,\mathrm{e}^{-i\hat{H}_S^{(0)}t/\hbar} = \hat{H}_S^{(0)}\,\mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,\mathrm{e}^{-i\hat{H}_S^{(0)}t/\hbar} = \hat{H}_S^{(0)}\ .$$

Therefore, $\hat{H}_I^{(0)}$ is the same as $\hat{H}_S^{(0)}$. Substituting this result into Eqn. (11.24) yields

$$\frac{\mathrm{d}\hat{O}_I}{\mathrm{d}t} = \frac{i}{\hbar}[\hat{H}_I^{(0)},\hat{O}_I]\ . \tag{11.25}$$

Therefore, in the interaction picture the evolution of the states is ruled by $\hat{V}$ while the evolution of the observables is ruled by $\hat{H}^{(0)}$.

Chapter 12

TIME-DEPENDENT PERTURBATION THEORY

12.1 General derivation

This is a perturbative technique to obtain solutions of the ordinary Schrödinger equation

$$i\hbar \frac{\mathrm{d}\,|\psi_S(t)\rangle}{\mathrm{d}t} = \hat{H}\,|\psi_S(t)\rangle \ . \tag{12.1}$$

In the previous sections, we dealt with *stationary* perturbation theory, *i.e.*, our goal was to find approximated solutions of the stationary Schrödinger equation

$$\hat{H}\,|\psi_n\rangle = E_n\,|\psi_n\rangle \ . \tag{12.2}$$

Solving Eqn. (12.1) in an approximated way means we have to find an approximated form for the evolution operator

$$\hat{U}(t,t_0) = \mathrm{e}^{-i\hat{H}(t-t_0)/\hbar} \ . \tag{12.3}$$

By taking the time-derivative of both sides with respect to t we get

$$i\hbar \frac{\mathrm{d}\hat{U}(t,t_0)}{\mathrm{d}t} = \hat{H}\hat{U}(t,t_0) \ . \tag{12.4}$$

The standard initial condition which has to be fulfilled by the solution of this equation is

$$\hat{U}(t_0,t_0) = \mathbb{1} \ .$$

Now, consider the interaction picture. Working in the Schrödinger picture, we write down the Hamiltonian in the following way:

$$\hat{H} = \hat{H}_S^{(0)} + \hat{V}_S \ ,$$

where $\hat{H}_S^{(0)}$ is an Hamiltonian which we are able to solve exactly. $\hat{V}_S$ is an extra interaction in Schrödinger picture which can eventually depend explicitly on time.

We have already seen that the states in the interaction picture are found from those in the Schrödinger picture in the following way:

$$|\psi_I(t)\rangle = \mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,|\psi_S(t)\rangle = \hat{S}_0^\dagger(t)\,|\psi_S(t)\rangle \ , \tag{12.5}$$

while operators are found according to

$$\hat{O}_I = \mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{O}_S\,\mathrm{e}^{-i\hat{H}_S^{(0)}t/\hbar} = \hat{S}_0^\dagger(t)\hat{O}_S\hat{S}_0(t) \ , \tag{12.6}$$

where

$$\hat{S}_0(t) = \mathrm{e}^{-i\hat{H}_S^{(0)}t/\hbar} \ .$$

We have already seen that the states in the interaction picture evolve in time according to

$$i\hbar\frac{\mathrm{d}\,|\psi_I(t)\rangle}{\mathrm{d}t} = \hat{V}_I\,|\psi_I(t)\rangle \ , \tag{12.7}$$

where

$$\hat{V}_I = \mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{V}_S\,\mathrm{e}^{-i\hat{H}_S^{(0)}t/\hbar} \ .$$

Operators evolve as

$$\frac{\mathrm{d}\hat{O}_I}{\mathrm{d}t} = \frac{1}{i\hbar}[\hat{O}_I, \hat{H}_I^{(0)}] \ . \tag{12.8}$$

Notice that $\hat{H}_S^{(0)}$ and $\hat{H}_I^{(0)}$ are the same. In fact,

$$\hat{H}_I^{(0)} = \mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,\hat{H}_S^{(0)}\,\mathrm{e}^{-i\hat{H}_S^{(0)}t/\hbar} = \hat{H}_S^{(0)}\,\mathrm{e}^{i\hat{H}_S^{(0)}t/\hbar}\,\mathrm{e}^{-i\hat{H}_S^{(0)}t/\hbar} = \hat{H}_S^{(0)} \ .$$

Let us now write down the evolution operator $\hat{U}$ in the interaction picture. We have

$$\hat{U}_I(t) = \hat{S}_0^\dagger(t)\hat{U}(t)\hat{S}_0(t) \ . \tag{12.9}$$

By taking the time-derivative of both sides we can check that $\hat{U}_I(t)$ does not satisfy the usual equation of motion of the operators in the interaction picture, namely

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\hat{U}_I(t) = \hat{V}_I(t)\hat{U}_I(t) \ . \tag{12.10}$$

In fact, we get

$$\begin{aligned}
\frac{\mathrm{d}\hat{U}_I(t)}{\mathrm{d}t} &= \frac{\mathrm{d}\hat{S}_0^\dagger(t)}{\mathrm{d}t}\hat{U}(t)\hat{S}_0(t) + \hat{S}_0^\dagger(t)\frac{\mathrm{d}\hat{U}(t)}{\mathrm{d}t}\hat{S}_0(t) + \hat{S}_0^\dagger(t)\hat{U}(t)\frac{\mathrm{d}\hat{S}_0(t)}{\mathrm{d}t} \\
&= -\frac{1}{i\hbar}\hat{S}_0^\dagger\hat{H}^{(0)}\hat{U}\hat{S}_0 + \hat{S}_0^\dagger\frac{1}{i\hbar}\hat{H}(t)\hat{U}(t)\hat{S}_0(t) + \hat{S}_0^\dagger\hat{U}(t)\frac{1}{i\hbar}\hat{H}^{(0)}\hat{S}_0 \\
&= -\frac{1}{i\hbar}\hat{S}_0^\dagger\hat{H}^{(0)}\hat{U}\hat{S}_0 + \frac{1}{i\hbar}\hat{S}_0^\dagger\left(\hat{H}^{(0)}+\hat{V}\right)\hat{U}(t)\hat{S}_0(t) + \hat{S}_0^\dagger\hat{U}(t)\frac{1}{i\hbar}\hat{H}^{(0)}\hat{S}_0 \\
&= -\frac{1}{i\hbar}\cancel{\hat{S}_0^\dagger\hat{H}^{(0)}\hat{U}\hat{S}_0} + \frac{1}{i\hbar}\cancel{\hat{S}_0^\dagger\hat{H}^{(0)}\hat{U}\hat{S}_0} + \frac{1}{i\hbar}\hat{S}_0^\dagger\hat{V}\hat{U}\hat{S}_0 + \frac{1}{i\hbar}\hat{S}_0^\dagger\hat{U}(t)\hat{H}^{(0)}\hat{S}_0 \\
&= \frac{1}{i\hbar}\underbrace{\hat{S}_0^\dagger\hat{V}\hat{S}_0}_{\hat{V}_I}\underbrace{\hat{S}_0^\dagger\hat{U}\hat{S}_0}_{\hat{U}_I} + \frac{1}{i\hbar}\underbrace{\hat{S}_0^\dagger\hat{U}\hat{S}_0}_{\hat{U}_I}\underbrace{\hat{S}_0^\dagger\hat{H}^{(0)}\hat{S}_0}_{\hat{H}_I^{(0)}} \\
&= \frac{1}{i\hbar}\hat{V}_I\hat{U}_I + \frac{1}{i\hbar}\hat{U}_I\hat{H}_I^{(0)} \ .
\end{aligned}$$

We can see that in the expression above there is an extra term $\hat{U}_I\hat{H}_I^{(0)}/i\hbar$ which does not appear in Eqn. (12.10).

Consider instead the following operator

$$\hat{U}_I'(t) \equiv \hat{S}_0^\dagger\hat{U}(t) \ . \tag{12.11}$$

Let us derive the evolution equation for this operator in the interaction picture. By taking the time-derivative of both sides of Eqn. (12.11) we get

$$\begin{aligned}
\frac{\mathrm{d}\hat{U}_I'(t)}{\mathrm{d}t} &= \frac{\mathrm{d}\hat{S}_0^\dagger}{\mathrm{d}t}\hat{U}(t) + \hat{S}_0^\dagger\frac{\mathrm{d}\hat{U}(t)}{\mathrm{d}t} = -\frac{1}{i\hbar}\hat{S}_0^\dagger\hat{H}_0\hat{U}(t) + \hat{S}_0^\dagger\frac{1}{i\hbar}\hat{H}\hat{U} \\
&= -\frac{1}{i\hbar}\hat{S}_0^\dagger\hat{H}_0\hat{U}(t) + \frac{1}{i\hbar}\hat{S}_0^\dagger\left(\hat{H}_0+\hat{V}\right)\hat{U} = -\cancel{\frac{1}{i\hbar}\hat{S}_0^\dagger\hat{H}_0\hat{U}} + \cancel{\frac{1}{i\hbar}\hat{S}_0^\dagger\hat{H}_0\hat{U}} + \frac{1}{i\hbar}\hat{S}_0^\dagger\hat{V}\hat{U} \\
&= \frac{1}{i\hbar}\hat{S}_0^\dagger\hat{V}\hat{U} = \frac{1}{i\hbar}\hat{S}_0^\dagger\hat{V}\hat{S}_0\hat{S}_0^\dagger\hat{U} = \frac{1}{i\hbar}\hat{V}_I\hat{U}_I'(t) \ .
\end{aligned}$$

Thus, the evolution is

$$\frac{\mathrm{d}\hat{U}_I'(t)}{\mathrm{d}t} = \frac{1}{i\hbar}\hat{V}_I\hat{U}_I'(t) \ . \tag{12.12}$$

This equation tells us that it is $\hat{U}_I'$ the operator which run the evolution of the states $|\psi_I(t)\rangle$. In fact,

$$\begin{aligned}
\frac{\mathrm{d}\hat{U}_I'(t)}{\mathrm{d}t}|\psi_I(0)\rangle &= \frac{1}{i\hbar}\hat{V}_I\hat{U}_I'(t)|\psi_I(0)\rangle \\
\frac{\mathrm{d}}{\mathrm{d}t}\left[\hat{U}_I'(t)|\psi_I(0)\rangle\right] &= \frac{1}{i\hbar}\hat{V}_I\left[\hat{U}_I'(t)|\psi_I(0)\rangle\right]
\end{aligned} \tag{12.13}$$

By comparing Eqn. (12.13) to Eqn. (12.7), it is clear that we can identify

$$|\psi_I(t)\rangle = \hat{U}'_I(t)\,|\psi_I(0)\rangle \ . \tag{12.14}$$

Eqn. (12.12) can be integrated to find

$$\hat{U}'_I(t,t_0) = \mathbb{1} - \frac{i}{\hbar}\int_{t_0}^{t} \hat{V}_I(\tau)\hat{U}'_I(\tau,t_0)\,\mathrm{d}\tau \ . \tag{12.15}$$

Taking the derivative of Eqn. (12.15) yields Eqn. (12.12). Eqn. (12.15) is called integral equation, and it can be solved iteratively. We get

$$\begin{aligned}\hat{U}'_I(t,t_0) = \mathbb{1} + (i\hbar)^{-1}\int_{t_0}^{t} \underbrace{\hat{V}_I(\tau)}_{(1)}\,\mathrm{d}\tau \\ + (i\hbar)^{-2}\int_{t_0}^{t}\mathrm{d}\tau\int_{t_0}^{\tau}\mathrm{d}\tau'\,\underbrace{\hat{V}_I(\tau)\hat{V}_I(\tau')}_{(2)}\,\hat{U}'_I(\tau',t_0) \ .\end{aligned} \tag{12.16}$$

From this new expression of $\hat{U}'_I$ we see that we have a term with one power of $\hat{V}_I$ [the term (1) in Eqn. (12.16)] and another term with two powers of $\hat{V}_I$ [the term (2) in Eqn. (12.16)]. If we think at $\hat{V}_I$ as a quantity proportional to some "small" parameter g, *i.e.*,

$$\hat{V}_I = g\hat{\tilde{V}} \ ,$$

then Eqn. (12.16) contains successive powers of g until the second one. By iterating again the procedure, *i.e.*, by inserting again $\hat{U}'_I(\tau',t_0)$ on the right-hand side of Eqn. (12.16) and using the expression of $\hat{U}'_I(t,t_0)$ given by that equation (see the left-hand side), we get

$$\hat{U}'_I(t,t_0) = \mathbb{1} + \sum_{n=1}^{\infty}\hat{U}^n_I(t,t_0) \ , \tag{12.17}$$

where

$$\hat{U}^n_I(t,t_0) = (i\hbar)^{-n}\int \mathrm{d}\tau_n\,\mathrm{d}\tau_{n-1}\cdots\mathrm{d}\tau_1\left[\hat{V}_I(\tau_n)\cdots\hat{V}_I(\tau_1)\right] \ ,$$

with $t > \tau_n > \tau_{n-1} > \ldots > \tau_1 > t_0$. All the integrals are computed between t_0 and the corresponding τ_i. Notice that in these integrals it appears only $\hat{V}_I$, *i.e.*, the

interaction part. Furthermore, if $\hat{V}_I$ is proportional to g, Eqn. (12.17) is practically an expansion on g, and thus it is a *perturbative* series.

Let us remember that we are interested in $\hat{U}(t)$ and not $\hat{U}'_I$. To obtain the former, we can simply use Eqn. (12.11) which can also be written as $\hat{U}(t) = \hat{S}_0 \hat{U}'_I$.

By using Eqn. (12.17) and the relation among $\hat{U}(t)$ and $\hat{U}'_I(t)$ mentioned above and the definition of $\hat{V}_I$, we get

$$\hat{U}(t, t_0) = \hat{U}^{(0)}(t, t_0) + \sum_{n=1}^{\infty} \hat{\tilde{U}}^n(t, t_0) \ , \tag{12.18}$$

where

$$\hat{\tilde{U}}^n(t, t_0) = (i\hbar)^{-n} \int \mathrm{d}\tau_n \, \mathrm{d}\tau_{n-1} \cdots \mathrm{d}\tau_1 \ \hat{S}^{(0)}(t, \tau_n) \hat{V}(\tau_n) \hat{S}^{(0)}(\tau_n, \tau_{n-1}) \\ \cdots \hat{S}^{(0)}(\tau_2, \tau_1) \hat{V}(\tau_1) \hat{S}^{(0)}(\tau_1, t_0) \ . \tag{12.19}$$

Here we have indicated with $\hat{S}^{(0)}$ what previously we wrote as $\hat{S}_0$ and also notice that $\hat{V}_I$ has been turned into $\hat{V}$ via the associated transformations. Note that the whole expression in Eqn.(12.19) above depends only on V and $\hat{S}^{(0)}$ but calculated at different instants and in different intervals of time.

Now, let us come back to the original unperturbed Hamiltonian $\hat{H}^{(0)}$ which can be solved exactly. Suppose that $\hat{H}^{(0)}$ has a discrete set of eigenvalues

$$\begin{cases} \hat{H}^{(0)} \left| a \right\rangle = E_a^{(0)} \left| a \right\rangle \\ \hat{H}^{(0)} \left| b \right\rangle = E_b^{(0)} \left| b \right\rangle \\ \qquad \vdots \\ \hat{H}^{(0)} \left| k \right\rangle = E_k^{(0)} \left| k \right\rangle \\ \qquad \vdots \end{cases} \tag{12.20}$$

Now, let us define the following two quantities:

$$\omega_{k,l} \equiv \frac{1}{\hbar} \left(E_k^{(0)} - E_l^{(0)} \right) \ , \\ V_{k,l} \equiv \langle k | \hat{V}(t) | l \rangle \ .$$

$\omega_{k,l}$ is referred to as "Bohr frequency". Assume that at the time t_0 the system is in the state $|a\rangle$. We want to compute the probability that at time t the system can be found in the state $|b\rangle$, *i.e.*,

$$P_{a\to b} = \left|\langle b|\hat{U}(t,t_0)|a\rangle\right|^2 \ .$$

If the perturbation were zero, *i.e.*, $\hat{V}=0$, then we would have

$$\hat{U}(t,t_0) = \mathrm{e}^{-i\hat{H}^{(0)}(t-t_0)/\hbar}$$

and the probability would be

$$P_{a\to b} = \left|\left\langle b\left|\,\mathrm{e}^{-i\hat{H}^{(0)}(t-t_0)/\hbar}\,\right|a\right\rangle\right|^2 = \left|\mathrm{e}^{-iE_a^{(0)}(t-t_0)/\hbar}\,\langle b|a\rangle\right|^2 = 0\ ,$$

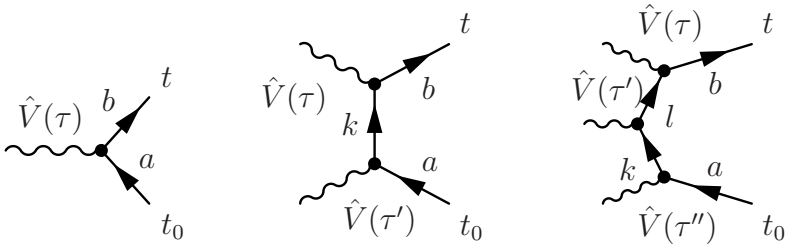
since $\langle b|a\rangle = 0$. If $\hat{V}\neq 0$, we can use Eqn. (12.18) and we obtain

$$\langle b|\hat{U}(t,t_0)|a\rangle = \sum_{n=1}^{\infty}\langle b|\hat{\tilde{U}}^n|a\rangle \ .$$

The first terms in this expression are:

$$\begin{aligned}
\langle b|\hat{\tilde{U}}^{(1)}|a\rangle &= -i\hbar^{-1}\int_{t_0}^{t}\mathrm{d}\tau\left[\underbrace{\mathrm{e}^{-iE_b^{(0)}(t-\tau)/\hbar}}_{\hat{S}^{(0)}(t,\tau)}V_{b,a}(\tau)\,\mathrm{e}^{-iE_a^{(0)}(\tau-t_0)/\hbar}\right]\\
\langle b|\hat{\tilde{U}}^{(2)}|a\rangle &= (i\hbar)^{-2}\sum_k\int_{t_0}^{t}\mathrm{d}\tau\int_{t_0}^{\tau}\mathrm{d}\tau'\Big[\,\mathrm{e}^{-iE_b^{(0)}(t-\tau)/\hbar}\,V_{b,k}(\tau)\,\mathrm{e}^{-iE_k^{(0)}(\tau-\tau')/\hbar}\\
&\quad\times V_{k,a}(\tau')\,\mathrm{e}^{-iE_a^{(0)}(\tau'-t_0)\hbar}\Big]\\
\langle b|\hat{\tilde{U}}^{(3)}|a\rangle &= (i\hbar)^{-3}\sum_k\sum_l\int_{t_0}^{t}\mathrm{d}\tau\int_{t_0}^{\tau}\mathrm{d}\tau'\int_{t_0}^{\tau'}\mathrm{d}\tau''\Big[\,\mathrm{e}^{-iE_b^{(0)}(t-\tau)/\hbar}\,V_{b,k}(\tau)\\
&\quad\times\mathrm{e}^{-iE_k^{(0)}(\tau-\tau')/\hbar}\,V_{k,l}(\tau')\,\mathrm{e}^{-iE_l^{(0)}(\tau'-\tau'')/\hbar}\\
&\quad\times V_{l,a}(\tau'')\,\mathrm{e}^{-iE_a^{(0)}(\tau''-t_0)/\hbar}\Big].
\end{aligned}$$

and so on. In order to obtain the various exponential terms in the expressions above we have made use of the states $\langle b|$ and $|a\rangle$ and the completeness relations $\sum_k$ in between.



(a) $\langle b|\hat{U}^{(1)}|a\rangle$ (b) $\langle b|\hat{U}^{(2)}|a\rangle$ (c) $\langle b|\hat{U}^{(3)}|a\rangle$

FIG. 12.1. Feynman diagrams

The matrix elements $\langle b|\hat{U}^{(1)}|a\rangle$, $\langle b|\hat{U}^{2}|a\rangle$, etc., have a graphical representation which is given in Fig. 12.1.

The intermediate states $|k\rangle$, $|l\rangle$ are called *virtual* (and we sum over these states) while $|a\rangle$ and $|b\rangle$ are named *physical* states.

Notice that the perturbation seems to occur at the time τ but we integrate all over the τ in between t_0 and t, and analogously for τ'.

If in the expansion above we stop at the n-th order, we get

$$P_{a\to b} \approx \left|\langle b|\hat{\tilde{U}}^{(1)}|a\rangle + \langle b|\hat{\tilde{U}}^{(2)}|a\rangle + \ldots + \langle b|\hat{\tilde{U}}^{(n)}|a\rangle\right|^2 .$$

Thus at the first order we have

$$P^{(1)}_{a\to b} = \left|\langle b|\hat{\tilde{U}}^{(1)}|a\rangle\right|^2 = \hbar^{-2}\left|\int_{t_0}^{t} \mathrm{e}^{i\omega_{b,a}\tau}\, V_{b,a}(\tau)\,\mathrm{d}\tau\right|^2 . \tag{12.21}$$

In the integral above there should be also an extra term $\mathrm{e}^{-iE_b^{(0)}t}\,\mathrm{e}^{iE_a^{(0)}t}$. This is a phase which does not depend on τ and it can be pushed outside the integral since the integral is done with respect to τ. Next we have to take the modulus of this phase but and its modulus is one.That is the reason why we neglected it.

Homework 12.1. Prove that $P^{(1)}_{a\to b} = P^{(1)}_{b\to a}$. This relation does not hold for higher perturbative orders.

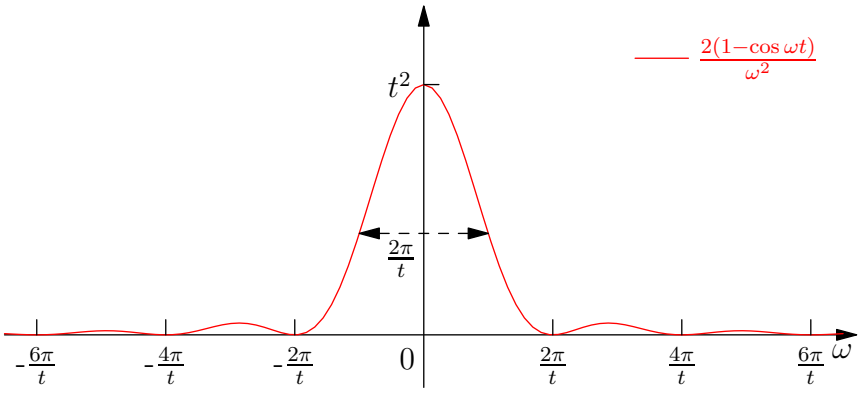


FIG. 12.2. Plot of $f(t,\omega)$.

If $\hat{V}$ does not depend on τ, it can be extracted from the integral and we get

$$P^{(1)}_{a\to b} \approx |V_{b,a}|^2 \, f(t,\omega_{b,a})/\hbar^2 \,, \tag{12.22}$$

where

$$f(t,\omega) \equiv \left| \int_0^t \mathrm{e}^{i\omega\tau} \, \mathrm{d}\tau \right|^2 = \frac{2\,(1-\cos\omega t)}{\omega^2} \,. \tag{12.23}$$

The plot of $f(t,\omega)$ as a function of ω is given in Fig. 12.2. The function has a narrow maximum at $\omega = 0$ with width $2\pi/t$, furthermore

$$\int_{-\infty}^{+\infty} f(t,\omega) \, \mathrm{d}\omega = 2\pi t \,. \tag{12.24}$$

Homework 12.2. Prove Eqn. (12.24).

As $t \to \infty$, the width $2\pi/t$ goes to zero and the value of the maximum at $t = 0$ goes to ∞. Thus the plot of $f(t,\omega)$ becomes like a Dirac's delta, and in order to reproduce Eqn. (12.24) we must have

$$f(t,\omega) \overset{t\to\infty}{\sim} 2\pi t \delta(\omega) \,.$$

By considering Eqn. (12.22), it can be seen that in order to obtain the transition probability we have to compute the square modulus of $\langle b|\hat{V}|a\rangle$ weighted by $f(t,\omega_{b,a})$, the latter being peaked around $\omega_{a,b} = 0$ and with $2\pi/t$ as width. As a consequence,

the transition is more likely to occur between states whose energy is near that of $|a\rangle$ (because the maximum of $f(t, \omega_{b,a})$ is at $\omega_{a,b} = 0$) and in an energy band given by

$$\delta E_0 \approx 2\pi\hbar/t \ .$$

So we can conclude that the pertubation leaves the energy unchanged within a bound of $\approx 2\pi\hbar/t$.

12.1.1 Structure of $f(t, \omega_{b,a})$

If $\omega_{b,a} = 0$, *i.e.*, if the state $|b\rangle$ has the same energy of $|a\rangle$, the perturbation preserves the energy and $f(t, \omega = 0) \sim t^2$, which means the longer is the time t the larger will be the probability of having a transition.

If $\omega_{b,a} \neq 0$, then $f(t, \omega)$ as a function of t is an oscillating function between 0 and $4/\omega_{b,a}^2$ with period $2\pi/\omega_{b,a}$, see Eqn. (12.23). So, $P_{a\to b}$ will oscillate with the same period [see, Eqn. (12.22)] around $2\left|V_{b,a}\right|^2/(E_{0,b} - E_{0,a})^2$. The factor of two comes from ω^2 of the denominator of $[1 - \cos\omega t]$. For small values of t with respect to $2\pi/\omega_{b,a}$, the behavior of $(1 - \cos\omega t)$ is like t^2.

12.1.2 Transition to nearby states

In this section we consider transitions to a set of nearby states.

Of course, so far we have considered only the case of a discrete spectrum, but nothing changes if we suppose that some states are surrounded by a continuous spectrum. Furthermore, let us suppose such states to be normalized according to

$$\langle b|b'\rangle = \frac{1}{n(b)}\delta(b - b') \ , \tag{12.25}$$

where $n(b)$ is a positive real-valued function of b, and b is some parameter used to label the states. The case $n(b) = 1$ gives the usual normalization condition

$$\langle b|b'\rangle = \delta(b - b') \ .$$

The projector onto a domain B is

$$\hat{P}_B = \int |b\rangle \langle b| \, n(b) \, \mathrm{d}b \, , \tag{12.26}$$

Homework 12.3. Show that the term $n(b)$ must be present in Eqn. (12.26) in order for $\hat{P}_B$ to fulfil the relation

$$\hat{P}_B \hat{P}_B = \hat{P}_B \, ,$$

which must hold because $\hat{P}_B$ is a projector.

Let $E(b)$ be the energy of the state $|b\rangle$. By changing the integration variable in Eqn. (12.26) from b to $E(b)$ we get

$$\hat{P}_B = \int_{B(E)} |b\rangle \langle b| \, \rho_b(E) \, \mathrm{d}E \, , \tag{12.27}$$

where

$$\rho_b(E) = n(b) \frac{\mathrm{d}b}{\mathrm{d}E} \, .$$

$\rho_b(E)$ is the density of states $|b\rangle$ at energy E, *i.e.*, the number of states $|b\rangle$ per unit of energy.

The transition probability from $|a\rangle$ to one of the states belonging to B is

$$\mathcal{P}_{a\to B} = \left| \hat{P}_B \hat{U}(t,0) \, |a\rangle \right|^2 \, ,$$

where $\hat{U}(t,0)$ is the time-evolution operator from time 0 to time t, and $\hat{P}_B$ is the projector onto B. Thus,

$$\mathcal{P}_{a\to B} = \left\langle a \left| \, \hat{U}^\dagger(t,0) \hat{P}_B^\dagger \, \right| \hat{P}_B \hat{U}(t,0) \, \right| a \Big\rangle = \langle a | \hat{U}^\dagger(t,0) \hat{P}_B \hat{U}(t,0) | a \rangle \, , \tag{12.28}$$

the latter equality follows from

$$\hat{P}_B^\dagger = \hat{P}_B \, , \qquad \hat{P}_B \hat{P}_B = \hat{P}_B \, .$$

Let us remember that $\hat{P}_B = \int_{B(E)} |b\rangle \langle b| \, \rho_b(E) \, \mathrm{d}E$, hence

$$\mathcal{P}_{a\to B} = \int_{B(E)} P_{a\to b} \rho_b(E) \, \mathrm{d}E \, ,$$

where

$$P_{a\to b} = \left|\langle b|\hat{U}(t,0)|a\rangle\right|^2 \ .$$

Now, for $P_{a\to b}$ we use Eqn. (12.22) calculated at first order in perturbation theory:

$$\mathcal{P}_{a\to B} \approx \frac{1}{\hbar^2}\int_{B(E)} |V_{b,a}|^2\,\rho_b(E) f(t,\omega_{b,a})\,\mathrm{d}E \ , \tag{12.29}$$

where $V_{b,a}$ depends on E via the parameter b.

Example

The frequency of oscillation in ω of the function f is $\frac{2\pi}{t}$ as can be easily seen by plotting f versus ω. Let us consider the transition to the levels b contained in the energy interval $(E_1 - \frac{\varepsilon}{2}, E_1 + \frac{\varepsilon}{2})$, and suppose ε to be small enough that both $V_{b,a}$ and ρ_b are constant in the interval. As a consequence we can pull them outside the integration in Eqn. (12.29).

Let us assume t to be sufficiently large to satisfy: $t \gg \frac{2\pi\hbar}{\varepsilon}$ or

$$\varepsilon \gg \frac{2\pi\hbar}{t} \ . \tag{12.30}$$

So the relation above indicates that ε is much bigger than the frequency of oscillation of f as a function of ω (if we put $\hbar = 1$). Let us turn to the integral :

$$\mathcal{P}_{a\to B} \approx \frac{1}{\hbar^2}\,|V_{b,a}|^2\,\rho_b \int_{B(E)} f(t,\omega_{b,a})\,\mathrm{d}E \ . \tag{12.31}$$

There are two possible cases:

1. that the maximum of f at $\omega_{b,a} = 0$ lies outside the integration domain . In this case, f inside the integral can be replaced by its mean value on several oscillations in ω. Remember the expression of f:

$$f(t,\omega) = \frac{2(1-\cos\omega t)}{\omega^2} \ .$$

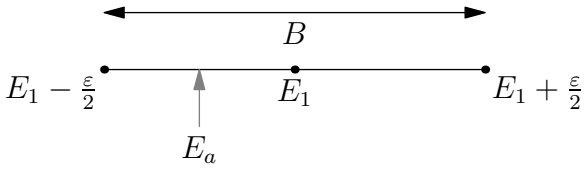


FIG. 12.3.

By taking the average of $\cos\omega t$ over many complete oscillations we get zero, thus the average of f is $2/\omega^2$. Hence,

$$\mathcal{P}_{a\to B} \approx 2\varepsilon\rho_b(E_1)\frac{|V_{a,b}|^2}{(E_1 - E_a)^2} \,, \tag{12.32}$$

where E_1 is the center of the interval

$$E_B = \left(E_1 - \frac{1}{2}\varepsilon, E_1 + \frac{1}{2}\varepsilon\right) \,.$$

2. that the maximum of f lies within the domain of integration, see Fig. 12.3. In this case, the leading contribution to the integral comes from the peak of f and only a small error results from taking the integral between $-\infty$ and $+\infty$. We know that

$$\int_{-\infty}^{+\infty} f(t,\omega)\,\mathrm{d}\omega = 2\pi t \,,$$

so

$$\mathcal{P}_{a\to B} \approx 2\pi\hbar^{-1}\left|V_{b,a}(E_a)\right|^2\rho_b(E_a)t \,. \tag{12.33}$$

In this case, $E_a \approx E_b$ since $E_a - E_b = \omega_{a,b}$ and $\omega \approx 0$ is the peak of f.

12.1.3 Transition probability per unit of time and Fermi's golden rule

The transition probability per unit of time is defined by

$$\tilde{\mathcal{P}}_{a\to B} = \frac{\mathrm{d}\mathcal{P}_{a\to B}}{\mathrm{d}t} \,.$$

Consider the two previous cases. In the first case, this quantity is zero because the transition probability is time-independent, Eqn. (12.32). These are the transitions which do not preserve the energy.

In the second case by taking the derivative with respect to t of Eqn. (12.33) we get

$$\tilde{\mathcal{P}}_{a\to B} = \frac{2\pi}{\hbar} \left|V_{b,a}\right|^2 \rho_b \ , \tag{12.34}$$

these are the transitions which preserve the energy of the states because $\omega_{ba} = 0$ is inside the integral we performed before. Eqn. (12.34) is called *Fermi's golden rule.*

ρ_b is the density of states whose energy is equal to that of the initial state, since everything is concentrated at the peak of f.

If we consider (12.30), *i.e.*,

$$\frac{t}{2\pi} \gg \frac{\hbar}{\varepsilon} \ , \tag{12.35}$$

t must be very long in order for Eqn. (12.34) to be valid. But, on the other end, we have used the perturbation theory at the first order and so it must be $\mathcal{P}_{a\to B} \ll 1$, *i.e.*,

$$\mathcal{P}_{a\to B} = \tilde{\mathcal{P}}^{(1)}_{a\to B} t \ll 1 \ . \tag{12.36}$$

So we conclude that indeed t cannot be too large.

Chapter 13

PATH INTEGRAL APPROACH TO QUANTUM MECHANICS

13.1 General derivation

This approach to quantum mechanics was presented by R. F. Feynman in his Ph.D. thesis of 1942. It was later published (1948) in Rev. Mod. Phys. with the title *"Space-time approach to non-relativistic Quantum Mechanics".*

Feynman wanted somehow a formulation in which "space-time" played a role and not just the Hilbert space, like in the traditional approach to quantum mechanics. His approach is very intuitive and helped in "visualizing" many quantum mechanical phenomena and in developing various techniques, like the Feynman diagrams, non-perturbative methods ($\hbar \to 0$, $N \to \infty$), etc. Somehow, Dirac had got close to the Feynman formulation of quantum mechanics in a paper in which he asked himself which is the role of the Lagrangian in quantum mechanics.

Let us first review the concept of *action* which everybody has learned in classical mechanics. Its definition is

$$S[x(t)] = \int_{(x_0,t_0)}^{(x_1,t_1)} \mathcal{L}(x(t), \dot{x}(t))\, \mathrm{d}t \ , \tag{13.1}$$

where $x(t)$ is *any* trajectory between (x_0, t_0) and (x_1, t_1), *not* necessarily the classical one, and $\mathcal{L}$ is the Lagrangian of the system. The action $S[x(t)]$ is what in mathematical terms is known as a *functional*. Remember that a functional is a map between a space of functions $x(t)$ and a set of numbers (the real or complex numbers, or similar structures). From Eqn. (13.1) one sees that $S[x(t)]$ is a functional because, once we

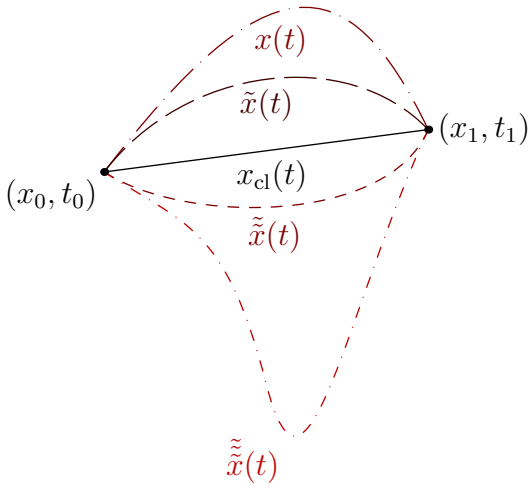


FIG. 13.1.

insert the function $x(t)$ on the right-hand side of Eqn. (13.1) (and perform the integration) we get a real number which is the value of the action on that trajectory. If we change the trajectory, we get a different number. A functional is indicated with square brackets, $S[x(t)]$, differently than a *function* whose argument is indicated with round brackets: $f(x)$. A function is a map between the set of numbers (real, complex, or similar) to another set of numbers (real, complex, etc.). So, if we restrict to the real numbers

$$\begin{aligned} &\text{Function:} \quad \mathbb{R} \to \mathbb{R} \,, \\ &\text{Functional:} \quad [\text{functions}] \to \mathbb{R} \,. \end{aligned}$$

Given these definitions, let us now see what is the path integral formulation of quantum-mechanics given by Feynman.

We know that in quantum mechanics a central element is the *transition kernel* to go from (x_0, t_0) to (x_1, t_1) which is defined as

$$K(x_1, t_1 | x_0, t_0) \equiv \langle x_1 | \, \mathrm{e}^{-i \frac{\hat{H}(t_1 - t_0)}{\hbar}} \, | x_0 \rangle \ . \tag{13.2}$$

What Feynman proved is the following formula:

$$K(x_1, t_1 | x_0, t_0) = \int_{(x_0, t_0)}^{(x_1, t_1)} \mathcal{D}\left[x(t)\right] \mathrm{e}^{\frac{i}{\hbar} S[x(t)]} \,, \tag{13.3}$$

where, on the right-hand side of Eqn. (13.3), the symbol $\int_{(x_0,t_0)}^{(x_1,t_1)} \mathcal{D}\left[x(t)\right]$ stays for a *functional integration* which "roughly" means the sum over all trajectories between (x_0, t_0) and (x_1, t_1).

So, in Eqn. (13.3) we insert a trajectory in $\mathrm{e}^{\frac{i}{\hbar}S[x(t)]}$, calculate this quantity and "sum" it to the same expression with a different trajectory inserted and so on for all trajectories between (x_0, t_0) and (x_1, t_1). This is the reason why this method is called *path-integral.* Note that all trajectories enter Eqn. (13.3) and not just the classical one.

We shall give a rigorous derivation of Eqn. (13.3) but for the moment let us try to understand a "more physical" reason why *trajectories* enter the expression of the quantum transition kernel.

Let us recall the double slit experiment, see Fig. 13.2. In Fig. 13.2(a) both slits A and B are open while in the other two figures only one is open. We know that the probabilities P_{AB}, P_A, P_B satisfy the inequality

$$P_{AB} \neq P_A + P_B \ ,$$

while for the probability amplitudes ψ_{AB}, ψ_A, ψ_B we have

$$\psi_{AB} = \psi_A + \psi_B \ . \tag{13.4}$$

Let us now put more screens with different openings, like in Fig. 13.3. Let us suppose we close the slits $1B$, $2A$, $2C$, $2D$,$3A$, $3B$ and let us call the associated wave function as

$$\psi_{\left[\begin{smallmatrix}1A\\2B\\3C\end{smallmatrix}\right]}$$

where we have indicated on the wave-functions which slits are open. For example for the wave-function above only the slits $1A$, $2B$ and $3C$ are open as shown in Fig. 13.3.

We can "associate" this amplitude with the path that join the slits $1A$, $2B$, $3C$. Let us also remember that Eqn. (13.4) can be generalized to any set of screens with

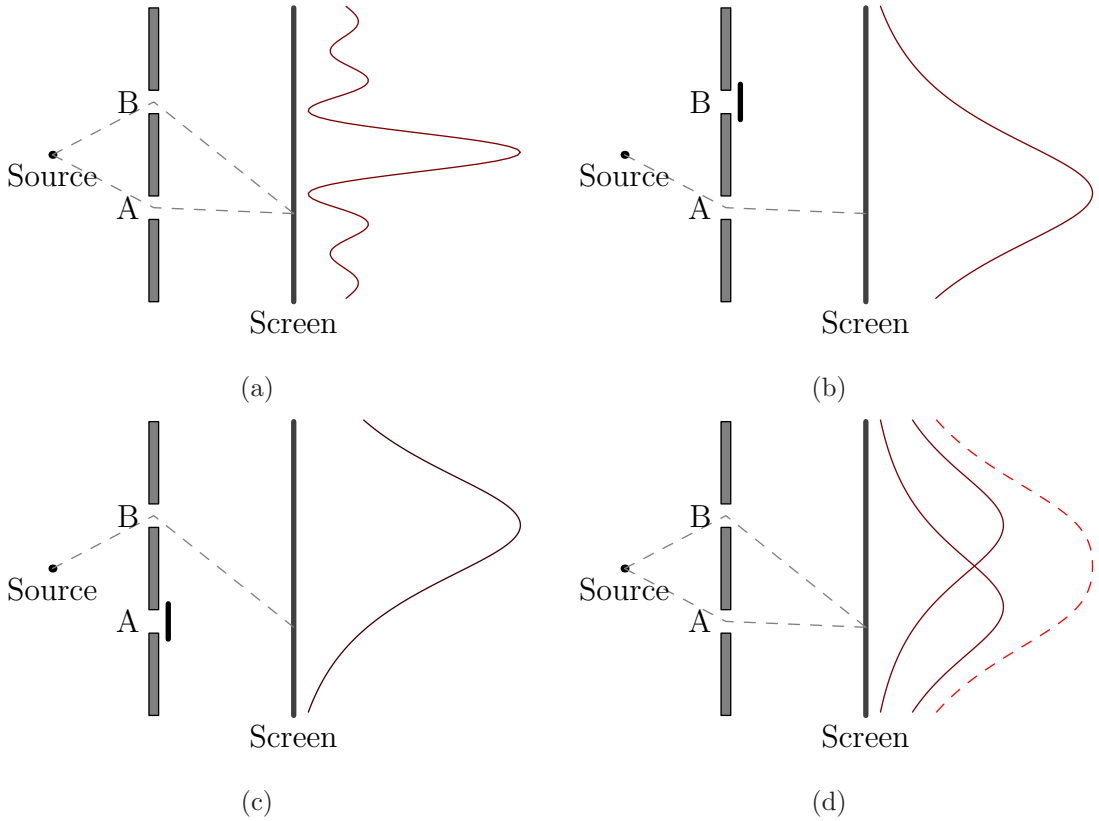


FIG. 13.2. (a) The probability P_{AB} with both slits open. (b) The probability P_A obtained keeping open the slit A (c) The probability P_B obtained keeping open the slit B (d) The probability $P_A + P_B$. We see that $P_{AB} \neq P_A + P_B$.

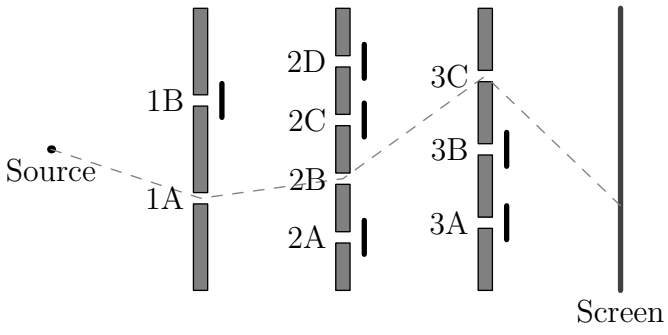


FIG. 13.3. More screens with different openings.

any set of open and closed slits, so:

$$\psi_{(\text{all slits open})} = \sum_i \psi^i_{(\text{some slits closed})} \ . \tag{13.5}$$

In turn the amplitudes on the right-hand side of Eqn. (13.5) can be written as the sum of all the amplitudes which have some of the remaining slits closed. The basic blocks of all these amplitudes will be those which have only one slit open per screen and to these we can associate a *path* running through the open slits. So Eqn. (13.5) can be formally written as

$$\psi = \sum_{(\text{paths})} \psi_{(\text{path})} \ , \tag{13.6}$$

where we have substituted the labels on the wave functions with the paths which join the open-slits on the various screens. Using this trick Feynman had the following idea: the open space between a source and a screen can be thought as if it were filled by an infinite set of screens each with an infinite set of slits. So now if we want the transition amplitudes from x_0 to x_1, *i.e.*, ψ_{x_0,x_1} we could write it as

$$\psi_{x_0,x_1} = \sum_{(\text{paths})} \psi_{(\text{path})} \ , \tag{13.7}$$

where the "(paths)" are the labels of the amplitude associated to a configuration of screens with only one slit open through which the path passes. Of course, the paths will be all possible paths between x_0 and x_1 because the screens have infinite slits. Let us stress that the paths in Eqn. (13.6) and (13.7) are nothing else than a "*symbol*" to indicate a set of slits open.

This somehow gives a physical intuition of why paths—even if they are only symbols or labels—enter the transition amplitudes. Of course, one cannot say that the particles follow one path or the other because, to check that, one should do a set of measurements along the whole path while in the transition $\langle x_0, t_0 | x_1, t_1 \rangle$ the only measurements are made at x_0 and x_1. What we can say from Eqn. (13.3) is that, if we do measurements only at x_0 and x_1 then the transition amplitude is the *sum*

of transition amplitudes each one associated to a different path between x_0 and x_1.If instead we do a set of infinte measures on all points of a path to find out if the particle follows that particular path, then the interference effect, between the various amplitudes entering the path-integral, will be destroyed.

Let us now give a rigorous derivation of Eqn. (13.3). For small time intervals it was first done by Dirac in 1932, for arbitrary time interval it was derived by Feynman in 1942 in his Ph.D. thesis.

Before proceeding in the derivation let us review some formula regarding the exponential of operators. If $\hat{A}$ and $\hat{B}$ are two operators, then $\mathrm{e}^{\hat{A}}\,\mathrm{e}^{\hat{B}}$ is *not* equal to $\mathrm{e}^{\hat{A}+\hat{B}}$ in general, like it would happen if $\hat{A}$ and $\hat{B}$ were numbers. The general formula was derived by Baker and Hausdorff and is

$$\mathrm{e}^{\hat{A}}\,\mathrm{e}^{\hat{B}} = \mathrm{e}^{H(\hat{A},\hat{B})}\ , \tag{13.8}$$

where

$$H(\hat{A},\hat{B}) = \hat{A}+\hat{B}+\frac{1}{2}[\hat{A},\hat{B}]+\frac{1}{12}\left[\hat{A},[\hat{A},\hat{B}]\right]+\frac{1}{12}\left[\hat{B},[\hat{B},\hat{A}]\right]+\ldots \tag{13.9}$$

if $\hat{A}$ and $\hat{B}$ commutes then $H(\hat{A},\hat{B}) = \hat{A}+\hat{B}$ like in the case of c-numbers.

Let us now go back to physics and calculate the *transition kernel* which is defined as

$$K(x,t|x_0,0) \equiv \langle x|\,\mathrm{e}^{-\frac{i}{\hbar}t\hat{H}}\,|x_0\rangle\ , \tag{13.10}$$

where $\hat{H}$ is the Hamiltonian of the system* If we divide the interval of time t in N sub-intervals we can write, using the Baker-Hausdorff formula, the following equality

$$\exp\left[-\frac{i}{\hbar}t\hat{H}\right] = \left\{\exp\left[-\frac{it}{\hbar N}\hat{H}\right]\right\}^N\ . \tag{13.11}$$

This is so because the operators $t\hat{H}/N$ commute among themselves in the Baker-Hausdorff formula.

* Note that this is not the *transition amplitude* which would be $\langle xt|x_0 0\rangle = \langle x|\,\mathrm{e}^{\frac{i}{\hbar}t\hat{H}}\,|x_0\rangle$.

Let us now remember that the Hamiltonian $\hat{H}$ is the sum of two parts $\hat{H} = \hat{A} + \hat{B}$, which do not commute because $\hat{A} = \hat{p}^2/2m$ and $\hat{B} = \hat{V}(\hat{x})$. So, using again the Baker-Hausdorff formula we can write

$$\begin{aligned} \exp\left[-\frac{i}{\hbar}\frac{t}{N}\hat{H}\right] &= \exp\left[-\frac{i}{\hbar}\frac{t}{N}\left(\hat{A}+\hat{B}\right)\right] \\ &= \exp\left[-\frac{i}{\hbar}\frac{t}{N}\hat{B}\right]\exp\left[-\frac{i}{\hbar}\frac{t}{N}\hat{A}\right] + \mathcal{O}\left(\left(\frac{t}{N}\right)^2\right) , \end{aligned} \quad (13.12)$$

where the terms $\mathcal{O}((t/N)^2)$ are those that come from the commutators of $t\hat{A}/N$ and $t\hat{B}/N$ present in Eqn. (13.9). Of course, if we take the limit $N \to \infty$ those terms are negligible with respect to the first. So combining Eqn. (13.12) with Eqn. (13.11) we can write

$$\left\langle x \left| \, \mathrm{e}^{-\frac{i}{\hbar}t\hat{H}} \, \right| x_0 \right\rangle = \lim_{N\to\infty} \left\langle x \left| \left[\mathrm{e}^{-\frac{it}{\hbar N}\hat{B}}\,\mathrm{e}^{-\frac{it}{\hbar N}\hat{A}}\right]^N \right| x_0 \right\rangle \quad (13.13)$$

Let us now write all terms of the operator

$$\left[\exp\left(-\frac{it\hat{B}}{\hbar N}\right)\exp\left(-\frac{it}{\hbar N}\hat{A}\right)\right]^N ,$$

and so Eqn. (13.13) reads

$$\left\langle x \left| \underbrace{\left[\mathrm{e}^{-\frac{it}{\hbar N}\hat{B}}\,\mathrm{e}^{-\frac{it}{\hbar N}\hat{A}}\right] \times \left[\mathrm{e}^{-\frac{it}{\hbar N}\hat{B}}\,\mathrm{e}^{-\frac{it}{\hbar N}\hat{A}}\right] \times \cdots \times \left[\mathrm{e}^{-\frac{it}{\hbar N}\hat{B}}\,\mathrm{e}^{-\frac{it}{\hbar N}\hat{A}}\right]}_{N \text{ times}} \right| x_0 \right\rangle . \quad (13.14)$$

Next let us divide the interval from x_0 to x in N intervals labelled by the points $x_1, x_2 \cdots x_{N-1}$ and corresponding to the N intervals of time. Let us now insert in Eqn. (13.14) after the first square bracket a completeness of the form $\int \mathrm{d}x_{N-1} \, |x_{N-1}\rangle \langle x_{N-1}|$, where x_{N-1} is the point before x in Fig. 13.4, and we continue by inserting the completeness $\int \mathrm{d}x_{N-2} \, |x_{N-2}\rangle \langle x_{N-2}|$ after the second square bracket in Eqn. (13.14) and so on. What we get is that Eqn. (13.13) can be written as

$$\begin{aligned} \langle x| \, \mathrm{e}^{-\frac{i}{\hbar}t\hat{H}} \, |x_0\rangle = \lim_{N\to\infty} \int \mathrm{d}x_{N-1} \cdots \mathrm{d}x_1 \Bigg\{ & \left\langle x \left| \, \mathrm{e}^{-\frac{it}{\hbar N}\hat{B}}\,\mathrm{e}^{-\frac{it}{\hbar N}\hat{A}} \, \right| x_{N-1} \right\rangle \\ & \times \left\langle x_{N-1} \left| \, \mathrm{e}^{-\frac{it}{\hbar N}\hat{B}}\,\mathrm{e}^{-\frac{it}{\hbar N}\hat{A}} \, \right| x_{N-2} \right\rangle \times \cdots \times \left\langle x_1 \left| \, \mathrm{e}^{-\frac{it}{\hbar N}\hat{B}}\,\mathrm{e}^{-\frac{it}{\hbar N}\hat{A}} \, \right| x_0 \right\rangle \Bigg\} . \end{aligned} \quad (13.15)$$

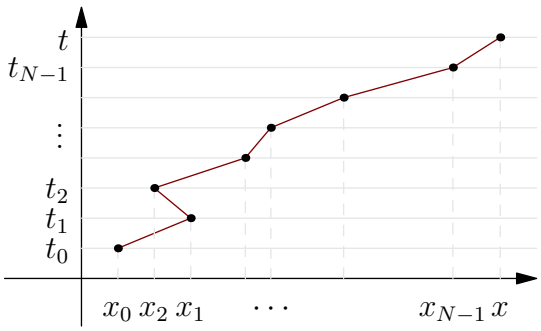


FIG. 13.4. The discretized approximation of the path integral

All this can be put in the compact form

$$\langle x|\, \mathrm{e}^{-\frac{i}{\hbar}t\hat{H}}\, |x_0\rangle = \lim_{N\to\infty} \int \prod_{j=1}^{N} Q_{j,j-1} \prod_{j=1}^{N-1} dx_j \ ,$$

where $Q_{j,j-1}$ is

$$Q_{j,j-1} \equiv \left\langle x_j \left| \mathrm{e}^{-\frac{it}{\hbar N}\hat{B}}\, \mathrm{e}^{-\frac{it}{\hbar N}\hat{A}} \right| x_{j-1} \right\rangle \ .$$

In the expression above we can apply $\hat{B} = \hat{V}(\hat{x})$ to the state $\langle x_j|$ and get

$$Q_{j,j-1} = \mathrm{e}^{-\frac{it}{\hbar N}V(x_j)} \left\langle x_j \left| \mathrm{e}^{-\frac{it}{\hbar N}\hat{A}} \right| x_{j-1} \right\rangle \ . \tag{13.16}$$

The operator $\hat{A}$ is instead $\hat{p}^2/2m$ so we cannot apply it directly to the state $|x_{j-1}\rangle$. What we will do is to insert a completeness $\int |p\rangle\, \langle p|\, \mathrm{d}p = \mathbb{1}$ before the state $|x_{j-1}\rangle$ in Eqn. (13.16). We get in this way

$$\left\langle x_j \left| \mathrm{e}^{-\frac{it}{\hbar N}\hat{A}} \right| x_{j-1} \right\rangle = \int \left\langle x_j \left| \mathrm{e}^{-\frac{it}{\hbar N}\frac{\hat{p}^2}{2m}} \right| p \right\rangle \langle p|x_{j-1}\rangle\, \mathrm{d}p \tag{13.17}$$

Remembering that

$$\langle p|x\rangle = \frac{1}{\sqrt{2\pi\hbar}}\, \mathrm{e}^{-\frac{i}{\hbar}px} \ ,$$

and applying $\hat{p}^2$ on $|p\rangle$ we get from Eqn. (13.17) the following expression

$$\left\langle x_j \left| \mathrm{e}^{-\frac{it}{\hbar N}\hat{A}} \right| x_{j-1} \right\rangle = \sqrt{\frac{1}{2\pi\hbar}} \int \mathrm{e}^{-\frac{i}{\hbar}\left\{\frac{t}{N}\frac{p^2}{2m} - p(x_j - x_{j-1})\right\}}\, \mathrm{d}p$$

which, by performing the integration in p, becomes

$$\left\langle x_j \left| \, \mathrm{e}^{-\frac{it}{\hbar N}\hat{A}} \, \right| x_{j-1} \right\rangle = \left[\frac{m}{2\pi i\hbar\frac{t}{N}}\right]^{\frac{1}{2}} \mathrm{e}^{\frac{im}{2\hbar}\frac{N}{t}(x_j - x_{j-1})^2} \ . \tag{13.18}$$

Inserting this in Eqn. (13.16) and next in Eqn. (13.15) we get

$$K(x,t|x_0,0) = \lim_{N\to\infty}\left(\frac{m}{2\pi i\hbar\varepsilon}\right)^{\frac{N}{2}} \int \mathrm{e}^{\frac{i}{\hbar}\sum_{j=1}^{N}\left\{\frac{m}{2\varepsilon}(x_j - x_{j-1})^2 - \varepsilon V(x_j)\right\}} \prod_{j=1}^{N-1} \mathrm{d}x_j \ , \tag{13.19}$$

where $\varepsilon = t/N$. Let us note two things:

1. In the exponential of Eqn. (13.19) we have the discretized form of the action. In fact,

$$\begin{aligned}
\int \mathcal{L}\,\mathrm{d}t \ &\Rightarrow \ \sum_j \mathcal{L}(x_j,\dot{x}_j)\varepsilon \\
&= \ \sum_j \left\{\frac{m}{2}\dot{x}_j^2 - V(x_j)\right\}\varepsilon \\
&= \ \sum_j \left\{m\frac{(x_j - x_{j-1})^2}{2\varepsilon^2}\varepsilon - V(x_j)\varepsilon\right\} \\
&= \ \sum_j \left\{m\frac{(x_j - x_{j-1})^2}{2\varepsilon} - V(x_j)\varepsilon\right\} \ .
\end{aligned}$$

2. If we look at Fig. 13.4 we see that a trajectory in the discretized form is the broken line we have drawn between x_0 and x. If we keep x_0 and x fixed and move the intermediate points (x_1,t_1), (x_2,t_2), $\ldots$, (x_{N-1},t_{N-1}), we get all possible trajectories between x_0 and x. This is exactly what is achieved by the integrations over x_j, $j = 1,2,\ldots,N-1$ in Eqn. (13.19). So the sum over all trajectories indicated by the functional integration $\int \mathcal{D}\,[x(t)]$ in Eqn. (13.3) is actually realized by the measure of integration contained in Eqn. (13.19), *i.e.*,

$$\int \mathcal{D}\,[x(t)] = \lim_{N\to\infty}\left(\frac{m}{2\pi i\hbar\varepsilon}\right)^{\frac{N}{2}} \int \prod_{j=1}^{N-1} \mathrm{d}x_j \ .$$

So we can conclude that Eqn. (13.19) is nothing else than the discretized form of the expression (13.3) and that the *functional* integral can be reduced to an infinite set of standard Lebesque integrals.

The path-integral not only brings to light the role of the *non-classical* trajectories in quantum mechanics but also the role of the action in quantum mechanics. In fact the action had played a role in classical mechanics but never in quantum mechanics. This idea was the one which triggered in 1932 the work of Dirac.

The last thing we want to bring to the attention of the reader is which are the paths which contribute most in the path-integral. We shall show that are the paths which are *continuous* but *nowhere-differentiable*.

The proof goes as follows. If we look at the kinetic piece in Eqn. (13.19) we see that in the limit of $\varepsilon \to 0$ we must have that $(x_j - x_{j-1})^2/\varepsilon$ remains finite otherwise the exponent would not make sense. This means

$$\frac{(x_j - x_{j-1})^2}{\varepsilon} \to \text{ finite} ,$$

i.e.,

$$(x_j - x_{j-1})^2 \sim \varepsilon , \qquad (\Delta x)^2 \sim \Delta t . \tag{13.20}$$

From Eqn. (13.20) we notice two things:

1. When $\Delta t \to 0$ we have $\Delta x \to 0$, which means that the paths are continuous;

2. the velocities $\Delta x/\Delta t$ goes as $1/\Delta x$ because, using Eqn. (13.20) , we have:

$$\frac{\Delta x}{\Delta t} \sim \frac{\Delta x}{(\Delta x)^2} = \frac{1}{\Delta x} ,$$

so when $\Delta x \to 0$ the velocity has a singularity. As this happens at every point it means the path is non-differentiable.

The paths which are differentiable, *i.e.*:

$$\frac{\Delta x}{\Delta t} \to \text{ finite}$$

i.e., $\Delta x \sim \Delta t$ have a kinetic piece in the action which goes as follows:

$$\frac{(\Delta x)^2}{\Delta t} \sim \frac{(\Delta t)^2}{\Delta t} \sim \Delta t \to 0 \, .$$

So in the continuum $\Delta t \to 0$ these kinetic pieces go as e^0. These are constants that can be factorized out of the path-integral and we can get rid of them in the normalization. So these paths do not give any phase which may interphere with the other paths and create typical quantum mechanical effects.

The non-differentiable paths are a typical indicator of quantum mechanical effects. In classical mechanics the paths are smooth instead. More details and examples on the path-integrals can be found in the many books on the subject.

13.2 Time dependent perturbation theory via path integrals and proto-Feynman's diagrams

In this section we will derive the time-dependent perturbative expansion for the time-evolution kernel via path-integrals, without using the interaction picture.

We have seen that the transition Kernel between $x = x_a$ and $x = x_b$ from time t_a to time t_b has the following path integral expression* :

$$K_V(b,a) = \int_a^b \mathcal{D}\left[x(t)\right] \exp\left\{\frac{i}{\hbar}\left[\int_{t_a}^{t_b}\left(\frac{m\dot{x}^2}{2} - V(x,t)\right)\mathrm{d}t\right]\right\} \, . \tag{13.21}$$

If we suppose that the potential is small or better that $\int_{t_a}^{t_b} V(x,t)\,\mathrm{d}t$ is small with respect to $\hbar$, then we can expand the integrand as follows:

$$\exp\left[-\frac{i}{\hbar}\int_{t_a}^{t_b} V(x,t)\,\mathrm{d}t\right] = 1 - \frac{i}{\hbar}\int_{t_a}^{t_b} V(x,t)\,\mathrm{d}t + \frac{1}{2!}\left(\frac{i}{\hbar}\right)^2\left[\int_{t_a}^{t_b} V(x,t)\,\mathrm{d}t\right]^2 + \ldots \tag{13.22}$$

* If we want to get the *transition amplitude* and not the *transition kernel*, we should just put a minus sign in front of the Lagrangian.

Of course, $\int_{t_a}^{t_b} V(x,t)\,\mathrm{d}t$ is a functional, so the statement that $\int_{t_a}^{t_b} V(x,t)\,\mathrm{d}t$ should be "small" with respect to $\hbar$ needs to be better clarified but we will leave this clarification for more advanced courses.

Using the above expansion, we can rewrite $K_V(b,a)$ in the following way:

$$K_V(b,a) = K^{(0)}(b,a) + K^{(1)}(b,a) + K^{(2)}(b,a)\ldots\ , \tag{13.23}$$

where

$$\begin{aligned}
K^{(0)}(b,a) &= \int_a^b \mathcal{D}\left[x(t)\right]\left[\exp\left(\frac{i}{\hbar}\int_{t_a}^{t_b}\frac{m\dot{x}^2}{2}\,\mathrm{d}t\right)\right]\ , \\
K^{(1)}(b,a) &= -\frac{i}{\hbar}\int \mathcal{D}\left[x(t)\right]\left\{\left(\int_{t_a}^{t_b} V(x(s),s)\,\mathrm{d}s\right)\exp\left[\frac{i}{\hbar}\int_{t_a}^{t_b}\frac{m\dot{x}^2}{2}\,\mathrm{d}t\right]\right\}\ , \\
K^{(2)}(b,a) &= -\frac{1}{2\hbar^2}\int \mathcal{D}\left[x(t)\right]\left\{\left[\int_{t_a}^{t_b} V(x(s),s)\,\mathrm{d}s\right]\left[\int_{t_a}^{t_b} V(x(s'),s')\,\mathrm{d}s'\right]\right. \\
&\qquad \left.\times\exp\left[\frac{i}{\hbar}\int_{t_a}^{t_b}\frac{m\dot{x}^2}{2}\,\mathrm{d}t\right]\right\}
\end{aligned}$$

Let us now proceed to evaluate the various terms. $K^{(0)}(b,a)$ is the free particle transition kernel. $K^{(1)}(b,a)$ can instead be written as

$$K^{(1)}(b,a) = -\frac{i}{\hbar}\int_{t_a}^{t_b} F(s)\,\mathrm{d}s\ , \tag{13.24}$$

where

$$F(s) = \int_a^b \mathcal{D}\left[x(t)\right] V(x(s),s)\exp\left(\frac{i}{\hbar}\int_{t_a}^{t_b}\frac{m\dot{x}^2}{2}\,\mathrm{d}t\right)\ . \tag{13.25}$$

Basically, $F(s)$ is the path-integral of the free particle but with inserted at $t=s$ the potential. So the time-evolution before $t=s$ is the one of the free particle, at $t=s$ it gets "perturbed" by $V(x(s),s)$, afterwards it is again the evolution of the free particle. The picture that we can associate to Eqn. (13.25) is the one in Fig. 13.5. There, we have supposed that the point reached at time s is x_c. In Eqn. (13.25) we are effectively integrating over x_c and as a consequence $F(s)$ can be written in the following manner:

$$F(t_c) = \int_{-\infty}^{+\infty} K^{(0)}(b,c)V(x_c,t_c)K^{(0)}(c,a)\,\mathrm{d}x_c\ .$$

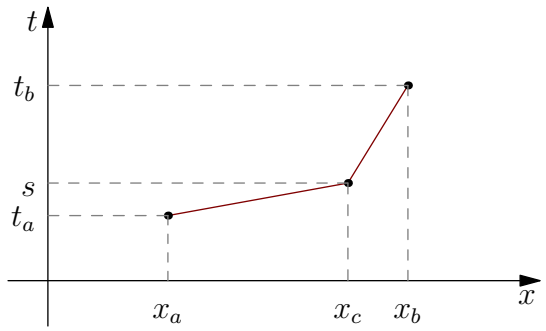


FIG. 13.5. Schematic representation of $F(s)$.

Using this, $K^{(1)}(b,a)$ can be represented as

$$K^{(1)}(b,a) = -\frac{i}{\hbar}\int_{t_a}^{t_b}\int_{-\infty}^{+\infty} K^{(0)}(b,c)V(x_c,t_c)K^{(0)}(c,a)\,\mathrm{d}x_c\,\mathrm{d}t_c\ . \tag{13.26}$$

The pictures that can be drawn associated to the perturbation series of Eqn. (13.23) are, in analogy to that in Fig. 13.5, the ones in Fig. 13.6. These are *proto-Feynman diagrams* in a potential theory.

We shall now derive an integral equation for $K_V(b,a)$. Using Eqn. (13.26) and the analog for higher orders we get easily the following expression (where $\int \mathrm{d}\tau_c = \int \mathrm{d}x_c\,\mathrm{d}t_c$):

$$\begin{aligned} K_V(b,a) = K^{(0)}(b,a) &- \frac{i}{\hbar}\int K^{(0)}(b,c)V(x_c,t_c)K^{(0)}(c,a)\,\mathrm{d}\tau_c \\ &+ \left(-\frac{i}{\hbar}\right)^2 \int\int K^{(0)}(b,c)V(x_c,t_c)K^{(0)}(c,d)V(x_d,t_d)K^{(0)}(d,a)\,\mathrm{d}\tau_c\,\mathrm{d}\tau_d \\ &+\cdots\ , \end{aligned} \tag{13.27}$$

which can also be written as

$$\begin{aligned} K_V(b,a) = K^{(0)}(b,a) &- \frac{i}{\hbar}\int K^{(0)}(b,c)V(x_c,t_c) \\ &\times\left[K^{(0)}(c,a) - \frac{i}{\hbar}\int K^{(0)}(c,d)V(x_d,t_d)K^{(0)}(d,a)\,\mathrm{d}\tau_d + \ldots\right]\mathrm{d}\tau_c + \ldots \end{aligned} \tag{13.28}$$

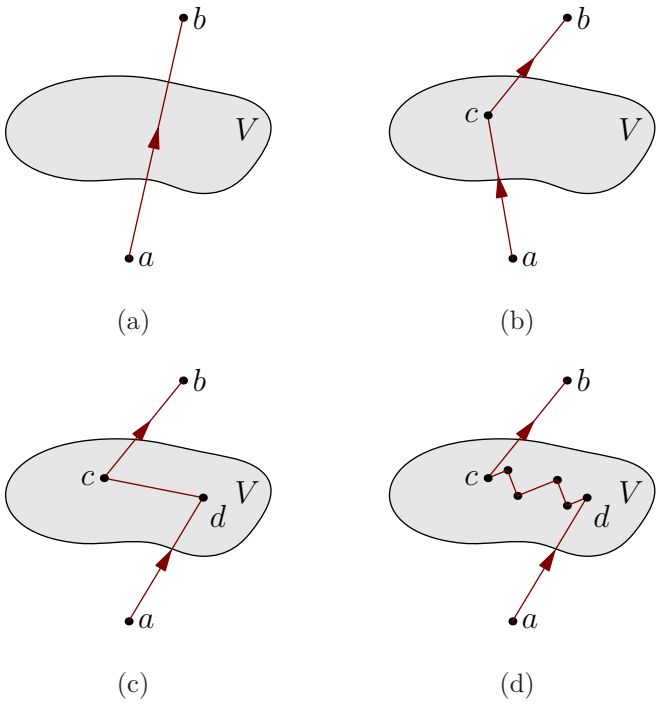


FIG. 13.6. Schematic representation of series expansion Eqn. (13.23).

If we look at Eqn. (13.28) and in particular at the quantity inside the square brackets, we notice that Eqn. (13.28) can be re-written as

$$K_V(b,a) = K^{(0)}(b,a) - \frac{i}{\hbar}\int K^{(0)}(b,c)V(x_c,t_c)K_V(c,a)\,\mathrm{d}\tau_c\,. \tag{13.29}$$

This is an integral equation for K_V. It is completely equivalent to the differential Schrödinger equation: is basically the *integral* version of it

Chapter 14

SEMICLASSICAL (WKB) QUANTUM MECHANICS

This chapter is mostly based on the Bachelor Thesis of Taracchini Andrea, University of Trieste.

14.1 Introduction

The wave function $\psi(\boldsymbol{x},t)$ of a physical system is a complex-valued quantity and so it can be written in general as

$$\psi(\boldsymbol{x},t) = A(\boldsymbol{x},t)\,\mathrm{e}^{\frac{i}{\hbar}S(\boldsymbol{x},t)}\ , \tag{14.1}$$

where $A(\boldsymbol{x},t)$ and $S(\boldsymbol{x},t)$ are real-valued functions. The presence of the factor $\hbar$ at exponent is just a convention and in this manner S acquires the dimension of an action.

Let us now calculate various expressions entering the Schroedinger equation:

$$\begin{aligned}
\nabla^2\psi(\boldsymbol{x},t) &= \nabla^2\left[A(\boldsymbol{x},t)\,\mathrm{e}^{\frac{i}{\hbar}S(\boldsymbol{x},t)}\right] \\
&= \left\{\nabla^2 A(\boldsymbol{x},t) + 2\frac{i}{\hbar}\,\boldsymbol{\nabla}A(\boldsymbol{x},t)\cdot\boldsymbol{\nabla}S(\boldsymbol{x},t) + \frac{i}{\hbar}A(\boldsymbol{x},t)\nabla^2 S(\boldsymbol{x},t)\right. \\
&\quad \left. - \frac{1}{\hbar^2}A(\boldsymbol{x},t)\left[\boldsymbol{\nabla}S(\boldsymbol{x},t)\right]^2\right\}\mathrm{e}^{\frac{i}{\hbar}S(\boldsymbol{x},t)}\ , \\
\frac{\partial\psi(\boldsymbol{x},t)}{\partial t} &= \frac{\partial}{\partial t}\left[A(\boldsymbol{x},t)\,\mathrm{e}^{\frac{i}{\hbar}S(\boldsymbol{x},t)}\right] \\
&= \left\{\frac{\partial A(\boldsymbol{x},t)}{\partial t} + \frac{i}{\hbar}A(\boldsymbol{x},t)\frac{\partial S(\boldsymbol{x},t)}{\partial t}\right\}\mathrm{e}^{\frac{i}{\hbar}S(\boldsymbol{x},t)}\ .
\end{aligned}$$

Inserting Eqn. (14.1) into the time-dependent Schrödinger equation

$$i\hbar\frac{\partial\psi(\boldsymbol{x},t)}{\partial t} = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{x},t)\right]\psi(\boldsymbol{x},t)\ , \tag{14.2}$$

we get

$$i\hbar\frac{\partial A(\boldsymbol{x},t)}{\partial t} - A(\boldsymbol{x},t)\frac{\partial S(\boldsymbol{x},t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2 A(\boldsymbol{x},t) - \frac{i\hbar}{m}\boldsymbol{\nabla}A(\boldsymbol{x},t)\cdot\boldsymbol{\nabla}S(\boldsymbol{x},t) - \frac{i\hbar}{2m}A(\boldsymbol{x},t)\nabla^2 S(\boldsymbol{x},t) + \frac{1}{2m}A(\boldsymbol{x},t)\left[\boldsymbol{\nabla}S(\boldsymbol{x},t)\right]^2 + V(\boldsymbol{x},t)A(\boldsymbol{x},t). \quad (14.3)$$

Equating the real and imaginary parts we get the two equations

$$\frac{\partial S(\boldsymbol{x},t)}{\partial t} + \frac{1}{2m}\left[\boldsymbol{\nabla}S(\boldsymbol{x},t)\right]^2 + V(\boldsymbol{x},t) = \frac{\hbar^2}{2m}\frac{\nabla^2 A(\boldsymbol{x},t)}{A(\boldsymbol{x},t)}\ , \quad (14.4)$$

$$m\frac{\partial A(\boldsymbol{x},t)}{\partial t} + \boldsymbol{\nabla}A(\boldsymbol{x},t)\cdot\boldsymbol{\nabla}S(\boldsymbol{x},t) + \frac{1}{2}A(\boldsymbol{x},t)\nabla^2 S(\boldsymbol{x},t) = 0\ . \quad (14.5)$$

Eqn. (14.5) is nothing but the continuity equation for the probability associated to the Schrödinger equation:

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0\ . \quad (14.6)$$

In fact, defining the probability ρ and the corresponding current $\boldsymbol{j}$ as

$$\rho = |\psi(\boldsymbol{x},t)|^2 = \left[A(\boldsymbol{x},t)\right]^2\ , \quad (14.7)$$

$$\boldsymbol{j} = \Re\left[\psi^*(\boldsymbol{x},t)\frac{\hbar}{im}\boldsymbol{\nabla}\psi(\boldsymbol{x},t)\right] = \left[A(\boldsymbol{x},t)\right]^2\frac{\boldsymbol{\nabla}S(\boldsymbol{x},t)}{m}\ , \quad (14.8)$$

Eqn. (14.5) becomes

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0\ .$$

Let us now turn to Eqn. (14.4). This equation is similar to an equation of classical mechanics known as Hamilton-Jacobi equation, which is

$$\frac{\partial S_0(\boldsymbol{x},t)}{\partial t} + \frac{\left[\boldsymbol{\nabla}S_0(\boldsymbol{x},t)\right]^2}{2m} + V(\boldsymbol{x},t) = 0\ . \quad (14.9)$$

Notice that Eqn. (14.4) for $S(\boldsymbol{x},t)$ reduces to the Hamilton-Jacobi equation in the limit $\hbar\to 0$.

People sometimes say that the Schrödinger equation is the quantum correction to the Hamilton-Jacobi, but this statement is not correct because the Schrödinger equation involves both $A(\boldsymbol{x},t)$ and $S(\boldsymbol{x},t)$ while the Hamilton-Jacobi involves only S. We have chosen to put the sub index "0" in the solution S_0 of the Hamilton-Jacobi equation, since it is the solution at order zero in $\hbar^2$ of Eqn. (14.4).

14.2 Review of the Hamilton-Jacobi formalism of classical mechanics

Let us briefly review the Hamilton-Jacobi formalism of classical mechanics. The Hamiltonian description of a system with f degrees of freedom involves f configurational variables $q_1, \ldots, q_f$ and their conjugated momenta $p_1, \ldots, p_f$. The dynamics is governed by the Hamilton equations of motions, which can be written as*

$$\dot{q}_k = \{q_k, H\}_{\text{P.B.}} , \tag{14.10}$$

$$\dot{p}_k = \{p_k, H\}_{\text{P.B.}} , \tag{14.11}$$

$k = 1, \ldots, f$, where H is the Hamiltonian.

An important feature of the Hamiltonian formalism is that it allows transformations on the phase space of the system

$$Q_k = Q_k(q_1, \ldots, q_f, p_1, \ldots, p_f, t) , \tag{14.12}$$

$$P_k = P_k(q_1, \ldots, q_f, p_1, \ldots, p_f, t) , \tag{14.13}$$

which mix the q's and p's yet preserve the Hamiltonian form of the equations of motion, *i.e.*, the equations of motion have the same form also with respect to the new set of transformed variables. Such transformations are referred to as *canonical* transformations.

For a transformation to be a canonical the fundamental condition (called *symplectic* condition) that must be satisfied is

$$\mathsf{MJM}^T = \mathsf{J} , \tag{14.14}$$

* Here, we have used the Poisson brackets, defined as

$$\{F, G\}_{\text{P.B.}} = \sum_{k=1}^{f} \left(\frac{\partial F}{\partial q_k} \frac{\partial G}{\partial p_k} - \frac{\partial F}{\partial p_k} \frac{\partial G}{\partial q_k} \right) ,$$

for arbitrary differentiable functions F and G defined on the phase space of the system.

where M is the Jacobian of the transformation, *i.e.*:

$$\mathsf{M} = \frac{\partial\left(Q_1,\ldots,Q_f,P_1,\ldots,P_f\right)}{\partial\left(q_1,\ldots,q_f,p_1,\ldots,p_f\right)} = \begin{pmatrix} \frac{\partial Q_1}{\partial q_1} & \cdots & \frac{\partial Q_f}{\partial q_1} & \frac{\partial P_1}{\partial q_1} & \cdots & \frac{\partial P_f}{\partial q_1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial Q_1}{\partial q_f} & \cdots & \frac{\partial Q_f}{\partial q_f} & \frac{\partial P_1}{\partial q_f} & \cdots & \frac{\partial P_f}{\partial q_f} \\ \frac{\partial Q_1}{\partial p_1} & \cdots & \frac{\partial Q_f}{\partial p_1} & \frac{\partial P_1}{\partial p_1} & \cdots & \frac{\partial P_f}{\partial p_1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial Q_1}{\partial p_f} & \cdots & \frac{\partial Q_f}{\partial p_f} & \frac{\partial P_1}{\partial p_f} & \cdots & \frac{\partial P_f}{\partial p_f} \end{pmatrix} ,$$

and J is called Poisson matrix:

$$\mathsf{J} = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} ,$$

which is written in an $f \times f$ blocks form.

The well-known variational principle tells that the solutions of Eqs. (14.10)–(14.11) satisfy the variational equation

$$\begin{aligned} &\delta \int_{t_1}^{t_2} \mathcal{L}(q_1,\ldots,q_f,\dot{q}_1,\ldots,\dot{q}_f,t)\,\mathrm{d}t \\ &\qquad = \delta \int_{t_1}^{t_2} \left[\sum_{k=1}^{f} \dot{q}_k p_k - H(q_1,\ldots,q_f,p_1,\ldots,p_f,t)\right] \mathrm{d}t = 0 \ . \end{aligned} \tag{14.15}$$

An analogous expression is valid for the new phase-space variables, *i.e.*,

$$\delta \int_{t_1}^{t_2} \left[\sum_{k=1}^{f} \dot{Q}_k P_k - K(Q_1,\ldots,Q_f,P_1,\ldots,P_f,t)\right] \mathrm{d}t = 0 \ , \tag{14.16}$$

where we have denoted with K the Hamiltonian with respect to the new set of variables. In order for Eqs. (14.15) and (14.16) to describe the same system, the two integrands must differ at most by the total derivative of a function $F(\boldsymbol{q},\boldsymbol{P},t)$. F is called the generating function of the canonical transformation. Because of this total

derivative we get:

$$\frac{\partial F(\boldsymbol{q},\boldsymbol{P},t)}{\partial q_k} = p_k \ , \tag{14.17}$$

$$\frac{\partial F(\boldsymbol{q},\boldsymbol{P},t)}{\partial P_k} = Q_k \ , \tag{14.18}$$

$$\frac{\partial F(\boldsymbol{q},\boldsymbol{P},t)}{\partial t} + H(\boldsymbol{q},\boldsymbol{p},t) = K(\boldsymbol{Q},\boldsymbol{P},t) \ . \tag{14.19}$$

By properly choosing the canonical transformation, it is possible to obtain a transformation such that $K(Q_1,\ldots,Q_f,P_1,\ldots,P_f)=0$, *i.e.*, the Hamiltonian in the new coordinates is simply zero. In this case, the new coordinates evolve according to

$$\dot{Q}_k = \{Q_k, K\}_{\text{P.B.}} = 0 \ ,$$
$$\dot{P}_k = \{P_k, K\}_{\text{P.B.}} = 0 \ ,$$

i.e., the new variables are constants of motion.

According to Eqn. (14.19), a canonical transformation of this type is generated by an $F(\boldsymbol{q},\boldsymbol{P},t)$ that satisfies

$$\frac{\partial F(\boldsymbol{q},\boldsymbol{P},t)}{\partial t} + H(\boldsymbol{q},\boldsymbol{p},t) = 0 \ . \tag{14.20}$$

If in this equation we rewrite everything as a function of only $(\boldsymbol{q},\boldsymbol{P},t)$, by using Eqn. (14.17), we get*

$$\frac{\partial F(\boldsymbol{q},\boldsymbol{P},t)}{\partial t} + H\left(\boldsymbol{q}, \frac{\partial F(\boldsymbol{q},\boldsymbol{P},t)}{\partial \boldsymbol{q}}, t\right) = 0 \ . \tag{14.21}$$

This is exactly the Hamilton-Jacobi equation: it is a non-linear partial differential equation for the generating function $F(\boldsymbol{q},\boldsymbol{P},t)$. Using Hamilton variational methods (with variations also of the time at the extremes) it is possible to prove that the Hamilton-Jacobi equation is satisfied by the classical action *function*

$$F(\boldsymbol{q},\boldsymbol{P},t) = S_{\text{cl}}(\boldsymbol{q},\boldsymbol{P},t) = \int_{t_1}^{t} \mathcal{L}\left(\boldsymbol{q}_{\text{cl}}(t'), \dot{\boldsymbol{q}}_{\text{cl}}(t'), t'\right) \mathrm{d}t' \ , \tag{14.22}$$

* Here and in the following with the symbol $\partial/\partial\boldsymbol{q}$ (*i.e.*, partial derivative with respect to a vector) we mean the gradient with respect to that vector, *i.e.*, $\partial/\partial\boldsymbol{q} = \boldsymbol{\nabla}_{\boldsymbol{q}}$.

where $\boldsymbol{q}_{\rm cl}(t')$ are the solutions of the equations of motion with boundary conditions

$$\begin{cases} \boldsymbol{q}_{\rm cl}(t_1) &= \boldsymbol{q} \\ \boldsymbol{P}_{\rm cl}(t_1) &= \boldsymbol{P} \end{cases} .$$

Remark. $S_{\rm cl}(\boldsymbol{q}, \boldsymbol{P}, t)$ should *not* be confused with the action functional, it is instead a *function* since we have already plugged in a particular trajectory.

Let us consider for simplicity the case of a point-like particle of mass m moving in a time-independent potential $V(\boldsymbol{x})$. We can write its Lagrangian in cartesian coordinates as

$$\mathcal{L} = \frac{1}{2} m \dot{\boldsymbol{x}}^2 - V(\boldsymbol{x}) \; , \tag{14.23}$$

so we get

$$\boldsymbol{p} = \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}} = m \dot{\boldsymbol{q}} = \frac{\partial S_{\rm cl}}{\partial \boldsymbol{q}} = \boldsymbol{\nabla} S_{\rm cl} \; , \tag{14.24}$$

and therefore the Hamilton-Jacobi equation becomes simply

$$\frac{\partial S_{\rm cl}}{\partial t} + \frac{\left[\boldsymbol{\nabla} S_{\rm cl}\right]^2}{2m} + V = 0 \; . \tag{14.25}$$

14.3 WKB method

14.3.1 General considerations

Let us turn back to Quantum Mechanics and let us consider the one-dimensional *stationary* Schrödinger equation

$$\frac{\mathrm{d}^2 \psi(x)}{\mathrm{d}x^2} + \frac{2m}{\hbar^2} \left[E - V(x)\right] \psi(x) = 0 \; . \tag{14.26}$$

Remember that if $V(x) = \text{const.} = V$ and $E > V$, Eqn. (14.26) has solutions of the form

$$\psi(x) = A \, \mathrm{e}^{\pm i k x} \; , \qquad k = \sqrt{\frac{2m}{\hbar^2}(E - V)} \; .$$

They are oscillating wave functions with constant amplitude A and constant wavelength $\lambda = 2\pi/k$.

Now, let us assume that $V(x)$ varies "slowly" as a function of the position x with respect to λ. Then, it seems resonable to assume that the wave function is still an oscillating one but with slowly-varying position-dependent amplitude and wavelength.

Analogously, if $V(x) = \text{const.} = V$ but with $E < V$, then the solutions of Eqn. (14.26) take the form

$$\psi(x) = A\,\mathrm{e}^{\pm\kappa x}\,, \qquad \kappa = \sqrt{\frac{2m}{\hbar^2}(V-E)}\,,$$

i.e., a pure exponential. Note that the κ above is different than the k of the previous formula. Again, if the potential $V(x)$ varies "slowly" as a function of x with respect to $1/\kappa$, then we may assume that the wave function solution is still of exponential type but with A and k slowly dependent on x.

It should already be clear that our considerations will give a reasonable approximation only within certain regions of the domain of $\psi(x)$. (We will be more precise later on.) In fact, at the points where $E = V$ both λ and $1/\kappa$ diverge to $+\infty$ and it does no longer make sense to assume $V(x)$ to be a slowly varying function of x on the scale of λ in the neighbor of such points. Such points are called *turning* points, since classically those are the points at which the particle changes the direction of its motion. The mathematical treatment of the solution in a neighbor of these points is rather cumbersome and we will invite the interested reader to study it in more advanced books.

14.3.2 WKB solutions and their validity

Let us come back to the stationary Schrödinger equation and write $\psi(x)$ as

$$\psi(x) = A(x)\,\mathrm{e}^{\frac{i}{\hbar}S(x)}\,, \tag{14.27}$$

where A and S are time-independent real-valued functions. Plugging Eqn. (14.27) into the stationary Schrödinger equation and equating real and imaginary parts we

get two coupled equations for A and S which are

$$(S'(x))^2 - 2m[E - V(x)] = \hbar^2 \frac{A''(x)}{A(x)} , \tag{14.28}$$

$$2A'(x)S'(x) + A(x)S''(x) = 0 , \tag{14.29}$$

where the prime denotes the derivative with respect to x, *i.e.*, $S'(x) = \mathrm{d}S(x)/\,\mathrm{d}x$, etc.

From Eqn. (14.29) by integration we get immediately that:

$$A(x) = \frac{\text{const}}{(S')^{\frac{1}{2}}} . \tag{14.30}$$

Substituting back into Eqn. (14.28) we obtain

$$(S')^2 = 2m(E - V) + \hbar^2 \left[\frac{3}{4} \left(\frac{S''}{S'} \right)^2 - \frac{1}{2} \frac{S'''}{S'} \right] . \tag{14.31}$$

This equation is rigorously equivalent to the Schrödinger equation we started from, Eqn. (14.26).

So far, no approximation has been made. The WKB approximation is basically an expansion of S in powers of $\hbar^2$, *i.e.*,

$$S = S_0 + \hbar^2 S_1 + \dots . \tag{14.32}$$

Of course, we shall keep only the lowest order terms in $\hbar^2$. The reason why the expansion Eqn. (14.32) has been done in $\hbar^2$ and not simply in $\hbar$ is that in Eqn. (14.31) $\hbar^2$ (and not $\hbar$ itself) makes its appearance.

At this point we can rewrite Eqn. (14.27) as

$$\psi(x) = A\, \mathrm{e}^{\frac{i}{\hbar}S} = A\, \mathrm{e}^{\frac{i}{\hbar}\left(S_0 + \hbar^2 S_1 + \dots\right)} . \tag{14.33}$$

Inserting the Eqn. (14.32) into (14.31) and looking only at the zero-th order in $\hbar^2$ we get

$$(S_0')^2 = 2m(E - V) . \tag{14.34}$$

Using the following definitions

$$\lambda\!\!\!^{-}(x) = \frac{1}{k(x)} = \frac{\hbar}{\sqrt{2m[E - V(x)]}} , \qquad \text{for } E > V(x) , \tag{14.35}$$

$$\ell(x) = \frac{1}{\kappa(x)} = \frac{\hbar}{\sqrt{2m[V(x) - E]}} , \qquad \text{for } E < V(x) , \tag{14.36}$$

Eqn. (14.34) becomes

$$S_0'(x) = \pm\frac{\hbar}{\lambda\!\!\!^{-}(x)} = \pm\hbar k(x) , \qquad \text{for } E > V(x) , \tag{14.37}$$

$$S_0'(x) = \pm i\frac{\hbar}{\ell(x)} = \pm i\hbar\kappa(x) , \qquad \text{for } E < V(x) . \tag{14.38}$$

So the WKB solutions have the form

$$\psi(x) \approx A\, \mathrm{e}^{\frac{i}{\hbar}S_0} , \tag{14.39}$$

which using Eqn. (14.30) becomes

$$\psi(x) \approx \psi_{\mathrm{WKB}}(x) = \frac{C}{\sqrt{k(x)}}\, \mathrm{e}^{\pm i\int k(x)\,\mathrm{d}x} , \qquad \text{for } E > V(x) , \tag{14.40}$$

$$\psi(x) \approx \psi_{\mathrm{WKB}}(x) = \frac{C}{\sqrt{\kappa(x)}}\, \mathrm{e}^{\pm \int \kappa(x)\,\mathrm{d}x} , \qquad \text{for } E < V(x) . \tag{14.41}$$

The general approximate solutions will be a linear combinations of the two above (with the plus and the minus sign). It is interesting to observe that, in the classically allowed region (*i.e.*, for $E > V$), we have that

$$\rho = |\psi(x)|^2 \approx \frac{|C|^2}{k(x)} \propto \frac{1}{p(x)} , \tag{14.42}$$

This equation shows that the probability to find the particle at the point x is inversely proportional to its linear momentum $p(x)$, *i.e.*, to its velocity, exactly as in Classical Mechanics.

When one uses approximation methods, it is always necessary to specify mathematically the limits of applicability of the method. To state that the WKB works well for slowly varying potentials is not a mathematical precise statement. If we want to

be more precise, let us note from Eqn. (14.33) that the corrections of order $\hbar^2$ brings into the WKB solutions only a further phase factor $\mathrm{e}^{i\hbar S_1}$. This factor can be neglected only if $\hbar S_1 \ll 1$. At this point it is enough to insert the expansion Eqn. (14.32) into Eqn. (14.31) and equate the terms of order $\hbar^2$. For example, for $E > V$ we find after some simple calculations

$$\hbar S_1'(x) = \pm \left[\frac{1}{4} \lambda\!\!\!^{-}''(x) - \frac{1}{8} \frac{\left[\lambda\!\!\!^{-}'(x)\right]^2}{\lambda\!\!\!^{-}(x)} \right] , \tag{14.43}$$

and integrating both sides of this expression we get

$$\hbar S_1 = \pm \left[\frac{1}{4} \lambda\!\!\!^{-}'(x) - \frac{1}{8} \int \frac{\left[\lambda\!\!\!^{-}'(x)\right]^2}{\lambda\!\!\!^{-}(x)} \, \mathrm{d}x \right] . \tag{14.44}$$

The criterion for the WKB approximation, *i.e.*, $\hbar S_1 \ll 1$, becomes then:

$$\lambda\!\!\!^{-}'(x) \ll 1 , \qquad \text{for } E > V(x) .$$

In the same way it is possible to show that the criterion is

$$\ell'(x) \ll 1 , \qquad \text{for } E < V(x) .$$

If we make explicit the dependence from the potential we get the validity condition

$$\frac{|m\hbar V'(x)|}{|2m\left[E - V(x)\right]|^{\frac{3}{2}}} \ll 1 . \tag{14.45}$$

Eqn. (14.45) is a more precise characterization of the statement that the potential must be slowly varying with x.

14.3.3 Connection formulae

At the turning points where $E = V(x)$ the quantities previously defined, namely, $\lambda\!\!\!^{-}(x)$ and $\ell(x)$, diverge and the approximations used before are not valid anymore.

Let us suppose that $E \gtrless V(x)$ for $x \gtrless a$. If the WKB approximation can be used everywhere except in a neighbor of the turning point $x = a$, then

$$\psi_{\rm WKB}(x) = \frac{A}{\sqrt{\kappa(x)}}\, \mathrm{e}^{-\int_x^a \kappa(x')\,\mathrm{d}x'} + \frac{B}{\sqrt{\kappa(x)}}\, \mathrm{e}^{+\int_x^a \kappa(x')\,\mathrm{d}x'}\ , \quad \text{for } x > a\ , \tag{14.46}$$

$$\psi_{\rm WKB}(x) = \frac{C}{\sqrt{k(x)}}\, \mathrm{e}^{-i\int_x^a k(x')\,\mathrm{d}x'} + \frac{D}{\sqrt{k(x)}}\, \mathrm{e}^{+i\int_a^a k(x')\,\mathrm{d}x'}\ , \quad \text{for } x < a\ . \tag{14.47}$$

The problem now is to find out how the coefficients C and D can be related to A and B. If we manage to do that via some abstract reasoning, we can bypass the problem of finding the exact WKB solution in the neighbor of the turning points where the WKB solution fails. We will only report here the "connection formulae" which connects the "oscillating WKB solutions" to the "exponentialy decaying WKB solutions" and let the interested reader study their derivation in more advanced books:

$$\begin{aligned}\frac{A}{\sqrt{\kappa(x)}}\, \mathrm{e}^{-\int_x^a \kappa(x')\,\mathrm{d}x'} + \frac{B}{\sqrt{\kappa(x)}}\, \mathrm{e}^{+\int_x^a \kappa(x')\,\mathrm{d}x'} \leftrightarrow \\ \leftrightarrow \frac{2A}{\sqrt{k(x)}} \cos\left[\int_a^x k(x')\,\mathrm{d}x' - \frac{\pi}{4}\right] - \frac{B}{\sqrt{k(x)}} \sin\left[\int_a^x k(x')\,\mathrm{d}x' - \frac{\pi}{4}\right]\ .\end{aligned} \tag{14.48}$$

Clearly, the oscillating solutions are valid for $x \ll a$ while the exponential solutions are valid for $x \gg a$. Analogous formulae hold when $V(x) \lesseqgtr E$ for $x \lesseqgtr a$.

14.4 α decay

The WKB belongs to the class of so called *non-perturbative* approximation methods. As we have already seen, the basic idea is to consider $\hbar$ "small", while the perturbative coupling of the potential can be any value, even not small. Other non-perturbative methods are the variational method and the adiabatic method, which we shall not discuss here. Since non-perturbative methods do not require the coupling constant to be small, they are the obvious choice when the interactions are "strong", as it happens for instance in nuclear physics.

In this section, we shall see an application of the WKB method in that sector of physics. In particular we shall study the α decay of the nuclei. A nucleus of radius

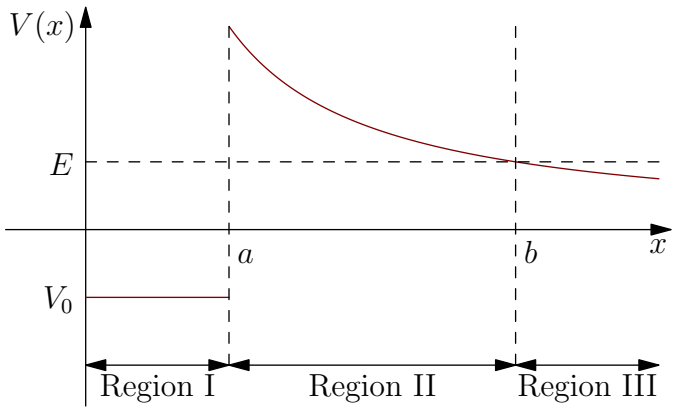


FIG. 14.1. Model of potential barrier in the α decay of the nuclei.

a and charge Z can be roughly represented by a potential well of depth V_0. (See Fig. 14.1.) For $x > a$, the potential is of Coulomb type and goes to zero for $x \to \infty$. Let us suppose that inside the nucleus there is a small bunch of matter, *i.e.*, α particles with charge z, and energy $E > 0$. Let us indicate with $x = b$ the point at which the Coulomb potential is equal to the energy, *i.e.*,

$$\frac{zZe^2}{4\pi\varepsilon_0 b} = E \, .$$

Our goal is to evaluate the probability amplitude that the α particle escapes the nucleus. We have to build a solution of the Schrödinger equation which in the region III behaves as a wave (transmitted wave). In that region, using the WKB approximation the solution has the form

$$\psi_{\rm III}(x) \sim \frac{C}{\sqrt{k(x)}} \exp\left[\pm i \int k(x)\, {\rm d}x\right] \, . \tag{14.49}$$

To simplify the calculations, let us choose a wave which propagates in the direction of positive x. Let us insert a phase of $\pi/4$ (ψ's are defined up to an overall phase)

and C is given by choosing the integration range, so we obtain:

$$\begin{aligned}\psi_{\rm III}(x) &= \frac{1}{\sqrt{k(x)}}\exp\left[i\int_b^x k(x')\,{\rm d}x' + i\frac{\pi}{4}\right] \\ &= \frac{1}{\sqrt{k(x)}}\left[\cos\left(\int_b^x k(x')\,{\rm d}x' + \frac{\pi}{4}\right) + i\sin\left(\int_b^x k(x')\,{\rm d}x' + \frac{\pi}{4}\right)\right] \ . \end{aligned} \quad (14.50)$$

Now, we make use of the connection formulae in order to determine ψ in the region II. We extend only the sin function, since the cos function would give a exponentially increasing function which would be not normalizable. We have

$$\begin{aligned}\psi_{\rm II}(x) &= \frac{-i}{\sqrt{\kappa(x)}}\exp\left[\int_x^b \kappa(x')\,{\rm d}x'\right] \\ &= \frac{-i}{\sqrt{\kappa(x)}}\exp\left[\int_a^b \kappa(x')\,{\rm d}x'\right]\exp\left[-\int_a^x \kappa(x')\,{\rm d}x'\right] \end{aligned} \quad (14.51)$$

Defining

$$\gamma = \int_a^b \kappa(x)\,{\rm d}x = \int_a^b \sqrt{\frac{2m}{\hbar^2}\left[V(x)-E\right]}\,{\rm d}x \ , \quad (14.52)$$

Eqn. (14.51) becomes

$$\psi_{\rm II}(x) = -\frac{i}{\sqrt{\kappa(x)}}\,{\rm e}^{\gamma}\exp\left[-\int_a^x \kappa(x')\,{\rm d}x'\right] \ . \quad (14.53)$$

In region I, the solution of the Schrödinger equation is always an oscillating function. To make the calculations easier, and indicating as usual $k_0 = \sqrt{2m\,(E-V_0)}/\hbar$, we can write

$$\psi_{\rm I}(x) = A\sin\left[k_0(x-a)+\varphi\right] = \frac{A}{2i}\left\{{\rm e}^{i[k_0(x-a)+\varphi]} - {\rm e}^{-i[k_0(x-a)+\varphi]}\right\} \ , \quad (14.54)$$

where the constants A and φ are determined by imposing the continuity of the wave function and of its first-order derivative in $x = a$. In particular, we find that

$$\left\{\begin{array}{rcl} k_0\cot\varphi &=& -\kappa(a) \\ A\sin\varphi &=& -\frac{i}{\sqrt{\kappa(a)}}\,{\rm e}^{\gamma} \end{array}\right. \ .$$

As we want to study the tunnel effect through the potential barrier, we need to calculate the transmission coefficient T, given by the ratio between the transmitted

probability current $J_{\rm III}^{\rm tr}$ which reaches region III and the incident current $J_{\rm I}^{\rm inc}$ coming from region I, *i.e.*,

$$T = \frac{J_{\rm III}^{\rm tr}}{J_{\rm I}^{\rm inc}} \, .$$

In order to calculate $J_{\rm I}^{\rm inc}$ let is note that $\psi_{\rm I}$ is the sum of an incident and a reflected wave. Let us remember that in one dimension the probability current is

$$J = \Re \left[\psi^\dagger(x) \frac{\hat{p}}{m} \psi(x) \right] = \Re \left[\psi^\dagger(x) \frac{\hbar}{im} \frac{\rm d}{{\rm d}x} \psi(x) \right] \, .$$

It is easy to prove that

$$J_{\rm I}^{\rm inc} = \frac{1}{4} \left| A \right|^2 \frac{\hbar k_0}{m} \qquad \text{and} \qquad J_{\rm III}^{\rm tr} = \frac{\hbar}{m} \, .$$

Thus

$$T = 4 \frac{\sqrt{(V_a - E)(E - V_0)}}{V_a - V_0} \, {\rm e}^{-2\gamma} \, ,$$

which means that

$$T \propto \exp \left[-\frac{2}{\hbar} \int_a^b \sqrt{2m \left[V(x) - E \right]} \, {\rm d}x \right] \, .$$

This one is the famous *Gamow* formula.

The reader may ask why we have used the WKB approximation. The reason is that we could not solve exactly the Schrödinger equation associated with the potential drawn in the previous figure. At the same time there was no small perturbative parameter associated with the potential which would allow us to use perturbative methods. Moreover, we had potentials which (especially the Coulomb potential felt outside the nucleus) were changing slowly on the scale of the wavelength involved, so it was the ideal ground to implement the WKB method.

Chapter 15

INTRODUCTION TO SCATTERING

15.1 introduction

The material for this chapter will be provided by the teacher via handwritten notes.

Chapter 16

THE SCHRÖDINGER FIELD AND SECOND QUANTIZATION

This chapter is taken from the Bachelor Thesis of Serafin Francesco, University of Trieste.

One of the main difference between Classical Mechanics (CM) and Quantum Mechanics (QM) is in the way a physical system is described. In QM we have the well-known postulate:

Postulate: In QM the state of a system is described by an element of a Hilbert Space: $\mathbb{H}$.

When the elements of $\mathbb{H}$ (indicated with $|\,\rangle$ in the Dirac notation) are *represented* on a basis they can take different forms: for example in the representation on the eigenstates of the position operator ($\hat{\mathbf{r}}$) , they are given by square integrable functions mapping points of $\mathbb{R}^n$ into $\mathbb{C}$ (or even $\mathbb{C}^m$):

$$\mathbb{H}^* \times \mathbb{H} \ni \langle \mathbf{r} \,|\, \psi(t) \rangle \equiv \psi(\mathbf{r}, t) \quad : \quad \mathbb{R}^n \times \mathbb{R} \to \mathbb{C}^m.$$

In what follows we will use for simplicity the position representation and so, besides the states, also the operators ($\hat{\mathbf{p}}, \hat{\mathbf{r}}, \hat{V}, ecc.$) will be given in that basis.

We know that the state of a quantum system in a potential $V(\mathbf{r})$ evolves in time according to the Schrödinger equation:

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right] \psi(\mathbf{r}, t) = i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) \tag{16.1}$$

Notice that $\psi(\mathbf{r},t)$, being a function, can be interpreted as a field , and we can ask ourselves if the evolution equation (16.1) can be derived via a variational principle from some appropriate field-Lagrangian. As $\psi(t,x)$ is a complex field, it is equivalent to two real and independent fields: $\Re e\psi$ and $\Im m\psi$. Equivalently we can describe the system using the field and its complex coniugate:

$$\begin{cases} \psi = \Re e\psi + i\Im m\psi \\ \psi^* = \Re e\psi - i\Im m\psi. \end{cases}$$

We shall assume that ψ and ψ^* go to zero at infinity. It is easy to see that in the *free* case the field Lagrangian, density $\mathcal{L} = \mathcal{L}(\psi,\psi^*)$, which reproduces the Schrödinger equation for the fields ψ and ψ^* is:

$$\mathcal{L} = \psi^*\left(i\hbar\frac{\partial}{\partial t} + \frac{\hbar^2}{2m}\nabla^2\right)\psi \quad . \tag{16.2}$$

To get the Schrödinger equation for a particle in a a potential V , one just need to add to the previous Lagrangian the term: $V\psi\psi^* = \psi^* V\psi$.

Let us now check that the Euler−Lagrange field equation derived from the Lagrangian above is the (16.1). From the literature * we know that the Euler−Lagrange equations for the fields η_ρ (with "ρ" labelling the set of fields if they are more than one) involve also its space-time derivative indicated as $\eta_{\rho,\nu}$ (with "ν" labelling the space-time index) and these equations have the form:

$$\frac{d}{dx^\nu}\frac{\partial\mathcal{L}}{\partial\eta_{\rho,\nu}} - \frac{\partial\mathcal{L}}{\partial\eta_\rho} = 0 \quad . \tag{16.3}$$

The Schrödinger fields are just two (ψ and ψ^*), and so we have $\rho = 1,2$. Doing the variation with respect to the field associated to $\rho = 2$, i.e ψ^*, for the free case we get :

$$\frac{d}{dx^\nu}\frac{\partial\mathcal{L}}{\partial\psi^*_{,\nu}} - \frac{\partial\mathcal{L}}{\partial\psi^*} = 0 - \left(i\hbar\frac{\partial}{\partial t} + \frac{\hbar^2}{2m}\nabla^2\right)\psi = 0 \tag{16.4}$$

* The students who do not know this should study the very first chapters on fields equations contained in many books like, for example, Goldstein or Landau's book.

In this manner we have eproduced (16.1).

To build the Hamiltonian density $\mathcal{H}$ associated to $\mathcal{L}$, we have to build the momenta associated to ψ and ψ^*. A standard convention is to indicate with π^* the momentum coniugated to ψ and with π the momentum coniugated to ψ^*:

$$\pi = \frac{\partial \mathcal{L}}{\partial \psi^*_{,\nu}} = 0 \qquad\qquad \pi^* = \frac{\partial \mathcal{L}}{\partial \psi_{,\nu}} = i\hbar\psi^* \quad .$$

Using all this the Hamiltonian density turns out to be :

$$\mathcal{H} \equiv (\pi^*, \pi) \cdot \frac{\partial}{\partial t}\begin{pmatrix}\psi \\ \psi^*\end{pmatrix} - \mathcal{L} = i\hbar\psi^*\frac{\partial\psi}{\partial t} - i\hbar\psi^*\frac{\partial\psi}{\partial t} - \psi^*\frac{\hbar^2}{2m}\nabla^2\psi = \psi^*\left(-\frac{\hbar^2}{2m}\nabla^2\right)\psi \quad . \tag{16.5}$$

If we integrate the expression above, the *free* Hamiltonian is :

$$H = \int_{\mathcal{Q}} \mathcal{H}\, d^3r = \int_{\mathcal{Q}} \psi^*\left(-\frac{\hbar^2}{2m}\nabla^2\right)\psi\, d^3r \quad . \tag{16.6}$$

Note that this expression coincides with the definition, in standard QM, of the expectation value of the kinetic operator $\hat{T}$ on the state ψ, which in the bra and ket notation of Dirac can be written as:

$$\langle \hat{T} \rangle = \langle \psi |\, \hat{T}\, | \psi \rangle \,. \tag{16.7}$$

Inserting the decomposition of the identity $\mathbb{1}$ in position space on the left and on the right of the operator $\hat{T}$, we get the position picture representation of this expectation value:

$$\langle \hat{T} \rangle = \int_{\mathcal{Q}}\int_{\mathcal{Q}'} \langle \psi \,|\, x' \rangle \, \langle x' |\, \hat{T}\, | x \rangle \, \langle x \,|\, \psi \rangle \; d^3x\, d^3x' = \int_{\mathcal{Q}} \psi^*(x)\left(-\frac{\hbar^2}{2m}\nabla^2\right)\psi(x)\, d^3x. \tag{16.8}$$

This is exactly the expression (16.6).

In the most general case in which the potential is not zero, the lagrangian (16.2) becomes:

$$\mathcal{L} = \psi^*\left(i\hbar\frac{\partial}{\partial t} - H\right)\psi \quad . \tag{16.9}$$

Now let us concentrate on the Schrödinger equation for a *free* system and on the interpretation of the wave function as a field. Fields are now, like in the electromagnetic case, our degrees of freedom but up to now we have not quantized these fields $\psi(x,t)$, but only the operators $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$. Because of this, we can say that (16.2) is the equation of motion of a *classical* field. To work in analogy with the electromagnetic field, we can say that the Schrödinger equation corresponds to the Maxwell equations and the $\psi(x,t)$ corresponds to the electromagnetic field.

Let us take the Schrödinger equation for a free system and let us indicate with ϕ_n^+ the eigenfunctions associated to the positive eigenvalues E_n^0:

$$-\frac{\hbar^2}{2m}\nabla^2\phi_n^+ = E_n^0\phi_n^+. \tag{16.10}$$

In this way we get the numerable set $\{\phi_n^+\}$ which is a basis of the Hilbert space where the eq. (16.10) is defined. A generic state $\psi \in \mathbb{H}$ can be written as a linear combinations of the states above because they make up a basis. The states ϕ_n can be represented on the position eigenstates. Moreover imposing periodic boundary conditions and supposing we work in a finite space of volume V, they become:

$$\phi_n^+(\mathbf{r},t) = \left(\frac{1}{V}\right)^{\frac{1}{2}} e^{i(\mathbf{p}_n\cdot\mathbf{r}-E_n^0 t)} \tag{16.11}$$

where $\mathbf{p}_n = 2\pi(V)^{-1/3}(n_x,n_y,n_z)$ and $E_n^0 = p_n^2/(2m)$. Remember that the ortogonality relation $\int_V \phi_i^*\phi_j\, d\mathbf{r} = \delta_{i,j}$ holds. So a generic state ψ can be written as:

$$\psi(\mathbf{r},t) = \sum_n a_n\phi_n^+(\mathbf{r},t) \tag{16.12}$$

where the expansion coefficients a_n are complex numbers. Using the expansion above for the ψ and inserting it into the Hamiltonian (16.6), we get:

$$\begin{aligned} H &= \int_{\mathcal{Q}\subset\mathbb{R}^3}\left(\sum_n a_n^*\phi_n^*\right)\left[-\frac{\hbar^2}{2m}\nabla^2\right]\left(\sum_k a_k\phi_k\right)d^3x = \\ &\int_{\mathcal{Q}}\left(\sum_n a_n^*\phi_n^*\right)\left(\sum_k a_k\left[-\frac{\hbar^2}{2m}\nabla^2\right]\phi_k\right)d^3x = \int_{\mathcal{Q}}\left(\sum_n a_n^*\phi_n^*\right)\left(\sum_k a_k E_k^0\phi_k\right)d^3x = \\ &= \sum_n\sum_k E_k^0 a_n^* a_k\int_{\mathcal{Q}}\phi_n^*\phi_k\, d^3x = \sum_n\sum_k E_k^0 a_n^* a_k\delta_{n,k} = \sum_n E_n^0 a_n^* a_n \end{aligned} \tag{16.13}$$

We can interpret this final relation by saying that the energy of the generic *field* ψ, is a combination of the energies of the single eigenstates of $\hat{H}$, and each of them is taken with a weight related to the coefficients a_n of the projections of ψ on the eigenstates ϕ_n.

Next thing we want to do is quantize the *field* that describes the system, that means we must promote the field $\psi(x)$ to be an operator. This procedure is called *second quantization*. We can prove that what we obtain is equivalent to the *first quantization* of a system described by our Hamiltonian but containing an arbitrary number of particles, this is a so called "*many body*" system.

The rest of the chapter is organized as follows: In section 1.1 we attack the "*many body*" problem by building the wave function of the system $\Psi(\mathbf{r}_1, ..., \mathbf{r}_N, t)$ and the associated evolution equation. Next we perform a change of basis in order to write the over-all Hamiltonian as a function of some creation and annihilation operators: $\hat{a}^\dagger$ e $\hat{a}$.

In 1.2 we show that the Hamiltonian derived in section 1.1 is formally equivalent to the one built using the second quantization procedure. As we said earlier this procedure turn the field ψ into an operator ($\psi \to \hat{\psi}$), and turn the Hamiltonian into the following operator: $\hat{\hat{H}} \equiv \int \hat{\psi}^\dagger(x) H(x) \hat{\psi}(x)\, dx$.

The second quantization is widely used as the only correct technique to treat quantum *relativistic* fields. This is a sector of physics where we always have to handle from the beginning a many body system. In fact for example a photon, with its natural creation of the many particle-antiparticle pairs from the vacuum along its motion , is automatically a many body system.

16.1 The "many body" *problem.*

Let us consider a system of N *identical* particles interacting via a two body potential $V(\mathbf{r}_i, \mathbf{r}_j)$. In first quantization the associated Hamiltonian has the following form:

$$\hat{H}_{(1^{st})} = \sum_{i=1}^{N} \hat{T}_i + \frac{1}{2} \sum_{i \neq j=1}^{N} \hat{V}(\mathbf{r}_i, \mathbf{r}_j) \quad . \tag{16.14}$$

The wave function, in position representation, which describes the system and its associated Schrödinger equation are:

$$\Psi(\mathbf{r}_1, ..., \mathbf{r}_N, t) \equiv \Psi(\{\mathbf{r}\}_N, t) \tag{16.15}$$

$$i\hbar \frac{\partial}{\partial t} \Psi(\{\mathbf{r}\}_N, t) = H_{(1^{st})} \Psi(\{\mathbf{r}\}_N, t) \quad . \tag{16.16}$$

Ψ can be build from the i*single* particle, time *independent* states, which we shall indicate with $\varphi_{(E)}(\mathbf{r}_i)$ for the $i-$particle. We choose these states which obey the stationary Shrödinger eq.: $H\varphi_{(E)}(\mathbf{r}_i) = i\hbar\partial_t \varphi_{(E)}(\mathbf{r}_i)$ because they make up a basis and have a clear physical meaning. As the particles are identical ones they all have the same spectrum. Taking all of this into account we can then expand the many-body wave function as:

$$\Psi(\{\mathbf{r}\}_N, t) = \sum_{E_i} C(E_1, ..., E_N, t) \varphi_{E_1}(\mathbf{r}_1) \cdot .... \cdot \varphi_{E_N}(\mathbf{r}_N) \quad , \tag{16.17}$$

where the sum over E_i indicates the sum over all the energies of all single particles. We have confined the time-dependence into the coefficients C.

Let us insert the expansion (16.17) in (16.16), and multiply the equation for the complex coniugate of a fixed state: $\varphi_{E_1^*}(\mathbf{r}_1) \cdot .... \cdot \varphi_{E_N^*}(\mathbf{r}_N)$. Next let us integrate what we get over $\{\mathbf{r}\}_N$. In this way we get an equation for the coefficients C associated to a particular choice of the energies $E_1^*, ..., E_N^*$. We will call this equation as "*equation for the coefficients*". Let us note that for a particular choice of the set of energies, many different products of states $\varphi_{E_1^*}(\mathbf{r}_1) \cdot .... \cdot \varphi_{E_N^*}(\mathbf{r}_N)$ may corresponds to that set of energies, i.e. there may be degeneracies on that set of energies. To solve this problem

let us define the so called *occupation number* n_i^* as the number of single-particle states which happens to have energy E_i^*.

At this point let us notice that it is possible to order, inside the coefficients, the energies which are equal to each other without having to change signs (this is possible at least for *bosoni*) and pass from *a sum over the energies to a sum over the occupation numbers*:

$$\boxed{\sum_{E_i} \rightarrow \sum_{(n_1^*,....,n_x^*)}} \tag{16.18}$$

where we have indicated with n_x $(x \neq N)$, the last occupation number which is different from zero. We can also imagine to do a sum over an infinitly long strings of occupation numbers $(n_1^*, ...; n_\infty^*)$, where all numbers are zero after the l'$x-$position. This trick allows us to remove the constraint of having a *finite* number of particles and so we can treat arbitrary large systems. Let us now define some new coefficients which are related to the C of (16.17) as :

$$\bar{C}(n_1^*, ..., n_x^*, t) \equiv C(E_1^*, ..., E_N^*, t) = C(\underbrace{E_1^*}_{n_1^*}, ..., \underbrace{E_x^*}_{n_x^*}, ..., E_N^*, t). \tag{16.19}$$

We can re-write the $\bar{C}$ as :

$$\bar{C}(n_1^*, ..., n_x^*, t) \equiv \left(\frac{N!}{n_1^*! \cdot .... \cdot n_x^*!} \right)^{\frac{1}{2}} f(n_1^*, ..., n_x^*, t) \quad . \tag{16.20}$$

In the last step we have introduced some abstract functions f related to the $C(E_1^*, ..., E_N^*, t)$. In this manner we can re-write the total wave function Ψ in a new basis as:

$$\Psi(\{\mathbf{r}\}_N, t) = \sum_{(n_1^*,....,n_x^*)} f(n_1^*, ..., n_x^*, t) \cdot \Phi_{(n_1^*,....,n_x^*)}(\{\mathbf{r}\}_N) \tag{16.21}$$

where

$$\Phi_{(n_1^*,....,n_x^*)}(\{\mathbf{r}\}_N) \equiv \sum_{E_i \Longleftrightarrow (n_1^*,....,n_x^*)} \left(\frac{n_1^*! \cdot .... \cdot n_x^*!}{N!} \right)^{\frac{1}{2}} \varphi_{E_1}(\mathbf{r}_1) \cdot .... \cdot \varphi_{E_N}(\mathbf{r}_N) \quad . \tag{16.22}$$

The sum is done over all energies which are compatible with a given string of occupation numbers $(n_1^*, ..., n_x^*)$. It is for this reason that we put on Φ a label given by a

string of occupation numbers. The states Φ are in fact in a one to one correspondence with the states $|n_1^*, ..., n_x^*\rangle$ which belong to what is called a Fock space which will be defined later on.

Let us remember that the C used in (16.17) have to obey what we called the "equation for the coefficients', and as a consequence also $f(n_1^*, ..., n_x^*, t)$ of equation (16.21) must obey a similar equation modulo some factors in front. In order to be able to use the dependence of f on the occupation numbers, we need to have only sum over the energies. The sum over the numbers of particles in the Hamiltonian (16.14) can be turned into a different sum as:

$$\boxed{\sum_{i=1}^{N} \to \sum_{E_i} n_i \equiv \sum_i n_i} \tag{16.23}$$

This replacement is very easy for the kinetic term $\hat{T}$ while in the case of $\hat{V}$ we have to impose the constraint $i \neq j$ which is present when we do the sum over the particles number and we have to find an analog one when we sum over the energies. This new constraint is derived via the following reasoning : if $E_i^* \neq E_j^*$, then the condition is automatically satisfied and the occupation are simply n_i^* ed n_j^*. If instead the energies are the same for two different particles, i.e. $E_i^* = E_j^*$, in that case we have to count only once the particles labelled by different indices but that occupy that same energy level. This means that we have to group together in couples, without repetition, the number of particles at energy E_i^* i.e, :

$$\binom{n_i^*}{2} = n_i^*(n_i^* - 1) \quad . \tag{16.24}$$

To unify the two cases $E_i^* = E_j^*$ e $E_i^* \neq E_j^*$, we can write:

$$\sum_{i \neq j}^{N} \longrightarrow \sum_{E_i} \sum_{E_j} n_i^*(n_j^* - \delta_j^i) \equiv \sum_i \sum_j n_i^*(n_j^* - \delta_j^i) \quad , \tag{16.25}$$

where the sums are over all admissible energies. The final equation for the coefficients

f is:

$$i\hbar\frac{\partial}{\partial t}f(n_1^*,....,n_x^*,t)=\sum_i\langle i|\hat{T}|i\rangle\, n_i^* f(n_i^*)+\sum_{i\neq j}\langle i|\hat{T}|j\rangle\,(n_i^*)^{\frac{1}{2}}(n_j^*+1)^{\frac{1}{2}}f(n_i^*-1,n_j^*+1)$$
$$+\frac{1}{2}\sum_{i=j}\sum_k\sum_l\langle ii|\hat{V}|kl\rangle\,(n_i^*)^{\frac{1}{2}}(n_i^*-1)^{\frac{1}{2}}(n_k^*+1)^{\frac{1}{2}}(n_l^*+1)^{\frac{1}{2}}f(n_i^*-2,n_k^*+1,n_l^*+1)$$
$$+\frac{1}{2}\sum_{i\neq j}\sum_k\sum_l\langle ij|\hat{V}|kl\rangle\,(n_i^*)^{\frac{1}{2}}(n_j^*)^{\frac{1}{2}}(n_k^*+1)^{\frac{1}{2}}(n_l^*+1)^{\frac{1}{2}}f(n_i^*-1,n_j^*-1,n_k^*+1,n_l^*+1)$$
$$+\text{remaining terms for which } k\neq l\neq(i\text{ o }j). \quad (16.26)$$

The compact notation used for $f(n_i^*+1)$, indicates that all n_j^* , with j from 1 to x, are the same as in the $f(n_1^*,....,n_x^*,t)$ which appears on the L.H.S of eq. (16.26) except the l'$i-$term which is turned into the (n_i^*+1).

The crucial point embodied in (16.22) is that we have passed from the basis of the energy eigenstates, (i.e. a sum over the energies), to an occupation number basis (i.e. a sum over $(n_1^*,....,n_x^*)$). It is possible to show that actually these are the position picture representation of the abstract states which make up what is called the *Fock space*. So we can write the total wave function Ψ as a linear combinations of states of the Fock space as :

$$|\Psi(t)\rangle=\sum_{(n_1,....,n_x)}f(n_1,....,n_x,t)\,|n_1,....,n_x\rangle \quad (16.27)$$

In order to derive (16.27) from (16.21) we just have to write $\varphi_{E_i}(\mathbf{r}_i)=\langle\mathbf{r}_i|E_i\rangle$ and check that the coefficients of the change of basis $|E_1,....,E_N\rangle\rightarrow|n_1,....,n_x\rangle$ are the same as in (16.22).

Let us now see how we can write the Hamiltonian in this new basis. Let us apply $i\hbar\frac{\partial}{\partial t}$ to both sides of (16.27), and note that on the R.H.S the only dependence on time is in the factors f, and so we can use the equation (16.26). Doing this and working out the calculations in details, we can notice that it is possible to reconstruct on the R.H.S the abstract state $|\Psi(t)\rangle$ with the new Hamiltonian applied on it. This one is basically all the R.H.S of (16.26) where each coefficient f is multiplied by the same

ket $|n_1, ..., n_x\rangle$. Next we operate a change of variables $n_i - 1 \equiv n'_i$ etc. and we get coefficients $f(n_1, ..., n_x)$ which are all equal and we get kets whose occupation numbers corresponds exactly to the factors of the type $(n'_i)^{\frac{1}{2}}$ by which they are multiplied

$$\text{example:} \qquad f(n'_1, ..., n'_x) \cdot (n'_i + 1)^{\frac{1}{2}} \, |n'_1, ..., n'_{i-1}, (n'_i + 1), n'_{i+1}....\rangle \tag{16.28}$$

Let us remember that in the Fock space we can define creation and annihilation operators $\hat{a}^\dagger$ e $\hat{a}$, which raises or lower of one unit the occupation numbers of the particle state on which they act. Their properties are:

$$\hat{a}_i^\dagger \, |...n_i...\rangle = (n_i + 1)^{\frac{1}{2}} \, |...n_i + 1...\rangle$$
$$\hat{a}_i \, |...n_i...\rangle = (n_i)^{\frac{1}{2}} \, |...n_i - 1...\rangle$$

and they derive from the well-known commuation relations.

As the new Hamiltonian contains terms of the form (16.28), we can rewrite it as a function of the operators $\hat{a}^\dagger$ e $\hat{a}$, and obtain the final form of the Hamiltonian for the many body problem in the Fock space. It turns out to be:

$$\boxed{\hat{H}_{(\text{many body})} \equiv \sum_{i,j} \hat{a}_i^\dagger \, \langle i|\hat{T}|j\rangle \, \hat{a}_j + \sum_{i,j,k,l} \hat{a}_i^\dagger \hat{a}_j^\dagger \, \langle ij|\hat{V}|kl\rangle \, \hat{a}_k \hat{a}_l \quad .} \tag{16.29}$$

Note that the matrix elements $\langle i|\hat{T}|j\rangle$ and $\langle ij|\hat{V}|kl\rangle$ are complex numbers and the only operators are $\hat{a}^\dagger$ and $\hat{a}$.

The Schrödinger equation for the *many body* problem, re-written in the Fock space, is equivalent to the eq. eq:Schrodmany where we started from, i.e:

$$\boxed{i\hbar \frac{\partial}{\partial t} \, |\Psi(t)\rangle_E = \hat{H}_{(1^{st})} \, |\Psi(t)\rangle_E} \iff \boxed{i\hbar \frac{\partial}{\partial t} \, |\Psi(t)\rangle_n = \hat{H}_{(\text{many body})} \, |\Psi(t)\rangle_n} \tag{16.30}$$

where with the symbols $|\Psi(t)\rangle_E$ and $|\Psi(t)\rangle_n$ we had indicated the ket $|\Psi(t)\rangle$ in the two different basis, respectively in the energy eigenstates and in the Fock space one. In this last space the only information we need regarding the state $|\Psi(t)\rangle_n$ is its ground state necause all the other states are obtained by applying the $\hat{a}^\dagger$ and $\hat{a}$ operators.

16.2 Second Quantization

Second Quantization is basically a quick recipe to build an Hamiltonian equivalent to the $\hat{H}_{(\text{many body})}$ presented in the previous section. For semplicity we will use the position representation. The prescriptions of that recipe are:

- Build the *field operator* which is defined at every point in space as:

$$\hat{\psi}(\{\mathbf{r}\}_N) \equiv \sum_i \varphi_i(\mathbf{r}_i)\hat{a}_i \tag{16.31}$$

  where the $\varphi_i(\mathbf{r}_i)$ are the single particle eigenstate of energy E_i (see the previous sections).

- Build the Hamiltonian $\hat{\hat{H}}$ of second quantization as:

$$\hat{\hat{H}} \equiv \sum_{i,j} \langle i|\hat{H}_{(1^{st})}|j\rangle\, \hat{a}_i^\dagger \hat{a}_j = \int \hat{\psi}^\dagger(\{\mathbf{r}\}_N) H_{(1^{st})}(\{\mathbf{r}\}_N)\hat{\psi}(\{\mathbf{r}\}_N)\, d\{\mathbf{r}\}_N \tag{16.32}$$

  Remember that $H_{(1^{st})}(\{\mathbf{r}\}_N)$ is a complex number.

In general, the second quantization rule modifies any operator not only the Hamiltonian and it does that using the analog of the (16.32).

The physical meaning of (16.32) is the following: let us suppose we take as $|i\rangle$, $|j\rangle$ two *eigenstate* of $\hat{H}_{(1^{st})}$. Doing $\langle i|\hat{H}_{(1^{st})}|j\rangle$ we get something proportional to δ_i^j and so only a sum over one index is left over:

$$\hat{\hat{H}} \equiv \sum_i \langle i|\hat{H}_{1^{st}}|j\rangle\, \hat{a}_i^\dagger \hat{a}_i = \sum_i E_i(\mathbf{r})\hat{n}_i. \tag{16.33}$$

So the operator $\hat{\hat{H}}$ somehow counts how many particles are in the $i-$state in the point $\{\mathbf{r}\}_N$ and it associates to it the corresponding energy E_i. As a consequence this Hamiltonian represents the total energy of a system with an arbitrary number of particles.

Via this Hamiltonian $\hat{\hat{H}}$ the evolution of the abstract state $|\Psi(t)\rangle_n \equiv \sum_{(n_1,...,n_x)} f(n_1,...,n_x,t)\,|n_1,...,n_x\rangle$ is given again by the Schroedinger equation:

$$\boxed{i\hbar\frac{\partial}{\partial t}\,|\Psi(t)\rangle_n = \hat{\hat{H}}\,|\Psi(t)\rangle_n = \int \hat{\psi}^\dagger(\{\mathbf{r}\}_N) H_{1^{st}}(\{\mathbf{r}\}_N)\hat{\psi}(\{\mathbf{r}\}_N)\,d\{\mathbf{r}\}_N\,|\Psi(t)\rangle_n} \quad . \tag{16.34}$$

Let us now prove the equivalence between the two formulations: the many body and the second quantization one. This proof will throw light also on the definition (16.31). In eq. (16.29) let us insert some completness relation in position space between the first quantization operators $\hat{T}$ and $\hat{V}$ in order to represent everything in position space and let us use the definition $\langle \mathbf{r}_i|i\rangle = \varphi_i(\mathbf{r}_i)$. Moreover let us use the fact that $\langle \mathbf{r}|\hat{T}|\mathbf{r}'\rangle = T(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}')$. By doing so, we get:

$$\begin{aligned}
\hat{H}_{(\text{many body})} &= \int d\mathbf{r}\left[\sum_i \varphi_i(\mathbf{r}_i)\hat{a}_i^\dagger\right] T(\mathbf{r}) \left[\sum_i \varphi_j(\mathbf{r}_j)\hat{a}_j\right] \\
&+ \int \left[\sum_i \varphi_i(\mathbf{r}_i)\hat{a}_i^\dagger\right]\left[\sum_j \varphi_j(\mathbf{r}_j)\hat{a}_j^\dagger\right] V(\mathbf{r}_i,\mathbf{r}_j)\left[\sum_k \varphi_k(\mathbf{r}_k)\hat{a}_k\right]\left[\sum_l \varphi_l(\mathbf{r}_l)\hat{a}_l\right] d\mathbf{r}\,d\mathbf{r}' \\
&\equiv \int \hat{\psi}^\dagger(\mathbf{r})T(\mathbf{r})\hat{\psi}(\mathbf{r})\,d\mathbf{r} + \int \hat{\psi}^\dagger(\mathbf{r})\hat{\psi}^\dagger(\mathbf{r}')V(\mathbf{r},\mathbf{r}')\hat{\psi}(\mathbf{r}')\hat{\psi}(\mathbf{r})\,d\mathbf{r}\,d\mathbf{r}' \equiv \hat{\hat{H}} \quad .
\end{aligned} \tag{16.35}$$

This is what we wanted to show. So beside the relation (16.30), we also have:

$$\boxed{i\hbar\frac{\partial}{\partial t}\,|\Psi(t)\rangle_n = \hat{H}_{(\text{many body})}\,|\Psi(t)\rangle_n} \iff \boxed{i\hbar\frac{\partial}{\partial t}\,|\Psi(t)\rangle_n = \hat{\hat{H}}\,|\Psi(t)\rangle_n} \tag{16.36}$$

16.3 Summary

We can say that we have obtained the Hamiltonian of second quantization following two different roads:

- From the "many body" problem in first quantization, via the change of basis $|E_1,....,E_N\rangle \to |n_1,....,n_x\rangle$, we have built the Hamiltonian (16.29) as a function of the creation and annihilation operators. As we showed in (16.35), it is then sufficient one more manipulation to get $\hat{\hat{H}}$.

- Following the recipe of second quantization, we build the field operator $\hat{\psi}$ and with this the Hamiltonmian $\hat{\hat{H}}$.

From the proof of the equivalence betyween $\hat{H}_{(\text{many body})}$ and $\hat{\hat{H}}$, we gather that the introduction of the field operator $\hat{\psi}$ and the machinery of second quantization seems to be, at least, a useful and quick tool to get the correct Hamiltonian for a many-body system. Beside quickness a further advantage brought in by second quantization is that we are not bounded anymore to treat systems with a finite number of particles.

The same approach of second quantization used here for bosons can also be introduced for fermions. With the handling of fermions in second quantization an important result that will pop up, via the comparison with bosons, is the so called *spin-statistics theorem* that in first quantization had to be taken as a *postulate* . We will not present this theorem here and advice the students to read more complete presentations in proper textbooks. Besides this important result, new physical ideas brought to light by second quantization will appear when we treat *relativistic* systems and they are the phenomena of pair creation, vacuum polarization etc..