Quantum Physics I

Giuseppe Carleo*

October 4, 2023

*Institute of Physics, Ecole Polytechnique FÃ©dÃ©rale de Lausanne (EPFL), CH-1015, Lausanne, Switzerland

Contents

Chapter 1

Basic Concepts

Those who are not shocked when they first come across quantum theory cannot possibly have understood it. — Niehls Bohr

1.1 The problems with classical mechanics

Towards the end of the 19th century and the beginning of the 20th century, several new experiments were performed that could not be simply explained using classical mechanics, and that in some cases were completely challenging the expectations based on Newtonian physics. In this course we will not attempt to give an historical account of the development of Quantum Mechanics, but rather present its core concepts and tools. It is however important to know some of the most stringent experimental work that ruled out the possibility of Classical Mechanics as a "good" theory of atoms and light. Here I choose to focus on two experiments, that will also allows us to start developing the core concepts of the theory.

1.2 The Geiger-Marsden experiment

1.2.1 Classical atoms cannot exist

After the experimental discovery of the atomic nucleus in 1911, performed in his group, Rutherford proposed a model in an attempt to explain the properties of the atom. Inspired by the orbiting motion of the planets around the sun, Rutherford considered the atom to consist of electrons orbiting around a positively charged massive center, the nucleus. It was soon recognized that, within the context of classical physics, this model suffers from two serious issues: (a) atoms are unstable and (b) atoms radiate energy over a continuous range of frequencies.

The first deficiency results from the application of Maxwell's electromagnetic theory to Rutherford's model: as the electron orbits around the nucleus, it accelerates and hence radiates energy. It must therefore lose energy. The radius of the orbit should then decrease continuously (spiral motion) until the electron collapses onto the nucleus; the typical time for such a collapse is about $10^{-8}s$.

Second, since the frequency of the radiated energy is the same as the orbiting frequency, and as the electron orbit collapses, its orbiting frequency increases continuously. Thus, the spectrum of the radiation emitted by the atom should be continuous. These two conclusions completely disagree with experiment, since atoms are stable and radiate energy over discrete frequency ranges.

These findings led N. Bohr and others to the first formulation of quantum mechanics (informally known as the "old quantum theory"), largely empirical, that attempted to fix both issues. Here we will not discuss the old formulation of quantum mechanics since, while it has great historical merit, it is entirely replaced by the modern quantum theory we will introduce in this course and it is essentially never used in any modern application.

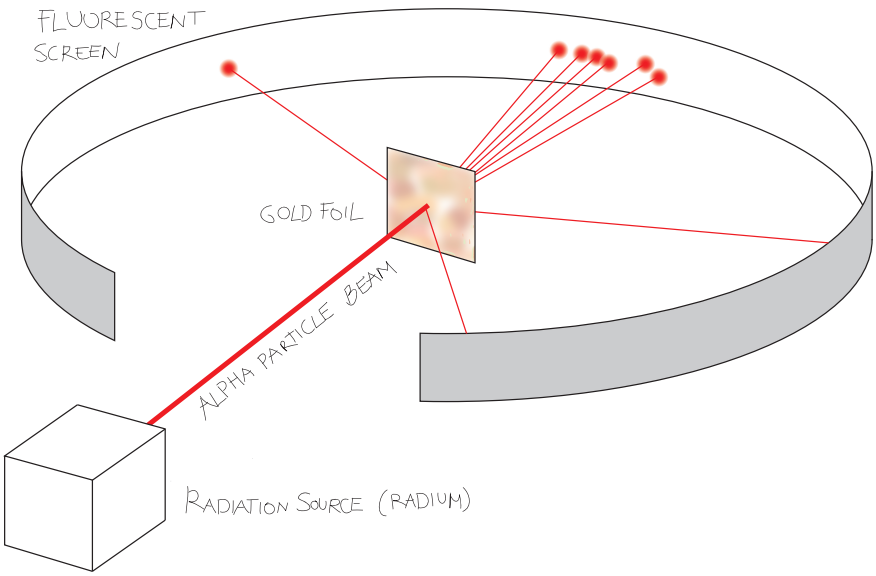

Figure 1.2.1: Schematics of the Geiger-Marsden experiment

1.3 The Stern-Gerlach Experiment

The Stern-Gerlach experiment (1921-1922) shows clearly a physical behavior that Classical Mechanics cannot explain. We will now describe it in a phenomenological way, and in the next chapter we will give the postulates of Quantum Mechanics that provide a correct interpretation of the results.

The goal of the experiment was to measure the *magnetic moment* of a silver atom.

In order to understand what factors can contribute to the magnetic moment, let us concentrate for a moment on a simpler case. Consider the classical model of an hydrogen atom: an electron orbiting a proton of opposite charge with angular momentum $\vec{L}$. Since the electron mass m is much smaller than the proton mass, the magnetic moment of this system is simply given by

$$\vec{\mu}_L = -\frac{e}{2m}\vec{L}. \tag{1.3.1}$$

Now, we can imagine that we send the hydrogen atom through a magnetic field $\vec{B}$. Since the atom has no net electric charge, the interaction energy between the magnetic field and the electric moment is just $V = -\vec{\mu}\cdot\vec{B}$. Assuming for simplicity that the magnetic field points in the z direction, the force acting on the atom is therefore

$$F_z = \frac{\partial}{\partial z}\vec{\mu}\cdot\vec{B} \simeq \mu_z\frac{\partial B_z}{\partial z}. \tag{1.3.2}$$

In addition to the orbital angular momentum, electrons and protons happen to have an intrinsic angular momentum or spin $\vec{S}$. In classical terms, we could think of this degree of freedom as a rotation around their proper axis. The spin angular momentum also induces a magnetic dipole moment

$$\vec{\mu}_S = g\frac{-e}{2m}\vec{S}, \tag{1.3.3}$$

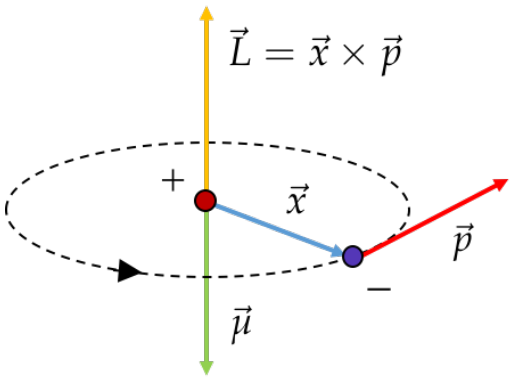

Figure 1.3.1: Magnetic moment of a classical hydrogen atom

where $g \simeq 2$ (for the electron) is known as the gyromagnetic ratio, and this magnetic moment couples to the magnetic field as before. In general, we expect that the total magnetic moment is the sum of the spin and orbital angular momenta.

In their experiment, Stern and Gerlach used silver atoms, passing through an inhomogeneous magnetic field. Silver is made up of a nucleus and 47 electrons, where 46 out of the 47 electrons can be visualized as forming a spherically symmetrical electron cloud with no net angular momentum (both spin and orbital components). The only contribution to the total angular momentum of the atom is due to the intrinsic spin of the 47th electron. Thus, with very good approximation

$$\vec{\mu} \quad \propto \quad \vec{S}.$$

Because of the force (1.3.2), the SG apparatus is therefore an effective way to measure the z component of $\vec{\mu}$, since atoms with $S_z < 0$ will experience a downward force, whereas atoms with $S_z > 0$ will experience an upward force. By measuring how many atoms emerge at which vertical position we therefore have an indirect measurement of S_z.

The atoms entering the apparatus are randomly oriented (basically, because they come from a high-temperature source), thus we could expect that on average each atom will have a random value of $-|\mu| < \mu_z < |\mu|$. We would therefore expect that at the exit of the SG apparatus, a vertical continuous stripe would appear. However, it is experimentally observed that only two isolated spots appear on the screen, implying that S_z can take only two possible values. Experimentally, it is found that these two values are $S_z^{(+)} = +\hbar/2$ and $S_z^{(-)} = +\hbar/2$, where the constant $\hbar$ (h bar) is a fundamental constant in quantum mechanics (Planck's constant) with numerical value $\hbar = 1.054571817 \times 10^{-34}[J \cdot s]$. The fact that the z component of the intrinsic spin of the electron can take only two values cannot be explained by classical mechanics, and is already a striking form of *quantization*, i.e. the fact that microscopic objects subjected to Quantum Mechanics often exhibit discrete values when measured, rather than the continuous values we would expect from classical theory, as in this case.

1.4 Sequential Stern-Gerlach Experiments

We now present a few *Gedankenexperiment* (thought experiments), that, while never realized in the lab by Stern and Gerlach, allow us to understand some of the fundamental concepts we will need to develop a full quantum theory.

To this purpose, we will use denote Stern-Gerlach experiments by the direction chosen for the magnetic field. For example, our previous example would correspond to a SG$\hat{z}$ device, where the magnetic field is aligned along the z direction. We will also use an SG$\hat{x}$

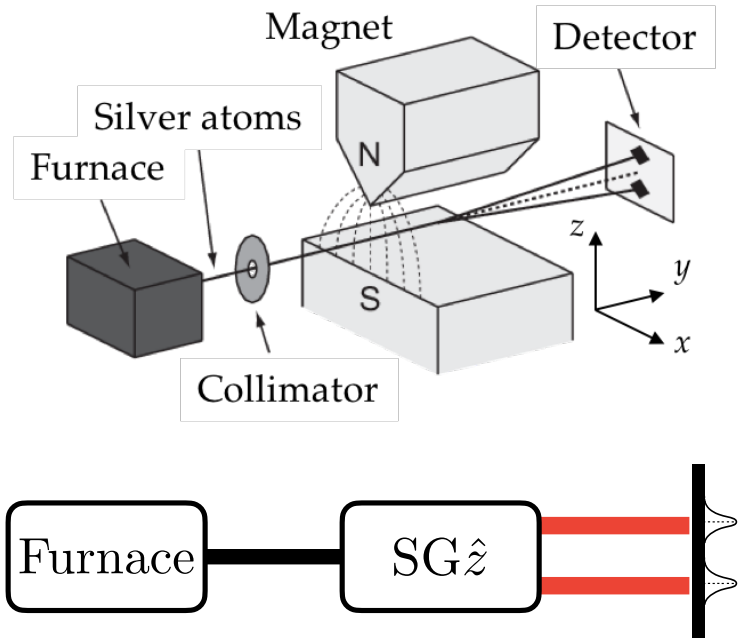

Figure 1.3.2: (Top) Schematic representation of the Stern and Gerlach apparatus. High-temperature leave the furnace on the left and are collimated inside a magnet with a strong magnetic field. The apparatus pushes electrons with positive spin upwards and electrons with negative spins downwards. The two isolated spots appearing on the detector are an indication of the quantization of the spin electron into two possible values instead of a continuum of values expected from classical theory. (Bottom) Schematic representation of the result obtained, where the SG apparatus with a magnet along the z direction is depicted.

device, obtained rotating the original device in a way that it is aligned with the x axis instead.

Additionally, we will consider the case in which we block one of the two beams coming out of the SG device, see Figure 1.4.1, where the down spins are blocked after they exit an SG$\hat{z}$ device.

1.4.1 First Experiment

In the first of our sequential experiments, we stack two SG$\hat{z}$, while blocking out the down component of the spins after the first device, see Figure 1.4.1. Assuming that N spins emerge from the upper part of the first detector, then when they enter the second SG$\hat{z}$ device we observe something quite intuitive for the final measurement outcome: we observe that all spins are up (thus with $S_z = +\hbar/2$) and none is found with $S_z = -\hbar/2$). This is quite intuitive, since the blocking has effectively acted as a filter for the down spins.

Notice that there is nothing special associated with the $\hat{z}$ direction. If we had carried out the same kind of blocking experiment but with two other SG$\hat{d}$ polarizers, where $\hat{d}$ is some arbitrary direction, the same outcome would have been observed. We just stick to the $\hat{z}$ direction as a conventional reference frame.

1.4.2 Second Experiment

The second experiment is already more interesting. In this case, we rotate the last SG device of the previous experiment, and align it in the x direction, thus we have a SG$\hat{x}$ device, see Fig. 1.4.2. In this case, we observe that the output of the last device contains

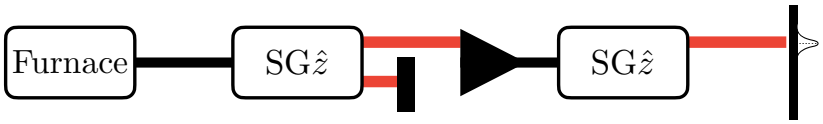

Figure 1.4.1: Experiment with two identical SG devices, and blocking one component before entering the second device. All atoms emerge with the unfiltered state.

again two distinct beams, thus corresponding to two values of the spin $S_x^{(+)} = +\hbar/2$ and $S_x^{(-)} = -\hbar/2$. Why is this the case? One possibility might be that 50% of the $S_z^{(+)}$ atom coming out of the first device are made up of atoms with both $S_z^{(+)}$ and $S_x^{(+)}$ and the other half has instead $S_z^{(+)}$ and $S_x^{(-)}$. This would explain the findings of this experiment, but we will see in the following that this is not what happens.

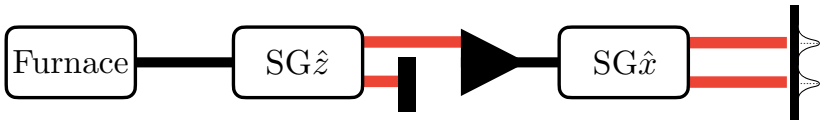

Figure 1.4.2: Experiment with one SG$\hat{z}$ device followed by a SG$\hat{x}$ device, and filtering one component before entering the second device. Both up and down spins are recorded at the output, with equal counts.

1.4.3 Third Experiment

The last experiment is certainly the most disturbing from a *classical* point of view. In this case, we simply add a further SG$\hat{z}$ device to the previous setup, while also blocking the $S_x^{(-)}$ component, see Fig. 1.4.3. What observe now, is that there is *again* a $S_z^{(-)}$ component appearing at the end of the experiment, and that its intensity is identical to that of $S_z^{(+)}$. This is very surprising from the classical standpoint, since we thought that by placing the first filter we had already blocked all particles with $S_z^{(-)}$, however this is not the case. In this sense, the previous hypothesis where the atoms entering the second device do not have a $S_z^{(-)}$ component is not consistent with this experiment.

This example is quite crucial in understanding already one of the peculiarities of quantum mechanics: there are *some* observables that cannot be measured simultaneously. Specifically, what happens is that the selecting the well determined value of $S_x^{(+)}$ destroys any previous information we had on the other direction of the spin. Notice that this phenomenon in an intrinsic *feature* of these observations, and even improving the experimental quality of the apparatus would not solve this issue.

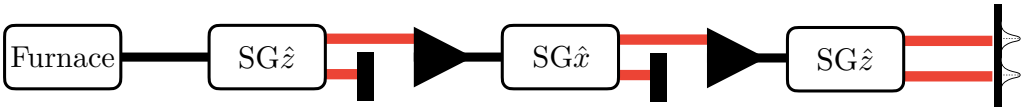

Figure 1.4.3: Experiment with two orthogonal $SG\hat{z}$ and $SG\hat{x}$ filters each blocking their down streams, followed by a SG$\hat{z}$ device. The last device outputs two streams of equal intensity.

1.4.4 Analogy with the polarization of light

While the results of these 3 sequential experiments cannot be explained using a simple classical picture of classical spinning objects, we can find an *analogy* in classical physics

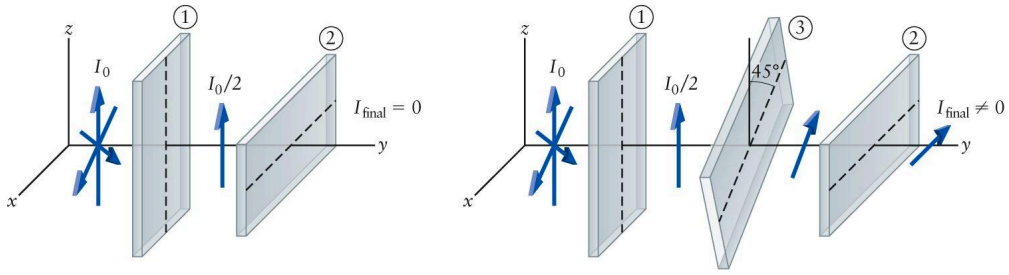

Figure 1.4.4: Unpolarized light entering a sequence of polaroid filters

that can help us materialize also the mathematical tools we need to describe the quantum world.

If you consider monochromatic light with wave vector k and frequency ω traveling in the z direction, then using a polarizer (for example a Polaroid filter) it is possible to obtain linearly polarized light. We can for example have a x-polarizing filter resulting in a time-dependent electric field:

$$\vec{E} = E_0 \hat{x} e^{i(kz-\omega t)},$$

or a y-polarizing filter resulting in

$$\vec{E} = E_0 \hat{y} e^{i(kz-\omega t)}.$$

If we stack two of this filter in a sequence, then the resulting amplitude will be proportional to the scalar product of the two polarization vectors. If we have an x-polarizing filter followed by a y-polarizing filter, then it is well known that no light will come out of the last filter.

If we now insert a third filter in the middle, polarized along a direction we call x' and forming a 45 degrees angle with the x axis, then there is again light coming out of the filter. A similar result would be obtained placing a y' filter forming a 90 degree angle with x', in the plane. The intensity of light can be reconstructed considering the electric fields coming out of the rotated intermediate filters. They read, respectively,

$$\begin{aligned} E_0 e^{i(kz-\omega t)} \hat{x}' &= E_0 e^{i(kz-\omega t)} \left(\frac{\hat{x}}{\sqrt{2}} + \frac{\hat{y}}{\sqrt{2}} \right), \\ E_0 e^{i(kz-\omega t)} \hat{y}' &= E_0 e^{i(kz-\omega t)} \left(-\frac{\hat{x}}{\sqrt{2}} + \frac{\hat{y}}{\sqrt{2}} \right), \end{aligned}$$

thus if we place an x' filter in the middle, then there will be a finite component in the $\hat{y}$ direction that will in turn lead to a finite light intensity at the end of the 3 filters.

This situation is analogous to what seen in the last SG experiment, in the sense that despite the first $\hat{x}$ filter has selected a certain polarization, the second filter "resets" this information and rotates the electric field in a direction that has now a finite $\hat{y}$ direction, that can be detected in the last step.

We can therefore think of an *analogy* between these results with polarizers and those obtained with SG. Specifically, these correspondences lead to similar results in both cases:

Spin state	Polarization state
$S_z^{(+)}$	x
$S_z^{(-)}$	y
$S_x^{(+)}$	$\mathrm{x}' = \frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}}$
$S_x^{(-)}$	$\mathrm{y}' = -\frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}}$

Applying these correspondence relations, we can further argue that the state of the quantum spin could be described by some two-dimensional vector, as much as the results of the light polarizers can be described at all stages, using the x and y components of the electric field. These two-dimensional vectors though live in an *abstract state space* rather than the physical space of the electric field. We therefore postulate that the $S_x^{(+)}$ and $S_x^{(-)}$ are described by vectors that are linear combinations of $S_z^{(+)}$ and $S_z^{(-)}$:

$$\begin{aligned}\overrightarrow{S_x}(+) &= \frac{1}{\sqrt{2}}\left[\overrightarrow{S_z}(+)+\overrightarrow{S_z}(-)\right],\\ \overrightarrow{S_x}(-) &= \frac{1}{\sqrt{2}}\left[-\overrightarrow{S_z}(+)+\overrightarrow{S_z}(-)\right].\end{aligned}$$

In this sense, we can see that when the state $\overrightarrow{S_x}(+)$ emerges from the second SG apparatus (remember that we have blocked $\overrightarrow{S_x}(-)$) it contains components of both $\overrightarrow{S_z}(+)$ and $\overrightarrow{S_z}(-)$ so that when the last apparatus measures the two components, it will find a finite intensity for both Z+ and Z-.

Finally, we can also address the question of representing the quantum state corresponding to $\overrightarrow{S_y}(+)$ and $\overrightarrow{S_y}(-)$. Indeed, we can always think of doing experiments including SGy apparata, with the magnetic field of the detector aligned along the y direction.

For symmetry reasons, we can expect that the situation is analogous to our study of the $z-x$ case, in the sense that we also expect $\overrightarrow{S_y}(+/-)$ to be linear combinations of $\overrightarrow{S_z}(+/-)$. Another analogy with light polarization is helpful here. In general, the electric field of a polarized beam propagating along the z axis can be written as

$$E = E_0 e^{i(kz-\omega t)}\left(\hat{x}+e^{i\phi}\hat{y}\right),$$

where ϕ is a phase difference, and the complex-valued components in the x and y directions are known as Jones vectors.

Circularly polarized light is the other relevant case of light polarization we haven't considered yet, thus it is the natural candidate for the last correspondence. Specifically, the electric field of right circularly polarized light and left circularly polarized light can be written as:

$$\begin{aligned}E_R &= E_0 e^{i(kz-\omega t)}\left(\frac{\hat{x}}{\sqrt{2}}+\frac{i}{\sqrt{2}}\hat{y}\right),\\ E_L &= E_0 e^{i(kz-\omega t)}\left(\frac{\hat{x}}{\sqrt{2}}-\frac{i}{\sqrt{2}}\hat{y}\right).\end{aligned}$$

The analogy we were missing is thus

Spin direction	Polarized light
$S_y^{(+)}$	$R=\frac{x}{\sqrt{2}}+i\frac{y}{\sqrt{2}}$
$S_y^{(-)}$	$\mathrm{L}=\frac{x}{\sqrt{2}}-i\frac{y}{\sqrt{2}}$

and we can therefore postulate that

$$\begin{aligned}\overrightarrow{S_y}(+) &= \frac{1}{\sqrt{2}}\left[\overrightarrow{S_z}(+)+i\overrightarrow{S_z}(-)\right],\\ \overrightarrow{S_y}(-) &= \frac{1}{\sqrt{2}}\left[\overrightarrow{S_z}(+)-i\overrightarrow{S_z}(-)\right].\end{aligned}$$

1.4.5 Beams intensity

So far, we have not discussed the problem of determining the bean intensity (i.e. how many atoms are recorded in each of the SG experiments). Also in this case, the analogy with classical optics is illuminating. In the case of light polarizers, the intensity is found taking the modulus squared of the electric field, at the exit of the filters: $I=|E_x|^2+|E_y|^2$. Amazingly, the same rules apply also for the quantum spin.

1.5 What we learned

From this analogy with optics, we have already discovered several important (and more general) features of quantum physics:

1. Quantization of the observables: contrary to what expected in the classical case, each of the three components of the spin S_x, S_y, S_z can only take integer values $\pm\hbar/2$
2. Indeterminacy of measurements: some observables (for example S_z and S_x) cannot be measured simultaneously.
3. Superposition: like for waves, a quantum state is a superposition of different elementary states.
4. State vector: quantum states are described by complex vectors in an *abstract state space*
5. The average of a quantum measurement is found squaring the components of the quantum state vector.

In the next Chapter, we will mathematically formalize these findings, and set up the tools that will allow us to we need to study quantum phenomena. It is also important to stress here that while the analogy with classical optics is enough to introduce the basic concepts of quantum mechanics, the abstract state space (Hilbert space, as we will see soon) of quantum physics is dramatically different from the state space of electric fields, especially when dealing with composite quantum systems (studied in the last Chapter).

References and Further Reading

1. J. J. Sakurai and J. Napolitano, "Modern Quantum Mechanics" (2017), Chapter 1.1
2. R. P. Feynman, R. Leighton, M. Sands, "The Feynman lectures on physics - Vol. III" (1965), Chapter 1
3. W. Gerlach, and O. Stern, "Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld. Z. Phys. 9, 349–352 (1922)

Chapter 2

Axioms and Tools

This Chapter contains the mathematical basis of the general concepts seen in Chapter 1.

2.1 Axiom 1: State Vectors

In quantum mechanics a physical state—for example, the spin of an electron—is represented by a state vector in a complex vector space. The dimensionality of this vector space in general is unrelated to the physical dimension of the system in exam (say, the familiar 3-dimensional space of classical mechanics coordinates, for example) but instead is an abstract space. In Stern-Gerlach–type experiments where the only quantum-mechanical degree of freedom is the spin of an atom, the dimensionality is determined by the number of alternative paths the atoms can follow when subjected to a SG apparatus; in the case of the silver atoms of the previous section, the dimensionality is just two, corresponding to the two possible values S_z can assume. Later, we will consider the case of continuous degrees of freedom—for example, the position (coordinate) or momentum of a particle—where the number of alternatives is infinite, in which case the vector space in question is known as a HILBERT *space*.

Following DIRAC, we call a vector in this space a *ket* and denote it by $|\psi\rangle \equiv \overrightarrow{\psi}$. This *state ket* is postulated to contain complete information about the physical state; everything we are allowed to ask about the state is contained in the ket.

Essentially, for finite-dimensional spaces, quantum states obey familiar linear algebra properties.

2.1.1 Properties of a vector (Hilbert) space:

1. It is a vector space over the complex numbers $\mathcal{C}$. Vectors in this space are conventionally called kets and denoted by $|\psi\rangle$.

2. The dual of *kets* are called *bras* and denoted by $\langle\psi|$. In the following we denote the dual correspondence with $\leftrightarrow$, and the following relations are postulated

$$|\psi\rangle \quad \leftrightarrow \quad \langle\psi| \tag{2.1.1}$$
$$c|\psi\rangle \quad \leftrightarrow \quad \langle\psi|c^\star, \tag{2.1.2}$$

   thus kets correspond to bras, and a constant c times a ket corresponds to a bra times the complex conjugate of that constant, $c^\star$. Informally speaking, for finite-dimensional spaces, kets correspond to column vectors ($n \times 1$ matrices), and bras correspond to row vectors ($1 \times n$ matrices) . In this case bras thus correspond to $\langle\psi| \equiv \left(\overrightarrow{\psi}\right)^\dagger$, where $\dagger$ denotes the conjugate transpose of the vector.

3. It has an inner product $\langle\psi|\phi\rangle$ that maps an ordered pair of vectors (in this case to a complex number), and that has the properties:

a) Positivity: $\langle\psi|\psi\rangle > 0$ for $|\psi\rangle \neq 0$.
b) Linearity: $\langle\varphi|(a|\psi_1\rangle + b|\psi_2\rangle) = a\langle\varphi|\psi_1\rangle + b\langle\varphi|\psi_2\rangle$
c) Symmetry: $\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^\star$, where the $\star$ denotes complex conjugation.

Notice again that all these three properties are very natural for complex vectors, once the inner product is identified with the usual dot product: $\langle\psi|\phi\rangle \equiv \vec{\psi}\cdot\vec{\phi}$.

2.1.2 Quantum states are rays

What is a ray? It is an equivalence class of vectors that differ by multiplication by a nonzero complex scalar. For any nonzero ray, we can by convention choose a representative of the class, denoted $|\psi\rangle$, that has unit norm:

$$\langle\psi|\psi\rangle = 1, \tag{2.1.3}$$

and all other states obtained multiplying this state by an arbitrary non-zero constant represent the same physical state:

$$c|\psi\rangle \quad \equiv \quad |\psi\rangle. \tag{2.1.4}$$

Since every ray corresponds to a possible state, given two states $|\varphi\rangle$, $|\psi\rangle$, another state can be constructed as the linear superposition of the two:

$$|\psi'\rangle = a|\varphi\rangle + b|\psi\rangle. \tag{2.1.5}$$

Notice that the *global* phase of the state is irrelevant, thus $e^{i\alpha}|\psi'\rangle \equiv |\psi'\rangle$, because of Eq. (2.1.4). The *relative* phase in this superposition is however physically significant. For example, the state $a|\varphi\rangle + b|\varphi\rangle$ is the same ray as $e^{i\alpha}(a|\varphi\rangle + b|\psi\rangle)$ but it is different from $a|\varphi\rangle + e^{i\alpha}b|\psi\rangle$.

2.2 Operators

Operators are the natural companion of state vectors. They are the tool used to do all physically meaningful *manipulations* of a quantum state. In the following, we denote operators with $\hat{A}$, to distinguish them from scalars. An operator acts on a ket from the left,

$$\hat{A}(|\psi\rangle) \quad = \quad \hat{A}|\psi\rangle, \tag{2.2.1}$$

resulting into another ket. In the familiar case of finite-dimensional vector spaces, operators are nothing but matrices acting on vectors. Thus the action of a matrix onto a vector results into another vector. In the more general case of Hilbert spaces, observables are linear maps taking vectors to vectors:

$$\hat{A} : |\psi\rangle \quad \rightarrow \quad \hat{A}|\psi\rangle \tag{2.2.2}$$

$$\hat{A}(a|\psi\rangle + b|\phi\rangle) \quad = \quad a\hat{A}|\psi\rangle + b\hat{A}|\phi\rangle. \tag{2.2.3}$$

2.2.1 Eigen-Kets

In general when an operator acts on a ket, it produces a distinct ket. However, there are special cases in which the application of an operator leads to a constant times the initial ket. Those are known as eigen-kets, $|A_1\rangle$, $|A_2\rangle$, $|A_3\rangle$, ...$|A_n\rangle$, and have the property that

$$\begin{aligned} \hat{A}|A_1\rangle &= a_1|A_1\rangle \\ \hat{A}|A_2\rangle &= a_2|A_2\rangle \\ \cdots \quad &\cdots \quad \cdots \\ \hat{A}|A_n\rangle &= a_n|A_n\rangle, \end{aligned} \tag{2.2.4}$$

where the action of the operator is to return the same kets multiplied by scalars (in general, complex-valued) $a_1, a_2, \ldots a_n$. Those are eigenvalues of the operator A, and the corresponding states $|A_1\rangle, |A_2\rangle, \ldots |A_n\rangle$, are eigen-kets. Again, for finite-dimensional spaces the notion of eigen-ket is strictly equivalent to that of eigen-vectors in linear algebra.

2.2.2 Adjoints and Hermitian Operators

The adjoint of the operator is denoted as $\hat{A}^\dagger$ and is defined by the dual relationship

$$\hat{A}|\psi\rangle \quad \leftrightarrow \quad \langle\psi|\hat{A}^\dagger, \tag{2.2.5}$$

for all states $|\psi\rangle$. A special class of operators is that of Hermitian operators $\hat{A}$, for which $\hat{A} = \hat{A}^\dagger$. As we will clarify in the following, Hermitian operators in quantum physics play an important role in the measurement process.

Theorem 1. *An Hermitian operator satisfies* $\langle\phi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}|\phi\rangle^\star$.

Proof. By the symmetry property of the inner product we have $\langle\phi|\hat{A}|\psi\rangle = \langle\phi|\hat{A}\psi\rangle = \langle\psi\hat{A}|\phi\rangle^\star$. By the definition of adjoint, we have that the dual of $|\hat{A}\psi\rangle \leftrightarrow \langle\psi|\hat{A}^\dagger = \langle\psi\hat{A}|$, thus $\langle\phi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}^\dagger|\phi\rangle^\star$. If $\hat{A} = \hat{A}^\dagger$ then $\langle\phi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}|\phi\rangle^\star$. □

The spectrum of eigenvalues associated with Hermitian operators has special properties, as given by the following Theorem.

Theorem 2. *The eigenvalues of an Hermitian operator* $\hat{A}$ *are real valued. The eigen-kets of* $\hat{A}$ *corresponding to different eigenvalues are orthogonal.*

Proof. Consider

$$\hat{A}|A_i\rangle = a_i|A_i\rangle, \tag{2.2.6}$$

and its dual equivalent (for some other eigen-ket j) :

$$\langle A_j|\hat{A}^\dagger = \langle A_j|a_j^\star \tag{2.2.7}$$

$$= \langle A_j|\hat{A}, \tag{2.2.8}$$

where the last line follows from the the fact that the operator is Hermitian. Multiplying the first equation on the left by $\langle A_j|$ and the second equation on the right by $|A_i\rangle$, and subtract the results we get

$$\langle A_j|A_i\rangle(a_i - a_j^\star) \quad = \quad 0. \tag{2.2.9}$$

In the case in which $i = j$ (same eigenvector), we then must have $a_i = a_i^\star$, thus the eigenvalues are real. In the case $i \neq j$, under the theorem assumption that the the eigenvalues are different (and real, because of the just proven result), we have that $(a_i - a_j) \neq 0$, thus the only possibility to satisfy equation (2.2.9) is that $\langle A_j|A_i\rangle = 0$, thus the two eigenvectors must be orthogonal. □

In the following we will assume that the eigen-kets of operators are taken to be orthonormal, i.e. satisfy the condition

$$\langle A_j|A_i\rangle \quad = \quad \delta_{ij}. \tag{2.2.10}$$

Notice that this condition can always be enforced, because of the orthogonality condition of the previous theorem and because, as we have seen before, the normalization of each $|A_i\rangle$ is arbitrary (we can thus take it to be 1).

2.2.3 Representing state kets with eigen-kets

Choosing an operator $\hat{A}$ and forming its eigen-kets is in general a very important conceptual and practical step needed to represent arbitrary ket states. We will use eigen-kets of operators as base kets to expand arbitrary kets, as much as for an Euclidean space one uses orthogonal unit vectors (coordinates) to represent an arbitrary vector.

Concretely, an arbitrary ket $|\psi\rangle$ can be represented as a linear combination of eigen-kets of some operator $\hat{A}$ in this way:

$$|\psi\rangle \quad = \quad \sum_i c_i |A_i\rangle, \tag{2.2.11}$$

where the complex-valued coefficients c_i are to be determined. Multiplying (2.2.11) on the left by $\langle A_j|$ and using the orthonormality condition, Eq. (2.2.10), we obtain

$$c_i \quad = \quad \langle A_i|\psi\rangle. \tag{2.2.12}$$

The expansion coefficients c_i (often referred to as *amplitudes* of the state on the eigen-kets of $\hat{A}$) are then formally computed as an inner product of a bra $\langle A_i|$ with a ket $|\psi\rangle$. We also notice that the normalization of the state reads

$$\langle\psi|\psi\rangle \quad = \quad \sum_{ij} \langle A_j|c_j^\star c_i|A_i\rangle = \sum_i |c_i|^2, \tag{2.2.13}$$

thus an equivalent condition for the state to be normalized is that $\langle\psi|\psi\rangle = \sum_i |c_i|^2 = 1$.

We can also rewrite the expansion in Eq. (2.2.11) as

$$|\psi\rangle \quad = \quad \sum_i |A_i\rangle\langle A_i|\psi\rangle. \tag{2.2.14}$$

Very interestingly, we can interpret $\hat{P}_i = |A_i\rangle\langle A_i|$ to be an operator acting on ket $|\psi\rangle$, from which we also deduce a very important relationship called completeness relation or closure:

$$\sum_i |A_i\rangle\langle A_i| \quad = \quad \hat{1}. \tag{2.2.15}$$

This relation follows from Eq. (2.2.14), that must be verified for arbitrary $|\psi\rangle$. The operator $\hat{\Lambda}_i = |A_i\rangle\langle A_i|$ is also an important tool in quantum mechanics, and it is known as *projector* operator, since it projects an arbitrary state vector onto a certain eigen-ket i of $\hat{A}$.

Exercise 3. Show that the projector operator satisfies $\hat{\Lambda}_i^2 = \hat{\Lambda}_i$.

2.2.4 Matrix Representation

The projection operator previously introduced is also important to highlight the direct connection between operators and matrices. Specifically, given an arbitrary operator $\hat{B}$ we can insert the completeness relation twice:

$$\hat{B} \quad = \quad \underbrace{\sum_i |A_i\rangle\langle A_i|}_{\hat{1}} \hat{B} \tag{2.2.16}$$

$$= \quad \sum_{ij} |A_i\rangle\langle A_i|\hat{B}|A_j\rangle\langle A_j|, \tag{2.2.17}$$

and identify $B(i,j) \equiv \langle A_i|\hat{B}|A_j\rangle$ as the *matrix element* of the operator $\hat{B}$ in the basis of the eigenkets i and j of $\hat{A}$. We can then explicitly write $\hat{B}$ as a $N \times N$ matrix with elements

$$\hat{B} \quad = \quad \begin{pmatrix} \langle A_1|\hat{B}|A_1\rangle & \langle A_1|\hat{B}|A_2\rangle & \dots & \langle A_1|\hat{B}|A_n\rangle \\ \langle A_2|\hat{B}|A_1\rangle & \langle A_2|\hat{B}|A_2\rangle & \dots & \langle A_2|\hat{B}|A_n\rangle \\ \dots & \dots & \dots & \dots \\ \langle A_n|\hat{B}|A_1\rangle & \langle A_n|\hat{B}|A_2\rangle & \dots & \langle A_n|\hat{B}|A_n\rangle \end{pmatrix}, \tag{2.2.18}$$

where we have assumed that the dimensionality of the vector space is finite and equal to n. For example, for our spin 1/2 case, we have that the dimensionality is $n = 2$.

This explicit matrix representation also allows to better clarify the concept of Hermitian conjugate (adjoint) operator $\hat{A}^\dagger$, since the result of Theorem 1 (namely $\langle\phi|\hat{B}|\psi\rangle = \langle\psi|\hat{B}|\phi\rangle^\star$ is true if $\hat{B}$ is an Hermitian operator) directly translates into a condition for the matrix elements

$$B(i,j) \quad = \quad B(j,i)^\star,$$

thus for finite-dimensional Hilbert spaces an Hermitian operator is nothing but an Hermitian matrix.

Further notice that the representation of the operator $\hat{A}$ itself in its eigen-ket basis is nothing but a diagonal matrix, whose elements are the eigenvalues, i.e. $A(i,j) = \delta_{ij} a_i$. It also immediately follows that $\hat{A} = \sum_i a_i |A_i\rangle\langle A_i|$.

Example 4. Dirac's notation is a powerful way of writing expressions involving linear operators. For example, consider the case of an operator $\hat{D} = \hat{B}\hat{C}$, then the matrix elements can be obtained using again the completeness relation:

$$\begin{aligned} D(i,j) &= \langle A_i|\hat{B}\hat{C}|A_j\rangle && (2.2.19)\\ &= \sum_k \langle A_i|\hat{B}\underbrace{|A_k\rangle\langle A_k|}_{\hat{1}}\hat{C}|A_j\rangle && (2.2.20)\\ &= \sum_k B(i,k)C(k,j), && (2.2.21) \end{aligned}$$

which correspond to the usual notion of matrix multiplication.

2.2.5 Finding eigen-kets

The explicit representation of operators in terms of matrix elements is also very useful to explicitly find eigen-kets and eigenvalues, given a certain operator. Let us consider again the matrix elements of the operator $\hat{B}$ in the $\hat{A}$ basis, namely: $B(i,j) \equiv \langle A_i|\hat{B}|A_j\rangle$. The eigenvalue equation is

$$\hat{B}|B_k\rangle \quad = \quad b_k|B_k\rangle, \tag{2.2.22}$$

for the unknown b_k and $|B_k\rangle$. We can rewrite this as

$$\begin{aligned} \hat{B}|B_k\rangle &= b_k|B_k\rangle && (2.2.23)\\ \hat{B}\sum_j |A_j\rangle\langle A_j|B_k\rangle &= b_k|B_k\rangle && (2.2.24)\\ \langle A_i|\hat{B}\sum_j |A_j\rangle\langle A_j|B_k\rangle &= b_k\langle A_i|B_k\rangle && (2.2.25)\\ \sum_j \langle A_i|\hat{B}|A_j\rangle\langle A_j|B_k\rangle &= b_k\langle A_i|B_k\rangle, && (2.2.26) \end{aligned}$$

which in matrix notation is

$$\underbrace{\begin{pmatrix} \langle A_1|\hat{B}|A_1\rangle & \cdots & \langle A_1|\hat{B}|A_n\rangle \\ \vdots & \ddots & \vdots \\ \langle A_n|\hat{B}|A_1\rangle & \cdots & \langle A_n|\hat{B}|A_n\rangle \end{pmatrix}}_{\hat{M}} \underbrace{\begin{pmatrix} \langle A_1|B_k\rangle \\ \vdots \\ \langle A_n|B_k\rangle \end{pmatrix}}_{\vec{v}^{(k)}} = b_k \begin{pmatrix} \langle A_1|B_k\rangle \\ \vdots \\ \langle A_n|B_k\rangle \end{pmatrix}, \tag{2.2.27}$$

thus the eigenvalues are found, as in standard linear algebra, as solutions of the equation

$$\det(\hat{M} - b_k\hat{I}) \quad = \quad 0, \tag{2.2.28}$$

and once the b_k are found, we solve the homogenous linear system

$$(\hat{M} - b_k\hat{I})\vec{v}^{(k)} \quad = \quad 0, \tag{2.2.29}$$

for the unknown vectors $\vec{v}^{(k)}$. This procedure is described in detail in all linear algebra books.

2.3 Axiom 2: Measurement

We now come to one of the most fundamental, yet counterintuitive (because it wildly departs from the classical world) axioms of quantum mechanics, related to how a measurement is performed. Quite generally, a measurement is a process in which information about the state of a physical system is acquired by an observer. An observable is a property of a physical system that in principle can be measured. Such property could be for example momentum and spin components, etc. In quantum mechanics, it is postulated that an observable is represented by an Hermitian (also known as *self-adjoint*) operator acting on the vector space of quantum states. The fundamental axiom of quantum mechanics (that cannot be proven) is that the measurement of an observable $\hat{A}$ *prepares* an eigenstate of the hermitian operator $\hat{A}$, and the observer learns the value of the corresponding eigenvalue.

In essence, let us assume that we have a certain quantum system described by a ket $|\psi\rangle$, whose expansion in the eigen-kets of $\hat{A}$ reads

$$|\psi\rangle = \sum_i c_i |A_i\rangle, \tag{2.3.1}$$

where, as shown before, the expansion coefficients (amplitudes) are $c_i = \langle A_i|\psi\rangle$. We also assume here that the state $|\psi\rangle$ is a *ray* chosen to be normalized, thus $\sum_i |c_i|^2 = 1$.

The measurement axiom means that when we measure the operator $\hat{A}$, the state $|\psi\rangle$ *immediately* collapses into one of the possible eigenstates $|A_i\rangle$ of $\hat{A}$, and the result of that specific measurement will be the associated eigenvalue, a_i.

The most important aspect of the measurement process, is that *which* of the several eigen-kets is obtained is determined only probabilistically. Specifically, we say that the outcome a_i for the measurement is obtained with a priori probability

$$\begin{aligned} \text{Prob}(a_i) &= |\langle A_i|\psi\rangle|^2 \\ &= |c_i|^2 \\ &= \langle\psi|A_i\rangle\langle A_i|\psi\rangle, \end{aligned} \tag{2.3.2}$$

known as Born's probability rule, introduced by Max Born in a seminal 1926 paper. We can immediately verify that the probability defined by Born's rule is a *bona fide* probability, in the sense that it is correctly normalized:

$$\sum_i \text{Prob}(a_i) = \sum_i |c_i|^2 \tag{2.3.3}$$

$$= \langle\psi|\psi\rangle \tag{2.3.4}$$

$$= 1, \tag{2.3.5}$$

where the last equality comes from the normalization condition of the state $|\psi\rangle$.

2.3.1 Repeated measurements

If many identically prepared systems are measured, each described by the same state $|\psi\rangle$, then the expectation value of the outcomes is what you would expect from standard probability theory, namely

$$\langle A\rangle \equiv \sum_i a_i \text{Prob}(a_i) \tag{2.3.6}$$

$$= \sum_i a_i \langle\psi|A_i\rangle\langle A_i|\psi\rangle \tag{2.3.7}$$

$$= \sum_i \langle\psi|\hat{A}|A_i\rangle\langle A_i|\psi\rangle \tag{2.3.8}$$

$$= \langle\psi|\hat{A}\underbrace{\sum_i |A_i\rangle\langle A_i|}_{\hat{1}}\psi\rangle \tag{2.3.9}$$

$$= \langle\psi|\hat{A}|\psi\rangle. \tag{2.3.10}$$

The latter equation is one of the most important equations of quantum mechanics, since it relates the average result for repeated measurements to the quantity $\langle\psi|\hat{A}|\psi\rangle$, known as the "expectation value" of the corresponding operator.

It should be understood here that by "repeated measurement" we mean, strictly, preparing the state $|\psi\rangle$ several times, and each time measuring the observable A. Each observation $k = 1, 2, \dots M$ will result in a random result $r_k \in \{a_1, \dots a_n\}$. An experimental observer can then estimate $\langle A\rangle$ with the simple mean

$$\langle A\rangle \simeq \frac{1}{M}\sum_k r_k, \tag{2.3.11}$$

and in the limit $M \to \infty$ this will coincide with the computed expression $\langle\psi|\hat{A}|\psi\rangle$.

A dramatically different scenario is instead obtained if we prepare the state $|\psi\rangle$ only once, and we perform a measurement over the *same* state over and over again. In this case, after the first measurement, the state will collapse to a corresponding random eigen-ket, say

$$|\psi\rangle \quad \underrightarrow{\text{measurement}} \quad |A_i\rangle, \tag{2.3.12}$$

with probability $P_i = |\langle A_i|\psi\rangle|^2$, and resulting in the value a_i for the measurement outcome. However, since the new state resulting from the measurement is just an eigenstate of the measurement operator, we have that the new state has $c'_i = 1$, thus if we measure the operator $\hat{A}$ *again*, the result of the measurement will be again a_i, with probability 1, a deterministic measurement!

2.4 Analysis of the Stern-Gerlach experiments

Having introduced the most fundamental postulates of quantum mechanics, we are now in position to resolve one unsatisfactory argument that we had to introduced at the beginning of these lectures: the form of the eigenstates of the spin operators.

First of all, it is very important to realize that the kind of experiment we are analyzing here falls under the second type described above (what we have called, repeated measurements). The reason is that the SG apparatus does not measure the spin of an individual electron, but rather of a large number of electrons at once. Schematically, we can think that independent electrons pass through the analyzer and each of them is deflected either upwards or downwards. What we are doing then is essentially equivalent to preparing the same state and measuring many times.

2.4.1 The operator $\hat{S}_z$

When we measure the z component of the spin in our SG apparatus, we postulate the existence of a corresponding measure operator that we call $\hat{S}_z$. When the spin is measured along the z direction, the system will then immediately collapse in one of the two eigen-kets of this operator: $|+\rangle$ or $|-\rangle$, and the result of the measurement will be, respectively, one of the two eigenvalues of the $\hat{S}_z$ operator, thus either $a_+ = +\hbar/2$ or $a_- = -\hbar/2$. In order to find an explicit expression for the spin operators, we start observing that the identity operator for a 2-dimensional finite vector space can be written as

$$\hat{1} = |+\rangle\langle+| + |-\rangle\langle-|, \tag{2.4.1}$$

and the representation of $\hat{S}_z$ in the basis of its eigenvectors is just

$$\hat{S}_z = \frac{\hbar}{2}\left(|+\rangle\langle+| - |-\rangle\langle-|\right). \tag{2.4.2}$$

The operator is also just a 2×2 diagonal matrix in this basis

$$\hat{S}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.4.3}$$

and the notion of eigen-kets being *vectors* is particularly clear when writing them as the algebraic eigenvectors of this matrix, namely

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.4.4}$$

$$|-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.4.5}$$

This also clarifies why we have previously remarked that kets are column vectors, whereas bras are row vectors:

$$\langle +| = \begin{pmatrix} 1 & 0 \end{pmatrix} \tag{2.4.6}$$

$$\langle -| = \begin{pmatrix} 0 & 1 \end{pmatrix}. \tag{2.4.7}$$

2.4.2 The operator $\hat{S}_x$

From Fig. 1.4.3, we have seen that when a beam of type $|S_x;+\rangle$ goes again through a S_z measurement, the beam is deflected in both directions, and we that that the counts observed in the SG experiment are equal: $N(+) = N(-)$. This means that, in general,

$$|S_x;+\rangle = \frac{1}{\sqrt{2}}|S_z;+\rangle + \frac{1}{\sqrt{2}}e^{i\delta_1}|S_z;-\rangle, \tag{2.4.8}$$

where δ_1 is a real-valued phase that we will determine in a moment. You can verify that this form is correct because:

1. When measuring S_z in the last stage, we apply the operator $\hat{S}_z$, thus the state collapses to one of the two eigenstates of $\hat{S}_z$, with equal probabilities $P(+) = \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2}$ and $P(-) = \left|\frac{1}{\sqrt{2}}e^{i\delta_1}\right|^2 = \frac{1}{2}$.
2. The state (2.4.8) is a correctly normalized ket, indeed: $\langle S_x;+|S_x;+\rangle = 1$, as it is easy to verify.

We can also find $|S_x;-\rangle$ only using the postulates we have introduced above. Indeed, we know that different eigen-kets of the same operator are orthogonal, thus we must have $\langle S_x;+|S_x;-\rangle = 0$, as well as $\langle S_x;-|S_x;-\rangle = 1$, these two conditions fix the form of the other eigen-ket:

$$|S_x;-\rangle = \frac{1}{\sqrt{2}}|S_z;+\rangle - \frac{1}{\sqrt{2}}e^{i\delta_1}|S_z;-\rangle, \tag{2.4.9}$$

Now, it should be remarked that using only the results of Experiment 3, we cannot determine the value of the phase factor δ_1, since the information we have from the counts allows us only to reconstruct the square modulus of the amplitudes, and not the amplitudes themselves. To find δ_1 we need more information.

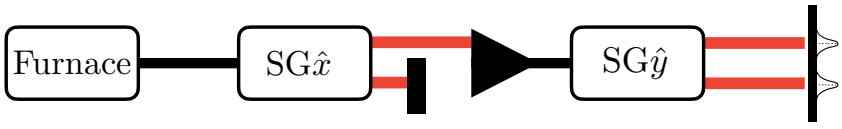

Figure 2.4.1: Experiment with two SG devices aligned in the x and y directions, respectively and blocking one component before entering the second device. All atoms emerge with the unfiltered state. Notice that in this case atoms exiting the furnace arrive from the z direction.

2.4.3 The operator $\hat{S}_y$

A similar analysis can be carried out for the y component of the spin, leading to:

$$|S_y;\pm\rangle = \frac{1}{\sqrt{2}}|S_z;+\rangle \pm \frac{1}{\sqrt{2}}e^{i\delta_2}|S_z;-\rangle, \tag{2.4.10}$$

where we have introduced yet another phase factor, δ_2, to be determined. In order to determine both phase factors, we consider the experiment as in Fig. 2.4.1. This class of experiments gives the same that was found before when considering the z and x directions, and not surprisingly so, because of symmetry reasons. This however implies that

$$|\langle S_x;+|S_y;+\rangle|^2 = |\langle S_x;+|S_y;-\rangle|^2 = \frac{1}{2}, \tag{2.4.11}$$

when blocking the $x-$ component of the spin (as in Figure) and

$$|\langle S_x;-|S_y;+\rangle|^2 = |\langle S_x;-|S_y;-\rangle|^2 = \frac{1}{2}, \tag{2.4.12}$$

when blocking the $x+$ component of the spin. These conditions allow to fix the phase factors, indeed we find the condition

$$\begin{aligned} |\langle S_x;\pm|S_y;+\rangle|^2 &= \left|\left(\frac{1}{\sqrt{2}}\langle S_z;+| \pm \frac{1}{\sqrt{2}}e^{-i\delta_1}\langle S_z;-|\right)\times\right. && (2.4.13)\\ &\quad \left.\left(\frac{1}{\sqrt{2}}|S_z;+\rangle + \frac{1}{\sqrt{2}}e^{i\delta_2}|S_z;-\rangle\right)\right|^2 && (2.4.14)\\ &= \left|\left(\frac{1}{2} \pm \frac{1}{2}e^{i(\delta_2-\delta_1)}\right)\right|^2 && (2.4.15)\\ &= \frac{1}{2}, && (2.4.16) \end{aligned}$$

which has a solution for $\delta_2 - \delta_1 = \pm\pi/2$, since $e^{\pm i\frac{\pi}{2}} = \pm i$, thus $|\frac{1}{2}(1 \mp i)|^2 = \frac{1}{2}$. While the phase difference is physical, there is no way (but for conventional reasons) to fix separately δ_1 and δ_2. The usual convention is to take $\delta_1 = 0$ and $\delta_2 = \pi/2$, yielding:

$$\begin{aligned} |S_x;\pm\rangle &= \frac{1}{\sqrt{2}}|S_z;+\rangle \pm \frac{1}{\sqrt{2}}|S_z;-\rangle, && (2.4.17)\\ |S_y;\pm\rangle &= \frac{1}{\sqrt{2}}|S_z;+\rangle \pm \frac{i}{\sqrt{2}}|S_z;-\rangle. && (2.4.18) \end{aligned}$$

The corresponding operators are

$$\begin{aligned} \hat{S}_x &= \frac{\hbar}{2}\left(|S_x;+\rangle\langle S_x;+| - |S_x;-\rangle\langle S_x;-|\right) && (2.4.19)\\ &= \frac{\hbar}{2}\left(|S_z;+\rangle\langle S_z;-| + |S_z;-\rangle\langle S_z;+|\right), && (2.4.20) \end{aligned}$$

and

$$\begin{aligned} \hat{S}_y &= \frac{\hbar}{2}\left(|S_y;+\rangle\langle S_y;+| - |S_y;-\rangle\langle S_y;-|\right) && (2.4.21)\\ &= \frac{\hbar}{2}i\left(-|S_z;+\rangle\langle S_z;-| + |S_z;-\rangle\langle S_z;+|\right). && (2.4.22) \end{aligned}$$

2.4.4 Pauli matrices

Putting together the results we have found for the three spin operators, we can finally also compute the corresponding matrix representation of these operators. They read:

$$\hat{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{2.4.23}$$

$$\hat{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{2.4.24}$$

$$\hat{S}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.4.25}$$

The matrices that appear in these equations are very famous, and called Pauli matrices $\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$, such that $\hat{S}_\alpha = \frac{\hbar}{2}\hat{\sigma}_\alpha$.

2.5 Commuting Observables

Two observables $\hat{A}$ and $\hat{B}$ are said to be *compatible* when their commutator $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$ is vanishing, i.e.

$$[\hat{A}, \hat{B}] = 0, \tag{2.5.1}$$

and incompatible when

$$[\hat{A}, \hat{B}] \neq 0. \tag{2.5.2}$$

Considering our spin example, we see that $\hat{S}_x$ and $\hat{S}_z$ are incompatible, since

$$[\hat{S}_z, \hat{S}_x] = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{2.5.3}$$

$$= -\frac{\hbar}{2i}\hat{\sigma}_y \tag{2.5.4}$$

$$= i\hbar\hat{S}_y, \tag{2.5.5}$$

whereas $\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2$ is compatible with $\hat{S}_z$ (and the other spin operators along the other directions).

Why do we make such a distinction between compatible and incompatible observables?

Theorem 5. *Suppose that A and B are compatible observables, and the eigenvalues of A are non-degenerate. Then the matrix representation of $\hat{B}$ is diagonal in the basis of A, thus $\langle A_i|\hat{B}|A_j\rangle = \delta_{ij}b_i$, where the lower script A here denotes that the matrix elements are w.r.t. to the eigenkets of A.*

Proof. Using the fact that $\hat{A}$ and $\hat{B}$ commute, we have

$$\langle A_i|[\hat{A}, \hat{B}]|A_j\rangle = \langle A_i|\hat{A}\hat{B} - \hat{B}\hat{A}|A_j\rangle \tag{2.5.6}$$

$$= (a_i - a_j)\langle A_i|\hat{B}|A_j\rangle \tag{2.5.7}$$

$$= 0, \tag{2.5.8}$$

thus $\langle A_i|\hat{B}|A_j\rangle$ must vanish for $i \neq j$. □

Important consequence of this theorem is that $|A_i\rangle$ and $|B_i\rangle$ are eigen-kets of both $\hat{A}$ and $\hat{B}$. This can shown using the fact that $\hat{B}$ is diagonal in the $\hat{A}$ eigen-basis, thus the decomposition

$$\hat{B} = \sum_i b_i|A_i\rangle\langle A_i|, \tag{2.5.9}$$

holds, which immediately implies that if we apply this operator to an eigen-ket of $\hat{A}$, we get

$$\hat{B}|A_j\rangle = \sum_i b_i |A_i\rangle\langle A_i|A_j\rangle \tag{2.5.10}$$

$$= b_j|A_j\rangle, \tag{2.5.11}$$

thus it is an eigen-ket of $\hat{B}$, and we also identify the diagonal matrix elements as the eigenvalues b_j. In general, when a certain ket is an eigen-ket of more than one operator, we typically denote it as $|A_i, B_i\rangle$ or, often, with a collective name $|K_i\rangle$.

Fundamental consequence of the commutativity of observables is therefore that the measurement process in this case is familiarly similar to what *would* happen in the classical case. For example, imagine again a state $|\psi\rangle$, and that we measure the observable A, then the result of the measurement will yield some random value a_j and the state will collapse in the corresponding eigenstate $|A_j\rangle$. Measuring now B will result in the value b_j, with probability 1 (recall that $|A_j\rangle$ is also an eigenvalue of B). Further measuring A would again return a_j, thus the measurement done with B has not destroyed (or affected in any way) the state of the system, as per the observable B is concerned.

This is a familiar situation in classical mechanics, in the sense that we can expect to be able to measure different quantities (say, velocity and position of a particle, for example) *without* changing the state of the system itself. This notion however breaks dramatically when considering non-commuting observables.

2.6 Non-Commuting observables

We now consider the important case in which observables are not commuting, thus $[\hat{A}, \hat{B}] \neq 0$. In this case, $\hat{A}$ and $\hat{B}$ do not share a set of common eigen-kets.

Theorem 6. *When $[\hat{A}, \hat{B}] \neq 0$, then it is not possible to find, in general, a complete set of simultaneous eigen-kets of $\hat{A}$ and $\hat{B}$*

Proof. Let us suppose that the converse is true, then we have that $\hat{A}\hat{B}|A_iB_i\rangle = \hat{A}b_i|A_iB_i\rangle = a_ib_i|A_iB_i\rangle$, and also $\hat{B}\hat{A}|A_iB_i\rangle = \hat{B}a_i|A_iB_i\rangle = a_ib_i|A_iB_i\rangle$. Since $\hat{A}\hat{B}|A_iB_i\rangle = \hat{B}\hat{A}|A_iB_i\rangle$ for all eigen-kets i, then this implies $\hat{A}\hat{B} = \hat{B}\hat{A}$, which is in contradiction of the assumption. □

2.7 The uncertainty principle

In Sec. 2.3.1, we have analyzed the case of repeated measurements, and came to the conclusion that expectation value of a given operator over many experiments is given by:

$$\langle A\rangle \equiv \langle\psi|\hat{A}|\psi\rangle. \tag{2.7.1}$$

In addition to the expectation value, we can also compute the variance associated with the measurement of the operator. To this end, we introduce the displacement operator defined as :

$$\widehat{\Delta A} = \hat{A} - \langle A\rangle\hat{I}, \tag{2.7.2}$$

such that the expectation value of its square is the variance:

$$\langle \Delta A^2\rangle \equiv \langle\psi|\left(\hat{A} - \langle A\rangle\hat{I}\right)^2|\psi\rangle \tag{2.7.3}$$

$$= \langle\psi|\hat{A}^2|\psi\rangle - \langle\psi|\hat{A}|\psi\rangle^2. \tag{2.7.4}$$

This variance really measures how much the outcome of a given measurement is different from its average, exactly following the definition of variance in statistics. For example, if

you take the case in which $|\psi\rangle$ is an eigenstate of $\hat{A}$ (say $|A_i\rangle$), it easy to see that

$$\begin{aligned} \langle \Delta A^2 \rangle &\equiv \langle A_i|\hat{A}^2|A_i\rangle - \langle A_i|\hat{A}|A_i\rangle^2 && (2.7.5)\\ &= A_i^2 - (A_i)^2 && (2.7.6)\\ &= 0, && (2.7.7) \end{aligned}$$

thus we recover the fundamental measurement postulate, telling us that if we repeatedly measure an eigenstate, we always find the same result (we have zero variance).

The uncertainty principle is an important result connecting the amount of intrinsic uncertainty (variance) associated with the measurement of two observables. It states that, for two observables $\hat{A}$ and $\hat{B}$, we have :

$$\boxed{\langle \Delta A^2 \rangle \langle \Delta B^2 \rangle \geq \frac{1}{4} \left| \langle [\hat{A}, \hat{B}] \rangle \right|^2}. \qquad (2.7.8)$$

Before proving it, let us discuss the consequences of this inequality. There are two cases:

1. The two observables commute, thus $[\hat{A}, \hat{B}] = 0$. In this case, then there is no intrinsic limit on the *precision* we can attain when measuring the two observables on the same state. $\langle \Delta A^2 \rangle$ and $\langle \Delta B^2 \rangle$ can be as small as we want.

2. The two observables do not commute, thus $[\hat{A}, \hat{B}] \neq 0$. In this case, there is an intrinsic limit on the precision we can attain when measuring the two observables.

In order to prove Eq. (2.7.8), we first need two intermediate results.

Theorem. *The Cauchy-Schwarz inequality*

$$\langle \alpha|\alpha\rangle\langle \beta|\beta\rangle \geq |\langle \alpha|\beta\rangle|^2, \qquad (2.7.9)$$

which is a generalization of the triangle inequality to other metric spaces with a given inner product.

Proof. The inequality can be proven in a variety of ways. Here we consider the ket

$$|C_\lambda\rangle = |\alpha\rangle + \lambda|\beta\rangle, \qquad (2.7.10)$$

obtained as a linear combination, with complex λ, of the two given kets. The norm of this ket is obviously positive, thus

$$\begin{aligned} \langle C_\lambda|C_\lambda\rangle &= (\langle \alpha| + \langle \beta|\lambda^\star)\,(|\alpha\rangle + \lambda|\beta\rangle) && (2.7.11)\\ &= \langle \alpha|\alpha\rangle + \langle \beta|\alpha\rangle\lambda^\star + \langle \alpha|\beta\rangle\lambda + |\lambda|^2\langle \beta|\beta\rangle && (2.7.12)\\ &\geq 0. && (2.7.13) \end{aligned}$$

This inequality holds for all values of λ, and the Cauchy-Schwarz inequality is found considering $\lambda = -\langle \beta|\alpha\rangle/\langle \beta|\beta\rangle$, since we have

$$\langle \alpha|\alpha\rangle - |\langle \alpha|\beta\rangle|^2/\langle \beta|\beta\rangle \geq 0, \qquad (2.7.14)$$

which proves the original inequality (notice that the case in which $\langle \beta|\beta\rangle = 0$ can be easily proven separately). □

Using the Cauchy-Schwarz inequality, with

$$\begin{aligned} |\alpha\rangle &= \widehat{\Delta A}|\psi\rangle && (2.7.15)\\ |\beta\rangle &= \widehat{\Delta B}|\psi\rangle, && (2.7.16) \end{aligned}$$

we get

$$\langle \Delta A^2 \rangle \langle \Delta B^2 \rangle \geq |\langle \psi|\widehat{\Delta B}\widehat{\Delta A}|\psi\rangle|^2, \qquad (2.7.17)$$

where we have used the fact that the displacement operators are Hermitian. We now evaluate the r.h.s noticing that for two arbitrary hermitian operators we have

$$\hat{O}_1\hat{O}_2 = \frac{1}{2}[\hat{O}_1,\hat{O}_2] + \frac{1}{2}\left\{\hat{O}_1,\hat{O}_2\right\}. \tag{2.7.18}$$

We also notice that

$$\begin{aligned}
\mathrm{Re}\left(\langle[\hat{O}_1,\hat{O}_2]\rangle\right) &= \frac{1}{2}\left(\langle[\hat{O}_1,\hat{O}_2]\rangle + \langle[\hat{O}_1,\hat{O}_2]\rangle^\star\right) && (2.7.19)\\
&= \frac{1}{2}\left(\langle\hat{O}_1\hat{O}_2\rangle - \langle\hat{O}_2\hat{O}_1\rangle + \langle\hat{O}_1\hat{O}_2\rangle^\star - \langle\hat{O}_2\hat{O}_1\rangle^\star\right) && (2.7.20)\\
&= 0, && (2.7.21)
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Im}\left(\langle\{\hat{O}_1,\hat{O}_2\}\rangle\right) &= \frac{1}{2i}\left(\langle\{\hat{O}_1,\hat{O}_2\}\rangle - \langle\{\hat{O}_1,\hat{O}_2\}\rangle^\star\right) && (2.7.22)\\
&= \frac{1}{2i}\left(\langle\hat{O}_1\hat{O}_2\rangle + \langle\hat{O}_2\hat{O}_1\rangle - \langle\hat{O}_1\hat{O}_2\rangle^\star - \langle\hat{O}_2\hat{O}_1\rangle^\star\right) && (2.7.23)\\
&= 0, && (2.7.24)
\end{aligned}$$

thus

$$\begin{aligned}
|\langle\psi|\widehat{\Delta B}\widehat{\Delta A}|\psi\rangle|^2 &= \frac{1}{4}\left\{\left|\langle[\widehat{\Delta A},\widehat{\Delta B}]\rangle\right|^2 + \left|\langle\{\widehat{\Delta A},\widehat{\Delta B}\}\rangle\right|^2\right\} && (2.7.25)\\
&= \frac{1}{4}\left\{\left|\langle[\widehat{A},\widehat{B}]\rangle\right|^2 + \left|\langle\{\widehat{\Delta A},\widehat{\Delta B}\}\rangle\right|^2\right\}, && (2.7.26)
\end{aligned}$$

thus omitting the second term, we get the uncertainty inequality.

2.8 Change of Basis

Non-commuting operators define a set of distinct eigen-kets and eigenvalues that can be independently used to describe the same physical system. For example, consider two operators $\hat{A}$ and $\hat{B}$ with eigenvalues a_i,b_i and eigen-kets $|A_i\rangle$,$|B_i\rangle$. We therefore have the usual eigenvalue relations:

$$\begin{aligned}
\hat{A}|A_i\rangle &= a_i|A_i\rangle && (2.8.1)\\
\hat{B}|B_i\rangle &= b_i|B_i\rangle, && (2.8.2)
\end{aligned}$$

with the orthonormality condition

$$\begin{aligned}
\langle A_i|A_j\rangle &= \delta_{ij} && (2.8.3)\\
\langle B_i|B_j\rangle &= \delta_{ij}. && (2.8.4)
\end{aligned}$$

Since they both form a complete basis for our vector space, it means that an arbitrary ket state can be written either as

$$|\psi\rangle = \sum_j c_j|A_j\rangle, \tag{2.8.5}$$

or

$$|\psi\rangle = \sum_j d_j|B_j\rangle. \tag{2.8.6}$$

Notice that the physical state here is exactly the same, what changes is just *how* we are mathematically representing it. In quantum physics it is very often necessary to relate different representations, we thus need to have a way to go from one basis to the other.

This is again something that is standard in linear algebra, however it is worth recalling it here using the bra-ket formalism.

The main result is that there is a unitary operator $\hat{U}$ that connects the two representations. More specifically,

$$|B_i\rangle = \hat{U}|A_i\rangle, \tag{2.8.7}$$

and unitarity here means that

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} \tag{2.8.8}$$
$$= \hat{1}. \tag{2.8.9}$$

The operator $\hat{U}$ in bra-ket notation takes a very elegant form

$$\hat{U} = \sum_j |B_j\rangle\langle A_j|, \tag{2.8.10}$$

which can be verified computing the explicit action of this operator on both the eigenstates of $\hat{A}$ and $\hat{B}$. For example, we have:

$$\hat{U}|A_i\rangle = \sum_j |B_j\rangle \underbrace{\langle A_j|A_i\rangle}_{\delta_{ij}} \tag{2.8.11}$$
$$= |B_i\rangle, \tag{2.8.12}$$

and the inverse transformation is found using the conjugate operator

$$\hat{U}^\dagger|B_i\rangle = \sum_j |A_j\rangle \underbrace{\langle B_j|B_i\rangle}_{\delta_{ij}} \tag{2.8.13}$$
$$= A_i, \tag{2.8.14}$$

where in both cases we have used the orthonormality conditions, Eqs. (2.8.3) and (2.8.4).

It is also straightforward to verify that the operator is unitary, using the completeness relations for $\hat{B}$:

$$\hat{U}\hat{U}^\dagger = \left(\sum_j |B_j\rangle\langle A_j|\right)\left(\sum_k |A_k\rangle\langle B_k|\right) \tag{2.8.15}$$
$$= \sum_{jk} |B_j\rangle \underbrace{\langle A_j|A_k\rangle}_{\delta_{jk}}\langle B_k| \tag{2.8.16}$$
$$= \sum_j |B_j\rangle\langle B_j| \tag{2.8.17}$$
$$= \hat{1}, \tag{2.8.18}$$

and $\hat{A}$:

$$\hat{U}^\dagger\hat{U} = \left(\sum_j |A_j\rangle\langle B_j|\right)\left(\sum_k |B_k\rangle\langle A_k|\right) \tag{2.8.19}$$
$$= \sum_{jk} |A_j\rangle \underbrace{\langle B_j|B_k\rangle}_{\delta_{jk}}\langle A_k| \tag{2.8.20}$$
$$= \sum_j |A_j\rangle\langle A_j| \tag{2.8.21}$$
$$= \hat{1}. \tag{2.8.22}$$

In the basis of $\hat{A}$ we can also work out the matrix elements of the operator $\hat{U}$

$$\begin{aligned} \langle A_i|\hat{U}|A_j\rangle &= \langle A_i|\left(\sum_k |B_k\rangle\langle A_k|\right)|A_j\rangle & (2.8.23)\\ &= \langle A_i|B_j\rangle, & (2.8.24)\end{aligned}$$

there the matrix elements are just the amplitudes, or scalar product, between the two sets of eigenstates. Similarly, the matrix elements of $\hat{U}^\dagger$ in the $\hat{B}$ basis are

$$\begin{aligned} \langle B_i|\hat{U}^\dagger|B_j\rangle &= \langle B_i|\left(\sum_k |A_k\rangle\langle B_k|\right)|B_j\rangle & (2.8.25)\\ &= \langle B_i|A_j\rangle, & (2.8.26)\end{aligned}$$

thus

$$\langle A_i|\hat{U}|A_j\rangle = \left(\langle B_j|\hat{U}^\dagger|B_i\rangle\right)^\star. \quad (2.8.27)$$

2.8.1 Transforming States

Given an explicit form for the transformation matrix, $\hat{U}$, we are therefore in position to solve the problem of finding the coefficients (amplitudes) of a given state in a certain basis (say, $|B_i\rangle$), once its coefficients in another basis, say $|A_i\rangle$ are known. We have

$$\begin{aligned} d_j &= \langle B_j|\psi\rangle & (2.8.28)\\ &= \langle B_j|\left(\sum_k c_k|A_k\rangle\right) & (2.8.29)\\ &= \sum_k \langle B_j|A_k\rangle c_k & (2.8.30)\\ &= \sum_k \langle B_j|\hat{U}^\dagger|B_k\rangle c_k & (2.8.31)\end{aligned}$$

2.8.2 Transforming Operators

Similar rules can be derived to transform matrix elements of operators, when passing from one representation to another. For example, we can derive rules to obtain matrix elements of a certain operator in two different bases:

$$\langle A_i|\hat{O}|A_j\rangle \rightarrow \langle B_i|\hat{O}|B_j\rangle.$$

This is

$$\begin{aligned} \langle B_i|\hat{O}|B_j\rangle &= \langle B_i|\underbrace{\sum_l |A_l\rangle\langle A_l|}_{\hat{I}}\hat{O}\underbrace{\sum_k |A_k\rangle\langle A_k|}_{\hat{I}}B_j\rangle & (2.8.32)\\ &= \sum_{lk}\langle B_i|A_l\rangle\langle A_l|\hat{O}|A_k\rangle\langle A_k|B_j\rangle & (2.8.33)\\ &= \sum_{lk}\langle A_i|\hat{U}^\dagger|A_l\rangle\langle A_l|\hat{O}|A_k\rangle\langle A_k|\hat{U}|A_j\rangle, & (2.8.34)\end{aligned}$$

which can be written also as a matrix multiplication

$$\hat{O}_{(B)} = \hat{U}^\dagger_{(A)}\hat{O}_{(A)}\hat{U}_{(A)}. \quad (2.8.35)$$

2.8.3 What's changing? States or operators?

We have seen that the unitary matrix corresponding to the rotation operator $\hat{U}$ can be used either to transform *states* or to transform *operators*. The two point of views go as follows:

1. On one side, there is only one observable ($\hat{A}$) and measurements of some other observable (say, $\hat{B}$) are obtained *rotating* the state vector to the corresponding new references basis. In this case then one first prepares a state $|\psi_B\rangle = \hat{U}|\psi\rangle$, and then measures $\hat{A}$ on the new state, such that

$$\langle B\rangle \quad = \quad \langle\psi_B|\hat{A}|\psi_B\rangle. \tag{2.8.36}$$

2. On the other side, there is only one physical state ($|\psi\rangle$) and measurements of different observables, for example of $\hat{B}$, are found applying distinct measurement operators on the reference state. Thus we have $\hat{B} = \hat{U}^\dagger\hat{A}\hat{U}$, and

$$\langle B\rangle \quad = \quad \langle\psi|\hat{B}|\psi\rangle. \tag{2.8.37}$$

The two formulations are completely equivalent, and it is often the application that tells us which way of thinking makes solving a certain problem easier. When studying time evolution, we will see more prominently what the differences between these two formulations can give us in terms of intuition on the physical systems. In that context, we will identify the first viewpoint as Schroedinger's view on quantum physics, and the second one as Heisenberg's viewpoint.

2.8.4 Example: measurements in different spin bases

As an example, we can compute the matrix elements of the transformation matrix connecting the $\hat{S}_z$ basis to the $\hat{S}_x$ basis. The transformation operator is

$$\hat{\mathcal{H}} \quad = \quad \sum_{i=\pm} |S_x;i\rangle\langle S_z;i|, \tag{2.8.38}$$

and recalling that $|S_x;\pm\rangle = \frac{1}{\sqrt{2}}|S_z;+\rangle \pm \frac{1}{\sqrt{2}}|S_z;-\rangle$, we have that the transformation matrix has the following matrix elements in the $\hat{S}_z$ basis:

$$\hat{\mathcal{H}} \quad = \quad \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

This transformation matrix is of fundamental importance, for example, in quantum computing and it is also known there as Hadamard gate. It has the property $\hat{U}\hat{U}^\dagger = \hat{1}$, as we expect from the general properties of the transformation matrices.

Measuring the $\hat{S}_x$ operator can be done then in one of two mathematically equivalent ways:

1. There is only one spin operator ($\hat{S}_z$) and measurements of the x component of the spin are found rotating the state vector to the corresponding new references basis [1]. In this case then one prepares a state $|\psi_x\rangle = \hat{\mathcal{U}}|\psi\rangle$, and then measures $\hat{S}_z$ on the new state, such that

$$\langle S_x\rangle \quad = \quad \langle\psi_x|\hat{S}_z|\psi_x\rangle.$$

2. There is only one state ($|\psi\rangle$) and measurements of different observables, for example of $\hat{S}_x$, are found applying distinct measurement operators on the reference state. Thus we have $\hat{S}_x = \hat{\mathcal{H}}^\dagger\hat{S}_z\hat{\mathcal{H}}$, as it can be easily checked, and

$$\langle S_x\rangle \quad = \quad \langle\psi|\hat{S}_x|\psi\rangle.$$

[1] Notice that rotating to the x basis, which by definition is a rotation in an *abstract vector space*, in this special case corresponds to an actual *physical* rotation in real space of the SG device. This is a quite special coincidence that happens almost exclusively for spin observables.

Chapter 3

Time Evolution

In the previous Chapters we have introduced the general formalism of quantum mechanics, and concentrated on the description of static quantum phenomena. Starting with this Chapter, we will introduce time dependence to quantum states and one of the fundamental tenets of non-relativistic quantum theory: SchrÃ¶dinger's equation.

3.1 Transforming Quantum States in Time

In the second Chapter we have introduced the first three Axioms of quantum mechanics, essentially concerning how a quantum system is represented as a state vector, how it is manipulated through operators, and how the measurement process takes place.

How do quantum states evolve in time though? We can formalize this question saying that we want to determine an operator $\hat{U}(t_0, t_1)$ that takes a given state at an initial time t_0: $|\Psi(t_0)\rangle$ and transforms it into

$$|\Psi(t_1)\rangle = \hat{U}(t_0, t_1)|\Psi(t_0)\rangle. \tag{3.1.1}$$

Notice that here the notation $\hat{U}(t_0, t_1)$ means that, in general, we expect the operator form to depend parametrically on both the initial (t_0) and the final time (t_1).

3.1.1 Conditions on $\hat{U}$

We can already derive several interesting properties of the operator $\hat{U}$, just requiring some fundamental properties related to time evolution. The operator $\hat{U}$ transform physical states into physical states, thus a first property we should expect is that it should preserve the normalization of $|\Psi\rangle$, since Axiom 2 requires that we have a consistent probabilistic interpretation of state vectors at each instant of time. This translates into the requirement:

$$\langle\Psi(t_1)|\Psi(t_1)\rangle = \langle\Psi(t_0)|\hat{U}^\dagger(t_0, t_1)\hat{U}(t_0, t_1)|\Psi(t_0)\rangle \tag{3.1.2}$$

$$= \langle\Psi(t_0)|\Psi(t_0)\rangle. \tag{3.1.3}$$

The second property we expect from this operator, is that it can be arbitrarily *composed*, in the sense that successive time evolutions of $t_0 \to t_1 \to t_2 \to t_3, \ldots$ must be equivalent to a single time evolution of the total time. For example we must have:

$$\hat{U}(t_2, t_3)\hat{U}(t_1, t_2)\hat{U}(t_0, t_1) = \hat{U}(t_0, t_3), \tag{3.1.4}$$

where it is crucial to notice that early times appear in the right of the expression above, since:

$$|\Psi(t_3)\rangle = \hat{U}(t_2, t_3)|\Psi(t_2)\rangle \tag{3.1.5}$$

$$= \hat{U}(t_2, t_3)\left[\hat{U}(t_1, t_2)|\Psi(t_1)\rangle\right] \tag{3.1.6}$$

$$= \hat{U}(t_2, t_3)\hat{U}(t_1, t_2)\left[\hat{U}(t_0, t_1)|\Psi(t_0)\rangle\right]. \tag{3.1.7}$$

Furthermore, if we go *back* in time , this operation should be equivalent to applying the inverse transformation:

$$\hat{U}(t_0, t_1) = \hat{U}^{-1}(t_1, t_0), \tag{3.1.8}$$

where $\hat{U}^{-1}$ denotes the inverse of the operator.

The last property that we can intuitively expect is that in the limit of small time evolution the operator $\hat{U}$ should strictly reduce to the identity,

$$\lim_{|t_1| \to t_0} \hat{U}(t_0, t_1) = \hat{1}, \tag{3.1.9}$$

since the state is unchanged when not time evolved.

3.1.2 Time-Invariant Case

In order to determine a concrete form for the time evolution operator, we start with a simpler case, in which we imagine the system to be completely isolated from the environment, and make the assumption that the time evolution depends only on time *differences* and not on the absolute values of times. In other words, assuming for example that we consider equispaced time intervals $t_0 \to t_0 + \Delta_t \to t_0 + 2\Delta_t \ldots$, we have

$$|\Psi(t_0 + \Delta_t)\rangle = \hat{U}(\Delta_t)|\Psi(t_0)\rangle, \tag{3.1.10}$$
$$|\Psi(t_0 + 2\Delta_t)\rangle = \hat{U}(\Delta_t)|\Psi(t_0 + \Delta_t)\rangle, \tag{3.1.11}$$

where it should be noticed that we are using the same operator to evolve the state of an interval Δ_t, regardless of the initial time. This assumption is, in practice, very well verified in the great majority of quantum systems. Later in this Chapter we will discuss how to go beyond this.

For the moment, let us show that all the conditions previously described are satisfied if we take the time evolution operator to be described by the following unitary operator

$$\hat{U}(\Delta_t) = e^{-i\hat{\Omega}\Delta_t}, \tag{3.1.12}$$

where $\hat{\Omega}$ is an Hermitian operator. The exponential of the operator indeed has essentially the same meaning it would have for regular numbers, and it is understood in terms of its Taylor expansion:

$$e^{-i\hat{\Omega}\Delta_t} = \hat{1} - i\hat{\Omega}\Delta_t - \frac{1}{2}\hat{\Omega}^2\Delta_t^2 + \mathcal{O}\left[(\Delta_t)^3\right], \tag{3.1.13}$$

thus we can also clearly see that $\hat{\Omega}$ carries the units of a frequency. From the Taylor expansion (3.1.13), we immediately see that (3.1.9) is verified. The unitarity assumption is also quick to verify, since it is an elementary property of the exponential of an operator that $e^{X^+} = \left(e^X\right)^\dagger$, thus

$$\hat{U}^\dagger(\Delta_t) = e^{i(\hat{\Omega})^\dagger \Delta_t} \tag{3.1.14}$$
$$= e^{i\hat{\Omega}\Delta_t}, \tag{3.1.15}$$

where in the last line we have used the fact that $\hat{\Omega}$ is Hermitian and

$$\hat{U}^\dagger(\Delta_t)\hat{U}(\Delta_t) = e^{i\hat{\Omega}\Delta_t}e^{-i\hat{\Omega}\Delta_t} \tag{3.1.16}$$
$$= \hat{1}. \tag{3.1.17}$$

The composition property is also a natural consequence of the exponential structure

$$\hat{U}(\Delta_1 + \Delta_2) = \hat{U}(\Delta_2)\hat{U}(\Delta_1) \tag{3.1.18}$$
$$e^{-i\hat{\Omega}(\Delta_1+\Delta_2)} = e^{-i\hat{\Omega}\Delta_2}e^{-i\hat{\Omega}\Delta_1} \tag{3.1.19}$$

as well as the time inversion property noticing that

$$\hat{U}(-\Delta_t) = e^{i\hat{\Omega}\Delta_t}, \tag{3.1.20}$$

thus we recover the time inversion property as a direct consequence of unitarity:

$$e^{i\hat{\Omega}\Delta_t}e^{-i\hat{\Omega}\Delta_t} = \hat{1}, \tag{3.1.21}$$

$$\hat{U}(-\Delta_t) = \hat{U}^{-1}(\Delta_t). \tag{3.1.22}$$

The exponential form therefore satisfies all the requirements, and it is actually part of a broader mathematical result known as Stone's theorem.

3.2 The SchrÃ¶dinger equation

In the previous Section we have used very general arguments to determine that the time evolution operator should take the form:

$$\hat{U}(\Delta_t) = e^{-i\hat{\Omega}\Delta_t}. \tag{3.2.1}$$

The operator $\hat{\Omega}$, which is an Hermitian operator, is what we can immediately identify as the *generator of the time evolution*. In classical mechanics, the generator of the time evolution is the Hamiltonian of the system. By analogy, it is therefore natural to identify $\hat{\Omega}$ also as the Hamiltonian *operator* of the quantum system. $\hat{\Omega}$, as previously introduced, however has the dimensions of a frequency, whereas in classical mechanics the Hamiltonian has the units of an energy. We thus *define* the actual Hamiltonian operator of the system with the correct units:

$$\hat{H} = \hbar\hat{\Omega}, \tag{3.2.2}$$

through Planck's reduced constant, $\hbar$. This units rescaling, strictly speaking, is not of fundamental importance for the development of the theory at this stage, however we will show later on that this definition allows to consistently recover classical physics at the macroscopic scale. We stress however that the Hamiltonian operator concept is a broader concept than what found in classical mechanics, and that is is often the case that quantum Hamiltonians do not have a direct classical counterpart. In this sense, the analogy should always be taken with a grain of salt, and strictly speaking, only the definition through the operator $\hat{U}$ is always correct. From the previous reasonings and definitions, we thus have that a quantum state evolves according to:

$$|\Psi(t+\Delta_t)\rangle = e^{-\frac{i}{\hbar}\hat{H}\Delta_t}|\Psi(t)\rangle. \tag{3.2.3}$$

This equation is, in essence, all we need to perform the time evolution of the system. However, we can derive a more famous equation when considering an infinitesimal time step δ_t. Using the Taylor series for the exponential, we get

$$|\Psi(t+\delta_t)\rangle = |\Psi(t)\rangle - i\delta_t\frac{\hat{H}}{\hbar}|\Psi(t)\rangle + \mathcal{O}(\delta_t^2), \tag{3.2.4}$$

and slightly rearranging the terms we get:

$$i\hbar\frac{|\Psi(t+\delta_t)\rangle - |\Psi(t)\rangle}{\delta_t} = \hat{H}|\Psi(t)\rangle. \tag{3.2.5}$$

By taking the limit $\delta_t \to 0$ in the previous equation, this leads to the famous SchrÃ¶dinger equation for quantum states:

$$\boxed{i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle} \tag{3.2.6}$$

This equation fully specifies how a quantum state evolves in time, once an initial condition $|\Psi\rangle$ is given and the Hamiltonian operator is established.

3.3 Energy Eigenstates

So far, we have determined a very important differential equations governing the time evolution of a quantum state, Eq. (3.2.6). We now analyze more in detail the connection between the spectrum of $\hat{H}$ and the dynamics of the quantum system.

A first observation is that we are considering the case in which $\hat{H}$ is time independent, as a direct consequence of the time-invariance assumption we have previously made (we will see later that when this assumption is violated the Hamiltonian can be thought of changing in time). For the moment, for time-independent $\hat{H}$ we can always define time-independent eigen-kets

$$\hat{H}|E_k\rangle = E_k|E_k\rangle, \tag{3.3.1}$$

where for the sake of simplicity we are assuming there that they are labeled by an integer index k. In this case, we can always expand the time-dependent state in terms of the energy eigen-kets:

$$|\Psi(t)\rangle = \sum_k |E_k\rangle\langle E_k|\Psi(t)\rangle \tag{3.3.2}$$

$$= \sum_k |E_k\rangle c_k(t). \tag{3.3.3}$$

The action of Schrodinger's equation in this representation is then particularly easy to visualize:

$$i\hbar\frac{\partial}{\partial t}\sum_k |E_k\rangle c_k(t) = \sum_k \hat{H}|E_k\rangle c_k(t), \tag{3.3.4}$$

$$i\hbar\sum_k |E_k\rangle\frac{\partial}{\partial t}c_k(t) = \sum_k E_k|E_k\rangle c_k(t). \tag{3.3.5}$$

Further multiplying on the left by $\langle E_{k'}|$ and using the orthonormality condition we get that each coefficient c_k satisfies

$$i\hbar\partial_t c_k(t) = E_k c_k(t), \tag{3.3.6}$$

which has the solution:

$$c_k(t) = e^{-iE_k/\hbar t}c_k(0). \tag{3.3.7}$$

We therefore already see the importance of energy eigenstates: if we know the expansion coefficients of the initial state in terms of the energy eigen-kets, then we can determine at all times the behavior of the quantum state. Notice that in this basis, dynamics is just a simple phase multiplication, thus it is also clear that by construction $\sum_k |c_k(t)|^2 = \sum_k |c_k(0)|^2$, the norm conservation condition. It is also instructive to derive the same result just using the form of the exponential operator $\hat{U}(t) = e^{-i\hat{H}t/\hbar}$, and noticing that it acts trivially on the eigenstates of the Hamiltonian

$$\hat{U}(t)|\Psi\rangle = e^{-i\hat{H}t/\hbar}|\Psi\rangle \tag{3.3.8}$$

$$= e^{-i\hat{H}t/\hbar}\sum_k c_k(0)|E_k\rangle \tag{3.3.9}$$

$$= \sum_k e^{-iE_k t/\hbar}c_k(0)|E_k\rangle \tag{3.3.10}$$

$$= \sum_k c_k(t)|E_k\rangle. \tag{3.3.11}$$

3.4 Time-Dependence of observables

Having established the fundamental equation governing the dynamics of a given quantum state, Eq. (3.2.6), we can now also determine how observables behave. To this purpose, we can compute the expectation value of some observable $\hat{A}$ at time t using the standard rules for expectation values:

$$\langle \hat{A} \rangle(t) \quad = \quad \langle \Psi(t) | \hat{A} | \Psi(t) \rangle, \tag{3.4.1}$$

we thus have

$$\langle \Psi(t) | \hat{A} | \Psi(t) \rangle \quad = \quad \langle \Psi(0) | e^{i\frac{\hat{H}}{\hbar}t} \hat{A} e^{-i\frac{\hat{H}}{\hbar}t} | \Psi(0) \rangle. \tag{3.4.2}$$

In order to explicitly compute this expectation value, it is once more convenient to consider the eigenstates of the Hamiltonian, and insert twice a completeness relation:

$$\begin{aligned} \langle \Psi(0) | e^{i\frac{\hat{H}}{\hbar}t} \hat{A} e^{-i\frac{\hat{H}}{\hbar}t} | \Psi(0) \rangle \quad &= \quad \sum_{k,k'} \langle \Psi(0) | e^{i\frac{\hat{H}}{\hbar}t} | E_k \rangle \langle E_k | \hat{A} | E_{k'} \rangle \langle E_{k'} | e^{-i\frac{\hat{H}}{\hbar}t} | \Psi(0) \rangle \\ &= \quad \sum_{k,k'} e^{i\frac{E_k - E_{k'}}{\hbar}t} c_k(0)^\star A(k,k') c_{k'}(0), \end{aligned} \tag{3.4.3}$$

where we have introduced the matrix elements of the operator $\hat{A}$ in the energy basis: $A(k,k') \equiv \langle E_k | \hat{A} | E_{k'} \rangle$, as well as the expansion coefficients for the initial state at time $t = 0$, $c_k(0) \equiv \langle E_k | \Psi(0) \rangle$. This expression basically tells us that the expectation value of an arbitrary operator in general is found as the summation of characteristic oscillations in time, with frequency depending solely on the energy differences

$$\omega_{k,k'} \quad = \quad \frac{E_k - E_{k'}}{\hbar}. \tag{3.4.4}$$

We will see soon that these kind of oscillations, first predicted by Niels Bohr, can be experimentally observed, and, among other things, they allow for the most precise measurements of $\hbar$ available so far.

3.4.1 Special case: conserved quantities

While, in general, expectation values of observables will oscillate in time according to the expression (3.4.3), there is however a very important case in which measuring the same quantity at later times will yield the same average results. In other words, we can find specific observables for which $\langle \hat{A} \rangle(t) = \langle \hat{A} \rangle(0)$ at all times. These are called *conserved* quantities, since their expectation value is conserved as a function time.

In general, a conserved quantity is associated to an operator $\hat{A}_c$ that commutes with the Hamiltonian, i.e.

$$[\hat{A}_c, \hat{H}] \quad = \quad 0, \tag{3.4.5}$$

thus implying, as we have seen before, that the eigenstates of the Hamiltonian are also eigenstates of $\hat{A}_c$. In turn, this means that the matrix elements of the operator in the energy eigen-basis greatly simplify:

$$\langle E_k | \hat{A}_c | E_{k'} \rangle \quad = \quad \delta_{k,k'} a_k, \tag{3.4.6}$$

and that all the Bohr frequencies are vanishing, resulting in

$$\langle \Psi(t) | \hat{A}_c | \Psi(t) \rangle \quad = \quad \sum_k |c_k(0)|^2 a_k, \tag{3.4.7}$$

which in turn is identical to the initial value

$$\begin{aligned} \langle \Psi(0) | \hat{A}_c | \Psi(0) \rangle \quad &= \quad \sum_{k,k'} c_k^\star(0) \langle E_k | \hat{A}_c | E_{k'} \rangle c_{kk'}(0) \\ &= \quad \sum_k |c_k(0)|^2 a_k \\ &= \quad \langle \Psi(t) | \hat{A}_c | \Psi(t) \rangle. \end{aligned} \tag{3.4.8}$$

3.5 Heisenberg picture

We have seen before that the expectation value of a given operator reads

$$\langle\Psi(0)|e^{i\frac{\hat{H}}{\hbar}t}\hat{A}e^{-i\frac{\hat{H}}{\hbar}t}|\Psi(0)\rangle \quad = \quad \langle\Psi(t)|\hat{A}|\Psi(t)\rangle, \tag{3.5.1}$$

this way of looking at expectation values of *static* operators over time-dependent states is principally due to Schrodinger. An alternative view, due to Dirac, but most well known as *Heisenberg picture*, consists in considering time-evolved operators

$$\hat{A}_h(t) \quad \equiv \quad e^{i\frac{\hat{H}}{\hbar}t}\hat{A}e^{-i\frac{\hat{H}}{\hbar}t}, \tag{3.5.2}$$

and *static states*. In this section we will denote Heisenberg operators with the subscript h. Thus, in this picture, all quantum states are static, and operators change in time. This picture is completely equivalent to Schrodinger's interpretation, however it can be sometimes useful to reason in terms of operators rather than states. This is best understood when obtaining the equations of motion for the Heisenberg operators $\hat{A}_h(t)$. Considering once more the case of time-independent Hamiltonians, as well as of observable without an intrinsic time dependence, we have

$$\hat{A}_h(t+\delta_t) \quad = \quad e^{i\frac{\hat{H}}{\hbar}(t+\delta_t)}\hat{A}e^{-i\frac{\hat{H}}{\hbar}(t+\delta_t)} \tag{3.5.3}$$

$$= \quad \hat{A}_h(t) + \delta_t\frac{i}{\hbar}\hat{H}\hat{A}_h - \delta_t\frac{i}{\hbar}\hat{A}_h\hat{H}, \tag{3.5.4}$$

thus

$$\frac{d\hat{A}_h(t)}{dt} \quad = \quad \frac{i}{\hbar}[\hat{H},\hat{A}_h]. \tag{3.5.5}$$

3.5.1 Conserved quantities

We have already seen when analyzing the time dependence of observables in Schroedinger's picture that if an observable commutes with the Hamiltonian then the Bohr frequencies are strictly vanishing, and the observable expectation value is time independent, Eq. (3.4.8). This notion of conservation law being associated with the commutator with the Hamiltonian is particularly transparent from Heisenberg's picture. We see from Eq. (3.5.5) that if $\hat{H}$ commutes with our observable, then the time derivative of the operators in the Heisenberg picture is vanishing, thus the observable is constant in time.

3.6 Example: Spin Precession

After having developed a great deal of the machinery necessary to study quantum dynamics, it is instructive to look at an example. Here we want to consider the case of a spin 1/2 particle (for example, an electron) inside a magnetic field. Classically, the potential energy of a magnetic dipole is

$$U = -\mu \cdot B, \tag{3.6.1}$$

where B is the magnetic field and μ is the magnetic dipole moment. Inspired by this form, and recalling that the Hamiltonian operator is related to measurements of the energy of the system, we *define* the quantum Hamiltonian operator describing the interaction of the spin with the magnetic field in this case to be

$$\hat{H} = -\frac{e}{m}\hat{\vec{S}} \cdot \vec{B}, \tag{3.6.2}$$

and considering the case in which the magnetic field is pointing in the z direction, we have

$$\hat{H} = -\frac{e}{m}\hat{S}_z B_z. \tag{3.6.3}$$

We see that since the Hamiltonian is proportional to $\hat{S}_z$, thus it is obvious that $[\hat{H}, \hat{S}_z] = 0$ and that the eigenstates of $\hat{S}_z$ are also eigenstates of the Hamiltonian, such that

$$\hat{H}|S_z;\pm\rangle = \left(\mp\frac{\hbar}{2}\frac{e}{m}B_z\right)|S_z;\pm\rangle \tag{3.6.4}$$

$$= \left(\pm\frac{\hbar}{2}\omega_B\right)|S_z;\pm\rangle, \tag{3.6.5}$$

where we have introduced the frequency $\omega_B \equiv \frac{|e|B_z}{m}$.

The time evolution then reads

$$\hat{U}(t) = e^{-i\frac{\hat{H}}{\hbar}t} \tag{3.6.6}$$

$$= e^{-i\omega_B t\frac{\hat{S}_z}{\hbar}}. \tag{3.6.7}$$

Thus if at time $t = 0$ the system is an arbitrary state:

$$|\Psi(0)\rangle = c_+|S_z;+\rangle + c_-|S_z;-\rangle, \tag{3.6.8}$$

at time t we have

$$|\Psi(t)\rangle = c_+ e^{-i\frac{\omega_B}{2}t}|S_z;+\rangle + c_- e^{i\frac{\omega_B}{2}t}|S_z;-\rangle, \tag{3.6.9}$$

following directly from Eq. (3.3.7), since the Hamiltonian is diagonal in the eigen-kets of S_z.

3.6.1 Starting from an eigenstate of S_z

Let us now analyze the specific case in which we start our time evolution from an eigenstate of S_z, say $|S_z;+\rangle$, thus $c_+ = 1$ and $c_- = 0$. We see that at time t the state remains always in the same state, since

$$|\Psi(t)\rangle = e^{-i\frac{\omega_B}{2}t}|S_z;+\rangle, \tag{3.6.10}$$

differs from the initial state only for an (irrelevant) global phase. This is a stationary state for the dynamics, as we have discussed before, and also since the operator $\hat{S}_z$ commutes with the Hamiltonian we clearly see that at all times the expectation value of the spin does not change:

$$\langle S_z(t)\rangle = \langle S_z(0)\rangle \tag{3.6.11}$$

$$= +\frac{\hbar}{2}. \tag{3.6.12}$$

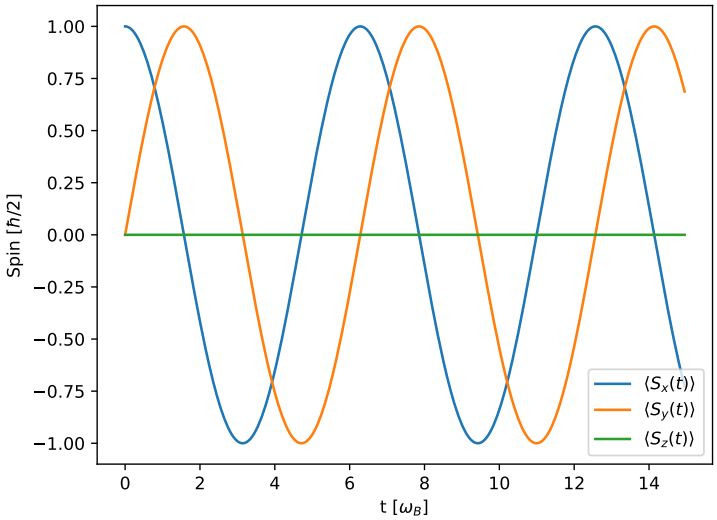

Figure 3.6.1: Values of the expectation values of the three components of the spin. The spin performs a precession in the xy plane, when the initial state is an eigenstate of $\hat{S}_x$ and the magnetic fields points in the z direction.

3.6.2 Starting from an eigenstate of S_x

The situation is more interesting if we start from an eigenstate of an operator that does not commute with the Hamiltonian. For example, we take

$$|\Psi(0)\rangle = |S_x;+\rangle, \tag{3.6.13}$$
$$= \frac{1}{\sqrt{2}}\left[|S_z;+\rangle + |S_z;-\rangle\right], \tag{3.6.14}$$

and we can for example compute the expectation value of S_x at later times:

$$\langle S_x(t)\rangle = \langle\Psi(t)|\hat{S}_x|\Psi(t)\rangle \tag{3.6.15}$$
$$= \frac{1}{2}\left[e^{i\frac{\omega_B}{2}t}\langle S_z;+| + e^{-i\frac{\omega_B}{2}t}\langle S_z;-|\right]\hat{S}_x \times \tag{3.6.16}$$
$$\times\left[e^{-i\frac{\omega_B}{2}t}|S_z;+\rangle + e^{i\frac{\omega_B}{2}t}|S_z;-\rangle\right] \tag{3.6.17}$$
$$= \frac{1}{2}\left[e^{i\omega_B t}\langle S_z;+|\hat{S}_x|S_z;-\rangle + e^{-i\omega_B t}\langle S_z;-|\hat{S}_x|S_z;+\rangle\right] \tag{3.6.18}$$
$$= \frac{\hbar}{4}\left[e^{i\omega_B t} + e^{-i\omega_B t}\right] \tag{3.6.19}$$
$$= \frac{\hbar}{2}\cos\omega_B t, \tag{3.6.20}$$

and similarly one can find

$$\langle S_y(t)\rangle = \frac{\hbar}{2}\sin\omega_B t, \tag{3.6.21}$$
$$\langle S_z(t)\rangle = 0. \tag{3.6.22}$$

We therefore see that the expectation value of the spin *precesses* with a frequency ω_B. This kind of phenomenon is very well established from the experimental point of view, and it is the basis also for very precise measurements of $\hbar$.

3.7 Time-Dependent Hamiltonians

We now generalize our discussion, and consider the case we have been so carefully avoiding until now, namely the case in which the time evolution operator depends not only on the time difference $\Delta_t = t_1 - t_0$ but also on both t_1 and t_0. In this case, it is non longer true that the exponential form we have derived before is valid:

$$\hat{U}(t_0, t_1) \neq e^{-i\hat{H}/\hbar(t_1-t_0)}.$$

This situation typically arises when the system interacts with an external degree of freedom that is explicitly time-dependent, and it is not isolated from the external environment[1]. To see why this is the case, let us consider the case we have considered so far, where we have identified the Hamiltonian operator $\hat{H}$ with the measurement of the energy of the system in exam. We have seen in 3.5.1 that if an operator commutes with the Hamiltonian, then its expectation value is a conserved quantity, i.e. it does not change with time. In the case of the Hamiltonian, it is therefore clear that in all the previous discussions we had

$$\begin{aligned} \langle \text{Energy} \rangle(t) &= \langle \Psi(t)|\hat{H}|\Psi(t)\rangle \\ &= \langle \text{Energy} \rangle(t=0), \end{aligned}$$

since obviously the Hamiltonian commutes with itself. The previous situation then transparently corresponds to the case in which the system is isolated and energy is conserved. To go beyond this assumption, it is customary to consider Hamiltonian operators that explicitly carry a time dependence (i.e., their matrix elements depend on time), for which, in general, we have

$$\begin{aligned} \langle \text{Energy} \rangle(t) &= \langle \Psi(t)|\hat{H}(t)|\Psi(t)\rangle \\ &\neq \langle \text{Energy} \rangle(t=0), \end{aligned}$$

and physically corresponds to a quantum system that is no longer isolated from the external environment and can freely exchange energy with it. In order to find an explicit expression for the time-evolution operator in this more general case, we consider small time increments δ_t, and look at the time evolution in this small interval:

$$t \rightarrow t + \delta_t \tag{3.7.1}$$

Since we are considering a small δ_t, we can make the assumption that within this time interval the Hamiltonian is constant and equal to the value it had on one of the two extremes, for example $\hat{H}(t)$. In this case, the previously determined time evolution still holds, and we have

$$\hat{U}_\delta(t) = e^{-i\hat{H}(t)\delta_t/\hbar}. \tag{3.7.2}$$

Following the same reasoning, we can also conclude that at each time the Schroedinger's equation still holds, provided that we take into account the explicit time dependence of the Hamiltonian, i.e. formally we have to solve:

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = \hat{H}(t)|\Psi(t)\rangle. \tag{3.7.3}$$

3.7.1 Time evolution operator

In order to derive explicit expressions for the full time operator $\hat{U}(t_0, t_1)$, it is very useful to divide the time interval $[t_0, t_1]$ into p smaller time slices of duration δ_t, such that $t_1 = t_0 + p \times \delta_t$. With the idea of splitting the time evolution into small chunks of times where the Hamiltonian is almost constant, and using the composition property of the time evolution

[1] In the previous discussion, the only allowed interaction with the environment was through the measurement process, otherwise the system was supposed to be perfectly isolated when undergoing time evolution.

operator, we can then write the total time evolution operator as a product of the "standard" time evolution operator for the small time steps:

$$\hat{U}(t_0,t_1) = \underbrace{\hat{U}_\delta(t_0+(p-1)\delta_t)\dots\hat{U}_\delta(t_0+\delta_t)\hat{U}_\delta(t_0)}_{p \text{ times}}, \tag{3.7.4}$$

$$\simeq e^{-i\hat{H}(t_0+(p-1)\delta_t)\delta_t/\hbar}\dots e^{-i\hat{H}(t_0+\delta_t)\delta_t/\hbar}e^{-i\hat{H}(t_0)\delta_t/\hbar} \tag{3.7.5}$$

where we have taken the number of time chunks to be $p = (t_1 - t_0)/\delta_t$, and it is to be understood that it is only in the limit $\delta_t \to 0$ that this expression becomes exact. These kind of small-time expansions for the time evolution operators are of fundamental importance in a variety of applications of quantum mechanics, ranging from path integrals to quantum computing.

While the expression (3.7.5) is rather general, there are however some special cases in which it simplifies. Let us consider now two important sub-cases.

3.7.1.1 Hamiltonians at different times commute

In this case, we assume that $[\hat{H}(t_1),\hat{H}(t_2)] = 0\ \forall t_1,t_2$. Recalling that for two generic operators $\hat{A}$ and $\hat{B}$ we have

$$\begin{cases} e^{\hat{A}}e^{\hat{B}} = e^{\hat{A}+\hat{B}} & \text{if } [\hat{A},\hat{B}] = 0 \\ e^{\hat{A}}e^{\hat{B}} \neq e^{\hat{A}+\hat{B}} & \text{if } [\hat{A},\hat{B}] \neq 0 \end{cases}, \tag{3.7.6}$$

and using the fact that the Hamiltonian commutes at different times, we can then put all the terms in Eq. (3.7.5) together in a single exponential and we have that

$$\hat{U}(t_0,t_1) \simeq \exp\left[-\frac{i}{\hbar}\sum_{k=0}^{p-1}\hat{H}(t_0+k\delta_t)\delta_t\right]. \tag{3.7.7}$$

In the limit of vanishing time step, the sum becomes and integral :

$$\hat{U}(t_0,t_1) = \exp\left[-\frac{i}{\hbar}\int_{t_0}^{t_1}dt'\hat{H}(t')\right]. \tag{3.7.8}$$

We will see when considering specific applications, that this expression is particularly useful.

3.7.1.2 Hamiltonians at different times do not commute

Eq. (3.7.5) is very general, therefore it is still a valid solution for the time evolution operator even in this case. It is possible to show (but we will not show the demonstration) that the total error of that approximation is of order δt, thus negligible in the limit in which $\delta_t \to 0$. The important difference however is that we cannot gather together all terms into a single exponential, since in general

$$e^{-\frac{i}{\hbar}\delta_t\hat{H}(t+\delta_t)}e^{-\frac{i}{\hbar}\delta_t\hat{H}(t)} \neq e^{-\frac{i}{\hbar}\delta_t\left(\hat{H}(t+\delta_t)+\hat{H}(t)\right)},$$

given that we are considering the case $[\hat{H}(t),\hat{H}(t')] \neq 0$. The formal solution to the time evolution operator is therefore given by

$$\hat{U}(t_0,t_1) = \lim_{\delta_t\to 0} e^{-i\hat{H}(t_0+(p-1)\delta_t)\delta_t/\hbar}\dots e^{-i\hat{H}(t_0+\delta_t)\delta_t/\hbar}e^{-i\hat{H}(t_0)\delta_t/\hbar}$$

$$= \lim_{p\to\infty}\Pi_{k=0}^{p-1}e^{-\frac{i}{\hbar}\hat{H}\left(t_0+k\frac{t_1-t_0}{p}\right)\frac{t_1-t_0}{p}}. \tag{3.7.9}$$

Alternative (and more complex) expressions for the time evolution operator can be obtained also using the Dyson series expansion or the Magnus expansion. During this course we won't consider these alternatives and, in general, we will not attempt to solve for the time

evolution operator directly when considering non-commuting time-dependent Hamiltonians. Instead, we will typically resort to solving the Schroedinger's equation (3.7.3) for the state $|\Psi(t)\rangle$, which can be recast as a system of linear ordinary differential equations for the expansion coefficients of the state vector :

$$i\hbar\dot{c}_k(t) = \sum_{k'}\langle A_k|H(t)|A_{k'}\rangle c_{k'}(t). \tag{3.7.10}$$

In this expression, it is very important to notice that we are representing our state using the eigen-kets of a *time-independent* operator $\hat{A}$:

$$\hat{A}|A_k\rangle = a_k|A_k\rangle \tag{3.7.11}$$
$$c_k(t) = \langle A_k|\Psi(t)\rangle, \tag{3.7.12}$$

since the eigen-kets of the Hamiltonian are also time dependent, thus cannot be used as a fixed reference set of "coordinates" for our state.

Chapter 4

Magnetic Resonance and Ammonia

In this Chapter we continue analyzing the dynamics of a spin in a magnetic field, extending our analysis to the case of time-dependent field. We also show an example of a complex physical system (the Ammonia molecule) that can be well approximated by two-dimensional state vectors, also called two-level systems.

4.1 Magnetic Resonance

Let us consider the situation in which a particle with spin 1/2 (for example an electron) is subjected to a time-dependent magnetic field.In the previous chapter, we have written the Hamiltonian in the case of a static magnetic field, and the situation is analogous also when $\vec{B}$ acquires a time dependence, with the notable exception that the Hamiltonian becomes time dependent:

$$\hat{H}(t) = -\frac{e}{m}\hat{\vec{S}}\cdot\vec{B}(t). \tag{4.1.1}$$

We imagine now that the magnetic field is such that

$$\vec{B}(t) = \vec{B_0} + \vec{B_1}(t) \tag{4.1.2}$$

$$= B_0\hat{z} + \hat{x}B_1\cos\omega t + \hat{y}B_1\sin\omega t, \tag{4.1.3}$$

thus there is a static magnetic field in the z direction (as in the previous exercise) but also a time-dependent magnetic field in the xy plane oscillating with a frequency ω. The Hamiltonian then reads

$$\hat{H}(t) = \frac{-e}{m}B_0\hat{\vec{S}}\cdot\hat{z} + \frac{-e}{m}B_1(\hat{\vec{S}}\cdot\hat{x} + \hat{\vec{S}}\cdot\hat{y}) \tag{4.1.4}$$

$$= \omega_0\hat{S}_z + \omega_1[\cos(\omega t)\hat{S}_x + \sin(\omega t)\hat{S}_y], \tag{4.1.5}$$

where we have defined

$$\omega_0 \equiv \frac{|e|\,B_0}{m}, \tag{4.1.6}$$

$$\omega_1 \equiv \frac{|e|\,B_1}{m}. \tag{4.1.7}$$

Using the expression for the Pauli operators, we can write the explicit matrix elements of the Hamiltonian in the basis of the eigen-kets of $\hat{S}_z$:

$$\hat{H}(t) \doteq \frac{\hbar}{2}\omega_0\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} + \frac{\hbar}{2}\omega_1\left[\cos(\omega t)\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} + i\sin(\omega t)\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\right] \tag{4.1.8}$$

$$= \frac{\hbar}{2}\begin{pmatrix}\omega_0 & \omega_1(\cos(\omega t) - i\sin(\omega t))\\ \omega_1(\cos(\omega t) + i\sin(\omega t)) & -\omega_0\end{pmatrix} \tag{4.1.9}$$

$$= \frac{\hbar}{2}\begin{pmatrix}\omega_0 & \omega_1 e^{-i\omega t}\\ \omega_1 e^{i\omega t} & -\omega_0\end{pmatrix}. \tag{4.1.10}$$

We immediately notice that in order to solve for the dynamics of the system we cannot use the solution of the time evolution operator we have found for the time-independent Hamiltonians case. Moreover, it is easy to prove (left as an exercise) that

$$[\hat{H}(t_1), \hat{H}(t_2)] \neq 0, \tag{4.1.11}$$

for $t_1 \neq t_2$. Since we cannot easily solve the time-dependence for quantum states using the unitary evolution operator $\hat{U}$, we resort to the last method outlined in the previous Chapter, namely directly attacking the Schroedinger's equation. We represent the state at time t as

$$|\Psi(t)\rangle = c_+(t)|S_z;+\rangle + c_-(t)|S_z;-\rangle, \tag{4.1.12}$$

and, following Eq. (3.7.10), we write the Schroedinger's equation in matrix form as

$$i\hbar \begin{pmatrix} \dot{c}_+(t) \\ \dot{c}_-(t) \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} \omega_0 & \omega_1 e^{-i\omega t} \\ \omega_1 e^{i\omega t} & -\omega_0 \end{pmatrix} \begin{pmatrix} c_+(t) \\ c_-(t) \end{pmatrix}, \tag{4.1.13}$$

thus the following system of coupled differential equations:

$$\begin{cases} i\partial_t c_+(t) = \frac{\omega_0}{2} c_+(t) + \frac{\omega_1}{2} e^{-i\omega t} c_-(t) \\ i\partial_t c_-(t) = \frac{\omega_1}{2} e^{i\omega t} c_+(t) - \frac{\omega_0}{2} c_-(t) \end{cases}. \tag{4.1.14}$$

4.1.1 Solving the Schroedinger Equation

Solving the differential equations (3.4.3) can be easily done numerically, however we focus here on finding an explicit analytic solution. For this purpose, we consider a change of variables:

$$\begin{aligned} a_+(t) &= e^{i\omega t/2} c_+(t) \\ a_-(t) &= e^{-i\omega t/2} c_-(t), \end{aligned} \tag{4.1.15}$$

and substituting these into the original differential equations we obtain for the first equation:

$$ie^{-i\omega t/2}\left[\partial_t a_+(t) - i\frac{\omega}{2} a_+(t)\right] = \frac{\omega_0}{2} e^{-i\omega t/2} a_+(t) + \frac{\omega_1}{2} e^{-i\omega t/2} a_-(t) \tag{4.1.16}$$

$$i\partial_t a_+(t) = \frac{\omega_0 - \omega}{2} a_+(t) + \frac{\omega_1}{2} a_-(t), \tag{4.1.17}$$

thus we see that the transformation we have done has removed the time dependence on the coefficients. Similarly, the same things happen for the other equation, and we get:

$$\begin{cases} i\partial_t a_+(t) = \frac{\omega_0 - \omega}{2} a_+(t) + \frac{\omega_1}{2} a_-(t), \\ i\partial_t a_-(t) = \frac{\omega_1}{2} a_+(t) - \frac{\omega_0 - \omega}{2} a_-(t). \end{cases} \tag{4.1.18}$$

Another way of looking at the change of variables (4.1.15) is that it defines a new state:

$$|\Psi(t)'\rangle = e^{i\omega t \frac{\hat{S}_z}{\hbar}} |\Psi(t)\rangle, \tag{4.1.19}$$

which also leads to a transformed *time-independent* Hamiltonian

$$\hat{H}' \doteq \frac{\hbar}{2} \begin{pmatrix} \omega_0 - \omega & \omega_1 \\ \omega_1 & \omega - \omega_0 \end{pmatrix}, \tag{4.1.20}$$

such that it satisfies the Schroedinger equation

$$i\hbar\partial_t |\Psi(t)'\rangle = \hat{H}' |\Psi(t)'\rangle. \tag{4.1.21}$$

Because the transformed Hamiltonian is time independent, we can easily solve this problem using the "standard" approach, namely diagonalizing $\hat{H}'$ and developing the time evolved states in the eigen-ket basis. The general solution is left as an exercise, here we consider a few special cases, similarly to what we did in the absence of the time-dependence in the external field.

4.1.2 At resonance

At the resonance condition ($\omega_0 = \omega$) we have that the effective Hamiltonian is simply

$$\hat{H}' \doteq \frac{\hbar}{2}\begin{pmatrix} 0 & \omega_1 \\ \omega_1 & 0 \end{pmatrix} \tag{4.1.22}$$

$$= \omega_1 \hat{S}_x, \tag{4.1.23}$$

Let us analyze now the specific case in which we start our time evolution from an eigenstate of S_z, say $|S_z;+\rangle$, thus $c_+ = 1$ and $c_- = 0$. This state corresponds, because of (4.1.15), to $a_+ = 1$ and $a_- = 0$, thus

$$|\Psi(0)'\rangle = |S_z;+\rangle. \tag{4.1.24}$$

and

$$|\Psi(t)'\rangle = \frac{1}{\sqrt{2}}\left(e^{-i\omega_1\frac{t}{2}}|S_x;+\rangle + e^{i\omega_1\frac{t}{2}}|S_x;-\rangle\right) \tag{4.1.25}$$

$$= \frac{1}{2}\left(e^{-i\omega_1\frac{t}{2}}\left(|S_z;+\rangle + |S_z;-\rangle\right) + e^{i\omega_1\frac{t}{2}}\left(|S_z;+\rangle - |S_z;-\rangle\right)\right) \tag{4.1.26}$$

$$= \cos(\omega_1 t/2)|S_z;+\rangle - i\sin(\omega_1 t/2)|S_z;-\rangle. \tag{4.1.27}$$

and inverting the transformation we get

$$|\Psi(t)\rangle = e^{-i\omega t/2}\cos(\omega_1 t/2)|S_z;+\rangle - ie^{i\omega t/2}\sin(\omega_1 t/2)|S_z;-\rangle. \tag{4.1.28}$$

We thus see that the probability of finding the state in the opposite state ($|S_z;-\rangle$) is

$$P_-(t) = |\langle S_z;-|\Psi(t)\rangle|^2 \tag{4.1.29}$$

$$= \sin\left(\frac{t}{2}\omega_1\right)^2. \tag{4.1.30}$$

Notice that these oscillations have a frequency $\omega_1/2$ and can then be used as a tool to accurately measure the transverse magnetic field B_1, since $\omega_1 = \frac{|e|B_1}{m}$. This is the principle used in Magnetic Resonance Imaging (MRI) to measure very small magnetic field in the body, useful to learn about the local chemical or physical composition of cells and tissues.

4.1.3 Off Resonance

To solve the more general case, we notice that

$$P_-(t) = |\langle S_z;-|\Psi(t)\rangle|^2 \tag{4.1.31}$$

$$= \left|\langle S_z;-|e^{-i\omega t\frac{\hat{S}_z}{\hbar}}|\Psi(t)'\rangle\right|^2 \tag{4.1.32}$$

$$= \left|e^{i\omega t/2}\langle S_z;-|\Psi(t)'\rangle\right|^2 \tag{4.1.33}$$

$$= |\langle S_z;-|\Psi(t)'\rangle|^2, \tag{4.1.34}$$

where we have seen that the primed state is found by time evolution with a time-independent Hamiltonian:

$$|\Psi(t)'\rangle = e^{-i\frac{\hat{H}'}{\hbar}t}|\Psi(0)\rangle, \tag{4.1.35}$$

where we also have used that $|\Psi(0)\rangle = |\Psi(0)'\rangle$. In order to compute this state explicitly, we use the general identity

$$e^{i|\vec{a}|\left(\frac{a_x}{|\vec{a}|}\hat{\sigma}_x + \frac{a_y}{|\vec{a}|}\hat{\sigma}_y + \frac{a_z}{|\vec{a}|}\hat{\sigma}_z\right)} = \hat{I}\cos(|\vec{a}|) + i\left(\frac{a_x}{|\vec{a}|}\hat{\sigma}_x + \frac{a_y}{|\vec{a}|}\hat{\sigma}_y + \frac{a_z}{|\vec{a}|}\hat{\sigma}_z\right)\sin(|\vec{a}|), \tag{4.1.36}$$

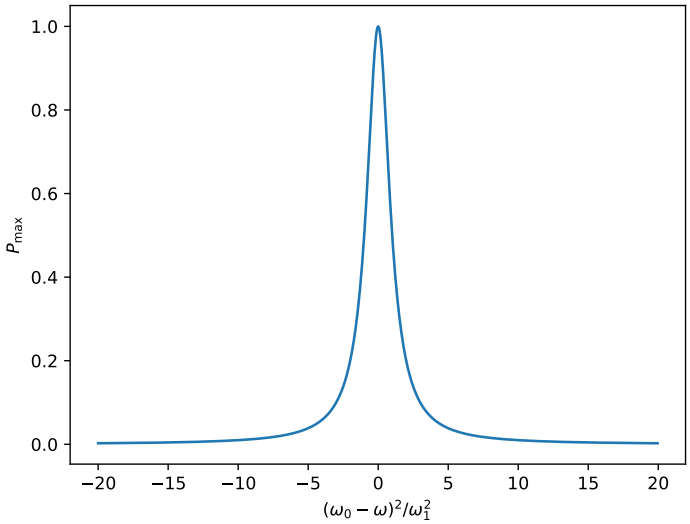

Figure 4.1.1: Maximum probability of flipping a spin using a time-dependent transverse field. At resonance, the maximum probability approaches unity.

for a vector $\vec{a} = (a_x, a_y, a_z)$. Proving this identity is left as an exercise, and relies on the Taylor series expansion of the exponential, in combination with the fact that the Pauli matrices square to the identity ($\hat{\sigma}_\alpha^2 = \hat{1}$). We therefore see that the unitary operator associated with the time evolution has exactly the same form:

$$e^{-i\frac{\hat{H}'}{\hbar}t} = \exp\left\{-\frac{i}{2}t\left[(\omega_0-\omega)\right]\hat{\sigma}_z + \omega_1\hat{\sigma}_x\right]\right\} \tag{4.1.37}$$

$$= \hat{I}\cos\left(\frac{t}{2}\Omega\right) - \frac{i}{\Omega}\left(\omega_1\hat{\sigma}_x + (\omega_0-\omega)\,\hat{\sigma}_z\right)\sin\left(\Omega\frac{t}{2}\right), \tag{4.1.38}$$

where $\Omega^2 = (\omega_0-\omega)^2 + \omega_1^2$. Thus if the initial state is $|\Psi(0\rangle = |S_z;+\rangle$, we see that the only term that matters for the transition probability is:

$$P_-(t) = \left|\frac{i\omega_1}{\Omega}\sin\left(\Omega\frac{t}{2}\right)\langle S_z;-|\hat{\sigma}_x|\Psi(0)\rangle\right|^2 \tag{4.1.39}$$

$$= \frac{\omega_1^2}{\Omega^2}\left[\sin\left(\Omega\frac{t}{2}\right)\right]^2 \tag{4.1.40}$$

$$= \frac{\omega_1^2}{\omega_1^2+(\omega_0-\omega)^2}\left[\sin\left(\sqrt{\omega_1^2+(\omega_0-\omega)^2}\frac{t}{2}\right)\right]^2. \tag{4.1.41}$$

This expression, first derived by Rabi, clearly shows that unless we are spot on the resonance (i.e. $\omega \simeq \omega_0$) the maximum probability of flipping the spin is small and given by a Lorentzian curve:

$$\max_t P_-(t) = \frac{\omega_1^2}{\omega_1^2+(\omega_0-\omega)^2}, \tag{4.1.42}$$

as also shown in Figure 4.1.1

Alternatively, it is possible to find the same result using the eigen-kets of the hamiltonian $\hat{H}'$. Calling $\Delta = \frac{\hbar}{2}(\omega_0-\omega)$, and $T = \frac{\hbar}{2}\omega_1$, we can diagonalize the Hamiltonian matrix explicitly, solving:

$$\det\begin{pmatrix}\Delta-E & T\\ T & -\Delta-E\end{pmatrix} = \left(E^2-\Delta^2\right)-T^2 \tag{4.1.43}$$

$$= 0 \tag{4.1.44}$$

which has the two solutions

$$E_\pm = \pm\frac{\hbar}{2}\sqrt{\omega_1^2 + (\omega_0 - \omega)^2}. \tag{4.1.45}$$

We see that this result is in agreement with what we previously found, in the sense that the Bohr frequency is

$$\Omega = \frac{E_+ - E_-}{\hbar}.$$

The corresponding two eigen-kets can also be found, and read:

$$|\pm\rangle \doteq \frac{1}{\sqrt{T^2 + (E_\pm - \Delta)^2}} \begin{pmatrix} T \\ E_\pm - \Delta \end{pmatrix}. \tag{4.1.46}$$

We leave as an Exercise to derive Rabi's formula using these eigen-kets, which is a more laborious route.

4.2 The Ammonia Molecule

The Ammonia molecule (NH_3) is an interesting case of complex system that nonetheless admits a relatively simple description in terms of a two level system. In the following we will make the assumption that we have a single molecule at rest, that is spinning around the axis formed by the triangular plane containing the Hydrogen atoms and the Nitrogen atom (see Fig. 4.2.1 for a visualization of the pyramidal structure of this molecule). In this situation, there are two equivalent positions for the Nitrogen atom: either in the upper part of the Hydrogen plane or in the lower part. Restricting our attention to only these two possible geometric configurations, we are going to associate two basis states $|1\rangle$ and $|2\rangle$, respectively to the up and down Nitrogen configurations, thus

$$|1\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{4.2.1}$$

$$|2\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{4.2.2}$$

Once determined the relevant basis states for our state vector, in general we will have that the state of molecule is given, as usual, by:

$$|\Psi\rangle = c_1|1\rangle + c_2|2\rangle, \tag{4.2.3}$$

and in general we expect the molecule to be in a linear superposition of the two geometric configurations.

Next, we would like to make an educated guess for the Hamiltonian of the system. Since we have claimed that the system is completely symmetric under exchanging the states $|1\rangle$ and $|2\rangle$, we expect that the average energy of these two states must be the same, thus

$$\langle 1|\hat{H}|1\rangle = \langle 2|\hat{H}|2\rangle \tag{4.2.4}$$

$$= E_0, \tag{4.2.5}$$

where E_0 is some experimentally measurable value. Moreover, in nature both configurations are observed, thus there must be a process that allows the two states to transform into each other, i.e. an up state to become a down state and vice-versa. Since the dynamics of the system is induced by the Hamiltonian, as we have seen in the previous Chapter, it must be that the matrix elements of the Hamiltonian connecting different states are non-vanishing (otherwise, as we will see, these transitions wouldn't be possible). We call the off-diagonal elements of the Hamiltonian

$$\langle 1|\hat{H}|2\rangle = \langle 2|\hat{H}|1\rangle^\star \tag{4.2.6}$$

$$= -A, \tag{4.2.7}$$

where $A \geq 0$ is a phenomenological, real-valued constant to be determined from the experiments. Also notice that $\langle 1|\hat{H}|2\rangle = \langle 2|\hat{H}|1\rangle^\star$ follows, in general, from the fact that the Hamiltonian is an Hermitian operator. The Hamiltonian then has the following matrix elements:

$$\hat{H} \doteq \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix}. \tag{4.2.8}$$

The energy eigenvalues are found diagonalizing this matrix, solving

$$\det \begin{pmatrix} E_0 - E & -A \\ -A & E_0 - E \end{pmatrix} = (E_0 - E)^2 - A^2 \tag{4.2.9}$$

$$= 0 \tag{4.2.10}$$

which has the two solutions

$$E_\pm = E_0 \pm A. \tag{4.2.11}$$

The corresponding two eigen-kets are also easily found (Exercise), and read:

$$|+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \tag{4.2.12}$$

$$|-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{4.2.13}$$

The Hamiltonian we are considering is time-independent, thus the time evolution operator is

$$\hat{U}(t) = e^{-\frac{i}{\hbar}\hat{H}t}, \tag{4.2.14}$$

and we can also find its matrix elements in the $|1\rangle$,$|2\rangle$ basis. For example, we can use the completeness of the energy eigen-kets:

$$\langle 1|e^{-\frac{i}{\hbar}\hat{H}t}|1\rangle = \langle 1|e^{-\frac{i}{\hbar}\hat{H}t}|+\rangle\langle +|1\rangle + \langle 1|e^{-\frac{i}{\hbar}\hat{H}t}|-\rangle\langle -|1\rangle \tag{4.2.15}$$

$$= e^{-\frac{i}{\hbar}E_+ t}\,|\langle +|1\rangle|^2 + e^{-\frac{i}{\hbar}E_- t}\,|\langle -|1\rangle|^2 \tag{4.2.16}$$

$$= \frac{1}{2} e^{-\frac{i}{\hbar}E_0 t}\left(e^{-\frac{i}{\hbar}At} + e^{+i\frac{E_0}{\hbar}t}\right) \tag{4.2.17}$$

$$= e^{-\frac{i}{\hbar}E_0 t}\cos(At/\hbar), \tag{4.2.18}$$

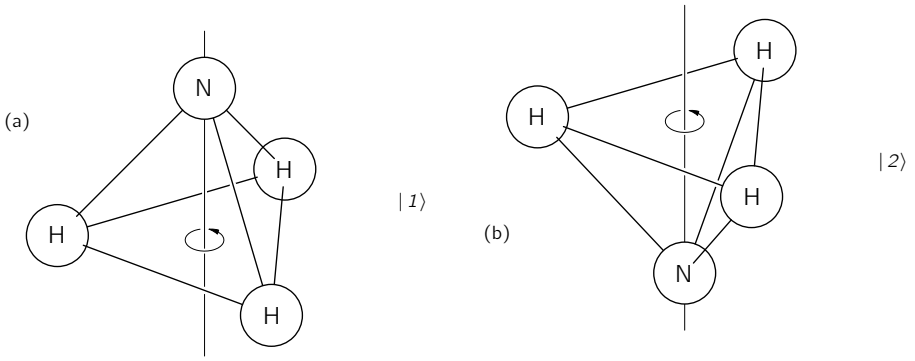

Figure 4.2.1: The ammonia molecule has two equivalent geometrical arrangements, one with the nitrogen up (left) and one with the nitrogen down (right). [Credits: Feynman's Lectures on Physics]

$$\langle 1|e^{-\frac{i}{\hbar}\hat{H}t}|2\rangle = \langle 1|e^{-\frac{i}{\hbar}\hat{H}t}|+\rangle\langle +|2\rangle + \langle 1|e^{-\frac{i}{\hbar}\hat{H}t}|-\rangle\langle -|2\rangle \quad (4.2.19)$$

$$= \frac{1}{2}\left(e^{-\frac{i}{\hbar}E_+ t} - e^{-\frac{i}{\hbar}E_- t}\right) \quad (4.2.20)$$

$$= \frac{1}{2}e^{-\frac{i}{\hbar}E_0 t}\left(e^{-\frac{i}{\hbar}At} - e^{+i\frac{E_0}{\hbar}t}\right) \quad (4.2.21)$$

$$= e^{-\frac{i}{\hbar}E_0 t} i\sin(At/\hbar), \quad (4.2.22)$$

and similarly for the other two matrix elements. The full matrix representation is therefore:

$$\hat{U}(t) \doteq e^{-\frac{i}{\hbar}E_0 t}\begin{pmatrix} \cos(At/\hbar) & i\sin(At/\hbar) \\ i\sin(At/\hbar) & \cos(At/\hbar) \end{pmatrix}, \quad (4.2.23)$$

which, as expected, can be checked to be a valid unitary matrix (Exercise).

Now, for a generic state as in Eq. (4.2.3), we can determine the time evolution of the amplitudes, i.e. $c_1(t)$ and $c_2(t)$ simply applying the time-evolution matrix:

$$\begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix} = e^{-\frac{i}{\hbar}E_0 t}\begin{pmatrix} \cos(At/\hbar) & i\sin(At/\hbar) \\ i\sin(At/\hbar) & \cos(At/\hbar) \end{pmatrix}\begin{pmatrix} c_1(0) \\ c_2(0) \end{pmatrix} \quad (4.2.24)$$

$$= e^{-\frac{i}{\hbar}E_0 t}\begin{pmatrix} \cos(At/\hbar)c_1(0) + i\sin(At/\hbar)c_2(0) \\ i\sin(At/\hbar)c_1(0) + \cos(At/\hbar)c_2(0) \end{pmatrix}, \quad (4.2.25)$$

which fully specify the time evolution of the state. For example, if we started from a state pointing in the up direction (i.e. $c_1(0) = 1$ and $c_2(0) = 0$), we would have

$$\begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix} = e^{-\frac{i}{\hbar}E_0 t}\begin{pmatrix} \cos(At/\hbar) & i\sin(At/\hbar) \\ i\sin(At/\hbar) & \cos(At/\hbar) \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (4.2.26)$$

$$= e^{-\frac{i}{\hbar}E_0 t}\begin{pmatrix} \cos(At/\hbar) \\ i\sin(At/\hbar) \end{pmatrix}, \quad (4.2.27)$$

thus the system acquires a finite probability of flipping into the down position, which can be computed as

$$P(2) = |\langle 2|\Psi(t)\rangle|^2 \quad (4.2.28)$$

$$= |c_2(t)|^2 \quad (4.2.29)$$

$$= \left[\sin\left(\frac{At}{\hbar}\right)\right]^2, \quad (4.2.30)$$

whereas the probability of finding the Nitrogen down is

$$P(1) = |\langle 1|\Psi(t)\rangle|^2 \quad (4.2.31)$$

$$= |c_1(t)|^2 \quad (4.2.32)$$

$$= \left[\cos\left(\frac{At}{\hbar}\right)\right]^2. \quad (4.2.33)$$

We therefore see that the time evolution of the molecule is analogous to what we found previously for the spin precession in the presence of a time-dependent magnetic field.

4.3 Static Electric Field

The charge distribution in the ammonia molecule is not spatially symmetric, specifically the electronic charge tends to be closer to the nitrogen atom. As a consequence, there is en effective electric dipole moment $\mu_{\mathbf{E}}$ pointing from the nitrogen atom to the hydrogens plane (see Figure 4.3.1).

Because of this finite electric dipole, the molecule will interact with an external electric field $\mathcal{E}$. The interaction energy is

$$U_E = -\mu_E \cdot \mathcal{E}, \quad (4.3.1)$$

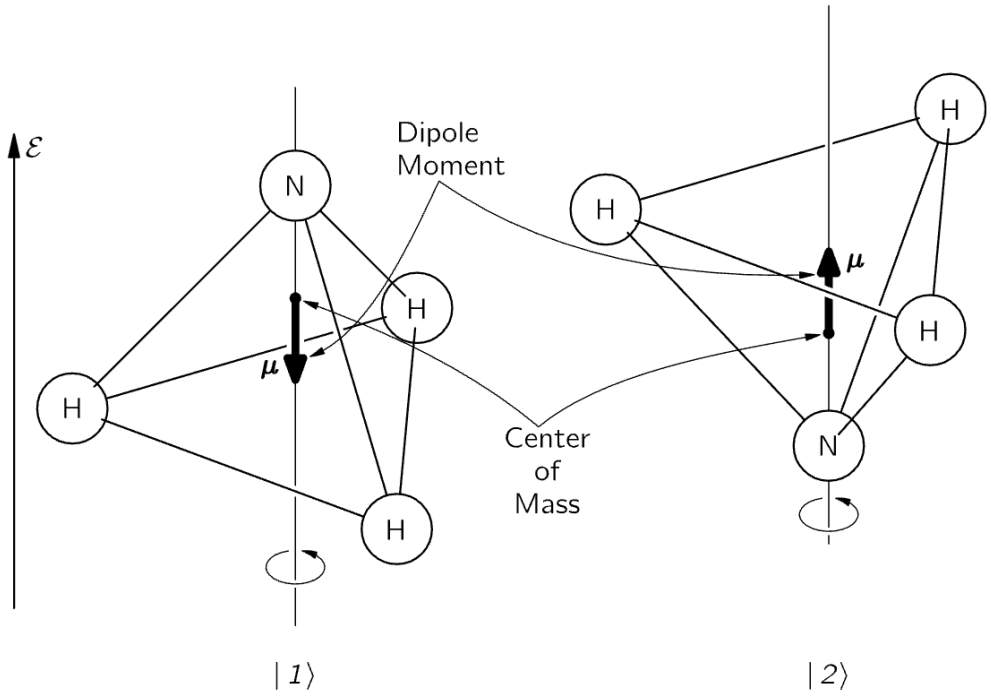

Figure 4.3.1: The ammonia molecule in the presence of an electric field in the z direction. The figure also shows that the dipole moment has two opposite directions in the states $|1\rangle$ and $|2\rangle$. [Credits: Feynman's Lectures on Physics]

thus for an electric field in the z direction we have that the two states $|1\rangle$ and $|2\rangle$ must have different energies (in one case the electric dipole is aligned with the electric field, in the other case it is anti-aligned). The Hamiltonian then becomes

$$\hat{H} \doteq \begin{pmatrix} E_0 + \mu_E \mathcal{E}_z & -A \\ -A & E_0 - \mu_E \mathcal{E}_z \end{pmatrix}. \tag{4.3.2}$$

The energy eigenvalues of this modified Hamiltonian can be found solving

$$\begin{aligned} \det \begin{pmatrix} E_0 + \mu_E \mathcal{E}_z - E & -A \\ -A & E_0 - \mu_E E_z - E \end{pmatrix} &= (E_0 - E + \mu_E \mathcal{E}_z)(E_0 - E - \mu_E \mathcal{E}_z) \quad (4.3.3) \\ &= 0, \quad (4.3.4) \end{aligned}$$

which has solutions

$$E_\pm = E_0 \pm \sqrt{(\mu_E \mathcal{E}_z)^2 + A^2}. \tag{4.3.5}$$

In the interesting case in which the electric field is small (i.e. $\mu_E \mathcal{E}_z \ll E_0$) we can approximate the eigenvalues to be

$$E_\pm \simeq E_0 \pm \left(A + \frac{(\mu_E \mathcal{E}_z)^2}{2A} \right), \tag{4.3.6}$$

where we have used the expansion $\sqrt{1+\epsilon} \simeq 1 + \frac{\epsilon}{2}$.

An interesting consequence of the presence of the electric field, is that we can use it to separate the two states $|1\rangle$ and $|2\rangle$ with a device similar to the Stern and Gerlach device (with the important difference that here we are using the electric field rather than the

magnetic field to perform the separation). This can be understood thinking of a spatially inhomogeneous electric field, such that the force acting on the two states is

$$F_z \simeq -\frac{\partial}{\partial z}E_{\pm} \tag{4.3.7}$$

$$= \mp\frac{(\mu_E)^2}{A}\frac{\partial}{\partial z}\mathcal{E}_z, \tag{4.3.8}$$

thus at the output of the analyzer we measure the states $|+\rangle$ and $|-\rangle$ going, respectively, in the upper and lower sides of the detector.

This selection mechanism is used, for example, in the Ammonia MASER, where the molecule with the higher energy are filtered out thanks to an external electric field, as outlined above.

4.4 Quantum Zeno Effect

We conclude our discussion on the dynamics of two-level systems showcasing one peculiarity of the interplay between quantum time evolution and the measurement process. For this case, we consider an idealized situation in which we have a single electron source, immersed in a magnetic field in the x direction, of strength B_x. We have already seen that in this case the Hamiltonian is

$$\hat{H} = \omega_x \hat{S}_x, \tag{4.4.1}$$

with $\omega_x = \frac{|e|B_x}{m}$. This situation is analogous to the resonant case studied for the time-dependent field, and the time evolved state is:

$$|\Psi(t)\rangle = \cos(\omega_x t/2)|S_z;+\rangle - i\sin(\omega_x t/2)|S_z;-\rangle. \tag{4.4.2}$$

This implies that if we start from a state $|\Psi(0)\rangle = |S_z;+\rangle$, we can find a total time T such that the probability of flipping the spin is equal to one, since:

$$P_-(T) = |\langle S_z;-|\Psi(T)\rangle|^2 \tag{4.4.3}$$

$$= |\sin(\omega_x T/2)|^2 \tag{4.4.4}$$

$$= 1, \tag{4.4.5}$$

implies (for example) $T = \pi/\omega_x$. An experimental setup with Stern-Gerlach devices realizing this scenario is shown in Fig. 4.4.1.

Figure 4.4.1: Polarized up spins (in the state $|S_z;+\rangle$) entering a magnetic field for a time $T = \pi/\omega_x$ exit with probability one in the down state, $|S_z;-\rangle$.

We then imagine to place a sequence of M Stern-Gerlach devices, all measuring the z component of the spin, as shown in Figure 4.4.2. We also make the assumption that the spin measurement is faster than any other time scale. What happens now is that we have M measurements, at a set of discrete times $t_k = kT/M$ and after each measurement the spin is "reset" to either $|S_z;+\rangle$ or $|S_z;-\rangle$, depending on the value of the amplitudes $\langle S_z;\pm|\Psi(t_k)\rangle$.

Now, if we take a large number of measurements, we have that the system evolves N times under the action of the magnetic-field Hamiltonian for a small time $\delta_t = T/M$. For

example, if the previously measured state was $|S_z;+\rangle$ at step t_{k-1}, then the probability of getting the same state after the measurement at step t_k is

$$P_+^{\delta_t} = |\cos(\omega_x \delta_t/2)|^2 \tag{4.4.6}$$

$$= 1 - \frac{\delta_t^2 \omega_x^2}{8} + \mathcal{O}(\delta_t^4). \tag{4.4.7}$$

The probability of getting a sequence of M states of type $|S_z;+\rangle$ following the measurements is then the product of these individual probabilities:

$$P_+(T) \simeq \left(1 - \frac{\delta_t^2 \omega_x^2}{8}\right)^M \tag{4.4.8}$$

$$= \left(1 - \frac{T^2 \omega_x^2}{8M^2}\right)^M \tag{4.4.9}$$

$$= \left(1 - \frac{\pi^2}{8M^2}\right)^M \tag{4.4.10}$$

thus we see that in the limit of a very large number of measurements,

$$\lim_{M\to\infty} P_+(T) = 1. \tag{4.4.11}$$

This results therefore is the exact contrary of what we would have obtained in the absence of intermediate measurements, where had found $P_+(T) = 0$!

This phenomenon, also observed experimentally, is known as Quantum Zeno Effect, because it's the quantum analog of the many paradoxes that the greek philosopher Zeno of Elea had invented (for example, "Achilles and the tortoise", is a similar classical paradox involving infinitesimal changes in time).

Figure 4.4.2: Experimental Setup for the Quantum Zeno Effect: a very long sequence of M measurements followed by the application of magnetic fields for shorter times $\delta_t = T/M$ leads to an output state that is unchanged with respect to the input.

4.5 References and Further Reading

A complete treatment of magnetic resonance can be found in full detail in Cohen-Tannoudji (complement F-IV of Tome 1). The Ammonia example has been first presented by Feynman in his lectures on Physics, Vol. III, Chapter 8 and Chapter 9.

Chapter 5

Continuous Degrees of Freedom

In the previous Chapters we have focused on the theoretical description of discrete degrees of freedom. These typically arise in the case of spin wave functions, when measurements only result in a discrete and finite set of possibilities. There are however very important cases in which observables are intrinsically continuous, even in the quantum case. This is the case for example of quantities such as position and momenta of particles. While it is always possible to think about these cases as specific limit of finite-dimensional vector spaces (see also Exercise 7), it is more natural to extend the previous formalism to account for intrinsically continuous degrees of freedom.

5.1 Bra-Ket formalism for continuous degrees of freedom

The formalism in this case is very close to what already discussed about finite vector spaces. We consider for example some continuous degree of freedom described by the operator $\hat{\xi}$, so that it possesses a set of eigenvalues ξ and eigenvectors $|\xi\rangle$ satisfying:

$$\hat{\xi}|\xi'\rangle \quad = \quad \xi'|\xi'\rangle, \tag{5.1.1}$$

notice that in this case there will be infinitely many kets $|\xi'\rangle$ satisfying this relationship, each of those with some associated eigenvalue.

Other than this important distinction, we also need to generalize our formalism to accommodate for orthonormality relationships and closure relationships that are well suited for continuous variables. Table 5.1 summarizes the main correspondences, that should be fairly intuitive to understand.

Property	Discrete case	Continuous case
Operator	$\hat{A}$	$\hat{\xi}$
Eigenvalues	$\hat{A}\|A_i\rangle = a_i\|A_i\rangle$	$\hat{\xi}\|\xi'\rangle = \xi'\|\xi'\rangle$
Completeness	$\sum_i \|A_i\rangle\langle A_i\| = \hat{1}$	$\int d\xi' \|\xi'\rangle\langle\xi'\| = \hat{1}$
State Expansion	$\|\Psi\rangle = \sum_i \|A_i\rangle\langle A_i\|\Psi\rangle$	$\|\Psi\rangle = \int d\xi' \|\xi'\rangle\langle\xi'\|\Psi\rangle$
State Normalization	$\sum_i \|\langle A_i\|\Psi\rangle\|^2 = 1$	$\int d\xi' \|\langle\xi'\|\Psi\rangle\|^2 = 1$
Orthonormality	$\langle A_i\|A_j\rangle = \delta_{ij}$	$\langle\xi'\|\xi''\rangle = \delta(\xi' - \xi'')$
Operators Matrix Elements	$\langle A_i\|\hat{A}\|A_j\rangle = \delta_{ij}a_j$	$\langle\xi'\|\hat{\xi}\|\xi''\rangle = \delta(\xi' - \xi'')\xi'$

Table 5.1: Correspondence between discrete and continuous kets formalism.

5.1.1 Dirac's Delta

The only more subtle point in this correspondence concerns the introduction of Dirac's delta $\delta(x)$, a generalized function that plays a very important role in the study of Hilbert spaces. Here we just recall that Dirac delta can be seen as the limit of an infinitely tight Gaussian:

$$\delta(x) \quad = \quad \lim_{\sigma \to 0} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}, \tag{5.1.2}$$

as also shown in Fig. 5.1.1.

Figure 5.1.1: Dirac's delta as the limit of an infinitely narrow Gaussian, $\delta_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$.

From this limiting expression, it follows directly that the delta is an even function of its argument:

$$\delta(x) \quad = \quad \delta(-x), \tag{5.1.3}$$

and that it integrates to one

$$\int_{-\infty}^{\infty} dx \delta(x) \quad = \quad 1. \tag{5.1.4}$$

Another distinctive feature of the delta is that

$$\int_{-\infty}^{\infty} dx F(x)\delta(x) \quad = \quad F(0), \tag{5.1.5}$$

which is quite natural when thinking of the delta as a very sharp gaussian, that is zero almost everywhere but close to the origin. This property of course generalizes to arbitrary arguments of the delta, that correspond to shifting the mean value of the corresponding limiting gaussian:

$$\int_{-\infty}^{\infty} dx F(x)\delta(x - x_0) \quad = \quad F(x_0). \tag{5.1.6}$$

Other important properties of Dirac's delta can be found in math textbooks.

5.1.2 The state function

Since, in general, the variable ξ' associated to the eigen-kets is continuous, we can identify the expansion coefficients of an arbitrary quantum state in this basis as a complex-valued *function* (also known as "wave function" or "state function"):

$$\Psi(\xi) \quad \equiv \quad \langle \xi | \Psi \rangle, \tag{5.1.7}$$

with the property of being L^2, integrable, i.e. it can be correctly normalized in a way that

$$\int d\xi |\Psi(\xi)|^2 \quad = \quad 1. \tag{5.1.8}$$

Notice that this property is a fundamental property of wave functions, and it is a direct consequence of the Born interpretation for $P(\xi) = |\Psi(\xi)|^2$, that is the probability *density* of measuring a given value of ξ. For example, armed with Born's probability density, if we had to compute the probability P_δ of measuring a value $\xi_0 < \xi < \xi_0 + \delta$, we would use the standard rules of probability theory:

$$P_\delta(\xi_0) \quad = \quad \int_{\bar{\xi}_0}^{\bar{\xi}_0 + \delta} d\xi \, |\Psi(\xi)|^2 \tag{5.1.9}$$

$$\simeq \quad \delta \, |\Psi(\xi_0)|^2 \tag{5.1.10}$$

$$= \quad \delta \, |\langle \xi_0 | \Psi \rangle|^2 . \tag{5.1.11}$$

Notice that this also marks a slight but important difference with respect to the case of discrete variables, since the amplitude $|\langle\xi_0|\Psi\rangle|^2$ is *not* the probability of obtaining the measurement value ξ_0, as it would be in the discrete case, but rather the probability *density* of obtaining ξ_0. Strictly speaking, for a continuous variable the probability of obtaining *exactly* ξ_0 is infinitesimally small in the window δ around it, thus it is zero in the limit $\delta \to 0$. In most of the applications we will never compute point-wise probability densities to evaluate physical quantities, but rather integrals over finite windows of values.

5.2 Position operator

Let us now focus on the common case in which we are interested in measuring or just characterizing theoretically the position of a given particle (say, an electron). For simplicity, we first focus on the case in which the particle is constrained to be in one dimension. In this case, the eigen-kets are just one-dimensional coordinates:

$$\hat{x}|x'\rangle \quad = \quad x'|x'\rangle. \tag{5.2.1}$$

As much as done when considering the measurement postulates for spin systems, a very similar situation is found when considering continuous variables. Specifically, we can imagine that we can measure the position of our particle taking a snapshot of it. Every time we take a picture of this particle, we will see a spot in our picture at a given position x', and the wave-function collapses into the corresponding eigenstate

$$|\Psi\rangle \quad \to \quad |x'\rangle\langle x'|\Psi\rangle. \tag{5.2.2}$$

We can easily extend this description also to higher dimensions, i.e. we lift the constraint of having purely one-dimensional particles. In this case the wave-function is therefore a complex-valued function of the vector $\mathbf{r} = (x, y, z)$:

$$\langle x, y, z|\Psi\rangle \quad \equiv \quad \Psi(x, y, z), \tag{5.2.3}$$

where we have postulated that Ψ is an eigenstate of all coordinates. This hypothesis is verified experimentally. As a result of the discussion in the previous Chapter, this implies that position operators commute

$$[\hat{x}, \hat{y}] \quad = \quad 0 \tag{5.2.4}$$
$$[\hat{x}, \hat{z}] \quad = \quad 0 \tag{5.2.5}$$
$$[\hat{y}, \hat{z}] \quad = \quad 0. \tag{5.2.6}$$

5.3 Translation operator

In addition to the concept of *position* for a quantum particle, the other major observable concerning particles in continuous space is the *momentum*. In order to derive a consistent form for the momentum operator, we first need to introduce the concept of translation operator, since this will be instrumental in defining the form that the momentum operator takes in quantum mechanics.

We start by considering an infinitesimal translation operator, $\hat{T}(\delta\mathbf{r})$ parameterized by a certain 3-dimensional infinitesimal translation $\delta\mathbf{r} = (\delta x, \delta y, \delta z)$, whose job is to translate a certain eigen-ket of the position operator:

$$\hat{T}(\delta\mathbf{r})|\mathbf{r}\rangle \quad = \quad |\mathbf{r} + \delta\mathbf{r}\rangle. \tag{5.3.1}$$

The action of this operator is quite simple, since it takes a certain eigen-ket of the position operator, $|\mathbf{r}\rangle$, and returns another eigen-ket of the position operator, $|\mathbf{r}' + \delta\mathbf{r}'\rangle$. From this expression we also see that $|\mathbf{r}'\rangle$ is not an eigen-ket of the translation operator, since it is transformed into another eigen-ket and not into itself.

Applied on an arbitrary state, $|\Psi\rangle$, the action of the infinitesimal translation operator is

$$\begin{aligned}\hat{T}(\delta\mathbf{r})|\Psi\rangle &= \hat{T}(\delta\mathbf{r})\int d\mathbf{r}|\mathbf{r}\rangle\Psi(\mathbf{r}) && (5.3.2)\\ &= \int d\mathbf{r}\hat{T}(\delta\mathbf{r})|\mathbf{r}\rangle\Psi(\mathbf{r}) && (5.3.3)\\ &= \int d\mathbf{r}|\mathbf{r}+\delta\mathbf{r}\rangle\Psi(\mathbf{r}) && (5.3.4)\\ &= \int d\mathbf{r}|\mathbf{r}\rangle\Psi(\mathbf{r}-\delta\mathbf{r}), && (5.3.5)\end{aligned}$$

where in the last line we have considered the change of variable $\mathbf{r} \to \mathbf{r} - \delta\mathbf{r}$, that does not affect the value of the integral, since we are already integrating over the full space. This expression also shows that, in position space, the effect of the translation operator is effectively $\Psi(\mathbf{r}) \to \Psi(\mathbf{r} - \delta\mathbf{r})$.

We can already derive several interesting properties of the operator $\hat{T}$, just looking at how the state transforms under its action. Specifically, we should have that the translated state $|\Psi'\rangle = \hat{T}(\delta\mathbf{r})|\Psi\rangle$, is still correctly normalized, i.e.

$$\begin{aligned}\langle\Psi'|\Psi'\rangle &= \langle\Psi|\hat{T}^\dagger(\delta\mathbf{r})\hat{T}(\delta\mathbf{r})|\Psi\rangle && (5.3.6)\\ &= \langle\Psi|\Psi\rangle. && (5.3.7)\end{aligned}$$

This condition is satisfied if the translation operator is unitary:

$$\hat{T}^\dagger(\delta\mathbf{r})\hat{T}(\delta\mathbf{r}) = \hat{1}. \quad (5.3.8)$$

The second property we expect from this operator, is that it can be arbitrarily *composed*, in the sense that subsequent translations of $\delta\mathbf{r}_1, \delta\mathbf{r}_2, \delta\mathbf{r}_3, \ldots$ must be equivalent to a single translation of the sum vector:

$$\hat{T}(\delta\mathbf{r}_1)\hat{T}(\delta\mathbf{r}_2)\hat{T}(\delta\mathbf{r}_2)\ldots = \hat{T}(\delta\mathbf{r}_1 + \delta\mathbf{r}_2 + \delta\mathbf{r}_3 + \ldots). \quad (5.3.9)$$

Furthermore, if the translate a certain system *back* to its original position, this operation should be equivalent to applying the inverse transformation:

$$\hat{T}(-\delta\mathbf{r}) = \hat{T}^{-1}(\delta\mathbf{r}), \quad (5.3.10)$$

where $\hat{T}^{-1}$ denotes the inverse of the operator.

The last property that we can intuitively expect is that in the limit of vanishing translations the operator $\hat{T}$ should strictly reduce to the identity

$$\lim_{|\delta\mathbf{r}|\to 0} \hat{T}(\delta\mathbf{r}) = \hat{1}. \quad (5.3.11)$$

As we have already seen for the case of the time evolution operator, and as a consequence of Stone's theorem, all these conditions are satisfied if we take the infinitesimal translation operator to be described by the following unitary operator

$$\hat{T}(\delta\mathbf{r}) = e^{-i\hat{K}\cdot\delta\mathbf{r}}, \quad (5.3.12)$$

where $\hat{K}$ is a vector operator $\hat{K} = (\hat{K}_x, \hat{K}_y, \hat{K}_z)$ where each of the individual components are Hermitian operators. Here, the exponential of the operator has exactly the same meaning it would have for finite vector spaces, and it is again understood in terms of its Taylor expansion:

$$e^{-i\hat{K}\cdot\delta\mathbf{r}} = \hat{1} - i\hat{K}\cdot\delta\mathbf{r} - \frac{1}{2}(\hat{K}\cdot\delta\mathbf{r})(\hat{K}\cdot\delta\mathbf{r}) + \mathcal{O}\left[(\delta\mathbf{r})^3\right]. \quad (5.3.13)$$

From this expansion, we immediately see that (5.3.11) is verified. The unitarity assumption is also quick to verify, since it is an elementary property of the exponential of an operator that $e^{X^+} = \left(e^X\right)^\dagger$, thus

$$\hat{T}^\dagger(\delta\mathbf{r}) = e^{i\left(\hat{K}\right)^\dagger\cdot\delta\mathbf{r}} \tag{5.3.14}$$

$$= e^{i\hat{K}\cdot\delta\mathbf{r}}, \tag{5.3.15}$$

where in the last line we have used the fact that $\hat{K}$ is Hermitian. The composition property is also a consequence of the exponential structure

$$e^{-i\hat{K}\cdot\delta\mathbf{r}_1}e^{-i\hat{K}\cdot\delta\mathbf{r}_2}\ldots = e^{-i\hat{K}\cdot(\delta\mathbf{r}_1+\delta\mathbf{r}_2+\ldots)}, \tag{5.3.16}$$

as well as the inversion property

$$\hat{T}(-\delta\mathbf{r}) = e^{i\hat{K}\cdot\delta\mathbf{r}}, \tag{5.3.17}$$

$$e^{i\hat{K}\cdot\delta\mathbf{r}}e^{-i\hat{K}\cdot\delta\mathbf{r}} = \hat{1}, \tag{5.3.18}$$

$$\hat{T}(-\delta\mathbf{r})\hat{T}(\delta\mathbf{r}) = \hat{1}, \tag{5.3.19}$$

$$\hat{T}(-\delta\mathbf{r}) = \hat{T}^{-1}(\delta\mathbf{r}). \tag{5.3.20}$$

5.3.1 Commutation relations of $\hat{K}$

The operator $\hat{K}$ we have introduced before is Hermitian thus, by the fundamental axioms of the quantum theory, it is also a physical *observable*. A natural question we can ask already at this stage is whether this observable is compatible or not with measurements of the position operator. For this purpose, we know that it is enough to compute the commutator $[\hat{K}, \hat{r}]$ and verify if it is vanishing or not.

When we first apply an infinitesimal translation and *then* the position operator, we have:

$$\hat{\mathbf{r}}\hat{T}(\delta\mathbf{r})|\mathbf{r}'\rangle = \hat{\mathbf{r}}|\mathbf{r}'+\delta\mathbf{r}\rangle \tag{5.3.21}$$

$$= (\mathbf{r}'+\delta\mathbf{r})|\mathbf{r}'+\delta\mathbf{r}\rangle, \tag{5.3.22}$$

on the other hand when we first apply the position operator and *then* translate, we have

$$\hat{T}(\delta\mathbf{r})\hat{\mathbf{r}}|\mathbf{r}'\rangle = \mathbf{r}'\hat{T}(\delta\mathbf{r})|\mathbf{r}'\rangle \tag{5.3.23}$$

$$= \mathbf{r}'|\mathbf{r}'+\delta\mathbf{r}\rangle, \tag{5.3.24}$$

subtracting these two equations we therefore get

$$[\hat{\mathbf{r}}, \hat{T}(\delta\mathbf{r})]|\mathbf{r}'\rangle = \delta\mathbf{r}'|\mathbf{r}'+\delta\mathbf{r}\rangle \tag{5.3.25}$$

$$= \delta\mathbf{r}'|\mathbf{r}'\rangle + \mathcal{O}[(\delta\mathbf{r})^2]. \tag{5.3.26}$$

Now, since this equation must be verified for all kets $|\mathbf{r}'\rangle$, we conclude that this commutator identity between operators must hold

$$[\hat{\mathbf{r}}, \hat{T}(\delta\mathbf{r})] = \delta\mathbf{r}\hat{1}. \tag{5.3.27}$$

Remark. In the following, we will often avoid explicitly writing the r.h.s. of this equation as $\delta\mathbf{r}\hat{1}$ and omit the identity operator. It should be however always clear that the result of a commutator of two operators is *always* an operator, and never a scalar.

Given the definition of the translation operator in terms of $\hat{K}$, we also have

$$[\hat{\mathbf{r}}, \hat{1} - i\hat{K}\cdot\delta\mathbf{r}] = -i[\hat{\mathbf{r}}, \hat{K}\cdot\delta\mathbf{r}] \tag{5.3.28}$$

$$= \delta\mathbf{r}. \tag{5.3.29}$$

Component-wise, we have (say for the x component):

$$[\hat{x}, \hat{K}_x] = i,$$

and more generally

$$[\hat{\mathbf{r}}_\alpha, \hat{\mathbf{K}}_\beta] = i\delta_{\alpha\beta}.$$

5.4 The Momentum operator

Similarly to what we have done for the case of the time evolution, where we had identified the operator $\hat{\Omega}$ with the Hamiltonian of the system, through the units rescaling $\hat{H} = \hbar\hat{\Omega}$, also in the case of the $\hat{\mathbf{K}}$ operator we can perform a similar analysis. Specifically, and also in analogy with classical mechanics, we want to identify the generator of the spatial translations with the momentum operator $\hat{\mathbf{p}}$. From a dimensional point of view, we see then this is achieved posing

$$\hat{\mathbf{p}} = \hbar\hat{\mathbf{K}}, \tag{5.4.1}$$

thus the commutation relations read

$$[\hat{\mathbf{r}}_\alpha, \hat{\mathbf{p}}_\beta] = i\hbar\delta_{\alpha\beta}. \tag{5.4.2}$$

We will show a bit later on that the factor $\hbar$ is not only necessary for dimensional reasons, but also necessary to both recover classical mechanics, in an appropriately chosen limit, and to explain experimental evidence from atomic physics.

5.4.1 Correspondence Principle

It was remarked by Dirac, that the commutation relations:

$$[\hat{\mathbf{r}}_\alpha, \hat{\mathbf{r}}_\beta] = 0 \tag{5.4.3}$$
$$[\hat{\mathbf{p}}_\alpha, \hat{\mathbf{p}}_\beta] = 0 \tag{5.4.4}$$
$$[\hat{\mathbf{r}}_\alpha, \hat{\mathbf{p}}_\beta] = i\hbar\delta_{\alpha\beta}. \tag{5.4.5}$$

are formally similar to the classical Poisson bracket relations between positions and momenta, that naturally emerge in the Hamiltonian formalism of classical mechanics. The analogy, also known as "correspondence principle", is found through the replacement

$$[,]_{\text{classical}} \rightarrow \frac{[,]}{i\hbar}, \tag{5.4.6}$$

where the classical Poisson bracket of functions of the canonical coordinates r and p is

$$[A(\mathbf{r},\mathbf{p}), B(\mathbf{r},\mathbf{p})]_{\text{classical}} = \sum_\alpha \left(\frac{\partial A}{\partial r_\alpha}\frac{\partial B}{\partial p_\alpha} - \frac{\partial A}{\partial p_\alpha}\frac{\partial B}{\partial r_\alpha} \right). \tag{5.4.7}$$

Most notably, in classical mechanics we have that

$$[\mathbf{r}_\alpha, \mathbf{p}_\beta]_{\text{classical}} = \delta_{\alpha\beta}, \tag{5.4.8}$$

which through the formal replacement, Eq. (5.4.6), reduces to the canonical quantum commutation relations. This analogy is valid essentially because classical Poisson brackets have algebraic rules that are analogous to operator commutators (for example, you can check that $[A,B]_{\text{classical}} = -[B,A]_{\text{classical}}$ etc. This analogy also carries along to the case of dynamics, where a classical observable of the canonical coordinates, $A(r,q)$ satisfies the equation of motion:

$$\frac{d}{dt}A(r,q) = [A,H]_{\text{classical}}, \tag{5.4.9}$$

where H is the classical Hamiltonian function, satisfying Hamilton's equations of motion:

$$\begin{cases} \dot{r} = \frac{\partial H}{\partial p} = [r,H]_{\text{classical}} \\ \dot{p} = -\frac{\partial H}{\partial r} = [p,H]_{\text{classical}} \end{cases}, \tag{5.4.10}$$

we thus immediately see that Eq. (5.4.9) is formally equivalent to Heisenberg's equations of motion, provided that the replacement, Eq. (5.4.6), is performed. Incidentally, this also

provides a further justification of why we have previously called the operator $\hat{H}$, generating the time evolution, the *Hamiltonian* of the system. Nonetheless, it should be remarked that these analogies are somewhat only formal (essentially due to the algebraic properties of the Poisson brackets), and that is often the case that quantum mechanical quantities do not have a classical analogous through the correspondence (5.4.6). For example, the spin observable is not in formal correspondence to some function of the classical variables, and an approach based on the time or space translation operators are more general to derive equations of motion and commutation relations for the quantum observables, rather than relying on the classical correspondence.

5.5 Position representation

We are now ready to discuss some important properties of the position representation, namely of how quantum states expanded in the basis of the position operator look like. We start the discussion first with the one-dimensional case, to avoid proliferation of unnecessary bold faces and indices. We have already seen at the beginning of this Chapter how an arbitrary state can be expended in a continuous basis written as

$$|\Psi\rangle = \int dx|x\rangle\langle x|\Psi\rangle \tag{5.5.1}$$

$$= \int dx|x\rangle\psi(x). \tag{5.5.2}$$

Overlaps between states expressed in the position representation are also obtained integrating over space:

$$\langle\Phi|\Psi\rangle = \int dx\langle\Phi|x\rangle\langle x|\Psi\rangle \tag{5.5.3}$$

$$= \int dx\Phi(x)^\star\Psi(x), \tag{5.5.4}$$

and matrix elements of some some operator $\hat{A}$ are just

$$\begin{aligned}\langle\Phi|\hat{A}|\Psi\rangle &= \int dx\langle\Phi|x\rangle\langle x|\hat{A}\Psi\rangle \\ &= \int dxdx'\langle\Phi|x\rangle\langle x|\hat{A}|x'\rangle\langle x'|\Psi\rangle \\ &= \int dxdx'\Phi(x)^\star A(x,x')\Psi(x'). \end{aligned}\tag{5.5.5}$$

All of these are really just immediate consequences of the definition of completeness for kets defined on continuous variables. As it can be seen from Eq. (5.5.5), in general for an arbitrary operator we need to evaluate its matrix elements $\langle x|\hat{A}|x'\rangle \equiv A(x,x')$ in the position representation basis.

A simplification occurs when the operator $\hat{A}$ is diagonal in this basis. This happens for any function of the coordinates only, say for operators $\hat{A}_f = f(\hat{x})$, where f is some arbitrary analytical function of the coordinates (for example $f(\hat{x}) = a\hat{x}^2 - b\hat{x}$ etc). By taking a Taylor series expansion of the function, in this case it is easy to see (exercise) that

$$[\hat{A}_f, \hat{x}] = 0, \tag{5.5.6}$$

thus the operator is diagonal in the position basis, implying that

$$\hat{A}_f|x\rangle = f(x)|x\rangle, \tag{5.5.7}$$

and for the matrix elements of the operators in the position basis we have

$$\langle x|\hat{A}_f|x'\rangle = \delta(x-x')f(x). \tag{5.5.8}$$

In this special case then the matrix elements of the operator among two general states simplify to

$$\langle\Phi|\hat{A}_f|\Psi\rangle = \int dx\Phi(x)^\star f(x)\Psi(x). \tag{5.5.9}$$

5.5.1 Momentum operator in the position basis

We now want to compute matrix elements of the momentum operator in the position basis. For simplicity, we work again in one spatial dimension and consider only the component x of the momentum, $\hat{p}_x$. Knowing the action of the momentum operator in this basis is fundamental, for example, to compute expectation values of momenta, kinetic energy etc, as we will see.

To start off, we consider the action of the translation operator on an arbitrary ket, Eq. (5.3.5), in the case of a small displacement δ_x :

$$e^{-i\hat{K}_x\delta_x}|\Psi\rangle = \int dx'|x'\rangle\psi(x'-\delta_x) \tag{5.5.10}$$

$$= \int dx'|x'\rangle\left(\psi(x') - \delta_x\frac{\partial\psi}{\partial x'}\right) + \mathcal{O}(\delta_x^2), \tag{5.5.11}$$

and multiplying on the left by a basis ket we deduce that the wave-function amplitudes after applying the translation operator read

$$\langle x|e^{-i\hat{K}_x\delta_x}|\Psi\rangle = \psi(x) - \delta_x\frac{\partial\psi(x)}{\partial x} + \mathcal{O}(\delta_x^2). \tag{5.5.12}$$

On the other hand,

$$e^{-i\hat{K}_x\delta_x}|\Psi\rangle = (\hat{I} - i\hat{K}_x\delta_x)|\Psi\rangle + \mathcal{O}(\delta_x^2), \tag{5.5.13}$$

from which we also deduce that

$$\langle x|e^{-i\hat{K}_x\delta_x}|\Psi\rangle = \psi(x) - \frac{i}{\hbar}\delta_x\langle x|\hat{p}_x|\Psi\rangle + \mathcal{O}(\delta_x^2). \tag{5.5.14}$$

Equating equations (5.5.12) and (5.5.14) at the given order in δ_x we then get

$$\langle x|\hat{p}_x|\Psi\rangle = -i\hbar\frac{\partial\psi(x)}{\partial x}. \tag{5.5.15}$$

This expression basically tells us that the effect of the momentum operator (apart from factors) is to take the derivative of the wave function. If now consider the special case in which $|\Psi\rangle = |x'\rangle$, and $\psi(x) = \langle x|\Psi\rangle = \langle x|x'\rangle = \delta(x-x')$, we get the sought-after matrix elements:

$$\langle x|\hat{p}_x|x'\rangle = -i\hbar\frac{\partial}{\partial x}\delta(x-x'). \tag{5.5.16}$$

Also, matrix elements of the momentum operator between states known in the position basis are then just

$$\langle\Phi|\hat{p}_x|\Psi\rangle = -i\hbar\int dxdx'\Phi(x)^\star\frac{\partial}{\partial x}\delta(x-x')\Psi(x') \tag{5.5.17}$$

$$= -i\hbar\int dx\Phi(x)^\star\frac{\partial}{\partial x}\Psi(x). \tag{5.5.18}$$

Arbitrary analytic functions of the momentum can also be obtained, using the corresponding Taylor series, and knowing that each power of the momentum performs a derivative with respect to the coordinates. An important higher order function is just the square of the

momentum, giving rise to the kinetic energy $E_T = \hat{p}^2/2m$, for a massive particle. In this case,

$$\langle x|\hat{p}_x^2|\Psi\rangle = \int dx' \langle x|\hat{p}_x|x'\rangle\langle x'|\hat{p}_x|\Psi\rangle \tag{5.5.19}$$

$$= (-i\hbar)^2 \int dx' \left(\frac{\partial}{\partial x}\delta(x-x')\right)\left(\frac{\partial}{\partial x'}\psi(x')\right) \tag{5.5.20}$$

$$= -\hbar^2 \frac{\partial^2}{\partial x^2}\psi(x). \tag{5.5.21}$$

Exercise 7. Instead of using the operatorial representation of the momentum operator as a derivative operator acting on a Hilbert space, consider instead an approximation of the momentum operator in a finite-dimensional vector space. To this end, consider a finite one-dimensional system living in a box of width L, and discretize the space into small "bins" of width Δ. What is the dimensionality of the state vector in this discretized position representation? Using then the definition of the translation operator, show that the momentum operator in this finite-dimensional vector space is approximated by a tridiagonal matrix.

5.6 Momentum representation

Until now we have worked solely with eigenstates of the position operator, however it is interesting to look at the eigenstates of the momentum operator as well. These are defined by the usual eigenvalue relation:

$$\hat{p}|\hat{p}\rangle = p|\hat{p}\rangle, \tag{5.6.1}$$

and can be useful, for example, if we wanted to represent a certain wave function in this basis. In order to avoid cluttering the notation, in this section we will omit the lower index x to characterize the x component of the momentum, thus it is assumed, starting from the equation above, that $\hat{p} \equiv \hat{p}_x$ Since in the previous discussion we have already derived the action of the momentum operator on an arbitrary ket (Eq. (5.5.15)), we can rewrite the eigenvalue equation as

$$\langle x|\hat{p}|p\rangle = p\langle x|p\rangle \tag{5.6.2}$$

$$= -i\hbar\frac{\partial}{\partial x}\langle x|p\rangle. \tag{5.6.3}$$

We therefore see that the eigenfunctions of the momentum satisfy this simple differential equation

$$p\langle x|p\rangle = -i\hbar\frac{\partial}{\partial x}\langle x|p\rangle. \tag{5.6.4}$$

It is easy to see that we are after an exponential form:

$$\langle x|p\rangle = Ne^{i\frac{px}{\hbar}}, \tag{5.6.5}$$

where N is a normalization that should be fixed imposing the usual constraint:

$$\langle p|p'\rangle = \delta(p-p'), \tag{5.6.6}$$

$$\int dx \langle p|x\rangle\langle x|p'\rangle = \delta(p-p') \tag{5.6.7}$$

We can explicitly write the l.h.s. of this equation and notice that is is just a representation of the delta function:

$$|N|^2 \int dx e^{i\frac{(p'-p)x}{\hbar}} = |N|^2 2\pi\hbar\delta(p-p'), \tag{5.6.8}$$

we therefore conclude that $N = 1/\sqrt{2\pi\hbar}$ is a good normalization (there is an arbitrary phase to be picked in choosing N and the convention is just to take N to be real and positive). We thus have that the eigenstates of the momentum operators are plane waves:

$$\langle x|p\rangle = \frac{e^{i\frac{px}{\hbar}}}{\sqrt{2\pi\hbar}}. \tag{5.6.9}$$

This also allows us to find the relationship between wave functions in different bases. For example, a wave-function in real space, $\psi(x)$, has a representation $\tilde{\psi}(p)$ in momentum space:

$$\langle p|\psi\rangle = \int dx \langle p|x\rangle\langle x|\psi\rangle \tag{5.6.10}$$

$$\tilde{\psi}(p) = \int dx \frac{e^{-i\frac{px}{\hbar}}}{\sqrt{2\pi\hbar}}\psi(x), \tag{5.6.11}$$

and vice-versa:

$$\langle x|\psi\rangle = \int dp \langle x|p\rangle\langle p|\psi\rangle \tag{5.6.12}$$

$$\psi(x) = \int dp \frac{e^{i\frac{px}{\hbar}}}{\sqrt{2\pi\hbar}}\tilde{\psi}(p). \tag{5.6.13}$$

The correspondence between real-space and momentum-space wave-function is now evident in its beauty: transforming a given quantum state between these two bases requires performing (inverse and direct) Fourier transforms of the corresponding wave functions. It should be stressed that all of these results have been obtained using the few postulates of completeness for quantum states, and the connection between translation operator and the momentum operator.

5.7 Quantum and classical particles

We are now already in position to make an intermediate summary of the results we have obtained so far, and clarify the fundamentally different description of particles arising from quantum mechanics. The summary is presented in Table 5.2.

	Classical	Quantum
State	Two vector quantities : $\mathbf{r} = (x, y, z)$ and $\mathbf{p} = (p_x, p_y, p_z)$	A ray in Hilbert space : the state vector $\|\Psi\rangle$
Quantities	The values of $\mathbf{r}$ and $\mathbf{p}$ can be measured directly	$\|\Psi(\mathbf{r})\|^2$ and $\|\Psi(\mathbf{p})\|^2$ are the probability densities of observing a certain $\mathbf{r}$ or $\mathbf{p}$
Uncertainty	No constraint	Heisenberg principle
Time Evolution	$\dot{\mathbf{r}} = [\mathbf{r}, H]_{\text{classical}}$ $\dot{\mathbf{p}} = [\mathbf{p}, H]_{\text{classical}}$	$\dot{\hat{\mathbf{r}}} = \frac{1}{i\hbar}[\hat{\mathbf{r}}, \hat{H}]$ $\dot{\hat{\mathbf{p}}} = \frac{1}{i\hbar}[\hat{\mathbf{p}}, \hat{H}]$

Table 5.2: Comparing the classical and quantum description of a particle.

5.8 Gaussian wave packet

In the following Chapter we will see how the wave functions can be obtained from first principles, solving the Schroedinger equation. For the moment however, it is already interesting to look at specific cases of wave functions that can help us familiarize with the basic concepts of the theory.

An important example is called the *gaussian wave packet* and allows us, in a certain limit, also to connect to the classical behavior we would expect from a point-like particle. The wave function in position space takes the form:

$$\Psi(x;k,d) = \frac{1}{\sqrt[4]{2\pi}\sqrt{d}} e^{ikx-\frac{x^2}{4d^2}}, \tag{5.8.1}$$

thus this state is parameterized by two constants k and d we can vary at will. We will drop the explicit parametric dependence on these two parameters in the following. First important observation is that the Born probability density in real space is:

$$|\Psi(x)|^2 = \frac{1}{d\sqrt{2\pi}} e^{-\frac{x^2}{2d^2}}, \tag{5.8.2}$$

thus it is a *gaussian* centered at the origin and variance d^2. This is the reason why we referred to this before as a gaussian wave packet. The first consequence of this observation is that the parameter d controls how localized the particle is around the origin. The smaller d, the more localized the particle position will be, and a measurement of the position operator will result in small variations across different measurements. On the other hand, the larger d the more delocalized it is, and an observation of the position operator will result in wildly different values for x at each measurement outcome.

The expectation values of the position operator is 0 for symmetry reasons (it is also just the mean of the Gaussian):

$$\begin{aligned} \langle \hat{x} \rangle &= \langle \Psi | \hat{x} | \Psi \rangle && (5.8.3)\\ &\int dx \left(x |\Psi(x)|^2 \right) && (5.8.4)\\ &= 0. && (5.8.5) \end{aligned}$$

The spread of the measurement of x can be quantified by the expectation value of $\hat{x}^2$, which in turn coincides with the variance of the gaussian:

$$\begin{aligned} \langle \hat{x}^2 \rangle &= \int dx \left(x^2 |\Psi(x)|^2 \right) && (5.8.6)\\ &= d^2. && (5.8.7) \end{aligned}$$

Thus the *intrinsic* uncertainty related to the probabilistic nature of the measurement process (quantum *noise*, if you wish) is given by

$$\begin{aligned} \langle \delta \hat{x}^2 \rangle &= \langle \hat{x}^2 \rangle - \langle \hat{x} \rangle^2 && (5.8.8)\\ &= d^2. && (5.8.9) \end{aligned}$$

Again this expression should be interpreted in terms of many repeated measurements on identical systems, imagining that each outcome for the measurement of x is recorded, and that in the limit of a large number of measurements the variance of the observed x approaches d^2.

The expectation value of the momentum operator is conveniently computed recalling that $\langle x|\hat{p}|\Psi\rangle = -i\hbar \frac{\partial}{\partial x}\Psi(x)$ thus

$$\begin{aligned} \langle \hat{p} \rangle &= \langle \Psi | \hat{p} | \Psi \rangle && (5.8.10)\\ &= -i\hbar \int dx \Psi(x)^\star \frac{\partial}{\partial x} \Psi(x) && (5.8.11)\\ &= \hbar k, && (5.8.12) \end{aligned}$$

the detailed derivation of the last line is left as an exercise. From this expression we see then that the parameter k also has a transparent physical meaning: it is the "average" wave

vector of the quantum particle described by this wave packet. One can also show that

$$\langle \hat{p}^2 \rangle = \frac{\hbar^2}{4d^2} + \hbar^2 k^2 \tag{5.8.13}$$

$$\langle \delta\hat{p}^2 \rangle = \frac{\hbar^2}{4d^2}. \tag{5.8.14}$$

From these two expressions we therefore see that the Gaussian wave packet *saturates* the Heisenberg indetermination principle:

$$\langle \delta\hat{x}^2 \rangle \langle \delta\hat{p}^2 \rangle \quad = \quad \frac{\hbar^2}{4}. \tag{5.8.15}$$

That is why the gaussian wave packet is often called the minimum uncertainty wave function, in the sense that it is not possible to find other states with less uncertainty, as quantified by the Heisenberg bound. This also tells us that if we are in a limit in which $d \to 0$ there will be huge indeterminacy in the value of the momentum, and vice versa. To form some intuition of this behavior, we can think that the more we try to spatially *squeeze* the particle, the "hotter" it gets, with its kinetic energy increasing. While useful, as all analogies with the classical world (in this case with thermodynamics) this analogy too should be taken with a grain of salt. In the quantum case there is absolutely no dynamics (yet) involved, and these rapid oscillations are just a result of the intrinsic probabilistic nature of quantum mechanics.

Exercise 8. Find the momentum-space representation of the gaussian wave-packet state.

5.8.1 Classical observer and precision

This result also also tells us that, as long as our classical observer has an experimental *resolution* (intrinsic precision of the instrument) on the momentum significantly worse than $\frac{\hbar^2}{4d^2}$ (i.e. they are not able to resolve features below that scale), and as long as the experimental resolution on the position is significantly worse than d^2, both position and momentum will *appear* to take constant values , every time their measurement is performed. This tells us that in order to see quantum mechanical effects, it is often the case that we need to go at scales (both in space and momentum) that have for long been not accessible to experimentalists before the beginning of the 20th century.

5.9 References and Further Reading

The discussion done in this Chapter is adapted from Sakurai's "Modern Quantum Mechanics" (Chapter 1, sections 1.6 and 1.7). A detailed treatment of the position and momentum representations can be found also in Cohen-Tannoudji's book (Chapter 2 in general, and also complement DII).

Chapter 6

The Quantum Harmonic Oscillator

In this Chapter we start our analysis of a number of simple Hamiltonians that are both practically and conceptually relevant to understand some of the emergent phenomena in quantum mechanics. The main task will be to solve the Schrodinger equation for these Hamiltonians, both in the static and in the time-dependent case.

6.1 Stationary States

We start our analysis with the simple harmonic oscillator in one dimension. The Hamiltonian in this case is given by

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2\hat{x}^2}{2}, \tag{6.1.1}$$

where the first term is the usual kinetic energy and the second one represents the energy of a "spring" connected to the mass m. While this is a highly idealized model, it is often very useful to understand molecular vibrations in solids or even the behavior of diluted atoms confined with light. It is therefore essential, for *any* physicist, to know how to solve this quantum Hamiltonian and understand its qualitative and quantitative features.

We start determining the eigenstates of the Hamiltonian. There are several techniques to derive them, in this Chapter we will follow the so-called "ladder method". We define for this purpose a pair of so-called annihilation and creation operators (for reason to be clarified later) with the following form :

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + \frac{i\hat{p}}{m\omega}\right) \tag{6.1.2}$$

$$\hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} - \frac{i\hat{p}}{m\omega}\right). \tag{6.1.3}$$

These are non-Hermitian operators (thus, they cannot be directly measured) however they are conceptually very useful to determine the eigenstates of our Hamiltonian. We first notice that

$$[\hat{a}, \hat{a}^\dagger] = \frac{m\omega}{2\hbar}\left[\hat{x} + \frac{i\hat{p}}{m\omega}, \hat{x} - \frac{i\hat{p}}{m\omega}\right] \tag{6.1.4}$$

$$= \frac{m\omega}{2\hbar}\frac{i}{m\omega}\left(-[\hat{x}, \hat{p}] + [\hat{p}, \hat{x}]\right) \tag{6.1.5}$$

$$= \frac{m\omega}{2\hbar}\frac{2\hbar}{m\omega} = 1. \tag{6.1.6}$$

We further define the so-called "number operator"

$$\hat{N} = \hat{a}^\dagger\hat{a}, \tag{6.1.7}$$

this is manifestly Hermitian, and it is directly connected to the Hamiltonian. We can see that

$$\begin{aligned}
\hat{N} &= \frac{m\omega}{2\hbar}\left(\hat{x}^2 - \frac{i}{m\omega}\hat{p}\hat{x} + \frac{i}{m\omega}\hat{x}\hat{p} + \frac{\hat{p}^2}{m^2\omega^2}\right) && (6.1.8)\\
&= \frac{m\omega}{2\hbar}\left(\frac{\hat{p}^2}{m^2\omega^2} + \hat{x}^2 - \frac{\hbar}{m\omega}\right) && (6.1.9)\\
&= \frac{1}{\hbar\omega}\left(\frac{\hat{p}^2}{2m} + \frac{m\omega^2}{2}\hat{x}^2\right) - \frac{1}{2} && (6.1.10)\\
&= \frac{\hat{H}}{\hbar\omega} - \frac{1}{2}. && (6.1.11)
\end{aligned}$$

Thus the Hamiltonian is (essentially, apart from scaling factors and a constant shift) identical to the number operator:

$$\hat{H} = \hbar\omega\left(\hat{N} + \frac{1}{2}\right), \quad (6.1.12)$$

therefore it is also true that $[\hat{H}, \hat{N}] = 0$ and that the eigenvalues and eigenvectors of $\hat{N}$ are essentially the same as those of $\hat{H}$. We denote these eigenvalues and eigenvectors with

$$\hat{N}|n\rangle = n|n\rangle, \quad (6.1.13)$$

and it will be shown in a moment that the eigenvalues n are positive integers. We start showing that $n \geq 0$, indeed we use the fact that

$$|\hat{a}|n\rangle|^2 \geq 0, \quad (6.1.14)$$

and the eigenvalue equation for the number operator

$$\begin{aligned}
|\hat{a}|n\rangle|^2 &= \langle n|\hat{a}^\dagger\hat{a}|n\rangle && (6.1.15)\\
&= n && (6.1.16)\\
&\geq 0. && (6.1.17)
\end{aligned}$$

It is also useful to derive the commutation relations satisfied between the number operator and the creation/annihilation operators:

$$\begin{aligned}
[\hat{N}, \hat{a}^\dagger] &= [\hat{a}^\dagger\hat{a}, \hat{a}^\dagger] && (6.1.18)\\
&= \hat{a}^\dagger, && (6.1.19)
\end{aligned}$$

and

$$\begin{aligned}
[\hat{N}, \hat{a}] &= [\hat{a}^\dagger\hat{a}, \hat{a}] && (6.1.20)\\
&= -\hat{a}. && (6.1.21)
\end{aligned}$$

With these in hand, we are now ready to show that the states $a^\dagger|n\rangle$ and $\hat{a}|n\rangle$ are eigenstates of $\hat{N}$ with eigenvalues $n+1$ and $n-1$ respectively. We see this directly:

$$\begin{aligned}
\hat{N}(\hat{a}^\dagger|n\rangle) &= [\hat{N}, \hat{a}^\dagger]|n\rangle + \hat{a}^\dagger\hat{N}|n\rangle && (6.1.22)\\
&= \hat{a}^\dagger|n\rangle + n\hat{a}^\dagger|n\rangle && (6.1.23)\\
&= (n+1)\hat{a}^\dagger|n\rangle, && (6.1.24)
\end{aligned}$$

and

$$\begin{aligned}
\hat{N}(\hat{a}|n\rangle) &= [\hat{N}, \hat{a}]|n\rangle + \hat{a}\hat{N}|n\rangle && (6.1.25)\\
&= -\hat{a}|n\rangle + n\hat{a}|n\rangle && (6.1.26)\\
&= (n-1)\hat{a}|n\rangle. && (6.1.27)
\end{aligned}$$

We thus see that

$$\hat{a}^\dagger|n\rangle = \alpha_n|n+1\rangle \tag{6.1.28}$$
$$\hat{a}|n\rangle = \beta_n|n-1\rangle, \tag{6.1.29}$$

where the constants α and β can be determined imposing the normalization conditions

$$\langle n+1|n+1\rangle = \frac{\langle n|\hat{a}\hat{a}^\dagger|n\rangle}{|\alpha_n|^2} \tag{6.1.30}$$
$$= \frac{\langle n|\hat{N}+[\hat{a},\hat{a}^\dagger]|n\rangle}{|\alpha_n|^2} \tag{6.1.31}$$
$$= \frac{n+1}{|\alpha_n|^2}, \tag{6.1.32}$$

thus $|\alpha_n|^2 = n+1$. With a similar reasoning we can determine the other normalization constant

$$\langle n-1|n-1\rangle = \frac{\langle n|\hat{a}^\dagger\hat{a}|n\rangle}{|\beta_n|^2} \tag{6.1.33}$$
$$= \frac{\langle n|\hat{N}|n\rangle}{|\beta_n|^2} \tag{6.1.34}$$
$$= \frac{n}{|\beta_n|^2}. \tag{6.1.35}$$

For simplicity, we thus take $\alpha_n = \sqrt{n+1}$ and $\beta_n = \sqrt{n}$, and we have the fundamental equations characterizing our creation and annihilation operators:

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \tag{6.1.36}$$
$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle. \tag{6.1.37}$$

These two equations also imply that n must a positive integer. If this was not the case, by repeated applications of the destruction operator we could find a negative eigenvalue, which is not possible, as we have already demonstrated. If instead $n \geq 0$, then the destruction operator terminates with the null state $|0\rangle$, corresponding to $n = 0$. The sequence of eigenstates is then simply given as

$$|0\rangle \tag{6.1.38}$$
$$|1\rangle = a^\dagger|0\rangle \tag{6.1.39}$$
$$|2\rangle = \frac{a^\dagger}{\sqrt{2}}|1\rangle = \frac{(a^\dagger)^2}{\sqrt{2}}|0\rangle \tag{6.1.40}$$
$$|3\rangle = \frac{a^\dagger}{\sqrt{3}}|2\rangle = \frac{(a^\dagger)^3}{\sqrt{6}}|0\rangle \tag{6.1.41}$$
$$|n\rangle = \frac{a^\dagger}{\sqrt{n}}|n-1\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle, \tag{6.1.42}$$

the eigenvalues of the Hamiltonian are instead simply

$$E(n) = \hbar\omega\left(n+\frac{1}{2}\right), \tag{6.1.43}$$

thus since n is an integer, we see once more a *quantization* of the energy levels, whereas in the classical case we would expect the internal energy of the harmonic oscillator to be continuous.

6.1.1 Eigenstates in real space

We have now determined the full spectrum of eigenvalues and eigenvectors of the harmonic oscillator hamiltonian, however we haven't really explicitly found the functional form for the eigenstates. To this end, we need to specify a convenient basis. We will now consider the position basis, since it is quite natural to reason in terms of position amplitudes, at least to start developing some intuition about quantum behavior. Consider for example the ground state amplitude

$$\langle x|0\rangle \quad = \quad \psi_0(x), \tag{6.1.44}$$

we can solve for the function $\psi_0(x)$ using the fact that

$$\hat{a}|0\rangle \quad = \quad 0, \tag{6.1.45}$$

and recalling the definition of the destruction operator in terms of position and momentum operators: $\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + \frac{i\hat{p}}{m\omega}\right)$, thus

$$\langle x|\hat{a}|0\rangle \quad = \quad \sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{\hbar}{m\omega}\frac{\partial}{\partial x}\right)\psi_0(x) = 0. \tag{6.1.46}$$

We can also define a convenient length scale $x_0 \equiv \sqrt{\hbar/m\omega}$, and explicitly solve this differential equation. It is easily verified that the normalized ground state wave function that solves the differential equation reads

$$\psi_0(x) \quad = \quad \frac{1}{\pi^{1/4}\sqrt{x_0}}\exp\left[-\frac{1}{2}\left(\frac{x}{x_0}\right)^2\right], \tag{6.1.47}$$

which is a special case of the Gaussian wave packed we have previously analyzed, Eq. (5.8.1).

6.1.1.1 Uncertainty product

In that context, we have already evaluated expectation values of physical quantities over this type of wave function, however it is also interesting to see how one can compute expectation values of functions of $\hat{x}$ and $\hat{p}$ using the creation and annihilation operators $\hat{a}$ and $\hat{a}^\dagger$. Considering again the definition of these operators,

$$\hat{a} \quad = \quad \sqrt{\frac{1}{2}}x_0\left(\hat{x} + \frac{i\hat{p}}{m\omega}\right) \tag{6.1.48}$$

$$\hat{a}^\dagger \quad = \quad \sqrt{\frac{1}{2}}x_0\left(\hat{x} - \frac{i\hat{p}}{m\omega}\right). \tag{6.1.49}$$

We have that

$$\hat{x} \quad = \quad \frac{\sqrt{2}}{2}x_0\left(\hat{a} + \hat{a}^\dagger\right) \tag{6.1.50}$$

$$\hat{p} \quad = \quad \frac{\sqrt{2}}{2i}\frac{\hbar}{x_0}\left(\hat{a} - \hat{a}^\dagger\right). \tag{6.1.51}$$

From these expressions it immediately follows that

$$\langle 0|\hat{x}|0\rangle \quad = \quad 0 \tag{6.1.52}$$

$$\langle 0|\hat{p}|0\rangle \quad = \quad 0, \tag{6.1.53}$$

since, in general, only expectations of operators containing products of an equal number of $\hat{a}$ and $\hat{a}^\dagger$ will be different from zero. The situation is different for higher orders of the particle coordinates and momenta, since

$$\hat{x}^2 \quad = \quad \frac{1}{2}x_0^2\left(\hat{a}\hat{a} + \hat{a}^\dagger\hat{a}^\dagger + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\right), \tag{6.1.54}$$

yielding

$$\langle 0|\hat{x}^2|0\rangle = \frac{x_0^2}{2}\langle 0|\hat{a}\hat{a} + \hat{a}^\dagger\hat{a}^\dagger + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}|0\rangle \tag{6.1.55}$$

$$= \frac{x_0^2}{2}\langle 0|\hat{a}\hat{a}^\dagger|0\rangle \tag{6.1.56}$$

$$= \frac{x_0^2}{2}\langle 0|\hat{1} + \hat{N}|0\rangle \tag{6.1.57}$$

$$= \frac{x_0^2}{2}, \tag{6.1.58}$$

as expected then x_0 plays the role of the variance of the density distribution. Similarly,

$$\hat{p}^2 = -\frac{1}{2}\frac{\hbar^2}{x_0^2}\left(\hat{a}\hat{a} + \hat{a}^\dagger\hat{a}^\dagger - \hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\right), \tag{6.1.59}$$

thus

$$\langle 0|\hat{p}^2|0\rangle = -\frac{1}{2}\frac{\hbar^2}{x_0^2}\langle 0|\hat{a}\hat{a} + \hat{a}^\dagger\hat{a}^\dagger - \hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a}|0\rangle \tag{6.1.60}$$

$$= \frac{1}{2}\frac{\hbar^2}{x_0^2}\langle 0|\hat{a}\hat{a}^\dagger|0\rangle \tag{6.1.61}$$

$$= \frac{1}{2}\frac{\hbar^2}{x_0^2}\langle 0|\hat{1} + \hat{N}|0\rangle \tag{6.1.62}$$

$$= \frac{1}{2}\frac{\hbar^2}{x_0^2}, \tag{6.1.63}$$

$$= \frac{1}{2}\hbar m\omega. \tag{6.1.64}$$

It then follows that

$$\langle 0|\Delta x^2|0\rangle\langle 0|\Delta p^2|0\rangle = \frac{\hbar^2}{4}, \tag{6.1.65}$$

thus the ground state of the harmonic oscillator is also the state with minimum uncertainty.

6.1.2 Excited States

With the explicit form of the ground-state wave function given by Eq. (6.1.47), we can use the creation operators to generate also explicit expressions for the excited (higher-energy) states. For example,

$$\psi_1(x) = \langle x|\hat{a}^\dagger|0\rangle \tag{6.1.66}$$

$$= \frac{1}{\sqrt{2}x_0}\langle x|\left(\hat{x} - \frac{i\hat{p}}{m\omega}\right)|0\rangle \tag{6.1.67}$$

$$= \frac{1}{\sqrt{2}x_0}\left[x\psi_0(x) - x_0^2\frac{\partial}{\partial x}\psi_0(x)\right] \tag{6.1.68}$$

$$= \frac{1}{\sqrt{2}x_0}\psi_0(x)\left[x - x_0^2\frac{\partial}{\partial x}\log\psi_0(x)\right] \tag{6.1.69}$$

$$= \frac{1}{\sqrt{2}}2\frac{x}{x_0}\psi_0(x) \tag{6.1.70}$$

$$= \frac{1}{\pi^{1/4}\sqrt{2x_0}}2\left(\frac{x}{x_0}\right)\exp\left[-\frac{1}{2}\left(\frac{x}{x_0}\right)^2\right] \tag{6.1.71}$$

Higher excited states can be found similarly, by repeated application of the $\hat{a}^\dagger$ operator. We report here the result for completeness

$$\psi_n(x) = \langle x|\frac{\left(\hat{a}^\dagger\right)^n}{n!}|0\rangle \tag{6.1.72}$$

$$= \left(\frac{1}{\pi^{1/4}\sqrt{2^n n!}x_0^{n+\frac{1}{2}}}\right)\left(x - x_0^2\frac{\partial}{\partial x}\right)^n \exp\left[-\frac{1}{2}\left(\frac{x}{x_0}\right)^2\right]. \tag{6.1.73}$$

Other equivalent expressions for the excited states, in terms of Hermite polynomials, can be found in textbooks.

6.2 Time evolution of the Harmonic Oscillator

In order to evaluate the time evolution induced by the harmonic oscillator hamiltonian, it is convenient to work in the Heisenberg picture. We have:

$$\begin{aligned}\frac{d}{dt}\hat{x}_h(t) &= \frac{i}{\hbar}[\hat{H}, \hat{x}_h(t)]\\ \frac{d}{dt}\hat{p}_h(t) &= \frac{i}{\hbar}[\hat{H}, \hat{p}_h(t)],\end{aligned}$$

and evaluating the commutators:

$$[\hat{H}, \hat{x}_h(t)] = [\hat{H}, e^{\frac{i}{\hbar}t\hat{H}}\hat{x}e^{-\frac{i}{\hbar}t\hat{H}}] \tag{6.2.1}$$

$$= \hat{H}e^{\frac{i}{\hbar}t\hat{H}}\hat{x}e^{-\frac{i}{\hbar}t\hat{H}} - e^{\frac{i}{\hbar}t\hat{H}}\hat{x}e^{-\frac{i}{\hbar}t\hat{H}}\hat{H} \tag{6.2.2}$$

$$= e^{\frac{i}{\hbar}t\hat{H}}\hat{H}\hat{x}e^{-\frac{i}{\hbar}t\hat{H}} - e^{\frac{i}{\hbar}t\hat{H}}\hat{x}\hat{H}e^{-\frac{i}{\hbar}t\hat{H}} \tag{6.2.3}$$

$$= e^{\frac{i}{\hbar}t\hat{H}}[\hat{H}, \hat{x}]e^{-\frac{i}{\hbar}t\hat{H}} \tag{6.2.4}$$

$$= e^{\frac{i}{\hbar}t\hat{H}}[\frac{\hat{p}^2}{2m}, \hat{x}]e^{-\frac{i}{\hbar}t\hat{H}} \tag{6.2.5}$$

$$= -e^{\frac{i}{\hbar}t\hat{H}}i\frac{\hbar}{m}\hat{p}e^{-\frac{i}{\hbar}t\hat{H}} \tag{6.2.6}$$

$$= -i\frac{\hbar}{m}\hat{p}_h(t), \tag{6.2.7}$$

and similarly

$$[\hat{H}, \hat{p}_h(t)] = e^{\frac{i}{\hbar}t\hat{H}}[\hat{H}, \hat{p}]e^{-\frac{i}{\hbar}t\hat{H}} \tag{6.2.8}$$

$$= e^{\frac{i}{\hbar}t\hat{H}}[m\omega^2\frac{\hat{x}^2}{2}, \hat{p}]e^{-\frac{i}{\hbar}t\hat{H}} \tag{6.2.9}$$

$$= m\omega^2 i\hbar\hat{x}_h(t). \tag{6.2.10}$$

Thus

$$\frac{d}{dt}\langle x(t)\rangle = \frac{1}{m}\langle p(t)\rangle \tag{6.2.11}$$

$$\frac{d}{dt}\langle p(t)\rangle = -m\omega^2\langle x(t)\rangle. \tag{6.2.12}$$

These equations for the expectation values are completely equivalent to the classical expressions, since in the classical case:

$$\frac{d}{dt}x(t) = \frac{p(t)}{m}, \tag{6.2.13}$$

$$\frac{d}{dt}p(t) = -\frac{\partial}{\partial x}V(x) \tag{6.2.14}$$

$$= -m\omega^2 x(t). \tag{6.2.15}$$

We therefore see that the dynamics of the quantum expectation values is given by expressions that are the same as for the classical values, with oscillations of frequency ω:

$$\langle x(t)\rangle = \langle x(0)\rangle \cos\omega t + \frac{\langle p(0)\rangle}{\omega m}\sin\omega t \tag{6.2.16}$$

$$\langle p(t)\rangle = \langle p(0)\rangle \cos\omega t - m\omega\langle x(0)\rangle \sin\omega t. \tag{6.2.17}$$

While these expressions are the same as their classical counterparts, it is important to realize that the expectation values of the position and momentum will not *always* oscillate with frequency ω. For example, analogously to what what we have seen for the ground state wave function, for all the eigen-states of the harmonic oscillator we have:

$$\langle n|\hat{x}|n\rangle = 0 \tag{6.2.18}$$

$$\langle n|\hat{p}|n\rangle = 0, \tag{6.2.19}$$

thus if, for example, at time $t = 0$ our system is in one of the eigenstates $|n\rangle$, then the expectation values just stay equal to zero at all times $\langle x\rangle(t) = \langle p\rangle(t) = 0$. In order to see oscillations, one needs to prepare the initial state in, at least, a superposition of two distinct eigenstates. For example:

$$|\Psi(0)\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right),$$

is such that $\hat{x} = \frac{\sqrt{2}}{2}x_0\left(\hat{a} + \hat{a}^\dagger\right)\hat{p} = \frac{\sqrt{2}}{2i}\frac{\hbar}{x_0}\left(\hat{a} - \hat{a}^\dagger\right)$.

$$\begin{aligned}\langle\Psi(0)|\hat{x}|\Psi(0)\rangle &= \frac{1}{2}\left(\langle 0| + \langle 1|\right)\frac{\sqrt{2}}{2}x_0\left(\hat{a} + \hat{a}^\dagger\right)\left(|0\rangle + |1\rangle\right)\\ &= \frac{\sqrt{2}}{4}x_0\left[\langle 0|\hat{a}|1\rangle + \langle 1|\hat{a}^\dagger|0\rangle\right]\\ &= \frac{\sqrt{2}}{2}x_0,\end{aligned}$$

and

$$\begin{aligned}\langle\Psi(0)|\hat{p}|\Psi(0)\rangle &= \frac{1}{2}\left(\langle 0| + \langle 1|\right)\frac{\sqrt{2}}{2i}\frac{\hbar}{x_0}\left(\hat{a} - \hat{a}^\dagger\right)\left(|0\rangle + |1\rangle\right)\\ &= \frac{\sqrt{2}}{4i}\frac{\hbar}{x_0}\left[\langle 0|\hat{a}|1\rangle - \langle 1|\hat{a}^\dagger|0\rangle\right]\\ &= 0,\end{aligned}$$

thus

$$\langle x(t)\rangle = \sqrt{\frac{\hbar}{2m\omega}}\cos\omega t \tag{6.2.20}$$

$$\langle p(t)\rangle = -\sqrt{\frac{\hbar m\omega}{2}}\sin\omega t. \tag{6.2.21}$$

6.3 Ehrenfest's Theorem

In the previous section we have explicitly solved the Heisenberg's equations of motion, and found out that the dynamics of the expectation values of positions and momenta for the harmonic oscillator follow the classical equations of motion. This result is not a coincidence, and it is actually a consequence of a deeper and more general result. Through Heisenberg's equations of motion we can demonstrate the connection between quantum expectation values and classical equations of motion, for a larger class of Hamiltonians. This connection, knows as Ehrenfest's Theorem concerns the time dependence of expectation values of general continous-space Hamiltonians that depend on momentum and position:

$$\hat{H}(\hat{x},\hat{p}) = \frac{\hat{p}^2}{2m} + V(\hat{x}). \tag{6.3.1}$$

The time-dependent expectation value of the position can be found using Heisenberg's equations of motion we have previously derived:

$$\begin{aligned}
\frac{d}{dt}\langle\hat{x}\rangle &= \frac{i}{\hbar}\langle[\hat{H},\hat{x}]\rangle && (6.3.2)\\
&= \frac{i}{\hbar}\frac{1}{2m}\langle[\hat{p}^2,\hat{x}]\rangle && (6.3.3)\\
&= \frac{i}{\hbar}\frac{1}{2m}\left\{\langle\hat{p}[\hat{p},\hat{x}]\rangle + \langle[\hat{p},\hat{x}]\hat{p}\rangle\right\} && (6.3.4)\\
&= \frac{i}{\hbar}\frac{1}{2m}\left\{-2i\hbar\langle\hat{p}\rangle\right\} && (6.3.5)\\
&= \frac{\langle\hat{p}\rangle}{m}. && (6.3.6)
\end{aligned}$$

We therefore see that, for what averages are concerned, this is exactly the result we would expect for a classical system. Similarly, we can compute the time-dependence of the momentum operator

$$\begin{aligned}
\frac{d}{dt}\langle\hat{p}\rangle &= \frac{i}{\hbar}\langle[\hat{H},\hat{p}]\rangle && (6.3.7)\\
&= \frac{i}{\hbar}\langle[V(\hat{x}),\hat{p}]\rangle && (6.3.8)\\
&= -\left\langle\frac{\partial}{\partial x}V(x)\right\rangle && (6.3.9)
\end{aligned}$$

And see that it also coincides, in expectation, with what you would expect from Newton's equation. In order to obtain this commutator explicitly, we have used these very general commutator results:

Theorem 9. *The following commutator relations holds:* $[\hat{x}^n,\,\hat{p}] = i\hbar n\hat{x}^{n-1}$

Proof. The proof is obtained by recursive application of the elementary commutation relation. Since

$$\hat{C}(n) = [\hat{x}^n,\hat{p}] = \hat{x}^{n-1}[\hat{x},\hat{p}] + [\hat{x}^{n-1},\hat{p}]\hat{x}, \tag{6.3.10}$$

we see then that

$$\begin{aligned}
\hat{C}(n) &= i\hbar\hat{x}^{n-1} + \hat{C}(n-1)\hat{x} && (6.3.11)\\
&= i\hbar\hat{x}^{n-1} + i\hbar\hat{x}^{n-1} + \hat{C}(n-2)\hat{x}^2 && (6.3.12)\\
&= (n-1)i\hbar\hat{x}^{n-1} + \hat{C}(1)\hat{x}^{n-1}, && (6.3.13)
\end{aligned}$$

but since $\hat{C}(1) = [\hat{x},\hat{p}] = i\hbar$, it follows that $\hat{C}(n) = i\hbar n\hat{x}^{n-1}$. □

Corollary 10. *For regular functions of the coordinates, we have* $[V(\hat{x}),\,\hat{p}] = i\hbar V'(\hat{x})$

Proof. The proof is obtained considering the Taylor series of the function $V(\hat{x}) = \sum_{n=0}^{\infty}\hat{x}^n g_n$, and its derivative $V'(\hat{x}) = \sum_{n=0}^{\infty} n\hat{x}^{n-1}g_n$. Thus

$$\begin{aligned}
\left[\sum_{n=0}^{\infty}\hat{x}^n g_n,\hat{p}\right] &= \sum_{n=0}^{\infty} g_n[\hat{x}^n,\hat{p}] && (6.3.14)\\
&= i\hbar\sum_{n=0}^{\infty} g_n n\hat{x}^{n-1} && (6.3.15)\\
&= i\hbar V'(\hat{x}). && (6.3.16)
\end{aligned}$$

□

Chapter 7

Wave Mechanics

In the previous Chapter, we have seen how to solve a paradigmatic model (the harmonic oscillator) using an algebraic technique based (essentially) only on commutators, both for the stationary states and for the dynamics of the system. In the Chapter we explore instead more directly Schroedinger's equation in real space.

7.1 Schroedinger's wave equation

We have seen that the Hamiltonian of a particle subjected to an external potential reads

$$\hat{H} = \frac{\hat{p}^2}{2m} + \widehat{V(x)}. \tag{7.1.1}$$

In the following we look into solutions of Schroedinger's equation in position representation, thus time-dependent amplitudes:

$$\langle x|\Psi(t)\rangle \equiv \Psi(x,t). \tag{7.1.2}$$

The amplitudes satisfy the usual equation:

$$i\hbar\frac{\partial}{\partial t}\langle x|\Psi(t)\rangle = \langle x|\hat{H}|\Psi(t)\rangle, \tag{7.1.3}$$

and recalling that

$$\hat{p}^2 = \left(-i\hbar\frac{\partial}{\partial x}\right)^2 \tag{7.1.4}$$

$$= -\hbar^2\frac{\partial^2}{\partial x^2}, \tag{7.1.5}$$

we have the famous Schroedinger's wave equation:

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x,t) + V(x)\Psi(x,t). \tag{7.1.6}$$

As we have discussed in the Chapter on time evolution, the solution to this equation takes the formal solution:

$$\langle x|\Psi(t)\rangle = \langle x|e^{-i\frac{\hat{H}}{\hbar}t}|\Psi(0)\rangle \tag{7.1.7}$$

where we have used the fact that the Hamiltonian is time independent. Since the exponential of the Hamiltonian acts trivially on the eigenstates of the Hamiltonian $|\phi_E\rangle$:

$$e^{-\frac{i}{\hbar}\hat{H}t}|\phi_E\rangle = e^{-\frac{i}{\hbar}Et}|\phi_E\rangle, \tag{7.1.8}$$

it is very useful to solve for the them explicitly, giving rise to the so-called time-independent Schroedinger's equation:

$$\langle x|\hat{H}|\phi_E\rangle = E\langle x|\phi_E\rangle, \tag{7.1.9}$$

which in functional form is equivalently written as the following differential equation:

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\phi_E(x) + V(x)\phi_E(x) = E\phi_E(x). \tag{7.1.10}$$

The generalization of these equations to 3 dimensions is straightforward, as it is sufficient to recall that the kinetic energy in 3 dimensions is the sum of the three cartesian components, thus

$$\hat{H} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + V(x,y,z) \tag{7.1.11}$$

$$= -\frac{\hbar^2}{2m}\nabla^2_{\mathbf{r}} + V(\mathbf{r}), \tag{7.1.12}$$

and a 3-dimensional wave function $\Psi(\mathbf{r},t)$ satisfies the corresponding Schroedinger's equation with this Hamiltonian.

7.1.1 Probability flux

We have seen that a fundamental postulate of quantum mechanics is that physical wave functions are always normalized, thus:

$$\langle\Psi(t)|\Psi(t)\rangle = 1. \tag{7.1.13}$$

Similarly to the dynamics of fluids, this conservation law implies the existence of a continuity equation associated to the conservation of the integral of the probability density, $\rho(x,t) = |\Psi(x,t)|^2$. In order to find out explicitly the form of the continuity equation, we work out the time derivative of the probability density:

$$\frac{\partial}{\partial t}|\Psi(x,t)|^2 = \left(\frac{\partial}{\partial t}\Psi(x,t)^\star\right)\Psi(x,t) + \Psi(x,t)^\star\left(\frac{\partial}{\partial t}\Psi(x,t)\right), \tag{7.1.14}$$

and substituting the time derivative of the wave function with Schroedinger's equation we have

$$\frac{\partial}{\partial t}|\Psi(x,t)|^2 = \frac{1}{i\hbar}\left[\left(\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi^\star\right)\Psi + \Psi^\star\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi\right)\right] \tag{7.1.15}$$

$$= -\frac{i\hbar}{2m}\left[\left(\frac{\partial^2}{\partial x^2}\Psi^\star\right)\Psi - \Psi^\star\left(\frac{\partial^2}{\partial x^2}\Psi\right)\right] \tag{7.1.16}$$

$$= -\frac{i\hbar}{2m}\frac{\partial}{\partial x}\left[\left(\frac{\partial}{\partial x}\Psi^\star\right)\Psi - \Psi^\star\left(\frac{\partial}{\partial x}\Psi\right)\right]. \tag{7.1.17}$$

We then have the following continuity equation

$$\frac{\partial}{\partial t}\rho(x,t) + \frac{\partial}{\partial x}J(x,t) = 0, \tag{7.1.18}$$

where have defined the probability *current*

$$J(x,t) = \frac{i\hbar}{2m}\left[\left(\frac{\partial}{\partial x}\Psi^\star(x,t)\right)\Psi(x,t) - \Psi(x,t)^\star\left(\frac{\partial}{\partial x}\Psi(x,t)\right)\right]. \tag{7.1.19}$$

Equation (7.1.18) is a continuity equation in the sense that it is similar to what is used in hydrodynamics to express the conservation of mass, whereas in this context it is derived

assuming conservation of probability. The continuity equation also expresses a local conservation law for the probability. If we integrate the square modulus of the wave function in a finite interval $[a.b]$, we have

$$\frac{\partial}{\partial t}\int_a^b dx|\Psi(x,t)|^2 \quad = \quad J(a,t) - J(b,t), \tag{7.1.20}$$

thus the probability in the region delimited by a and b increases or decreases as a result of the flux differences at the two extremes.

Since the probability current expresses the flow of probability density in time, we might intuitively expect that it is also related to the momentum of the particle, as much the mass current in a fluid is related to the velocity of the fluid. In his first works, Schroedinger was actually led to conjecture that at each point in space one could identify this current with the actual *velocity field* of the particles, however this leads to complications in the theory and contradictions with the experiments. The connection between the probability flux and the momentum is only valid *on average*, i.e. when considering expectation values of the momentum operator. To show this, we recall that the average momentum is given by:

$$\langle p(t)\rangle \quad = \quad -i\hbar\int dx\Psi^\star(x,t)\frac{\partial\Psi(x,t)}{\partial x} \tag{7.1.21}$$

$$\quad = \quad i\hbar\int dx\Psi(x,t)\frac{\partial\Psi^\star(x,t)}{\partial x}, \tag{7.1.22}$$

where in the last line we have used the fact that the expectation value of the operator is always real, thus $\langle p\rangle(t) = \langle p\rangle^\star(t)$. We can then conclude that

$$\int dxJ(x,t) \quad = \quad \frac{\langle p(t)\rangle}{m}, \tag{7.1.23}$$

thus showing that only on average the current is equal to the particle's velocity.

7.2 General Properties of Wave Functions

We now want to establish a few general properties of the solutions of Schroedinger's wave equation. We concentrate, for simplicity, on the one-dimensional case, but all the results presented here are easily generalized to three dimensions.

7.2.1 Normalization

The first condition we have already seen, is that physically valid wave functions must be normalizable, thus

$$\int_{-\infty}^{+\infty} dx|\Psi(x)|^2 \quad = \quad 1, \tag{7.2.1}$$

which is a strict requirement due to the probabilistic interpretation of the wave-function squared. It should be remarked, however, that Schroedinger's equation generally admits both normalizable solutions and un-normalizable solutions. In fact, the only requirement we asked for eigen-kets of continuous variables is

$$\langle\xi|\xi'\rangle \quad = \quad \delta(\xi-\xi') \tag{7.2.2}$$

which is the completeness requirement of the basis but does not imply the normalization condition of the eigen-kets (for discrete variables, instead, the two conditions are the same). An important consequence is that, for example, the eigen-kets of the momentum are not physically valid quantum states, since they are not normalizable! This is because they correspond to the eigen-kets of the Hamiltonian of a free particle

$$\hat{H} \quad = \quad \hat{p}^2/2m, \tag{7.2.3}$$

whose probability density of being in a certain position of space is expected to be independent on the specific position. In turn, this implies that the particle, in principle, could be *anywhere* in the universe with uniform probability. This is clearly an absurd requirement, and it is also the reason why the wave function is not normalizable. In general, however, free particles do not exist in Nature, since they are always confined by some external potential or by the interactions with other particles. In this sense, we can always think that there is an intrinsic scale L beyond which ($|x| > L$) it is essentially impossible to find the particle. This finite length scale makes the wave function normalizable, as we will also show in the following examples.

While some solutions of the Schroedinger equation are not normalizable, thus unphysical, they are still very much useful to analyze the dynamics of physical wave functions.

7.2.2 Continuity

On physical grounds, it is reasonable to expect that the probability density is a continuous function of x, thus $\Psi(x)$ is also expected to be continuous. From a mathematical point of view, if the potential $V(x)$ is bound and analytic, the solutions of the Schroedinger's equation must also be continuous and analytical. The situation is a bit different if the potential is singular. Roughly speaking, however, a discontinuity in the wave function can lead to a finite energy E only if it is *compensated* by a potential $V(x) \sim \delta'(x)$ that cancels out the singularity due to the kinetic energy at the discontinuity. For example, imagine that around $x \sim 0$ the wave function has a sign-like discontinuity $\Psi(x) = \text{sign}(x)$, then around the origin:

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x) = -\frac{\hbar^2}{m}\delta'(x), \tag{7.2.4}$$

thus only if the potential is $V(x) = \frac{\hbar^2}{2m}\delta'(x)$ can cancel out this singularity and lead to a finite energy E.

7.2.3 Continuity conditions of the Derivative

In order to establish wether the first derivative $\Psi'(x)$ of the wave function is continuous or not at some point $x = a$, we integrate the Schroedinger's equation around that region with a small ϵ:

$$\int_{a-\epsilon}^{a+\epsilon} dx \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x) + V(x)\Psi(x)\right] = E\int_{a-\epsilon}^{a+\epsilon} dx\Psi(x) \tag{7.2.5}$$

$$-\frac{\hbar^2}{2m}\left[\Psi'(a+\epsilon) - \Psi'(a-\epsilon)\right] + \int_{a-\epsilon}^{a+\epsilon} V(x)\Psi(x)dx = \mathcal{O}(\epsilon). \tag{7.2.6}$$

From this equation we see that there are, essentially, two important sub-cases.

1. $V(a)$ is finite, thus $\int_{a-\epsilon}^{a+\epsilon} V(x)\Psi(x)dx = \mathcal{O}(\epsilon)$ and we can conclude that $\Psi'(a+\epsilon) = \Psi'(a-\epsilon) + \mathcal{O}(\epsilon)$, thus the first derivative is continuous in $x = a$.

2. $V(a)$ is infinite, singular, etc (for example, $V(x) = \delta(x-a)$), then the first derivatives is not continuous, and the discontinuity in the derivative is fixed by the equation above, thus

$$\Psi'(a+\epsilon) - \Psi'(a-\epsilon) = \frac{2m}{\hbar^2}\int_{a-\epsilon}^{a+\epsilon} V(x)\Psi(x)dx. \tag{7.2.7}$$

   We have therefore seen that a discontinuity in the wave function appears if the potential, for example, $V(x) \sim \delta(x)$. Later on we will see an example application with this idealized, yet reasonable potential.

7.2.4 Minimum Allowed Energy

The energy eigenvalue E can, in general, be an arbitrary number. However, for a given potential $V(x)$ there is a minimum energy value that can be taken. Specifically, we must have

$$\begin{aligned} E &\geq \min_x V(x) && (7.2.8)\\ &= V_{\min}. && (7.2.9) \end{aligned}$$

This condition follows from the fact that the Hamiltonian contains two terms, and the first one (the kinetic energy) is positive definite. In order to see this, consider

$$|\Phi\rangle = \hat{p}|\Psi\rangle, \tag{7.2.10}$$

then

$$\begin{aligned} \langle\Phi|\Phi\rangle &\geq 0 && (7.2.11)\\ \langle\Psi|\hat{p}^2|\Psi\rangle &\geq 0. && (7.2.12) \end{aligned}$$

Thus for an eigen-ket of energy E we have

$$\begin{aligned} E &= \langle E|\frac{\hat{p}^2}{2m} + V(x)|E\rangle && (7.2.13)\\ &\geq \langle\Psi|V(x)|\Psi\rangle && (7.2.14)\\ &\geq V_{\min}. && (7.2.15) \end{aligned}$$

Another very important bound, known as variational bound, is found when considering the expectation value of the Hamiltonian over a generic state. In this case, and assuming that the Hamiltonian has a discrete spectrum ($E_0 \leq E_1 \leq E_2 \dots$) for simplicity, we have:

$$\begin{aligned} \langle H\rangle &= \langle\Psi|\hat{H}|\Psi\rangle && (7.2.16)\\ &= E_0|\langle E_0|\Psi\rangle|^2 + E_1|\langle E_1|\Psi\rangle|^2 + \dots && (7.2.17)\\ &\geq E_0|\langle E_0|\Psi\rangle|^2 && (7.2.18)\\ &\geq E_0. && (7.2.19) \end{aligned}$$

Therefore the expectation value of the energy on an arbitrary state cannot be lower than the ground-state energy, E_0, and it is exactly equal to E_0 only if $|\Psi\rangle = |E_0\rangle$.

7.2.5 Bound states and scattering eigenstates

The solutions of the time-independent Schroedinger equation, $|E\rangle$, can be generally classified into two kind of states, depending on the value of the energy. The main criterion is

$$\begin{cases} E < V(\pm\infty) & \text{, bound state} \\ E > V(\pm\infty) & \text{, scattering state} \end{cases} \tag{7.2.20}$$

thus the type of possible eigenstates depends on the type of potential considered. Physically speaking, bound states correspond to cases when the wave-function goes to zero in the limit $x \to \pm\infty$, thus we say that the particle is *bound* inside a certain finite region of space. The harmonic oscillator is a clear case where this happens, since there $V(\pm\infty) = \pm\infty$, and we have seen that the wave-function vanishes exponentially when approaching $x \to \pm\infty$.

Scattering states are a very different *beast*, since in general they do not correspond to physical states. The reason is that, since they are not required to vanish at $x \to \pm\infty$, they are not normalizable. Nonetheless, albeit not being physical states, their are still valid solutions of the time-independent Schroedinger equation and, as we will see in the following, they play an important mathematical role in solving the time-dependent Schroedinger equation.

7.2.6 Parity operator

We define the parity operator

$$\hat{\Pi}|x\rangle = |-x\rangle, \tag{7.2.21}$$

which transforms coordinates into their negative values, thus acting on wave-functions as:

$$\langle x|\hat{\Pi}|\Psi\rangle = \Psi(-x). \tag{7.2.22}$$

The eigenvalues λ of the parity operator and the corresponding eigen-kets $|u_\lambda\rangle$ are found easily, since:

$$\begin{aligned}
\langle x|\hat{\Pi}|u_\lambda\rangle &= \lambda\langle x|u_\lambda\rangle && (7.2.23)\\
\langle x|\hat{\Pi}\hat{\Pi}|u_\lambda\rangle &= \lambda\langle x|\hat{\Pi}|u_\lambda\rangle && (7.2.24)\\
\langle x|u_\lambda\rangle &= \lambda^2\langle x|u_\lambda\rangle, && (7.2.25)
\end{aligned}$$

thus $\lambda = \pm 1$, and the eigenfunctions with positive/negative eigenvalue are all the even/odd functions:

$$\begin{aligned}
u_+(-x) &= u_+(x), && (7.2.26)\\
u_-(-x) &= -u_-(x). && (7.2.27)
\end{aligned}$$

For potential energies that are spatially symmetric, $(V(x) = V(-x))$ we can see that the Hamiltonian commutes with the parity operator. This is easily seen considering the matrix elements of the commutator in coordinate representation:

$$\begin{aligned}
\langle x|[\hat{H},\hat{\Pi}]|x'\rangle &= \langle x|\left(\frac{\hat{p}^2}{2m}\hat{\Pi} + V(x)\hat{\Pi}\right) - \left(\hat{\Pi}\frac{\hat{p}^2}{2m} + \hat{\Pi}V(x)\right)|x'\rangle && (7.2.28)\\
&= \langle x|\left(\frac{\hat{p}^2}{2m} + V(x)\right)|-x'\rangle - \langle -x|\left(\frac{\hat{p}^2}{2m} + V(x)\right)|x'\rangle && (7.2.29)\\
&= -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\delta(x+x') + V(x)\delta(x+x') + && (7.2.30)\\
&\quad +\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\delta(-x-x') - V(-x)\delta(x+x') && (7.2.31)\\
&= (V(x) - V(-x))\,\delta(x+x') && (7.2.32)\\
&= 0. && (7.2.33)
\end{aligned}$$

Thus when $V(x) = V(-x)$, we can diagonalize the Hamiltonian and the parity operator simultaneously, and the solutions of the Schroedinger's equation at some energy E must be also eigenvectors of the parity, thus they either satisfy

$$\Psi_+^E(x) = \Psi_+^E(-x), \tag{7.2.34}$$

or

$$\Psi_-^E(x) = -\Psi_-^E(-x). \tag{7.2.35}$$

The harmonic oscillator is one example of potential that is symmetric, and indeed we have seen that in that case the even eigenvalues $n = 0, 2, 4, \ldots$ correspond to spatially even functions, whereas the odd ones $n = 1, 3, 5, \ldots$ correspond to spatially odd functions. While we haven't used this symmetry explicitly to solve the harmonic oscillator, reflection symmetry $(V(x) = V(-x))$ is in general a very useful tool, since it allows us to solve Schroedinger's equation more easily, since we can treat even and odd functions independently, as we will see in the examples below.

7.3 Bound States

We now analyze of a series of Hamiltonians that illustrate some of the general properties of wave functions we have derived above, as well as show some remarkable properties of the quantum world. We first consider a few examples of systems with bound states, thus studying eigenstates of the Hamiltonian with energies $E < V(\pm\infty)$. As we have discussed previously, these are also physical states, since they can be normalized. We will consider scattering states in a separate section.

7.3.1 Particle in a box

The first system we study here consists of a particle inside a one-dimensional box with hard walls, that prevent the particle from escaping. We can model this situation with a symmetric potential:

$$V(x) = \begin{cases} 0, & -\frac{L}{2} \leq x \leq \frac{L}{2} \\ \infty & \text{otherwise,} \end{cases} \tag{7.3.1}$$

where we have called L the linear size of the box, as also shown in Fig. 7.3.1. Since the potential is infinite at $x \to \pm\infty$, we expect to find only physically valid, bound states, as also found for the harmonic oscillator case.

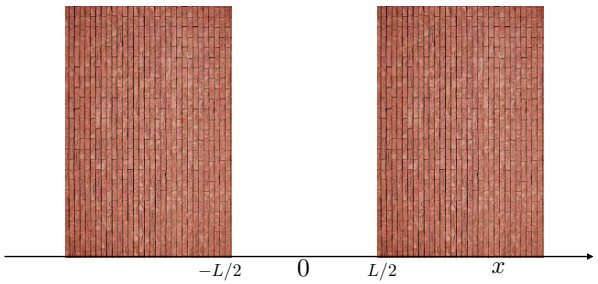

Figure 7.3.1: Potential of a particle in a box with hard walls. The potential energy is here taken to be infinite beyond the limits $0 \leq x \leq L$.

Outside the box, we must have that $|\Psi(x)|^2 = 0$, since we have assumed that the potential is infinite, and the particle is not found outside. Inside the box, the potential energy is 0, thus the time-independent Schroedinger's equation reads

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} = E\Psi, \tag{7.3.2}$$

which can also be written

$$\frac{\partial^2\Psi}{\partial x^2} = -k^2\Psi, \tag{7.3.3}$$

where we have introduced $k = \sqrt{2mE}/\hbar$. From the general bound we have derived in the previous section, we know that the energy must be $E \geq 0$, since in this case $V_{\min} = 0$. It also follows that the quantity k is real and that $k \geq 0$, thus the solution to the differential equation (7.3.3) is

$$\Psi(x) = A\sin kx + B\cos kx, \tag{7.3.4}$$

where A and B are two constants to be determined. From the discussion on the parity operator, we know that the eigenfunctions are of two types in this case, since the potential is symmetric. We thus have:

$$\Psi_+(x) = B\cos k_+ x, \tag{7.3.5}$$
$$\Psi_-(x) = A\sin k_- x, \tag{7.3.6}$$

where $E^{\pm} = \hbar^2 k_{\pm}/2m$ are the energies of, respectively, even and odd states. We now need to fix the free constants imposing appropriate boundary conditions and the normalization for the wave function. Specifically, continuity of the wave function implies

$$\Psi_{\pm}(-L/2) = 0 \tag{7.3.7}$$
$$\Psi_{\pm}(L/2) = 0. \tag{7.3.8}$$

These condition yield

$$B\cos k_+ L/2 = 0 \tag{7.3.9}$$
$$A\sin k_- L/2 = 0, \tag{7.3.10}$$

which are satisfied if

$$k_+ \frac{L}{2} = \left(n_+ + \frac{1}{2}\right)\pi \tag{7.3.11}$$
$$k_- \frac{L}{2} = n_-\pi, \tag{7.3.12}$$

thus

$$k_- = 2n_- \frac{\pi}{L} \tag{7.3.13}$$
$$= (0, 2\frac{\pi}{L}, 4\frac{\pi}{L}, \dots) \tag{7.3.14}$$
$$k_+ = (2n_+ + 1)\frac{\pi}{L} \tag{7.3.15}$$
$$= (\frac{\pi}{L}, 3\frac{\pi}{L}, \dots). \tag{7.3.16}$$

The solution with $n_- = 0$ can be discarded, since it corresponds to a null wave function. Overall the energies of the system are then given by

$$k_n = n\frac{\pi}{L}, \tag{7.3.17}$$
$$n = (1, 2, \dots), \tag{7.3.18}$$

thus yielding a set of discrete values of k, and with states of even/odd n corresponding to spatially odd/even wave functions. This quantization of k directly implies also quantization of the energies, that can now be labelled with the integer index n:

$$E_n = \frac{\hbar^2 k_n^2}{2m} \tag{7.3.19}$$
$$= \frac{\hbar^2\pi^2}{2mL^2} n^2, \tag{7.3.20}$$

thus the quantum particle can take only discrete values of the energy, in radical contrast with the classical case. In order to determine the normalization constants, we need to

impose the normalization condition. For example, for odd states:

$$\begin{aligned}
\langle \Psi^- | \Psi^- \rangle &= \int_{-\infty}^{\infty} dx |\Psi_\pm(x)|^2 && (7.3.21)\\
&= 2\int_0^{L/2} dx |\Psi_\pm(x)|^2 && (7.3.22)\\
&= 2A^2 \int_0^{L/2} dx \sin^2(kx) && (7.3.23)\\
&= A^2 \int_0^{L/2} dx \left(1 - \cos(2kx)\right) && (7.3.24)\\
&= A^2 \int_0^{kL} dx' \left(1 - \cos(x')\right)/2k && (7.3.25)\\
&= \frac{A^2}{2k} \left(x' - \sin(x')\right)\Big|_0^{kL} && (7.3.26)\\
&= \frac{A^2}{2k} (kL) && (7.3.27)\\
&= A^2 \frac{L}{2} && (7.3.28)\\
&= 1. && (7.3.29)
\end{aligned}$$

In summary, the eigenstates of the Hamiltonian are

$$\Psi_n(x) = \begin{cases} \sqrt{\frac{2}{L}} \cos\left(n\frac{\pi}{L}x\right) & n \text{ odd} \\ \sqrt{\frac{2}{L}} \sin\left(n\frac{\pi}{L}x\right). & n \text{ even} \end{cases} \qquad (7.3.30)$$

These states vanish at the edges of the box, and have a number of nodes (values where the wave-function vanishes inside the box) equal to $n-1$, as also shown in Fig. 7.3.2.

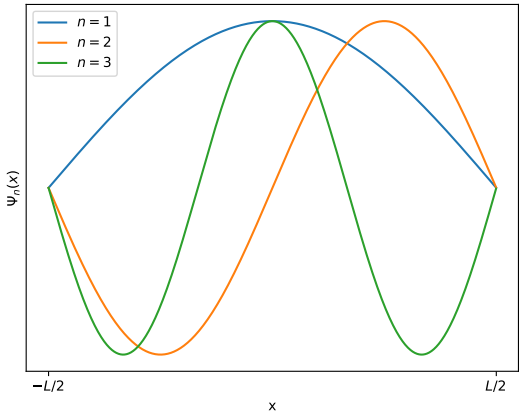

Figure 7.3.2: Energy eigenstates of a particle in a box with hard walls. The wave function is shown in the region $0 \le x \le L$, and it is vanishing elsewhere.

From the general theory we also know that a solution of the time-dependent Schroedinger equation can be written as linear superposition of these basis states (since the Hamiltonian is time independent), thus

$$\Psi(x,t) = \sum_{n=1}^{\infty} c_n(t)\Psi_n(x), \qquad (7.3.31)$$

where

$$c_n(t) = \langle \Psi_n | \Psi(t=0) \rangle e^{-i\frac{E_n}{\hbar}t}, \tag{7.3.32}$$

and the initial amplitude $\langle \Psi_n | \Psi(t=0) \rangle$ is found using the standard tools we have introduced in the previous Chapter. For example, if the initial state is known to be an even state in position representation: $\Psi(x,0) = \Psi(-x,0)$, then only the coefficients with n even are non-zero and given by

$$\langle \Psi_n | \Psi(t=0) \rangle = \int_{-\infty}^{\infty} dx \Psi_n(x) \Psi(x,0) \tag{7.3.33}$$

$$= \sqrt{\frac{2}{L}} \int_{-L/2}^{L/2} dx \cos\left(n\frac{\pi}{L}x\right) \Psi(x,0). \tag{7.3.34}$$

7.3.2 Finite potential well

We now generalize the previous case and consider a finite potential well, described by the potential:

$$V(x) = \begin{cases} V_0, & x < \frac{L}{2} \\ 0 & -\frac{L}{2} \leq x \leq \frac{L}{2} \\ V_0 & x > \frac{L}{2}. \end{cases} \tag{7.3.35}$$

There are thus three distinct regions in space denoted I, II and III in Fig. 7.3.3. We con-

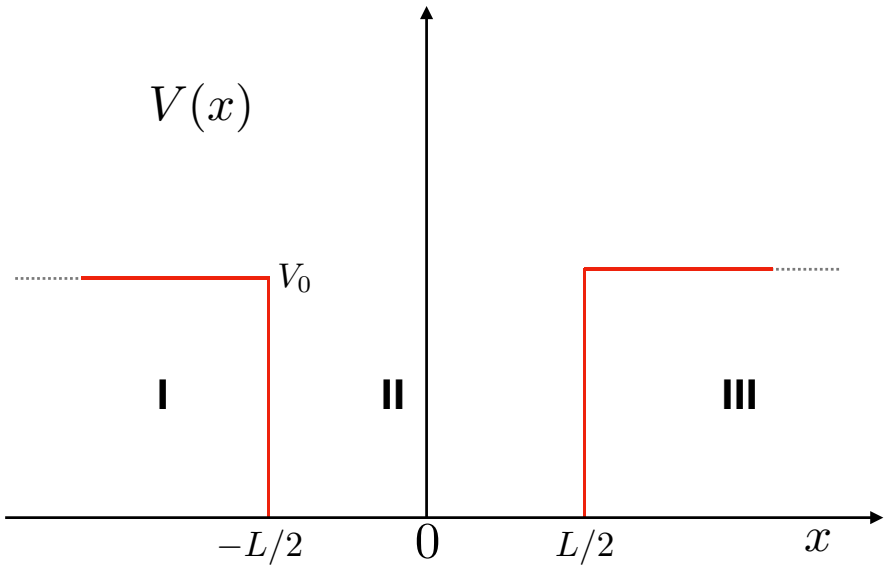

Figure 7.3.3: Potential of a particle in a finite well. The potential energy is vanishing inside the box (region II), and equal to V_0 beyond the box limits $-\frac{L}{2} \leq x \leq \frac{L}{2}$, regions I and III.

centrate here on the bound states solutions, thus for solutions of the Schroedinger equation with $E < V_0$. The Schroedinger's equation in the 3 regions reads

$$\begin{cases} \Psi''(x) + k^2 \Psi(x) = 0, & |x| \leq \frac{L}{2} \\ \Psi''(x) - \rho^2 \Psi(x) = 0 & |x| > \frac{L}{2}, \end{cases} \tag{7.3.36}$$

where we have defined

$$k = \sqrt{2mE}/\hbar \tag{7.3.37}$$

$$\rho = \sqrt{2m(V_0 - E)}/\hbar. \tag{7.3.38}$$

The solution is therefore

$$\begin{cases} \Psi_I(x) = N_I e^{\rho x}, & x < -\frac{L}{2} \\ \Psi_{II}(x) = N_{II}\cos(kx) + N'_{II}\sin(kx) & -\frac{L}{2} \leq x \leq \frac{L}{2} \\ \Psi_{III}(x) = N_{III}e^{-\rho x} & x > \frac{L}{2}, \end{cases} \tag{7.3.39}$$

as before, we have discarded un-normalizable solutions in regions I and II, thus the sign of the exponents follows this choice. In this case the potential is always finite, thus we require both continuity of the wave function and of its first derivative. This yields the following four conditions:

$$\begin{cases} \Psi_I(-\frac{L}{2}) = \Psi_{II}(-\frac{L}{2}) \\ \Psi'_I(-\frac{L}{2}) = \Psi'_{II}(-\frac{L}{2}) \\ \Psi_{II}(\frac{L}{2}) = \Psi_{III}(\frac{L}{2}) \\ \Psi'_{II}(\frac{L}{2}) = \Psi'_{III}(\frac{L}{2}) \end{cases}. \tag{7.3.40}$$

We first concentrate on the even solution:

$$\Psi(x) = \Psi(-x), \tag{7.3.41}$$

thus yielding $N_{III} = N_I$ and $N'_{II} = 0$. The four continuity conditions above then reduce to the following:

$$\begin{cases} N_I e^{-\rho L/2} = N_{II}\cos(kL/2) \\ N_I \rho e^{-\rho L/2} = N_{II} k \sin(kL/2) \end{cases}, \tag{7.3.42}$$

thus yielding

$$\frac{\rho}{k} = \tan\left(\frac{kL}{2}\right). \tag{7.3.43}$$

Recalling that

$$k = \sqrt{2mE}/\hbar \tag{7.3.44}$$

$$\rho = \sqrt{2m(V_0 - E)}/\hbar, \tag{7.3.45}$$

we have

$$\rho^2 + k^2 = \frac{2m}{\hbar^2}\left[E^2 + (V_0 - E)^2 + 2(V_0 - E)E\right] \tag{7.3.46}$$

$$= \frac{2m}{\hbar^2}V_0^2 \tag{7.3.47}$$

$$= k_0^2 \tag{7.3.48}$$

Since $\rho/k > 0$, we can solve this taking

$$\frac{\rho^2}{k^2} = \left(\tan\left(\frac{kL}{2}\right)\right)^2 \tag{7.3.49}$$

$$\frac{k_0^2}{k^2} - 1 = \frac{1}{(\cos(kL/2))^2} - 1, \tag{7.3.50}$$

thus we have that the values of k satisfy

$$\left|\frac{k}{k_0}\right| = |\cos(kL/2)| \tag{7.3.51}$$

$$\tan\frac{kL}{2} > 0 \tag{7.3.52}$$

Similarly, for the odd solutions:

$$\Psi(x) = -\Psi(-x), \tag{7.3.53}$$

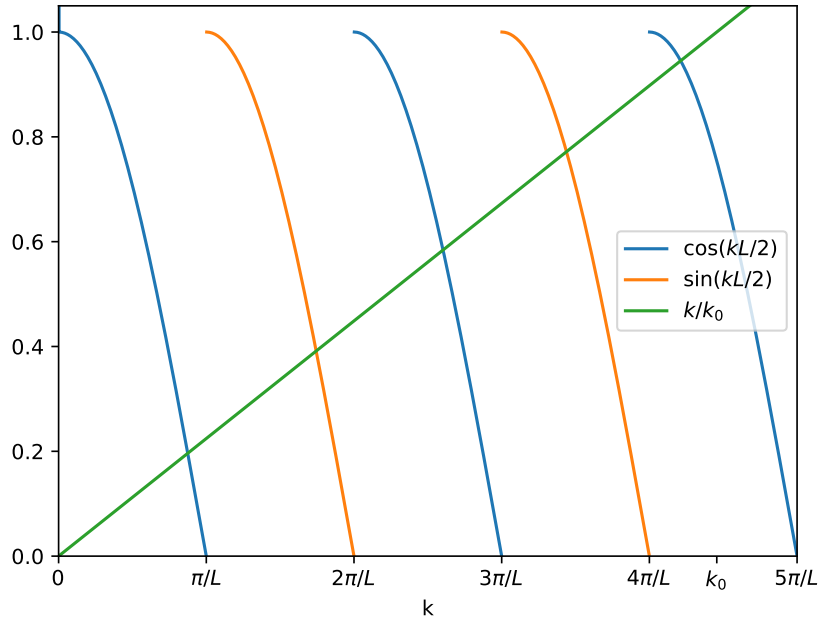

Figure 7.3.4: Graphical solution of the non-linear equations (7.3.51) and (7.3.55). Depending on the value of k_0, a number of intersections between the arcs and the straight line can be found. In the case depicted, there are 3 even solutions (intersections of the straight line with the blue lines) and 2 odd solutions (intersections of the straight line with the orange line).

we must have $N_{III} = -N_I$ and $N_{II} = 0$, thus

$$\begin{cases} N_I e^{-\rho L/2} = -N_{II} \sin(kL/2) \\ N_I \rho e^{-\rho L/2} = N_{II} k \cos(kL/2) \end{cases}, \tag{7.3.54}$$

and the condition to be satisfied is

$$\left| \frac{k}{k_0} \right| = |\sin(kL/2)| \tag{7.3.55}$$

$$\tan \frac{kL}{2} < 0. \tag{7.3.56}$$

The non-linear equations (7.3.51) and (7.3.55) can be solved numerically. Graphically, the solutions corresponds to the intersections of a straight line with slope $1/k_0$ and arcs given by the trigonometric functions $|\cos(kL/2)|$ and $|\sin(kL/2)|$. These are shown in Fig. 7.3.4. Example wave functions corresponding to the even solutions are shown in Fig. 7.3.5. A very remarkable feature is that these wave functions all have finite support in the region $|x| > L/2$, meaning that there is a finite probability of finding the particle outside of the classically allowed region, since classically a particle with $E < E_0$ wouldn't have enough energy to escape the well. This phenomenon is one manifestation of quantum tunneling, i.e. the ability of quantum particles to "escape" potential wells, effectively *tunneling* through them rather than crossing them like for classical systems.

7.3.3 Delta Potential

$$V(x) = -\alpha \delta(x), \tag{7.3.57}$$

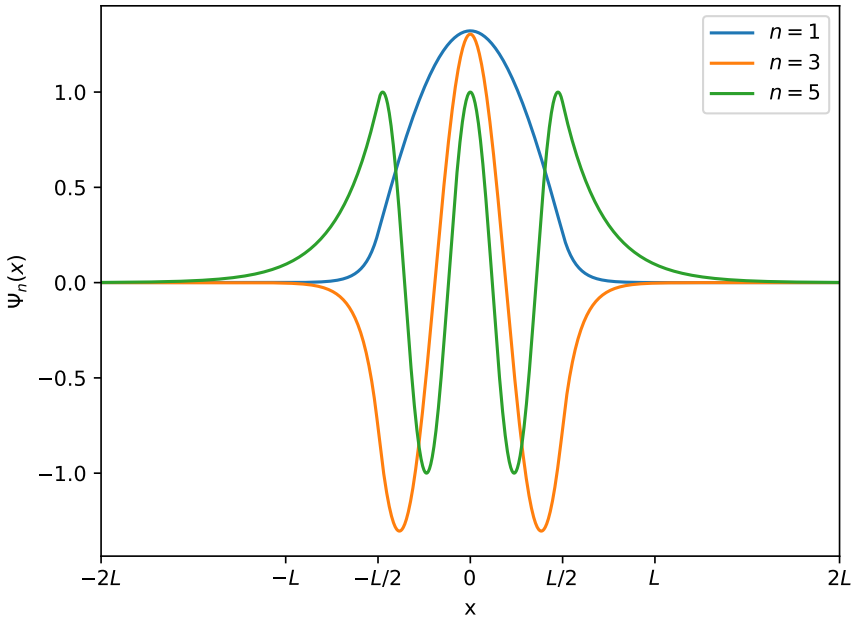

Figure 7.3.5: Example of spatially even stationary states for the finite-well potential. The 3 states correspond to the same value of V_0 shown in Fig. 7.3.4.

with $\alpha > 0$, and the Schroedinger Equation for the stationary states is

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x)}{\partial x^2} - \alpha\delta(x)\Psi(x) = E\Psi(x). \tag{7.3.58}$$

We now consider the case of bound states (thus, $E < 0$), thus in the region $x < 0$, where $V(x) = 0$ we have an exponential form for the wave function. Posing $\rho = \sqrt{2m|E|}/\hbar$, we have :

$$\Psi_<(x) = Ae^{-\rho x} + Be^{\rho x}. \tag{7.3.59}$$

From this equation we can immediately conclude that $A = 0$, otherwise $\Psi(x \to -\infty) \sim \infty$ would lead to a non-normalizable state. Similarly, in the region $x > 0$ the wave function must have only the component

$$\Psi_>(x) = Fe^{-\rho x}. \tag{7.3.60}$$

We now use the boundary conditions we have derived in the previous discussions, and first enforce continuity at $x = 0$, thus leading to $B = F$. We then use the discontinuity condition on the first derivative:

$$\Psi'(\epsilon) - \Psi'(-\epsilon) = \frac{2m}{\hbar^2}\int_{a-\epsilon}^{a+\epsilon} V(x)\Psi(x)dx, \tag{7.3.61}$$

$$= -\alpha\frac{2m}{\hbar^2}\int_{a-\epsilon}^{a+\epsilon}\delta(x)\Psi(x)dx, \tag{7.3.62}$$

$$= -\alpha\frac{2m}{\hbar^2}\Psi(0) \tag{7.3.63}$$

$$= -\alpha\frac{2m}{\hbar^2}B, \tag{7.3.64}$$

and using the explicit form for the derivatives we get

$$-2\rho B = -\alpha\frac{2m}{\hbar^2}B,$$

fixing the energy through the only admissible value of ρ :

$$\rho = \alpha\frac{m}{\hbar^2}, \tag{7.3.65}$$

$$E = -\frac{m\alpha^2}{2\hbar^2}. \tag{7.3.66}$$

We have thus found that this potential admits only a single bound state, and $\Psi(x) = Be^{-\rho|x|}$, with the constant B determined through normalization, leading to:

$$\Psi(x) = \frac{\sqrt{m\alpha}}{\hbar}e^{-\frac{\alpha m}{\hbar^2}|x|}. \tag{7.3.67}$$

7.4 Scattering states

We now analyze solutions of the Schroedinger equation that are not normalizable, yet play an important role to analyze the dynamics of quantum systems.

7.4.1 Wave Packets

Consider the Hamiltonian of a free particle

$$\hat{H} = \frac{\hat{p}^2}{2m}, \tag{7.4.1}$$

whose eigenstates are just the eigenstates of the momentum with energies $\frac{p^2}{2m}$:

$$\hat{H}|p\rangle = \frac{p^2}{2m}|p\rangle, \tag{7.4.2}$$

where

$$\langle x|p\rangle = \frac{e^{i\frac{px}{\hbar}}}{\sqrt{2\pi\hbar}}. \tag{7.4.3}$$

The difficulty with these eigenstates is that they are legitimate solutions of the time-independent Schroedinger equation, but they are not legitimate physical states, since they are not square-normalizable. However, we have seen that a generic solution of the time-dependent Schroedinger equation for time-independent hamiltonians is

$$|\Psi(t)\rangle = e^{-i\frac{\hat{H}t}{\hbar}}|\Psi(0)\rangle. \tag{7.4.4}$$

The time evolution for the free particle hamiltonian is particularly easy in momentum representation,

$$\langle p|\Psi(t)\rangle = \int dp'\langle p|e^{-i\frac{\hat{H}t}{\hbar}}|p'\rangle\langle p'|\Psi(0)\rangle \tag{7.4.5}$$

$$= \int dp'\langle p|e^{-i\frac{\hat{p}^2t}{2m\hbar}}|p'\rangle\langle p'|\Psi(0)\rangle \tag{7.4.6}$$

$$= e^{-i\frac{p^2}{2m\hbar}t}\langle p|\Psi(0)\rangle. \tag{7.4.7}$$

In coordinate representation the expressions are a bit more complicated, but can be easily found using the representation of the momentum eigenstates in coordinate space:

$$\begin{aligned}
\langle x|\Psi(t)\rangle &= \int dp \langle x|p\rangle\langle p|\Psi(t)\rangle && (7.4.8)\\
&= \int dp \frac{e^{i\frac{px}{\hbar}}}{\sqrt{2\pi\hbar}} e^{-i\frac{p^2}{2m\hbar}t}\langle p|\Psi(0)\rangle && (7.4.9)\\
&= \int dp \frac{e^{i\frac{px}{\hbar}}}{\sqrt{2\pi\hbar}} e^{-i\frac{p^2}{2m\hbar}t}\int dx' \frac{e^{-i\frac{px'}{\hbar}}}{\sqrt{2\pi\hbar}}\psi(x',0) && (7.4.10)\\
&= \frac{1}{2\pi\hbar}\int dx'\psi(x',0)\int dp e^{i\frac{p(x-x')}{\hbar}} e^{-i\frac{p^2}{2m\hbar}t} && (7.4.11)\\
&= \sqrt{\frac{m}{2\pi\hbar t}} e^{-i\pi/4}\int dx'\psi(x',0) e^{i\frac{(x-x')^2}{2\hbar t}m}. && (7.4.12)
\end{aligned}$$

We thus see that, in real space, what happens is that the initial wave functions is convolved with a time-dependent kernel $e^{i\frac{(x-x')^2}{2\hbar t}m}$. Both in real and momentum space however we immediately notice that if the initial state is normalized then it will stay normalized at all subsequent times. In this sense, the unnormalized states $\langle x|p\rangle$ only play a mathematical role in the solution of the time-evolution of a given initial and physically valid wave function.

7.4.2 Group velocity

The time evolution of a free wave packet, as we have seen, depends on the specific form of the initial state, $\langle p|\Psi(0)\rangle = \Phi(p)$, the time-evolved packet is given by

$$\Psi(x,t) = \int dp \frac{e^{i\left(\frac{px}{\hbar}-\frac{p^2}{2m\hbar}t\right)}}{\sqrt{2\pi\hbar}}\Phi(p). \quad (7.4.13)$$

While the details will depend on the function $\Phi(p)$, there are however some general properties that we can deduce in the case in which we assume that $\Phi(p)$ is a smooth function centered around some momentum $\bar{p}$ and with a relatively small width Δp. Since the integral is then dominated by points around $\bar{p}$, we can start by considering a Taylor expansion of the energies around that momentum:

$$\begin{aligned}
E(p) &= \frac{p^2}{2m} && (7.4.14)\\
&\simeq \frac{\bar{p}^2}{2m} + \frac{\bar{p}}{m}(p-\bar{p}) + \mathcal{O}(\Delta p^2) && (7.4.15)\\
&= E(\bar{p}) + E'(\bar{p})(p-\bar{p}) + \mathcal{O}(\Delta p^2), && (7.4.16)
\end{aligned}$$

thus

$$\begin{aligned}
\Psi(x,t) &\simeq \frac{e^{-iE(\bar{p})t/\hbar}}{\sqrt{2\pi\hbar}}\int dp e^{i\left(\frac{px}{\hbar}-\frac{E'(\bar{p})(p-\bar{p})}{\hbar}t\right)}\Phi(p) && (7.4.17)\\
&= \frac{e^{-iE(\bar{p})t/\hbar}}{\sqrt{2\pi\hbar}}\int ds e^{i\left(\frac{(s+\bar{p})x}{\hbar}-\frac{E'(\bar{p})}{\hbar}st\right)}\Phi(s+\bar{p}) && (7.4.18)\\
&= \frac{e^{i\left(E'(\bar{p})\bar{p}-E(\bar{p})\right)\frac{t}{\hbar}}}{\sqrt{2\pi\hbar}}\int ds e^{i\left(\frac{(s+\bar{p})}{\hbar}(x-E'(\bar{p})t)\right)}\Phi(s+\bar{p}) && (7.4.19)\\
&\quad \frac{e^{i\left(E'(\bar{p})\bar{p}-E(\bar{p})\right)\frac{t}{\hbar}}}{\sqrt{2\pi\hbar}}\int dp e^{i\left(\frac{p}{\hbar}(x-E'(\bar{p})t)\right)}\Phi(p) && (7.4.20)\\
&= e^{-i\left(E(\bar{p})-E'(\bar{p})\bar{p}\right)\frac{t}{\hbar}}\Psi\left(x-E'(\bar{p})t,0\right). && (7.4.21)
\end{aligned}$$

This expression is particularly interesting because it tells us that the form of the time-evolved wave packed (apart from a phase factor) is approximately equal to the initial state

but in a moving frame $x'(t) = x - E'(\bar{p})t$. This means that, for example, the maximum of the wave function will move with a velocity:

$$\begin{aligned} \mathrm{v_g} &= E'(\bar{p}) && (7.4.22)\\ &= \frac{\bar{p}}{m}, && (7.4.23) \end{aligned}$$

this velocity is known as group velocity, and it corresponds to the velocity that a classical free particle of initial momentum $\bar{p}$ would have. It should be also noticed that this expression is valid as long as the quadratic corrections to the energy can be neglected, thus when

$$\Delta p^2 \frac{t}{m\hbar} \ll 1, \tag{7.4.24}$$

thus for time scales given by the inverse of the momentum spread of the initial state

$$t \ll \frac{m\hbar}{\Delta p^2}. \tag{7.4.25}$$

7.4.3 Step Potential

We now analyze the step potential

$$V(x) = \begin{cases} 0, & x < 0 \\ V_0 & x > 0, \end{cases} \tag{7.4.26}$$

which admits scattering states only. We will consider two cases $E > V_0$ and $E < V_0$, and in both cases the solutions of the Schroedinger equations will be extending infinitely away for $x \to -\infty$, and are thus non-normalizable. Eventually, we will be interested in considering the physical situation in which the (unnormalized) stationary states of this potential are used to determine the dynamics of physically valid (normalizable) states, through a superposition of the stationary states, as previously done for free particles. Before doing that, we need to study in some detail the stationary states.

7.4.3.1 Stationary states for $E > V_0$

We start our analysis with the case $E > V_0$. Calling

$$\begin{aligned} k_1 &= \sqrt{2mE/\hbar^2} && (7.4.27)\\ k_2 &= \sqrt{2m(E-V_0)/\hbar^2}, && (7.4.28) \end{aligned}$$

we see that the Schroedinger equation in both regions of space has the same form and reads,

$$\begin{aligned} \Psi''_<(x) + k_1^2 \Psi_<(x) &= 0 && (7.4.29)\\ \Psi''_>(x) + k_2^2 \Psi_>(x) &= 0, && (7.4.30) \end{aligned}$$

thus

$$\begin{aligned} \Psi_<(x) &= Ae^{ik_1 x} + Be^{-ik_1 x} && (7.4.31)\\ \Psi_>(x) &= Ce^{ik_2 x} + De^{-ik_2 x}, && (7.4.32) \end{aligned}$$

where we have 4 constants (A, B, C, D) to be determined (in addition to the energy E). Physically, we will be interested in studying the dynamics of wave packets in this potential, and we will consider the situation in which we have an initial wave packet, with finite positive momentum, thus effectively moving from the left ($x = -\infty$) to the right ($x > 0$). For this reason, we will consider only solutions with $D = 0$. Using then the usual conditions of continuity of the wave function at $x = 0$ we get:

$$A + B = C, \tag{7.4.33}$$

whereas the continuity of the first derivative yields

$$ik_1 A - ik_1 B = ik_2 C, \tag{7.4.34}$$

thus combining the two equations we have

$$\frac{B}{A} = \frac{k_1 - k_2}{k_1 + k_2}, \tag{7.4.35}$$

$$\frac{C}{A} = \frac{2k_1}{k_1 + k_2}. \tag{7.4.36}$$

We can get further insight into the meaning of these coefficients by evaluating the probability current to the left and to the right of the step. We recall the expression for the probability current:

$$J(x,t) = \frac{i\hbar}{2m}\left[\left(\frac{\partial}{\partial x}\Psi^\star(x,t)\right)\Psi(x,t) - \Psi(x,t)^\star\left(\frac{\partial}{\partial x}\Psi(x,t)\right)\right]. \tag{7.4.37}$$

$$= \frac{\hbar}{m}\mathrm{Im}\left[\Psi(x,t)^\star\left(\frac{\partial}{\partial x}\Psi(x,t)\right)\right], \tag{7.4.38}$$

thus

$$J_<(x,t) = \frac{\hbar}{m}\mathrm{Im}\left[\left(A^\star e^{-ik_1 x} + B^\star e^{ik_1 x}\right)\left(Aik_1 e^{ik_1 x} - Bik_1 e^{-ik_1 x}\right)\right] \tag{7.4.39}$$

$$= \frac{\hbar k_1}{m}\left(|A|^2 - |B|^2\right), \tag{7.4.40}$$

and

$$J_>(x,t) = \frac{\hbar}{m}\mathrm{Im}\left[\left(C^\star e^{-ik_2 x}\right)\left(Cik_2 e^{ik_2 x}\right)\right] \tag{7.4.41}$$

$$= \frac{\hbar k_2}{m}|C|^2. \tag{7.4.42}$$

The left current can also be written as the difference of two currents,

$$J_< = J_A - J_B, \tag{7.4.43}$$

where J_A and J_B are the currents of the two individual components. It can also be remarked already that currents in this case are all time independent, as it should be since we are here considering time-independent states. In particular, because in the continuity equation

$$\partial_x J + \partial_t |\Psi(x,t)|^2 = 0 \tag{7.4.44}$$

there is no time dependence (the second term is zero) we see that also that the current should be space independent $\partial_x J = 0$ for stationary solutions. We can verify this checking that the current does not have any spatial dependence, for example

$$J_< = \frac{\hbar k_1}{m}|A|^2\left(1 - \frac{|B|^2}{|A|^2}\right) \tag{7.4.45}$$

$$= \frac{\hbar k_1}{m}|A|^2\left(1 - \left(\frac{k_1 - k_2}{k_1 + k_2}\right)^2\right) \tag{7.4.46}$$

$$= \frac{\hbar k_1}{m}|A|^2\left(\frac{4k_1 k_2}{(k_1 + k_2)^2}\right) \tag{7.4.47}$$

and

$$J_> = |A|^2\frac{\hbar k_2}{m}\left(\frac{2k_1}{k_1 + k_2}\right)^2, \tag{7.4.48}$$

thus $J_< = J_>$. Since the probability fluxes are made out of three different components, J_A (an incoming wave), J_B (a reflected wave), J_C (a transmitted wave) it is often useful to reason in terms of reflection and transmission coefficients. The reflection coefficient is defined as

$$\begin{aligned} R &\equiv \frac{J_B}{J_A} && (7.4.49)\\ &= \frac{|B|^2}{|A|^2} && (7.4.50)\\ &= \left(\frac{k_1 - k_2}{k_1 + k_2}\right)^2 && (7.4.51)\\ &\leq 1, && (7.4.52) \end{aligned}$$

whereas the transmission coefficient is

$$\begin{aligned} T &\equiv \frac{J_C}{J_A} && (7.4.53)\\ &= \frac{k_2}{k_1}\frac{|C|^2}{|A|^2} && (7.4.54)\\ &= \frac{4k_1k_2}{(k_1 + k_2)^2} && (7.4.55)\\ &\leq 1. && (7.4.56) \end{aligned}$$

It should be remarked that in general these coefficients are not just the ratios of the waves amplitudes, but are really defined as ratios of probability currents (this is particularly clear of the transmission coefficient, which is $T \neq |C|^2/|A|^2$). Also, as expected,

$$R + T = 1. \quad (7.4.57)$$

7.4.3.2 Stationary states for $E < V_0$

In this case we replace the wave vector k_2 with the decay constant,

$$\rho_2 = \sqrt{2m(V_0 - E)/\hbar^2} \quad (7.4.58)$$

such that

$$\begin{aligned} \Psi_<(x) &= Ae^{ik_1x} + Be^{-ik_1x} && (7.4.59)\\ \Psi_>(x) &= Ce^{\rho_2 x} + De^{-\rho_2 x}, && (7.4.60) \end{aligned}$$

and we can already set $C = 0$ to have a bound solution when $x \to \infty$. Imposing the usual continuity conditions at $x = 0$ we now get

$$\begin{aligned} \frac{B}{A} &= \frac{k_1 - i\rho_2}{k_1 + i\rho_2} && (7.4.61)\\ \frac{D}{A} &= \frac{2k_1}{k_1 + i\rho_2}. && (7.4.62) \end{aligned}$$

In order to evaluate the reflection coefficient, we write the first ratio as:

$$\begin{aligned} \frac{B}{A} &= \frac{i(k_1 - i\rho_2)}{i(k_1 + i\rho_2)} && (7.4.63)\\ &= -\frac{\rho_2 + ik_1}{\rho_2 - ik_1} && (7.4.64)\\ &= -\frac{(\rho_2^2 + k_1^2)e^{i\delta(E)}}{(\rho_2^2 + k_1^2)e^{-i\delta(E)}} && (7.4.65)\\ &= -e^{2i\delta(E)}, && (7.4.66) \end{aligned}$$

with

$$\delta(E) = \arctan\frac{k_1}{\rho_2} \tag{7.4.67}$$

$$= \arctan\sqrt{\frac{E}{V_0 - E}}. \tag{7.4.68}$$

Notice that in this case the reflection coefficient is

$$R = \left|\frac{B}{A}\right|^2 \tag{7.4.69}$$

$$= \left|-e^{i\delta(E)}\right|^2 \tag{7.4.70}$$

$$= 1. \tag{7.4.71}$$

The fact that the reflection coefficient is one means that there is no transmitted *current*, however it should be remarked that this is not the same as saying that the probability of finding a particle over the barrier is zero. Finite currents in this context are rather to be interpreted as the existence of non-normalizable plane-wave solutions, thus since on the right of the barrier we have a normalizable exponentially decaying wave function, we also have that these solutions do not carry momentum, thus they have zero current:

$$J_> = \frac{\hbar}{m}\mathrm{Im}\left[\Psi_>^\star \Psi_>'\right] \tag{7.4.72}$$

$$= -\frac{\hbar}{m}|D|^2\mathrm{Im}\left[\rho_2 e^{-2\rho_2 x}\right] \tag{7.4.73}$$

$$= 0, \tag{7.4.74}$$

implying also that $J_< = J_A - J_B = 0$, thus

$$J_A = J_B, \tag{7.4.75}$$

$$J_C = 0. \tag{7.4.76}$$

7.4.4 Wave Packets

We can now form physical solutions of the time-dependent Schroedinger equation, considering the time evolution of some initial wave function $|\Psi(0)\rangle$, that in momentum space we take to be $\Phi(p) = \langle p|\Psi(0)\rangle$, such that is sharply peaked around a certain value $\bar{p}$ (you can think for example that we are taking a gaussian wave packet in momentum space, with mean given by $\bar{p}$, but the precise form is not too important, provided that the wave function is peaked around $\bar{p}$ and that other momenta are suppressed. Furthermore, we will concentrate our analysis on the two cases in which either the initial state is exactly non-zero either for $p < p_0$ or for $p > p_0$, where we have introduced the characteristic momentum of the barrier,

$$p_0 = \sqrt{2mV_0}, \tag{7.4.77}$$

$$k_0 = \sqrt{\frac{2mV_0}{\hbar^2}}. \tag{7.4.78}$$

7.4.4.1 $\bar{p} > p_0$

In general we have that

$$\Psi(x,t) = \int_0^\infty dE e^{-iEt/\hbar}\langle E|\Psi(0)\rangle\langle x|E\rangle \tag{7.4.79}$$

however we use the fact that the initial state is vanishing for $E < E_0$, thus

$$\Psi(x,t) = \begin{cases} \int_{V_0}^\infty dE\left(Ae^{ik_1 x} + Be^{-ik_1 x}\right)e^{-iE(k_1)t/\hbar}\langle E|\Psi(0)\rangle & x < 0 \\ \int_{V_0}^\infty dE\left(Ce^{ik_2 x}\right)e^{-iE(k_2)t/\hbar}\langle E|\Psi(0)\rangle & x > 0 \end{cases} \tag{7.4.80}$$

and changing integration variables we have

$$\Psi(x,t) = \begin{cases} \frac{\hbar^2}{m}\int_{k_0}^{\infty} dk_1 k_1 \left(Ae^{ik_1x} + Be^{-ik_1x}\right) e^{-iE(k_1)t/\hbar}\Phi(k_1) & x<0 \\ \frac{\hbar^2}{m}\int_{k_0}^{\infty} dk_2 k_2 \left(Ce^{ik_2x}\right) e^{-iE(k_2)t/\hbar}\Phi(k_2) & x>0 \end{cases} \tag{7.4.81}$$

We can split the wave function into three contributions (incoming, reflected and transmitted):

$$\Psi(x,t) \propto \begin{cases} \Psi_{\rm inc}(x,t) + \Psi_{\rm ref}(x,t) & x<0 \\ \Psi_{\rm tran}(x,t) & x>0 \end{cases}, \tag{7.4.82}$$

that read

$$\Psi_{\rm inc}(x,t) = \int_{k_0}^{\infty} dk_1 \tilde{A}(k_1) e^{ik_1x} e^{-iE(k_1)t/\hbar}\Phi(k_1) \tag{7.4.83}$$

$$\Psi_{\rm ref}(x,t) = \int_{k_0}^{\infty} dk_1 \tilde{B}(k_1) e^{-ik_1x} e^{-iE(k_1)t/\hbar}\Phi(k_1) \tag{7.4.84}$$

$$\Psi_{\rm tran}(x,t) = \int_{k_0}^{\infty} dk_2 \tilde{C}(k_2) e^{-ik_1x} e^{-iE(k_2)t/\hbar}\Phi(k_2), \tag{7.4.85}$$

where $\tilde{A}(k), \tilde{B}(k), \tilde{C}(k)$ are smooth functions of momentum. The first term describes an incoming wave packet, whose group velocity can be found again in the case when $\Phi(p)$ is strongly peaked around some momentum $\bar{p}$, thus the wave packet maximum moves as

$$x = \frac{\bar{p}}{m}t, \tag{7.4.86}$$

with the constant velocity $\bar{p}/m$. However notice that $\Psi_{\rm inc}(x,t)$ is defined only for $x<0$, thus it is clear that this solution is valid only if $t<0$.

For $\Psi_{\rm ref}(x,t)$ we have that $\Psi_{\rm ref}(-x,t)$ has the usual functional form we have studied so far for wave-packets, thus implying that

$$x = -\frac{\bar{p}}{m}t. \tag{7.4.87}$$

We have then a peak when $x<0$, as before, but in this case we must have $t>0$. Thus the image is that this term indeed corresponds to a wave that stems from the reflection from the barrier at $t=0$ and starts moving after the reflection happens.

Finally, for $\Psi_{\rm tran}(x,t)$ we recall that $k_2^2 + k_0^2 = k_1^2$, thus the wave-packet moves at the group velocity

$$v_{\rm trans} = \partial_{k_2} E(k_2)|_{\bar{p}(k_2)} \tag{7.4.88}$$

$$= \frac{\hbar k_2(\bar{p})}{m} \tag{7.4.89}$$

$$= \frac{\sqrt{\bar{p}^2 - p_0^2}}{m}. \tag{7.4.90}$$

This solution is valid for $x>0$ and $t>0$, thus it coexists with the reflected wave that travels in the opposite direction and the wave packet travels with a positive velocity $\frac{\sqrt{\bar{p}^2-p_0^2}}{m}$.

In conclusion, in this case we have that the incoming wave packet is partially reflected (with the same velocity, in modulus, of the incoming packet) and that the transmitted wave packet instead moves beyond the barrier with a reduced velocity.

7.4.4.2 $\bar{p} < p_0$

The last case we consider are wave packets such that the initial state $\Phi(p>p_0)=0$ thus we only take superpositions of eigenstates with $E<V_0$. The time dependent solution in momentum space is then found using superpositions of the stationary states we found for

$E < V_0$, however we will only consider the behavior of the wave-packet in the region $x < 0$, where the analysis can still be carried with the approximate tools we have developed so far. In that region, we have

$$\Psi(x,t) \propto \left\{ \int_0^{k_0} dk_1 k_1 \left(e^{ik_1 x} - e^{2i\delta(E)} e^{-ik_1 x} \right) e^{-iE(k_1)t/\hbar} \Phi(k_1) \quad x < 0 \right. , \tag{7.4.91}$$

thus

$$\Psi_{\rm inc}(x,t) = \int_0^{k_0} dk_1 k_1 e^{ik_1 x} e^{-iE(k_1)t/\hbar} \Phi(k_1) \tag{7.4.92}$$

$$\Psi_{\rm ref}(x,t) = -\int_0^{k_0} dk_1 k_1 e^{2i\delta(E)} e^{-ik_1 x} e^{-iE(k_1)t/\hbar} \Phi(k_1). \tag{7.4.93}$$

The most important difference with respect to the previous case is that the coefficient of the reflected wave (B, in the previous section) is not real as before but it is actually a pure phase: $B/A = -e^{2i\delta(E)}$. This will change the previous analysis of the wave-packet velocity we have done, since we need to expand also $\delta(E(k))$ around p_0 and not only the energy, thus

$$2\delta(E) - k_1 x - E(k_1)t/\hbar \simeq \text{const} + 2\partial_{k_1}\delta(E)|_{\bar{k}} k_1 - k_1 x - E'(\bar{k})t/\hbar, \tag{7.4.94}$$

thus this is equivalent to the transformation

$$x \to -x + 2\partial_{k_1}\delta(E)|_{\bar{k}} \tag{7.4.95}$$

$$= -x + \Delta_x \tag{7.4.96}$$

Recalling that

$$\delta(E) = \arctan \sqrt{\frac{E}{V_0 - E}}, \tag{7.4.97}$$

we have

$$\partial_{k_1}\delta(E) = \partial_E \delta(E) \frac{\hbar^2}{m} k_1, \tag{7.4.98}$$

and

$$\partial_E \delta(E) = \frac{1}{2}\sqrt{\frac{1}{E(V_0 - E)}} \tag{7.4.99}$$

$$\geq 0 \tag{7.4.100}$$

thus

$$\Delta_x = 2\partial_E \delta(E)|_{\bar{p}} \frac{\bar{p}}{m} \tag{7.4.101}$$

$$= \Delta_t \frac{\bar{p}}{m} \tag{7.4.102}$$

$$= \frac{p_0}{m}\sqrt{\frac{1}{\bar{E}(V_0 - \bar{E})}}, \tag{7.4.103}$$

where $\bar{E} = \bar{p}^2/2m$, giving a displacement for the peak of the wave-packet:

$$-x + \Delta_x = \frac{\bar{p}}{m} t, \tag{7.4.104}$$

$$x = -\frac{\bar{p}}{m}(t - \Delta_t). \tag{7.4.105}$$

This result is therefore telling us that there is a *time delay* (remember that $\Delta_t > 0$) in the propagation of the reflected wave. The reflected wave thus moves back with the same speed of the incoming wave, but at a later time than expected.

7.5 References and Further Reading

A general discussion on Schroedinger's formulation of wave mechanics is found in Sakurai's "Modern Quantum Mechanics" (Chapter 2, sections 2.4 and 2.5), even though the discussion in there does not cover some of the more basic details we have covered in our lectures. Cohen-Tannoudji's book contains instead a comprehensive study of one-dimensional problems, both in the case of bound and scattering states. (Chapter 1 in general and complements H1 and J1 are recommended reads).

Chapter 8

Angular Momentum

In this Chapter we study angular momentum, which is a concept of fundamental importance in mechanics. When we have introduced the momentum operator as well as the Hamiltonian operator, we have made the important observation that these operators emerge as *generators* of certain physical operations. This is the case of the translations (for the momentum) and of time evolution (for the Hamiltonian). Another fundamental operation that we can perform on a physical system is the ensemble of *rotations*. We will see in this Chapter, that the very concept of angular momentum in quantum mechanics emerges from the properties of rotations. In this Chapter we first introduce the angular momentum and its commutation relations, and then study the general properties of its spectrum and eigenstates.

8.1 Rotation matrices

We start by recalling some important properties of rotation matrices, that will be instrumental in deriving angular momentum in the quantum setting. Let us consider a 3-dimensional system, specified by a vector of coordinates $\mathbf{v} = (v_x, v_y, v_z) \equiv (v_1, v_2, v_3)$. In general, any rotation can be expressed as the action of a 3×3 matrix $\hat{R}$, such that the transformed coordinates read:

$$\mathbf{v}' = \hat{R}\mathbf{v}. \tag{8.1.1}$$

We can find conditions on the matrix R noticing that the scalar product between two rotated vectors must be preserved by the rotation:

$$\mathbf{v}' \cdot \mathbf{w}' = \mathbf{v} \cdot \mathbf{w}, \tag{8.1.2}$$

thus

$$\sum_i \left(\sum_j R_{ij} v_j \right) \left(\sum_k R_{ik} w_k \right) = \sum_i v_i w_i \tag{8.1.3}$$

$$\sum_{jk} (v_j w_k) \sum_i R_{ij} R_{ik} = \sum_i v_i w_i, \tag{8.1.4}$$

which is satisfied if the matrix is an orthogonal matrix:

$$\sum_i R_{ij} R_{ik} = \delta_{jk} \tag{8.1.5}$$

$$\hat{R}\hat{R}^T = \hat{I}. \tag{8.1.6}$$

In general, all orthogonal matrices have $\det \hat{R} = \pm 1$, but in the following we will study only rotations such that $\det \hat{R} = 1$. These are called orientation-preserving rotations, since they do not involve flipping axes, for example, but rather only continuously changing the coordinate system.

In order to understand the general properties of $\hat{R}$ it is very useful to expand the matrix $\hat{R}$ for small rotations:

$$\hat{R} = \hat{I} + \delta\hat{\rho}, \tag{8.1.7}$$

where $\hat{\rho}$ is a 3×3 matrix we wish to determine and that characterizes both the rotation *direction* and *angle* of rotation. We can compute the inverse infinitesimal rotation to be just $\hat{R}^{-1} = \hat{I} - \delta\hat{\rho}$, as it can be easily checked

$$\hat{R}\hat{R}^{-1} = (\hat{I} + \delta\hat{\rho})(\hat{I} - \delta\hat{\rho}) \tag{8.1.8}$$
$$= \hat{I} + \mathcal{O}(\delta^2). \tag{8.1.9}$$

Since the rotation matrix is orthogonal: $\hat{R}\hat{R}^T = \hat{I}$, or equivalently $\hat{R}^T = \hat{R}^{-1}$, we then have

$$\hat{R}^T = \hat{R}^{-1} \tag{8.1.10}$$
$$\left(\hat{I} + \delta\hat{\rho}\right)^T = \hat{I} - \delta\hat{\rho} \tag{8.1.11}$$
$$\hat{\rho}^T = -\hat{\rho}, \tag{8.1.12}$$

thus the infinitesimal rotation matrix must be antisymmetric $\rho_{ij} = -\rho_{ji}$, or explicitly written:

$$\hat{\rho} \doteq \begin{pmatrix} 0 & \rho_{12} & \rho_{13} \\ -\rho_{12} & 0 & \rho_{23} \\ -\rho_{13} & -\rho_{23} & 0 \end{pmatrix}. \tag{8.1.13}$$

The three independent components of the matrix $\hat{\rho}$ then fully determine the infinitesimal action of the rotation operator. We can now conveniently define a vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$ in terms of the 3 independent components:

$$\theta_1 \equiv -\rho_{23} \tag{8.1.14}$$
$$\theta_2 \equiv \rho_{13} \tag{8.1.15}$$
$$\theta_3 \equiv -\rho_{12}. \tag{8.1.16}$$

Notice that she sign convention relating the components of $\boldsymbol{\theta}$ to the matrix elements of $\hat{\rho}$ is arbitrary and boils down to choosing whether the rotations are to be taken clockwise or anti-clockwise. Were we have used the widely adopted anti-clockwise choice that allows to write the matrix elements compactly as:

$$\rho_{ij} = -\epsilon_{ijk}\theta_k, \tag{8.1.17}$$

where we have introduced the antisymmetric Levi-Civita symbol

$$\epsilon_{ijk} = \begin{cases} 1 & \text{cyclic permutations of } (1,2,3) \\ -1 & \text{cyclic permutations of } (1,3,2) \\ 0 & i = j \ , \ j = k \text{ or } i = k. \end{cases} \tag{8.1.18}$$

With this convention, an infinitesimal rotation then reads

$$R_{ij} = \delta_{ij} - \sum_k \epsilon_{ijk}\delta\theta_k + \mathcal{O}(\delta^2), \tag{8.1.19}$$

and the transformed vector

$$v_i' = \sum_j R_{ij} v_j \tag{8.1.20}$$
$$= v_i - \sum_{jk} \epsilon_{ijk} v_j \delta\theta_k, \tag{8.1.21}$$

but the last term is nothing but a cross product, since in general

$$(\mathbf{a}\times\mathbf{b})_k = \sum_{ij}\epsilon_{ijk}a_i b_j \tag{8.1.22}$$

$$(\mathbf{a}\times\mathbf{b})_i = -\sum_{jk}\epsilon_{ijk}b_j a_k \tag{8.1.23}$$

Vector-wise we can write the infinitesimal rotation as

$$\mathbf{v}' = \mathbf{v} + (\delta\boldsymbol{\theta}\times\mathbf{v})\,, \tag{8.1.24}$$

thus the meaning of the vector $\delta\boldsymbol{\theta}$ is really what we would expect from a rotation along the $\boldsymbol{\theta}$ direction, and the vector is rotated by an infinitesimal angle $|\delta\boldsymbol{\theta}|$ along that direction, as also shown in Fig. 8.1.1.

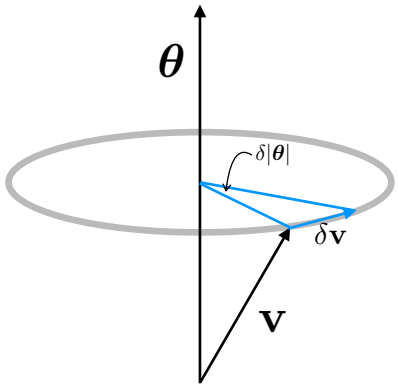

Figure 8.1.1: Geometrical meaning of formula (8.1.24).

8.1.1 Rotations do not commute

A very important feature of rotations is that rotations across different directions do not commute. This can visually understood very easily when considering rotations across different axes, as shown in Figure 8.1.2.

In order to explicitly compute the commutator, we can consider two infinitesimal rotations through the directions $\delta\boldsymbol{\alpha}$ and $\delta\boldsymbol{\beta}$. If we apply first the $\boldsymbol{\alpha}$ rotation and then the $\boldsymbol{\beta}$ rotation we get

$$\hat{R}(\delta\boldsymbol{\beta})\hat{R}(\delta\boldsymbol{\alpha})\mathbf{v} = \hat{R}(\delta\boldsymbol{\beta})(\mathbf{v}+\delta\boldsymbol{\alpha}\times\mathbf{v}+\frac{\delta^2}{2}\hat{S}(\boldsymbol{\alpha})\mathbf{v}+\dots) \tag{8.1.25}$$

$$= \mathbf{v}+\delta\boldsymbol{\alpha}\times\mathbf{v}+\delta\boldsymbol{\beta}\times\mathbf{v}+\delta\boldsymbol{\beta}\times(\delta\boldsymbol{\alpha}\times\mathbf{v})+ +\frac{\delta^2}{2}\left(\hat{S}(\boldsymbol{\alpha})\mathbf{v}+\hat{S}(\boldsymbol{\beta})\mathbf{v}\right)+\mathcal{O}(\delta^3), \tag{8.1.26}$$

where we have formally introduced the second order development of the rotation matrix $\hat{S}(\boldsymbol{\theta})$, that we won't compute explicitly for the moment. The composition of the two rotations in the other order gives the same expression but with $\boldsymbol{\alpha}\leftrightarrow\boldsymbol{\beta}$, i.e. :

$$\hat{R}(\delta\boldsymbol{\alpha})\hat{R}(\delta\boldsymbol{\beta})\mathbf{v} = \mathbf{v}+\delta\boldsymbol{\alpha}\times\mathbf{v}+\delta\boldsymbol{\beta}\times\mathbf{v}+\delta\boldsymbol{\alpha}\times(\delta\boldsymbol{\beta}\times\mathbf{v})+ +\frac{\delta^2}{2}\left(\hat{S}(\boldsymbol{\alpha})\mathbf{v}+\hat{S}(\boldsymbol{\beta})\mathbf{v}\right)+\mathcal{O}(\delta^3), \tag{8.1.27}$$

so we see that the commutator applied on the vector $\mathbf{v}$ is given by

$$[\hat{R}(\delta\boldsymbol{\beta}),\hat{R}(\delta\boldsymbol{\alpha})]\mathbf{v} = \delta\boldsymbol{\beta}\times(\delta\boldsymbol{\alpha}\times\mathbf{v})-\delta\boldsymbol{\alpha}\times(\delta\boldsymbol{\beta}\times\mathbf{v})+\mathcal{O}(\delta^3). \tag{8.1.28}$$

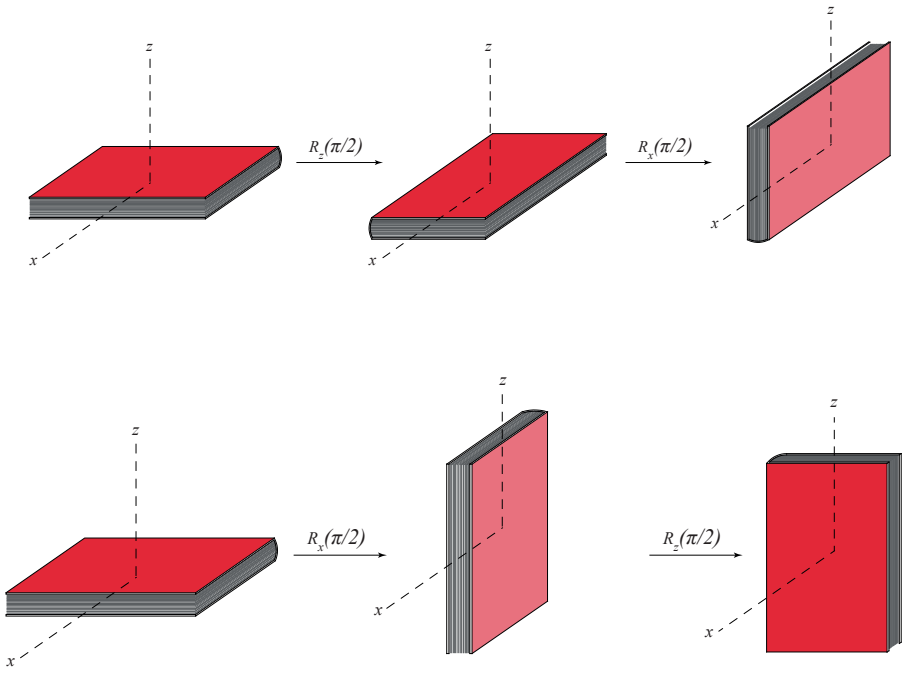

Figure 8.1.2: Rotations do not commute: in the top Figure, a book is first rotated by 90 degrees in the z direction and then by 90 degrees in x direction. In the bottom Figure, the book is first rotated by 90 degrees in the x direction and then by 90 degrees in the z direction. The final outcomes are clearly different, showing that the order in which rotations are performed is important, thus rotations along different axes do not commute.

To further evaluate this expression we recall that the cross product is not associative but it rather satisfies the Jacobi identity:

$$A \times (B \times C) + B \times (C \times A) + C \times (A \times B) \quad = \quad 0, \tag{8.1.29}$$

and also the elementary antisymmetric property

$$A \times B \quad = \quad -B \times A, \tag{8.1.30}$$

therefore

$$\begin{aligned} \delta\boldsymbol{\beta} \times (\delta\boldsymbol{\alpha} \times \mathbf{v}) - \delta\boldsymbol{\alpha} \times (\delta\boldsymbol{\beta} \times \mathbf{v}) &= \delta\boldsymbol{\beta} \times (\delta\boldsymbol{\alpha} \times \mathbf{v}) + \delta\boldsymbol{\alpha} \times (\mathbf{v} \times \delta\boldsymbol{\beta}) && (8.1.31)\\ &= -\mathbf{v} \times (\delta\boldsymbol{\beta} \times \delta\boldsymbol{\alpha}) && (8.1.32)\\ &= (\delta\boldsymbol{\beta} \times \delta\boldsymbol{\alpha}) \times \mathbf{v}. && (8.1.33) \end{aligned}$$

The commutator among the infinitesimal rotations then gives an interesting result:

$$\begin{aligned} [\hat{R}(\delta\boldsymbol{\beta}), \hat{R}(\delta\boldsymbol{\alpha})]\mathbf{v} &= (\delta\boldsymbol{\beta} \times \delta\boldsymbol{\alpha}) \times \mathbf{v} && (8.1.34)\\ &= \hat{R}(\delta\boldsymbol{\beta} \times \delta\boldsymbol{\alpha})\mathbf{v} - \mathbf{v}, && (8.1.35) \end{aligned}$$

and is therefore equivalent to a rotation of an angle $\mathcal{O}(\delta^2)$ in the direction $\boldsymbol{\beta} \times \boldsymbol{\alpha}$. We also see, instead, that rotations in the same direction ($\delta\boldsymbol{\beta}$ parallel to $\delta\boldsymbol{\alpha}$) clearly commute, this means that a non-infinitesimal rotation along a fixed direction can be found by many (N)

repeated applications of an infinitesimal rotation of size $\delta = 1/N$:

$$\begin{aligned}
\hat{R}(\boldsymbol{\theta}) &= \lim_{N\to\infty} \hat{R}(\delta\boldsymbol{\theta}) \dots \hat{R}(\delta\boldsymbol{\theta}) && (8.1.36)\\
&= \lim_{N\to\infty} \Pi_{i=1}^{N} \hat{R}(\boldsymbol{\theta}/N) && (8.1.37)\\
&= \lim_{N\to\infty} \Pi_{i=1}^{N} \left(\hat{1} + \hat{\rho}/N\right) && (8.1.38)\\
&= \exp(\hat{\rho}), && (8.1.39)
\end{aligned}$$

thus an exponential of the 3×3 matrix $\hat{\rho}$. For example, if $\boldsymbol{\theta} = (0, 0, \theta_3)$ one can compute the rotation matrix explicitly as

$$\begin{aligned}
\hat{R}(\theta_3) &\doteq \exp\theta_3 \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} && (8.1.40)\\
&= \begin{pmatrix} \cos(\theta_3) & -\sin(\theta_3) & 0 \\ \sin(\theta_3) & \cos(\theta_3) & 0 \\ 0 & 0 & 1 \end{pmatrix}, && (8.1.41)
\end{aligned}$$

which exactly coincides with the familiar expression for rotations in the z direction. Similarly for the other directions:

$$\begin{aligned}
\hat{R}(\theta_1) &\doteq \exp\theta_1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} && (8.1.42)\\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_1) & -\sin(\theta_1) \\ 0 & \sin(\theta_1) & \cos(\theta_1) \end{pmatrix}, && (8.1.43)
\end{aligned}$$

which coincides with rotations in the x direction of an angle θ_1, and finally

$$\begin{aligned}
\hat{R}(\theta_2) &\doteq \exp\theta_2 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} && (8.1.44)\\
&= \begin{pmatrix} \cos(\theta_2) & 0 & \sin(\theta_2) \\ 0 & 1 & 0 \\ -\sin(\theta_2) & 0 & \cos(\theta_2) \end{pmatrix}, && (8.1.45)
\end{aligned}$$

coinciding with rotations along the y direction. It should be remarked here that we have derived these matrices using purely algebraic techniques and only making use of the elementary property of rotations, namely condition (8.1.2).

8.2 Rotation operator

In quantum mechanics, we assume that every rotation described by $\hat{R}$ is in one to one correspondence to some operator that transforms quantum states accordingly, $\hat{D}(R)$. Since we have seen that rotations matrices are themselves uniquely identified by the vector $\boldsymbol{\theta}$, we will just call the quantum mechanical operator $\hat{D}(\boldsymbol{\theta})$. As we have done for the momentum operator and for the time evolution operator, we can then also associate a quantum mechanical rotation operator to the rotations, such that a rotated ket is written

$$|\psi_{\boldsymbol{\theta}}\rangle = \hat{D}(\boldsymbol{\theta})|\psi\rangle. \qquad (8.2.1)$$

It should be noticed that, while defined in terms of a 3×3 matrix, the operator $\hat{D}(\boldsymbol{\theta})$ however acts on the Hilbert space spanned by state vectors and not, in general, on the regular three-dimensional cartesian space. The action of this rotation operator therefore needs to be specified in terms of the requested properties on the kets.

As much as done for the other transformations we have studied so far, also for rotations we expect:

$$\begin{aligned} \langle \psi_{\boldsymbol{\theta}} | \psi_{\boldsymbol{\theta}} \rangle &= \langle \Psi | \hat{D}^\dagger(\boldsymbol{\theta}) \hat{D}(\boldsymbol{\theta}) | \Psi \rangle & (8.2.2) \\ &= \langle \Psi | \Psi \rangle. & (8.2.3) \end{aligned}$$

thus the operator is unitary:

$$\hat{D}^\dagger(\boldsymbol{\theta}) \hat{D}(\boldsymbol{\theta}) = \hat{I}. \quad (8.2.4)$$

The second property we expect from this operator, is that it can be arbitrarily *composed* :

$$\hat{D}(\boldsymbol{\theta}_1) \hat{D}(\boldsymbol{\theta}_2) = \hat{D}(\boldsymbol{\theta}_1 \boldsymbol{\theta}_2), \quad (8.2.5)$$

where $\boldsymbol{\theta}_1 \boldsymbol{\theta}_2$ denotes the composite rotation $\hat{R}(\boldsymbol{\theta}_1 \boldsymbol{\theta}_2) \equiv \hat{R}(\boldsymbol{\theta}_1) \hat{R}(\boldsymbol{\theta}_2)$. Furthermore, if we rotate a certain system *back* to its original state, this operation should be equivalent to applying the inverse transformation:

$$\hat{D}(-\boldsymbol{\theta}) = \hat{D}^{-1}(\boldsymbol{\theta}), \quad (8.2.6)$$

where $\hat{D}^{-1}$ denotes the inverse of the operator.

The last property that we can intuitively expect is that in the limit of vanishing rotations the operator $\hat{D}$ should strictly reduce to the identity

$$\lim_{\delta \to 0} \hat{D}(\delta \boldsymbol{\theta}) = \hat{I}. \quad (8.2.7)$$

As we have already seen for the case of the time evolution operator, and as a consequence of Stone's theorem, all these conditions are satisfied if we take an infinitesimal rotation operator to be described by an exponential of a Hermitian operator:

$$\hat{D}(\delta \boldsymbol{\theta}) = e^{-i \frac{\hat{\mathbf{J}} \cdot \delta \boldsymbol{\theta}}{\hbar}}, \quad (8.2.8)$$

where once again this transformation *defines* the angular momentum operators

$$\begin{aligned} \hat{\mathbf{J}} &\equiv (\hat{J}_1, \hat{J}_2, \hat{J}_3) & (8.2.9) \\ &\equiv (\hat{J}_x, \hat{J}_y, \hat{J}_z). & (8.2.10) \end{aligned}$$

8.3 Commutation Relations

In order to establish the commutation relations among the different components of the angular momentum, we recall that for infinitesimal rotation matrices we have found

$$[\hat{R}(\delta \boldsymbol{\beta}), \hat{R}(\delta \boldsymbol{\alpha})] = \hat{R}(\delta \boldsymbol{\beta} \times \delta \boldsymbol{\alpha}) - \hat{I}. \quad (8.3.1)$$

Since we have postulated a one to one correspondence between rotation matrices and rotation operators this also implies that for infinitesimal rotation operators we must have:

$$[\hat{D}(\delta \boldsymbol{\beta}), \hat{D}(\delta \boldsymbol{\alpha})] = \hat{D}(\delta \boldsymbol{\beta} \times \delta \boldsymbol{\alpha}) - \hat{I} \quad (8.3.2)$$

The left hand side of this equation is

$$\begin{aligned} [\hat{D}(\delta \boldsymbol{\beta}), \hat{D}(\delta \boldsymbol{\alpha})] &= \left(\hat{I} - \frac{i\delta}{\hbar} \hat{\mathbf{J}} \cdot \boldsymbol{\beta} \right) \left(\hat{I} - \frac{i\delta}{\hbar} \hat{\mathbf{J}} \cdot \boldsymbol{\alpha} \right) + \\ &\quad - \left(\hat{I} - \frac{i\delta}{\hbar} \hat{\mathbf{J}} \cdot \boldsymbol{\alpha} \right) \left(\hat{I} - \frac{i\delta}{\hbar} \hat{\mathbf{J}} \cdot \boldsymbol{\beta} \right) & (8.3.3) \\ &= -\frac{\delta^2}{\hbar^2} [\hat{\mathbf{J}} \cdot \boldsymbol{\beta}, \hat{\mathbf{J}} \cdot \boldsymbol{\alpha}], & (8.3.4) \end{aligned}$$

and the right hand-side is

$$\hat{D}(\delta\boldsymbol{\beta}\times\delta\boldsymbol{\alpha})-\hat{I} \;=\; -\frac{i}{\hbar}\delta^2\hat{\mathbf{J}}\cdot(\boldsymbol{\beta}\times\boldsymbol{\alpha}), \tag{8.3.5}$$

thus equating the two

$$[\hat{\mathbf{J}}\cdot\boldsymbol{\beta},\hat{\mathbf{J}}\cdot\boldsymbol{\alpha}] \;=\; i\hbar\hat{\mathbf{J}}\cdot(\boldsymbol{\beta}\times\boldsymbol{\alpha}). \tag{8.3.6}$$

This equation must hold for arbitrary directions $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, thus we can take for example them to be parallel to two of the unit vectors $\boldsymbol{\alpha}=\alpha_j\mathbf{e}_j$ and $\boldsymbol{\beta}=\beta_i\mathbf{e}_i$, yielding

$$\begin{aligned}[\hat{J}_i,\hat{J}_j] &= i\hbar\hat{\mathbf{J}}\cdot(\mathbf{e}_i\times\mathbf{e}_j) &(8.3.7)\\ &= i\hbar\hat{\mathbf{J}}\cdot\left(\sum_{i'j'k}\epsilon_{i'j'k}e_{i'}e_{j'}\mathbf{e}_k\right) &(8.3.8)\\ &= i\hbar\sum_k\epsilon_{ijk}\hat{J}_k. &(8.3.9)\end{aligned}$$

This relation is most often written dropping the explicit summation over k, since that is redundant, and it yields the following compact expression for the commutators of the different components of the angular momentum:

$$[\hat{J}_i,\hat{J}_j] \;=\; i\hbar\epsilon_{ijk}\hat{J}_k. \tag{8.3.10}$$

Equation (8.3.10) is the fundamental commutator relation for angular momentum operators and, together with the canonical commutation relations, is one of the cornerstones of quantum theory.

A remarkable difference with respect to the case of linear momentum, $\hat{\mathbf{P}}=(\hat{p}_x,\hat{p}_y,\hat{p}_z)$, is that in this case the generators of the angular momentum do not commute among themselves. The lack of commutativity of the rotation matrices $\hat{R}$ then directly implies also the lack of commutation of the $\hat{\mathbf{J}}$ operators.

8.3.1 Finite Rotations

We have analyzed so far the rotation operator for infinitesimal transformations, of order δ. As a consequence of the non-commutativity of the angular momentum operators in the different directions, also rotations along different directions do not commute and the order in which they are realized is important. However, rotations along the same direction clearly commute, thus we can write the finite rotation operator as a product of many infinitesimal transformations, in such a way that for finite rotations:

$$\hat{D}(\boldsymbol{\theta}) \;=\; e^{-i\frac{\hat{\mathbf{J}}\cdot\boldsymbol{\theta}}{\hbar}}. \tag{8.3.11}$$

8.3.2 Rotations of vector operators

In order to further understand the properties of the angular momentum operator, let us consider some vector observable

$$\begin{aligned}\hat{\mathbf{V}} &= (\hat{V}_x,\hat{V}_y,\hat{V}_z) &(8.3.12)\\ &\equiv (\hat{V}_1,\hat{V}_2,\hat{V}_3), &(8.3.13)\end{aligned}$$

with three components such as, for example, the spin operator $\hat{\mathbf{S}}=(\hat{S}_x,\hat{S}_y,\hat{S}_z)$ or the position operator $\hat{\mathbf{r}}=(\hat{x},\hat{y},\hat{z})$. Intuitively, physical properties obtained rotating states or, alternatively, rotating observables should be the same. We then require that, for two arbitrary kets $|\psi\rangle$ and $|\phi\rangle$:

$$\langle\phi_{\boldsymbol{\theta}}|\hat{\mathbf{V}}|\psi_{\boldsymbol{\theta}}\rangle \;=\; \langle\phi|\hat{\mathbf{V}}_{\boldsymbol{\theta}}|\psi\rangle, \tag{8.3.14}$$

or, in words, that the matrix elements of the operators are the same regardless of whether we rotate the kets or we rotate the operator. In components, this reads

$$\langle \phi_{\boldsymbol{\theta}} | \hat{V}_i | \psi_{\boldsymbol{\theta}} \rangle = \sum_j R_{ij}(\theta) \langle \phi | \hat{V}_j | \psi \rangle, \tag{8.3.15}$$

and expanding the rotated states in terms of $\hat{D}(\boldsymbol{\theta})$ we have

$$\langle \phi | \hat{D}^\dagger(\boldsymbol{\theta}) \hat{V}_i \hat{D}(\boldsymbol{\theta}) | \psi \rangle = \sum_j R_{ij}(\boldsymbol{\theta}) \langle \phi | \hat{V}_j | \psi \rangle, \tag{8.3.16}$$

and since this must be for all $|\psi\rangle$ and $|\phi\rangle$, this implies that

$$\hat{D}^\dagger(\boldsymbol{\theta}) \hat{V}_i \hat{D}(\boldsymbol{\theta}) = \sum_j R_{ij}(\boldsymbol{\theta}) \hat{V}_j. \tag{8.3.17}$$

If we now consider again the infinitesimal form of the rotation operator, we have that the left hand side of Eq. (8.3.17) reads

$$\hat{D}^\dagger(\boldsymbol{\theta}) \hat{V}_i \hat{D}(\boldsymbol{\theta}) \simeq \left(1 + \frac{i}{\hbar} \sum_k \delta\boldsymbol{\theta}_k \hat{J}_k\right) \hat{V}_i \left(1 - \frac{i}{\hbar} \sum_k \delta\boldsymbol{\theta}_k \hat{J}_k\right) \tag{8.3.18}$$

$$= \hat{V}_i + \frac{i}{\hbar} \sum_k \delta\boldsymbol{\theta}_k [\hat{J}_k, \hat{V}_i]. \tag{8.3.19}$$

In order to compute the right-hand side of of Eq. (8.3.17), we recall that

$$\sum_j R_{ij}(\boldsymbol{\theta}) \hat{V}_j \simeq \hat{V}_i + \left(\delta\boldsymbol{\theta} \times \hat{\mathbf{V}}\right) \cdot \mathbf{e}_i \tag{8.3.20}$$

$$= \hat{V}_i + \left(\sum_{i'j'k} \epsilon_{i'j'k} \delta\theta_{i'} \hat{V}_{j'} \mathbf{e}_k\right) \cdot \mathbf{e}_i \tag{8.3.21}$$

$$= \hat{V}_i + \sum_{i'j'} \epsilon_{i'j'i} \delta\theta_{i'} \hat{V}_{j'} \tag{8.3.22}$$

$$= \hat{V}_i + \sum_{kj} \epsilon_{kji} \delta\theta_k \hat{V}_j \tag{8.3.23}$$

$$= \hat{V}_i - \sum_{kj} \epsilon_{ijk} \delta\theta_k \hat{V}_j \tag{8.3.24}$$

Equating the two we then get

$$\frac{i}{\hbar} [\hat{J}_k, \hat{V}_i] = -\sum_j \epsilon_{ijk} \hat{V}_j \tag{8.3.25}$$

$$\frac{i}{\hbar} [\hat{J}_i, \hat{V}_k] = \sum_j \epsilon_{ijk} \hat{V}_j \tag{8.3.26}$$

$$[\hat{J}_i, \hat{V}_j] = -i\hbar \sum_j \epsilon_{ikj} \hat{V}_k \tag{8.3.27}$$

and finally exchanging $k \longleftrightarrow j$ and removing the redundant sum, we recover a more familiar form:

$$[\hat{J}_i, \hat{V}_j] = i\hbar \epsilon_{ijk} \hat{V}_k. \tag{8.3.28}$$

We therefore see that the commutator relations among different components of the angular momentum, Eq. (8.3.10), are just a special case of Eq. (8.3.28) when we take the vector operator to be the angular momentum itself: $\hat{\mathbf{V}} = \hat{\mathbf{J}}$.

8.4 General properties of Angular Momentum eigenstates

Having defined the characteristic commutation relations for the angular momentum operators, Eq. (8.3.10), we are now ready to study some general properties of its eigenvalues and eigenvectors. It is useful to introduce another operator, defined as the sum of the squares of the different components of the angular momentum:

$$\hat{J}^2 = \hat{J}_x\hat{J}_x + \hat{J}_y\hat{J}_y + \hat{J}_z\hat{J}_z \tag{8.4.1}$$
$$= \hat{J}_1^2 + \hat{J}_2^2 + \hat{J}_3^2. \tag{8.4.2}$$

This operator has the property that it commutes with all the components of the angular momentum.

Theorem 11. $[\hat{J}^2, \hat{J}_{\boldsymbol{\alpha}}] = 0$ *for any component of the angular momentum* $\boldsymbol{\alpha} = (x, y, z)$.

Proof. Consider for example,

$$\begin{aligned}
[\hat{J}^2, \hat{J}_z] &= [\hat{J}_x\hat{J}_x + \hat{J}_y\hat{J}_y + \hat{J}_z\hat{J}_z, \hat{J}_z] && (8.4.3)\\
&= [\hat{J}_x\hat{J}_x + \hat{J}_y\hat{J}_y, \hat{J}_z] && (8.4.4)\\
&= \hat{J}_x[\hat{J}_x, \hat{J}_z] + [\hat{J}_x, \hat{J}_z]\hat{J}_x + \hat{J}_y[\hat{J}_y, \hat{J}_z] + [\hat{J}_y, \hat{J}_z]\hat{J}_y && (8.4.5)\\
&= -i\hbar\hat{J}_x\hat{J}_y - i\hbar\hat{J}_y\hat{J}_x + i\hbar\hat{J}_y\hat{J}_x + i\hbar\hat{J}_x\hat{J}_y && (8.4.6)\\
&= 0. && (8.4.7)
\end{aligned}$$

The other cases are left as an exercise. □

Since $\hat{J}^2$ commutes with all the components of the angular momentum, it is possible to find simultaneous eigenkets of $\hat{J}^2$ and one of the three components. Mostly for historical reasons, it is the case that one choses to find simultaneous eigenkets of $\hat{J}^2$ and $\hat{J}_z$, but other choices are possible. In the following we will concentrate on the task of determining the eigenvalues of $\hat{J}^2$ and $\hat{J}_z$.

8.4.1 Ladder operators

Similarly to what already done when studying the harmonic oscillator, it is convenient to work here with non-hermitian operators, called ladder operators, taken to be linear combinations of the x and y components of the angular momentum:

$$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y. \tag{8.4.8}$$

These operators do not commute, and satisfy

$$\begin{aligned}
[\hat{J}_+, \hat{J}_-] &= [\hat{J}_x + i\hat{J}_y, \hat{J}_x - i\hat{J}_y] && (8.4.9)\\
&= -i[\hat{J}_x, \hat{J}_y] + i[\hat{J}_y, \hat{J}_x] && (8.4.10)\\
&= 2i[\hat{J}_y, \hat{J}_x] && (8.4.11)\\
&= 2\hbar\hat{J}_z, && (8.4.12)
\end{aligned}$$

and also

$$\begin{aligned}
[\hat{J}_z, \hat{J}_\pm] &= [\hat{J}_z, \hat{J}_x \pm i\hat{J}_y] && (8.4.13)\\
&= [\hat{J}_z, \hat{J}_x] \pm i[\hat{J}_z, \hat{J}_y] && (8.4.14)\\
&= i\hbar\hat{J}_y \pm i(-i\hbar\hat{J}_x) && (8.4.15)\\
&= \hbar(\pm\hat{J}_x + i\hat{J}_y) && (8.4.16)\\
&= \pm\hbar\hat{J}_\pm. && (8.4.17)
\end{aligned}$$

It is also easy to check that the ladder operators commute with $\hat{J}^2$, since they are just a linear combination of operators that commute with $\hat{J}^2$, thus

$$\begin{aligned}
[\hat{J}^2, \hat{J}_\pm] &= [\hat{J}^2, \hat{J}_x] \pm i[\hat{J}^2, \hat{J}_y] && (8.4.18)\\
&= 0. && (8.4.19)
\end{aligned}$$

8.4.2 Eigenvalues of $\hat{J}^2$ and $\hat{J}_z$

Armed with the ladder operators, we are now in position to derive the spectrum of eigenvalues of $\hat{J}^2$ and $\hat{J}_z$. Since we are looking for common eigenkets of these two operators, we write the eigenvalue equation as

$$\hat{J}^2|a,m\rangle = \hbar^2 a|a,m\rangle \tag{8.4.20}$$
$$\hat{J}_z|a,m\rangle = \hbar m|a,m\rangle, \tag{8.4.21}$$

where the $\hbar$ factors are introduced for dimensionality consistence. We are thus labeling the common eigenkets with $|a,m\rangle$ and the real numbers a and m are the eigenvalues of the two operators, and need to be determined from our analysis.

First of all, we analyze the action of the ladder operators on these eigenkets. For example, we can ask what is the ket resulting from the action of the ladder operators on the eigenkets. We first analyze the action of $\hat{J}_z$ on such states:

$$\hat{J}_z\hat{J}_\pm|a,m\rangle = \left([\hat{J}_z,\hat{J}_\pm]+\hat{J}_\pm\hat{J}_z\right)|a,m\rangle \tag{8.4.22}$$
$$= \pm\hbar\hat{J}_\pm|a,m\rangle + \hbar m\hat{J}_\pm|a,m\rangle \tag{8.4.23}$$
$$= \hbar(m\pm 1)\hat{J}_\pm|a,m\rangle \tag{8.4.24}$$

thus we see that the ladder operators increase or decrease the eigenvalue of $\hat{J}_z$ by $\hbar$. They are called ladder operators precisely for this reason: they go up and down in the ladder of (discrete, as we will see) eigenvalues of the z component of the angular momentum.

The behavior of $\hat{J}^2$ on the "laddered" states is different though :

$$\hat{J}^2\hat{J}_\pm|a,m\rangle = \left([\hat{J}^2,\hat{J}_\pm]+\hat{J}_\pm\hat{J}^2\right)|a,m\rangle \tag{8.4.25}$$
$$= \hbar^2 a\hat{J}_\pm|a,m\rangle, \tag{8.4.26}$$

thus the ladder operators *do not change* the eigenvalue associated to $\hat{J}^2$.

Another property we can readily identify is that

$$a \geq m^2, \tag{8.4.27}$$

as it can be demonstrated noticing that

$$\langle a,m|\hat{J}_x^2|a,m\rangle = |\hat{J}_x|a,m\rangle|^2 \tag{8.4.28}$$
$$\geq 0, \tag{8.4.29}$$

and similarly for the y component, thus

$$\hbar^2 a = \langle a,m|\hat{J}^2|a,m\rangle \tag{8.4.30}$$
$$= \langle a,m|\hat{J}_x^2|a,m\rangle + \langle a,m|\hat{J}_y^2|a,m\rangle + \hbar^2 m^2 \tag{8.4.31}$$
$$\geq \hbar^2 m^2. \tag{8.4.32}$$

Because of this inequality, it follows that for fixed a, there must be a maximum allowed value of m, that we call $m_{\max}$. Specifically, this is defined by the condition that the raising ladder operator cannot produce a state with larger eigenvalue $m_{\max}+\hbar$, otherwise this would violate the inequality (8.4.27). We thus have

$$\hat{J}_+|a,m_{\max}\rangle = 0. \tag{8.4.33}$$

Multiplying this equation from the left with $\hat{J}_-$ and using the fact that

$$\hat{J}_-\hat{J}_+ = (\hat{J}_x - i\hat{J}_y)(\hat{J}_x + i\hat{J}_y) \tag{8.4.34}$$
$$= \hat{J}_x^2 + \hat{J}_y^2 + i\hat{J}_x\hat{J}_y - i\hat{J}_y\hat{J}_x \tag{8.4.35}$$
$$= \hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z, \tag{8.4.36}$$

we have

$$\begin{aligned}\hat{J}_-\hat{J}_+|a,m_{\max}\rangle &= \left(\hat{J}^2-\hat{J}_z^2-\hbar\hat{J}_z\right)|a,m_{\max}\rangle && (8.4.37)\\ &= \hbar^2(a-m_{\max}^2-m_{\max})|a,m_{\max}\rangle && (8.4.38)\\ &= 0, && (8.4.39)\end{aligned}$$

which is verified only if $(a-m_{\max}^2-m_{\max})=(a-m_{\max}(1+m_{\max}))=0$, thus for fixed a we have that the maximum allowed value of m is given by

$$m_{\max}(1+m_{\max}) = a. \quad (8.4.40)$$

Using a similar argument, we can show that there must be also a minimum value $m_{\min}$, thus

$$\hat{J}_-|a,m_{\min}\rangle = 0, \quad (8.4.41)$$

and in this case

$$\begin{aligned}\hat{J}_+\hat{J}_- &= (\hat{J}_x+i\hat{J}_y)(\hat{J}_x-i\hat{J}_y)\\ &= \hat{J}_x^2+\hat{J}_y^2-i\hat{J}_x\hat{J}_y+i\hat{J}_y\hat{J}_x\\ &= \hat{J}^2-\hat{J}_z^2+\hbar\hat{J}_z, && (8.4.42)\end{aligned}$$

thus

$$\begin{aligned}\hat{J}_+\hat{J}_-|a,m_{\min}\rangle &= \left(\hat{J}^2-\hat{J}_z^2+\hbar\hat{J}_z\right)|a,m_{\min}\rangle && (8.4.43)\\ &= \hbar^2(a-m_{\min}^2+m_{\min})|a,m_{\min}\rangle\\ &= 0,\end{aligned}$$

yielding

$$m_{\min}(m_{\min}-1) = a. \quad (8.4.44)$$

We therefore see that we must have $m_{\max}=-m_{\min}$ in order to satisfy both Eq. (8.4.40) and Eq. (8.4.44). Not only m takes quantized values, but it is bound in

$$-j \le m \le j, \quad (8.4.45)$$

where we have defined $j \equiv m_{\max}$. Since the difference between the maximum and minimum eigenvalue must be an integer, we conclude that $2j =$ integer. Examples of possible values are

$$\begin{aligned}j=0 &\quad , \quad m=0 && (8.4.46)\\ j=\frac{1}{2} &\quad , \quad m=-\frac{1}{2},\frac{1}{2} && (8.4.47)\\ j=1 &\quad , \quad m=-1,0,1 && (8.4.48)\\ j=\frac{3}{2} &\quad , \quad m=-\frac{3}{2},-\frac{1}{2},\frac{1}{2},\frac{3}{2} && (8.4.49)\\ \dots &\quad , \quad \dots && (8.4.50)\end{aligned}$$

Eigenstates of J^2 are conventionally denoted by the integer or semi-integer j rather than a itself, and we have

$$\begin{aligned}\hat{J}^2|j,m\rangle &= \hbar^2 j(j+1)|j,m\rangle && (8.4.51)\\ \hat{J}_z|j,m\rangle &= \hbar m|j,m\rangle. && (8.4.52)\end{aligned}$$

8.4.3 Matrix Elements

We can now also determine other matrix elements for the angular momentum operators. Specifically, we have constructed the ket $|j,m\rangle$ in such a way that is an eigen-ket of $\hat{J}^2$ and $\hat{J}_z$, thus the matrix representation of these operators will be diagonal. The main remaining matrix elements to be computed are between the other components of the angular momentum, $\hat{J}_x$ and $\hat{J}_y$. It is more convenient to work again with the ladder operators. We have seen that the $\hat{J}_+$ operator increases the m component by one, thus

$$\hat{J}_+|j,m\rangle = C_+|j,m+1\rangle, \tag{8.4.53}$$

where C_+ is a constant to be determined. We can determine C_+ considering that

$$|C_+|^2 = \left|\hat{J}_+|j,m\rangle\right|^2 \tag{8.4.54}$$

$$= \langle j,m|\hat{J}_-\hat{J}_+|j,m\rangle, \tag{8.4.55}$$

and using Eq. 8.4.36, we have that

$$|C_+|^2 = \hbar^2\left[j(j+1)-m(m+1)\right], \tag{8.4.56}$$

similarly we have

$$\hat{J}_-|j,m\rangle = C_-|j,m-1\rangle, \tag{8.4.57}$$

and

$$|C_-|^2 = \hbar^2\left[j(j+1)-m^2+m\right], \tag{8.4.58}$$

$$= \hbar^2\left[j(j+1)-m(m-1)\right]. \tag{8.4.59}$$

Thus, taking a convention where these matrix elements are real, we get

$$\hat{J}_+|j,m\rangle = \hbar\sqrt{j(j+1)-m(m+1)}|j,m+1\rangle, \tag{8.4.60}$$

$$\hat{J}_-|j,m\rangle = \hbar\sqrt{j(j+1)-m(m-1)}|j,m-1\rangle. \tag{8.4.61}$$

8.5 Spin 1/2 case

In the previous sections we have developed the general theory of angular momentum operators, satisfying the general commutation relation, Eq. (8.3.10). As we have also discussed in sec. (8.3.2), the concept of angular momentum is rather general and concerns the transformation of all 3-dimensional vector operators, not only real space coordinates. In fact, we have already seen such an example, albeit in disguise. As we will now show, the spin operator is a specific example of momentum operator.

We recall that we have defined the spin 1/2 operator to be the 3-dimensional vector of operators:

$$\hat{\mathbf{S}} = \frac{\hbar}{2}(\hat{\sigma}_x,\hat{\sigma}_y,\hat{\sigma}_z), \tag{8.5.1}$$

where we recall the definition of the Pauli matrices:

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{8.5.2}$$

$$\hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{8.5.3}$$

$$\hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8.5.4}$$

We can immediately identify the three components of the spin with those of a legitimate angular momentum operator: $(\hat{S}_x, \hat{S}_y, \hat{S}_j) \equiv (\hat{J}_x, \hat{J}_y, \hat{J}_z)$. We can in fact verify that

$$[\hat{S}_\alpha, \hat{S}_\beta] = i\hbar \hat{S}_\gamma \epsilon_{\alpha\beta\gamma}. \tag{8.5.5}$$

For example, explicit computation of this commutator in one case yields:

$$\begin{aligned} \frac{\hbar^2}{4}(\hat{\sigma}_x\hat{\sigma}_y - \hat{\sigma}_y\hat{\sigma}_x) &= \frac{\hbar^2}{4}\left[\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} - \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}\right] && (8.5.6)\\ &= \frac{\hbar^2}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} && (8.5.7)\\ &= i\hbar\hat{S}_z. && (8.5.8) \end{aligned}$$

The total angular momentum is

$$\begin{aligned} \hat{S}^2 &= \frac{\hbar^2}{4}\left(\hat{\sigma}_x^2 + \hat{\sigma}_y^2 + \hat{\sigma}_z^2\right) && (8.5.9)\\ &= \frac{3}{4}\hbar^2\hat{I}, && (8.5.10) \end{aligned}$$

where we have used the fact (easy to check) that each of the Pauli matrices squares to unity. We therefore see that in this specific case the spin operator corresponds to a momentum operator with $j = \frac{1}{2}$:

$$\begin{aligned} \hat{S}^2\left|j=\frac{1}{2}, m=\pm\frac{1}{2}\right\rangle &= \hbar^2\frac{3}{4}\left|j=\frac{1}{2}, m=\pm\frac{1}{2}\right\rangle && (8.5.11)\\ \hat{S}_z\left|j=\frac{1}{2}, m=\pm\frac{1}{2}\right\rangle &= \pm\frac{\hbar}{2}\left|j=\frac{1}{2}, m=\pm\frac{1}{2}\right\rangle. && (8.5.12) \end{aligned}$$

Analogously, the ladder operators for spins 1/2 $\hat{S}_\pm = \hat{S}_x \pm i\hat{S}_y$ can also be explicitly computed:

$$\begin{aligned} \hat{S}_+ &= \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} && (8.5.13)\\ \hat{S}_- &= \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, && (8.5.14) \end{aligned}$$

and it is left as an exercise to show that these increase and decrease, respectively, the eigenvalues of $\hat{S}_z$.

8.5.1 Spins and Rotations

When we introduced the notion of spin, in the first lecture, we were already expecting the spin to be "quantum equivalent" of the intrinsic angular momentum of an object. In this sense, reconnecting the properties of the spin operator to those of the angular momentum is certainly reassuring. However, we have seen that one of the fundamental notions in quantum physics is that observables are connected to symmetry operations, and that the angular momentum is intrinsically connected to the notion of rotational operators. What kind of rotations then are connected to the spin degrees of freedom. To answer this question, let us recall that the rotation operator defining angular momentum was given by Eq. (8.3.11), thus for spin 1/2

$$\hat{D}(\boldsymbol{\theta}) = e^{-\frac{i}{\hbar}\hat{\mathbf{S}}\cdot\boldsymbol{\theta}}. \tag{8.5.15}$$

To see that this is truly a rotation of the system, we can for example consider a rotation along the z direction, thus $\boldsymbol{\theta} = (0, 0, \theta_z)$ such that

$$\hat{D}(\theta_z) = e^{-\frac{i}{\hbar}\hat{S}_z\theta_z}. \tag{8.5.16}$$

We expect that the 3 components of the spin to rotate according to the rotation matrix along the z direction, which has the well known form

$$\hat{R}(\theta_z) = \begin{pmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{8.5.17}$$

in such a way that

$$\langle \mathbf{S} \rangle_{\theta_z} = \hat{R} \langle \mathbf{S} \rangle. \tag{8.5.18}$$

We can easily verify that this equation is non-trivially satisfied by the spin operators. For example, if we take the expectation value of $\hat{S}_x$, after the rotation we have

$$\langle S_x \rangle_{\theta_z} = \langle \Psi_{\theta_z} | \hat{S}_x | \Psi_{\theta_z} \rangle \tag{8.5.19}$$

$$= \langle \Psi | e^{\frac{i}{\hbar}\hat{S}_z\theta_z} \hat{S}_x e^{-\frac{i}{\hbar}\hat{S}_z\theta_z} | \Psi \rangle, \tag{8.5.20}$$

where the form of the rotated operator is computed explicitly using the matrix representations:

$$e^{\frac{i}{\hbar}\hat{S}_z\theta_z} \hat{S}_x e^{-\frac{i}{\hbar}\hat{S}_z\theta_z} \doteq \frac{\hbar}{2} \begin{pmatrix} e^{i\frac{\theta_z}{2}} & 0 \\ 0 & e^{-i\frac{\theta_z}{2}} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} e^{-i\frac{\theta_z}{2}} & 0 \\ 0 & e^{i\frac{\theta_z}{2}} \end{pmatrix} \tag{8.5.21}$$

$$= \frac{\hbar}{2} \begin{pmatrix} e^{i\frac{\theta_z}{2}} & 0 \\ 0 & e^{-i\frac{\theta_z}{2}} \end{pmatrix} \begin{pmatrix} 0 & e^{i\frac{\theta_z}{2}} \\ e^{-i\frac{\theta_z}{2}} & 0 \end{pmatrix} \tag{8.5.22}$$

$$= \frac{\hbar}{2} \begin{pmatrix} 0 & e^{i\theta_z} \\ e^{-i\theta_z} & 0 \end{pmatrix} \tag{8.5.23}$$

$$= \frac{\hbar}{2} \begin{pmatrix} 0 & \cos\theta_z \\ \cos\theta_z & 0 \end{pmatrix} + \frac{\hbar}{2} i \begin{pmatrix} 0 & \sin\theta_z \\ -\sin\theta_z & 0 \end{pmatrix} \tag{8.5.24}$$

$$= \hat{S}_x \cos(\theta_z) - \hat{S}_y \sin(\theta_z), \tag{8.5.25}$$

thus

$$\langle \Psi_{\theta_z} | \hat{S}_x | \Psi_{\theta_z} \rangle = \langle \Psi | \hat{S}_x | \Psi \rangle \cos(\theta_z) - \langle \Psi | \hat{S}_y | \Psi \rangle \sin(\theta_z) \tag{8.5.26}$$

$$\langle S_x \rangle_{\theta_z} = \langle S_x \rangle \cos(\theta_z) - \langle S_y \rangle \sin(\theta_z), \tag{8.5.27}$$

as expected for the x component of Eq. (8.5.18).

8.6 References and Further Reading

The discussion in this Chapter is adapted from Sakurai's "Modern Quantum Mechanics" (Chapter 3, sections 3.1, 3.2, 3.5), that presents a remarkably modern way of introducing angular momentum. Cohen-Tannoudji's book also contains a short discussion on the connection between rotations and Angular Momentum in its Chapter 6, at the beginning of complement B_{VI}, even though the main topic (Chapter 6, is not presented as in Sakurai's). As you might also see from other more "traditional" textbooks, one way of presenting angular momentum is by means of orbital angular momentum, a topic we won't introduce in this course before the next Chapter. As we will see, orbital angular momentum is just a special case of the angular momentum operators presented in this Chapter. Thus, the more fundamental (and maybe more elegant, depending on your taste) way of introducing angular momentum is not through orbital angular momentum but rather through rotation operators, as we have done in the previous discussion.

Chapter 9

Orbital Angular Momentum

Orbital angular momentum is maybe the most intuitive form of angular momentum. In this Chapter we will study the properties of orbital angular momentum, using the general theory we have developed in the previous Chapter.

9.1 Orbital Angular Momentum operator

Classically, orbital angular momentum is defined as

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}, \tag{9.1.1}$$

In the quantum mechanical setting, we can use the correspondence principle to find the operator equivalent of (9.1.1). This is achieved promoting dynamical variables to operators $\mathbf{r} = (x, y, z) \to \hat{\mathbf{r}} = (\hat{x}, \hat{y}, \hat{z})$, and similarly for the momentum operator, $\mathbf{p} = (p_x, p_y, p_z) \to \hat{\mathbf{p}} = (\hat{p}_x, \hat{p}_y, \hat{p}_z)$. The components of the orbital angular momentum operator $\hat{\mathbf{L}} = (\hat{L}_x, \hat{L}_y, \hat{L}_z)$ can be found remembering that

$$(\mathbf{a} \times \mathbf{b})_i = \sum_{jk} \epsilon_{ijk} a_j b_k, \tag{9.1.2}$$

yielding

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \tag{9.1.3}$$
$$\hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \tag{9.1.4}$$
$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x. \tag{9.1.5}$$

We can then compute explicitly commutators between different components of the orbital angular momentum, just using the fundamental commutation relations between positions and momenta. For example:

$$[\hat{L}_x, \hat{L}_y] = [\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z] \tag{9.1.6}$$
$$= [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x] - [\hat{y}\hat{p}_z, \hat{x}\hat{p}_z] - [\hat{z}\hat{p}_y, \hat{z}\hat{p}_x] + [\hat{z}\hat{p}_y, \hat{x}\hat{p}_z] \tag{9.1.7}$$
$$= \hat{y}\hat{p}_x[\hat{p}_z, \hat{z}] + \hat{x}\hat{p}_y[\hat{z}, \hat{p}_z] \tag{9.1.8}$$
$$= i\hbar(-\hat{y}\hat{p}_x + \hat{x}\hat{p}_y) \tag{9.1.9}$$
$$= i\hbar\hat{L}_z. \tag{9.1.10}$$

In general, it can be shown that

$$[\hat{L}_\alpha, \hat{L}_\beta] = i\hbar\epsilon_{\alpha\beta\gamma}\hat{L}_\gamma, \tag{9.1.11}$$

thus orbital angular momentum, as an operator, satisfies the same commutator relations (8.3.10) we expect from a general angular momentum operator.

Exercise 12. Demonstrate Eq. (9.1.11).

9.2 Rotation Operator

By explicitly computing the commutator relations of the components of the orbital angular momentum, we have shown that $\hat{\mathbf{L}}$ satisfies the properties of a rotation operator. We haven't explicitly shown however what *kind* of rotations this operator performs. We will now show that $\hat{\mathbf{L}}$ is associated to rotations of the coordinate system. Let us consider for example a rotation by an angle θ_z around the z direction, such that the rotation vector reads $\boldsymbol{\theta} = (0, 0, \theta_z)$. At the beginning of the previous Chapter, we have recalled that real-space rotations along a certain direction are fully encoded by 3×3 matrices. For the case of rotations along the z axis, we have that this matrix takes the explicit form:

$$\hat{R}(\boldsymbol{\theta}) = \begin{pmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{9.2.1}$$

This matrix acts on coordinates, thus we can write rotated eigen-kets of the position operator as

$$\begin{aligned} |\mathbf{r}(\boldsymbol{\theta})\rangle &= \hat{R}(\boldsymbol{\theta})|\mathbf{r}\rangle & (9.2.2)\\ &= |\left(\cos\theta_z\right) x - \left(\sin\theta_z\right) y, \left(\sin\theta_z\right) x + \left(\cos\theta_z\right) y, z\rangle, & (9.2.3) \end{aligned}$$

which in the limit of a small rotation angle becomes

$$\begin{aligned} |\mathbf{r}(\delta\boldsymbol{\theta})\rangle &= \hat{R}(\delta\boldsymbol{\theta})|\mathbf{r}\rangle & (9.2.4)\\ &= |x - \delta\theta_z y, y + \delta\theta_z x, z\rangle, & (9.2.5) \end{aligned}$$

thus the amplitudes of a quantum state in this rotated frame read:

$$\begin{aligned} \langle\mathbf{r}(\delta\boldsymbol{\theta})|\psi\rangle &= \psi(x - \delta\theta_z y, y + \delta\theta_z x, z) & (9.2.6)\\ &= \psi(\mathbf{r}) + \delta\theta_z \left(-y\frac{\partial\psi(\mathbf{r})}{\partial x} + x\frac{\partial\psi(\mathbf{r})}{\partial y}\right) + \dots & (9.2.7) \end{aligned}$$

Now, we would like to compare this expression to what we would obtain considering the rotation operator defined in terms of the orbital angular momentum

$$\hat{D}(\boldsymbol{\theta}) = e^{-i\frac{\hat{L}_z}{\hbar}\theta_z}. \tag{9.2.8}$$

The action of the rotation operator on a basis ket is:

$$\begin{aligned} |\mathbf{r}'(\boldsymbol{\theta})\rangle &= \hat{D}(\boldsymbol{\theta})|\mathbf{r}\rangle & (9.2.9)\\ &= e^{-i\frac{\hat{L}_z}{\hbar}\theta_z}|\mathbf{r}\rangle, & (9.2.10) \end{aligned}$$

thus the amplitudes of a given quantum state in this basis (rotated by $\hat{D}$) are

$$\langle\mathbf{r}'(\boldsymbol{\theta})|\psi\rangle = \langle\mathbf{r}|e^{i\frac{\hat{L}_z}{\hbar}\theta_z}|\psi\rangle. \tag{9.2.11}$$

Notice that the expression for the amplitudes above can be interpreted in two equivalent ways: either we rotate the basis eigen-kets $|\mathbf{r}'(\boldsymbol{\theta})\rangle = \hat{D}(\boldsymbol{\theta})|\mathbf{r}\rangle$ and keep the state $|\psi\rangle$ unchanged, or we keep the basis eigen-kets unchanged and rotate the state in the *opposite direction*, thus $|\psi\rangle \to \hat{D}(-\boldsymbol{\theta})|\psi\rangle = \hat{D}^\dagger(\boldsymbol{\theta})|\psi\rangle$. In the limit of small angle, the rotation operator $\hat{D}^\dagger(\delta\boldsymbol{\theta})$ defined in terms of the z component on the angular momentum reads

$$\begin{aligned} e^{i\frac{\hat{L}_z}{\hbar}\delta\theta_z} &= e^{i\frac{\hat{L}_z}{\hbar}\delta\theta_z} & (9.2.12)\\ &= \hat{1} + \frac{i}{\hbar}\delta\theta_z\hat{L}_z + \dots & (9.2.13)\\ &= \hat{1} + \frac{i}{\hbar}\delta\theta_z\left(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x\right) + \dots & (9.2.14)\\ &= \hat{1} + \delta\theta_z\left(\hat{x}\partial_y - \hat{y}\partial_x\right) + \dots & (9.2.15) \end{aligned}$$

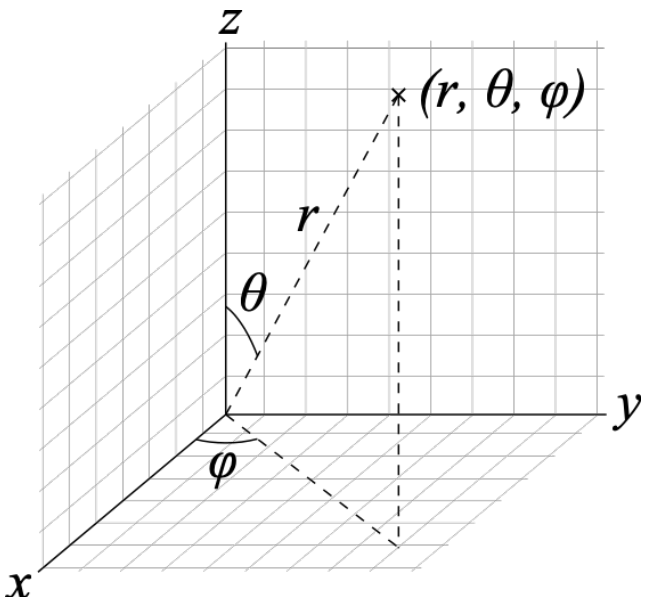

Figure 9.3.1: Spherical coordinate system adopted in these notes.

thus

$$\langle \mathbf{r}'(\delta\boldsymbol{\theta})|\psi\rangle = \langle \mathbf{r}|e^{i\frac{\hat{L}_z}{\hbar}\delta\theta_z}|\psi\rangle, \tag{9.2.16}$$

$$= \psi(\mathbf{x}) + \delta\theta_z\left(-y\frac{\partial\psi(\mathbf{r})}{\partial x} + x\frac{\partial\psi(\mathbf{r})}{\partial y}\right) \tag{9.2.17}$$

which is identical to the expression found using the rotation matrix, Eq. (9.2.7), also implying that $|\mathbf{r}'(\delta\boldsymbol{\theta})\rangle = |\mathbf{r}(\delta\boldsymbol{\theta})\rangle$. We therefore identified rotations of the coordinate system along the z axis with the action of the operator $\hat{D}(\boldsymbol{\theta})$ with $\boldsymbol{\theta} = (0, 0, \theta_z)$ and the orbital angular momentum $\hat{L}_z$.

The same argument can be repeated for all the other directions x, y and it is easy to verify that the other components of the orbital angular momentum correspond to rotations in the respective directions.

9.3 Representation of L in Spherical Coordinates

In order to analyze the eigenfunctions of the orbital angular momentum, it is much more convenient to consider the representation of $\hat{\mathbf{L}}$ in spherical coordinates (r, θ, ϕ), rather than cartesian ones (x, y, z). There are several possible conventions for spherical coordinates, here we adopt the following definition

$$\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta. \end{aligned} \tag{9.3.1}$$

A sketch of this coordinate system is found in Fig. 9.3.1.

We can then express the three components of the angular momentum operator in this system. To this end, it is necessary to consider the form of the gradient operator in spherical coordinates. Consider for example derivatives with respect to the angle ϕ, that controls

rotations around the z axis:

$$\frac{\partial}{\partial\phi} = \frac{\partial x}{\partial\phi}\frac{\partial}{\partial x} + \frac{\partial y}{\partial\phi}\frac{\partial}{\partial y} + \frac{\partial z}{\partial\phi}\frac{\partial}{\partial z} \tag{9.3.2}$$

$$= -r\sin\theta\sin\phi\frac{\partial}{\partial x} + r\sin\theta\cos\phi\frac{\partial}{\partial y} \tag{9.3.3}$$

$$= -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}. \tag{9.3.4}$$

On the other hand, we can immediately connect this result to the representation of $\hat{L}_z$ in cartesian coordinates

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \tag{9.3.5}$$

$$= -i\hbar\left(\hat{x}\frac{\partial}{\partial_y} - \hat{y}\frac{\partial}{\partial_x}\right) \tag{9.3.6}$$

$$= -i\hbar\frac{\partial}{\partial\phi}. \tag{9.3.7}$$

The z component of the orbital angular momentum has therefore a very simple expression in terms of gradients with respect to the azimuthal angle, and, in this coordinate system, it closely resemble the action of a linear momentum operator.

Deriving the other components is a straightforward, yet laborious extension of what we have already seen for the z component. The first step is to consider the gradients in polar coordinates as linear combinations of gradients in cartesian coordinates, differentiating Eq. (9.3.1) we have:

$$\begin{pmatrix}\partial_r\\ \partial_\theta\\ \partial_\phi\end{pmatrix} = \begin{pmatrix}\frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} & \frac{\partial z}{\partial r}\\ \frac{\partial x}{\partial\theta} & \frac{\partial y}{\partial\theta} & \frac{\partial z}{\partial\theta}\\ \frac{\partial x}{\partial\phi} & \frac{\partial y}{\partial\phi} & \frac{\partial z}{\partial\phi}\end{pmatrix}\begin{pmatrix}\partial_x\\ \partial_y\\ \partial_z\end{pmatrix} \tag{9.3.8}$$

$$= \begin{pmatrix}\sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta\\ r\cos\theta\cos\phi & r\cos\theta\sin\phi & -r\sin\theta\\ -r\sin\theta\sin\phi & r\sin\theta\cos\phi & 0\end{pmatrix}\begin{pmatrix}\partial_x\\ \partial_y\\ \partial_z\end{pmatrix}, \tag{9.3.9}$$

where you can notice that the last line of this matrix is what we had explicitly derived in (9.3.2). The second step is to consider the inverse transformation, so to express the cartesian derivatives as a linear combination of the spherical derivatives. This is obtained inverting the 3×3 (Jacobian) matrix above, and, after a lengthy and heartless calculation we omit here, we have:

$$\begin{pmatrix}\partial_x\\ \partial_y\\ \partial_z\end{pmatrix} = \begin{pmatrix}\sin\theta\cos\phi & \frac{\cos\theta\cos\phi}{r} & -\frac{\sin\phi}{r\sin\theta}\\ \sin\theta\sin\phi & \frac{\cos\theta\sin\phi}{r} & \frac{\cos\phi}{r\sin\theta}\\ \cos\theta & -\frac{\sin\theta}{r} & 0\end{pmatrix}\begin{pmatrix}\partial_r\\ \partial_\theta\\ \partial_\phi\end{pmatrix}. \tag{9.3.10}$$

The third and final step is to express the x and y components of the orbital angular momentum in terms of these derivatives, finding

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \tag{9.3.11}$$

$$= -i\hbar\left(\hat{y}\partial_z - \hat{z}\partial_y\right) \tag{9.3.12}$$

$$= -i\hbar\left(r\sin\theta\sin\phi\right)\left(\cos\theta\partial_r - \frac{\sin\theta}{r}\partial_\theta\right) + \tag{9.3.13}$$

$$+i\hbar\left(r\cos\theta\right)\left(\sin\theta\sin\phi\partial_r + \frac{\cos\theta\sin\phi}{r}\partial_\theta + \frac{\cos\phi}{r\sin\theta}\partial_\phi\right) \tag{9.3.14}$$

$$= i\hbar\left(\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right), \tag{9.3.15}$$

for the x component, and similarly for the y component

$$\begin{aligned}
\hat{L}_y &= \hat{z}\hat{p}_x - \hat{x}\hat{p}_z && (9.3.16)\\
&= -i\hbar\left(\hat{z}\partial_x - \hat{x}\partial_z\right) && (9.3.17)\\
&= -i\hbar\left(r\cos\theta\right)\left(\sin\theta\cos\phi\partial_r + \frac{\cos\theta\cos\phi}{r}\partial_\theta - \frac{\sin\phi}{r\sin\theta}\partial_\phi\right) + && (9.3.18)\\
&\quad +i\hbar\left(r\sin\theta\cos\phi\right)\left(\cos\theta\partial_r - \frac{\sin\theta}{r}\partial_\theta\right) && (9.3.19)\\
&= -i\hbar\left(\cos\phi\frac{\partial}{\partial\theta} - \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right). && (9.3.20)
\end{aligned}$$

With these definitions, we can also find explicitly expressions for the ladder operators

$$\begin{aligned}
\hat{L}_\pm &= \hat{L}_x \pm i\hat{L}_y\\
&= \pm\hbar e^{\pm i\phi}\left(\frac{\partial}{\partial\theta} \pm i\cot\theta\frac{\partial}{\partial\phi}\right). && (9.3.21)
\end{aligned}$$

Similarly, using the definition of $\hat{L}^2$ in terms of the ladder operators:

$$\hat{L}^2 = \hat{L}_z^2 + \frac{1}{2}\left(\hat{L}_+\hat{L}_- + \hat{L}_-\hat{L}_+\right), \quad (9.3.22)$$

and after another lengthy calculation we omit here one gets

$$\hat{L}^2 = -\hbar^2\left[\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\right]. \quad (9.3.23)$$

Exercise 13. Derive the expression in Eq. (9.3.10) explicitly as well as the expression in (9.3.23) for $\hat{L}^2$ in spherical coordinates.

9.4 Eigenfunctions

Armed with the representation of the orbital angular momentum in spherical coordinates, we are now ready to study its eigenstates. As done for the general theory of angular momentum, we consider again common eigenstates of $\hat{L}_z$ and $\hat{L}^2$, such that

$$\begin{aligned}
\hat{L}^2|l,m\rangle &= \hbar^2 l(l+1)|l,m\rangle && (9.4.1)\\
\hat{L}_z|l,m\rangle &= \hbar m|l,m\rangle. && (9.4.2)
\end{aligned}$$

As we have seen from their explicit expressions, Eqs. (9.3.23) and (9.3.7), both operators depend only on the angles θ and ϕ and are completely independent on the radial component r. This implies that also the eigenstates, in polar coordinates, will have a factorized form. At fixed value of l and m, the eigenstates of the orbital angular momentum are then the product of a function of θ and ϕ times a radial function. They are conventionally written as:

$$\langle\mathbf{r}|l,m\rangle = \Phi_m^l(r)Y_m^l(\theta,\phi), \quad (9.4.3)$$

where $\Phi_m^l(r)$ is a radial function and the functions $Y_m^l(\theta,\phi)$, encoding the angular part, are called *spherical harmonics*. The normalization condition that the eigenfunctions satisfy is, in general,

$$\begin{aligned}
\langle l,m|l,m\rangle &= \int d\mathbf{r}|\Phi_m^l(r)|^2|Y_m^l(\theta,\phi)|^2 && (9.4.4)\\
&= \left(\int_0^\infty dr|\Phi_m^l(r)|^2r^2\right)\times\left(\int_0^{2\pi}d\phi\int_0^\pi d\theta\sin(\theta)|Y_m^l(\theta,\phi)|^2\right) && (9.4.5)\\
&= N_r^{lm}\times N_{\theta,\phi}^{lm} && (9.4.6)\\
&= 1, && (9.4.7)
\end{aligned}$$

thus taking the form of a product of two normalizations, one for the radial part, and one for the angular part. The conventionally adopted choice, which is also quite convenient for all calculations, is to take the two factors identically equal to 1, thus we require

$$\int_0^\infty dr|\Phi_m^l(r)|^2 r^2 \quad = \quad 1, \tag{9.4.8}$$

for the radial part, and the normalization condition for the spherical harmonics is instead

$$\int_0^{2\pi} d\phi \int_0^\pi d\theta \sin(\theta)|Y_m^l(\theta,\phi)|^2 \quad = \quad 1 \tag{9.4.9}$$

Notice that the radial function $\Phi_m^l(r)$ cannot be determined from the general eigenvalue equations we have written above, and it is thus arbitrary, provided that the normalization condition, Eq. (9.4.8), is verified. In the following we will concentrate then only on the non-trivial angular part, and study the properties of the spherical harmonics, as well as the associated spectrum of eigenvalues l and m.

9.4.1 Eigenvalue equation for $\hat{L}_z$

We start with the case of $\hat{L}_z$, for which the eigenvalue equation projected onto spherical coordinates takes the form

$$\langle \mathbf{r}|\hat{L}_z|l,m\rangle \quad = \quad \hbar m\langle \mathbf{r}|l,m\rangle. \tag{9.4.10}$$

Recalling the representation of the $\hat{L}_z$ operator in the spherical coordinates representation,

$$\hat{L}_z \quad = \quad -i\hbar\frac{\partial}{\partial\phi}, \tag{9.4.11}$$

we have that the spherical harmonics satisfy the following differential equation

$$-i\hbar\frac{\partial}{\partial\phi}\Phi_m^l(r)Y_m^l(\theta,\phi) \quad = \quad \hbar m\Phi_m^l(r)Y_m^l(\theta,\phi), \tag{9.4.12}$$

thus we see that it is independent of the radial part,

$$-i\hbar\frac{\partial}{\partial\phi}Y_m^l(\theta,\phi) \quad = \quad \hbar m Y_m^l(\theta,\phi). \tag{9.4.13}$$

Moreover, this equation does not carry any differential dependence on θ, thus it is satisfied by separation of variables

$$Y_m^l(\theta,\phi) \quad = \quad \chi_m^l(\theta)e^{im\phi}. \tag{9.4.14}$$

From this expression we can also make a very important deduction on the possible values taken by m and l. The general theory of the angular momentum tells us that m is either integer or semi-integer and takes values in

$$-l \leq m \leq l. \tag{9.4.15}$$

However, for orbital angular momentum there is a little surprise! If we assume that the eigenfunctions of the angular momentum are single-valued (an assumption which is essential if we wish to use these functions as a basis in which to expand arbitrary wave-functions) we must have that

$$Y_m^l(\theta,\phi+2\pi) \quad = \quad Y_m^l(\theta,\phi), \tag{9.4.16}$$

thus $e^{im2\pi} = 1$ and m must be an integer, ruling out the possibility of a semi integer value. In turn, this implies that l itself is an integer, and the allowed eigenvalues are

$$\boxed{m = -l\cdots -3,-2,-1,0,1,2,3,\ldots l}. \tag{9.4.17}$$

It should remarked that this spectrum of eigenvalues is in stark contrast with what happens for spins, that instead are allowed to take also semi-integer values of orbital angular momentum and not only integer values. This also explains why the Stern and Gerlach experiment was a "smoking-gun" (direct proof) for the existence of an intrinsic angular momentum, the spin, of the electron as opposed to the orbital angular momentum. The observation of an *even* number of possible values of m ($m = \pm 1/2$,in the SG experiment) indeed is not compatible with orbital angular momentum, that for any value of l allows only for an *odd number* of m states.

9.4.2 General eigenvalue equation for $\hat{L}^2$

The other equation satisfied by the spherical harmonics is the eigenfunction condition for $\hat{L}^2$

$$\langle \theta, \phi | \hat{L}^2 | l, m \rangle = \hbar^2 l(l+1) \langle \theta, \phi | l, m \rangle, \tag{9.4.18}$$

and using the explicit form for $\hat{L}^2$ in spherical coordinates, Eq. (9.3.23), we get the following differential equation

$$\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} Y_m^l(\theta,\phi) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial}{\partial\theta} Y_m^l(\theta,\phi)\right) + \\ + l(l+1) Y_m^l(\theta,\phi) = 0 \tag{9.4.19}$$

however recalling that the ϕ dependence is fixed by Eqs. (9.4.13), we have

$$\frac{\partial^2}{\partial\phi^2} Y_m^l(\theta,\phi) = im \frac{\partial}{\partial\phi} Y_m^l(\theta,\phi) \tag{9.4.20}$$

$$= -m^2 Y_m^l(\theta,\phi), \tag{9.4.21}$$

thus we can completely remove the ϕ dependence from Eq. (9.4.19), leading to

$$\sin\theta \frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial}{\partial\theta}\chi_m^l(\theta)\right) + \left[l(l+1)\sin^2\theta - m^2\right]\chi_m^l(\theta) = 0. \tag{9.4.22}$$

This differential equation is equivalent to the associated Legendre equation, and its solution is denoted $P_l^m(\cos\theta) \equiv \chi_m^l(\theta)$ and can be found in many textbooks. Then, apart from a normalization factor, we have $Y_m^l(\theta) \propto e^{im\phi} P_l^m(\cos\theta)$. The overall normalization can be found recalling the orthonormality conditions

$$\langle l, m | l', m' \rangle = \delta_{ll'}\delta_{mm'}, \tag{9.4.23}$$

implying

$$\int_0^{2\pi}\int_0^{\pi} Y_\ell^m(\theta,\phi)^* Y_{\ell'}^{m'}(\theta,\phi)\sin\theta d\theta d\phi = \delta_{\ell,\ell'}\delta_{m,m'} \tag{9.4.24}$$

and can be used to fix the overall normalization of the spherical harmonics. Furthermore, it is customary in physics literature to take a phase convention such that the spherical harmonics are complex valued and satisfy

$$Y_l^{-m}(\theta,\phi) = (-1)^m Y_l^m(\theta,\phi)^\star. \tag{9.4.25}$$

Overall, these two conditions fix the final form to be

$$\boxed{Y_\ell^m(\theta,\phi) = (-1)^{(m+|m|)/2}\left[\frac{(2\ell+1)(\ell-|m|)!}{4\pi(\ell+|m|)!}\right]^{1/2} e^{im\phi} P_\ell^{|m|}(\cos\theta)} \tag{9.4.26}$$

where $P_l^m(\cos\theta)$ are associated Legendre functions.

We quote some of the first few spherical harmonics, that can be useful in exercises. The lowest spherical harmonic is just a constant,

$$Y_0^0(\theta,\phi) = \frac{1}{\sqrt{4\pi}}, \tag{9.4.27}$$

then for $l=1$ we have

$$\begin{cases} Y_{\pm1}^1(\theta,\varphi) = \mp\sqrt{\dfrac{3}{8\pi}}\sin\theta \mathrm{e}^{\pm i\varphi} \\ Y_0^1(\theta,\varphi) = \sqrt{\dfrac{3}{4\pi}}\cos\theta \end{cases}, \tag{9.4.28}$$

and for $l=2$

$$\begin{cases} Y_{\pm2}^2(\theta,\varphi) = \sqrt{\dfrac{15}{32\pi}}\sin^2\theta \mathrm{e}^{\pm 2i\varphi} \\ Y_{\pm1}^2(\theta,\varphi) = \mp\sqrt{\dfrac{15}{8\pi}}\sin\theta\cos\theta \mathrm{e}^{\pm i\varphi} \\ Y_0^2(\theta,\varphi) = \sqrt{\dfrac{5}{16\pi}}\left(3\cos^2\theta - 1\right). \end{cases} \tag{9.4.29}$$

Higher spherical harmonics can be found in books, if necessary.

9.4.3 Recursive relation

An alternative approach to derive explicit expressions for $\chi_m^l(\theta)$, as in Eq. (9.4.14), is based on the ladder operators, similar to what we have already done for the harmonic oscillator. Specifically, we know that for the maximum allowed value of m, $(m=l)$ we must have

$$\hat{L}_+|l,l\rangle = 0, \tag{9.4.30}$$

thus

$$\langle \mathbf{r}|\hat{L}_+|l,l\rangle = \hbar e^{i\phi}\Phi_m^l(r)\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right)\chi_l^l(\theta)e^{il\phi} \tag{9.4.31}$$

$$= \mathrm{F}_m^l(r,\phi)\left(\frac{\partial}{\partial\theta}\chi_l^l(\theta) - l\cot\theta\chi_l^l(\theta)\right) \tag{9.4.32}$$

$$= 0. \tag{9.4.33}$$

We then have that the differential equation satisfied by χ_l^l is

$$\left(\frac{\partial}{\partial\theta}\chi_l^l(\theta) - l\cot\theta\chi_l^l(\theta)\right) = 0, \tag{9.4.34}$$

which has solution (easy to check)

$$\chi_l^l(\theta) = c_{ll}(\sin\theta)^l, \tag{9.4.35}$$

where c_{ll} is a normalization constant that can be determined imposing the normalization condition, Eq. (9.4.9). Omitting the explicit calculation of the normalization constant, the spherical harmonic in this case reads:

$$Y_l^l(\theta,\phi) = c_{ll}e^{il\phi}(\sin\theta)^l. \tag{9.4.36}$$

The spherical harmonics for smaller values of m can then be found by repeated applications of $\hat{L}_-$, since we know from the general theory of angular momentum that

$$\hat{L}_-|l,m\rangle = C_-(l,m)|l,m-1\rangle, \tag{9.4.37}$$

with $C_-(l,m) = \hbar\sqrt{l(l+1) - m(m-1)}$. We then find:

$$Y^l_{m-1}(\theta,\phi) = \frac{1}{C_-(l,m)} \times e^{-i\phi}\left(\frac{\partial}{\partial\theta} - i\cot\theta\frac{\partial}{\partial\phi}\right) Y^l_m(\theta,\phi) \tag{9.4.38}$$

$$\chi^l_{m-1}(\theta) = \frac{c_{lm}}{C_-(l,m)} \times \left(\frac{\partial}{\partial\theta}\chi^l_m(\theta) + m\cot\theta\chi^l_m(\theta)\right). \tag{9.4.39}$$

The latter expression is a recursive relation that allows us to systematically compute all the spherical harmonics, starting from the explicit expression we found for $Y^l_l(\theta,\phi)$.

9.4.4 Some properties of Spherical Harmonics

While in the previous discussion we have only quoted the final result for the spherical harmonics, since its derivation is not conceptually interesting beyond the mathematical aspect, it is important though to know some general properties of the orbital angular momentum eigenfunctions.

One important property is that spherical harmonics are orthonormal, which implies

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin(\theta) Y^l_m(\theta,\phi)^\star Y^{l'}_{m'}(\theta,\phi) = \delta_{ll'}\delta_{mm'} \tag{9.4.40}$$

and that all functions $F(\theta,\phi)$ of the solid angles θ and ϕ can be written as a linear combination of these basis functions:

$$F(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_{lm} Y^l_m(\theta,\phi), \tag{9.4.41}$$

or in ket form

$$\langle\phi,\theta|F\rangle = \sum_{lm}\langle\phi,\theta|l,m\rangle\langle l,m|F\rangle \tag{9.4.42}$$

thus the coefficients $c_{lm} = \langle l,m|F\rangle$ read

$$\langle l,m|F\rangle = \int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta d\theta Y^l_m(\theta,\varphi)^\star F(\theta,\varphi). \tag{9.4.43}$$

From the general expressions for the spherical harmonics, Eq. (9.4.26), we can also immediately notice that spherical harmonics with $m = 0$ are purely real. This results from the fact that the normalization constant c_{ll} defined above has an arbitrary phase, which is traditionally fixed in such a way that

$$Y_l^{-m}(\theta,\phi) = (-1)^m Y_l^m(\theta,\phi)^\star. \tag{9.4.44}$$

Since the square modulus of the spherical harmonics does not depend on the angle ϕ, a useful way of plotting them is presented in Fig. 9.4.1. From this Figure it can be noticed that $l = 0$ state, also known as "s state", is spherically symmetric, thus it has no preferential angular direction. The $l = 1$ states, known as "p states", instead have different θ dependent shapes. For $l = 1$, $m = 0$ for example we can see that that there are two lobes, such that they have a zero in the xy plane.

9.5 References and Further Reading

The discussion in this Chapter shows the main conceptual steps required to construct the spherical harmonics, but some more laborious (and less interesting) steps have not been reproduced. Cohen-Tannoudji's book contains, in Chapter 6, a thorough discussion of all these technical aspects. The interested reader is then invited to look in there for more details.

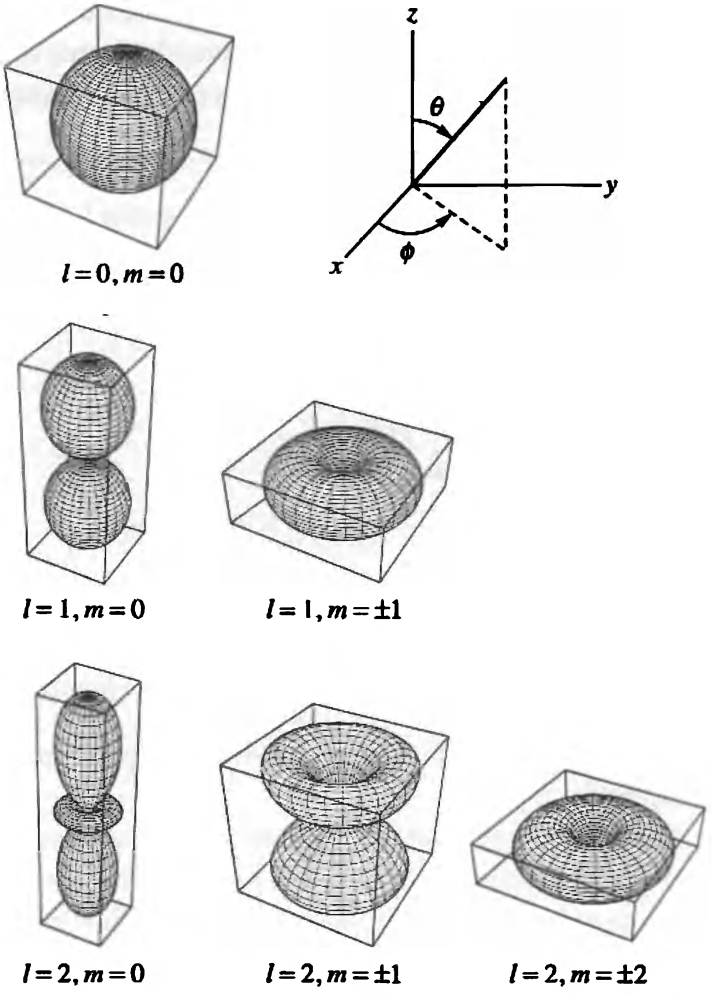

Figure 9.4.1: Polar plots of $|Y_l^m(\theta, \phi)|^2$

Chapter 10

Central Potentials

In this Chapter we will study the Schroedinger equation for a particle in 3 dimensions and subjected to a central potential $V(r)$, thus depending only on the distance from the center of coordinates $r = \sqrt{x^2 + y^2 + z^2}$. This is a case of paramount importance in physics, since potentials such as the Coulomb interaction intrinsically have radial symmetry as in this case. Analogously to the classical case, the Hamiltonian takes the form

$$\hat{H} = \frac{\hat{p}^2}{2M} + V(|\hat{\mathbf{r}}|), \tag{10.0.1}$$

with the first term corresponding the kinetic energy,

$$\begin{aligned} \hat{K} &= \frac{\hat{p}^2}{2M} && (10.0.2)\\ &= -\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) && (10.0.3)\\ &= -\frac{\hbar^2}{2M}\nabla^2, && (10.0.4) \end{aligned}$$

where we have called M the mass of the particle.

Since the potential depends only on the distance $r = |\hat{\mathbf{r}}|$, and not on the angle, it is clear that when rotating the coordinate system of an arbitrary angle the potential energy will not change. This implies that the potential energy commutes with the rotation operator $\hat{D}(\boldsymbol{\theta})$ and in turn that also the components of the orbital angular momentum must commute with $\hat{V}$:

$$[\hat{L}_\alpha, \hat{V}] = 0, \tag{10.0.5}$$

and as a consequence

$$[\hat{L}^2, \hat{V}] = 0. \tag{10.0.6}$$

Moreover, the kinetic term also commutes with the rotation operator $\hat{D}(\boldsymbol{\theta})$, since $\hat{p}^2$ is the norm of the vector $\hat{\mathbf{p}} = (\hat{p}_x, \hat{p}_y, \hat{p}_z)$ and by construction we have seen that rotations in any direction do not change the norm of vectors, thus

$$\begin{aligned} [\hat{L}_\alpha, \hat{p}^2] &= 0 && (10.0.7)\\ [\hat{L}^2, \hat{p}^2] &= 0. && (10.0.8) \end{aligned}$$

It then follows that the whole Hamiltonian commutes with the orbital angular momentum operators

$$\begin{aligned} [\hat{L}_\alpha, \hat{H}] &= 0 && (10.0.9)\\ [\hat{L}^2, \hat{H}] &= 0. && (10.0.10) \end{aligned}$$

The reader interested in a formal proof of these commutators can find it in 10.6.1.

Because of the last two relations, we can immediately deduce that the states $|l,m\rangle$ we have found in the previous Chapter are also eigenstates of the Hamiltonian, and that a generic eigenstate of the Hamiltonian is therefore of the form we previously found:

$$\langle \mathbf{r}|l,m\rangle = \Phi^l(r)Y^l_m(\theta,\phi). \tag{10.0.11}$$

In the previous Chapter we had seen that the radial part, $\Phi^l(r)$, is an arbitrary function that cannot be determined diagonalizing the orbital angular momentum only. In this Chapter we will show that the radial function satisfies a one-dimensional Schroedinger equation (for the radial variable r) with a modified potential energy, that we can call $V_{\text{eff}}(r)$ and that we will determine in the following.

10.1 Kinetic energy

The expression of the kinetic energy can be greatly simplified using the angular momentum operator. In order to do so, we consider the important identity

$$\hat{L}^2 = \hat{r}^2\hat{p}^2 - (\hat{\mathbf{r}}\cdot\hat{\mathbf{p}})^2 + i\hbar\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}, \tag{10.1.1}$$

whose demonstration is straightforward but a bit lengthy, and can be found in 10.6.2. Identity (10.1.1) can be projected onto a ket in spherical coordinates space, $|\mathbf{r}\rangle = |r,\theta,\phi\rangle$ and we obtain

$$\left\langle \mathbf{r}\left|\hat{L}^2\right|\psi\right\rangle = \left\langle \mathbf{r}\left|\left[\hat{r}^2\hat{p}^2 - (\hat{\mathbf{r}}\cdot\hat{\mathbf{p}})^2 + i\hbar\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}\right]\right|\psi\right\rangle. \tag{10.1.2}$$

However the last term is easily expressed in spherical coordinates, recalling that

$$r\frac{\partial}{\partial r} = r\left(\frac{\partial x}{\partial r}\frac{\partial}{\partial x} + \frac{\partial y}{\partial r}\frac{\partial}{\partial y} + \frac{\partial z}{\partial r}\frac{\partial}{\partial z}\right) \tag{10.1.3}$$

$$= x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z} \tag{10.1.4}$$

$$= \mathbf{r}\cdot\nabla, \tag{10.1.5}$$

thus

$$\langle \mathbf{r}|\hat{\mathbf{r}}\cdot\hat{\mathbf{p}}|\psi\rangle = -i\hbar\mathbf{r}\cdot\nabla\langle\mathbf{r}\mid\psi\rangle \tag{10.1.6}$$

$$= -i\hbar\left(r\frac{\partial}{\partial r}\langle\mathbf{r}\mid\psi\rangle\right), \tag{10.1.7}$$

and similarly for the second term

$$\left\langle \mathbf{r}\left|(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}})^2\right|\psi\right\rangle = \left\langle \mathbf{r}\left|(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}})(\hat{\mathbf{r}}\cdot\hat{\mathbf{p}})\right|\psi\right\rangle \tag{10.1.8}$$

$$= -\hbar^2\left(r\frac{\partial}{\partial r}\right)\left(r\frac{\partial}{\partial r}\right)\langle\mathbf{r}\mid\psi\rangle. \tag{10.1.9}$$

Combining all these results we arrive to the following expression for the kinetic energy operator in the Hamiltonian

$$\frac{1}{2M}\left\langle \mathbf{r}\left|\hat{p}^2\right|\psi\right\rangle = -\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right)\langle\mathbf{r}\mid\psi\rangle + \frac{\left\langle \mathbf{r}\left|\hat{L}^2\right|\psi\right\rangle}{2Mr^2}. \tag{10.1.10}$$

Physically speaking, this kinetic energy contains two parts that are easy to recognize: a term corresponding to the rotational kinetic energy (the second term) and another term that corresponds to the radial kinetic energy.

10.2 Radial equation

We have previously seen that the Hamiltonian commutes with $\hat{L}^2$, thus the eigenfunctions of the Hamiltonian must be also eigenfunctions of $\hat{L}^2$ and all the components $\hat{L}_\alpha$ of the angular momentum. The time-independent Schroedinger equation

$$-\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)\psi(\mathbf{r})+\frac{\left\langle \mathbf{r}\left|\hat{L}^2\right|\psi\right\rangle}{2Mr^2}+V(r)\psi(\mathbf{r}) \quad = \quad E\psi(\mathbf{r}), \tag{10.2.1}$$

can be then simplified noticing that $\hat{L}^2$ is acting on its eigenstate, thus

$$-\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)\psi(\mathbf{r})+\frac{\hbar^2 l(l+1)}{2Mr^2}\psi(\mathbf{r})+V(r)\psi(\mathbf{r}) \quad = \quad E\psi(\mathbf{r}). \tag{10.2.2}$$

Moreover, we have seen that all eigenfunction of the Hamiltonians are of the separable form

$$\psi(\mathbf{r}) \quad = \quad \Phi^l(r)Y^l_m(\theta,\phi), \tag{10.2.3}$$

where $\Phi^l(r)$ is a radial function to be determined, and $Y^l_m(\theta,\phi)$ are the spherical harmonics we have derived previously.

Substituting then Eq. (10.2.3) into Eq. (10.2.2) we have that the following differential equation has to be satisfied by the radial part:

$$\left[-\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)+\frac{l(l+1)\hbar^2}{2Mr^2}+V(r)\right]\Phi^l(r) \quad = \quad E\Phi^l(r). \tag{10.2.4}$$

Notice that from the beginning we have dropped the index m in the radial part ($\Phi^l_m(r)\to\Phi^l(r)$), and it is clear from the equation above that indeed Φ depends only on l. This differential equation can be further simplified making the substitution

$$\Phi^l(r) \quad = \quad \frac{u^l(r)}{r}, \tag{10.2.5}$$

yielding an effective one-dimensional Schrodinger equation:

$$\left[-\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial r^2}\right)+\left(\frac{l(l+1)\hbar^2}{2Mr^2}+V(r)\right)\right]u^l(r)=Eu^l(r). \tag{10.2.6}$$

10.3 Properties of the radial wave function

The 1D Schroedinger equation for the radial wave function $u^l(r)$ is an important result of the previous section, because it allows us to map the complicated 3D Schroedinger equation on a case that is easier to analyze, both for numerical and analytical studies. The radial equation (10.2.6) is entirely equivalent to a one-dimensional Schrodinger equation in the radial variable, for a particle moving in an effective potential

$$V^l_{\text{eff}}(r) \quad = \quad \frac{l(l+1)\hbar^2}{2Mr^2}+V(r). \tag{10.3.1}$$

The term $\frac{l(l+1)\hbar^2}{2Mr^2}$ is known as *centrifugal barrier* and it is a repulsive force that increases for larger angular momenta l. Intuitively, what happens then is that at higher angular momenta the effective force felt by the atom increases, and, as we will see, states of large l correspond to wave functions in which the atomic density is "pushed away" from the origin.

The radial wave function has also other general properties that we can deduce here.

10.3.1 Normalization

We start recalling that the overall wave function of the problem has been separated into a radial and an angular part, in a way that

$$\psi(\mathbf{r}) = \Phi^l(r)Y_m^l(\theta,\phi) \tag{10.3.2}$$

$$= \frac{u^l(r)}{r}Y_m^l(\theta,\phi), \tag{10.3.3}$$

and that the overall normalization for the wave function is

$$\int d\mathbf{r}|\psi(\mathbf{r})|^2 = \int_0^\infty dr r^2|\Phi^l(r)|^2 \int d\Omega|Y_m^l(\theta,\phi)|^2, \tag{10.3.4}$$

and that since the spherical harmonics are already normalized to 1 over the angular variables (see discussion in the previous Chapter), we are left with the normalization condition

$$\int_0^\infty dr|u^l(r)|^2 = 1. \tag{10.3.5}$$

We therefore see that $u^l(r)$ is, to all practical purposes (including normalization), a one-dimensional wave-function for a particle constrained in the region $0 \leq r < \infty$.

10.3.2 Small r limit

In the limit of small distances $r \to 0$, the radial equation (10.2.6) can be used to find general properties of the radial wave function. We start making the assumption that $u^l(r) \sim r^s$ for small r, where s is some power we want to determine. With this assumption, we have

$$-\frac{\hbar^2}{2M}s(s-1)r^{s-2} + \frac{l(l+1)\hbar^2}{2M}r^{s-2} + V(r)r^s = Er^s. \tag{10.3.6}$$

So far we haven't made any assumption about the central potential itself, but we need to make sure that it has some reasonably good behavior when approaching the origin. Here we make the assumption that the central potential is sufficiently regular, i.e. it can be divergent at the origin, but with a power such that at least

$$\lim_{r\to 0} r^2V(r) = 0, \tag{10.3.7}$$

notice that this is the case, for example, for the Coulomb potential $V(r) \sim 1/r$. If this is the case, then the dominating term for small values of r is the powers r^{s-2}, thus in order to *cancel* the divergence in (10.3.6), we need to have that the terms proportional r^{s-2} cancel out, thus requiring

$$s(s-1) = l(l+1), \tag{10.3.8}$$

which is satisfied for two possible values:

$$s = l+1 \tag{10.3.9}$$

$$s = -l. \tag{10.3.10}$$

This in turn implies that either $u^l(r) \sim r^{l+1}$ or $u^l(r) \sim \frac{1}{r^l}$, for small distances. However, recalling that $l \geq 0$, we now argue that only the first solution is acceptable. If, for $l \geq 1$, the second possibility is considered, we would then have a divergence at the origin that would violate the normalization condition (10.3.5), which is clearly unacceptable. For $l = 0$, the normalization could be still satisfied, however this would imply that $\Phi^l(r) \sim 1/r$, and this comes with other issues. An intuitive argument for this is that the 3-dimensional kinetic energy would then behave as

$$\nabla^2\frac{1}{r} = -4\pi\delta^3(\mathbf{r}), \tag{10.3.11}$$

thus for the energy E to be finite, we would need the central potential to compensate the delta singularity. However, because we are considering here only regular potentials satisfying (10.3.7), we argue that $u^l(r) \sim 1/r^l$ is not an allowed behavior. In conclusion, we must have

$$u^l(r) \sim r^{l+1}, \tag{10.3.12}$$

for small r, and the probability density of finding the particle at the origin is always vanishing, since for $l \geq 0$

$$|u^l(0)|^2 = 0. \tag{10.3.13}$$

10.4 Coulomb Potential

We now specialize our discussion to the very important case of a Coulomb potential. We consider here the case in which a single electron (we ignore its spin degrees of freedom) feels the interaction potential due to a nucleus of charge Ze. If we make the assumption that the nucleus is much heavier than the electron, we can neglect the motion of the nucleus and only consider the electron motion in the field of the nucleus. Notice that this assumption is very well verified in practice ,since a proton is almost 2000 times heavier than an electron ($m_p/m_e = 1836.15\ldots$). The Hamiltonian then simply reads

$$\hat{H} = \frac{\hat{P}^2}{2m_e} - \frac{Ze^2}{|\hat{r}|}, \tag{10.4.1}$$

where m_e is the electron mass. We therefore see that this is a special case of a central potential, and that the radial equation to be satisfied is

$$\left[-\frac{\hbar^2}{2m_e}\frac{\partial^2}{\partial r^2} + \frac{l(l+1)\hbar^2}{2m_e r^2} - \frac{Ze^2}{r}\right] u^l(r) = E u^l(r). \tag{10.4.2}$$

The potential energy goes to zero at infinity, thus bound states must have negative energy. Focusing then on bound states only, with $E = -|E|$, we can get rid of dimensional factors through the following dimensionless distance:

$$\rho = \sqrt{\frac{8m_e|E|}{\hbar^2}}r. \tag{10.4.3}$$

With this substitution the radial equation becomes

$$\frac{d^2u}{d\rho^2} - \frac{l(l+1)}{\rho^2}u + \left(\frac{\lambda}{\rho} - \frac{1}{4}\right)u = 0, \tag{10.4.4}$$

where we have also introduced

$$\lambda = \frac{Ze^2}{\hbar}\sqrt{\frac{m_e}{2|E|}}. \tag{10.4.5}$$

The differential equation (10.4.4) is still not easy to solve analytically (it's very easy to solve on a computer though), but we can still deduce the spectrum of eigenvalues just looking at the asymptotic behaviors of the radial wave function.

10.4.1 Behavior at $\rho \to \infty$

In this regime, the differential equation (10.4.4) simplifies to

$$\frac{d^2u}{d\rho^2} - \frac{1}{4}u = 0, \tag{10.4.6}$$

which has solutions

$$u(\rho) = Ae^{-\rho/2} + Be^{\rho/2}. \tag{10.4.7}$$

However, we have previously seen that $u(\rho)$ must be normalizable, thus we must have $B = 0$.

10.4.2 Regular solutions

Since we know both the large ρ behavior, and the small ρ behavior (from the general discussion on the properties of the radial wave-function), we can attempt to find a solution that satisfies both boundary conditions, of the form

$$u(\rho) = \rho^{l+1} e^{-\rho/2} F(\rho), \tag{10.4.8}$$

where we can expect now $F(\rho)$ to be a smooth, non-singular function. With this further substitution we have

$$\frac{d^2F}{d\rho^2} + \left(\frac{2l+2}{\rho} - 1\right)\frac{dF}{d\rho} + \left(\frac{\lambda}{\rho} - \frac{l+1}{\rho}\right)F = 0, \tag{10.4.9}$$

which we can now attempt to solve considering a series expansion for the function F

$$F(\rho) = \sum_{k=0}^{\infty} c_k \rho^k, \tag{10.4.10}$$

with the only constraint that $c_0 \neq 0$, so to guarantee that $u(\rho \to 0) \sim \rho^{l+1}$. With this series expansion, we see that the coefficients c_k must satisfy

$$\sum_{k=2}^{\infty} k(k-1)c_k\rho^{k-2} + \sum_{k=1}^{\infty}(2l+2)kc_k\rho^{k-2} + \\ + \sum_{k=0}^{\infty}\left[-k+\lambda-(l+1)\right]c_k\rho^{k-1} = 0. \tag{10.4.11}$$

Making the change of indices $k-1 = k'$ in the first two summations and further renaming $k' \to k$, we get

$$\sum_{k=0}^{\infty}[k(k+1) + (2l+2)(k+1)]c_{k+1}\rho^{k-1} + \\ + \sum_{k=0}^{\infty}[-k+\lambda-(l+1)]c_k\rho^{k-1} = 0, \tag{10.4.12}$$

which is satisfied if

$$\frac{c_{k+1}}{c_k} = \frac{k+l+1-\lambda}{(k+1)(k+2l+2)}. \tag{10.4.13}$$

This Equation behaves like

$$\frac{c_{k+1}}{c_k} \underset{k\to\infty}{\sim} \frac{1}{k}, \tag{10.4.14}$$

however this is the same asymptotic behavior that you can expect from the function e^{ρ}, for which $c_k = \frac{1}{k!}$. This behavior is therefore again forbidden, since it would imply that $u(\rho) \sim e^{\rho}$ and thus the radial wave function would be not normalizable. The only possibility here is therefore that the series expansion, Eq. (10.4.10), terminates at some value of $k_{\max}$ such that $c_k = 0, \forall k > k_{\max}$. The termination condition implies

$$k_{\max} + l + 1 - \lambda = 0, \tag{10.4.15}$$

where $k_{\max} = 0, 1, 2\ldots$ and recalling the definition of the factor λ in terms of the energy, we finally have that the spectrum of energy eigenvalues is given by

$$E = -\frac{m_e Z^2 e^4}{2\hbar^2\left(1+l+k_{\max}\right)^2}. \tag{10.4.16}$$

Figure 10.4.1: Energy levels for the Hydrogen atom. States are labelled with the notation nl, where n is the principal quantum number and l is the orbital angular momentum index. Traditionally, instead of using $l = (0, 1, 2, 3 \dots)$ one uses the spectroscopy notation $l = (\mathrm{s}, \mathrm{p}, \mathrm{d}, \mathrm{f} \dots)$. In this diagram then the ground state is the 1s state (with principal number $n = 1$ and $l = s = 0$). We also see that there are several excited states corresponding to the same l, since for fixed $n = 1 + l + k_{\max}$ there can be multiples values of integers l and $k_{\max}$ that add up to the same n. For example, for $n = 2$ we have two states, 2s (with $l = 0$,$k_{\max} = 1$) and 2p (with $l = 1$,$k_{\max} = 0$).

This can be further simplified noticing that we can define an integer, called *principal quantum number*, such as

$$n = 1 + l + k_{\max}, \tag{10.4.17}$$

thus yielding the quantized eigenvalues

$$E_n = -\frac{m_e Z^2 e^4}{2\hbar^2 n^2}, \tag{10.4.18}$$

where now $n = 1, 2, \dots$.

For the Hydrogen atom ($Z = 1$), the numerical constant is given by

$$\frac{m_e e^4}{2\hbar^2} \simeq 13.6 \text{ [eV]}, \tag{10.4.19}$$

and a scheme of the energy levels is shown in Fig. 10.4.1.

10.5 Examples of eigenstates

To summarize, the eigenstates are indexed by the principal quantum number as well as by the angular momentum indices. It is also given by

$$\begin{aligned} \langle \mathbf{r}|n, l, m\rangle &= \Psi_{nlm}(\mathbf{r}) && (10.5.1)\\ &= \Phi_n^l(r) Y_m^l(\theta, \phi) && (10.5.2)\\ &= \frac{u_n^l(r)}{r} Y_m^l(\theta, \phi) && (10.5.3) \end{aligned}$$

where we have explicitly added the dependency of the radial functions Φ both on l and n, the principal quantum number, since we showed that

$$\begin{aligned} u_n^l(\rho) &= \rho^{l+1} e^{-\rho/2} F(\rho) && (10.5.4)\\ &= \rho^{l+1} e^{-\rho/2} \sum_{k=0}^{k_{\max}} c_k \rho^k && (10.5.5)\\ &= \rho^{l+1} e^{-\rho/2} \sum_{k=0}^{n-1-l} c_k \rho^k. && (10.5.6) \end{aligned}$$

The dimensionless radial variable is given by

$$\begin{aligned} \rho &= \sqrt{\frac{8 m_e |E|}{\hbar^2}} r && (10.5.7)\\ &= \frac{2}{n} \left(\frac{m_e e^2}{\hbar^2} \right) r && (10.5.8)\\ &= \frac{2}{n} \frac{r}{a_0} && (10.5.9) \end{aligned}$$

where we have introduced a characteristic length scale for the hydrogen atom introduced by Bohr, and known as the Bohr radius:

$$\begin{aligned} a_0 &= \frac{\hbar^2}{m_e e^2} && (10.5.10)\\ &\simeq 0.529\,[\tilde{\text{A}}]. && (10.5.11)\\ &= 52.9\,[\text{pm}]. && (10.5.12)\end{aligned}$$

For example, these expressions allow to easily find the ground state wave function, for which we have $n = 1$ and $l = 0$. The radial part for the ground state reads

$$\begin{aligned} u_1^0(\rho) &= c_0 \rho e^{-\rho/2} && (10.5.13)\\ \Phi_1^0(r) &= 2\left(\frac{1}{a_0}\right)^{3/2} e^{-\frac{r}{a_0}}, && (10.5.14)\end{aligned}$$

where the normalization constant c_0 has been explicitly computed imposing

$$\int_0^\infty dr |c_0|^2 |u_1^0(r)|^2 = 1. \quad (10.5.15)$$

Taking into account also the angular dependence through the spherical harmonic, the full wave function for the ground state is then

$$\begin{aligned} \Psi_{100}(\mathbf{r}) &= \Phi_1^0(r) Y_0^0(\theta,\phi) && (10.5.16)\\ &= \frac{1}{\sqrt{\pi} a_0^{\frac{3}{2}}} e^{-\frac{r}{a_0}}. && (10.5.17)\end{aligned}$$

From this expression we clearly see that Bohr's radius plays the role of a characteristic distance after which the probability of finding the electron is exponentially suppressed. Similarly, other low-energy eigenstates can be explicitly found. We quote here a few of them

$$\begin{aligned} \Psi_{200}(\mathbf{r}) &= \frac{1}{4\sqrt{2\pi} a_0^{\frac{3}{2}}}\left(2 - \frac{r}{a_0}\right) e^{-\frac{r}{2a_0}}, && (10.5.18)\\ \Psi_{210}(\mathbf{r}) &= \frac{1}{4\sqrt{2\pi} a_0^{\frac{3}{2}}}\left(\frac{r}{a_0}\right) e^{-\frac{r}{2a_0}} \cos\theta, && (10.5.19)\\ \Psi_{21\pm1}(\mathbf{r}) &= \frac{1}{8\sqrt{\pi} a_0^{\frac{3}{2}}}\left(\frac{r}{a_0}\right) e^{-\frac{r}{2a_0}} \sin\theta. && (10.5.20)\end{aligned}$$

10.6 Appendix: Detailed Proofs

10.6.1 Commutators of the Hamiltonian

Here we prove that the Hamiltonian for a particle in a central potential

$$\hat{H} = \frac{\hat{p}^2}{2M} + V(|\hat{\mathbf{r}}|), \quad (10.6.1)$$

commutes with $\hat{L}^2$ and all the components $\hat{L}_\alpha$ of the orbital angular momentum. To do so, we will show below that both the kinetic and the potential energy terms in the Hamiltonian commute with $\hat{L}^2$ and $\hat{L}_\alpha$.

10.6.1.1 Kinetic Part

We start by showing that

$$\begin{aligned} [\hat{L}_z, \hat{p}_x] &= [\hat{x}\hat{p}_y - \hat{y}\hat{p}_x, \hat{p}_x] && (10.6.2)\\ &= [\hat{x}\hat{p}_y, \hat{p}_x] && (10.6.3)\\ &= [\hat{x}, \hat{p}_x]\hat{p}_y && (10.6.4)\\ &= i\hbar\hat{p}_y, && (10.6.5)\end{aligned}$$

and similarly for the other orthogonal component

$$\begin{aligned}
[\hat{L}_z, \hat{p}_y] &= [\hat{x}\hat{p}_y - \hat{y}\hat{p}_x, \hat{p}_y] && (10.6.6)\\
&= -[\hat{y}\hat{p}_x, \hat{p}_y] && (10.6.7)\\
&= -[\hat{y}, \hat{p}_y]\hat{p}_x && (10.6.8)\\
&= -i\hbar\hat{p}_x. && (10.6.9)
\end{aligned}$$

Also, the z components obviously commute

$$\begin{aligned}
[\hat{L}_z, \hat{p}_z] &= [\hat{x}\hat{p}_y - \hat{y}\hat{p}_x, \hat{p}_z] && (10.6.10)\\
&= 0. && (10.6.11)
\end{aligned}$$

We therefore have that the commutator of $\hat{p}^2$ with the $\hat{L}_z$ component is

$$\begin{aligned}
[\hat{L}_z, \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2] &= \hat{p}_x[\hat{L}_z, \hat{p}_x] + [\hat{L}_z, \hat{p}_x]\hat{p}_x + \hat{p}_y[\hat{L}_z, \hat{p}_y] + [\hat{L}_z, \hat{p}_y]\hat{p}_y && (10.6.12)\\
&= i\hbar\hat{p}_x\hat{p}_y + i\hbar\hat{p}_y\hat{p}_x - i\hbar\hat{p}_y\hat{p}_x - i\hbar\hat{p}_x\hat{p}_y && (10.6.13)\\
&= 0. && (10.6.14)
\end{aligned}$$

Since the momentum squared commutes with the z component, it must commute also with the other components (for example you can think of renaming $z \to x$ and $x \to z$, such that the above commutator reads

$$\begin{aligned}
[\hat{L}_x, \hat{p}_z^2 + \hat{p}_y^2 + \hat{p}_x^2] &= [\hat{L}_x, \hat{p}^2] && (10.6.15)\\
&= 0, && (10.6.16)
\end{aligned}$$

and the same for $\hat{L}_y$. Since all components $\hat{L}_\alpha$ commute with $\hat{p}^2$ it also follows that

$$\begin{aligned}
[\hat{L}^2, \hat{p}^2] &= [\hat{L}_x^2, \hat{p}^2] + [\hat{L}_y^2, \hat{p}^2] + [\hat{L}_z^2, \hat{p}^2] && (10.6.17)\\
&= 0. && (10.6.18)
\end{aligned}$$

Moreover, the kinetic energy is just proportional (with the scalar $1/2M$) to $\hat{p}^2$, thus we have proven that the kinetic energy commutes with both $\hat{L}_\alpha$

$$\left[\hat{L}_\alpha, \frac{\hat{p}^2}{2M}\right] = 0, \tag{10.6.19}$$

and $\hat{L}^2$

$$\left[\hat{L}^2, \frac{\hat{p}^2}{2M}\right] = 0. \tag{10.6.20}$$

10.6.1.2 Potential Energy Part

The commutators of the $\hat{L}_\alpha$ operators with the coordinates are found

$$\begin{aligned}
[\hat{L}_z, \hat{x}] &= [\hat{x}\hat{p}_y - \hat{y}\hat{p}_x, \hat{x}] && (10.6.21)\\
&= -[\hat{y}\hat{p}_x, \hat{x}] && (10.6.22)\\
&= -\hat{y}[\hat{p}_x, \hat{x}] && (10.6.23)\\
&= i\hbar\hat{y}, && (10.6.24)
\end{aligned}$$

then

$$\begin{aligned}
[\hat{L}_z, \hat{y}] &= [\hat{x}\hat{p}_y - \hat{y}\hat{p}_x, \hat{y}] && (10.6.25)\\
&= [\hat{x}\hat{p}_y, \hat{y}] && (10.6.26)\\
&= \hat{x}[\hat{p}_y, \hat{y}] && (10.6.27)\\
&= -i\hbar\hat{x}, && (10.6.28)
\end{aligned}$$

and also

$$[\hat{L}_z, \hat{z}] = [\hat{x}\hat{p}_y - \hat{y}\hat{p}_x, \hat{z}] \quad (10.6.29)$$
$$= 0. \quad (10.6.30)$$

We then see that

$$[\hat{L}_z, \hat{r}^2] = [\hat{L}_z, \hat{r}^2] \quad (10.6.31)$$
$$= [\hat{L}_z, \hat{x}^2] + [\hat{L}_z, \hat{y}^2] + [\hat{L}_z, \hat{z}^2] \quad (10.6.32)$$
$$= 0, \quad (10.6.33)$$

and with a similar reasoning as before, the same must be true for all the other components, thus

$$[\hat{L}_x, \hat{r}^2] = [\hat{L}_y, \hat{r}^2] \quad (10.6.34)$$
$$= 0, \quad (10.6.35)$$

and

$$[\hat{L}^2, \hat{r}^2] = [\hat{L}_x^2, \hat{r}^2] + [\hat{L}_y^2, \hat{r}^2] + [\hat{L}_z^2, \hat{r}^2] \quad (10.6.36)$$
$$= 0. \quad (10.6.37)$$

Now if we consider that the potential is a function of the distance only, it is also a function of r^2, thus expanding the potential in a power series we see that it commutes with $\hat{L}^2$

$$[\hat{L}^2, V(r^2)] = \left[\hat{L}^2, \sum_k c_k \left(r^2\right)^k\right] \quad (10.6.38)$$
$$= \sum_k c_k \left[\hat{L}^2, \left(r^2\right)^k\right] \quad (10.6.39)$$
$$= 0, \quad (10.6.40)$$

and $\hat{L}_\alpha$

$$[\hat{L}_\alpha, V(r^2)] = \left[\hat{L}_\alpha, \sum_k c_k \left(r^2\right)^k\right] \quad (10.6.41)$$
$$= \sum_k c_k \left[\hat{L}_\alpha, \left(r^2\right)^k\right] \quad (10.6.42)$$
$$= 0. \quad (10.6.43)$$

10.6.2 $\hat{L}^2$ identity

Here we prove that

$$\hat{L}^2 = \hat{r}^2\hat{p}^2 - (\hat{\mathbf{r}} \cdot \hat{\mathbf{p}})^2 + i\hbar\hat{\mathbf{r}} \cdot \hat{\mathbf{p}}.$$

This identity can be obtained considering the definition of $\hat{L}^2$ and recalling the following relation for Levi-Civita symbols

$$\sum_i \epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}, \quad (10.6.44)$$

which can be easily proven. An explicit computation of $\hat{L}^2$ then gives :

$$
\begin{aligned}
\hat{L}^2 &= \hat{\mathbf{r}} \times \hat{\mathbf{p}} \cdot \hat{\mathbf{r}} \times \hat{\mathbf{p}} && (10.6.45)\\
&= \sum_i \sum_{jk} \epsilon_{ijk} \hat{r}_j \hat{p}_k \sum_{lm} \epsilon_{ilm} \hat{r}_l \hat{p}_m && (10.6.46)\\
&= \sum_{jklm} \left(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}\right) \hat{r}_j \hat{p}_k \hat{r}_l \hat{p}_m && (10.6.47)\\
&= \sum_{jk} \left(\hat{r}_j \hat{p}_k \hat{r}_j \hat{p}_k - \hat{r}_j \hat{p}_k \hat{r}_k \hat{p}_j\right) && (10.6.48)\\
&= \sum_{jk} \left(\hat{r}_j ([\hat{p}_k, \hat{r}_j] + \hat{r}_j \hat{p}_k) \hat{p}_k - \hat{r}_j \hat{p}_k \left([\hat{r}_k, \hat{p}_j] + \hat{p}_j \hat{r}_k\right)\right) && (10.6.49)\\
&= \sum_{jk} \left(\hat{r}_j (-i\hbar\delta_{kj} + \hat{r}_j \hat{p}_k) \hat{p}_k - \hat{r}_j \hat{p}_k \left(i\hbar\delta_{kj} + \hat{p}_j \hat{r}_k\right)\right) && (10.6.50)\\
&= \sum_{jk} \hat{r}_j \hat{r}_j \hat{p}_k \hat{p}_k - i\hbar \sum_j \hat{r}_j \hat{p}_j - i\hbar \sum_j \hat{r}_j \hat{p}_j - \sum_{jk} \hat{r}_j \hat{p}_k \hat{p}_j \hat{r}_k && (10.6.51)\\
&= \hat{r}^2 \hat{p}^2 - 2i\hbar \hat{\mathbf{r}} \cdot \hat{\mathbf{p}} - \sum_{jk} \hat{r}_j \hat{p}_j \hat{p}_k \hat{r}_k && (10.6.52)\\
&= \hat{r}^2 \hat{p}^2 - 2i\hbar \hat{\mathbf{r}} \cdot \hat{\mathbf{p}} - \sum_{jk} \hat{r}_j \hat{p}_j \left([\hat{p}_k, \hat{r}_k] + \hat{r}_k \hat{p}_k\right) && (10.6.53)\\
&= \hat{r}^2 \hat{p}^2 - 2i\hbar \hat{\mathbf{r}} \cdot \hat{\mathbf{p}} - \sum_j \hat{r}_j \hat{p}_j \left(-3i\hbar + \sum_k \hat{r}_k \hat{p}_k\right) && (10.6.54)\\
&= \hat{r}^2 \hat{p}^2 - 2i\hbar \hat{\mathbf{r}} \cdot \hat{\mathbf{p}} + 3i\hbar \hat{\mathbf{r}} \cdot \hat{\mathbf{p}} - (\hat{\mathbf{r}} \cdot \hat{\mathbf{p}})^2 && (10.6.55)\\
&= \hat{r}^2 \hat{p}^2 - (\hat{\mathbf{r}} \cdot \hat{\mathbf{p}})^2 + i\hbar \hat{\mathbf{r}} \cdot \hat{\mathbf{p}}, && (10.6.56)
\end{aligned}
$$

which concludes our proof.

10.7 References and Further Reading

The discussion in this Chapter discusses the eigenstates of central potentials, a topic which is found in all quantum mechanics books. For example, Cohen-Tannoudji's book contains, in Chapter 7, a detailed discussion.

Chapter 11

Composite Systems and Quantum Entanglement

Our discussion of quantum phenomena so far has been limited to the study of systems that consist of a single degree of freedom (for example, one spin or one particle). In this Chapter we extend the quantum formalism in order to analyze the behavior of quantum systems composed by many degrees of freedom. We will see that when the postulates of quantum mechanics are applied to systems of many particles, they give rise to interesting and counterintuitive phenomena such as quantum *entanglement*.

11.1 State space for many particles

Suppose we have two particles, labeled A and B. From the postulates of quantum mechanics we have enunciated at the beginning of this course, we know the state of the system comprising both particles, let's call it AB, must be described by a ray in a complex vector space, or in a complex Hilbert space, if A and B have continuous degrees of freedom. The natural question to ask is then, in what space does a generic state for the two particles, $|\psi_{AB}\rangle$ live in? If we call $\mathcal{H}_A$ and $\mathcal{H}_B$ the vector (Hilbert) spaces in which the quantum states of the individual particles live, then it is a postulate of quantum mechanics that a generic state vector describing the combined system lives in a space

$$\mathcal{H}_{AB} \quad = \quad \mathcal{H}_A \otimes \mathcal{H}_B. \tag{11.1.1}$$

The symbol $\otimes$ refers to a *tensor product*, a mathematical operation that combines two vector (Hilbert) spaces to produces another one. The meaning of the tensor product is more easily understood in terms of explicit basis vectors, in the case of discrete vector spaces. For this purpose, let us assume that $\mathcal{H}_A$ is spanned by a set of basis vectors $\{|\mu_1\rangle, |\mu_2\rangle, |\mu_3\rangle, \ldots |\mu_{n_A}\rangle\}$ and that $\mathcal{H}_B$ is spanned by a set of other basis vectors $\{|\nu_1\rangle, |\nu_2\rangle, |\nu_3\rangle, \ldots |\nu_{n_B}\rangle\}$. Then, the vector space $\mathcal{H}_{AB}$ is by construction spanned by basis vectors consisting of all the pairwise combinations of the basis vectors of A and B, and the basis states of the composite system are written as

$$|\mu_i\rangle \otimes |\nu_j\rangle \quad \forall \quad i \in [1, n_A], j \in [1, n_B]. \tag{11.1.2}$$

The symbol $\otimes$ is a mathematical operation known as "tensor product" or "outer product" of two vectors, that we are going to characterize more in detail in the following. For the moment, we can already see that the total number of basis states for the composite system is $n_A \times n_B$ and all quantum states in $\mathcal{H}_{AB}$ can be written as linear combinations of the composite basis states:

$$|\psi_{AB}\rangle \quad = \quad \sum_{ij} c_{ij} |\mu_i\rangle \otimes |\nu_j\rangle \tag{11.1.3}$$

$$= \quad \sum_{ij} c_{ij} |\lambda_{ij}\rangle \tag{11.1.4}$$

with c_{ij} some complex coefficients, and where we have defined the basis vectors $|\lambda_{ij}\rangle \equiv |\mu_i\rangle \otimes |\nu_j\rangle$.

In order to work with these states, we need to know how to perform inner products between states belonging to the tensor product space $\mathcal{H}_{AB}$. The inner product between two basis states is defined as

$$\begin{aligned} \langle \lambda_{ij}|\lambda_{kl}\rangle &= (\langle \mu_i| \otimes \langle \nu_j|)\,(|\mu_k\rangle \otimes |\nu_l\rangle) && (11.1.5)\\ &\equiv \langle \mu_i|\mu_k\rangle\langle \nu_j|\nu_l\rangle && (11.1.6)\\ &= \delta_{ik}\delta_{jl}. && (11.1.7) \end{aligned}$$

This definition is relatively easy to understand: the inner product is obtained as the product of the elementary (A or B) inner products. Also, it shows that the basis states of the composite system are orthogonal by construction. As a consequence, the inner product between two generic states of the composite system

$$\begin{aligned} |\phi\rangle &= \sum_{ij} b_{ij}|\lambda_{ij}\rangle, && (11.1.8)\\ |\psi\rangle &= \sum_{ij} c_{ij}|\lambda_{ij}\rangle, && (11.1.9) \end{aligned}$$

reads

$$\begin{aligned} \langle \phi|\psi\rangle &= \sum_{ij}\sum_{kl} b^{\star}_{ij} c_{kl}\langle \lambda_{ij}|\lambda_{kl}\rangle && (11.1.10)\\ &= \sum_{ij}\sum_{kl} b^{\star}_{ij} c_{kl}\delta_{ik}\delta_{jl} && (11.1.11)\\ &= \sum_{ij} b^{\star}_{ij} c_{ij}. && (11.1.12) \end{aligned}$$

We also see that the basis states of the composite system satisfy the closure relation:

$$\sum_{ij} |\lambda_{ij}\rangle\langle \lambda_{ij}| = \hat{I}. \qquad (11.1.13)$$

Formally speaking, the tensor product satisfies all the intuitive properties you might expect from a product, for example, given a scalar a and two arbitrary vectors $|v\rangle \in \mathcal{H}_A$ and $|w\rangle \in \mathcal{H}_B$ we have

$$a(|v\rangle \otimes |w\rangle) = (a|v\rangle) \otimes |w\rangle = |v\rangle \otimes (a|w\rangle)\,, \qquad (11.1.14)$$

also, it is distributive, thus

$$\begin{aligned} (|v_1\rangle + |v_2\rangle) \otimes |w\rangle &= |v_1\rangle \otimes |w\rangle + |v_2\rangle \otimes |w\rangle, && (11.1.15)\\ |v\rangle \otimes (|w_1\rangle + |w_2\rangle) &= |v\rangle \otimes |w_1\rangle + |v\rangle \otimes |w_2\rangle. && (11.1.16) \end{aligned}$$

Finally, the construction of the product state space can be generalized from the case of two particles to the case of many particles, $A, B, C \ldots$, since the composite vector (Hilbert) space will be simply given by the tensor product of the individual state spaces

$$\mathcal{H}_{ABC\ldots} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \ldots, \qquad (11.1.17)$$

and in general the resulting space will have a large dimension when we have many particles, since it is the product of the size of the individual dimensions

$$n_{ABC\ldots} = n_A \times n_B \times n_C \times \ldots. \qquad (11.1.18)$$

11.1.1 Example: two spins 1/2

Let us see an example of this formalism in the case of two spins 1/2, thus when $\mathcal{H}_A$ and $\mathcal{H}_B$ are both vector spaces of dimension 2. As basis states of the individual spins we take the eigenkets of S_z, thus the resulting tensor product space is given by the 4 states

$$|1\rangle = |+\rangle_A \otimes |+\rangle_B \tag{11.1.19}$$
$$|2\rangle = |+\rangle_A \otimes |-\rangle_B \tag{11.1.20}$$
$$|3\rangle = |-\rangle_A \otimes |+\rangle_B \tag{11.1.21}$$
$$|4\rangle = |-\rangle_A \otimes |-\rangle_B, \tag{11.1.22}$$

and a generic state of two spins is written as

$$|\psi\rangle = \sum_{k=1}^{4} c_k |k\rangle, \tag{11.1.23}$$

where, as always, by definition

$$c_k = \langle k|\psi\rangle. \tag{11.1.24}$$

For example, take

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|+\rangle_A \otimes |-\rangle_B - |-\rangle_A \otimes |+\rangle_B\right), \tag{11.1.25}$$
$$= \frac{1}{\sqrt{2}} \left(|2\rangle - |3\rangle\right). \tag{11.1.26}$$

we can can easily check that this is a physically valid state, since it is correctly normalized:

$$\langle\psi|\psi\rangle = \frac{1}{2} \left(\langle 2|2\rangle + \langle 3|3\rangle\right) \tag{11.1.27}$$
$$= 1. \tag{11.1.28}$$

11.2 Operators

So far we have introduced the state space for a system of many particles but we haven't talked about the operators that act on this space, and how they are related to the measurement process. If we have two operators $\hat{T}_A$ and $\hat{T}_B$ acting on the individual spaces, the resulting operator that acts on the product space is also written as a tensor product:

$$\hat{T}_{AB} = \hat{T}_A \otimes \hat{T}_B, \tag{11.2.1}$$

where the resulting operator $\hat{T}_{AB}$ now acts on vectors in the space $\mathcal{H}_A \otimes \mathcal{H}_B$. The composite operator acts as follows:

$$\hat{T}_{AB}|\lambda_{ij}\rangle = \left(\hat{T}_A \otimes \hat{T}_B\right) \left(|\mu_i\rangle \otimes |\nu_j\rangle\right) \tag{11.2.2}$$
$$\equiv \left(\hat{T}_A|\mu_i\rangle\right) \otimes \left(\hat{T}_B|\nu_j\rangle\right), \tag{11.2.3}$$

thus, quite naturally, each of the two operators in the product act on the kets that belong to the respective vector spaces. As a special case, notice that if we are given only an operator that acts on one of the two subsystem, this is to be understood as

$$\hat{T}'_{AB} = \hat{T}_A \otimes \hat{I}_B \tag{11.2.4}$$

if only $\hat{T}_A$ is given, and where $\hat{I}_B$ is the identity operator for subsystem B. Similarly,

$$\hat{T}''_{AB} = \hat{I}_A \otimes \hat{T}_B, \tag{11.2.5}$$

if only $\hat{T}_B$ is given. As a result, it is easy to see that these two operators, acting non-trivially only on one of the two subsystems, commute since:

$$\begin{aligned}
\hat{T}''_{AB}\hat{T}'_{AB}|\lambda_{ij}\rangle &= \left(\hat{I}_A \otimes \hat{T}_B\right)\left(\hat{T}_A \otimes \hat{I}_B\right)|\mu_i\rangle \otimes |\nu_j\rangle && (11.2.6)\\
&= \left(\hat{I}_A \otimes \hat{T}_B\right)\left(\hat{T}_A|\mu_i\rangle \otimes \hat{I}_B|\nu_j\rangle\right) && (11.2.7)\\
&= \hat{T}_A|\mu_i\rangle \otimes \hat{T}_B|\nu_j\rangle, && (11.2.8)\\
\hat{T}'_{AB}\hat{T}''_{AB}|\lambda_{ij}\rangle &= \left(\hat{T}_A \otimes \hat{I}_B\right)\left(\hat{I}_A \otimes \hat{T}_B\right)|\mu_i\rangle \otimes |\nu_j\rangle && (11.2.9)\\
&= \left(\hat{T}_A \otimes \hat{I}_B\right)\left(\hat{I}_A|\mu_i\rangle \otimes \hat{T}_B|\nu_j\rangle\right) && (11.2.10)\\
&= \hat{T}_A|\mu_i\rangle \otimes \hat{T}_B|\nu_j\rangle, && (11.2.11)
\end{aligned}$$

thus

$$[\hat{T}_A \otimes \hat{I}_B, \hat{I}_A \otimes \hat{T}_B] = 0. \quad (11.2.12)$$

11.2.1 Example: spin 1/2 operators

Let us give again a concrete example for two spins 1/2, and imagine that we are interested in studying the *total* z component of the spin. If we call $\hat{S}_z^{(A)}$ and $\hat{S}_z^{(B)}$ the spin operators for the individual spins, such that

$$\begin{aligned}
\hat{S}_z^{(A)}|m\rangle_A &= \hbar m|m\rangle_A && (11.2.13)\\
\hat{S}_z^{(B)}|m'\rangle_B &= \hbar m'|m'\rangle_B, && (11.2.14)
\end{aligned}$$

for $m, m' = \pm 1/2$, then it is natural to define the total spin as the sum of these two operators. In order to do so, however, we need to recall that these operators are acting on different spaces, thus before summing them up we need to "upgrade" them to be good operators for the composite vector space. Thus the total $\hat{S}_z^{(AB)}$ operator reads:

$$\hat{S}_z^{(AB)} = \hat{S}_z^{(A)} \otimes \hat{I}^{(B)} + \hat{I}^{(A)} \otimes \hat{S}_z^{(B)}. \quad (11.2.15)$$

It is then straightforward to see how this operator acts on a general state. For example, if we take a basis vector for the composite system, we have

$$\begin{aligned}
\hat{S}_z^{(AB)}\left(|m\rangle_A \otimes |m'\rangle_B\right) &= \left(\hat{S}_z^{(A)} \otimes \hat{I}^{(B)} + \hat{I}^{(A)} \otimes \hat{S}_z^{(B)}\right)\left(|m\rangle_A \otimes |m'\rangle_B\right) && (11.2.16)\\
&= \left(\hat{S}_z^{(A)}|m\rangle_A\right) \otimes |m'\rangle_B + |m\rangle_A \otimes \left(\hat{S}_z^{(B)}|m\rangle_B\right) && (11.2.17)\\
&= \hbar m\left(|m\rangle_A \otimes |m'\rangle_B\right) + \hbar m'\left(|m\rangle_A \otimes |m'\rangle_B\right) && (11.2.18)\\
&= \hbar(m+m')\left(|m\rangle_A \otimes |m'\rangle_B\right), && (11.2.19)
\end{aligned}$$

thus the composite state is an eigen-ket of the total spin, with an eigenvalue $\hbar(m + m')$ that is the sum of the individual eigenvalues.

11.3 Measurements

For the single-component case, we recall that the measurement process in quantum mechanics works as follows. If we have a state:

$$|\psi\rangle = \sum_i |A_i\rangle\langle A_i|\psi\rangle, \quad (11.3.1)$$

and we measure the operator $\hat{A}$ with eigenkets $|A_i\rangle$ and eigenvalues a_i then we get the result a_i with probability $P_i = |a_i|^2$. Also, after the measurement, the state collapses to the measured eigenstate, $|A_i\rangle$.

In the case of a two-particle system, there are two kind of measurements we can perform.

11.3.1 Global measurement

In the first case, we measure an operator $\hat{T} = \hat{T}_A \otimes \hat{T}_B$, thus intrinsically defined to act on the joint vector space, and in this sense corresponding to a measurement of the entire system AB. Similarly to the standard situation, then we can diagonalize the operator:

$$\hat{T}_i|T_i\rangle = t_i|T_i\rangle, \tag{11.3.2}$$

in such a way that (assuming the operator has non-degenerate spectrum)

$$|\psi\rangle = \sum_i |T_i\rangle\langle T_i|\psi\rangle, \tag{11.3.3}$$

thus a measurement will yield the state $|T_i\rangle$ with probability $P_i = |\langle T_i|\psi\rangle|^2$.

11.3.2 Partial measurement

In the second case, we can measure an operator that is defined only on one of the two subsystems, for example $\hat{T}_A$. In this sense, we are performing a partial measurement of the system, since we measure only the properties of one subpart, ignoring the rest of the system.

We can rewrite a generic state of two particles as

$$|\psi\rangle = \sum_{ij} c_{ij}|T_i^A\rangle \otimes |T_j^B\rangle \tag{11.3.4}$$

$$= \sum_i |T_i^A\rangle \otimes \left(\sum_j c_{ij}|T_j^B\rangle\right) \tag{11.3.5}$$

$$= \sum_i |\mu_i\rangle \otimes |\phi_i^B\rangle, \tag{11.3.6}$$

where we have defined

$$|\phi_i^B\rangle = \sum_j c_{ij}|T_j^B\rangle. \tag{11.3.7}$$

This expression then allows us to get a better intuition about what happens when we measure only the first subsystem (A). In that case, assuming that we measure the operator $\hat{T}_A$ with eigenvalues t_i^A, it is postulated that after the measurement the system collapses into

$$|\psi_i'\rangle \propto |T_i^A\rangle \otimes |\phi_i^B\rangle. \tag{11.3.8}$$

The probability for this to happen is postulated to be

$$P_i = \langle\phi_i^B|\phi_i^B\rangle = \sum_j c_{ij}^\star\langle T_j^B|\sum_{j'} c_{ij'}|T_{j'}^B\rangle. \tag{11.3.9}$$

$$= \sum_j |c_{ij}|^2, \tag{11.3.10}$$

which is a generalization of what we have seen for the single particle case. We can also explicitly compute the normalization of the state after the measurement, which reads

$$|\psi_i'\rangle = \frac{1}{\sqrt{\langle T_i^A|T_i^A\rangle\langle\phi_i^B|\phi_i^B\rangle}}|T_i^A\rangle \otimes |\phi_i^B\rangle. \tag{11.3.11}$$

$$= \frac{1}{\sqrt{\sum_j |c_{ij}|^2}}|T_i^A\rangle \otimes |\phi_i^B\rangle \tag{11.3.12}$$

$$= \sum_j \frac{c_{ij}}{\sqrt{P_i}}|T_i^A\rangle \otimes |T_j^B\rangle. \tag{11.3.13}$$

11.3.3 Example: Spin Measurements

Let us consider again an example for two spins. We consider the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle_A \otimes |-\rangle_B - |-\rangle_A \otimes |+\rangle_B\right), \tag{11.3.14}$$

and let us suppose that we are interested in measuring the value of the spin z on system A. As always, this measurement can yield only two possible outcomes, $m = +\hbar/2$ and $m = -\hbar/2$. This state is such that $c_{++} = c_{--} = 0$ and $c_{+-} = c_{-+} = 1/\sqrt{2}$.

The probability of obtaining the first outcome is $P_+ = |c_{++}|^2 + |c_{+-}|^2 = 1/2$, and if $m = +\hbar/2$ the system then collapses into the normalized state

$$|\psi'_+\rangle = \frac{1}{\sqrt{1/2}}\left(c_{+-}|+\rangle_A \otimes |-\rangle_B + c_{++}|+\rangle_A \otimes |+\rangle_B\right) \tag{11.3.15}$$

$$= |+\rangle_A \otimes |-\rangle_B. \tag{11.3.16}$$

In the other case, it is easy to see that the system collapses into

$$|\psi'_-\rangle = |-\rangle_A \otimes |+\rangle_B, \tag{11.3.17}$$

also with probability $P_- = |c_{-+}|^2 + |c_{--}|^2 = 1/2$.

11.3.4 Entanglement

In the previous discussion we have seen that the measurement of one part of the system directly influences the outcomes of a measurement of the other part. This is one manifestation of what is called quantum "entanglement". More specifically, a state of two spins is said to be entangled if its coefficients *cannot* be written as the product of two independent coefficients. If instead the global wave function can be written as the product of two wave-functions corresponding to the subsystems A and B, then we say that the system is "separable". For a separable state, the wave function then reads

$$|\psi\rangle_{\text{sep}} = \sum_{ij} c_{ij}|T_i^A\rangle \otimes |T_j^B\rangle \tag{11.3.18}$$

$$= \sum_{ij} c_i^{(A)} c_j^{(B)}|T_i^A\rangle \otimes |T_i^B\rangle \tag{11.3.19}$$

$$= \left(\sum_i c_i^{(A)}|T_i^A\rangle\right) \otimes \left(\sum_j c_j^{(B)}|T_j^B\rangle\right) \tag{11.3.20}$$

$$= |\phi^{(A)}\rangle \otimes |\phi^{(B)}\rangle. \tag{11.3.21}$$

If a system is separable, we also immediately see that a measurement performed on one part does not affect the other one. For example, if we measure $\hat{T}_A$ the system will collapse into some state

$$|\psi_i\rangle = |T_i^A\rangle \otimes |\phi^{(B)}\rangle, \tag{11.3.22}$$

with probability $|c_i^{(A)}|^2$, but the resulting state for the subsystem B will always be $|\phi^{(B)}\rangle$, independently on the outcome of the measurement on A.

To explicitly determine whether a state is separable or entangled, we have to check whether the matrix of coefficient factorizes or not. For example, for two spins the condition of separability reads

$$c_{++} = c_+^{(A)} c_+^{(B)} \tag{11.3.23}$$

$$c_{+-} = c_+^{(A)} c_-^{(B)} \tag{11.3.24}$$

$$c_{-+} = c_-^{(A)} c_+^{(B)} \tag{11.3.25}$$

$$c_{--} = c_-^{(A)} c_-^{(B)}, \tag{11.3.26}$$

which is verified if

$$c_{++}c_{--} - c_{+-}c_{-+} = \det\hat{c} \tag{11.3.27}$$
$$= 0. \tag{11.3.28}$$

Thus if the determinant of the coefficients c_{ij} for the state in the composite basis is zero, then the state is separable.

For example, our state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle_A \otimes |-\rangle_B - |-\rangle_A \otimes |+\rangle_B\right), \tag{11.3.29}$$

has a coefficient matrix

$$\hat{c} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & -\frac{1}{\sqrt{2}} \end{pmatrix}, \tag{11.3.30}$$

whose determinant is non-zero and it is thus an entangled state.

11.4 The Einstein-Podolsky-Rosen Paradox

In 1935, Einstein, Podolsky, and Rosen (EPR) used the counter-intuitive features of quantum entanglement to formulate a thought experiment known as the EPR paradox. They tried to use this thought experiment to argue that quantum theory cannot serve as a fundamental description of reality. Subsequently, however, it was shown that the EPR paradox is not an actual paradox; physical systems really do have the strange behavior that the thought experiment highlighted.

The EPR paradox starts with considering an entangled state, for example the state we have considered before

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle_A \otimes |-\rangle_B - |-\rangle_A \otimes |+\rangle_B\right), \tag{11.4.1}$$

for which we have seen that if we measure $\hat{S}_z$ on the first spin (A) there are two possible outcomes, both with the same probability of 1/2, and leading to the two states after the measurement

$$|\psi'_+\rangle = |+\rangle_A \otimes |-\rangle_B, \tag{11.4.2}$$
$$|\psi'_-\rangle = |-\rangle_A \otimes |+\rangle_B. \tag{11.4.3}$$

From these expressions it is clear that both $|\psi'_+\rangle$, $|\psi'_-\rangle$ are not only eigenstates of $\hat{S}_z^{(A)} \otimes \hat{I}^{(B)}$ but also of $\hat{I}^{(A)} \otimes \hat{S}^{(B)}$. It then follows that if we now measure $\hat{S}_z$ on the second spin (B), we will find (with probability one) $-\hbar/2$ when measuring the state $|\psi'_+\rangle$ and $+\hbar/2$ when measuring the state $|\psi'_-\rangle$. This thus tells us that the result of the first measurement (on spin A) has influenced directly the result of a subsequent measurement on the other spin B! In other words, if the result of the first measurement yields a $+\hbar/2$ for spin A, then with 100% probability, a subsequent measurement of spin B will always yield a value of $-\hbar/2$. If the result of first measurement yields instead $-\hbar/2$ for spin A, then with 100% probability, the spin B will later have a value of $+\hbar/2$. The Table 11.1 summarizes the situation.

The outcomes of this simple experiment are quite troubling if we recall that the axioms of quantum mechanics tell us that the wave function collapses instantaneously. Moreover, the instantaneous collapse of the wave function should happen *regardless* of the distance among the two particles, A and B, in our case. This means that, for example, we could have two particles described by the state $|\psi\rangle$ but very distant apart. We might have particle A on Earth, and particle B on a planet orbiting Alpha Centauri, about 4 light years away from us. This means that if an experimentalist measures particle A on Earth and finds $+\hbar/2$, then when the other experimentalist in Alpha Centauri measures *immediately after* its part

Measure of $\hat{S}_z^{(A)}$	Outcome of meas. $\hat{S}_z^{(A)}$	Measure of $\hat{S}_z^{(B)}$	Outcome of meas. $\hat{S}_z^{(B)}$
Possibility 1 (50% probability)	$\|+\rangle_A \otimes \|-\rangle_B$ Value $+\hbar/2$	One outcome (100% probability)	$\|+\rangle_A \otimes \|-\rangle_B$ Value $-\hbar/2$
Possibility 2 (50% probability)	$\|-\rangle_A \otimes \|+\rangle_B$ Value $-\hbar/2$	One outcome (100% probability)	$\|-\rangle_A \otimes \|+\rangle_B$ Value $+\hbar/2$

Table 11.1: Possible outcomes when measuring first $\hat{S}_z^{(A)}$ and then $\hat{S}_z^{(B)}$ on the state $|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle_A \otimes |-\rangle_B - |-\rangle_A \otimes |+\rangle_B)$. Overall, there are two possibilities, each happening with a 50% probability.

of the system, the result obtained will be $-\hbar/2$. The experimentalist in Alpha Centauri then will *immediately* know that the experimentalist on Earth has just found $+\hbar/2$ in their measurement! This leads to a paradox, since the Centaurian seems to have learned about something that has happened 4 light years aways at more or less exactly the same time that thing has happened, and not after 4 years, the time that an electromagnetic signal sent from Earth would take to reach Alpha Centauri. This experiment then seems to pose a serious issue because it is apparently incompatible with the theory of special relativity, according to which no information can travel faster than the light. This issue is also one of the main arguments of the criticism against quantum mechanics posed by the EPR paradox.

But is what we have found really a way to transmit information between two points faster than light? *Unfortunately*, it is not. Let us consider the case in which the experimentalist on A plays a trick to their friend on Alpha Centauri and *does not perform the measurement*. At the same time, B performs the measurement as always. In this case, the possible outcomes for the B measurement are two, $\pm\hbar/2$, and they have equal probabilities to happen. If the observer in B measures $-\hbar/2$ could then wrongly infer that A has measured $+\hbar/2$, as in the case before, but in this case the measurement in A has not even happened! This means that, in reality, the observer in Alpha Centauri has no practical way of determining wether the observer on Earth has performed a measurement or not, thus this scheme cannot be used as a way to transmit information. What happens instead is that in both cases (either A measures or not), the observer in B will measure $+\hbar/2$ or $-\hbar/2$ with a 50 percent probability. This discussion can be generalized also to measurements performed in other directions of the spins, and more complex measurement scenarios and, in all cases we know, there is no way to use quantum entanglement to transmit information faster than the light.

11.5 Bell's Inequality

In addition to the lack of faster-than-the-light communication, John Bell also found out that the experiment described above can be fully described in terms of some classical theory, and that the concept of "action at a distance" implied by entanglement not even necessary to describe the experimental results obtained on Earth and on Alpha Centauri! In this alternative classical theory, we can imagine that the state of each of the two particles is described, at all times, by some classical set of "hidden" variables $\mathbf{h} = [h_1, h_2, h_3, \ldots]$ that fully contain the outcome of any spin measurement that will be done on the system. Moreover, we assume the these hidden variables are "locally stored" in each particle, thus a particle on earth would have its own set of hidden variables and a particle on Alpha Centauri would have another set of hidden variables. In this "local hidden-variable theory" the measurement of a spin on Earth then just corresponds to reading the value of the corresponding components of the variables $\mathbf{h}$. In this purely classical theory then there is no intrinsic randomness in the measurement, as instead predicted by quantum mechanics.

For example, we can imagine that in practice the experiment we have described before, involving particles A and B is created locally (say on Earth) and then the particle B is set far apart, in a way careful enough not to break the properties of the physical state. If this is the case, then we can think that at the moment of the local creation of the state, each of

Possibility	$\mathbf{h}^{(A)} = [h_1^{(A)}, h_2^{(A)}, h_3^{(A)}]$	$\mathbf{h}^{(B)} = [h_1^{(B)}, h_2^{(B)}, h_3^{(B)}]$
1	$[+1, +1, +1]$	$[-1, -1, -1]$
2	$[+1, +1, -1]$	$[-1, -1, +1]$
3	$[+1, -1, +1]$	$[-1, +1, -1]$
4	$[+1, -1, -1]$	$[-1, +1, +1]$
5	$[-1, +1, +1]$	$[+1, -1, -1]$
6	$[-1, +1, -1]$	$[+1, -1, +1]$
7	$[-1, -1, +1]$	$[+1, +1, -1]$
8	$[-1, -1, -1]$	$[+1, +1, +1]$

Table 11.2: Possible outcomes of a local classical theory describing the measurements of two spins along 3 directions, $\hat{u}_1, \hat{u}_2, \hat{u}_3$.

the two particles picks up, randomly, some classical vectors

$$\mathbf{h}^{(A)} = [h_x^{(A)}, h_y^{(A)}, h_z^{(A)}] \tag{11.5.1}$$
$$\mathbf{h}^{(B)} = [h_x^{(B)}, h_y^{(B)}, h_z^{(B)}], \tag{11.5.2}$$

whose components $h_\alpha = \pm 1$ fully describe the state of each particle. A measurement of S_z on the first particle would then be simply given by $\hbar/2 \left(h_x^{(A)}\right)$, and so on. The variables h_α themselves *can* be random, but their value is assigned once for all at the moment in which we create the system, in a process we cannot control in detail, but entirely classical.

For example, to explain the experiment above, at the moment of creating the state we assign a vector of random hidden variables to A and a vector if hidden variables to B. There are just two possibilities for the assignment: the first possibility is that we have created a classical pair with

$$h_z^{(A)} = +1 \quad , \quad h_z^{(B)} = -1, \tag{11.5.3}$$

thus all measurements of S_z will yield $+\hbar/2$ over the first particle and $-\hbar/2$ over the second particle. The other possibility is that instead we have created a pair with

$$h_z^{(A)} = -1 \quad , \quad h_z^{(B)} = +1, \tag{11.5.4}$$

thus all measurements of spins will yield $-\hbar/2$ over the first particle and $+\hbar/2$ over the second particle. When A and B are far apart, then it is clear that measuring S_z it will be always such that $S_z^{(B)} = -S_z^{(A)}$, as also predicted by the quantum theory. The benefit of this classical theory is that, quite clearly, implies that there is not faster-than-light travel of information, since the information is just pre-shared, at the moment of the creation of the state, on Earth.

Does this mean that all we have studied in this course is not very useful and that instead everything is explained by some classical theory of hidden variables? Or, can we find instead a situation that is not explained by this local, classical theory but that *really* requires quantum mechanics? The answer is that we can, and it is due to fundamental work of John Bell in the 1960s. To prove it, we will need to slightly generalize our discussion and consider 3 general measurement axes, say $\hat{u}_1, \hat{u}_2, \hat{u}_3$ that are not necessarily orthogonal.

11.5.1 Predictions of the classical theory

In the classical theory, when we measure a spin in the $\hat{u}_i$ direction it means simply that we are measuring the associated hidden variable, h_i. Since we are still considering the case with the constraint $h_i^{(A)} + h_i^{(B)} = 0$, we see that there are in total 8 possible outcomes predicted by the classical theory, and summarized in Table 11.2. In general, at the moment of the creation of the system, we assume that each of these 8 possibilities is realized with some probability p_i. For example, we can imagine of having created the pair $\mathbf{h}^{(A)} = [+1, -1, -1]$

and $\mathbf{h}^{(B)} = [-1, +1, +1]$ with a probability p_4. From the Table we can also deduce some general results for the probabilities of observing some pair of values. For example, we can concentrate the case in which every time we make a measurement, we choose to measure two different directions for each observer, $u^{(A)} \neq u^{(B)}$. In general then there are 6 possible direction choices for the measurements that satisfy this condition:

$$(u^{(A)}, u^{(B)}) = (u_1^{(A)}, u_2^{(B)}), (u_1^{(A)}, u_3^{(B)}), (u_2^{(A)}, u_3^{(B)}), \tag{11.5.5}$$

$$, (u_2^{(A)}, u_1^{(B)}), (u_3^{(A)}, u_1^{(B)}), (u_3^{(A)}, u_2^{(B)}). \tag{11.5.6}$$

Among these 6, we can further restrict our attention to the probability that the outcomes of measurements along different direction yield opposite signs. Take the case 2 in the Table, here we see that out of the six possible measurement directions with different axes, only 2 out of six

$$[h_1^{(A)}, h_2^{(B)}] = [+, -] \tag{11.5.7}$$

$$[h_2^{(A)}, h_1^{(B)}] = [-, +] \tag{11.5.8}$$

also yield opposite outcomes ($h_i^{(A)} = -h_j^{(B)}$), thus there is a $\frac{1}{3}$ probability that the outcomes are anti-parallel. Cases 1 and 8 instead yield 6 out 6 measurements with opposite directions. It is easy to check that for all the cases listed above it holds that

$$P_{\text{opposite}} \geq \frac{1}{3}, \tag{11.5.9}$$

thus the probability of getting opposite outcomes when measuring the spins along two distinct directions $u^{(A)} \neq u^{(B)}$ is at least 1/3 for the classical model. This inequality is one simplified version of the inequalities deduced by Bell in his seminal paper of 1964.

11.5.2 Predictions of quantum theory

We have seen that a classical, purely local theory predicts the inequality (11.5.9). What are the results that we expect from the quantum theory for a similar experiment where we measure the components $\hat{u}_1, \hat{u}_2, \hat{u}_3$ on each of the two particles?

To simplify, let us take the case in which $\hat{u}_2 = \hat{y}$,$\hat{u}_3 = \hat{z}$ and the first component is in the $\hat{x}z$ plane, rotated by an angle θ with respect to the z axis:

$$(\hat{u}_1, \hat{u}_2, \hat{u}_3) = (\cos\theta\hat{z} + \sin\theta\hat{x}, \hat{y}, \hat{z})\,. \tag{11.5.10}$$

The spin operators in these three directions then are

$$\hat{S}_1 = \frac{\hbar}{2}(\cos\theta\hat{\sigma}_z + \sin\theta\hat{\sigma}_x) \tag{11.5.11}$$

$$\hat{S}_2 = \frac{\hbar}{2}\hat{\sigma}_y \tag{11.5.12}$$

$$\hat{S}_3 = \frac{\hbar}{2}\hat{\sigma}_z. \tag{11.5.13}$$

Let us imagine that we measure $\hat{S}_3^{(A)} \otimes \hat{I}^{(B)}$ and then $\hat{I}^{(A)} \otimes \hat{S}_1^{(B)}$. After the first measurement, with probability 1/2 the state $|\psi\rangle$ collapses into

$$|\psi'_-\rangle = |-\rangle_A \otimes |+\rangle_B. \tag{11.5.14}$$

We then compute the expectation value of the second operator on this state,

$$\langle\psi'_-|\hat{I}^{(A)} \otimes \hat{S}_1^{(B)}|\psi'_-\rangle = \frac{\hbar}{2}\cos\theta\,({}_B\langle-|\hat{\sigma}_z|-\rangle_B) + \sin\theta\,({}_B\langle-|\hat{\sigma}_x|-\rangle_B)\,, \tag{11.5.15}$$

$$= -\frac{\hbar}{2}\cos\theta \tag{11.5.16}$$

$$= \frac{\hbar}{2}P_+^- - \frac{\hbar}{2}P_-^-. \tag{11.5.17}$$

thus $P_+^- - P_-^- = -\cos\theta$, but since $P_+^- + P_-^- = 1$ we have that

$$P_-^- = (1+\cos\theta)/2. \tag{11.5.18}$$

Alternatively, with probability 1/2 the state $|\psi\rangle$ collapses into

$$|\psi'_+\rangle \quad = \quad |+\rangle_A \otimes |-\rangle_B. \tag{11.5.19}$$

We then compute the expectation value of the second operator on this state,

$$\begin{aligned}
\langle\psi'_+|\hat{I}^{(A)} \otimes \hat{S}_1^{(B)}|\psi'_+\rangle &= \frac{\hbar}{2}\cos\theta\,({}_B\langle+|\hat{\sigma}_z|+\rangle_B) + \sin\theta\,({}_B\langle+|\hat{\sigma}_x|+\rangle_B)\,, &(11.5.20)\\
&= \frac{\hbar}{2}\cos\theta &(11.5.21)\\
&= \frac{\hbar}{2}P_+^+ - \frac{\hbar}{2}P_-^+. &(11.5.22)
\end{aligned}$$

thus $P_+ - P_- = \cos\theta$, but since $P_+ + P_- = 1$ we have that

$$P_+^+ = (1+\cos\theta)/2. \tag{11.5.23}$$

thus the probability that the outcomes of the first measurement over A and the second measurement over B yield opposite results is

$$P^{31}_{\text{opposite}}(\theta) = \frac{1}{2}P_+^+ + \frac{1}{2}P_-^- = \frac{1}{2}(1+\cos\theta), \tag{11.5.24}$$

and we clearly see that there are many value of θ that violate Bell's inequality, i.e. such that

$$P^{31}_{\text{opposite}}(\theta) \quad < \quad \frac{1}{3}, \tag{11.5.25}$$

for example for $P^{31}_{\text{opposite}}(\frac{2}{3}\pi) = \frac{1}{4} < \frac{1}{3}$. Then, if we perform a measurement at such an angle (say at $2\pi/3$) and we observe a probability that is smaller than 1/3, we can conclude that the classical theory of hidden variables is not able to explain the outcome. Quite amazingly, this kind of experiments have been performed several times now, starting from the 1980s works of Alain Aspect and coworkers. In these experiments, typically done with photons, it has been shown that Nature violates Bell's inequalities and that therefore we cannot explain the measurements just in terms of local classical variables! This is quite good news, since it means that the time you have invested during this semester in learning quantum theory is well invested, after all.

11.6 References and Further Reading

The discussion in this Chapter presents the mathematical and physical structure of quantum systems with many components. Despite its fundamental importance for modern applications of quantum mechanics, as well as for the fundamental meaning of the quantum theory, this is a topic that is still relatively "young" and it is not well discussed in traditional textbooks. The interested reader can however look at Cohen-Tannoudji's Chapter 2, section F on the general aspects of tensor products. Instead, complement D_{III} of Chapter 3 contains the discussion on partial measurements. Our discussion on hidden variables and Bell's inequalities is instead adapted (and simplified) from what found in Sakurai (3.10). The reader can find more details and a slightly more general derivation in there.

Chapter 12

Addition of Angular Momenta

In this Chapter we study the problem of adding angular momenta. The way quantum angular momenta are added is more complex than what you are used to in the classical world, and also leads to several counterintuitive phenomena.

12.1 Generalities

Given two particles with given angular momenta operators, say $\hat{\mathbf{J}}_{(1)}$ and $\hat{\mathbf{J}}_{(2)}$, we would like to study the total angular momentum that these two particles have. The resulting angular momentum is the sum of the individual momenta, however we have to be careful when performing the sum and recall that the two operators act on distinct Hilbert spaces, $\mathcal{H}_{(1)}$ and $\mathcal{H}_{(2)}$. As discussed in the previous Chapter, the correct way of summing the two operators is then to first "upgrade" them to act on the same Hilbert space $\mathcal{H} = \mathcal{H}_{(1)} \otimes \mathcal{H}_{(2)}$, and then consider the sum. The total angular momentum operator is then to be defined as

$$\begin{aligned} \hat{\mathbf{J}} &= \hat{\mathbf{J}}_{(1)} \otimes \hat{I}_{(2)} + \hat{I}_{(1)} \otimes \hat{\mathbf{J}}_{(2)} && (12.1.1)\\ &= (\hat{J}_{(1)x} \otimes \hat{I}_{(2)} + \hat{I}_{(1)} \otimes \hat{J}_{(2)x}, \dots, \dots) && (12.1.2)\\ &= (\hat{J}_x, \hat{J}_y, \hat{J}_z). && (12.1.3) \end{aligned}$$

Moreover, the two angular momentum operators $\hat{\mathbf{J}}_{(1)}$ and $\hat{\mathbf{J}}_{(2)}$ act on different spaces thus their components commute:

$$[\hat{J}_{(1)\alpha} \otimes \hat{I}_{(2)}, \hat{I}_{(1)} \otimes \hat{J}_{(2)\beta}] = 0. \quad (12.1.4)$$

This is easily shown using the tensor product notation:

$$\begin{aligned} \left(\hat{I}_{(1)} \otimes \hat{J}_{(2)\beta}\right)\left(\hat{J}_{(1)\alpha} \otimes \hat{I}_{(2)}\right)|\Psi\rangle_1 \otimes |\Psi\rangle_2 &= \left(\hat{I}_{(1)} \otimes \hat{J}_{(2)\beta}\right)\left(\hat{J}_{(1)\alpha}|\Psi\rangle_1\right) \otimes |\Psi\rangle_2\\ &= \left(\hat{J}_{(1)\alpha}|\Psi\rangle_1\right) \otimes \left(\hat{J}_{(2)\beta}|\Psi\rangle_2\right) && (12.1.5)\\ \left(\hat{J}_{(1)\alpha} \otimes \hat{I}_{(2)}\right)\left(\hat{I}_{(1)} \otimes \hat{J}_{(2)\beta}\right)|\Psi\rangle_1 \otimes |\Psi\rangle_2 &= \left(\hat{J}_{(1)\alpha} \otimes \hat{I}_{(2)}\right)|\Psi\rangle_1 \otimes \left(\hat{J}_{(2)\beta}|\Psi\rangle_2\right)\\ &= \left(\hat{J}_{(1)\alpha}|\Psi\rangle_1\right) \otimes \left(\hat{J}_{(2)\beta}|\Psi\rangle_2\right). && (12.1.6) \end{aligned}$$

Since it is quite cumbersome to carry around the tensor product symbols, in the following we will use a slightly *wrong* but widely adopted notation, in which we write the total angular momentum operator as

$$\hat{\mathbf{J}} = \hat{\mathbf{J}}_{(1)} + \hat{\mathbf{J}}_{(2)}. \quad (12.1.7)$$

This notation is compact but possibly also dangerous, because you might tempted to assume (wrongly) that $\hat{\mathbf{J}}_{(1)}$ and $\hat{\mathbf{J}}_{(2)}$ act on the same Hilbert space, however we have stressed many times now that this is not the case. So, just be careful when using this notation, and if

in doubt, just go back to the tensor product one! In compact notation, the commutation relations for the components of the angular momenta read

$$[\hat{J}_{(1)\alpha}, \hat{J}_{(2)\beta}] = 0. \tag{12.1.8}$$

As a consequence of this relationship, we can also easily verify that the total angular momentum operator is still a valid angular momentum operator, in the sense that it satisfies the usual commutation relations. We can check this explicitly:

$$\begin{aligned}
[\hat{J}_\alpha, \hat{J}_\beta] &= [\hat{J}_{(1)\alpha} + \hat{J}_{(2)\alpha}, \hat{J}_{(1)\beta} + \hat{J}_{(2)\beta}] && (12.1.9)\\
&= [\hat{J}_{(1)\alpha}, \hat{J}_{(1)\beta}] + [\hat{J}_{(2)\alpha}, \hat{J}_{(2)\beta}] && (12.1.10)\\
&= i\hbar\epsilon_{\alpha,\beta,\gamma}\hat{J}_{(1)\gamma} + i\hbar\epsilon_{\alpha,\beta,\gamma}\hat{J}_{(2)\gamma} && (12.1.11)\\
&= i\hbar\epsilon_{\alpha,\beta,\gamma}\hat{J}_\gamma. && (12.1.12)
\end{aligned}$$

Physically speaking, the total angular momentum operator then must be also associated to a rotation operator

$$\hat{D}(\boldsymbol{\theta}) = e^{-\frac{i}{\hbar}\hat{\mathbf{J}}\cdot\boldsymbol{\theta}}. \tag{12.1.13}$$

The meaning of this rotation operator is clarified considering the product of the two rotation operators acting on each of the two subsystems, namely

$$\begin{aligned}
\hat{D}_{(1)}(\boldsymbol{\theta})\hat{D}_{(2)}(\boldsymbol{\theta}) &= e^{-\frac{i}{\hbar}\hat{\mathbf{J}}_{(1)}\cdot\boldsymbol{\theta}}e^{-\frac{i}{\hbar}\hat{\mathbf{J}}_{(2)}\cdot\boldsymbol{\theta}} && (12.1.14)\\
&= e^{-\frac{i}{\hbar}\hat{\mathbf{J}}\cdot\boldsymbol{\theta}}, && (12.1.15)
\end{aligned}$$

where in the last line we have used the fact that $[\hat{J}_{(1)\alpha}, \hat{J}_{(2)\beta}] = 0$, thus the product of the two exponentials can be absorbed into a single exponential of the sum. From this expression we also deduce that the rotation operator associated to the total angular momentum corresponds to taking rotations of the coordinate systems of both particles at the same time,

$$\hat{D}(\boldsymbol{\theta}) = \hat{D}_{(1)}(\boldsymbol{\theta}) \otimes \hat{D}_{(2)}(\boldsymbol{\theta}). \tag{12.1.16}$$

Moreover, since $\hat{\mathbf{J}}$ is just another angular momentum operator, it will also have a set of eigenvalues and eigenvectors of the "standard" form:

$$\begin{aligned}
\hat{\mathbf{J}}^2|j,m\rangle &= \hbar^2 j(j+1)|j,m\rangle && (12.1.17)\\
\hat{J}_z|j,m\rangle &= \hbar m|j,m\rangle. && (12.1.18)
\end{aligned}$$

This expression however does not tell the whole story, since there are other quantities that commute with $\hat{\mathbf{J}}^2$ and $\hat{J}_z$. We can verify for example that the total momentum squared commutes with the individual total momenta squared. To prove this, we start writing

$$\begin{aligned}
\hat{\mathbf{J}}^2 &= \left(\hat{J}_{(1)x} + \hat{J}_{(2)x}\right)^2 + \left(\hat{J}_{(1)y} + \hat{J}_{(2)y}\right)^2 + \left(\hat{J}_{(1)z} + \hat{J}_{(2)z}\right)^2 && (12.1.19)\\
&= \hat{\mathbf{J}}^2_{(1)} + \hat{\mathbf{J}}^2_{(2)} + 2\hat{J}_{(1)z}\hat{J}_{(2)z} + 2\hat{J}_{(1)x}\hat{J}_{(2)x} + 2\hat{J}_{(1)y}\hat{J}_{(2)y} && (12.1.20)\\
&= \hat{\mathbf{J}}^2_{(1)} + \hat{\mathbf{J}}^2_{(2)} + 2\hat{J}_{(1)z}\hat{J}_{(2)z} + \frac{1}{2}\left(\hat{J}^+_{(1)} + \hat{J}^-_{(1)}\right)\left(\hat{J}^+_{(2)} + \hat{J}^-_{(2)}\right) + && (12.1.21)\\
&\quad -\frac{1}{2}\left(\hat{J}^+_{(1)} - \hat{J}^-_{(1)}\right)\left(\hat{J}^+_{(2)} - \hat{J}^-_{(2)}\right) && (12.1.22)\\
&= \hat{\mathbf{J}}^2_{(1)} + \hat{\mathbf{J}}^2_{(2)} + 2\hat{J}_{(1)z}\hat{J}_{(2)z} + \hat{J}^+_{(1)}\hat{J}^-_{(2)} + \hat{J}^-_{(1)}\hat{J}^+_{(2)}, && (12.1.23)
\end{aligned}$$

thus since $[\hat{\mathbf{J}}^2_{(1)}, \hat{J}^\pm_{(1)}] = [\hat{\mathbf{J}}^2_{(1)}, \hat{J}^z_{(1)}] = 0$, and similarly for the particle 2, we have

$$\begin{aligned}
[\hat{\mathbf{J}}^2, \hat{\mathbf{J}}^2_{(1)}] &= 0, && (12.1.24)\\
[\hat{\mathbf{J}}^2, \hat{\mathbf{J}}^2_{(2)}] &= 0. && (12.1.25)
\end{aligned}$$

Moreover, two individual squared momenta also commute with the total $\hat{J}^z$, since

$$\begin{aligned}
[\hat{J}_z, \hat{\mathbf{J}}^2_{(1)}] &= [\hat{J}_{(1)z} + \hat{J}_{(2)z}, \hat{\mathbf{J}}^2_{(1)}] && (12.1.26)\\
&= [\hat{J}_{(1)z}, \hat{\mathbf{J}}^2_{(1)}] && (12.1.27)\\
&= 0, && (12.1.28)\\
[\hat{J}_z, \hat{\mathbf{J}}^2_{(2)}] &= [\hat{J}_{(1)z} + \hat{J}_{(2)z}, \hat{\mathbf{J}}^2_{(2)}] && (12.1.29)\\
&= [\hat{J}_{(2)z}, \hat{\mathbf{J}}^2_{(2)}] && (12.1.30)\\
&= 0. && (12.1.31)
\end{aligned}$$

This means that we have four mutually commuting quantities, $\hat{\mathbf{J}}^2$,$\hat{\mathbf{J}}^2_{(1)}$,$\hat{\mathbf{J}}^2_{(2)}$ and $\hat{J}_z$ whose eigenvalues can be used to index the eigen-kets of the total momentum such that

$$\begin{aligned}
\hat{\mathbf{J}}^2|j_1, j_2; j, m\rangle &= \hbar^2 j(j+1)|j_1, j_2; j, m\rangle, && (12.1.32)\\
\hat{J}_z|j_1, j_2; j, m\rangle &= \hbar m|j_1, j_2; j, m\rangle, && (12.1.33)\\
\hat{\mathbf{J}}^2_{(1)}|j_1, j_2; j, m\rangle &= \hbar^2 j_1(j_1+1)|j_1, j_2; j, m\rangle, && (12.1.34)\\
\hat{\mathbf{J}}^2_{(2)}|j_1, j_2; j, m\rangle &= \hbar^2 j_2(j_2+1)|j_1, j_2; j, m\rangle. && (12.1.35)
\end{aligned}$$

12.1.1 Tensor-Product Basis

While the representation we have introduced above is the "standard" representation of the composed momenta, it is not very convenient to work with. It is in fact more natural to introduce basis states that are simultaneous eigen-kets of the individual components, and that we have already analyzed in the previous Chapters. We therefore consider the tensor-product basis states

$$|j_1, j_2; m_1, m_2\rangle = |j_1, m_1\rangle \otimes |j_2, m_2\rangle, \quad (12.1.36)$$

that are simultaneous eigen-kets of $\hat{\mathbf{J}}^2_{(1)}$,$\hat{\mathbf{J}}^2_{(2)}$,$\hat{J}_{(1)z}$,$\hat{J}_{(2)z}$. These four operators are obviously mutually commuting, since operators with different particle indexes act on different Hilbert spaces (thus commute) and same-particle operators commute as well, i.e. we already know that $[\hat{\mathbf{J}}^2_{(1)}, \hat{J}_{(1)z}] = [\hat{\mathbf{J}}^2_{(2)}, \hat{J}_{(2)z}] = 0$. The basis eigen-kets we consider then satisfy

$$\begin{aligned}
\hat{\mathbf{J}}^2_{(1)}|j_1, j_2; m_1, m_2\rangle &= \hbar^2 j_1(j_1+1)|j_1, j_2; m_1, m_2\rangle, && (12.1.37)\\
\hat{\mathbf{J}}^2_{(2)}|j_1, j_2; m_1, m_2\rangle &= \hbar^2 j_2(j_2+1)|j_1, j_2; m_1, m_2\rangle, && (12.1.38)\\
\hat{J}_{(1)z}|j_1, j_2; m_1, m_2\rangle &= \hbar m_1|j_1, j_2; m_1, m_2\rangle, && (12.1.39)\\
\hat{J}_{(2)z}|j_1, j_2; m_1, m_2\rangle &= \hbar m_2|j_1, j_2; m_1, m_2\rangle. && (12.1.40)
\end{aligned}$$

While this basis is convenient, the tensor-product states however are not eigenstates of the total momentum squared. This is because we cannot diagonalize at the same time the four operators above ($\hat{\mathbf{J}}^2_{(1)}$,$\hat{\mathbf{J}}^2_{(2)}$,$\hat{J}_{(1)z}$,$\hat{J}_{(2)z}$) and also $\hat{\mathbf{J}}^2$. This can be checked noticing that for example $\hat{\mathbf{J}}^2$ does not commute with the single-particle $\hat{J}_z$ operators:

$$\begin{aligned}
[\hat{J}_{(1)z}, \hat{\mathbf{J}}^2] &= [\hat{J}_{(1)z}, \hat{\mathbf{J}}^2_{(1)} + \hat{\mathbf{J}}^2_{(2)} + 2\hat{J}_{(1)z}\hat{J}_{(2)z} + \hat{J}^+_{(1)}\hat{J}^-_{(2)} + \hat{J}^-_{(1)}\hat{J}^+_{(2)}] && (12.1.41)\\
&= [\hat{J}_{(1)z}, \hat{J}^+_{(1)}\hat{J}^-_{(2)} + \hat{J}^-_{(1)}\hat{J}^+_{(2)}] && (12.1.42)\\
&= \hat{J}^-_{(2)}[\hat{J}_{(1)z}, \hat{J}^+_{(1)}] + [\hat{J}_{(1)z}, \hat{J}^-_{(1)}]\hat{J}^+_{(2)} && (12.1.43)\\
&= \hbar\hat{J}^-_{(2)}\hat{J}^+_{(1)} - \hbar\hat{J}^-_{(1)}\hat{J}^+_{(2)} && (12.1.44)\\
&\neq 0. && (12.1.45)
\end{aligned}$$

Nonetheless, we can still use this convenient basis to express the eigen-kets of the total angular momentum squared, i.e. we can develop the eigen-kets as

$$|j_1, j_2; j, m\rangle = \sum_{m_1 m_2} |j_1, j_2; m_1, m_2\rangle\langle j_1, j_2; m_1, m_2|j_1, j_2; j, m\rangle. \quad (12.1.46)$$

The coefficients of this transformation,

$$C^{jm}_{j_1 m_1 j_2 m_2} \equiv \langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m \rangle, \tag{12.1.47}$$

are called Clebsch–Gordan coefficients an play the vital role of connecting the two representations.

12.1.2 Properties of Clebsch–Gordan coefficients

There is a number of important properties of these coefficients that we can already deduce at this stage. First of all, the coefficients vanish unless

$$\boxed{m = m_1 + m_2}. \tag{12.1.48}$$

This can be proven noticing that

$$\left(\hat{J}_z - \hat{J}_{(1)z} - \hat{J}_{(2)z}\right) |j_1, j_2; j, m\rangle = 0, \tag{12.1.49}$$

thus multiplying this equation by $\langle j_1, j_2; m_1, m_2|$ we have

$$(m - m_1 - m_2)\langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m\rangle = 0, \tag{12.1.50}$$

implying that the Clebsch–Gordan coefficients (appearing in the left hand side of this Equation) must vanish unless $m = m_1 + m_2$. This condition is quite natural, since it tells us that the total $\hat{J}_z$ has an eigenvalue which is the sum of the two individual eigenvalues of $\hat{J}_{(1)z}$ and $\hat{J}_{(2)z}$.

The other important condition is on the possible values that j can take, as it turns out that

$$\boxed{|j_1 - j_2| \leq j \leq j_1 + j_2}. \tag{12.1.51}$$

In order to see why this is the case, remember that

$$-j \leq m \leq j \tag{12.1.52}$$
$$-j_1 \leq m_1 \leq j_1 \tag{12.1.53}$$
$$-j_2 \leq m_2 \leq j_2. \tag{12.1.54}$$

Now, if we set $m = j$, and $j_1 = m_1$, the inequality for m_2 (which is $m_2 = m - m_1$, as we have seen before) becomes:

$$-j_2 \leq j - j_1 \leq j_2 \tag{12.1.55}$$
$$j_1 - j_2 \leq j \leq j_1 + j_2. \tag{12.1.56}$$

Also, if we take $m = j$ and $j_2 = m_2$, the inequality for m_1 ($m_1 = m - m_2$) becomes

$$-j_1 \leq j - j_2 \leq j_1 \tag{12.1.57}$$
$$j_2 - j_1 \leq j \leq j_1 + j_2, \tag{12.1.58}$$

thus we conclude with Eq. (12.1.51).

12.1.2.1 State counting

An alternative way of convincing ourselves that Eq. (12.1.51) must be true is by counting the bases states in the two representations. In other words, we can count how many basis states $|j_1, j_2; m_1, m_2\rangle = |j_1 m_1\rangle \otimes |j_2 m_2\rangle$ exist and how many states $|j_1, j_2; j, m\rangle$ exist. In the first case, we know that the total number of states is given by the product of the number of states spanned by the individual kets that are taken into the tensor product, thus

$$N_I = (2j_1 + 1)(2j_2 + 1). \tag{12.1.59}$$

In the other case, we know that for each j there are $2j+1$ states and if the inequality, Eq. (12.1.51), is satisfied we have that j must run between $(j_1 - j_2)$ and $(j_1 + j_2)$, assuming, as we can always do, that we pick $j_1 \geq j_2$. This means that the total number of states in the second count is (the full summation of the series is left as an exercise):

$$N_{II} = \sum_{j=j_1-j_2}^{j_1+j_2} (2j+1) \tag{12.1.60}$$

$$= (2j_1+1)(2j_2+1), \tag{12.1.61}$$

thus we find the consistent result $N_I = N_{II}$.

12.1.3 Explicit form of the Clebsch–Gordan coefficients

Determining explicit and general expressions for Clebsch–Gordan is a time consuming and not very productive exercise that is still reason of nightmares for generations of students who were forced to derive them. We just quote here the final result, so that you can understand the reason of such nightmares:

$$\begin{aligned} C^{jm}_{j_1 m_1 j_2 m_2} = {} & \delta_{m_1+m_2,m} \sqrt{\frac{2j+1}{(j_1+j_2-j)!(j_1-j_2+j)!(j_2-j_1+j)!(j_1+j_2+j+1)!}} \times \\ & \times \sqrt{(j+m)!(j-m)!(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!} \\ & \times \sum_k (-1)^k \binom{j_1+j_2-j}{k} \binom{j_1-j_2+j}{j_1-m_1-k} \binom{j_2-j_1+j}{j_2+m_2-k}. \end{aligned} \tag{12.1.62}$$

We also recall here a few more properties of the Clebsch-Gordan coefficients. They are real-valued by convention, and they satisfy the closure conditions

$$\sum_{jm} C^{jm}_{j_1 m_1 j_2 m_2} C^{jm}_{j_1 m_1' j_2 m_2'} = \delta_{m_1,m_1'} \delta_{m_2,m_2'} \tag{12.1.63}$$

$$\sum_{m_1 m_2} C^{jm}_{j_1 m_1 j_2 m_2} C^{j'm'}_{j_1 m_1 j_2 m_2} = \delta_{j,j'} \delta_{m,m'}. \tag{12.1.64}$$

12.2 Two Spins

We first consider a simple example of the formalism developed so far, where we can easily find an explicit representation for $|j_1, j_2; j, m\rangle$ bypassing the explicit calculation of the Clebsch–Gordan coefficients. The example consists in forming the total angular momentum resulting from two spins 1/2. Formally, we form the vector operator

$$\hat{\mathbf{S}} = \hat{\mathbf{S}}_{(1)} + \hat{\mathbf{S}}_{(2)} \tag{12.2.1}$$

$$= (\hat{S}_x, \hat{S}_y, \hat{S}_z). \tag{12.2.2}$$

The "convenient" basis in this case is then simply

$$|s_1, s_2; m_1, m_2\rangle = |s_1, m_1\rangle \otimes |s_2, m_2\rangle, \tag{12.2.3}$$

$$|m_1, m_2\rangle = |\pm, \pm\rangle, \tag{12.2.4}$$

where in the last line we omitted the $s_1 = s_2 = 1/2$ quantum numbers and just concentrated on the two possible values of $m_1 = \pm 1/2$ and $m_2 = \pm 1/2$. In total then we have four states

$$|1\rangle = |++\rangle \tag{12.2.5}$$

$$|2\rangle = |+-\rangle \tag{12.2.6}$$

$$|3\rangle = |-+\rangle \tag{12.2.7}$$

$$|4\rangle = |--\rangle. \tag{12.2.8}$$

In order to find the states $|j_1, j_2; j, m\rangle$ in the "standard" representation, we start by explicitly computing the total spin squared

$$\begin{aligned}\hat{\mathbf{S}}^2 &= \hat{\mathbf{S}}^2_{(1)} + \hat{\mathbf{S}}^2_{(2)} + 2\hat{S}_{(1)z}\hat{S}_{(2)z} + \hat{S}^+_{(1)}\hat{S}^-_{(2)} + \hat{S}^-_{(1)}\hat{S}^+_{(2)} && (12.2.9)\\ &= \frac{3}{4}\hbar^2\left(\hat{I}_{(1)} + \hat{I}_{(2)}\right) + 2\hat{S}_{(1)z}\hat{S}_{(2)z} + \hat{S}^+_{(1)}\hat{S}^-_{(2)} + \hat{S}^-_{(1)}\hat{S}^+_{(2)}. && (12.2.10)\end{aligned}$$

From this expression we see that the last two terms are vanishing when applied to the states $|++\rangle$ and $|--\rangle$ since, for example,

$$\begin{aligned}\hat{S}^+_{(1)}|++\rangle &= \hat{S}^+_{(1)} \otimes \hat{I}_{(2)}\left(|+\rangle_1 \otimes |+\rangle_2\right) && (12.2.11)\\ &= \left(\hat{S}^+_{(1)}|+\rangle_1\right) \otimes |+\rangle_2 && (12.2.12)\\ &= 0. && (12.2.13)\end{aligned}$$

Furthermore, we can easily verify that these two states are eigenstates of $\hat{\mathbf{S}}^2$, since

$$\begin{aligned}\hat{\mathbf{S}}^2|++\rangle &= \left(\frac{3}{4}\hbar^2 2 + 2\hbar^2\frac{1}{2}\frac{1}{2}\right)|++\rangle && (12.2.14)\\ &= 2\hbar^2|++\rangle, && (12.2.15)\\ \hat{\mathbf{S}}^2|--\rangle &= \left(\frac{3}{4}\hbar^2 2 + 2\hbar^2\left(-\frac{1}{2}\right)\left(-\frac{1}{2}\right)\right)|--\rangle && (12.2.16)\\ &= 2\hbar^2|--\rangle, && (12.2.17)\end{aligned}$$

with eigenvalue $2\hbar^2$. However we know from the general theory that the eigenvalues of $\hat{\mathbf{S}}^2$ are also equal to $\hbar^2 s(s+1)$, thus we conclude that these two states have $s = 1$. We have then successfully found the first two states we were looking for in the standard representation:

$$\begin{aligned}|j=1; m=1\rangle &= |++\rangle && (12.2.18)\\ |j=1; m=-1\rangle &= |--\rangle. && (12.2.19)\end{aligned}$$

From the general theory of angular momentum, however we know that the $j = 1$ states always come as a triplet of states ($m = -1, 0, 1$), thus there must be still a missing state we haven't found yet with $j = 1, m = 0$. In order to find it, we apply the lowering operator

$$\begin{aligned}\hat{S}^- &= \hat{S}_x - i\hat{S}_y && (12.2.20)\\ &= \hat{S}_{(1)x} + \hat{S}_{(2)x} - i\hat{S}_{(1)y} - i\hat{S}_{(2)y} && (12.2.21)\\ &= \hat{S}^-_{(1)} + \hat{S}^-_{(2)} && (12.2.22)\end{aligned}$$

to the state with highest m:

$$\begin{aligned}\hat{S}^-|j=1; m=1\rangle &= \hat{S}^-_{(1)}|++\rangle + \hat{S}^-_{(2)}|++\rangle && (12.2.23)\\ \hbar\sqrt{j(j+1)-m(m-1)}|j=1; m=0\rangle &= \hbar\sqrt{s_1(s_1+1)-m_1(m_1-1)}|-+\rangle + \\ &\quad +\hbar\sqrt{s_2(s_2+1)-m_2(m_2-1)}|+-\rangle && (12.2.24)\\ \sqrt{2}|j=1; m=0\rangle &= |-+\rangle + |+-\rangle. && (12.2.25)\end{aligned}$$

From the last line then we can read out the third state with $j = 1, m = 0$ we were missing before:

$$|j=1; m=0\rangle = \frac{1}{\sqrt{2}}\left(|-+\rangle + |+-\rangle\right).$$

To find the last and final state (remember that we started with four states for the "convenient" basis, so we need to find also 4 states in the standard basis) we realize that the missing state must be the one with $|j = 0; m = 0\rangle$ (that is the only allowed value of j

remaining, from the inequality condition (12.1.51)). This state is found imposing that it is orthogonal to all the other states we have already found. We start by imposing that it is orthogonal to the other state we found for $m = 0$:

$$\langle j = 0, m = 0 | j = 1; m = 0 \rangle \quad = \quad 0, \tag{12.2.26}$$

and we obtain:

$$|j = 0; m = 0\rangle \quad = \quad \frac{1}{\sqrt{2}} \left(|-+\rangle - |+-\rangle \right). \tag{12.2.27}$$

It can be easily checked that this state is also orthogonal to all the other states previously found, just because they carry different m thus all products such $\langle -- | -+ \rangle$ etc. give zero. To summarize, we have found the four states we were looking for in the "standard" representation:

$$|j = 1; m = 1\rangle \quad = \quad |++\rangle \tag{12.2.28}$$

$$|j = 1; m = 0\rangle \quad = \quad \frac{1}{\sqrt{2}} \left(|-+\rangle + |+-\rangle \right), \tag{12.2.29}$$

$$|j = 1; m = -1\rangle \quad = \quad |--\rangle, \tag{12.2.30}$$

$$|j = 0; m = 0\rangle \quad = \quad \frac{1}{\sqrt{2}} \left(|-+\rangle - |+-\rangle \right), \tag{12.2.31}$$

with the first three having $j = 1$ (also known as "triplet" of states) and the last one with $j = 0$, also known as "singlet".

12.3 Adding Spin and Orbital Momentum

Another example we propose here is the important case of adding spin and orbital angular momentum degrees of freedom. For example, we can form the total angular momentum of a particle with spin:

$$\hat{\mathbf{J}} \quad = \quad \hat{\mathbf{L}} + \hat{\mathbf{S}} \tag{12.3.1}$$

$$\quad = \quad \hat{\mathbf{L}} \otimes \hat{I}_S + \hat{I}_L \otimes \hat{\mathbf{S}}, \tag{12.3.2}$$

where the second line once more emphasizes that the two operators act on different Hilbert spaces. Notice that actually in this case the first operator (the orbital angular momentum) acts on an infinite Hilbert space, whereas the spin operator acts on a finite vector space. As before, we can interpret the resulting rotation operator as just the product of two independent rotations on the respective degrees of freedom:

$$\hat{D}(\boldsymbol{\theta}) \quad = \quad \exp\left(-\frac{i}{\hbar} \hat{\mathbf{L}} \cdot \boldsymbol{\theta} \right) \otimes \exp\left(-\frac{i}{\hbar} \hat{\mathbf{S}} \cdot \boldsymbol{\theta} \right). \tag{12.3.3}$$

A typical way of writing the wave function of a particle with spin (say, an electron) is by means of the tensor-product basis, $|\mathbf{r}\rangle \otimes |s, m\rangle$ such that the state vector is

$$\left(\langle \mathbf{r}| \otimes \langle s, m| \right) |\Psi\rangle \quad = \quad \Psi(\mathbf{r}, m), \tag{12.3.4}$$

where the first variable $\mathbf{r} = (x, y, z)$ is clearly continuous, whereas the second one $m = \pm 1/2$ is discrete. An alternative way to write the state is as a vector of two continuous-space wave-functions:

$$\begin{pmatrix} \Psi_+(\mathbf{r}) \\ \Psi_-(\mathbf{r}) \end{pmatrix}, \tag{12.3.5}$$

in such a way that $\Psi_\pm(\mathbf{r}) = \Psi(\mathbf{r}, \pm 1/2)$. From these notations, the meaning of $\Psi_\pm(\mathbf{r})$ is quite clear, for example $|\Psi_+(\mathbf{r})|^2$ would give the probability density of finding a spin "up" at position $\mathbf{r}$. This representation is also called "spin-orbital".

12.3.1 Standard basis

The spin-orbital representation is typically enough for most applications, however we might be interested in finding also the standard representation involving the eigenvalues of $\hat{\mathbf{J}}^2$, $\hat{J}_z$, $\hat{\mathbf{S}}^2$, $\hat{\mathbf{L}}^2$. Similarly to what done for the general case, we will call this basis

$$|l, s; j, m\rangle, \tag{12.3.6}$$

whereas the "convenient" tensor-product basis is given by

$$|l, m_l, s, m_s\rangle = |l, m_l\rangle \otimes |s, m_s\rangle. \tag{12.3.7}$$

In the following we will omit the explicit value of $s = 1/2$ and l from all kets, since they are fixed everywhere. From the general inequality, Eq. (12.1.51), we know that there are only two allowed values for j, namely $j_{\max} = l + 1/2$ and $j_{\min} = l - \frac{1}{2}$.

As also seen for the case of two spins, the state with $j_{\max} = m_{\max} = l + 1/2$ is always easily written in terms of tensor product basis:

$$|j = l + 1/2, m = l + 1/2\rangle = |l, l\rangle \otimes |+\rangle. \tag{12.3.8}$$

To obtain the remaining states in this "multiplet" of states with $j = l+1/2$, i.e. all the other states with $m = (l+1/2, l-1/2, l-3/2, \cdots - l+1/2, -l-1/2)$ we apply total spin lowering operator $\hat{J}^- = \hat{L}^- + \hat{S}^-$ and recall the identity $\hat{J}^-|j, m\rangle = \hbar\sqrt{j(j+1) - m(m-1)}|j, m-1\rangle$ giving

$$\begin{aligned}
\hat{J}^-|j = l + 1/2, m = l + 1/2\rangle &= \hbar\sqrt{2l+1}|j = l + 1/2, m = l - 1/2\rangle & (12.3.9)\\
&= \hat{L}^-|l, l\rangle \otimes |+\rangle + |l, l\rangle \otimes \hat{S}^-|+\rangle & (12.3.10)\\
&= \hbar\left(\sqrt{2l}|l, l-1\rangle \otimes |+\rangle + |l, l\rangle \otimes |-\rangle\right), & (12.3.11)
\end{aligned}$$

thus

$$\begin{aligned}
|j = l + 1/2, m = l - 1/2\rangle &= \sqrt{\frac{2l}{2l+1}}|l, l-1\rangle \otimes |+\rangle + \\
&\quad + \sqrt{\frac{1}{2l+1}}|l, l\rangle \otimes |-\rangle. \qquad (12.3.12)
\end{aligned}$$

To generate all the other states we apply again and again $\hat{J}^-$, finding the general relation (left as an exercise)

$$\begin{aligned}
|j = l + 1/2, m\rangle &= \sqrt{\frac{l + m + 1/2}{2l+1}}|l, m - \frac{1}{2}\rangle \otimes |+\rangle + \\
&\quad + \sqrt{\frac{l - m + 1/2}{2l+1}}|l, m + \frac{1}{2}\rangle \otimes |-\rangle. \qquad (12.3.13)
\end{aligned}$$

Once determined all the states with $j = l+1/2$, we then go on with states having $j = l-1/2$. The highest m state in this multiplet is, in general, a linear combination of spin states with up and down and the appropriate value of m_l such that $m_l + s_z = m$, thus $m_l = l - 1$ or $m_l = l$:

$$|j = l - 1/2, m = l - 1/2\rangle = c_1|l, l-1\rangle \otimes |+\rangle + c_2|l, l\rangle \otimes |-\rangle, \tag{12.3.14}$$

where the two coefficients are to be determined. However, we have already found a state with $m = l - 1/2$ before, in the multiplet with $j = l + 1/2$. We can then find the two coefficients imposing the orthogonality condition

$$\begin{aligned}
\langle j = l + 1/2, m = l - 1/2|j = l - 1/2, m = l - 1/2\rangle &= 0, & (12.3.15)\\
c_1\sqrt{\frac{2l}{2l+1}} + c_2\sqrt{\frac{1}{2l+1}} &= 0. & (12.3.16)
\end{aligned}$$

A simple solution (also recalling that we must have $|c_1|^2 + |c_2|^2 = 1$) is to take $c_1 = \sqrt{\frac{1}{2l+1}}$ and $c_2 = -\sqrt{\frac{2l}{2l+1}}$, therefore

$$\begin{aligned} |j = l - 1/2, m = l - 1/2\rangle &= \sqrt{\frac{1}{2l+1}}|l, l-1\rangle \otimes |+\rangle + \\ &\quad -\sqrt{\frac{2l}{2l+1}}|l, l\rangle \otimes |-\rangle. \end{aligned} \tag{12.3.17}$$

With this state at hand we can either further apply $\hat{J}^-$ and generate all the states in the multiplet with $j = l - 1/2$, or impose that the other states in the multiplet are orthogonal to those we already generated with the same values of m:

$$\langle j = l + 1/2, m | j = l - 1/2, m\rangle = 0, \tag{12.3.18}$$

finally yielding

$$\begin{aligned} |j = l - 1/2, m\rangle &= \sqrt{\frac{l - m + 1/2}{2l+1}}|l, m - \frac{1}{2}\rangle \otimes |+\rangle + \\ &\quad -\sqrt{\frac{l + m + 1/2}{2l+1}}|l, m + \frac{1}{2}\rangle \otimes |-\rangle. \end{aligned} \tag{12.3.19}$$

12.4 References and Further Reading

The discussion in this Chapter is mainly adapted and simplified from Sakurai, Chapter 3 (Section 3.8). Cohen-Tannoudji's book discusses the addition of angular momentum in Volume 2 (Chapter 10). The complements to that Chapter (especially A_X andB_X) contain a lot of details on the Clebsch-Gordan coefficients and also some additional examples that might be useful to gain further technical knowledge of the general topic.