ARNAU RIOS HUGUET

ADVANCED QUANTUM MECHANICS

LAST UPDATE: SEPTEMBER 13, 2023

Contents

1 Introduction to Quantum Mechanics

This Module is designed to provide you with all the key contents in Quantum Mechanics that you will need for the remainder of your Master's Programme. In some cases, this levelling process may take you through familiar contents. In others, you will be seeing new material. I will try to point out to specific references and books as we go along to provide you with further resources to explore the different topics.

1.1 Preliminaries

Generally speaking, quantum mechanics concerns the physics of systems at a microscopic scale. This does not necessarily mean that the systems are microscopic, as some macroscopic phenomena are influenced directly by quantum processes. There are several examples of such systems, ranging from molecules to atoms (and chemistry) to nuclei and to fundamental particles. There are also instances of cosmology and astrophysics where quantum mechanics plays an important role at a macroscopic scale.

This Module is not so much concerned with specific applications, but rather discusses *methods* and *techniques* that are useful and need to be understood to become an able practitioner of quantum mechanics. As a consequence, the Module focuses on abstract concepts and theoretical ideas and techniques. We won't pay particular attention to the experimental foundation of quantum mechanics or to the history of the discipline - as interesting as these areas are!

1.1.1 Hilbert spaces

The theory of quantum mechanics is formulated in the framework of Hilbert spaces. In short, a Hilbert space $\mathcal{E}_H$ is a vector space[1] that has an additional scalar (or inner) product. More specifically, Hilbert spaces are sets of complex objects, and possess a positive definite Hermitian scalar product[2]. Consider two vectors in this space, which we denote as $|\psi\rangle$ and $|\phi\rangle$. The Hermitian scalar product of $|\psi\rangle$ and $|\phi\rangle$ is written as $\langle\psi|\phi\rangle$. It possesses the important Hermitian

[1] A vector space is a set of elements that is closed under addition and under the multiplication by a scalar.

[2] We further assume that Hilbert spaces are complete and separable, but these mathematical subtleties are not relevant for the discussions below.

symmetry

$$\langle\phi|\psi\rangle = (\langle\psi|\phi\rangle)^* \,, \tag{1.1}$$

where z^* denotes the complex conjugate of the complex variable z. The product $\langle\psi|\phi\rangle$ is linear in ϕ, so that $\langle\psi|a\phi + b\phi'\rangle = a\langle\psi|\phi\rangle + b\langle\psi|\phi'\rangle$; and antilinear in ψ, so that $\langle a\psi + b\psi'|\phi\rangle = a^*\langle\psi|\phi\rangle + b^*\langle\psi'|\phi\rangle$.

The scalar product allows us to define the notion of a *distance* in Hilbert spaces, characterised by the (squared) norm of a vector, $\|\psi\|^2 = \langle\psi|\psi\rangle$. This norm is positive definite, $\|\psi\|^2 \geq 0$. We shall see that the wave functions have a probabilistic nature and, in consequence, their norm is usually chosen to equal unity, $\langle\psi(t)|\psi(t)\rangle = 1$.

The only vector satisfying $\|\psi\|^2 = 0$ is the null vector $\psi = 0$.

A key fundamental property of quantum mechanical wave functions is that they belong to a Hilbert space $\mathcal{E}_H$. For a particle in 3D space, the Hilbert space $\mathcal{E}_H$ is the space of square-integrable functions in three real variables, (x, y, z), called $\mathcal{L}^2(R^3)$. It is important to keep in mind that the description in real space, $\psi(\mathbf{r}, t)$, is not the only possible one. A Fourier transform of these variables to a function $\psi(\mathbf{p}, t)$ provides an equivalent description of the states. In fact, there are many such different *representations* of the same mathematical object, ψ, the vector of the Hilbert space $\mathcal{E}_H$ of interest.

The idea that a *ket* is more general than its representation is similar to the idea that a *vector* is a more abstract construct than the set of coordinates used to represent it.

We shall say that a physical system is described at time t by a *state vector*, which we shall call $|\psi(t)\rangle$. In Dirac notation, these are called *kets*. As we have just said, these objects are more general than their specific (real-space, momentum-space, mixed) representations.

We will encounter two types of Hilbert spaces: discrete (or countable) and continuous (infinite-dimensional) spaces. In a discrete finite-dimensional Hilbert space of dimension N, the vectors can be represented as column matrices,

$$|\psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix} \,, \quad |\phi\rangle = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_N \end{pmatrix} \,, \tag{1.2}$$

and the scalar product is simply

$$\langle\psi|\phi\rangle = \sum_{i=1}^{N} \psi_i^* \phi_i \,. \tag{1.3}$$

In other words, this is the matrix product of the row matrix $\left(\psi_1^*, \psi_2^*, \cdots \psi_N^*\right)$ with the column matrix $|\psi\rangle$. Countable, finite-dimensional systems are important in quantum mechanics. Typically, the dimension of a state N is finite in practical numerical applications. In other cases of interest, one may also have a countable, but infinite set of discrete states, like in the hydrogen atom.

For continuous Hilbert spaces, the discrete formulation above does not work. The vectors in Hilbert spaces are instead square-integrable

functions. In 3D real space, for instance, these vectors are functions of position. We express these projecting the ket state into real space, $\langle \mathbf{r}|\psi\rangle = \psi(\mathbf{r})$, and their scalar products are the integrals

$$\langle \psi|\phi\rangle = \int d^{(3)}\mathbf{r}\ \ \psi^*(\mathbf{r})\,\phi(\mathbf{r})\ . \tag{1.4}$$

Analogously, the momentum-state representation of the ket $|\psi\rangle$ is $\langle \mathbf{p}|\psi\rangle = \psi(\mathbf{p})$.

1.1.2 Operators

A (linear) operator $\hat{O}$ acting in a Hilbert space $\mathcal{E}_H$ transforms any given ket, $|\psi\rangle$, into another ket, $\hat{O}\,|\psi\rangle$, in the same Hilbert space . One often considers the scalar product of $\hat{O}\,|\psi\rangle$ on $|\phi\rangle$, $\langle\phi|\left(\hat{O}\,|\psi\rangle\right)$. These scalar products are computed differently depending on the discrete or continuous nature of the wave functions.

Perhaps the simplest possible operator is the identity operator, denoted $\hat{I}$, which transforms a quantum state into itself, either a ket $\hat{I}|\psi\rangle = |\psi\rangle$ or a bra $\langle\psi|\hat{I} = \langle\psi|$.

In a countable Hilbert space of dimension N, operators $\hat{O}$ transforming the column N-vector wave functions $|\phi\rangle$ are square matrices of dimension $N \times N$. The product $\langle\phi|\left(\hat{O}\,|\psi\rangle\right) = \langle\phi|\hat{O}|\psi\rangle$ is then the result of first multiplying the column vector $|\psi\rangle$ with the matrix O to obtain a second column vector. One can then take the scalar product with the row (and conjugated) vector, $|\phi\rangle$, to obtain a complex scalar quantity. Mathematically, this reads

$$\langle\phi|\hat{O}|\psi\rangle = \sum_{i,j=1}^{N} \phi_i^* O_{ij}\psi_j\,. \tag{1.5}$$

Note that the product with operators is associative, so we have the identity

$$\langle\phi|\hat{O}|\psi\rangle = \left(\langle\phi|\hat{O}\right)|\psi\rangle = \langle\phi|\left(\hat{O}|\psi\rangle\right). \tag{1.6}$$

The product $\langle\phi|\hat{O}|\psi\rangle$ is often called a *matrix element* of $\hat{O}$. When the two wave functions in the previous expression are the same, we say that we are computing the *expectation value* of the operator $\hat{O}$, which we denote with the non-capital letter associated to the operator,

$$\langle o\rangle = \langle\psi|\hat{O}|\psi\rangle\ . \tag{1.7}$$

In continuous spaces, operators cannot be expressed simply in terms of matrices. In some cases, for instance, they involve derivatives of or integrations over wave functions and hence need to be treated with care. An example of a continuous-space operator is the momentum operator, $\mathbf{p} = -i\hbar\nabla$. Consider for instance its effect on the wave function $\psi_{\mathbf{k}}(\mathbf{r}) = A\sin(\mathbf{k}\cdot\mathbf{r})$:

$$\mathbf{p}\psi_{\mathbf{k}}(\mathbf{r}) = -i\hbar\nabla A\sin(\mathbf{k}\cdot\mathbf{r}) = -i\hbar\mathbf{k}A\cos(\mathbf{k}\cdot\mathbf{r})\ .$$

The state $\mathbf{p}\psi_{\mathbf{k}}(\mathbf{r})$ is clearly very different from the initial state wave function $\psi_{\mathbf{k}}(\mathbf{r})$.

1.1.3 Hermitian operators

Consider an operator $\hat{O}$ acting in $\mathcal{E}_H$. The *adjoint* of $\hat{O}$ or $\hat{O}^\dagger$ is defined by the relation

$$\left\langle\psi\middle|\hat{O}^\dagger\middle|\phi\right\rangle = \left(\langle\phi|\hat{O}|\psi\rangle\right)^* \tag{1.8}$$

for any $|\psi\rangle$ and $|\phi\rangle$ in $\mathcal{E}_H$. An operator $\hat{O}$ is *Hermitian* or *self-adjoint* if $\hat{O} = \hat{O}^\dagger$ or, in other words, if $\langle\psi|\hat{O}^\dagger|\phi\rangle = \langle\phi|\hat{O}|\psi\rangle$ for any $|\psi\rangle$ and $|\phi\rangle$. If an operator $\hat{O}$ is Hermitian, the expectation value of the observable $\hat{O}$ defined in Eq. (1.7) is real. In other words:

$$\hat{O} = \hat{O}^\dagger \Rightarrow \langle o\rangle = \langle o\rangle^* \,. \tag{1.9}$$

It is important to remember how to take the Hermitian conjugate of an expression. This is obtained by following this set of rules:

- Reversing the order of the terms.
- Transforming
  1. operators into their adjoints, $\hat{O} \to \hat{O}^\dagger$.
  2. kets into bras and viceversa, $|\psi\rangle \to \langle\psi|$.
  3. numbers into their complex conjugates, $\lambda \to \lambda^*$.

Using these rules, the Hermitian conjugate of the expression $\lambda|\psi\rangle\hat{A}^\dagger\langle\phi|\hat{B}$ is $\lambda^*\hat{B}^\dagger|\phi\rangle\hat{A}\langle\psi|$.

1.1.4 Eigenvalues and eigenvectors

An operator $\hat{O}$ typically transforms a wave function into another. In some cases of special interest, however, a wave function is transformed into itself up to a factor,

$$\hat{O}\,|\psi_\alpha\rangle = o_\alpha\,|\psi_\alpha\rangle \,. \tag{1.10}$$

In these specific cases, the quantities o_α are called *eigenvalues*. The vectors $|\psi_\alpha\rangle$ are *eigenvectors*. In countable Hilbert spaces, α is a discrete index that typically labels individual states, $\alpha = 0, 1, \cdots, N$. This discrete index is what we call a *quantum number*, characterising the different states of the system. In continuous Hilbert spaces, α is a continuous index.

The eigenvalues and eigenvectors of Hermitian operators have special properties. The eigenvalues, in particular, are real. In addition, two eigenvectors associated to two different eigenvalues are necessarily orthogonal[3]:

$$\langle\psi_\alpha|\psi_\beta\rangle = 0, \text{ for } \alpha \neq \beta \,.$$

[3] Some of these questions will be explored in the tutorial sessions.

More importantly, the eigenstates of a Hermitian operator in a Hilbert space $\mathcal{E}_H$ are a *complete set*: we can represent any other state in this space as a weighted sum over such sets of states.

1.1.5 Dirac notation

Dirac, *bra-ket* or bracket notation is a general, abstract way of writing down the relevant equations of quantum mechanics. To every ket, $|\psi\rangle$, of the Hilbert space $\mathcal{E}_H$, we can associate an element of the dual conjugate space $\mathcal{E}_H^*$ denoted as *bra*, $\langle\psi|$. The action of the bra $\langle\psi|$ on the ket $|\phi\rangle$ is equal to the scalar product:

$$\langle\psi|(\phi\rangle) = \langle\psi|\phi\rangle \,. \tag{1.11}$$

In this sense, bras act on ket states as operators would do.

One can also use bras and kets to denote compactly eigenstates of any operator. For the Hamiltonian operator, $\hat{H}$, we can denote the eigenvalue problem as

$$\hat{H}\,|\psi_n\rangle = E_n\,|\psi_n\rangle\ . \tag{1.12}$$

Similarly, we can take the Hermitian conjugate this relation to have a Hamiltonian operator that acts on the left bra:

$$\langle\psi_n|\,\hat{H}^\dagger = \langle\psi_n|\,E_n^*\,. \tag{1.13}$$

Since the Hamiltonian is an Hermitian operator and its eigenvalues are real, this becomes

$$\langle\psi_n|\,\hat{H} = \langle\psi_n|\,E_n\,. \tag{1.14}$$

In some instances, the braket notation can be used further. One often labels the eigenstate kets by the eigenstate label itself, eg $|n\rangle$ instead of $|\psi_n\rangle$. For operators with continuous spectra, like the momentum operator $\hat{p}$, one may even label the eigenstate with the continuous variable,

$$\hat{p}\,|p\rangle = p\,|p\rangle\ . \tag{1.15}$$

In Dirac notation, the expression $|\psi\rangle\langle\phi|$ can be considered an operator, in the sense that it can act on a ket to deliver a different ket in the same Hilbert space,

$$(|\psi\rangle\langle\phi|)\,|\chi\rangle = |\psi\rangle\,\underbrace{\langle\phi|\chi\rangle}_{\lambda} = \lambda|\psi\rangle\,, \tag{1.16}$$

where λ is in general a complex number $\lambda = \langle\phi|\chi\rangle$.

1.1.6 Spectral theorems and decompositions

In nonrelativistic quantum mechanics, Hilbert spaces are separable, implying the existence of at least one countable orthonormal basis. Let $\{|n\rangle, n = 1, 2, \dots\}$ denote a basis (or a countable set of ket states) such that

$$\langle m|n\rangle = \delta_{n,m}\,. \tag{1.17}$$

Any ket $|\psi\rangle$ or bra $\langle\psi|$ can be decomposed in this basis

$$|\psi\rangle = \sum_n C_n|n\rangle\,, \quad \langle\psi| = \sum_n C_n^*\langle n|\,. \tag{1.18}$$

In general, there is an infinite number of nonvanishing coefficients C_n. The set $\{C_n\}$ in the basis $\{|n\rangle\}$ defines the state $|\psi\rangle$ completely and is therefore a new *representation* of the state. The product of $|\psi\rangle = \sum_n C_n|n\rangle$ and $|\phi\rangle = \sum_n B_n|n\rangle$ is given by

Choosing a basis that minimizes the number of coefficients is an important skill of the quantum mechanics practitioner!

$$\langle\phi|\psi\rangle = \sum_n B_n^* C_n\,. \tag{1.19}$$

The norm of any ket is given by the sum of squares of C_n,

$$\langle\psi|\psi\rangle = \sum_n |C_n|^2 \,, \tag{1.20}$$

which typically equals unity.

Consider the operator

$$\hat{P}_n = |n\rangle\langle n| \,. \tag{1.21}$$

This is a projection operator (or *projector*) onto the state $|n\rangle$. Indeed, acting with P_n on $|\psi\rangle$,

$$\hat{P}_n|\psi\rangle = \hat{P}_n \sum_k C_k|k\rangle = |n\rangle\langle n|\sum_k C_k|k\rangle = C_n|n\rangle \,, \tag{1.22}$$

we find that P_n filters out the component of $|\psi\rangle$ that belongs to the state $|n\rangle$. A projector is an *idempotent* operator, $\hat{P}_n^2 = \hat{P}_n$.

One can extend this property to sets of states, rather than individual states. Consider the subspaces $\mathcal{E}_\nu$ generated by a subset of different states $\{|n\rangle, n \in \nu\}$. The projector onto this subspace is simply $\hat{P}_\nu = \sum_{n\in\{\nu\}} \hat{P}_n$. The projection of a state $|\psi\rangle$ onto the entire Hilbert space is nothing but $|\psi\rangle$ itself. This gives the important *closure* relation

$$\sum_{\text{all } n} \hat{P}_n = \sum_{\text{all } n} |n\rangle\langle n| = \hat{I} \,, \tag{1.23}$$

where $\hat{I}$ is the identity operator.

The set of eigenvalues of a Hermitian operator $\hat{O}$ is denoted by $\{o_\alpha, \alpha = 0, 1, 2, \dots\}$, and the normalized eigenvectors are denoted as $|\alpha, r\rangle$, $\hat{O}|\alpha^r\rangle = o_\alpha|\alpha^r\rangle$. The extra label r accounts for the fact that, on occasion, different eigenvectors, labelled by different values of $r = 1, \dots, n_\alpha$, yield the same eigenvalue o_α. We say that the eigenvalues are *degenerate* and n_α is the degree of degeneracy. The eigenvectors are orthonormal so that

$$\langle\alpha'^{r'}|\alpha^r\rangle = \delta_{\alpha,\alpha'}\delta_{r,r'} \,. \tag{1.24}$$

We also have the equivalent decomposition of the identity,

$$\hat{I} = \sum_\alpha \sum_{r=1}^{n_\alpha} |\alpha^r\rangle\langle\alpha^r| \,. \tag{1.25}$$

The eigenvectors of a Hermitian operator play a special role because, when considered altogether, they can be used as a basis. This is guaranteed by the following theorem:

Spectral (or Riesz') theorem

The set $\{|\alpha^r\rangle\}$ of orthonormal eigenvectors of a Hermitian operator $\hat{O}$ is a complete orthonormal basis of $\mathcal{E}_H$.

There are several interesting consequences of this theorem. For instance, any vector $|\psi\rangle$ can be expanded in the set of eigenvectors

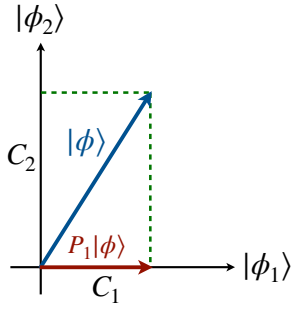

Figure 1.1: A state $|\psi\rangle$ can be represented as the combination of two state vectors, $|\psi\rangle = C_1|\phi_1\rangle + C_2|\phi_1\rangle$. The projector operator P_1 filters out the first component, $P_1|\psi\rangle = C_1|\phi_1\rangle$.

of any Hermitian operator $\hat{O}$, $\{|\alpha^r\rangle\}$. The operator $\hat{O}$ has a spectral decomposition, that can be written as:

$$\hat{O} = \sum_{\alpha} \sum_{r=1}^{n_\alpha} o_\alpha |\alpha^r\rangle\langle\alpha^r| \,. \tag{1.26}$$

This happens to be useful because it is expressed entirely in terms of diagonal objects in α. In other words, when working on the basis generated by $\hat{O}$, one only requires the diagonal matrix elements of the operator. In a countable space of dimension N, there are only N associated eigenvalues, as opposed to the full operator matrix which, in any other basis, would have a size $N \times N$.

1.2 *Postulates of quantum mechanics*

Quantum Mechanics can be formulated in terms of a few simple theoretical principles, or *postulates*. These principles are based on experimental observations, which we will not cover in this course. There are extensive descriptioms of such experiments in elementary quantum mechanics literature [4].

[4] D. J. Griffiths and D. F. Schroeder. *Introduction to Quantum Mechanics*. Cambridge University Press, third edition, 2018; and A. Messiah. *Quantum Mechanics*. Dover Press, 1999

The goal of this section is to introduce these postulates. We use the key mathematical concepts introduced in the previous section, that are necessary to operate with quantum mechanics. To keep the notation as simple as possible, expressions are written for a $1-$dimensional system, but the generalization to many dimensions is usually straightforward

The first postulate provides a framework for quantum mechanics. It sets up a basis for the description of the state of a system in terms of wave functions.

Postulate 1: the superposition principle

We can associate a Hilbert space $\mathcal{E}_H$ to any physical system. At each time t, the state of the system is completely determined by a normalized vector $|\psi(t)\rangle$ of $\mathcal{E}_H$.

This postulate states that a wave function is nothing but a vector in a Hilbert space. Moreover, the theorem does not exclude the possibility that normalized linear superpositions $|\psi\rangle = \sum_\alpha C_\alpha |\psi_\alpha\rangle$ of state vectors $|\psi_\alpha\rangle$, where C_α are complex, and such that $\sum_\alpha C_\alpha^2 = 1$, are accessible state vectors. Different linear combinations can provide descriptions of the system - hence the idea of *superposition*. In other words, if $|\psi_1\rangle$ and $|\psi_2\rangle$ are possible eigenstates of the system, so is the linear combination $|\psi\rangle = a_1 |\psi_1\rangle + a_2 |\psi_2\rangle$, with a_1 and a_2 complex constants.

The following postulates, 2 to 5, all relate to the *measurement* process. We do not define in detail what a measurement is, but suffice to say that it involves a repeated experimental process in which sufficient (in the ideal case, infinite) statistics are obtained to ascertain the properties of certain physical observables of the system.

We assume that the initial state is a quantum state $|\psi\rangle$ and that the measurement process will affect this state in some way. The following postulate clarifies how physical observables are represented in quantum mechanics.

Postulate 2

With each physical quantity A, one can associate a linear and Hermitian operator $\hat{A}$ acting in $\mathcal{E}_H$. $\hat{A}$ is the observable *which represents quantity A.*

The third postulate is a radical statement regarding the possible values of any given observable during a measurement.

Postulate 3: the quantization principle

The only possible result of the measurement of an observable A is one of the eigenvalues a_α of the corresponding operator $\hat{A}$.

For continuous operators, this statement does not have much interest, in the sense that observables could potentially be measured to have any real value. For discrete operators, however, this principle states that they are quantized. The measurement of a given set of observables (energies, spins) can only yield a very specific set of values.

The following postulate describes the probabilities associated to the different outcomes of the measurement of a given observable. Assume that the state of the system before the measurement is given by $|\psi\rangle$. We denote by $\hat{P}_\alpha$ the projector onto the subspace associated with the eigenvalue a_α.

The state $|\alpha\rangle$ could be degenerate and α could be a continuous variable too. We ignore these complications for the time being.

Postulate 4

The probability of finding the value of a_α in a measurement of A is $\mathcal{P}(a_\alpha) = \langle\psi|\hat{P}_\alpha|\psi\rangle = |\langle\psi|\psi_\alpha\rangle|^2$.

This postulate states that the *square* of the wave function can be interpreted as a probability. Accordingly, the expectation value of an observable is given by the expression

$$\langle a\rangle = \sum_\alpha a_\alpha \mathcal{P}(a_\alpha)\,. \tag{1.27}$$

Knowing that $\hat{A} = \sum_\alpha a_\alpha \hat{P}_\alpha$, it is easy to see that this is equivalent to Eq. (1.7).

The following postulate dictates what happens to a state right after the measurement has been performed.

Postulate 5 or the collapse of the wavepacket

Immediately after the measurement of A that gives the result a_α has been performed, the new state of system is the normalized projection of $|\psi\rangle$ onto the eigensubspace associated with a_α,

$$|\psi'\rangle = \frac{\hat{P}_\alpha|\psi\rangle}{\sqrt{\langle\psi|\hat{P}_\alpha|\psi\rangle}}.$$

This is the most controversial of the postulates of quantum mechanics, and the most difficult to get comfortable with. It is motivated by experience with repeated measurements. If an experimental sample is prepared in a state $|\psi\rangle$ then it is observed that a measurement of A can yield a variety of results, with probabilities $|\langle a_n|\psi\rangle|^2$. In other words, identically prepared initial states yield different, statistically distributed experimental outcomes, as dictated by the fourth postulate. However, if A is measured with outcome a_n on a given system, and then the state of the system is immediately remeasured after that, the results of the second measurement are **not** anymore statistically distributed anymore. Instead, the system is always in the state $|a_n\rangle$.

The final postulate introduces the time evolution of states so long as no measurements are performed.

Postulate 6 or the time evolution

The time evolution of the state vector $|\psi(t)\rangle$ is governed by the Schrödinger equation, where $\hat{H}(t)$ is the observable associated with the total energy of the system (called the Hamiltonian):

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}(t)\,|\psi(t)\rangle\ . \tag{1.28}$$

The Hamiltonian operator can be written as

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

where $V(x)$ is the potential energy and m is the mass of the system. Assuming an initial state $|\psi(t_0)\rangle$, the formal solution to this equation for a time-independent hamiltonian is

$$|\psi(t)\rangle = \exp\left(-\frac{i}{\hbar}\hat{H}(t-t_0)\right)|\psi(t_0)\rangle\,. \tag{1.29}$$

The unitary operator $\hat{U}(t,t_0) = \exp\left(-\frac{i}{\hbar}\hat{H}(t-t_0)\right)$ is usually called the time evolution operator.

Assuming an isolated system with a time-independent hamiltonian, the energy eigenstates at any time are given by $\hat{H}\,|\psi_\alpha\rangle = E_\alpha\,|\psi_\alpha\rangle$. The set $\{\psi_\alpha\}$ forms a complete basis in the Hilbert

space $\mathcal{E}_H$, and we can expand on it any vector $|\psi(t)\rangle$. At $t = 0$, we expand the eigenstate in this basis

$$|\psi(t=0)\rangle = \sum_\alpha C_\alpha |\psi_\alpha\rangle \text{ , with } C_\alpha = \langle\psi_\alpha|\psi(t=0)\rangle \text{ .}$$

At any later time, t, the expansion coefficients are different and, in general, time-dependent, so we write

$$|\psi(t)\rangle = \sum_\alpha \lambda_\alpha(t) |\psi_\alpha\rangle \tag{1.30}$$

, keeping in mind that $\lambda_\alpha(t=0) = C_\alpha$. The Schrödinger equation implies that

This can be seen by expanding the left-hand side of the equation, $i\hbar \frac{d}{dt} |\psi(t)\rangle = \sum_\alpha i\hbar \frac{d}{dt} \lambda_\alpha(t) |\psi_\alpha\rangle$, and then the right-hand side, $\hat{H} |\psi(t)\rangle = \sum_\alpha \hat{H} \lambda_\alpha(t) |\psi_\alpha\rangle = \sum_\alpha \lambda_\alpha(t) E_\alpha |\psi_\alpha\rangle$. Since $|\psi_\alpha\rangle$ is a complete basis, the coefficients in the sum must be the same one by one.

$$i\hbar \frac{d}{dt} \lambda_\alpha(t) = E_\alpha \lambda_\alpha(t) \text{ ,}$$

This is a trivial first-order differential equation for the coefficients, with solution $\lambda_\alpha(t) = e^{-i\frac{E_\alpha t}{\hbar}}$. This solution allows us to find the expansion of the state at *any* time, t:

$$|\psi(t)\rangle = \sum_\alpha C_\alpha^{-iE_\alpha t/\hbar} |\psi_\alpha\rangle \text{ .} \tag{1.31}$$

We stress that this explicit solution is possible because we are working with a time-independent Hamiltonian and expanding over its eigenstates. It is otherwise rather difficult to find close expressions for solutions of Eq. (1.28) when the Hamiltonian has an explicit time dependence.

1.2.1 *Transformations of states and observables*

We shall now look at an example of the importance of the postulates, which involves symmetries and invariance properties. We expect that the laws of nature shall be invariant under certain space-time operations. When such invariances appear, they tend to lead to important simplifications of the mathematical formulation of the problems at stake.

Consider one such transformation going from a reference frame $\mathcal{S}$ to a frame $\mathcal{S}'$. This must involve a transformation of the system's wave function, $\psi \to \psi'$. We also expect operators to change when a transformation is performed, so that $\hat{A} \to \hat{A}'$. These transformations are generic, but the postulates of quantum mechanics provide some important constraints on the mathematical structure of the transforms.

Examples of such transformations are the Galilean transformation; a (fixed) rotation of the axis; or a displacement of the origin of time.

The eigenvalue equation for the operator $\hat{A}$ in the frame $\mathcal{S}$ reads $\hat{A}|\phi_\alpha\rangle = a_\alpha|\phi_\alpha\rangle$. In the transformed frame, $\mathcal{S}'$, the equation instead reads $\hat{A}'|\phi'_\alpha\rangle = a_\alpha|\phi'_\alpha\rangle$. Importantly, the eigenvalues in this expression are not primed. The eigenvalues of $\hat{A}$ and $\hat{A}'$ must be the same because they represent equivalent observables from a quantum mechanical perspective. According to the second and third postulates, they must have the same set of possible values.

How are wave functions in different reference frames related to each other? Let us expand a given state into the eigenvectors of $\hat{A}$,

$|\psi\rangle = \sum_\alpha c_\alpha |\phi_\alpha\rangle$. The transformed ket will necessarily read $|\psi'\rangle = \sum_\alpha c'_\alpha |\phi'_\alpha\rangle$, where $|\phi'_\alpha\rangle$ are the eigenvectors of $\hat{A}'$. However, according to the fourth postulate, the probabilities associated to these projections must be the same, so that $|c_\alpha|^2 = |c'_\alpha|^2$ or, stated differently, $|\langle\phi_\alpha|\psi\rangle|^2 = |\langle\phi'_\alpha|\psi'\rangle|^2$. These statements must hold, because they encode the fact that the *same events* in two different frames of reference must be described by the same probabilities.

Clearly, the two statements regarding operators and states limit substantially the possible transformations between different frames. The mathematical character of these transformations is specified by the following theorem due to Wigner:

Wigner theorem

Any mapping of a Hilbert space onto itself that preserves the value of $|\langle\phi|\psi\rangle|$ *may be implemented by an operator* $\hat{U}$*,*

$$|\psi\rangle \to |\psi'\rangle = \hat{U}|\psi\rangle\,, \quad |\phi\rangle \to |\phi'\rangle = \hat{U}|\phi\rangle\,, \tag{1.32}$$

with $\hat{U}$ *being* unitary *or* antiunitary*.*

By far, the most common case is the *unitary* one. If $\hat{U}$ is unitary, then $\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}$. Clearly, then, $\langle\phi'|\psi'\rangle = \langle\phi|\hat{U}^\dagger\hat{U}|\psi\rangle = \langle\phi|\psi\rangle$. In other words, unitary transformations preserve the complex value of inner products. One can prove that any continuous transformation must be unitary[5]. Continuous transformations include rotations, translations and Galilean transformations. Some discrete transformations, like a parity transformation, are also unitary.

[5] L. E. Ballentine. *Quantum Mechanics: A Modern Development*. World Scientific, 1998

Antiunitary transformations are somewhat more exotic. If $\hat{U}$ is antiunitary, then $\hat{U}c|\psi\rangle = c^*\hat{U}|\psi\rangle$, where c is a complex number. If $\hat{U}$ is antiunitary, $\langle\phi'|\psi'\rangle = \langle\phi|\psi\rangle^*$ and hence the phase of the inner product of two vectors in Hilbert space can change. Antiunitary transformations must be discrete. Time reversal is an example of an antiunitary transformation.

We have postulated that a reference frame transformations must also involve operator transformations, $\hat{A} \to \hat{A}'$. To find the relation between $\hat{A}'$, $\hat{A}$ and $\hat{U}$, consider the equation $\hat{A}'|\phi'_\alpha\rangle = a_\alpha|\phi'_\alpha\rangle$ in the "primed" frame. Substituting $|\phi'_\alpha\rangle = \hat{U}|\phi_\alpha\rangle$ yields $\hat{A}'\hat{U}|\phi_\alpha\rangle = a_\alpha\hat{U}|\phi_\alpha\rangle$ and hence $U^{-1}\hat{A}'\hat{U}|\phi_\alpha\rangle = a_\alpha|\phi_\alpha\rangle = \hat{A}|\phi_\alpha\rangle$. Since this equation holds for the complete set ϕ_α, it must hold for any arbitrary vector and hence $U^{-1}\hat{A}'\hat{U} = \hat{A}$. Put differently, the operator transformation that we were looking for is

$$\hat{A} \to \hat{A}' = \hat{U}\hat{A}\hat{U}^{-1}\,. \tag{1.33}$$

As an example, consider the unitary transformation associated to a displacement along axis i, $x_i \to x_i + s_i$, where $i = x, y, z$. One can show that the unitary operator associated to such a transformation is $\hat{U} = e^{-is_i\hat{p}_i}$, where $\hat{p}$ is the ith component of the momentum

operator[6]. In other words, $|x_i'\rangle = \hat{U}|x_i\rangle = e^{-is_i\hat{p}_i}|x_i\rangle = |x_i + s_i\rangle$. Consider the following eigenvalue equation for the position operator in the unprimed frame, $\hat{x}_i|x_i\rangle = x_i|x_i\rangle$. The equivalent equation in the primed frame reads $\hat{x}_i'|x_i'\rangle = x_i|x_i'\rangle$. It is important to note that the eigenvalue here does not change (it is not primed!). We can see this from the following expansion on the left-hand-side term:

$$\begin{aligned}\hat{x}_i'|x_i'\rangle = \hat{U}\hat{x}_i\hat{U}^{-1}|x_i + s_i\rangle = \hat{U}\hat{x}_i|x_i\rangle = \hat{U}x_i|x_i\rangle = x_i\hat{U}|x_i\rangle \\ = x_i|x_i + s_i\rangle = x_i|x_i'\rangle.\end{aligned}$$

With this, we have seen that the postulates of quantum mechanics constrain substantially the structure of the allowed transformations of frames of reference. Their unitary or anti-unitary nature simplifies several considerations in terms of symmetry. We shall now turn our attention to a different application of the postulates.

[6] L. E. Ballentine. *Quantum Mechanics: A Modern Development*. World Scientific, 1998

1.2.2 *Uncertainty principle of operators*

Consider two quantities A and B and their corresponding observable operators $\hat{A}$ and $\hat{B}$. Let $|\psi\rangle$ be a state of the system. Suppose we perform (repeated) measurements of $\hat{A}$ and $\hat{B}$, yielding the expectation values $\langle a\rangle$ and $\langle b\rangle$. The root-mean-square deviations of these observables are defined as $\Delta a^2 = \langle\hat{A}^2\rangle - \langle a\rangle^2$ and $\Delta b^2 = \langle\hat{B}^2\rangle - \langle b\rangle^2$. We want to relate these two quantities, knowing the state $|\psi\rangle$ and the operators $\hat{A}$ and $\hat{B}$. We first define new variables centered around the expectation values: $\underline{\hat{A}} = \hat{A} - \langle a\rangle$ and $\underline{\hat{B}} = \hat{B} - \langle b\rangle$. It is easy to prove[7] that $\langle\underline{\hat{A}}^2\rangle = \Delta a^2$ (and likewise for B). In other words, averaging $\underline{\hat{A}}^2$ yields the root-mean-square deviation of $\hat{A}$ straight away.

Let us now consider an arbitrary state $|\psi\rangle$ and the vector $\left(\underline{\hat{A}} + i\lambda\underline{\hat{B}}\right)|\psi\rangle$, where λ is a real constant. Let's find the square of the norm of this vector:

$$\begin{aligned}\left\|\underline{\hat{A}} + i\lambda\underline{\hat{B}}|\psi\rangle\right\|^2 &= \langle\psi|\left(\underline{\hat{A}} - i\lambda\underline{\hat{B}}\right)\left(\underline{\hat{A}} + i\lambda\underline{\hat{B}}\right)|\psi\rangle \\ &= \langle\underline{\hat{A}}^2\rangle + \lambda^2\langle\underline{\hat{B}}\rangle + i\lambda\left\langle\left[\underline{\hat{A}}, \underline{\hat{B}}\right]\right\rangle \\ &= \lambda^2\Delta b^2 + i\lambda\left\langle\left[\underline{\hat{A}}, \underline{\hat{B}}\right]\right\rangle + \Delta a^2\,. \qquad (1.34)\end{aligned}$$

This norm is a real number. Clearly, the first two terms are scalars, because $\underline{\hat{A}}$ and $\underline{\hat{B}}$ are Hermitian operators. As for the last term, the commutator operator $i\left[\underline{\hat{A}}, \underline{\hat{B}}\right]$ is also Hermitian[8], and its expectation value, $i\left\langle\left[\underline{\hat{A}}, \underline{\hat{B}}\right]\right\rangle$, must be real. The expression in Eq. (1.34) is the norm of a vector and, as such, it must be positive (or at most zero) for any value of the real constant λ. In other words, as a function of λ, this expression should thus not change sign.

To ensure this, the second-order polynomial in λ in the last term of Eq. (1.34) must not have any roots. This immediately means that the discriminant $\left\langle i\left[\underline{A}, \underline{\hat{B}}\right]\right\rangle^2 - 4\Delta a^2\Delta b^2 \leq 0$ must be negative (or at most zero). Rephrasing this expression, we find $\Delta a^2\Delta b^2 \geq \frac{1}{4}\left\langle i\left[\underline{A}, \underline{\hat{B}}\right]\right\rangle^2$.

[7] For any operator $\hat{A}^2$, take the expectation value and expand $\langle\underline{\hat{A}}^2\rangle = \langle\hat{A}^2 - \hat{A}\langle a\rangle - \langle a\rangle\hat{A} + \langle a\rangle^2\rangle = \langle\hat{A}^2\rangle - 2\langle a\rangle^2 + \langle a\rangle^2 = \langle\hat{A}^2\rangle - \langle a\rangle^2 = \Delta a^2$.

[8] To show this, consider $\left(i\left[\underline{\hat{A}}, \underline{\hat{B}}\right]\right)^\dagger = \left(i\underline{\hat{A}}\underline{\hat{B}} - i\underline{\hat{B}}\underline{\hat{A}}\right)^\dagger = -i\underline{\hat{B}}^\dagger\underline{\hat{A}}^\dagger + i\underline{\hat{A}}^\dagger\underline{\hat{B}}^\dagger = -i\underline{\hat{B}}\underline{\hat{A}} + i\underline{\hat{A}}\underline{\hat{B}} = i\left[\underline{\hat{A}}, \underline{\hat{B}}\right]$.

And, taking square roots on both sides,

$$\Delta a \Delta b \geq \frac{1}{2} \left| \left\langle \left[\hat{A}, \hat{B} \right] \right\rangle \right| ,$$

For the quantities $\hat{x}$ and $\hat{p}$, with the commutator $[\hat{x}, \hat{p}] = i\hbar$, this leads to the well-known Heisenberg uncertainty principle $\Delta x \Delta p \geq \frac{\hbar}{2}$.

where we have used $\left[\underline{\hat{A}}, \underline{\hat{B}}\right] = \left[\hat{A}, \hat{B}\right]$. Note that we use an absolute value, that arises when taking the square root and to avoid any ambiguities in sign on the commutator. This indicates that the product of the two rms deviations must be larger than the expectation value of the commutator of operators. In other words, this uncertainty relation states that the two mean square deviations cannot be made simultaneously small (except for the special cases where the two observables commute). This is a general form of the uncertainty relations for any two observables sometimes called the Robertson uncertainty relation[9].

[9] H. P. Robertson. The uncertainty principle. *Phys. Rev.*, 34:163–164, Jul 1929

1.2.3 *Ehrenfest's theorem*

As a final application of the postulates, we discuss a theorem due to Ehrenfest that provides an insight into the nature of conserved quantities in quantum mechanics[10]. Ehrenfest's theorem allows us to calculate the time evolution of expectation values using the sixth postulate. Consider a physical quantity A and its associated expectation value $\langle a \rangle$. Taking the time derivative of this expression

[10] P. Ehrenfest. Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik. *Zeitschrift fur Physik*, 45(7-8):455–457, July 1927

$$\begin{aligned} \frac{d}{dt}\langle a \rangle &= \frac{d}{dt}\langle \psi | \hat{A} | \psi \rangle \\ &= \left(\frac{d}{dt}\langle \psi | \right) \hat{A} | \psi \rangle + \langle \psi | \left(\frac{d}{dt} \hat{A} \right) | \psi \rangle + \langle \psi | \hat{A} \left(\frac{d}{dt} | \psi \rangle \right) , \end{aligned}$$

we find three terms. Using the Schrödinger equation, $i\hbar \frac{d}{dt} |\psi\rangle = \hat{H} |\psi\rangle$, and its conjugate, $-i\hbar \frac{d}{dt} \langle \psi | = \langle \psi | \hat{H}$, we can simplify two of these terms to find

$$\frac{d}{dt}\langle a \rangle = \frac{1}{i\hbar} \langle \psi | [\hat{A}, \hat{H}] | \psi \rangle + \langle \psi | \frac{d\hat{A}}{dt} | \psi \rangle .$$

If the operator under consideration is time-independent, the last term cancels and the expression further simplifies to

$$\frac{d}{dt}\langle a \rangle = \frac{1}{i\hbar} \langle \psi | [\hat{A}, \hat{H}] | \psi \rangle .$$

This is Ehrenfest's theorem, which provides the time evolution of physical quantities using only the commutator of their operators and the Hamiltonian.

We consider now a few examples of applications of this theorem:

- **Conservation of the norm**: for $\hat{A} = \hat{I}$, we obtain that $\frac{d\langle \psi | \psi \rangle}{dt} = 0$. In other words, the norm of a wave function must be preserved.
- **Conservation of energy**: for an isolated system and a time-independent hamiltonian, the choice $\hat{A} = \hat{H}$ leads to $\frac{d\langle E \rangle}{dt} = 0$.

- **Conservation of momentum**: for a free particle with Hamiltonian $\hat{H} = \frac{\hat{p}}{2m}$, clearly $[\hat{p}, \hat{H}] = 0$ and so momentum is conserved: $\frac{d\langle p \rangle}{dt} = 0$. This statement can be generalized to any hamiltonian that only depends on momentum.

2 The Harmonic Oscillator

2.1 Real-space representation

An harmonic oscillator is an object that is subject to a quadratic potential, producing a restoring force which is proportional to the displacement. The Hamiltonian associated to such a system for one spatial dimension is

$$\hat{H} = -\frac{\hbar}{2m}\frac{\partial^2}{\partial X^2} + \frac{1}{2}m\omega^2 X^2 \,, \tag{2.1}$$

where $P = i\hbar\partial_X$ is the momentum; X is the displacement; m, the mass of the object and ω, the frequency of oscillation (in radians per unit of time). The Harmonic Oscillator is useful in a variety of quantum mechanics settings, particularly where vibrations play an important role. Our aim is to find the solutions of the eigenvalue problem

$$\left(-\frac{\hbar}{2m}\frac{\partial^2}{\partial X^2} + \frac{1}{2}m\omega^2 X^2\right)\psi_n(X) = E_n\psi_n(X) \,, \tag{2.2}$$

where E_n are the specific energy values for which the solutions are square integrable. The solution of this well-known system is available virtually in all quantum mechanics textbooks[1].

A combination of $\hbar$, m and ω provides the key scales for energies and lengths in this problem. The energy scale is $\hbar\omega$, whereas the lengthscale is $\bar{x} = \sqrt{\frac{\hbar}{m\omega}}$. We work with the dimensionless energy $\epsilon = \frac{E}{\hbar\omega}$ and dimensionless displacement $x = \frac{X}{\bar{x}}$. With these, the Schrödinger equation becomes

$$\frac{1}{2}\left(-\frac{\partial^2}{\partial x^2} + x^2\right)\psi_n(x) = \epsilon_n\psi_n(x) \,. \tag{2.3}$$

This is a second-order differential equation and its solutions are well-known. The square-integrable eigenvectors are

$$\psi_n(x) = \mathcal{N}_n e^{-\frac{x^2}{2}} H_n(x) \,, \tag{2.4}$$

with the normalization constants, $\mathcal{N}_n = \left(2^n n!\sqrt{\pi}\right)^{-1/2}$.We have introduced the Hermite polynomials $H_n(x) = (-1)^n e^{x^2}\frac{d^n}{d^n x}e^{-x^2}$ for $n = 0, 1, \ldots$[2]. Hermite polynomials of degree n involve only even (odd) powers of n if n is even (odd)[3].

The wave functions are orthonormal,

$$\langle n|m\rangle = \int \mathrm{d}x\psi_n(x)^*\psi_m(x) = \delta_{n,m} \,, \tag{2.5}$$

[1] E. Merzbacher. *Quantum Mechanics*. Wiley, 1998; J.-L. Basdevant and J. Dalibard. *Quantum Mechanics*. Springer, 2002; and L. E. Ballentine. *Quantum Mechanics: A Modern Development*. World Scientific, 1998

[2] The first few Hermite polynomials are

$$H_0(x) = 1 \,,$$
$$H_1(x) = 2x \,,$$
$$H_2(x) = 4x^2 - 2 \cdots$$

[3] G. B. Arfken, H. J. Weber, and F. E. Harris. *Mathematical Methods for Physicists*. Elsevier, 2005; and K. F. Riley and M. P. Hobson. *Essential Mathematical Methods for the Physical Sciences*. Cambridge University Press, 2011

and also fulfil the recursion relations:

$$\sqrt{2}x\psi_n(x) = \sqrt{n+1}\psi_{n+1}(x) + \sqrt{n}\psi_{n-1}(x)\,, \tag{2.6}$$

$$\sqrt{2}\frac{d}{dx}\psi_n(x) = \sqrt{n}\psi_{n-1}(x) - \sqrt{n+1}\psi_{n+1}(x)\,. \tag{2.7}$$

The $n = 0, 1, 2$ and 3 wave functions are shown in Fig. 2.1.

The corresponding eigenvalues are given by

$$\epsilon_n = n + \frac{1}{2}\,. \tag{2.8}$$

This can be obtained by plugging the eigenvectors of Eq. (2.4) into Eq. (2.3), and exploiting the fact that the Hermite polynomials are the solutions of Hermite's equation, $\partial_x^2 H_n - 2x\partial_x H_n = -2nH_n$.

Importantly, the Harmonic Oscillator has a non-null *zero point energy*, $\epsilon_0 = \frac{1}{2}$, unlike any classical system.

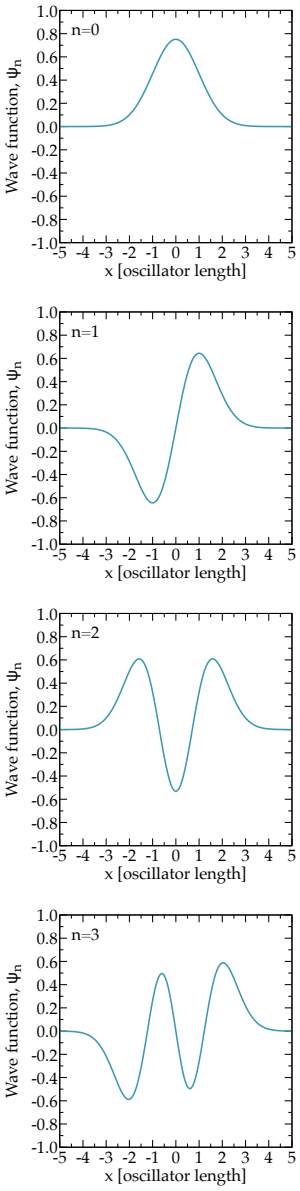

Figure 2.1: The $n = 0, 1, 2$ and 3 wavefunctions $\psi_n(x)$ of the harmonic oscillator. Note that n coincides with the number of nodes of each wavefunction.

2.2 *Algebraic method and ladder operators*

We now proceed to solve the Harmonic Oscillator Hamiltonian of Eq. (2.1) and find its eigenvalues and eigenvectors with a very different procedure. We do not rely on any differential equation knowledge. Instead, we will work with directly with operators in the Hilbert space generated by the solutions themselves. We start by introducing the momentum operator, $\hat{P} = -i\hbar\frac{\partial}{\partial X}$. Our solution relies on the fact that these conjugate variables fulfil the commutation relation

$$\left[\hat{X}, \hat{P}\right] = i\hbar\,. \tag{2.9}$$

Note that, strictly speaking, $\hat{X}$ does not need to represent a spatial dimension. In some cases of interest, $\hat{X}$ may be a distance variable in some abstract space (eg, an operator $\hat{Q}$). In such a situation, the momentum $\hat{P}$ should be the *canonical conjugate* variable of $\hat{Q}$ (in the sense of Hamiltonian dynamics).

We work in dimensionless variables to find our solution. We use the momentum scale $\bar{p} = (m\hbar\omega)^{-1/2} = \frac{\hbar}{\bar{x}}$ to introduce a dimensionless momentum operator $\hat{p} = \frac{\hat{P}}{\bar{p}} = (m\hbar\omega)^{-1/2}\hat{P}$. The dimensionless position and momentum satisfy the commutation relation $[\hat{x}, \hat{p}] = i$. With these, the Schrödinger equation can be manipulated,

$$\frac{1}{2}\left(-\frac{\partial^2}{\partial x^2} + x^2\right)\underbrace{\psi_n(x)}_{\langle x|n\rangle} = \langle x|\frac{1}{2}\left(\hat{p}^2 + x^2\right)|n\rangle\,, \tag{2.10}$$

to show that the Hamiltonian becomes

$$\hat{H} = \frac{1}{2}\left(\hat{p}^2 + \hat{x}^2\right)\,. \tag{2.11}$$

We now introduce two operators that will allow us to find a solution. These are the so-called *ladder operators*. The lowering

operator is the following linear combination of $\hat{p}$ and $\hat{x}$,

$$\hat{a} = \frac{\hat{x} + i\hat{p}}{\sqrt{2}}\,, \tag{2.12}$$

whereas the raising operator is

$$\hat{a}^\dagger = \frac{\hat{x} - i\hat{p}}{\sqrt{2}}\,. \tag{2.13}$$

These relations can be inverted to find the position and momentum opertors as a function of $\hat{a}$ and $\hat{a}^\dagger$:

$$\hat{x} = \frac{\hat{a} + \hat{a}^\dagger}{\sqrt{2}} \tag{2.14}$$

$$\hat{p} = i\frac{\hat{a}^\dagger - \hat{a}}{\sqrt{2}}\,. \tag{2.15}$$

The lowering and raising operators are Hermitian conjugates of each other[4]. They fulfil the commutation relation

$$\left[\hat{a}, \hat{a}^\dagger\right] = 1\,, \tag{2.16}$$

and we can rewrite the Hamiltonian as

$$\hat{H} = \hat{a}^\dagger\hat{a} + \frac{1}{2}\,. \tag{2.17}$$

Clearly, the spectrum of $\hat{H}$ and the spectrum of the so-called number operator,

$$\hat{N} = \hat{a}^\dagger\hat{a}\,, \tag{2.18}$$

are the same.

Let us thus attempt to find the spectrum of $\hat{N}$. The first key point is to obtain the commutation relations between $\hat{N}$, $\hat{a}$ and $\hat{a}^\dagger$. To do so, we exploit the relation $[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}$, so that

$$[\hat{N}, \hat{a}] = [\hat{a}^\dagger\hat{a}, \hat{a}] = \hat{a}^\dagger \underbrace{[\hat{a}, \hat{a}]}_{0} + \underbrace{[\hat{a}^\dagger, \hat{a}]}_{-1}\hat{a} = -\hat{a} \tag{2.19}$$

and, similarly, $[\hat{N}, \hat{a}^\dagger] = \hat{a}^\dagger$.

Now consider an eigenstate of the number operator with eigenvalue ν,

$$\hat{N}|\nu\rangle = \nu|\nu\rangle\,, \tag{2.20}$$

and consider the effect of inserting an operator $\hat{a}$ in between $\hat{N}$ and the state,

$$\hat{N}\hat{a}|\nu\rangle = \left(\hat{a}\hat{N} - \hat{a}\right)|\nu\rangle = \hat{a}\left(\hat{N} - 1\right)|\nu\rangle = (\nu - 1)\,\hat{a}|\nu\rangle\,, \tag{2.21}$$

where we have used the result of the commutator in Eq. (2.19). Reading the leftmost and the rightmost formulae above, it becomes clear that $\hat{a}|\nu\rangle$ is an eigenstate of $\hat{N}$ with eigenvalue $\nu - 1$, provided that $\hat{a}|\nu\rangle \neq 0$. Let us find the norm of the vector $\hat{a}|\nu\rangle$,

$$\left(\langle\nu|\hat{a}^\dagger\right)\left(\hat{a}\,|\nu\rangle\right) = \left\langle\nu\middle|\hat{N}\middle|\nu\right\rangle = \nu\,\langle\nu|\nu\rangle = \nu \geq 0\,, \tag{2.22}$$

[4] But they are not Hermitian operators. This will become clearer in the following chapter.

Since the norm is positive, so must ν - meaning the eigenvalues must necessarily be positive. Starting from an arbitrary real value of ν, and acting with the operator $\hat{a}$ repeatedly, one would get the eigenvalues $\nu-1$, $\nu-2$, $\nu-3$ and so on. This clarifies why $\hat{a}$ is called a *lowering* operator: it lowers the number of modes by one unit.

The action of the lowering operator on a state $|\nu\rangle$ is proportional to $|\nu-1\rangle$, so that $\hat{a}|\nu\rangle = C_\nu|\nu-1\rangle$. To determine C_ν, we again compute the norm of the vector, Eq. (2.22),

$$|C_\nu|^2 = (\langle\nu|\hat{a}^\dagger)(\hat{a}|\nu\rangle) = \langle\nu|\hat{N}|\nu\rangle = \nu, \tag{2.23}$$

and so

$$\hat{a}|\nu\rangle = \sqrt{\nu}|\nu-1\rangle,\ (\nu \neq 0). \tag{2.24}$$

In the last step, we are free to choose the irrelevant phase of C_ν and set it to a real constant.

What is the minimum value that ν can take? Starting from any arbitrary real value, and by successive lowering states, we would at some point reach a state $|\nu'\rangle$ with an eigenvalue between 0 and 1. By applying $\hat{a}$ again on $|\nu'\rangle$, we would get a negative eigenvalue - in contradiction to Eq. (2.22). In other words, if ν' was a non-integer value, we could construct negative eigenvalues of $\hat{N}$. Since this is not possible, we must assume ν to be a positive integer, $\nu = n$, with $n = 0, 1, 2, \ldots$. With this, the lowest step in the ladder is the state $\nu' = 0$, which simultaneously fulfils the equations

$$\hat{N}|0\rangle = 0|0\rangle, \tag{2.25}$$

$$\hat{a}|0\rangle = 0\,. \tag{2.26}$$

The latest equality is compatible with Eq. (2.24), which shows that the lowest eigenvalue $n = 0$ is such that $\hat{a}|0\rangle = 0$. This is also an eigenvector of the number operator with zero eigenvalue, $\hat{N}|0\rangle = 0$, and is typically called the *vacuum* (or the ground) state.

What is the effect of acting with the Hermitian conjugate operator. $\hat{a}^\dagger$? Following similar steps as earlier on, we find

$$\hat{N}\hat{a}^\dagger|\nu\rangle = \hat{a}^\dagger\left(\hat{N}+1\right)|\nu\rangle = (\nu+1)\,\hat{a}^\dagger|\nu\rangle\,. \tag{2.27}$$

$\hat{a}^\dagger|\nu\rangle$ is clearly an eigenvector of $\hat{N}$ with eigenvalue $\nu+1$. Acting repeatedly with $\hat{a}^\dagger$, one obtains a sequence of eigenvalues which is one unit greater than the predecessor. This is called a *raising* operator.

Since the sequence begins at $|0\rangle$, the allowed values of ν (and hence the spectrum of $\hat{N}$) are nonnegative integers $\nu = n$, so that

$$\hat{N}|n\rangle = n|n\rangle\,,\ n = 0, 1, 2, \ldots \tag{2.28}$$

We know that $\hat{a}^\dagger|n\rangle$ is proportional to $|n+1\rangle$, so we can write $\hat{a}^\dagger|n\rangle = C'_n|n+1\rangle$. To obtain the constant C'_n, we find again the norm of the vector

$$\left|C'_n\right|^2 = (\langle n|\hat{a})(\hat{a}^\dagger|n\rangle) = \langle n|\hat{N}+1|n\rangle = n+1. \tag{2.29}$$

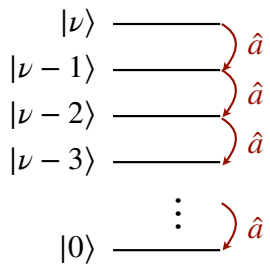

Figure 2.2: Acting with the lowering operator $\hat{a}$ repeatedly starting from $|\nu\rangle$, one obtains successive eigenstates with one lower unit, down to the vacuum state $|0\rangle$.

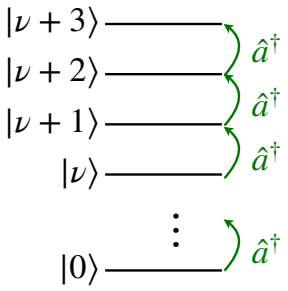

Figure 2.3: Acting with the raising operator $\hat{a}^\dagger$ repeatedly starting from $|\nu\rangle$, one obtains successive eigenstates with one more unit.

The phase of the vector $|n+1\rangle$ is arbitrary, and one may choose it to be real and positive so that $C'_n = \sqrt{n+1}$ and

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle. \tag{2.30}$$

Starting from the vacuum, we can build the different modes in the spectrum:

$$a^\dagger|0\rangle = |1\rangle, \tag{2.31}$$
$$a^\dagger|1\rangle = \sqrt{2}|2\rangle, \tag{2.32}$$
$$a^\dagger|2\rangle = \sqrt{3}|3\rangle \tag{2.33}$$

and so on. From this, it becomes clear that we can access any mode n by acting with $\hat{a}^\dagger$ on the vaccum n times:

$$|n\rangle = \frac{1}{\sqrt{n!}}\left(a^\dagger\right)^n|0\rangle. \tag{2.34}$$

We now have all the information we need to solve for the spectrum of the hamiltonian. Since, according to Eq. (2.17), we have $\hat{H} = \hat{N} + \frac{1}{2}$, and we know the eigenvectors of $\hat{N}$, it follows that:

$$\hat{H}|n\rangle = \epsilon_n|n\rangle\,, \text{ with } \epsilon_n = n + \frac{1}{2}. \tag{2.35}$$

This is the same result we obtained in real space, but this time relying only on the properties of the *ladder* operators, $\hat{a}$ and $\hat{a}^\dagger$. This fully algebraic solution of the harmonic oscillator does not rely on any differential equation knowledge.

2.2.1 *Degeneracy*

An underlying assumption in the previous derivations was that the eigenvalues ϵ_n are not degenerate. This becomes clear *a posteriori* when solving for the spectrum, Eq. (2.35). Can one prove this *a priori*? The answer is yes[5]. We start by looking at the vacuum state, which has $\epsilon_0 = \frac{1}{2}$. All the potentially degenerate states vacuum states must satisfy the equation $\hat{a}|0^r\rangle = 0$, where the index r denotes different degenerate states. To find the degeneracy of the ϵ_0 level, we can write:

$$\hat{a}|0^r\rangle = \frac{1}{\sqrt{2}}\left(\hat{x} + i\hat{p}\right)|0^r\rangle = 0\,. \tag{2.36}$$

In the real space representation, the previous expression becomes

$$\langle x|\hat{a}|0^r\rangle = \frac{1}{\sqrt{2}}\left(x - \frac{d}{dx}\right)\phi_0^r(x) = 0\,, \tag{2.37}$$

where $\langle x|0^r\rangle = \phi_0^r(x)$. The solution to this first-order differential equation is known and unique up to a constant, so that

$$\phi_0^r(x) = \mathcal{N}^r e^{-\frac{x^2}{2}}\,. \tag{2.38}$$

If $\phi_0^r(x)$ is normalized, then clearly $\mathcal{N}^r$ can only take one value and, in fact, $\mathcal{N}^r = \pi^{-1/4}$. There is no other possible source of r

[5] C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum Mechanics*, volume 1. Wiley-VCH, 1977

dependence on the right hand side of this equation - so the state cannot be degenerate. Another way of stating this is that the potentially different solutions to Eq. (2.37) are all proportional to one another.But proportional solutions are not physically different to each other - which means all the eigenstates must be the same (hence not degenerate).

We can now prove by recurrence that the remainder of the spectrum is non-degenerate. To do this, we need to prove that if the state with energy $\epsilon_n = n + \frac{1}{2}$ is non-degenerate, then the state with energy $\epsilon_{n+1} = n + \frac{3}{2}$ is not either. Let us assume that there exists only one level that satisfies

$$\hat{N}|n\rangle = n|n\rangle\,. \tag{2.39}$$

For the eigenvalue $n+1$, we could potentially have different eigenstates labelled by r:

$$\hat{N}|(n+1)^r\rangle = (n+1)|(n+1)^r\rangle\,. \tag{2.40}$$

We know (by hypothesis) that the state $\hat{a}|n+1^r\rangle$ is not zero; not degenerate, and an eigenstate of $\hat{N}$ with eigenvalue n. This implies the existence of a complex constant $\mathcal{C}^r$ such that $\hat{a}|(n+1)^r\rangle = \mathcal{C}^r|n\rangle$. Now applying $\hat{a}^\dagger$ on both sides of this equation, we find:

$$\hat{a}^\dagger\hat{a}|(n+1)^r\rangle = (n+1)|(n+1)^r\rangle = \mathcal{C}^r\hat{a}^\dagger|n\rangle\,, \tag{2.41}$$

so that the $n+1$ states can be written as:

$$|(n+1)^r\rangle = \frac{\mathcal{C}^r}{n+1}\hat{a}^\dagger|n\rangle\,. \tag{2.42}$$

We have isolated the expressions for the kets $|(n+1)^r\rangle$ and found that all of them are proportional to the unique, non-degenerate state $\hat{a}^\dagger|n\rangle$. If the states are proportional to each other, they are not physically different - implying that the eigenstates can not be degenerate.

So, since $|0\rangle$ is not degenerate, $|1\rangle$ cannot be degenerate. Nor can the state $|2\rangle$ or, for that matter, any of the successive states. We can therefore safely assume that the states are non-degenerate.

Keep in mind that the argument above works so long as r labels degenerate states that cannot be unambiguously distinguished. If r is a distinguishable quantum number (say, for instance, the spin of a particle), the argument above can be manipulated to account for the fact that the r different states live in different sectors of a tensor product. For instance, for a two-state system with $r = 1, 2$, the constant in Eq. (2.38) would read $\mathcal{N}^{r=1} \otimes \mathcal{N}^{r=2}$. One can then prove that for a given value of r, one can unambiguously connect the different states $|0^r\rangle$, $|1^r\rangle$, etc to each other using ladder operators.

2.2.2 *Fock states*

According to what we have just explained, the vacuum state $|0\rangle$ is non-degenerate and, hence, unique. Applying the raising operator $\hat{a}^\dagger$ on it, we can generate all the eigenstates of the number operator

$\hat{N}$. In other words, we are generating a complete basis $\{|n\rangle\}$, so that $\sum_{n=0}^{\infty}|n\rangle\langle n| = \hat{I}$. The states that define this basis, with a well defined n quantum number (or number of excitations) are the so-called *Fock states*.

We can provide the matrix elements of the different operators in the basis of Fock states. For the number operator in the Fock basis, the matrix is diagonal:

$$\hat{N} = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 2 & 0 & \cdots \\ 0 & 0 & 0 & 3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{2.43}$$

Clearly, in this basis the different eigenstates are represented by the column vectors:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}; \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}; \quad |2\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix}, \tag{2.44}$$

and so on. We can also express the raising and lowering operators in this basis. For the raising operators, we can use Eq. (2.30) to find the matrix elements:

$$\langle n'|\hat{a}^\dagger|n\rangle = \sqrt{n+1}\delta_{n',n+1}. \tag{2.45}$$

This shows that $\hat{a}^\dagger$ only has non-zero matrix elements immediately below the diagonal, so that

$$\hat{a}^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ \sqrt{1} & 0 & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{2.46}$$

In contrast, the matrix elements of the lowering operator can be computed from Eq. (2.24), and yield

$$\langle n'|\hat{a}|n\rangle = \sqrt{n}\delta_{n',n-1}, \tag{2.47}$$

which means that only the elements right above the diagonal are active:

$$\hat{a} = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & \cdots \\ 0 & 0 & 0 & \sqrt{3} & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{2.48}$$

As an example of the applications of this method, we can look into the expectation values of the position and momentum operators on

the Fock states. Let us look at the expectation value of the position operator, Eq. (2.14) in *any* eigenstate:

$$\langle \hat{x} \rangle_n = \frac{1}{\sqrt{2}} \underbrace{\langle n|\hat{a}|n\rangle}_{=0} + \frac{1}{\sqrt{2}} \underbrace{\langle n|\hat{a}^\dagger|n\rangle}_{=0} = 0 \tag{2.49}$$

Note that the same result in real space would have involved at least an integral over eigenstates! Similarly, we can compute the expectation value of the momentum operator, Eq. (2.15):

$$\langle \hat{p} \rangle_n = \frac{i}{\sqrt{2}} \underbrace{\langle n|\hat{a}^\dagger|n\rangle}_{=0} - \frac{i}{\sqrt{2}} \underbrace{\langle n|\hat{a}|n\rangle}_{=0} = 0 \tag{2.50}$$

Furthermore, this method allows us to compute more complex results. Let us now look at the expectation valus of $\hat{x}^2$ and $\hat{p}^2$ in Fock states. For the space operator,

$$\begin{aligned} \langle \hat{x}^2 \rangle_n &= \frac{1}{2}(\langle n|\hat{a}^2|n\rangle + \langle n|\hat{a}\hat{a}^\dagger|n\rangle + \langle n|\hat{a}^\dagger\hat{a}|n\rangle + \langle n|\hat{(a}^\dagger)^2|n\rangle) \\ &= \frac{1}{2}(0 + \langle n|\hat{N}+1|n\rangle + \langle n|\hat{N}|n\rangle + 0) = n + \frac{1}{2} \end{aligned} \tag{2.51}$$

which implies

$$\Delta x_n^2 = \langle \hat{x}^2 \rangle_n - \langle \hat{x} \rangle_n^2 = n + \frac{1}{2}. \tag{2.52}$$

We can perform an analogous calculation for the momentum operator:

$$\langle \hat{p}^2 \rangle_n = \frac{-1}{2}(\langle n|\hat{a}^2|n\rangle - \langle n|\hat{a}\hat{a}^\dagger|n\rangle - \langle n|\hat{a}^\dagger\hat{a}|n\rangle + \langle n|\hat{(a}^\dagger)^2|n\rangle) = n + \frac{1}{2},$$

and find the squared deviation

$$\Delta p_n^2 = \langle \hat{p}^2 \rangle_n - \langle \hat{p} \rangle_n^2 = n + \frac{1}{2}. \tag{2.53}$$

With this, we see that the Heisenberg uncertainty principle for the harmonic oscillator becomes:

$$\Delta x_n \Delta p_n = n + \frac{1}{2} \geq \frac{1}{2}. \tag{2.54}$$

When going back from dimensionless to dimensionfull position and momentum states, the standard $\hbar$ factor appears and the uncertainty relation is just $\Delta X_n \Delta P_n \geq \frac{\hbar}{2}\left(n + \frac{1}{2}\right)$.

3 Coherent States

Coherent states play an important role in quantum optics, especially in the description of lasers. They were introduced in pioneering work by Roy J. Glauber[1], who was awarded the 2005 Nobel prize for his contribution to the quantum theory of optical coherence. Broadly speaking, coherent states are characterised by the fact that they involve an indefinite number of photons but have a very precisely defined phase. This is contrast to a state in Fock space, which has a fixed particle number but a random phase.

[1] Roy J. Glauber. Coherent and incoherent states of the radiation field. *Phys. Rev.*, 131:2766, 1963; and Roy J. Glauber. Nobel Lecture: One Hundred Years of Light Quanta. https://www.nobelprize.org/uploads/2018/06/glauber-lecture.pdf. Visited on 22 September 2021

Consider an harmonic oscillator whose minimum is not anymore at $x = 0$ but rather at a different value $x_0 \neq 0$. You may obtain this by applying a constant force f to the oscillator, so that

$$\begin{aligned}\hat{H}' &= \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2 x^2 - fx \\ &= \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2 (x - x_0)^2 + \frac{1}{2}m\omega^2 x_0^2\end{aligned}$$

with $x_0 = \frac{f}{m\omega^2}$. We can produce an eigenstate of such Hamiltonian by displacing the ground state wave function of the harmonic oscillator by x_0,

$$\psi_z(x) = \psi_0(x - x_0) = \frac{1}{\pi^{1/4}} e^{-\frac{(x-x_0)^2}{2}} . \tag{3.1}$$

This state is not anymore an eigenstate of the original hamiltonian $\hat{H}$. It is, however, an eigenstate of $\hat{a}$! To see this, let us apply the operator, in its real-space form, to the state $\psi_z(x)$:

$$\begin{aligned}\hat{a}\psi_z(x) &= \frac{1}{\sqrt{2}}\left[x + \frac{d}{dx}\right]\psi_0(x - x_0) = \frac{1}{\sqrt{2}}\left[x - (x - x_0)\right]\psi_0(x - x_0) \\ &= \underbrace{\frac{x_0}{\sqrt{2}}}_{z}\psi_z(x) = z\psi_z(x) .\end{aligned}$$

We thus find that $\psi_z(x) = \langle x|z\rangle$ is an eigenstate of $\hat{a}$ with eigenvalue z. For the time being, z is just a generic complex parameter, and we will explore its physical interpretation later on. From the previous equation, we derive the general relation

$$\hat{a}|z\rangle = z|z\rangle , \tag{3.2}$$

which defines a generic *coherent state* $|z\rangle$. It is important to note here that $\hat{a}$ is **not** an Hermitian operator and, as a consequence, z need not

Equivalently, we write

$$\langle z|\hat{a}^\dagger = \langle z|z^*$$

for the action of the raising operator on a bra $\langle z|$.

be real. In general, z is a complex quantity that can be described in terms of a modulus and a phase, $z = |z|e^{i\theta}$. Likewise, the eigenstates $|z\rangle$ belonging to different z's may not form a complete basis or be orthonormal.

As we have just seen above, when z is real, it can represent a translation of the ground-state wave function of the harmonic oscillator to a new central position, $x_0 = \sqrt{2}z$. One of the aims of this chapter is to identify what a complex z actually means physically.

3.1 Expansion on Fock states

As a first step towards a mathematical understanding of coherent states, let us try to expand the state $|z\rangle$ in terms of known harmonic oscillator (Fock) states, $|z\rangle = \sum_n c_n|n\rangle$. Applying the lowering operator, we find

$$\hat{a}|z\rangle = \sum_{n=0}^{\infty} c_n \hat{a}|n\rangle = \sum_{n=1}^{\infty} c_n\sqrt{n}|n-1\rangle = \sum_{m=0}^{\infty} c_{m+1}\sqrt{m+1}|m\rangle\,, \quad (3.3)$$

where the last equality arises from the replacement $n = m+1$. We can now write the right hand side of the eigenvalue equation, Eq. (3.2), to get

$$\hat{a}|z\rangle = z|z\rangle = \sum_{m=0}^{\infty} c_m z|m\rangle\,, \quad (3.4)$$

and, consequently,

$$c_{m+1} = \frac{z}{\sqrt{m+1}}c_m\,. \quad (3.5)$$

To solve this recursive equation, we set $c_{m=0} = 1$. This value may be changed by an overall normalization. With this, we find the recursion $c_1 = z$, $c_2 = \frac{z^2}{\sqrt{2}}$, $c_3 = \frac{z^3}{\sqrt{3\times 2}}$, and so on, so that $c_n = \frac{z^n}{\sqrt{n!}}$. The coherent state can thus be represented by the series

$$|z\rangle = \mathcal{N}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}|n\rangle\,, \quad (3.6)$$

up to a normalization constant $\mathcal{N}$. This constant can be determined by imposing a normalization condition, $\langle z|z\rangle = 1$, so that

$$\langle z|z\rangle = |\mathcal{N}|^2\sum_{n=0}^{\infty}\frac{|z|^{2n}}{n!} = |\mathcal{N}|^2 e^{|z|^2} = 1 \Rightarrow \mathcal{N} = e^{-\frac{|z|^2}{2}}\,. \quad (3.7)$$

With this, we are finally ready to write down a formal expression providing coherent states in terms of Fock states,

$$|z\rangle = e^{-\frac{1}{2}|z|^2}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}|n\rangle\,. \quad (3.8)$$

This expression is formal, and it involves in principle an infinite number of terms. The gaussian prefactor depends on the absolute value of z, $|z|$, but the argument of the series is a complex number to the power of n.

The representation of coherent states in Fock space is a very useful tool. It will allow us to prove a series of interesting results. We start

by writing down the expectation value of the number operator over a coherent state,

$$\langle \hat{n} \rangle_z = \langle z|\hat{N}|z\rangle = \langle z|\hat{a}^\dagger \hat{a}|z\rangle = z^* z = |z|^2 , \tag{3.9}$$

which is nothing but the squared modulus of z.

This equation indicates that in a coherent state the average particle number, $\langle \hat{n} \rangle_z$, plays a role that is similar to the mode number n of a single Fock state, $\langle \hat{n} \rangle_n = n$. In a Fock state, however, the variance of the particle number over a Fock state is null, $\langle \Delta \hat{n} \rangle_n = \sqrt{\langle \hat{n}^2 \rangle_n - \langle \hat{n} \rangle_n^2} = \sqrt{n^2 - n^2} = 0$. In contrast, the variance of $\hat{N}$ over a coherent state is finite and, in fact, equal to the expectation value of $\hat{n}$, $\Delta n = |z|^2$. To find this result, consider the following:

$$\langle \hat{n}^2 \rangle_z = \langle z|\hat{N}^2|z\rangle = \langle z|\hat{a}^\dagger \hat{a}\hat{a}^\dagger \hat{a}|z\rangle = z^* \langle z|\hat{a}\hat{a}^\dagger|z\rangle z . \tag{3.10}$$

At first sight, one might thing that to evaluate this last expression one would require to evaluate $a^\dagger|z\rangle$ or $\langle z|\hat{a}$. Neither of these operations, however, have been (or can be) defined. One can get away from this issue by employing the commutator relation between $\hat{a}$ and $\hat{a}^\dagger$, Eq. (2.16), which can be expressed as $\hat{a}\hat{a}^\dagger = 1 + \hat{a}^\dagger \hat{a}$. With this in mind, we find

$$\langle \hat{n}^2 \rangle_z = |z|^2 \Big(1 + \langle z|\hat{a}^\dagger \hat{a}|z\rangle\Big) = |z|^2 \Big(1 + |z|^2\Big) . \tag{3.11}$$

The variance is finally found

$$\Delta n_z = \sqrt{\langle \hat{n}^2 \rangle_z - \langle \hat{n} \rangle_z^2} = \sqrt{|z|^2(1 + |z|^2) - |z|^2} = |z|^2 . \tag{3.12}$$

In a sense, this equation represents the connection between a "particle" property (the mean coherent state number, $\langle \hat{n} \rangle_z$) and a "wave" property (the complex amplitude squared).

We can now ask ourselves what is the probability of finding m excitations on a given coherent state. To do this, we project $|z\rangle$ on a state $|m\rangle$ and find:

$$P_z(m) = |\langle z|m\rangle|^2 = e^{-|z|^2} \frac{|z|^{2m}}{m!} = e^{-\langle \hat{n} \rangle_z} \frac{\langle \hat{n} \rangle_z^m}{m!} . \tag{3.13}$$

This is nothing but a Poisson distribution with an average number of events $\langle \hat{n} \rangle_z$. Figure 3.1 shows the Poisson distribution as a function of m for different values of $\langle \hat{n} \rangle_z$. Classical particles obey the same Poisson formula when they are taken at random from a pool with an average, known rate $|z|^2$. Thus counting the modes m in a coherent state, they behave like randomly distributed classical particles, which is somehow surprising since coherent states are wave-like.

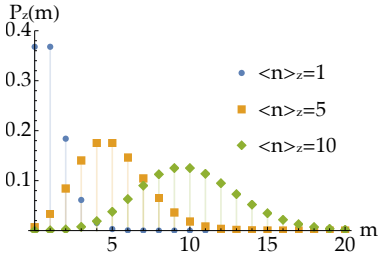

Figure 3.1: The Poisson distributions $P_z(m)$ as a function of m for different values of the mean value of modes, $\langle n \rangle_z = |z|^2$. As $\langle n \rangle_z$ increases, the peak of the distribution and its width shift to larger values.

3.1.1 *Orthogonality and completeness*

We have stressed that $\hat{a}$ is not Hermitian. A consequence of this fact is that its eigenvectors need not be orthonormal and may not form a complete basis. Coherent states are indeed non-orthonormal.

Consider, for instance, two different coherent states $|z\rangle$ and $|z'\rangle$ with $z \neq z'$. Let us compute their product:

$$\begin{aligned}\langle z|z'\rangle &= e^{-\frac{1}{2}|z|^2} e^{-\frac{1}{2}|z'|^2} \sum_{n,m=0}^{\infty} \frac{(z^*)^n}{\sqrt{n!}} \frac{z'^m}{\sqrt{m!}} \underbrace{\langle m|n\rangle}_{\delta_{m,n}} \\ &= e^{-\frac{1}{2}|z|^2} e^{-\frac{1}{2}|z'|^2} \sum_{n}^{\infty} \frac{(z^*)^n z'^n}{n!} = e^{-\frac{1}{2}|z|^2 - \frac{1}{2}|z'|^2 + z^* z'} \end{aligned} \tag{3.14}$$

We can rewrite the exponent to discriminate more transparently between the modulus and the phase of $\langle z|z'\rangle$:

$$\begin{aligned}-\frac{1}{2}|z|^2 - \frac{1}{2}|z'|^2 + z^* z' &= -\frac{1}{2}\left(|z|^2 + |z'|^2 - z^* z' - z^* z' - z z'^* + z z'^*\right) \\ &= -\frac{1}{2}\left|z - z'\right|^2 + \frac{1}{2}\left(z^* z' - z z'^*\right) .\end{aligned}$$

The first term in the last expression is real, and the second is purely imaginary[2]. With this, we can separate the product in the two components $\langle z|z'\rangle = |\langle z|z'\rangle| e^{i\theta}$, with a modulus $|\langle z|z'\rangle| = e^{-\frac{1}{2}|z-z'|^2}$ and a phase, $i\theta = \frac{1}{2}\left(z^* z' - z z'^*\right)$.

[2] To prove this, just see that the complex conjugate of this quantity, $(z^* z' - z z'^*)^* = z z'^* - z^* z' = -z^* z' + z z'^*$, is minus the quantity itself.

The first important point to make here is that for two distinct eigenvalues of $\hat{a}$, z and z', the product of their eigenvectors is not zero in general. In other words, they are not orthonormal. This was to be expected since $\hat{a}$ is not an Hermitian operator. Having said that, the modulus of the product, $|\langle z|z'\rangle|^2 = e^{-|z-z'|^2}$, decays exponentially with the distance between z and z'. In practical terms, if the two eigenvalues are far from each other, $|z - z'| \gg 1$, their eigenvectors may be considered orthogonal for all practical purposes.

Even though coherent states are not orthogonal, they do provide a complete basis in the sense that they fulfil a relation that is akin to completeness,

$$\frac{1}{\pi} \int \mathrm{d}^2 z |z\rangle\langle z| = \hat{I} . \tag{3.15}$$

To be more specific, the coherent states are *overcomplete*, which means that, as a consequence of their nonorthogonality, any coherent state can be expanded in terms of all the other coherent states.

To prove this "completeness" relation, we use $z = |z| e^{i\phi}$ and $\mathrm{d}^2 z = |z| d|z| d\phi$, so that

$$\begin{aligned}\frac{1}{\pi} \int \mathrm{d}^2 z |z\rangle\langle z| &= \frac{1}{\pi} \sum_{n,m} \int_0^{\infty} d|z| \int_0^{2\pi} d\phi e^{-|z|^2} |z| \frac{|z|^n}{\sqrt{n!}} \frac{|z|^m}{\sqrt{m!}} e^{i\phi(n-m)} |n\rangle\langle m| \\ &= \sum_n \underbrace{2 \int_0^{\infty} d|z| e^{-|z|^2} \frac{|z|^{2n+1}}{n!}}_{=1} |n\rangle\langle n| = \sum_n |n\rangle\langle n| = \hat{I} ,\end{aligned}$$

where we have used the fact that the angular integral equals $2\pi\delta_{n,m}$ and the gaussian integral equals unity.

To find the result of the gaussian integration, start with the simple integral $f(a) = \int_0^{\infty} \mathrm{d}x x e^{-ax^2} = \frac{1}{2a}$. Differentiating $f(a)$ n times yields, on the one hand, $\partial_a^{(n)} f(a) = (-1)^n \int_0^{\infty} \mathrm{d}x x^{2n+1} e^{-ax^2}$. On the other hand, the same derivative yields $\partial_a^{(n)} f(a) = \frac{1}{2}\partial_a^{(n)} a^{-1} = (-1)^n \frac{n!}{2a^{n+1}}$ and thus $\int_0^{\infty} \mathrm{d}x x^{2n+1} e^{-ax^2} = \frac{n!}{2a^{n+1}}$, which for $a = 1$ yields the result we need.

3.2 *Displacement operator*

We can rephrase the fundamental equation representing coherent states in Fock space, Eq. (3.8), to an equation involving only the

vacuum, $|0\rangle$. To do this, we use Eq. (2.34) and find:

$$|z\rangle = e^{-\frac{1}{2}|z|^2} \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}} \frac{\left(\hat{a}^\dagger\right)^n}{\sqrt{n!}} |0\rangle = e^{-\frac{1}{2}|z|^2} e^{z\hat{a}^\dagger} |0\rangle \,, \tag{3.16}$$

where we have used the standard definition of the exponential of an operator, $e^{\hat{A}} = \sum_{n=0}^{\infty} \frac{1}{n!} \hat{A}^n$. This expression provides a substantial formal simplification. Ideally, we would like to have an expression involving an exponential operator that is unitary. To this end, we note the following,

$$e^{-z^*\hat{a}}|0\rangle = \sum_n (-1)^n \frac{(z^*\hat{a})^n}{n!}|0\rangle = |0\rangle \,,$$

which arises from the expression $\hat{a}^n|0\rangle = 0$. We can use this identity in Eq. (3.16) to get the more symmetric expression

$$|z\rangle = e^{-\frac{1}{2}|z|^2} e^{z\hat{a}^\dagger} e^{-z^*\hat{a}} |0\rangle \,. \tag{3.17}$$

Taking $\hat{A} = z\hat{a}^\dagger$ and $\hat{B} = -z^*\hat{a}$, we find that

$$\left[\hat{A}, \hat{B}\right] = -\left[z\hat{a}^\dagger, z^*\hat{a}\right] = -|z|^2 \left[\hat{a}^\dagger, \hat{a}\right] = |z|^2 \,. \tag{3.18}$$

Thus $\hat{A}$ and $\hat{B}$ commute with their commutator, $\left[\hat{A}, \left[\hat{A}, \hat{B}\right]\right] = \left[\hat{B}, \left[\hat{A}, \hat{B}\right]\right] = 0$. In this specific case, we can exploit Glauber's relation[3], $e^{-\frac{1}{2}\left[\hat{A},\hat{B}\right]} e^{\hat{A}} e^{\hat{B}} = e^{\hat{A}+\hat{B}}$, to find:

$$|z\rangle = e^{-\frac{1}{2}|z|^2} e^{z\hat{a}^\dagger} e^{-z^*\hat{a}} |0\rangle = e^{z\hat{a}^\dagger - z^*\hat{a}} |0\rangle \,. \tag{3.19}$$

This inspires the definition of the **displacement operator**,

$$\mathcal{D}(z) = e^{z\hat{a}^\dagger - z^*\hat{a}} \,, \tag{3.20}$$

which, when operating on the vacuum, yields the coherent state, $|z\rangle = \mathcal{D}(z)|0\rangle$. As the exponential of an anti-unitary operator, the displacement operator is unitary[4]. This implies the relations

$$\mathcal{D}^\dagger(z) = \mathcal{D}(-z) \,, \tag{3.21}$$

$$\mathcal{D}(z)\mathcal{D}^\dagger(z) = \hat{I} \,. \tag{3.22}$$

Further, using Glauber's relation[5] on $\hat{A} = z\hat{a}^\dagger - z^*\hat{a}$ and $\hat{B} = z'\hat{a}^\dagger - z'^*\hat{a}$, one can also relate the displacement operator of the individual displacements, z and z', to the operator associated to the sum $z + z'$:

$$\begin{aligned} \mathcal{D}(z)\mathcal{D}(z') &= e^{z\hat{a}^\dagger - z^*\hat{a}} e^{z'\hat{a}^\dagger - z'^*\hat{a}} = e^{\frac{zz'^* - z^*z'}{2}} e^{z\hat{a}^\dagger - z^*\hat{a} + z'\hat{a}^\dagger - z'^*\hat{a}} \\ &= e^{\frac{zz'^* - z^*z'}{2}} \mathcal{D}(z + z'). \end{aligned} \tag{3.23}$$

As an example of the usefulness of the operator, we can re-derive the orthogonality relation of coherent states, Eq. (3.14), in a different but formally simpler way,

$$\begin{aligned} \langle z'|z\rangle &= \langle 0|\mathcal{D}^\dagger(z')\mathcal{D}(z)|0\rangle = \langle 0|\mathcal{D}(-z')\mathcal{D}(z)|0\rangle \\ &= e^{\frac{-z'z^* + z'^*z}{2}} \langle 0|\mathcal{D}(z - z')|0\rangle = e^{\frac{-z'z^* + z'^*z}{2}} e^{-\frac{|z'-z|}{2}} \,. \end{aligned}$$

The latest formula arises, from instance from the projection of Eq. (3.14) into an $n = 0$ state. We now proceed to provide a more physical insight of this operator.

[3] See Week 3 tutorials for a proof.

[4] An antiunitary operation is such that $\hat{A}^\dagger = -\hat{A}$, as fulfilled by $\hat{A} = z\hat{a}^\dagger - z^*\hat{a}$. The exponential of an antiunitary operator is $\hat{U} = e^{\hat{A}}$ and its Hermitian conjugate, $\hat{U}^\dagger = e^{\hat{A}^\dagger} = e^{-\hat{A}} = \hat{U}^{-1}$.

[5] Glauber's relation is fulfilled because $\left[\hat{A}, \hat{B}\right] = \left[z\hat{a}^\dagger - z^*\hat{a}, z'\hat{a}^\dagger - z'^*\hat{a}\right] = -zz'^*[\hat{a}^\dagger, \hat{a}] - z^*z'[\hat{a}, \hat{a}^\dagger] = zz'^* - z^*z'$ is a number, and hence $\left[\hat{A}, \hat{B}\right]$ commutes with $\hat{A}$ and $\hat{B}$.

3.3 *Physical interpretation of coherent states*

One can further manipulate the exponent of Eq. (3.20) by rewriting it in terms of the position and momentum operator, Eqs. (2.12) and (2.13),

$$\begin{aligned} z\hat{a}^\dagger - z^*\hat{a} &= \frac{1}{\sqrt{2}}\left[z\left(\hat{x}-i\hat{p}\right) - z^*\left(\hat{x}+i\hat{p}\right)\right] \\ &= \frac{1}{\sqrt{2}}\left[\left(z-z^*\right)\hat{x} - i\left(z+z^*\right)\hat{p}\right] \\ &= \sqrt{2}i\left[\mathrm{Im}(z)\hat{x} - \mathrm{Re}(z)\hat{p}\right] . \end{aligned}$$

Defining a new position variable, $x_0 = \sqrt{2}\,\mathrm{Re}(z)$, and a new momentum variable, $p_0 = \sqrt{2}\,\mathrm{Im}(z)$, the displacement operator is rewritten as:

$$\mathcal{D}(z) = e^{-i(x_0\hat{p} - p_0\hat{x})} . \tag{3.24}$$

As a unitary operator, we may be able to identify $\mathcal{D}$ with some sort of physical transformation between states. What transformation, exactly?

Let us start by considering the projection of an arbitrary operator into real space:

$$|\psi\rangle = \int \mathrm{d}x\, \psi(x)|x\rangle . \tag{3.25}$$

A translation of the previous state by an amount x_0 is achieved by an operator transformation $\hat{U}(x_0)$, such that

$$|\psi_{x_0}\rangle = \hat{U}(x_0)|\psi\rangle = \int \mathrm{d}x\, \psi(x)|x+x_0\rangle = \int \mathrm{d}x\, \psi(x-x_0)|x\rangle . \tag{3.26}$$

On the one hand, Taylor expanding the wave function around $\psi(x)$, we find

$$\psi(x-x_0) = \sum_n \frac{1}{n!}\partial_x^n\psi(x)\,(-x_0)^n . \tag{3.27}$$

On the other hand, applying the operator $\hat{U}(x_0) = e^{-\frac{ix_0\hat{p}}{\hbar}}$ to the original wave function, we see

$$\hat{U}(x_0)\psi(x) = \sum_n \frac{1}{n!}\frac{(-i\hat{p})^n}{\hbar^n}\psi(x)\,(x_0)^n . \tag{3.28}$$

The two expansions above can be reconciled if we identify the momentum operator with the derivative $i\hat{p} = \hbar\partial_x$ or $\hat{p} = -i\hbar\partial_x$. This indicates that the operator $\hat{U}$ is indeed implementing a translation of the state. In group theoretic language, we would say that momentum is the *generator* of the translation group. We have just proved that the term $\hat{U}(x_0) = e^{-\frac{ix_0\hat{p}}{\hbar}}$ in the displacement operator is a translation in real space by an amount x_0.

The displacement operator involves also a second term involving the exponential of the term $\frac{i\hat{x}p_0}{\hbar}$. The effect of this term can be found in complete analogy to the previous derivation but replacing all

positions by momenta, $x \to p$. In other words, $\hat{U}(p_0) = e^{\frac{i\hat{x}p_0}{\hbar}}$ provides a shift in momentum space, so that $\hat{U}(p_0)|p\rangle = |p + p_0\rangle$

This suggests that a coherent state generated as $|z\rangle = \mathcal{D}(z)|0\rangle$ can be interpreted as an harmonic oscillator ground state that has been shifted by an amount x_0 in the real axis and has been given a momentum p_0. In other words, the real part of z, which is related to x_0, corresponds to a shift in position space. The imaginary part of z, related to p_0, is a shift in momentum space. A general coherent state $z = \frac{x_0+ip_0}{\sqrt{2}}$ is a state that has shifts in both real and momentum space.

However, the displacement operator $\mathcal{D}(z)$ is not the product of $\hat{U}(x_0)\hat{U}(p_0)$. Instead, exploiting Glauber's formula with $\hat{A} = p_0\hat{x}$ and $\hat{B} = -x_0\hat{p}$, one finds

$$\mathcal{D}(z) = e^{-i(x_0\hat{p}-p_0\hat{x})} = e^{-\frac{1}{2}ix_0p_0}e^{ip_0\hat{x}}e^{-ix_0\hat{p}}, \tag{3.29}$$

which indicates that there is an additional phase involved in $\mathcal{D}$. This phase is typically irrelevant, but is necessary for situations where the mixing of states or dynamics play a role. Let us check that this operator provides the desired translation of oscillator ground states,

$$\begin{aligned}\psi_z(x) &= \mathcal{D}(z)\psi_0(x)e^{-\frac{1}{2}ix_0p_0}e^{ip_0\hat{x}}e^{-ix_0\hat{p}}\psi_0(x) = e^{-\frac{1}{2}ix_0p_0}e^{ip_0\hat{x}}\psi_0(x-x_0)\\ &= \frac{1}{\pi^{1/4}}e^{-i\frac{p_0}{2}(x_0-2x)}e^{-\frac{(x-x_0)^2}{2}}.\end{aligned} \tag{3.30}$$

Other than the irrelevant phase factor $e^{-\frac{1}{2}ix_0p_0}$ this is precisely the expression for a shifted and "kicked" gaussian wave packet.

3.4 *Minimum uncertainty states*

Some of the previous results are further confirmed by considering the expectation value of the position and the momentum operators over a coherent state. For the position, we find

$$\langle x\rangle_z = \langle z|\hat{x}|z\rangle = \langle z|\frac{\hat{a}^\dagger+\hat{a}}{\sqrt{2}}|z\rangle = \frac{z^*+z}{\sqrt{2}} = \sqrt{2}\mathrm{Re}(z) = x_0. \tag{3.31}$$

A completely analogous calculation for the momentum yields the expected result too,

$$\langle p\rangle_z = \langle z|\hat{p}|z\rangle = \langle z|i\frac{\hat{a}^\dagger-\hat{a}}{\sqrt{2}}|z\rangle = i\frac{z^*-z}{\sqrt{2}} = \sqrt{2}\mathrm{Im}(z) = p_0. \tag{3.32}$$

In other words, the expectation value of the position and the momentum of a coherent state are nothing but x_0 and p_0 and, hence, directly proportional to the real and imaginary parts of z.

We can now look for the dispersion of these observables. Let us start with the position operator and compute its square,

$$\langle x^2\rangle_z = \langle z|\hat{x}^2|z\rangle = \frac{1}{2}\langle z|\hat{a}^\dagger\hat{a}^\dagger + \hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger + \hat{a}\hat{a}|z\rangle.$$

The right-most term, $\langle z|\hat{a}\hat{a}|z\rangle$, is easy to compute, as it involves applying $\hat{a}$ twice on $|z\rangle$, thus yielding z^2. The left-most term,

$\langle z|\hat{a}^\dagger\hat{a}^\dagger|z\rangle$, can instead be calculated by applying $\hat{a}^\dagger$ twice to the bra, $\langle z|$, giving z^{*2}. Again, since we have not defined what the action of $\hat{a}^\dagger$ on the ket $|z\rangle$ does, the term $\langle z|\hat{a}\hat{a}^\dagger|z\rangle$ cannot be computed straight away. To find a result, we are required to put the lowering operator on the right using the commutation relation[6] just as we did in Eq. (3.11), $\langle z|\hat{a}\hat{a}^\dagger|z\rangle = \langle z|\hat{a}^\dagger\hat{a}|z\rangle + 1 = |z|^2 + 1$, and we use Eq. (3.9) to find $\langle z|\hat{a}^\dagger\hat{a}|z\rangle$. Collecting all the terms together, we find that

$$\langle x^2\rangle_z = \frac{1}{2}\left(z^{*2} + z^2 + 2|z|^2 + 1\right) = \frac{1}{2}\left((z+z^*)^2 + 1\right) = \langle x\rangle_z^2 + \frac{1}{2}\,.$$

The dispersion in position is then clearly $\Delta x^2 = \langle x^2\rangle_z - \langle x\rangle_z^2 = \frac{1}{2}$. We see that this is the same results that we obtained for the ground state $n = 0$ of the harmonic oscillator, Eq. (2.52). In a sense, this is not surprising because we have just interpreted coherent states as shifted, but shape-preserving, harmonic oscillators states. Having said that, it is somewhat surprising that the infinite superposition of states given by Eq. (3.8) provides a state with exactly the same properties as the $n = 0$ state. More importantly, this provides the *minimum* possible value.

A similar calculation can be performed for the square of the momentum operator:

$$\begin{aligned}\langle p^2\rangle_z &= \frac{1}{2}\langle z|-\hat{a}^\dagger\hat{a}^\dagger + \hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger - \hat{a}\hat{a}|z\rangle \\ &= \frac{1}{2}\left(-z^{*2} - z^2 + 2|z|^2 + 1\right) = \frac{1}{2}\left(-(z-z^*)^2 + 1\right) = \langle p\rangle_z^2 + \frac{1}{2}\,.\end{aligned}$$

We thus find that the square momentum operator also yields the minimum value of all the possible harmonic oscillator states, $\Delta p^2 = \frac{1}{2}$.

Combining these two results, we find that the coherent states saturate the uncertainty principle,

$$\Delta x \Delta p = \frac{1}{2}\,. \tag{3.33}$$

By this, we mean that there is no uncertainty relation as such, but rather an equal sign at precisely the minimum value allowed by the standard position-momentum Heisenberg principle. We have thus just proven that coherent states are *minimum uncertainty* states. Surprisingly, one can also prove the inverse of this statement - namely that the only states that saturate the position-momentum Heisenberg uncertainty principle are coherent states[7]. In this sense, by saturating the uncertainty relation, we say that coherent states are the closest possible to their classical counterparts. We shall now see that the dynamics of coherent states also follows closely the classical picture of particles oscillating on harmonic oscillator wells.

[6] For more complicated products of raising and lowering operators, the procedure of taking all lowering operators to the rights is called *normal ordering*. It is an essential tool in the second quantization philosophy of quantum many-body theory.

[7] See Complement III of C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum Mechanics*, volume 1. Wiley-VCH, 1977

3.5 *Dynamics of coherent states*

We may wonder what the dynamics of this state is. In other words, let us assume we can generate the state $\psi_z(x)$ at $t = 0$. Clearly, this

is not an eigenstate of the Hamiltonian and hence it should evolve and change as time t evolves. How does this evolution actually take place?

The dynamics can be easily solved by using the expansion in Fock space, $|z\rangle = e^{-\frac{1}{2}|z|^2}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}|n\rangle$, and the solution $\lambda_\alpha(t) = e^{-i\frac{E_\alpha t}{\hbar}}$ for the expansion coefficients of any time-dependent state, as we saw right at the end of Chapter 1, Eq. (1.30). With this, one finds that

$$|z(t)\rangle = e^{-\frac{1}{2}|z|^2}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}e^{-i\epsilon_n t}|n\rangle,$$

which can be further simplified by taking $\epsilon_n = n + \frac{1}{2}$,

$$|z(t)\rangle = e^{-\frac{1}{2}|z|^2}e^{-i\frac{1}{2}t}\sum_{n=0}^{\infty}\frac{\left(ze^{-it}\right)^n}{\sqrt{n!}}|n\rangle = e^{-i\frac{1}{2}t}\left|ze^{-it}\right\rangle. \tag{3.34}$$

We arrive to the simplification that coherent states evolve in time following a simple change in phase dictated by e^{-it}. In other words, starting at $t = 0$ with a coherent state $|z\rangle$, the corresponding time-evolved state is simply proportional to $|ze^{-it}\rangle$. This is remarkably simple and, in fact, is one of the few cases where the dynamics can be solve analytically!

Keep in mind that we set $\hbar\omega = 1$ in harmonic oscillator units, so the time exponents everywhere would read $i\omega T$ in standard units as opposed to it. In other words, the dimensionless unit of time in the harmonic oscillator is $t = \omega T/\hbar$.

Mathematically, there is a relevant analogy for the time evolution, as shown in Fig. 3.2. Consider the phase space of a given coherent state (eg the relation between its position and momentum). A given starting coherent state $|z\rangle$ is a point in this diagram with momentum p_0 and position x_0 at $t = 0$, shown with a red pointin Fig. 3.2. The corresponding time evolution does not change the modulus of the coherent state variable z. To see this, consider the average particle number of the time-evolved state, $\langle z(t)|\hat{n}|z(t)\rangle = |z(t)|^2 = z^*e^{+it}\times ze^{-it} = |z|^2$, which is clearly constant under the time evolution The relation $|z(t)\rangle = e^{-i\frac{1}{2}t}\left|ze^{-it}\right\rangle$ indicates that the corresponding time-evolved state follows a circular trajectories in phase space, with a fixed radius $|z|$. This uniform circular motion is shown in Fig. 3.2, and we stress that the period of rotation is $\Delta t = 2\pi$ which, in the original units, is nothing but the period of the oscillator itself, $\Delta T = \frac{2\pi}{\omega}$.

This could also be derived from the conservation of energy.

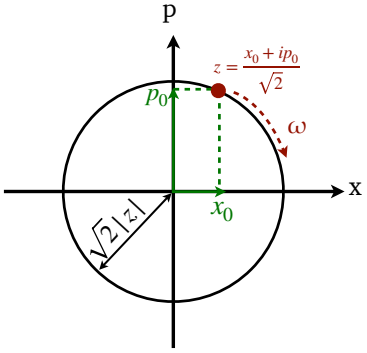

Figure 3.2: The dynamics in phase space of a coherent state which is initially at $z = \frac{x_0+ip_0}{2}$ is simply given by a circular motion of constant radius $\sqrt{2}z$ with the natural frequency of the oscillator ω.

Perhaps more importantly, the result of Eq. (3.34) allows us to straightforwardly get time-evolved coherent states from their "static" counterparts by just replacing $z \to z(t) = e^{-it}z$. With this trick, it is straightforward to compute the time evolution of the position and the momentum eigenstates for the time-dependent coherent states. We start with the average position of a coherent state which, exploiting Eq. (3.31), yields

$$\begin{aligned}\langle x(t)\rangle_z &= \langle z(t)|\hat{x}|z(t)\rangle = \sqrt{2}\mathrm{Re}(z(t)) = \sqrt{2}\mathrm{Re}(e^{-it}z)\\ &= \mathrm{Re}(e^{-it}(x_0 + ip_0)) = x_0\cos t + p_0\sin t\,.\end{aligned}$$

In a similar fashion, one can compute the average momentum using Eq. (3.32), and we get

$$\begin{aligned}\langle p(t)\rangle_z &= \langle z(t)|\hat{p}|z(t)\rangle = \sqrt{2}\mathrm{Im}(z(t)) = \sqrt{2}\mathrm{Im}(e^{-it}z)\\ &= \mathrm{Im}(e^{-it}(x_0 + ip_0)) = p_0\cos t - x_0\sin t\,.\end{aligned}$$

This clearly indicates that coherent states oscillate in time in a "coherent" manner, keeping an oscillation structure that is the natural frequency of the oscillator. Importantly, these equations are exactly the same that one would get for a classical particle with original position x_0 and momentum p_0. This again suggests that coherent states are, in a sense, the closest quantum states to the classical dynamics of the harmonic oscillator.

The dynamics of coherent states is far from trivial. If we were evolving in time a given Fock state, $\langle x|n\rangle$, it would be easy to show that the probability density of the state would not change, $|\langle x|n(t)\rangle|^2 = |\langle x|n(t=0)\rangle|^2$. In contrast, a coherent state wave function evolves in time according to $\langle x|z(t)\rangle$. Following the trick discussed above, it is clear that $\psi_{z(t)} = \langle x|z(t)\rangle$ and exploiting Eq. (3.30) with the obvious replacements $x_0 \to \langle x(t)\rangle_z$ and $p_0 \to \langle p(t)\rangle_z$, we find the time-evolved wave function

$$\psi_{z(t)}(x) = \frac{1}{\pi^{1/4}} e^{-i\frac{\langle p(t)\rangle_z}{2}(\langle x(t)\rangle_z - 2x)} e^{-\frac{(x-\langle x(t)\rangle_z)^2}{2}}. \tag{3.35}$$

The corresponding probability distribution is

$$\left|\psi_{z(t)}(x)\right|^2 = \frac{1}{\pi^{1/2}} e^{-(x-\langle x(t)\rangle_z)^2}. \tag{3.36}$$

This probability distribution does change with time! It is a gaussian state of fixed shape, whose peak follows the oscillatory trajectory of $\langle x(t)\rangle_z$. Snapshots for the time evolution of a coherent state without initial momentum, $p_0 = 0$, but initial position $x_0 = 1$, are shown in Fig. 3.3. These snapshots can be generated by evolving an harmonic oscillator ground state over time as well. An example of how to implement this time evolution is given in the Jupyter Notebook Dynamics_harmonic_oscillator.ipynb available in the master's github repository shorturl.at/fEXZ5.

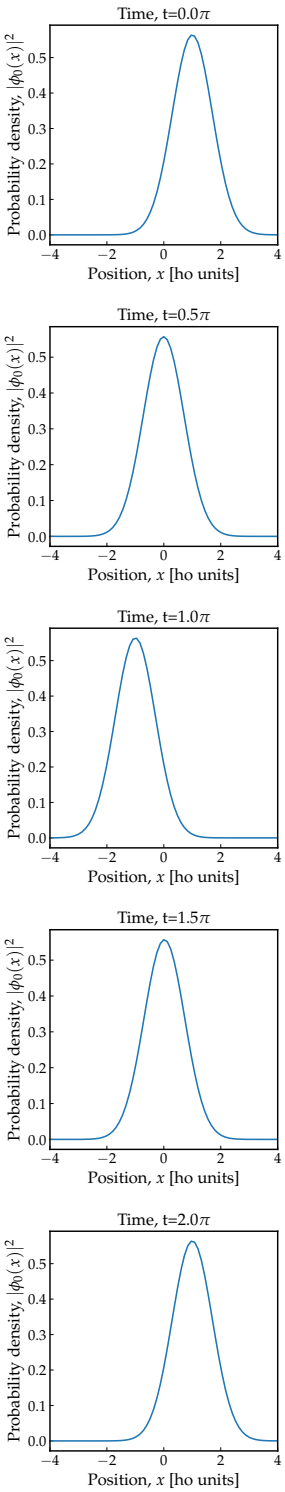

Figure 3.3: Snapshots of the time evolution of the probability density of a coherent state with $x_0 = 1$ and $p_0 = 0$. The time evolution is periodic on a timescale of 2π in harmonic oscillator units.

4 Second quantization

In several instances of physics, one is interested in treating systems that are made of not just one, but several particles. The quantum mechanical many-body problem requires a series of tools that are unique and have no counterpart to the one-body treatments that we have discussed so far. In this chapter, we briefly discuss the formalism of second quantization, that allows us to treat many-body systems from an abstract and powerful perspective.

4.1 Identical particles

In non-relativistic quantum mechanics, each particle is characterized by a set of parameters. These parameters are intrinsic to the particle - and they are the same for the same particles species. Think, for instance, of the mass; the charge; or the spin of an electron. All electrons have the same mass, charge and spin and, from these properties alone, we cannot distinguish two different electrons. Based on these properties, we can divide particles into classes depending on their intrinsic parameters, and particles belonging to the same class are *indistinguishable*.

The way this works is different in quantum than in classical mechanics. In classical mechanics, we can always distinguish particles depending on some of their dynamical properties, like their positions or momenta. The position of two identical quantum mechanical particles, in contrast, cannot be distinguished. To understand exactly how this comes about, let us consider what happens with the many-body wave function when particles are swapped around.

4.1.1 Permutations and many-body wave functions

Let us consider the wave function of $N-$particles in quantum mechanics. We characterise each one of these particles, $i = 1, \cdots, N$, by a complete set of single-particle variables, that can either be dynamical (eg momentum $\mathbf{p}$, spin projection m_s) or intrinsic (eg mass m, spin s). These variables are succinctly summarised for simplicity into a single, but possibly high-dimensional, variable $\mathbf{x}_i$,

$$\mathbf{x}_i = \{\mathbf{p}_i, \cdots; m_i, s_i, q_i, \cdots\} \ . \tag{4.1}$$

Since all the characteristics of the identical particles i are the same, the hamiltonian $\hat{H}(\mathbf{x}_1, \cdots, \mathbf{x}_N)$ governing the dynamics of these particles

should be completely symmetric with respect to any exchange $i \leftrightarrow j$. Mathematically, we can implement such exchanges by means of transposition or permutation operators, $\hat{P}_{ij}$.

One may wonder what happens to the wave function, $\Psi(\mathbf{x}_1, \cdots, \mathbf{x}_N)$, under similar exchanges. Let us first look at how one can build such wave functions. We assume we start with a given single particle basis. For discrete systems, this will provide a series of states $|\alpha\rangle$, $|\beta\rangle$, $|\gamma\rangle$, etc. Let us discuss the simplest two-body case. A first guess for a many-body state may be obtained from a tensor product:

For a continuum spectrum like that generated by momenta $\mathbf{p}$, the same idea corresponds to different values of momenta $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, ...

$$|\alpha, \beta) = |\alpha\rangle_1 \otimes |\beta\rangle_2 \,. \tag{4.2}$$

We use the notation $|\cdot)$ for these naive many-body states. A tensor product $\otimes$ essentially indicates that these two states live in two different, independent single-particle Hilbert spaces. More details about tensor products can be obtained form standard quantum mechanics references[1]. The subscripts are meant to represent the particle tag i.A generic two-body state must be a linear combination of such states,

[1] A. Messiah. *Quantum Mechanics*. Dover Press, 1999; and C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum Mechanics*, volume 1. Wiley-VCH, 1977

$$|\Psi\rangle = \sum_{\alpha,\beta} C_{\alpha\beta} |\alpha, \beta). \tag{4.3}$$

Consider the simplest possible case, in which the single-particle basis is constructed in real space and $\mathbf{x}_i = \mathbf{r}_i$. If we look at the spatial representation for the two-body wavefunction, we find

$$\begin{aligned}\Psi(\mathbf{r}_1, \mathbf{r}_2) &= \Big(\langle \mathbf{r}_1| \otimes |\langle \mathbf{r}_2|\Big)|\Psi\rangle = \sum_{\alpha,\beta} C_{\alpha\beta} \langle \mathbf{r}_1|\alpha\rangle_1 \langle \mathbf{r}_2|\beta\rangle_2 \\ &= \sum_{\alpha,\beta} C_{\alpha\beta} \phi_\alpha(\mathbf{r}_1) \phi_\beta(\mathbf{r}_2) \,.\end{aligned}$$

The interpretation of this two-body wave function is clear. The particle at position $\mathbf{r}_1$ lives on state α whereas the particle at position $\mathbf{r}_2$ sits on state β. We add up over all the possible available combinations of α and β.

In addition to wavefunctions, a quantum mechanical formulation requires operators. Any operator $\hat{A}$ acting on single particles is to be understood as a direct sum in which $\hat{A}$ acts on the subspace of every individual particle.

For instance, for the two-body case, this means that

$$\hat{A} = \hat{A}_1 \otimes I_2 + I_1 \otimes \hat{A}_2 \,.$$

The permutation[2] operator $\hat{P}_{12}$ interchanges **all** variables belonging to 1 and 2 in a wave function. For instance, acting with $\hat{P}_{12}$ on the tensor product state of Eq. (4.2), we find

$$\hat{P}_{12}|\alpha, \beta) = \hat{P}_{12}|\alpha\rangle_1 \otimes |\beta\rangle_2 = |\alpha\rangle_2 \otimes |\beta\rangle_1 = |\beta\rangle_1 \otimes |\alpha\rangle_2 = |\beta, \alpha) \,. \tag{4.4}$$

Similar exchange operators exist for *parts* of the wave function. For instance, the operator that exchanges **only** the positions of particles 1 and 2 is the Majorana operator, $\hat{P}^r_{12}$. The operators that permutes particle spin, $\hat{P}^s_{12}$, is typically called the Bartlett operator. For two particles with spin described in real space, the total permutation operator is the product of the two terms, $\hat{P}_{12} = \hat{P}^r_{12}\hat{P}^s_{12}$. In

[2] Some references call this the transposition operator or the Heisenberg operator. Here, we use the word *transposition* to define the exchange of only 2 particles or variables, whereas a *permutation* for more than 2 variables can involve more than one transposition. For $N = 3$, for instance, the cyclic permutation $\hat{P}_{123}|\alpha, \beta, \gamma) = |\gamma, \alpha, \beta)$ is a single permutation, but is the result of two transpositions $\hat{P}_{123} = \hat{P}_{12}\hat{P}_{13}$.

mathematics, the permutations of N objects form a group called the *symmetric group* S_N, with $N!$ elements. The properties of this group are well understood and provide a mathematical framework for many-body theory.

Permutation operators fulfil a series of useful properties. They are clearly idempotent, so that $\hat{P}_{ij}\hat{P}_{ij} = \hat{P}_{ij}^2 = \hat{I}$. As such, the permutation operator is its own inverse, $\hat{P}_{ij}^{-1} = \hat{P}_{ij}$. Permutation operators are also Hermitian and, hence, unitary. One may wonder what the eigenvalues, η, of this operator are. In other words, what are the η values such that

$$\hat{P}_{12}|\alpha,\beta) = \eta|\alpha,\beta)\,? \tag{4.5}$$

To find these, consider applying the operator twice,

$$\hat{P}_{12}^2|\alpha,\beta) = |\alpha,\beta) = \eta^2|\alpha,\beta)\,, \tag{4.6}$$

which clearly shows that the permutation operator can only have two possible eigenvalues, $\eta = +1$ or $\eta = -1$. The eigenfunctions of $\hat{P}_{12}$ are thus divided into two classes: symmetric eigenfunctions, with eigenvalues $\eta = +1$, and antisymmetric eigenfunctions, with $\eta = -1$.

Note that this result is generic and is not restricted to the transposition case. The result of any $N-$body permutation has a well-defined sign $\eta = \pm 1$.

For two particles, an arbitrary ket $|\alpha,\beta)$ can always be presented as a superposition of a symmetric and an antisymmetric function,

$$|\alpha,\beta) = \frac{1}{2}\left[|\alpha,\beta) + |\beta,\alpha)\right] + \frac{1}{2}\left[|\alpha,\beta) - |\beta,\alpha)\right]\,, \tag{4.7}$$

since

$$\hat{P}_{12}\left[|\alpha,\beta) \pm |\beta,\alpha)\right] = \pm\left[|\alpha,\beta) \pm |\beta,\alpha)\right]\,. \tag{4.8}$$

The situation is more complicated for $N > 2$ particles because, among other things, permutations can be made fo products of transpositions. In fact, any permutation $\hat{P}$ can be written as a product of r transpositions $i \leftrightarrow j$. The parity of a permutation is then $(-1)^r = \pm 1$, which we shall denote as $(-1)^P$. Consider the three-particle tensor product state $|\alpha,\beta,\gamma)$. There are 3! different permutations of 3 objects, so one defines six different permutation operators:

$$\hat{I}|\alpha,\beta,\gamma) = |\alpha,\beta,\gamma); \quad \hat{P}_{12}|\alpha,\beta,\gamma) = |\beta,\alpha,\gamma); \quad \hat{P}_{13}|\alpha,\beta,\gamma) = |\gamma,\beta,\alpha);$$
$$\hat{P}_{23}|\alpha,\beta,\gamma) = |\alpha,\gamma,\beta); \quad \hat{P}_{123}|\alpha,\beta,\gamma) = |\gamma,\alpha,\beta); \quad \hat{P}_{123}^2|\alpha,\beta,\gamma) = |\beta,\gamma,\alpha),$$

where $\hat{P}_{123}$ is a cyclic permutation. Some of these permutations are not pure transpositions, like the cyclic permutation $\hat{P}_{123} = \hat{P}_{12}\hat{P}_{13}$. This permutation has $r = 2$ and hence a parity of $(-1)^2 = +1$.

The difficulty for many-body physics arises from the fact that permutations affecting the same particles do not commute. For instance, consider the two products of permutations below

$$\hat{P}_{13}\hat{P}_{12}|\alpha,\beta,\gamma) = \hat{P}_{13}|\beta,\alpha,\gamma) = |\gamma,\alpha,\beta);$$
$$\hat{P}_{12}\hat{P}_{13}|\alpha,\beta,\gamma) = \hat{P}_{12}|\gamma,\beta,\alpha) = |\beta,\gamma,\alpha);$$

leading to two different states. This implies that we cannot build a complete common set of eigenfunctions of $\hat{H}$ and $\hat{P}$, which would of course provide several difficulties.

To find a way out, let us look at the states that can be build by applying permutation operators. Some of these are particularly simple. Consider for instance the fully-symmetric 3−body state:

$$|\alpha,\beta,\gamma\} = \frac{1}{\sqrt{3!}}\Big[|\alpha,\beta,\gamma) + |\beta,\alpha,\gamma) + |\gamma,\beta,\alpha) + |\alpha,\gamma,\beta) + |\gamma,\alpha,\beta) + |\beta,\gamma,\alpha)\Big] \tag{4.9}$$

This state is invariant under any permutation P_{ij}. In other words, the sign of all pertutations is always $\eta = +1$. In contrast, the antisymmetric state

$$|\alpha,\beta,\gamma\rangle = \frac{1}{\sqrt{3!}}\Big[|\alpha,\beta,\gamma) - |\beta,\alpha,\gamma) - |\gamma,\beta,\alpha) - |\alpha,\gamma,\beta) + |\gamma,\alpha,\beta) + |\beta,\gamma,\alpha)\Big] \tag{4.10}$$

changes sign under the individual permutations $\hat{P}_{ij}$ (but not under the cyclic permutations[3]). Note that we introduce here the notation $|\cdot\}$ for fully symmetric states and $|\cdot\rangle$ for fully antisymmetric states.

[3] For three particles, cyclic permutations are the result of applying two permutation operators successively $\hat{P}_{123} = \hat{P}_{12}\hat{P}_{13}$ and $\hat{P}^2_{123} = \hat{P}_{13}\hat{P}_{12}$.

Whereas fully symmetric and fully antrisymmetric states are appealing, they are not the only types of states that can be built from permutations. For instance, the state

$$\frac{1}{\sqrt{2!}}\Big[0 + 0 + |\gamma,\beta,\alpha) - |\alpha,\gamma,\beta) + |\gamma,\alpha,\beta) - |\beta,\gamma,\alpha)\Big]$$

is anti-symmetric with respect to the $\hat{P}_{12}$ permutation, but it is not an eigenstate of some of the other permutations. How can we work with such states, that cannot be in the same basis as the Hamiltonian?

4.1.2 Symmetrization postulate

The way out of this difficulty is found by postulating that not all the allowed states are physical. In principle, many-body wave functions could be formed by the different combinations of possible symmetries under permutations. However, by carefully analyzing the experimentally observed spectra of allowed states in many-body systems (like atoms, molecules or nuclei), one reaches the conclusion that, of all the possible symmetries, only two types of wave functions are allowed. This observation can be elevated to the postulate level, because no other postulate has been applied to the many-body case yet.

Symmetrization postulate

The states of a system of identical particles must be either totally symmetric or totally antisymmetric under the exchange of **any** *pair. Mixed symmetry states do not exist.*

How does this make the process easier? Consider the fully symmetric three-body state $|\alpha,\beta,\gamma\}$ in Eq. (4.9). This state is an eigenfunction of all the pair permutation (or transposition) operators,

$$\hat{P}_{ij}|\alpha,\beta,\gamma\} = +|\alpha,\beta,\gamma\}\,.$$

The state is also invariant under two transpositions and hence

$$\hat{P}_{12}\hat{P}_{13}|\alpha,\beta,\gamma\} = \hat{P}_{13}\hat{P}_{12}|\alpha,\beta,\gamma\} = +|\alpha,\beta,\gamma\}\,.$$

In other words, for this specific state, the two permutations do commute. The same result holds for the fully antisymmetric state, in which each permutation brings in a $-$ sign, but two permutations do not change the overall state. For the fully symmetric and the fully antisymmetric state, then, permutations do commute among themselves. Moreover, since the Hamiltonian of indistinguishable particles must also commute with the permutation operators, we can build a common basis for the many-body problem.

Moreover, quantum field theory provides an additional useful result that allows us to connect the symmetry or antrisymmetry of the wave function to particle spins.

Spin-statistics theorem

- *Particles whose spin is an integer multiple of $\hbar$ have only symmetric states.*
- *Particles whose spin is a half-integer multiple of $\hbar$ have only antisymmetric states.*

We name the particles of the first case *bosons* and at finite temperature these follow Bose-Einstein statistics. In contrast, particles with antrisymmetric wave functions are called *fermions* and follow Fermi-Dirac statistics.

4.1.3 Symmetrization of states: plane waves

Let us consider a system of $N = 2$ particles. If we consider the non-interacting hamiltonian, $\hat{H} = \frac{\hat{p}_1^2}{2m} + \frac{\hat{p}_2^2}{2m} = \sum_{i=1}^{2} \hat{h}_i$, we know that the single-particle basis can be diagonalized in terms of well-defined momentum states,

$$\hat{p}\,|p\rangle = p\,|p\rangle\,, \quad \hat{h}\,|p\rangle = \frac{p^2}{2m}\,|p\rangle\,. \tag{4.11}$$

Note that this single-particle solution is independent of the particle label $i = 1, 2$. A well-defined momentum state expressed in real space is nothing but a plane wave

$$\langle x|p\rangle = \mathcal{N} e^{ipx}\,. \tag{4.12}$$

normalization notwithstanding. The corresponding antisymmetrized (fermionic) state of two particles reads

$$\psi_{p_1,p_2}(x_1,x_2) = [\langle x_1|\otimes\langle x_1|]\,|p_1,p_2\} = \frac{\mathcal{N}^2}{\sqrt{2}}\left[e^{ip_1x_1}e^{ip_2x_2} - e^{ip_2x_1}e^{ip_1x_2}\right]. \tag{4.13}$$

Introducing the center-of-mass and relative coordinates $R = \frac{x_1+x_2}{2}$ and $r = x_1 - x_2$, this reduces to

$$\psi_{p_1,p_2}(x_1,x_2) = \frac{\mathcal{N}^2}{\sqrt{2}}e^{i(p_1+p_2)R}\left[e^{i\frac{(p_1-p_2)x}{2}} - e^{-i\frac{(p_1-p_2)x}{2}}\right].$$

Clearly, the center-of-mass term becomes a global phase, whereas the relative position term depends on the relative momentum $q = \frac{p_1-p_2}{2}$. Squaring this wavefunction to find a probability, we obtain

$$|\psi_{p_1,p_2}(x_1,x_2)|^2 = \frac{\mathcal{N}^4}{2}\left|e^{iqx} - e^{-iqx}\right|^2 = 2\mathcal{N}^4\,|\sin qx|^2\,. \tag{4.14}$$

The corresponding probability, ignoring the irrelevant normalization prefactor, is shown in the top panel of Fig. 4.1. The area where $r \approx 0$ and $q \approx 0$ displays a vanishingly small probability. In other words, this suggests that for such plane-wave eigenstates, two fermions cannot have the same momentum (or the same position, for that matter) if there is no other physical quantity to differentiate them.

One can also work out the same expression for bosons. In Eq. (4.13), the relative sign should be changed to a plus. The probability subsequently becomes

$$|\psi_{p_1,p_2}(x_1,x_2)|^2 = \frac{\mathcal{N}^4}{2}\left|e^{iqx} + e^{-iqx}\right|^2 = 2\mathcal{N}^4\,|\cos qx|^2\,. \tag{4.15}$$

This probability is shown in the bottom panel of Fig. 4.1. Importantly, the bosonic and fermionic probabilities are "out" of phase, in the sense that where one is large, the other is small. In particular, two bosons can sit at the same position ($r = 0$) with an arbitrarily large probability.

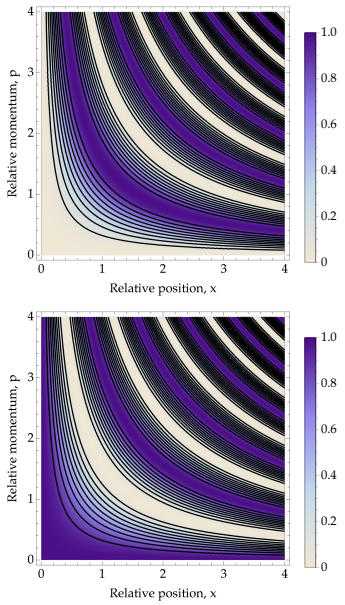

Figure 4.1: Top panel: probability (in arbitrary units) for finding a fermionic two-body plane wave wavefunction with relative position r and relative momentum q. In the region $r \approx 0$ and $q \approx 0$, the probability is small. Bottom panel: the same for a bosonic wavefunction.

4.1.4 *Symmetrization of states: harmonic oscillator*

Consider now a system of $N = 2$ particles with 2 distinct discrete single-particle states α and β. Classically, we could form 4 different equiprobable states: $(\alpha\alpha), (\alpha\beta), (\beta\alpha), (\beta\beta)$. In contrast, if the two particles are bosonic, we have 3 different possible states:

$$|\alpha,\alpha\} = |\alpha,\alpha\rangle; \quad |\beta,\beta\} = |\beta,\beta\rangle;$$
$$|\alpha,\beta\} = \frac{1}{\sqrt{2}}\left[|\alpha,\beta) + |\beta,\alpha)\right]. \tag{4.16}$$

The first two states indicate that two bosons may reside on the same state, α or β, or may reside in a symmetric combination of the two states. Fermions, in contrast, can only sit in an antisymmetric combination of α and β,

$$|\alpha,\beta\rangle = \frac{1}{\sqrt{2}}\left[|\alpha,\beta) - |\beta,\alpha)\right]. \tag{4.17}$$

A visual summary of the differences between these three types of states is provided in Fig. 4.2. Here, we show three different combinations of the $n = 0$ (α state) and $n = 1$ (β state) of the harmonic oscillator. The top panel shows the simple product state $\phi_0(x_1)\phi_1(x_2)$, which is not a physical state. Here, x_1 is the position of the first particle and x_2, of the second.

The middle panel shows instead the symmetric combination $\frac{1}{\sqrt{2}}\left(\phi_0(x_1)\phi_1(x_2) + \phi_0(x_2)\phi_1(x_1)\right)$ and the bottom panel shows the antisymmetric combination $\frac{1}{\sqrt{2}}\left(\phi_0(x_1)\phi_1(x_2) - \phi_0(x_2)\phi_1(x_1)\right)$. These are very different figures in the $x_1 - x_2$ plane. In particular, there is a clear line of zero probability in the fermionic case for $x_1 = x_2$. In contrast, for the bosonic case, the probability is largest when $x_1 = x_2$! All in all, this illustrates the differences between fermionic and bosonic systems.

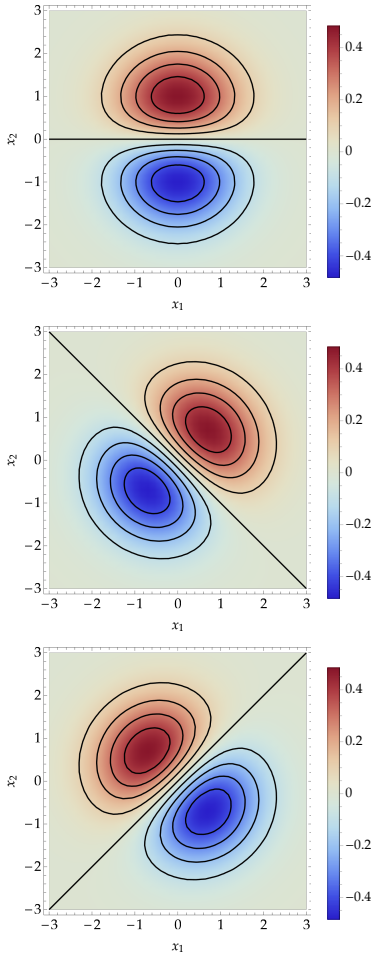

Figure 4.2: Examples of two-body wave functions formed by combining the $n = 0$ and $n = 1$ harmonic oscillator single-particle states. x_1 (x_2) is the position of particle 1 (2). The top panel is an unphysical product state, whereas the middle and bottom panels represent the symmetric (or bosonic) and antisymmetric (or fermionic) combinations. Note how different these look like along the $x_1 = x_2$ line.

4.1.5 *Symmetrization of states: generic states*

In a typical example for a many-body system, one would solve a single-particle problem to generate an orthonormal single-particle basis with states $|\alpha\rangle$, $|\beta\rangle$, $|\gamma\rangle$ and so on up until, say, $|\xi\rangle$. Note that, for bosons, some of these states may be repeated. It is then very easy to generate an arbitrary many-body product state

$$|\alpha, \beta, \gamma, \cdots, \xi) = |\alpha\rangle_1 \otimes |\beta\rangle_2 \otimes |\gamma\rangle_3 \otimes \cdots \otimes |\xi\rangle_N.$$

We can then generate fully symmetric or antisymmetric states from this arbitrary combination. Looking at the structure of the fully symmetric states of Eqs. (4.16) and (4.9), we see that

$$|\alpha, \beta, \gamma, \cdots, \xi\} = \frac{1}{\sqrt{N!}} \sum_P \hat{P}|\alpha, \beta, \gamma, \cdots, \xi), \tag{4.18}$$

where the sum is to be understood over all possible permutations $\hat{P}$, whether they are transpositions or their product. Similarly, a fully antisymmetric state may be built from the sum

$$|\alpha, \beta, \gamma, \cdots, \xi\rangle = \frac{1}{\sqrt{N!}} \sum_P (-1)^P \hat{P}|\alpha, \beta, \gamma, \cdots, \xi). \tag{4.19}$$

One can prove that if the initial single-particle basis was orthonormal, these symmetrized or antisymmetrized states are also orthonormal in the N-body Hilbert space $\mathcal{H}(N)$. They span the full space and fulfil associated completeness relations that we do not explore here in the interest of brevity, but can be found in other references[4].

We can project these wave functions into real space to see how they look like. We consider antisymmetric states and use $\alpha = 1$ and $\beta = 2$ to find

$$\begin{aligned}\Psi_A(\mathbf{r}_1, \mathbf{r}_2) &= \Big(\langle \mathbf{r}_1| \otimes \langle \mathbf{r}_2|\Big)|1, 2\rangle = \frac{1}{\sqrt{2}}\left[\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2) - \phi_1(\mathbf{r}_2)\phi_2(\mathbf{r}_1)\right] \\ &= \frac{1}{\sqrt{2}}\begin{vmatrix}\phi_1(\mathbf{r}_1) & \phi_1(\mathbf{r}_2) \\ \phi_2(\mathbf{r}_1) & \phi_2(\mathbf{r}_2)\end{vmatrix}.\end{aligned} \tag{4.20}$$

[4] V. Zelevinsky. *Quantum Physics*, volume 2. Wiley-VCH, 2011; and E.K.U. Gross, E. Runge, and O. Heinonen. *Many-Particle Theory,*. Taylor & Francis, 1991

The last line indicates that the two-body antisymmetric wave-function can be written as the determinant of a 2×2 matrix. In this matrix, the rows are selected by the state number and the columns, by the particle position. The generalization of this structure to the $N-$body case leads to a so-called *Slater determinant*,

$$\Psi_A(\mathbf{r}_1, \cdots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \det\{\phi_r(\mathbf{r}_c)\} , \tag{4.21}$$

where the row index r runs over N states and the column index c runs over N particles. The determinant automatically guarantees that anti-symmetry is kept, because it changes sign under a permutation of columns or rows. There is an additional important consequence of the antrisymmetric structure of wave functions that is particularly easy to see in a Slater determine. Imagine that two fermions, c and c', were allocated to the same state, r. This would mean that two rows in the determinant would be the same, which leads to a zero determinant. This indicates that two identical fermions cannot occupy the same state - a key, general result that is known as the *Pauli exclusion principle*.

Slater determinants can be shown to be the eigenstates of non-interacting $N-$body systems. When interactions between the particles are present, the ground state many-body wave function may not be a single, simple Slater determinant. The interactions allow transitions of particles from between different orbitals, so one may need more than one set of N single-particle wave functions to describe the system. One can use Slater determinants containing different single-particle states as a basis for many-body fermionic calculations.

It is important to note that the state index c in a Slater determinant runs from the lowest energy-single particle state, typically denoted by $c = 1$, to the second energy state, $c = 2$, all the way up to the N^{th} single-particle energy eigenvalue.

4.2 Fock space

The formalism we have described so far works for many-body systems, but it is not without flaws. For instance, the overlap between two different symmetrized $N-$body states (which is necessary to compute, say, expectation values or transition matrix elements) require the calculation of $(N!)^2$ different integrals. This is a taxing procedure as N increases. Moreover, the method we have described so far requires different calculations for different values of N. One would need to symmetrize or antisymmetrize wave functions differently for different values of N. Finally, one may also wonder how to extend this treatment to cases of relevance in statistical mechanics, where the particle number of N is not necessarily fixed. Think, for instance, of statistical mechanics and finite temperature situations, in the grandcanonical ensemble, where N is only determined on average. Of course, in particle physics there are also cases where particles are created or destroyed from energy requirements alone.

All of these arguments motivate the development of a different technical framework for quantum many-body physics. So far, we

have discussed the problem by working in the Hilbert space $\mathcal{H}(N)$ of $N-$particle systems. For systems in real space, this would correspond to the space of square-integrable antisymmetric combinations of N functions. If we want to let go of the idea of a fixed number of particles, we need to form a Hilbert space that is large enough to accommodate an undetermined number of particles. This is achieved by introducing the so-called *Fock space*,

$$\mathcal{F} = \bigoplus_{N=0}^{\infty} \mathcal{H}(N) \,.$$

This Fock space contains:

- a vacuum state, $|0\rangle$, with no particles;
- a complete set of single-particle states, $|\alpha\rangle$, with $\alpha = 1, 2, 3, \cdots$;
- a complete set of two-particle states, $|\alpha\beta\rangle$;
- a complete set of three-particle states, $|\alpha\beta\rangle$; and so on

These complete sets of many-particle states contain only vectors of the correct permutation symmetry, which we choose here to be fermionic as indicated by the $|\cdot\rangle$ notation. These could be bosonic states $|\cdot\}$, as well.

4.2.1 Creation and destruction operators

The key tool to exploit the extension into Fock space are the creation, $\hat{a}^\dagger$, and destruction, $\hat{a}$, operators. These act as mappings between many-particle Hilbert spaces of different particle numbers,

$$\hat{a}^\dagger_\alpha : \mathcal{H}(N-1) \to \mathcal{H}(N) \,,$$
$$\hat{a}_\alpha : \mathcal{H}(N) \to \mathcal{H}(N-1) \,.$$

Let us first discuss the creation operator, and we shall derive the properties of the annihilation operator from it. We also restrict the discussion to fermions in the first instance, although the extension to bosons is relatively straightforward and discussed in detail in several literature sources[5]. Acting on top of the vacuum state, with no particles, the operator $\hat{a}^\dagger_\alpha$ creates a particle in the state α,

$$\hat{a}^\dagger_\alpha |0\rangle = |\alpha\rangle \,. \tag{4.22}$$

Acting on top of a one-particle state $|\beta\rangle$, the same operator creates a two-particle state,

$$\hat{a}^\dagger_\alpha |\beta\rangle = |\alpha\beta\rangle \,. \tag{4.23}$$

It is useful to demand this state to be antisymmetric from the outset. For this to occur, we must have

$$\hat{a}^\dagger_\alpha \hat{a}^\dagger_\beta |0\rangle = |\alpha\beta\rangle = -|\beta\alpha\rangle = -\hat{a}^\dagger_\beta \hat{a}^\dagger_\alpha |0\rangle \,.$$

[5] L. E. Ballentine. *Quantum Mechanics: A Modern Development*. World Scientific, 1998; V. Zelevinsky. *Quantum Physics*, volume 2. Wiley-VCH, 2011; and E.K.U. Gross, E. Runge, and O. Heinonen. *Many-Particle Theory,*. Taylor & Francis, 1991

We can impose this condition through the anticommutation relation

$$\hat{a}_\alpha^\dagger \hat{a}_\beta^\dagger + \hat{a}_\beta^\dagger \hat{a}_\alpha^\dagger = 0 \Rightarrow \{\hat{a}_\alpha^\dagger, \hat{a}_\beta^\dagger\} = 0\,. \tag{4.24}$$

Adding a third particle to a two-body state immediately yields an anti-symmetrized three-body state,

$$\hat{a}_\alpha^\dagger |\beta\gamma\rangle = |\alpha\beta\gamma\rangle\,. \tag{4.25}$$

It is important to note the ordering of the states on antisymmetrized braket. Every time we create a state, we do it on the left-hand-side of the braket.

A generic state in the Fock space can be a linear combination of states with any number of particles. A state like this,

$$|\psi\rangle = \frac{1}{2}|0\rangle + \frac{1}{\sqrt{2}}|\alpha\rangle + \frac{1}{2}|\alpha\beta\rangle\,, \tag{4.26}$$

which includes a vacuum, single-particle and two-particle contribution, is perfectly possible.

Suppose we have an arbitrary $N-$body antisymmetrized ket where the state α has not been populated[6]. We shall use the notation

$$|\underset{\neq\alpha}{\cdots}\rangle$$

for such a state. Adding a particle in state α on top of this state leads somewhat unconsequentially to

$$\hat{a}_\alpha^\dagger |\underset{\neq\alpha}{\cdots}\rangle = |\alpha\ \underset{\neq\alpha}{\cdots}\rangle\,.$$

[6] This could be a single-particle state $|\beta\rangle$, but also a $4-$particle state like $|\beta\gamma\delta\eta\rangle$.

However, adding another particle on state α here yields

$$\hat{a}_\alpha^\dagger |\alpha\ \underset{\neq\alpha}{\cdots}\rangle = |\alpha\alpha\ \underset{\neq\alpha}{\cdots}\rangle\,.$$

Upon a permutation of the two first particles, the antisymmetric ket changes sign,

$$|\alpha\alpha\ \underset{\neq\alpha}{\cdots}\rangle = -|\alpha\alpha\ \underset{\neq\alpha}{\cdots}\rangle = 0,$$

Note that the position of the two α states is actually inconsequential for this argument. If one of them is n positions within the ket, we can perform $n-1$ permutations to bring it to the left-most position, changing only the sign. We can then do the same for the second α state, and the prove above holds.

but since the same result is recovered with a minus sign, the state itself must be zero. This is a mathematical proof of the fermionic Pauli principle,

$$\hat{a}_\alpha^\dagger |\alpha\ \underset{\neq\alpha}{\cdots}\rangle = 0, \tag{4.27}$$

which shows the usefulness of creation and destruction operators.

From the previous properties, one can also derive the two relations:

$$\langle\alpha\ \underset{\neq\alpha}{\cdots}|\hat{a}_\alpha^\dagger|\underset{\neq\alpha}{\cdots}\rangle = 1, \tag{4.28}$$

$$\langle\psi|\hat{a}_\alpha^\dagger|\underset{\neq\alpha}{\cdots}\rangle = 0,\ \text{if}\ \langle\psi|\alpha\ \underset{\neq\alpha}{\cdots}\rangle = 0. \tag{4.29}$$

One can summarize these relations by saying that creation operators always "create" on the leftmost side of the ket, and that destruction operators create on the leftmost side of the bra.

We can proceed and derive relations for the so-called annihilation operator, $\hat{a}_\alpha$, which is the adjoint operator of $\hat{a}_\alpha^\dagger$. We start from the most basic properties, Eqs. (4.22), (4.23) and (4.25) to find:

$$\hat{a}_\alpha^\dagger|0\rangle = |\alpha\rangle \;\rightarrow\; \langle 0|\hat{a}_\alpha = \langle\alpha|\,, \tag{4.30}$$

$$\hat{a}_\alpha^\dagger|\beta\rangle = |\alpha\beta\rangle \;\rightarrow\; \langle\beta|\hat{a}_\alpha = \langle\alpha\beta|\,, \tag{4.31}$$

$$\hat{a}_\alpha^\dagger|\beta\gamma\rangle = |\alpha\beta\gamma\rangle \;\rightarrow\; \langle\beta\gamma|\hat{a}_\alpha = \langle\alpha\beta\gamma|\,, \tag{4.32}$$

$$\hat{a}_\alpha^\dagger|\underset{\neq\alpha}{\cdots}\rangle = |\alpha\,\underset{\neq\alpha}{\cdots}\rangle \;\rightarrow\; \langle\alpha\,\underset{\neq\alpha}{\cdots}|\hat{a}_\alpha = \langle\underset{\neq\alpha}{\cdots}|\,. \tag{4.33}$$

Moreover, from the Pauli principle expression in Eq. (4.27), we find

$$\langle\alpha\,\underset{\neq\alpha}{\cdots}|\hat{a}_\alpha = 0. \tag{4.34}$$

Note that, by construction, the antisymmetrized kets are normalized so that: $\langle\alpha|\alpha\rangle = 1\,, \langle\alpha\beta|\alpha\beta\rangle = 1\,, \langle\alpha\beta\gamma|\alpha\beta\gamma\rangle = 1\,, \cdots$

Moreover, acting with a generic ket $|\psi\rangle$ on this equation, we also find that

$$\langle\alpha\,\underset{\neq\alpha}{\cdots}|\hat{a}_\alpha|\psi\rangle = 0. \tag{4.35}$$

The equivalent to Eqs. (4.28) and (4.28) for the annihilation operator read

$$\langle\underset{\neq\alpha}{\cdots}|\hat{a}_\alpha|\alpha\,\underset{\neq\alpha}{\cdots}\rangle = 1, \tag{4.36}$$

$$\langle\underset{\neq\alpha}{\cdots}|\hat{a}_\alpha|\psi\rangle = 0, \text{ if } \langle\psi|\alpha\,\underset{\neq\alpha}{\cdots}\rangle = 0. \tag{4.37}$$

To keep the antisymmetry of two-body kets in Eq. (4.31), annihilation operators must fulfil an anticommutation relation

$$\{\hat{a}_\alpha, \hat{a}_\beta\} = 0. \tag{4.38}$$

Consider now the case in which we choose the vacuum as the state ψ, $|\psi\rangle = |0\rangle$. From Eq. (4.37), we have

$$\langle\underset{\neq\alpha}{\cdots}|\hat{a}_\alpha|0\rangle = 0, \text{ if } \langle 0|\alpha\,\underset{\neq\alpha}{\cdots}\rangle = 0, \tag{4.39}$$

which indicates that the state $\hat{a}_\alpha|0\rangle$ is orthogonal to any vector where α is unoccupied. Moreover, from Eq. (4.35), we also find that

$$\langle\alpha\,\underset{\neq\alpha}{\cdots}|\hat{a}_\alpha|0\rangle = 0, \tag{4.40}$$

and so $\hat{a}_\alpha|0\rangle$ is orthogonal to any vector where α is occupied, too! The only way this can happen is if

$$\hat{a}_\alpha|0\rangle = 0, \tag{4.41}$$

To be clear, this means that the result of annihilating the vacuum state is null. The right-hand side of the equation is a strict zero, not a vacuum state.

and so the chain of annihilation operators is destructed at the vacuum.

Further, we can also use Eq. (4.35) with $|\psi\rangle = |\alpha\rangle$ to find

$$\langle\alpha\,\underset{\neq\alpha}{\cdots}|\hat{a}_\alpha|\alpha\rangle = 0, \tag{4.42}$$

which indicates that $\hat{a}_\alpha|\alpha\rangle$ is orthogonal to any vector where α is occupied. The only exception to this rule arises from Eq. (4.36) when $|\underset{\neq\alpha}{\cdots}\rangle = |0\rangle$, which indicates that

$$\langle 0|\hat{a}_\alpha|\alpha\rangle = 1, \tag{4.43}$$

and hence

$$a_\alpha|\alpha\rangle = |0\rangle. \tag{4.44}$$

Clearly, from Eq. (4.36) we also have that the action of an annihilation operator on any ket (not just a single-particle one) containing α

$$\hat{a}_\alpha|\alpha \underset{\neq\alpha}{\cdots}\rangle = |\underset{\neq\alpha}{\cdots}\rangle, \tag{4.45}$$

is to create a new ket with no α on it. In contrast, using $|\psi\rangle = |\underset{\neq\alpha}{\cdots}\rangle$ in Eqs. (4.37) and (4.35), we find that

$$\hat{a}_\alpha|\underset{\neq\alpha}{\cdots}\rangle = 0. \tag{4.46}$$

In other words, $\hat{a}_\alpha$ destroys the vector it the state α is unoccupied on it.

The last remaining part of second-quantization operator algebra is the anticommutation relation between creation and annihilation operators. Consider first the following case when $\alpha \neq \beta$:

$$\hat{a}_\alpha\hat{a}^\dagger_\beta + \hat{a}^\dagger_\beta\hat{a}_\alpha|\alpha \underset{\neq\beta}{\cdots}\rangle = \hat{a}_\alpha|\beta\alpha \underset{\neq\beta}{\cdots}\rangle + \hat{a}^\dagger_\beta|\underset{\neq\beta}{\cdots}\rangle = -|\beta \underset{\neq\beta}{\cdots}\rangle + |\beta \underset{\neq\beta}{\cdots}\rangle.$$

The choice of the state $|\alpha \underset{\neq\beta}{\cdots}\rangle$ is dictated by the anti-commutator. A state with β occupied or α empty will be zero straight away.

For $\alpha = \beta$, one can consider two separate cases. If the state contains α,

$$\hat{a}_\alpha\hat{a}^\dagger_\alpha + \hat{a}^\dagger_\alpha\hat{a}_\alpha|\alpha \underset{\neq\alpha}{\cdots}\rangle = \hat{a}_\alpha \overbrace{|\alpha\alpha \underset{\neq\beta}{\cdots}\rangle}^{=0} + \hat{a}^\dagger_\alpha|\underset{\neq\beta}{\cdots}\rangle = |\alpha \underset{\neq\alpha}{\cdots}\rangle$$

the anticommutator does not change the state. If it does not contain α,

$$\hat{a}_\alpha\hat{a}^\dagger_\alpha + \hat{a}^\dagger_\alpha\hat{a}_\alpha|\underset{\neq\alpha}{\cdots}\rangle = \hat{a}_\alpha|\alpha \underset{\neq\beta}{\cdots}\rangle + 0 = |\alpha \underset{\neq\alpha}{\cdots}\rangle,$$

the same thing happens! We conclude then that the following operator identity holds

$$\{\hat{a}_\alpha, \hat{a}^\dagger_\beta\} = \delta_{\alpha,\beta} \tag{4.47}$$

In this process, one sees that all Fock space vectors are actually eigenstates of the operator $\hat{a}^\dagger_\alpha\hat{a}_\alpha$:

$$\hat{a}^\dagger_\alpha\hat{a}_\alpha|\underset{\neq\alpha}{\cdots}\rangle = 0|\underset{\neq\alpha}{\cdots}\rangle$$
$$\hat{a}^\dagger_\alpha\hat{a}_\alpha|\alpha \underset{\neq\alpha}{\cdots}\rangle = 1|\alpha \underset{\neq\alpha}{\cdots}\rangle.$$

Note that the position of the α state in the second equality is inconsequential. If α is r states into the Fock ket, we can act with $r-1$ permutations to bring it to the leftmost position and act with $\hat{a}^\dagger_\alpha\hat{a}_\alpha$. If the result is not zero, we create α at the leftmost point and we can permute it $r-1$ times to the right, to fill the same position as in the initial ket. The phase acquired by these permutations is $(-1)^{2r-2} = +1$.

If $n_\alpha = 0$ or 1 is the occupation of fermionic states, the previous two equalities can be summarised as

$$\hat{a}^\dagger_\alpha\hat{a}_\alpha|\cdots\rangle = n_\alpha|\cdots\rangle,$$

where the state $|\cdots\rangle$ may or may not contain α. If we want to count all the fermions in our system, we should add up the n_α values over all states. This motivates the definition of the particle number operator,

$$\hat{N} = \sum_\alpha \hat{a}^\dagger_\alpha\hat{a}_\alpha. \tag{4.48}$$

This operator is different from other operators we have met before. To start with, it acts on Fock space rather than on a fixed $N-$body Hilbert space - since it creates and destroys a particle. The sum in this case goes over all the available single-particle states of the $1-$body Hilbert space, even though this operator counts many-body states. This is interesting because for a state with fixed N, $\hat{N}$ counts N but not by a sum over different particles, $i = 1, \cdots, N$, but over basis states (which is, in principle, an infinite sum $\alpha = 0, \cdots, \infty$). This procedure, in which we transform operators in the many-body system from sums over particles to sums over states, is useful for a variety of reasons and is the heart of the second quantization procedure.

A vector that is a linear combination of many-body basis vectors may not be an eigenvector of $\hat{N}$, in which case one expects a distribution of Ns. For instance, for the state in Eq. (4.26), one finds

$$\begin{aligned}\langle\psi|\hat{N}|\psi\rangle &= \frac{1}{4}\langle 0|\hat{N}|0\rangle + \frac{1}{2}\langle\alpha|\hat{N}|\alpha\rangle + \frac{1}{4}\langle\alpha\beta|\hat{N}|\alpha\beta\rangle \\ &= 0 + \frac{1}{2} + \frac{2}{4} = 1,\end{aligned}$$

even though the eigenstate is a mixture of 0, one- and two-body eigenstates.

4.3 *Operators in second quantization*

So far, we have seen that the creation and annihilation operators allow us to define practically many-body states. This in itself wouldn't be useful if we couldn't also transform operators into their Fock state counterparts. Let us first consider operators that act over one-body properties. Think, for instance, on the kinetic energy of an $N-$body system of identical fermions (and hence with the same mass m),

$$\hat{T} = \sum_{i=1}^{N} \frac{\hat{p}_i^2}{2m} = \sum_{i=1}^{N} \hat{t}_i. \tag{4.49}$$

The single-particle operators t_i act on one particle at a time[7]. We use $\hat{t}_i$ as a generic one-body operator in the following. Since it acts on individual particles, $\hat{t}_i$ has the matrix elements

$$t_{\alpha\beta} = \langle\alpha|\hat{t}_i|\beta\rangle\,. \tag{4.50}$$

[7] Similar expressions hold for other one-body properties like the total momentum, or the total external potential energy of a system of N particles.

A key point of this expression is that it does not depend on the particle index i. The single-particle states $|\alpha\rangle$ are predefined and the operator $\hat{t}_i$ is exactly the same for each particle - so the corresponding single-particle matrix elements do not know to which particle they need to be applied to. The spectral decomposition of the operator in the single-particle Hilbert space is then

$$\hat{t}_i = \sum_{\alpha,\beta} t_{\alpha\beta}|\alpha\rangle_i\langle\beta|_i, \tag{4.51}$$

where the particle information is stored in the bras and kets. With this, the total kinetic energy operator becomes

$$\hat{T} = \sum_{\alpha,\beta} t_{\alpha\beta} \sum_{i=1}^{N} |\alpha\rangle_i\langle\beta|_i. \tag{4.52}$$

Equation (4.49) above is somewhat limited, because there is a different number of terms for every physical system with a different number of particles. One would like to use it directly in a formalism involving creation and destruction operators. Consider how the second term of the one-body operator of Eq. (4.52) acts on a general

antisymmetrized ket:

$$\sum_{i=1}^{N}|\alpha\rangle_i\langle\beta|_i|\gamma\delta\cdots\eta\rangle = \sum_{i=1}^{N}|\alpha\rangle_i\langle\beta|_i\frac{1}{\sqrt{N!}}\sum_P(-1)^P P\left\{|\gamma\rangle_1\otimes|\delta\rangle_2\cdots\otimes|\eta\rangle_N\right\}.$$

The permutation operator can be safely taken out of the expression because the term $|\alpha\rangle_i\langle\beta|_i$ depends on a single particle (so there is really nothing to permute there!) and so

$$\sum_{i=1}^{N}|\alpha\rangle_i\langle\beta|_i|\gamma\delta\cdots\eta\rangle = \frac{1}{\sqrt{N!}}\sum_P(-1)^P\sum_{i=1}^{N}|\alpha\rangle_i\langle\beta|_i|\gamma\rangle_1\otimes|\delta\rangle_2\cdots\otimes|\eta\rangle_N.$$

On the sum over i, each term will act on one of the direct product kets on the right hand side. For the first term, for instance, one will find $\langle\beta|_1|\gamma\rangle_1 = \delta_{\beta,\gamma}$, which indicates that the term will only contribute if β and γ are the same. A similar thing happens for the second term - and all the way up to the Nth one. For fermionic systems, however, the ket can only contain at most one particle in the state β so, of all the N terms in the sum, only one remains. Suppose we had $\delta = \beta$ in the second particle. We would then find

$$\sum_{i=1}^{N}|\alpha\rangle_i\langle\beta|_i|\gamma\delta\cdots\eta\rangle = \frac{1}{\sqrt{N!}}\sum_P(-1)^P P\left\{|\gamma\rangle_1\otimes|\alpha\rangle_2\cdots\otimes|\eta\rangle_N\right\}.$$

There are two important things in this expression. First, the sum over i has been reduced to a single term. Second, the antisymmetrization guarantees the state becomes

$$\sum_{i=1}^{N}|\alpha\rangle_i\langle\beta|_i|\gamma\beta\cdots\eta\rangle = |\gamma\alpha\cdots\eta\rangle.$$

Now consider the following combination of creation and destruction operators acting on the same state,

$$\hat{a}_\alpha^\dagger\hat{a}_\beta|\gamma\beta\cdots\eta\rangle = |\gamma\alpha\cdots\eta\rangle,$$

which gives the same result! Of course, the same would have happened if, say, $\gamma = \beta$ or $\eta = \beta$ or, for that matter, any other state in the ket. If the ket had had no state β at all, the successive overlaps $\langle\beta|_i|\cdot\rangle_i$ would all cancel and the result would be zero. This would also happen by acting with $\hat{a}_\beta$ on a state not containing β. Overall, we conclude that the operator identity

$$\sum_{i=1}^{N}|\alpha\rangle_i\langle\beta|_i = \hat{a}_\alpha^\dagger\hat{a}_\beta$$

holds independently of the state under consideration. With this, one has from Eq. (4.52), that

$$\hat{T} = \sum_{\alpha,\beta}t_{\alpha,\beta}\hat{a}_\alpha^\dagger\hat{a}_\beta, \tag{4.53}$$

which is the second quantization representation of the operator. Compared to the original expression, this operator does not

To see this, consider the two-body example

$$\left(|\alpha\rangle_1\langle\beta|_1+|\alpha\rangle_2\langle\beta|_2\right)|\delta\gamma\rangle =$$
$$\left(|\alpha\rangle_1\langle\beta|_1+|\alpha\rangle_2\langle\beta|_2\right)\left(\frac{1}{\sqrt{2}}|\gamma\rangle_1|\delta\rangle_2-\frac{1}{\sqrt{2}}|\delta\rangle_1|\gamma\rangle_2\right)$$
$$=\frac{1}{\sqrt{2}}|\alpha\rangle_1\langle\beta|_1|\gamma\rangle_1|\delta\rangle_2-\frac{1}{\sqrt{2}}|\alpha\rangle_1\langle\beta|_1|\delta\rangle_1|\gamma\rangle_2$$
$$+\frac{1}{\sqrt{2}}|\alpha\rangle_2\langle\beta|_2|\gamma\rangle_1|\delta\rangle_2-\frac{1}{\sqrt{2}}|\alpha\rangle_2\langle\beta|_2|\delta\rangle_1|\gamma\rangle_2$$
$$=\frac{\delta_{\beta,\gamma}}{\sqrt{2}}|\alpha\rangle_1|\delta\rangle_2-\frac{\delta_{\beta,\delta}}{\sqrt{2}}|\alpha\rangle_1|\gamma\rangle_2$$
$$+\frac{\delta_{\beta,\delta}}{\sqrt{2}}|\gamma\rangle_1|\alpha\rangle_2-\frac{\delta_{\beta,\gamma}}{\sqrt{2}}|\delta\rangle_1|\alpha\rangle_2$$

where we have omitted the $\otimes$ symbol for brevity. We can group the terms coming from the different δ functions to find

$$\left(|\alpha\rangle_1\langle\beta|_1+|\alpha\rangle_2\langle\beta|_2\right)|\delta\gamma\rangle =$$
$$=\frac{\delta_{\beta,\gamma}}{\sqrt{2}}|\alpha\rangle_1|\delta\rangle_2-\frac{\delta_{\beta,\gamma}}{\sqrt{2}}|\delta\rangle_1|\alpha\rangle_2$$
$$+\frac{\delta_{\beta,\delta}}{\sqrt{2}}|\alpha\rangle_1|\gamma\rangle_2-\frac{\delta_{\beta,\delta}}{\sqrt{2}}|\gamma\rangle_1|\alpha\rangle_2$$
$$=\delta_{\beta,\gamma}\,|\alpha\delta\rangle-\delta_{\beta,\delta}\,|\gamma\alpha\,\rangle$$
$$=\delta_{\beta,\gamma}\,|\alpha\delta\rangle+\delta_{\beta,\delta}\,|\alpha\gamma\,\rangle.$$

This is the same result that one obtains from the second-quantization operation

$$\hat{a}_\alpha^\dagger\hat{a}_\beta\,|\delta\gamma\rangle = \delta_{\beta,\gamma}\,|\alpha\delta\rangle+\delta_{\beta,\delta}\,|\alpha\gamma\,\rangle.$$

explicitly depend on the number of particles since it acts and lives in Fock space. It is only when $\hat{T}$ is applied to an antisymmetric $N-$body ket that the dependence over N is restored. We should also point out that, while the derivation is a bit more cumbersome, the very same expression holds for bosons.

One-body operators are not all there is, though. In an interacting system, a key two-body operator is the interaction operator, which provides the potential between particles i and j,

$$\hat{V} = \sum_{i,j} \hat{V}_{i,j} = \frac{1}{2} \sum_{i \neq j} \hat{V}_{i,j}. \tag{4.54}$$

The second-quantized version of this operator is given by the expression,

$$\hat{V} = \frac{1}{2} \sum_{\alpha,\beta,\delta,\gamma} v_{\alpha\beta,\gamma\delta} \hat{a}^\dagger_\alpha \hat{a}^\dagger_\beta \hat{a}_\delta \hat{a}_\gamma, \tag{4.55}$$

where it is important to stress that the order of the last two destruction operators is inverted with respect to the matrix element,

$$v_{\alpha\beta,\gamma\delta} = (\alpha\beta|\hat{V}_{i,j}|\gamma\delta). \tag{4.56}$$

We have (at least theoretically!) constructed the Hamiltonian $\hat{H} = \hat{T} + \hat{V}$ in second quantization. The expression is independent of the number of particles and of the basis that we have chosen. Specifying a basis will immediately provide one with the single-particle states, which can be used to compute the matrix elements $t_{\alpha,\beta}$ and $v_{\alpha\beta,\gamma\delta}$. For a given system with N particles, one would then have to compute an expectation value of the type $\langle\chi_1 \cdots \chi_N|\hat{H}|\chi_1 \cdots \chi_N\rangle$, involving terms like $\langle\chi_1 \cdots \chi_N|\hat{a}^\dagger_\alpha \hat{a}_\beta|\chi_1 \cdots \chi_N\rangle$. There is whole technology set up to compute these expectation values, including Feynman diagrammatic techniques that arise from a Dyson series of the time evolution operator. The interested reader may see Ref.[8] for more details.

[8] E.K.U. Gross, E. Runge, and O. Heinonen. *Many-Particle Theory,*. Taylor & Francis, 1991

4.4 Bosons

For bosons, some of the properties above are the same and some are different. Creation and annihilation operators for bosons can be defined as well. However, demanding symmetry (rather than antisymmetry) in the two-body ket immediately implies that boson operators commute, and so

$$[\hat{a}^\dagger_\alpha, \hat{a}^\dagger_\beta] = 0, \quad [\hat{a}_\alpha, \hat{a}_\beta] = 0, \tag{4.57}$$

$$[\hat{a}_\alpha, \hat{a}^\dagger_\beta] = \delta_{\alpha,\beta}\,. \tag{4.58}$$

A key difference now is, of course, that more than two bosons can sit on the very same state. It is useful to work in this case with the occupation number representation. Suppose we had a symmetric ket in which state 1 was occupied twice; state 2 was not occupied; and

states 3 and 4 were occupied three times each. In the standard notation we have used so far, we would write this state as

$$|1,1,3,3,3,4,4,4\}.$$

There is clearly a lot of redundancy in this state representation. A more convenient notation can be used by specifying how many particles reside in each single-particle state. In this occupation number notation, the same state would read

$$|2_1, 0_2, 3_3, 3_4\}.$$

Thus a state in occupation number representation reads

$$|n_1, n_2, \cdots\},$$

This, of course, works for fermionic states too, so we write $|n_1, n_2, \cdots\rangle$, except here n_i can only be 0 or 1.

and typically there is a restriction $N = \sum_\alpha n_\alpha$.

The algebra of bosonic creation and destruction operators is isomorphous to the algebra of raising and lowering operators of the harmonic oscillator. In a sense, each bosonic state with n_α particles is equivalent to an harmonic oscillator in state $|n\rangle$. From this, one can easily deduce that the action of the creation and anihilation operators in bosonic many-body states can easily be represented in the occupation number representation as follows

$$\hat{a}_\alpha|n_1, n_2, \cdots, n_\alpha, \cdots\} = \sqrt{n_\alpha}|n_1, n_2, \cdots, n_\alpha - 1, \cdots\}, \tag{4.59}$$

$$\hat{a}^\dagger_\alpha|n_1, n_2, \cdots, n_\alpha, \cdots\} = \sqrt{n_\alpha + 1}|n_1, n_2, \cdots, n_\alpha + 1, \cdots\}. \tag{4.60}$$

The equivalent expression for fermions requires some more work to account for the phase acquired by permuting the state α to the leftmost position, $\theta_\alpha = (-1)^{\sum_{\beta<\alpha} n_\beta}$,

$$\hat{a}_\alpha|n_1, n_2, \cdots, n_\alpha, \cdots\rangle = n_\alpha\theta_\alpha|n_1, n_2, \cdots, 0_\alpha, \cdots\rangle, \tag{4.61}$$

$$\hat{a}^\dagger_\alpha|n_1, n_2, \cdots, n_\alpha, \cdots\rangle = (1 - n_\alpha)\theta_\alpha|n_1, n_2, \cdots, 1_\alpha, \cdots\rangle. \tag{4.62}$$

To understand how the phase comes about, we demand equivalence between the occupation number representation and the usual representation of fermionic Fock states. This requires a well-defined ordering of the states in the basis. For clarity, we now label our states with labels λ_i rather than $\alpha, \beta, \ldots$. The indices $i = 1, \cdots, i_{\max}$ may be dictated, for instance, by the ordering of single-particle energies. For eigenstates of $\hat{N}$, one usually has that $i_{\max} = N$. When the ordering is performed, a fermionic state with N particles can be generated out of the vacuum by acting with N creation operators[9]

$$|\lambda_1, \lambda_2, \cdots, \lambda_N\rangle = \hat{a}^\dagger_{\lambda_1}\hat{a}^\dagger_{\lambda_2}\cdots\hat{a}^\dagger_{\lambda_N}|0\rangle \tag{4.63}$$

[9] Note that the order in which operators are written is the same as the order of the states in the ket label.

Now, if one tries to remove the state λ_k (with $k = 1, \cdots, N$) from this Fock state, one needs to permute it all the way to the left-hand side of the ket. In this process, the ket acquires a phase θ_k which depends on how many creation operators are commuted through,

$$\begin{aligned} a_{\lambda_k}|\lambda_1, \lambda_2, \cdots, \lambda_{k-2}, \lambda_{k-1}, \lambda_k, \lambda_{k+1}, \cdots \lambda_N\rangle &= -a_{\lambda_k}|\lambda_1, \lambda_2, \cdots, \lambda_{k-2}, \lambda_k, \lambda_{k-1}, \lambda_{k+1}, \cdots \lambda_N\rangle \\ &= +a_{\lambda_k}|\lambda_1, \lambda_2, \cdots, \lambda_k, \lambda_{k-2}, \lambda_{k-1}, \lambda_{k+1}, \cdots \lambda_N\rangle \\ &= \theta_k a_{\lambda_k}|\lambda_k, \lambda_1, \cdots, \lambda_{k-2}, \lambda_{k-1}, \lambda_{k+1}, \cdots \lambda_N\rangle \\ &= \theta_k|\lambda_1, \cdots, \lambda_{k-2}, \lambda_{k-1}, \lambda_{k+1}, \cdots \lambda_N\rangle\,. \end{aligned}$$

Of course, if the original state represented an $N-$particle state, the outcome of the operation yields a state with $N-1$ particles.

In occupation number representation, the original state of Eq. (4.63) would read

$$|1_1, 1_2, \cdots, 1_N\rangle \equiv |\underbrace{1, 1, \cdots, 1}_{N \text{ times}}, 0, 0, 0\rangle\,, \tag{4.64}$$

where all occupied states have $n_i = 1_i$ because these are fermions. Clearly, if we want to remove the state at k we can do so, but we need to keep the same phase θ_k,

$$a_k|\underbrace{1, 1, \cdots, 1}_{N \text{ times}} 0, 0, 0\rangle = \theta_k|1_1, \cdots 1_{k-1}, 0_k, 1_{k+1}, \cdots, 1_N, 0, 0, 0\rangle\,. \tag{4.65}$$

Finally, we note that for both bosons and fermions, we can generate any Fock state of $N = \sum_\alpha n_\alpha$ particles by applying the creation operator on the vacuum several times:

$$|n_1, n_2, \cdots\rangle = \prod_\alpha \frac{1}{\sqrt{n_\alpha!}} \left(a_\alpha\right)^{n_\alpha} |0\rangle\,.$$

4.5 *Free Fermi gas*

As an example to the importance of many-body effects, we shall consider here an ideal, yet extremely model system. The non-interacting Fermi gas is a cornerstone of many-body theory and one that provides a fundamental understanding of several issues at play in fields that range from astronomy to nuclear physics. In view of the lack of interactions, the hamiltonian of a system of N identical fermions of mass m is simply

$$\hat{H} = \sum_{i=1}^{N} \frac{\hat{p}_i^2}{2m}.$$

It is natural to use a single-particle basis that diagonalises the single-particle hamiltonian. In this case, this corresponds to the momentum basis,

$$\frac{\hat{p}^2}{2m}|\mathbf{p}\rangle = \frac{p^2}{2m}|\mathbf{p}\rangle\,.$$

For simplicity, we will discuss the free Fermi gas in the one-dimensional case. We move later to the more realistic and important three-dimensional scenario.

4.5.1 *One dimension*

In one dimension, the momentum and position are scalar (rather than vector) operators. The eigenstates of the momentum operator projected in real space are nothing but plane waves,

$$\langle x|p\rangle = \psi_p(x) = \mathcal{N} e^{i\frac{px}{\hbar}}\,, \tag{4.66}$$

and $\mathcal{N}$ is a normalization constant. We assume that our system is embedded in a line of total length L. We need to consider boundary conditions for the wavefunction to quantize the momentum. For convenience, we choose these to be *periodic*. In this case,

$$\psi_p(x+L) = \psi_p(x) \Rightarrow e^{i\frac{pL}{\hbar}} = 1 \Rightarrow p_n = \frac{2\pi\hbar}{L} n\,, \tag{4.67}$$

and we see that the momentum becomes quantized. The quantum number $n = 0, \pm 1, \pm 2, \ldots$ provides a discrete value of momentum, p_n. We illustrate these in Fig. 4.3. An increase δn of momentum quanta involves a momentum increase of $\Delta p = \frac{2\pi\hbar}{L}\Delta n$. In other words, each momentum state occupies a size of $\delta p \approx \frac{2\pi\hbar}{L}$ in momentum space.

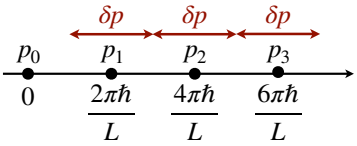

Figure 4.3: In a periodic system of length L, momenta are discretized, $p_n = \delta p n$, in quanta of size $\delta p = \frac{2\pi\hbar}{L}$.

The wave functions of discrete momenta must be normalized. By integrating the periodic wavefunction in the region $x \in (0, L)$, we find

$$\begin{aligned}\int_0^L \mathrm{d}x\, \psi^*_{p_m}(x)\psi_{p_n}(x) &= |\mathcal{N}|^2 \int_0^L \mathrm{d}x\, e^{i\frac{2\pi}{L}(m-n)x} \\ &= \begin{cases} |\mathcal{N}|^2 L, & n = m, \\ \frac{|\mathcal{N}|^2 L}{i2\pi(m-n)}\left(e^{i2\pi(m-n)} - 1\right), & n \neq m. \end{cases}\end{aligned} \tag{4.68}$$

For $n \neq m$, the previous integral clearly reduces to 0 because $e^{i2\pi(m-n)} = 1$. For $n = m$, we demand the wave functions to be normalized and find $\mathcal{N} = \frac{1}{\sqrt{L}}$. The integral then clearly reduces to the expected orthogonality condition

$$\int_0^L \mathrm{d}x\, \psi^*_{p_m}(x)\psi_{p_n}(x) = \delta_{n,m}\,. \tag{4.69}$$

All in all, we find that the eigenvalue equation for the single-particle basis is now particularly simple,

$$\frac{\hat{p}^2}{2m}\,|\mathbf{p_n}\rangle = \epsilon_n\,|\mathbf{p_n}\rangle\,,$$

with $\epsilon_n = \frac{p_n^2}{2m}$ and $p_n = n\delta p$. Consider now a set of $N = 3$ fermions filling up levels with positive n[10]. We consider first the case where such fermions do not have spin. The ground-state configuration for such fermions is one that minimises the energy, but also respects the Pauli principle. Clearly, the lowest-energy configuration in this setting is the one represented in the left panel of Fig. 4.4, with total energy $E = \epsilon_0 + \epsilon_1 + \epsilon_2$.

If the particles had spin $s = 1/2$, we could place $N = 6$ fermions in the same levels, because there would be 2 allowed states per single-particle level, as shown in in the right panel of Fig. 4.4. If these particles had $g > 2$ internal degrees of freedom (think, for instance, of isospin or color), we could place g fermions at each level[11]. Clearly, the lowest energy configuration in such a case would be $E = g(\epsilon_0 + \epsilon_1 + \epsilon_2)$.

In two situations we have just described, the single-particle states are filled in up until the state with $n = 2$. The last occupied state in

[10] The restriction to $n > 0$ can be lifted, but the final results in the. thermodynamic limit would be the same.

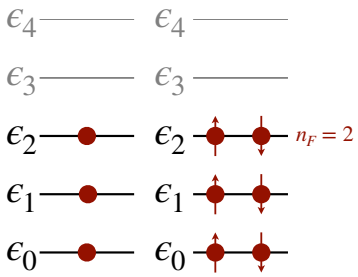

Figure 4.4: Left panel: minimum energy configuration of 3 spinless fermions Right panel: the same for a system of 6 fermions with spin 1/2. In the two cases, the last occupied state is the $n = 2$ state.

[11] We typically refer to g as the degeneracy of the system.

such a configuration is called the *Fermi level* and we denote it by n_F. The corresponding energy and momentum are the so-called Fermi energy, $\epsilon_F = \epsilon_{n_F}$, and Fermi momentum, $p_F = p_{n_F}$. It is easy to see that n_F is related to the number of particles and the degeneracy, $N = g n_F$. With this, we can write the Fermi momentum

We work for simplicity with an even number of particles N, but similar conclusions are reached for odd-particle-number systems.

$$p_F = \frac{2\pi\hbar}{L} n_F = \frac{2\pi\hbar}{g}\frac{N}{L} = \frac{2\pi\hbar}{g}\lambda \tag{4.70}$$

as a function of the number density, $\lambda = \frac{N}{L}$, of the system. This relation is rather remarkable. On the left-hand-side, we have a purely quantal property. On the right-hand-side, in contrast, we find λ, a thermodynamic property of the system. It is straightforward to see that Fermi energy is also a function of the density,

$$\epsilon_F = \frac{p_F^2}{2m} = \frac{\hbar^2}{2m}\frac{4\pi^2}{g^2}\lambda^2.$$

4.5.2 *Limit to the continuum*

This example shows that momentum modes in a periodic system are discretized. Adding over momentum modes, we should get a completeness relation,

$$\sum_n \psi^*_{p_n}(x)\psi_{p_n}(x') = \frac{1}{L}\sum_n e^{i\frac{2\pi n}{L}(x'-x)} \equiv \delta(x-x')\,. \tag{4.71}$$

This provides a definition of the $\delta-$function in real space. In the limit $L \to \infty$, one should recover a continuum of momentum states and the previous sum should be some sort of integral. We can indeed replace the sum by an integral if we take into account that each momentum mode has size Δp, since

$$\sum_n \to \frac{L}{2\pi\hbar}\int \mathrm{d}p. \tag{4.72}$$

This is perfectly compatible with Eq. (4.71) which, now in the continuum, reads

$$\frac{L}{2\pi\hbar}\int \mathrm{d}p\, \psi^*_p(x)\psi_p(x') = \frac{1}{2\pi\hbar}\int \mathrm{d}p\, e^{i\frac{p(x'-x)}{\hbar}} = \delta(x-x')\,. \tag{4.73}$$

Any sum over n can now be replaced by an integral. This is particularly useful if one wants to compute the total energy of the system,

$$E = g\sum_{n=0}^{n_F}\epsilon_n \to g\frac{L}{2\pi\hbar}\int_0^{p_F}\mathrm{d}p\epsilon_p = g\frac{L}{2\pi\hbar}\int_0^{p_F}\mathrm{d}p\frac{p^2}{2m} = g\frac{L}{2\pi\hbar}\frac{p_F^3}{6m}\,. \tag{4.74}$$

Inverting the relation in Eq. (4.70), we can find the density as a function of the Fermi momentum, $\lambda = \frac{g}{2\pi\hbar}p_F$, and using this in the previous expression one finds:

$$E = N\frac{p_F^2}{6m} = N\frac{\epsilon_F}{3}\,.$$

In other words, the energy per particle in the system

$$\frac{E}{N} = \frac{\epsilon_F}{3} \tag{4.75}$$

is just a third of the Fermi energy. Note that this energy is kinetic, but has a pure quantal origin. It is the quantization condition, together with the Pauli principle, that forces particles to have an increasingly large momentum value as they fill in the energy spectrum. This source of pressure is therefore purely quantal!

4.5.3 *Three dimensions*

This whole procedure can be straightforwardly extended to three dimensions. In that case, we consider a box of volume L^3 and a plane wave of the form $\psi_{\mathbf{p}}(\mathbf{r}) = \mathcal{N} e^{i\frac{\mathbf{p}\cdot\mathbf{r}}{\hbar}}$. We impose periodic boundary conditions along the three directions of space, $\psi_{\mathbf{p}}(x+L, y+L, z+L) = \psi_{\mathbf{p}}(x,y,z)$. This immediately dictates that momentum states be characterised by three indices, (n_x, n_y, n_z), so that the (vector) momentum becomes $\mathbf{p} = \frac{2\pi\hbar}{L}(n_x, n_y, n_z)$. Orthogonality of the wave functions over the box of size L^3,

$$\int \mathrm{d}^{(3)}\mathbf{r}\, \psi^*_{p_m}(r)\psi_{p_n}(r) = \delta_{n,m}\,,$$

allows us to determine the normalization constant. For $n = m$, following the procedure of Eq. (4.68), one finds that the integral yields $|\mathcal{N}|^2 L^3$ and hence $\mathcal{N} = \frac{1}{L^{3/2}}$.

In some references, the normalization is presented in terms of the volume $V = L^3$ and hence $\mathcal{N} = \frac{1}{\sqrt{V}}$.

We conclude that, in three dimensions, each quantum of momentum occupies a volume of size $\left(\frac{2\pi\hbar}{L}\right)^3$. In the infinite volume limit, the sums over momentum quanta should be replaced by integrals,

$$\sum_{n_x,n_y,n_z} \to \left(\frac{L}{2\pi\hbar}\right)^3 \int \mathrm{d}^{(3)}\mathbf{p}.$$

With the plane waves properly normalized to the volume,

$$\psi_{\mathbf{p}}(\mathbf{r}) = \frac{1}{L^{3/2}} e^{i\frac{\mathbf{p}\cdot\mathbf{r}}{\hbar}},$$

the completeness relation now reads

$$\left(\frac{L}{2\pi\hbar}\right)^3 \int \mathrm{d}^{(3)}\mathbf{p}\, \psi^*_{\mathbf{p}}(\mathbf{r})\psi_{\mathbf{p}}(\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}')\,,$$

where the δ is to be interpreted over three dimensions

$$\delta(\mathbf{r}-\mathbf{r}') = \delta(x-x')\delta(y-y')\delta(z-z')\,.$$

When filling in the lowest energy levels, particles will try to minimize the single-particle energies,

$$\epsilon_p = \frac{p^2}{2m} = \frac{1}{2m}\frac{2\pi\hbar}{L}\left(n_x^2 + n_y^2 + n_z^2\right). \tag{4.76}$$

The lowest energy configuration for a given number of particles is therefore a sphere in momentum space up to a given value of $n_x^2 + n_y^2 + n_z^2 < n_F$. One can therefore compute the total number of particles as

$$N = g\frac{L^3}{(2\pi\hbar)^3}\int_{p<p_F} d^{(3)}\mathbf{p} = g\frac{L^3}{(2\pi\hbar)^3}\frac{4\pi}{3}p_F^3\,.$$

In turn, one can isolate the Fermi momentum as a function of the density of the system $\rho = \frac{N}{L^3}$,

$$p_F = \left(\frac{6\pi^2}{g}\rho\right)^{1/3}\hbar\,.$$

In turn, the energy of the system can be computed analogously as the integral

$$E = g\frac{L^3}{(2\pi\hbar)^3}\int_{p<p_F} d^{(3)}\mathbf{p}\epsilon_p = g\frac{L^3}{(2\pi\hbar)^3}\frac{4\pi}{5}p_F^5 = \frac{3}{5}N\epsilon_F\,.$$

The intensive quantity, the energy per particle,

$$\frac{E}{N} = \frac{3}{5}\epsilon_F \tag{4.77}$$

is proportional to p_F^2 and hence to $\rho^{2/3}$. Again, we stress that this kinetic energy is purely due to quantum effects, since we have no temperature at all in these considerations. In fact, our treatment is only valid at zero temperature.

One can also compute the pressure from the derivative of the energy with respect to the volume, $V = L^3$,

$$P = -\left.\frac{\partial E}{\partial V}\right|_{N,T} = -\left.\frac{\partial E/N}{\partial V/N}\right|_{N,T} = -\left.\frac{\partial E/N}{\partial 1/\rho}\right|_{N,T} = \rho^2\left.\frac{\partial E/N}{\partial \rho}\right|_{N,T} \tag{4.78}$$

$$= \frac{3}{5}\frac{2}{3}\epsilon_F\rho = \frac{2}{5}\epsilon_F\rho. \tag{4.79}$$

Interestingly, the ideal gas relation $pV = \frac{2}{3}E$ is fulfilled, even though this system is not at all thermal!

5 Angular Momentum and Spin

5.1 Angular momentum

In classical mechanics, the orbital angular momentum $\mathbf{L}$ of a particle with momentum $\mathbf{p}$ with respect to the origin is defined as the vector product $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. In quantum mechanics, the very same angular momentum can be defined by introducing the operator $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$ This is a three-dimensional vector opertor with components $\hat{\mathbf{L}} = (\hat{L}_x, \hat{L}_y, \hat{L}_z)$. These do not commute with each other and instead we get

$$\left[\hat{L}_x, \hat{L}_y\right] = i\hbar\hat{L}_z, \quad \left[\hat{L}_y, \hat{L}_z\right] = i\hbar\hat{L}_x, \quad \left[\hat{L}_z, \hat{L}_x\right] = i\hbar\hat{L}_y. \tag{5.1}$$

To find the commutator $\left[\hat{L}_x, \hat{L}_y\right]$ use $\hat{L}_x = (\hat{\mathbf{r}} \times \hat{\mathbf{p}})_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y$ and $\hat{L}_y = (\hat{\mathbf{r}} \times \hat{\mathbf{p}})_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z$ to find $\left[\hat{L}_x, \hat{L}_y\right] = \left[\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z\right] = \left[\hat{y}\hat{p}_z, \hat{z}\hat{p}_x\right] + \left[\hat{z}\hat{p}_y, \hat{x}\hat{p}_z\right] = -i\hbar\hat{y}\hat{p}_x + i\hbar\hat{x}\hat{p}_y = i\hbar\hat{L}_z$.

One often uses $(1, 2, 3)$ to denote the space variables (x, y, z), and then the commutation relations can be written succinctly as

$$\left[\hat{L}_i, \hat{L}_j\right] = i\hbar\epsilon_{ijk}\hat{L}_k,$$

where ϵ_{ijk} is the Levi-Civita symbol.

Physically, the previous equations state that the different spatial components of the angular momentum cannot are not compatible observables. In other words, we cannot form a common basis for the three operators. We do know, however, that the observable $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$ commutes with all the individual angular momentum components,

$$\left[\hat{L}^2, \hat{\mathbf{L}}\right] = \mathbf{0}. \tag{5.2}$$

To show this, consider first the component $\hat{L}_x$, $\left[\hat{L}^2, \hat{L}_x\right] = \left[\hat{L}_y^2 + \hat{L}_z^2, \hat{L}_x\right]$. The first commutator yields $\left[\hat{L}_y^2, \hat{L}_x\right] = \hat{L}_y\left[\hat{L}_y, \hat{L}_x\right] + \left[\hat{L}_y, \hat{L}_x\right]\hat{L}_y = -i\hbar\hat{L}_y\hat{L}_z - i\hbar L_z\hat{L}_y$. The second commutator gives the exact same result with an opposite sign, $\left[\hat{L}_z^2, \hat{L}_x\right] = \hat{L}_z\left[\hat{L}_z, \hat{L}_x\right] + \left[\hat{L}_z, \hat{L}_x\right]\hat{L}_z = i\hbar\hat{L}_z\hat{L}_y + i\hbar L_y\hat{L}_z$, so that $\left[\hat{L}^2, \hat{L}_x\right] = 0$.

Starting from the operator $\hat{L}^2$ and *any* of the three independent components $\hat{L}_i$, we could construct a common basis. By convention, one chooses the third component $\hat{L}_z$. If we denote the eigenvalues of $\hat{L}^2$ by Λ_l and those of $\hat{L}_z$ by $|m\rangle$, we can defined a common basis, denoted by $|l, m\rangle$, such that

$$\hat{L}^2|l, m\rangle = \Lambda_l\hbar^2|l, m\rangle, \quad \hat{L}_z|l, m\rangle = m\hbar|l, m\rangle. \tag{5.3}$$

In keeping with previous sections, we could denote the state $|\Lambda_l, m\rangle$, labelling the eigenstate with the eigenvalue. It will become clear later on that l is a more adequate label for the eigenstate of $\hat{L}^2$. We note that we have introduced factors of $\hbar$ to keep the eigenvalues Λ_l and m dimensionless.

5.1.1 *Raising and lowering operators*

To do this we will need the angular momentum ladder operators, just as we did for the simple harmonic oscillator. These operators are defined as:

$$\hat{L}_+ = \hat{L}_x + i\hat{L}_y \tag{5.4}$$
$$\hat{L}_- = \hat{L}_x - i\hat{L}_y \tag{5.5}$$

To find the effect of $\hat{L}_\pm$, we look at their commutators with the operator $\hat{L}_z$. Let us start with $\hat{L}_+$:

$$\left[\hat{L}_z, \hat{L}_+\right] = \underbrace{\left[\hat{L}_z, \hat{L}_x\right]}_{i\hbar L_y} + i\underbrace{\left[\hat{L}_z, \hat{L}_y\right]}_{-i\hbar L_x} = i\hbar L_y + \hbar L_x = \hbar L_+ \,.$$

Similarly, for the lowering operator, one finds $\left[\hat{L}_z, \hat{L}_+\right] = -\hbar\hat{L}_-$. Consider now the action of $\hat{L}_+$ on an eigenstate:

$$\hat{L}_z\hat{L}_+|l,m\rangle = \hat{L}_+\hat{L}_z|l,m\rangle + \hbar\hat{L}_+|l,m\rangle = (m+1)\hbar\hat{L}_+|l,m\rangle \,. \tag{5.6}$$

Reading this expression from left to right, it becomes clear that $\hat{L}_+|l,m\rangle$ is an eigenvector of $\hat{L}_z$ with eigenvalue $(m+1)\hbar$. In other words, when $\hat{L}_+$ acts on an eigenstate of $\hat{L}_z$, m is raised by one unit,

$$\hat{L}_+|l,m\rangle = \mathcal{C}_{l,m}|l,m+1\rangle. \tag{5.7}$$

In complete analogy, $\hat{L}_-$ lowers m by one unit and hence

$$\hat{L}_-|l,m\rangle = \mathcal{D}_{l,m}|l,m-1\rangle. \tag{5.8}$$

It is important to see that $L_+|l,m\rangle = 0$ is also a valid option of Eq. (5.6).

5.1.2 *Eigenvalues*

From the definition of $\hat{L}^2$, we find that the expectation value of $\hat{L}^2$ is

$$\langle l,m|\hat{L}^2|l,m\rangle = \langle l,m|\hat{L}_x^2|l,m\rangle + \langle l,m|\hat{L}_y^2|l,m\rangle + \langle l,m|\hat{L}_z^2|l,m\rangle \,.$$

The expectation value of $\langle l,m|\hat{L}_x^2|l,m\rangle$ is the norm of a vector and hence necessarily positive, $\|L_x|l,m\rangle\|^2 \geq 0$. The same holds for the $y-$components, and hence the expectation values of $\hat{L}^2$ and L_z^2 are $\langle l,m|\hat{L}^2|l,m\rangle \geq \langle l,m|\hat{L}_z^2|l,m\rangle$. Employing the eigenvalue equations, Eqs. (5.3), one immediately finds

$$\Lambda_l \geq m^2 \,. \tag{5.9}$$

In other words, for a fixed value of l (and hence a fixed Λ_l), there must be a maximum and a minimum value of m. We denote the maximum value of m as l, so that $l = m_{\max}$. The condition to stop the tower of allowed eigenstates at the maximum must read

$$\hat{L}_+|l,m=l\rangle = 0 \,. \tag{5.10}$$

or otherwise the state that is generated would have $m = l + 1$ and hence overcome the bound of Eq. (5.9).

Let us now determine Λ_l. First, consider the product of L_+ and L_-:

$$\hat{L}_-\hat{L}_+ = \hat{L}_x^2 + \hat{L}_y^2 + i\left(\hat{L}_x\hat{L}_y - \hat{L}_y\hat{L}_x\right) = \hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z\,.$$

Applying this on the condition of Eq. (5.10), one finds:

$$\hat{L}_-\hat{L}_+|l, m = l\rangle = \hbar^2\left(\Lambda_l - l^2 - l\right) = 0 \Rightarrow \Lambda_l = l(l+1)\,.$$

Following the same reasoning that took us to the raising operator, we find that the lowering operator must be such that for $m = m_{\text{min}}$ one finds $\hat{L}_-|l, m_{\text{min}}\rangle = 0$. Using now the product $\hat{L}_+\hat{L}_- = \hat{L}^2 - \hat{L}_z^2 + \hbar\hat{L}_z$ on this state, we find the value of m_{min}:

$$\hat{L}_+\hat{L}_-|l, m_{\text{min}}\rangle = \hbar^2\left(l(l+1) - m_{\text{min}}^2 + m_{\text{min}}\right) = 0 \Rightarrow m_{\text{min}} = -l\,.$$

With this, we have shown the existence of a set of eigenvectors corresponding to integer-spaced values of m in the range $-l \leq m \leq l$. The difference between l and $-l$ must be an integer, and it follows that m must either be integer or half-integer. The allowed values of j and m are shown in the diagram of Fig. 5.1 and correspond to

$$\begin{aligned}
&j = 0, m = 0;\\
&j = \frac{1}{2}, m = -\frac{1}{2}, \frac{1}{2};\\
&j = 1, m = -1, 0, 1;\\
&j = \frac{3}{2}, m = -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2};\\
&\vdots
\end{aligned}$$

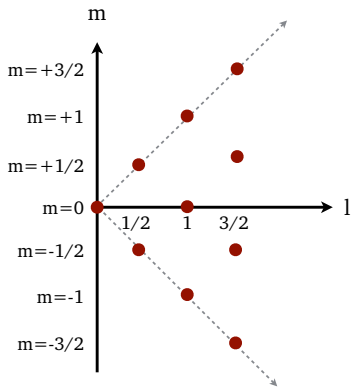

Figure 5.1: The filled circles represent the allowed values of m (vertical axis) for different values of l (horizontal axis).

We know that, starting with a state of definite m, $\hat{L}_\pm$ generate eigenstates with $m \pm 1$, see Eqs. (5.7) and (5.8). Let us determine the coefficient $\mathcal{C}_{l,m}$. We calculate the following norm,

$$\begin{aligned}
\|\hat{L}_+|l, m\rangle\|^2 &= \langle l, m|\hat{L}_-\hat{L}_+|l, m\rangle = \left|\mathcal{C}_{l,m}\right|^2\\
&= \langle l, m|\hat{L}^2 - \hat{L}_z^2 - \hbar L_z|l, m\rangle = l(l+1) - m(m+1)\,,
\end{aligned}$$

and hence find $\mathcal{C}_{l,m} = \sqrt{l(l+1) - m(m+1)}$. A similar calculation for the norm of $\hat{L}_-$ rather than $\hat{L}_+$ yields $\mathcal{D}_{l,m} = \sqrt{l(l+1) - m(m-1)}$. Note that, imoportantly, $\mathcal{C}_{l,m=l} = 0$ and $\mathcal{D}_{l,m=-l} = 0$, in keeping with previous results.

We choose the phase of $\mathcal{C}_{l,m}$ to be real for convenience.

5.1.3 *Quantum rotator*

A rigid body with moment of inertia I in classical mechanics has an associated kinetic energy $E = \frac{1}{2}I\omega^2 = \frac{L^2}{2I}$. This has a direct translation in quantum mechanics in terms of operators, so that the Hamiltonian of a rigidly rotating quantum object is $\hat{H} = \frac{\hat{L}^2}{2I}$. This is a simple operator that we can now easily solve with the tools of the previous subsection.

The eigenstates are the kets $|lm\rangle$ and the eigenstates,

$$\hat{H}|lm\rangle = E_l|lm\rangle \Rightarrow E_l = \frac{\hbar^2}{2I}l(l+1)\,. \tag{5.11}$$

Because the eigenvalues do not depend on m, the eigenvectors are $(2l+1)-$degenerate. These levels are shown in Fig. 5.2. A characteristic feature of a rotator's spectrum is the fact that the spacing of the levels increases by one unit for every level of l. In other words, if the ground state and the first excited state are separated by δE_1, the second excited state is $2\delta E_1$ units away from the first excited state. These types of spectra are ubiquitous in quantum systems - including nuclei, molecules and even hadron themselves.

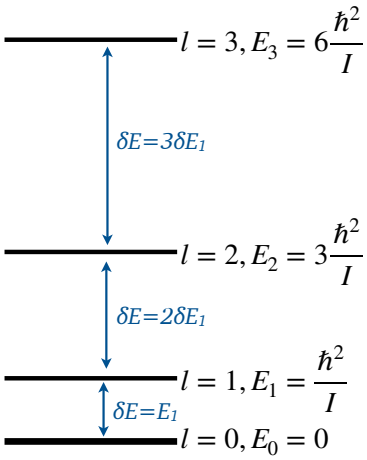

Figure 5.2: The spectrum of a rigidly rotating quantum object from $l=0$ to $l=3$.

5.2 *Spherical harmonics*

We now try to find explicit expressions for the eigenstates $|l,m\rangle$ in position space. It is convenient to work in spherical polar coordinates, (r,θ,ϕ), which are defined in terms of one magnitude and two angles. The magnitude is given by the total length along the radial direction, r. The azimuthal angle, ϕ, describes the distance of the vector from the $x-$axis in the $xy-$plane. The polar angle, θ, provides the inclination of the vector with respect to the zenith direction, traditionally taken along the $z-$axis

Fig. 5.3 provides a graphical illustration of these coordinates. The transform from spherical into Cartesian coordinates can be worked out by successive projections of the angles and reads:

$$x = r\cos\phi\sin\theta\,, \quad y = r\sin\phi\sin\theta\,, \quad z = r\cos\theta\,.$$

In spherical coordinates, the gradient of a field $\psi(r,\theta,\phi)$ reads

$$\nabla\psi = \frac{\partial\psi}{\partial r}\boldsymbol{e_r} + \frac{1}{r}\frac{\partial\psi}{\partial\theta}\boldsymbol{e_\theta} + \frac{1}{r\sin\theta}\frac{\partial\psi}{\partial\phi}\boldsymbol{e_\phi}\,.$$

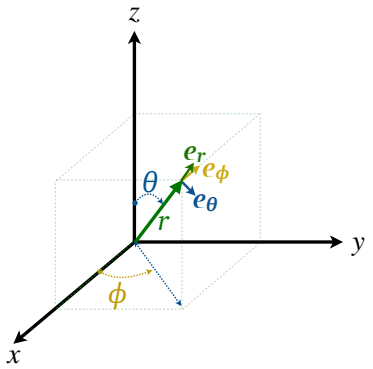

Figure 5.3: Spherical coordinates obtained from 3D Cartesian coordinates.

The spherical coordinates form a new orthonormal coordinate set, with associated unit vectors $\boldsymbol{e_r}$, $\boldsymbol{e_\theta}$ and $\boldsymbol{e_\phi}$ that point along the direction of increasing variables r, θ and ϕ. These unit vectors happen to be orthonormal so that, $\boldsymbol{e_r}\times\boldsymbol{e_\theta} = \boldsymbol{e_\phi}$, $\boldsymbol{e_\theta}\times\boldsymbol{e_\phi} = \boldsymbol{e_r}$ and $\boldsymbol{e_\phi}\times\boldsymbol{e_r} = \boldsymbol{e_\theta}$. The transformation of unit vectors is given by

$$\begin{pmatrix}\boldsymbol{e_r}\\ \boldsymbol{e_\theta}\\ \boldsymbol{e_\phi}\end{pmatrix} = \begin{pmatrix}\sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta\\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta\\ -\sin\phi & \cos\phi & 0\end{pmatrix}\begin{pmatrix}\mathbf{i}\\ \mathbf{j}\\ \mathbf{k}\end{pmatrix}\,, \tag{5.12}$$

With these, we can compute the angular momentum operator in polar coordinates:

$$\hat{\mathbf{L}} = \mathbf{r}\times\mathbf{p} = r\boldsymbol{e_r}\times(-i\hbar\nabla) = (-i\hbar)\left[\boldsymbol{e_\phi}\frac{\partial}{\partial\theta} - \boldsymbol{e_\theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}\right]\,.$$

We are looking for the eigenvectors of the two operators:

$$\hat{L}_z = \hat{\mathbf{L}}\cdot\boldsymbol{e_k} = -i\hbar\frac{\partial}{\partial\phi}\,, \tag{5.13}$$

and

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

We must solve the two couple differential equations for the eigenvalues $\langle \theta, \phi | l, m \rangle = Y(\theta, \phi)$,

$$\hat{L}_z Y(\theta,\phi) = \hbar m Y(\theta,\phi),$$
$$\hat{L}^2 Y(\theta,\phi) = \hbar^2 l(l+1) Y(\theta,\phi) .$$

Using Eq. (5.13) on the first equation above yields $\frac{\partial}{\partial\phi} Y = imY$ and hence

$$\frac{\partial^2 Y}{\partial\phi^2} = -m^2 Y. \tag{5.14}$$

Using this in the second equation, and multiplying by $\sin^2\theta$, we find

$$\sin\theta \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial Y}{\partial\theta} \right) + -m^2 Y = l(l+1)\sin^2\theta Y .$$

The two previous equations decouple the dependence on θ and ϕ, and hence we may take $Y(\theta,\phi) = F(\theta)G(\phi)$. The solution for $G(\phi)$ is trivial and such that $G_m(\phi) = e^{im\phi}$. We expect that a rotation of the system by 2π should leave the system unchanged, and hence we impose that the solutions are *single-valued* under rotation, $G_m(\phi + 2\pi) = G_m(\phi)$. This immediately implies that m should be integer. The standard solution of the second equation, which is an *associated Legendre equation*, is denoted $P_l^m(\cos\theta)$. One can prove that if the solutions are to be non-singular at $\theta = 0$ and $\theta = \pi$, then l must be a nonnegative integer, and such that relation $l \geq |m|$ is fulfilled.

The normalized solutions to the equations above are the *spherical harmonics*

$$Y_{l,m}(\theta,\phi) = (-1)^{\frac{m+|m|}{2}} \left[\frac{(2l+1)(l-|m|)!}{4\pi(l+|m|)!} \right] e^{im\phi} P_l^{|m|}(\cos\theta) , \tag{5.15}$$

where the Legendre polynomials are defined as

$$P_l^m(u) = \frac{(-1)^m}{2^l l!} \left(1-x^2\right)^{m/2} \frac{d^{l+m}}{du^{l+m}} \left(u^2-1\right)^l .$$

The arbitrary phase is chosen so that the following relationship holds:

$$Y_{l,-m}(\theta,\phi) = (-1)^m \left[Y_{l,m}(\theta,\phi) \right]^* . \tag{5.16}$$

Furthermore, they fulfil the orthonormality condition

$$\int_0^\pi \mathrm{d}\theta \, \sin\theta \int_0^{2\pi} \mathrm{d}\phi Y^*_{l,m}(\theta,\phi) Y_{l',m'}(\theta,\phi) = \delta_{l,l'}\delta_{m,m'} . \tag{5.17}$$

Under a parity transformation $(x,y,z) \to (-x,-y,-z)$, which in spherical coordinates reads $(r,\theta,\phi) \to (r,\theta-\pi,\phi+\pi)$, the spherical harmonics transform such that

$$Y_{l,m}(\theta,\phi) \to Y_{l,m}(\theta-\pi,\phi+\pi) = (-1)^l Y_{l,m}(\theta,\phi) . \tag{5.18}$$

In other words, spherical harmonics are of even parity for even l and of odd parity for odd values of l.

A few examples include:

$$Y_{0,0}(\theta,\phi) = \sqrt{\frac{1}{4\pi}},$$
$$Y_{1,0}(\theta,\phi) = \sqrt{\frac{3}{4\pi}}\cos\theta,\ Y_{1,\pm1}(\theta,\phi) = \mp\sqrt{\frac{3}{4\pi}}\sin\theta e^{\pm i\phi},$$
$$Y_{2,0}(\theta,\phi) = \sqrt{\frac{5}{16\pi}}(\cos^2\theta - 1),\ Y_{2,\pm1}(\theta,\phi) = \mp\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{\pm i\phi},$$
$$Y_{2,\pm2}(\theta,\phi) = \sqrt{\frac{15}{32\pi}}\sin^2\theta e^{\pm i2\phi}, \cdots$$

5.3 Particle in a central potential

Spherical harmonics are particularly useful when it comes to decoupling radial variables from angular ones. In problems where the potential is central, $V(\mathbf{r}) = V(r)$, this separation is achieved precisely by this means - so let us review this quickly. Consider the 3D Hamiltonian $\hat{H} = \frac{\hat{p}^2}{2m} + V(r)$. The Schrödinger equation is written:

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

It happens that $\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi}$ commutes with $V(r)$, but also with $\hat{p}^2$. The latter relation is a bit tricky to obtain, but important because it indicates that $\hat{L}_z$ commutes with $\hat{H}$.

We can define the radial component of the momentum as $\hat{p}_r = -i\hbar\frac{1}{r}\frac{\partial}{\partial r}r$, which is Hermitian unlike the "naive" counterpart $-i\hbar\frac{\partial}{\partial r}$. One can also prove[1] the operator identity $\hat{p}^2 = \hat{p}_r^2 + \frac{\hat{L}^2}{r^2}$. This provides the following expression for the Hamiltonian,

$$\left[\frac{p_r^2}{2m} + \frac{\hat{L}^2}{2mr^2} + V(r)\right]\psi(r,\theta,\phi) = E\psi(r,\theta,\phi)\,.$$

$\hat{H}$ commutes with both $\hat{L}^2$ and $\hat{L}_z$, indicating that we can build a common basis for these operators. We can thus solve the equation in two steps. We first solve for the eigenfunctions of $\hat{L}^2$. In this step, r is nothing but a parameter that can be ignored. Second, we solve the Schrödinger equation using the eigenfunctions in the angular variables. The form of $V(r)$ is only relevant for the second step.

The second step requires forming the common eigenfunctions of $\hat{H}$, $\hat{L}^2$ and $\hat{L}_z$. We guess that the solutions of the Schrödinger equation must be of the type

$$\psi_{n,l,m}(r,\theta,\phi) = R_{nl}(r)Y_{lm}(\theta,\phi)\,,$$

where l and m are quantum numbers associated to the angular dependence, whereas n labels the quantum numbers of the radial function. The functions $R_{nl}(r)$ are a solution of the equation

$$\left[-\frac{1}{2m}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{\hbar^2 l(l+1)}{2mr^2} + V(r)\right]R_{nl}(r) = ER_{nl}(r)\,,$$

[1] A. Messiah. *Quantum Mechanics*. Dover Press, 1999

The radial momentum on an arbitrary wavefuntion is

$$\hat{p}_r\psi = -i\hbar\frac{1}{r}\frac{\partial}{\partial r}r\psi = -i\hbar\frac{\psi}{r} + \frac{\partial}{\partial r}\psi$$

The square of this operator yields,

$$\hat{p}_r^2\psi = -\hbar^2\frac{\partial^2}{\partial r^2}\psi + \frac{2}{r}\frac{\partial}{\partial r}\psi\,,$$

which coincides with radial component of the laplacian in spherical coordinates,

$$\nabla^2\psi(r) = \frac{1}{r}\frac{\partial^2}{\partial r^2}r\psi(r)\,.$$

Since the spherical harmonics are normalized, the function R_{nl} is normalized such that

$$\int_0^\infty \mathrm{d}r r^2 \, |R_{nl}(r)|^2 = 1$$

This suggests that it is convenient to rewrite this equation in terms of the reduced wave function $u_{nl} = rR_{nl}$, so that the normalization becomes $\int_0^\infty \mathrm{d}r \, |u_l(r)|^2 = 1$. With this new variable equation reads

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2} + \frac{\hbar^2 l(l+1)}{2mr^2} + V(r) \right] u_{nl}(r) = E u_{nl}(r) \, . \tag{5.19}$$

One typically demands $u_{nl}(0) = 0$ at the origin. Solving this equation imposing this condition is what typically brings in the quantum number n (think of the radial quantum number n of the hydrogen atom).

It is important to stress that the introduction of the spherical harmonics allow one to reduce the original three-dimensional problem to an effective one-dimensional equation on the radial variable, r. In fact, if one adopts the new effective potential $\mathcal{V}_l(r) = l(l+1)\frac{\hbar^2}{2mr^2} + V(r)$, Eq. (5.19) is nothing but a one-dimensional Schrödinger equation for r,

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dr^2} + \mathcal{V}_l(r) \right] u_{n,l}(r) = E_{n,l} u_{n,l}(r) \, . \tag{5.20}$$

This is a substantial simplification compared to the original, three-dimensional starting point.

5.4 *Spin*

In addition to its orbital angular momentum, a particle may also have another form of angular momentum known as a *spin*. This is an intrinsic particle property (eg it does not depend on particle motion). Successive experiments have demonstrated that indeed fundamental fermionic particles like the electron, the neutron or the proton all have a spin quantum number that can be described like a half-integer angular momentum with $l = \frac{1}{2}$.

There are of course more complicated fermionic particles with spin $S = \frac{3}{2}$, like the Δ baryon. Bosons, like the pion π ($S = 1$), have pure integer spins.

We denote the vector operator spin as $\hat{\mathbf{S}}$, and it has three components $\hat{S}_x$, $\hat{S}_y$ and $\hat{S}_z$, satisfying the commutation relations $[\hat{S}_i, \hat{S}_j] = i\epsilon_{ijk}\hbar$. The eigenvalue equations for $\hat{S}^2$ and S_z are, just like in the angular momentum case,

$$\hat{S}^2|s,m\rangle = \hbar^2 s(s+1)|s,m\rangle, \quad \hat{S}_z|s,m\rangle = m\hbar|s,m\rangle. \tag{5.21}$$

Because a given species is characterized by the value of its spin s, it is usually sufficient to treat the spin operators $\hat{S}_x$, $\hat{S}_y$ and $\hat{S}_z$ acting on a space of dimension $2s+1$.

Let us treat the usual case of $s = \frac{1}{2}$. It is customary to define $\hat{\mathbf{S}} = \frac{1}{2}\hbar\hat{\boldsymbol{\sigma}}$, where $\hat{\sigma}_i$ are called Pauli spin operators. Explicit matrix

representations can be deduced in the basis formed by the eigenvectors of Eq. (5.21). These are given by the expressions:

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.22}$$

With the changes $x \to 1$, $y \to 2$ and $z \to 3$, one finds that these matrices satisfy the following important, frequently used relations:

$$\hat{\sigma}_i^2 = I, \quad [\hat{\sigma}_i, \hat{\sigma}_j] = 2i\epsilon_{ijk}\hat{\sigma}_k, \quad \{\hat{\sigma}_i, \hat{\sigma}_j\} = 2\delta_{ij}. \tag{5.23}$$

These can be written succinctly in the following matrix representation

$$\hat{\sigma}_i\hat{\sigma}_j = \delta_{ij} + i\epsilon_{ijk}\hat{\sigma}_k\,. \tag{5.24}$$

Do not confuse the subindex $i = x, y, z$ and the complex number i in front of the Levi-Civita symbol!

For two vectors **a** and **b**, this relation leads to

$$(\hat{\boldsymbol{\sigma}} \cdot \mathbf{a})\,(\hat{\boldsymbol{\sigma}} \cdot \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} + i\boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b})\,. \tag{5.25}$$

The Pauli matrices together with the identity matrix I form a basis for the vector space of matrices on the Hilbert space of dimension 2. Any 2×2 matrix A can be decomposed as

$$A = \lambda_0 I + \boldsymbol{\lambda} \cdot \hat{\boldsymbol{\sigma}}\,. \tag{5.26}$$

where the coefficients are

$$\lambda_0 = \frac{1}{2}\mathrm{Tr}A, \quad \lambda_i = \frac{1}{2}\mathrm{Tr}\left[A\hat{\sigma}_i\right]\,. \tag{5.27}$$

The coefficients λ_0 and λ_i are real so long as the matrix is Hermitian, $A = A^\dagger$. Since the Pauli matrices form a basis for matrices in any Hilbert space of dimension 2, they are often used in two-dimensional problems even if the physics has nothing to do with spin $s = \frac{1}{2}$. A classic two-level system is the two-level atom or, in its many-body context, the Lipkin model. We say that all dimension 2 Hilbert spaces are isomorphic to the spin $s = \frac{1}{2}$ case.

In a two-dimensional Hilbert space, the eigenstates of the spin Hamiltonian are:

$$|+\rangle \equiv \left|s = \frac{1}{2}, m = +\frac{1}{2}\right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{5.28}$$

$$|-\rangle \equiv \left|s = \frac{1}{2}, m = -\frac{1}{2}\right\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{5.29}$$

The most general vector in this space will be a linear combination of these two states, $|\xi\rangle = \xi_+|+\rangle + \xi_-|\xi-\rangle$. From the spectral decomposition of the spin operator $\hat{S}_z$, Eq. (1.26), the matrix σ_z can be seen as the difference of the projectors into the $|+\rangle$ and the $|-\rangle$ states,

$$\sigma_z = |+\rangle\langle+| - |-\rangle\langle-| = P_+ - P_-\,. \tag{5.30}$$

In general, however, the spin variable does not have to be aligned with the $z-$axis. There must exist an operator $\hat{S}_{\hat{n}} = \mathbf{S} \cdot \hat{\mathbf{n}}$, the spin

projector on the $\hat{n}$ axis, with eigenvalues $\pm\frac{\hbar}{2}$ and which takes the same form in the basis $\{|+,\hat{\mathbf{n}}\rangle, |-,\hat{\mathbf{n}}\rangle\}$. In this basis, the operator is written as

$$\hat{S}_{\hat{n}} \equiv \hat{\mathbf{S}} \cdot \hat{\mathbf{n}} = \frac{\hbar}{2} \left(|+,\hat{\mathbf{n}}\rangle\langle +,\hat{\mathbf{n}}| - |-,\hat{\mathbf{n}}\rangle\langle -,\hat{\mathbf{n}}| \right) . \tag{5.31}$$

The matrix representing this operator should be the most general Hermitian 2×2 matrix with eigenvalues $\pm\frac{\hbar}{2}$, reading

$$\hat{S}_{\hat{n}} = \frac{\hbar}{2} \begin{pmatrix} a & b \\ b^* & c \end{pmatrix} , \tag{5.32}$$

where a and c are real numbers. The equation for the eigenvalues is $\lambda^2 - (a+c)\lambda + ac - |b|^2 = 0$. Since $\lambda_+ + \lambda_- = 0$ and $\lambda_+\lambda_- = -1$, we must have $a + c = 0$ and $ac - |b|^2 = -1$. This implies $a^2 + |b|^2 = 1$ and suggests that these constants can be parametrized in terms of two angles α and β so that $a = \cos\beta$ and $b = e^{-i\alpha}\sin\beta$, so that

$$\hat{S}_{\hat{n}} = \frac{\hbar}{2} \begin{pmatrix} \cos\beta & e^{-i\alpha}\sin\beta \\ e^{+i\alpha}\sin\beta & -\cos\beta \end{pmatrix} . \tag{5.33}$$

The eigenvectors are then given by

$$|+,\hat{\mathbf{n}}\rangle = \begin{pmatrix} e^{-i\alpha/2}\cos\beta/2 \\ e^{i\alpha/2}\sin\beta/2 \end{pmatrix} , \quad |-,\hat{\mathbf{n}}\rangle = \begin{pmatrix} -e^{-i\alpha/2}\sin\beta/2 \\ e^{i\alpha/2}\cos\beta/2 \end{pmatrix} . \tag{5.34}$$

One can prove with some additional algebra that the angles $\alpha = \theta$ and $\beta = \phi$ correspond to the angles in spherical coordinates!

6 Rabi oscillations

6.1 Introduction and hamiltonian

We introduce here an exactly solvable time-dependent model. This time-dependent model has only 2 levels, and it can be solved with the spin formalism that we have just introduced. Moreover, the exact dynamical solution turns out to be useful when it comes to discussing time-dependent perturbation theory.

The exact model we shall solve is that of Rabi oscillations. Rabi oscillations are relevant in the context of magnetic resonances, which amplify experimental signal and allow for precise measurements of magnetic moments. In quantum computing, Rabi oscillations are also important in the sense that they allow for a full conversion of spin populations (and, hence, of qubit states). In this model, one places a spin $1/2$ magnetic moment into a known constant magnetic field, B_0, pointing along the $z-$axis. On top of this, one adds a time-dependent magnetic field B_1 that rotates at a variable angular frequency ω in the $xy-$plane. In other words, the field is $\mathbf{B}(t) = B_1 \cos\omega t\mathbf{i} + B_1 \sin\omega t\mathbf{i} + B_0\mathbf{k}$. We shall discuss the time evolution of spin states under such fields. As we have stressed earlier, spin and two-state systems are analogous, and we expect this description to be applicable to other time-dependent 2-level systems.

A magnetic dipole moment $\boldsymbol{\mu}$ under a magnetic field $\boldsymbol{B}$ has a hamiltonian $\hat{H} = -\boldsymbol{\mu}\cdot\boldsymbol{B}$. If we ignore angular momentum, the magnetic dipole moment is due entirely to the spin of the particle, $\boldsymbol{\mu} = \gamma\hat{\mathbf{S}}$, where γ is the gyromagnetic ratio of the particle. For a particle of charge q and mass m, this ratio is typically expressed as $\gamma = g\frac{q}{2m}$, with g a dimensionless number called $g-$factor. Since $\hat{\mathbf{S}} = \frac{\hbar}{2}\hat{\boldsymbol{\sigma}}$, we can also write $\hat{H} = -\frac{\hbar\gamma}{2}\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{B} = -\frac{g}{2}\frac{q\hbar}{2m}\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{B}$. For electrons, such as in Rabi's experiment, $g \approx -2$ and the prefactor in front of the magnetic field is numerically equivalent to μ_B, the so-called *Bohr magneton*, $\mu_0 = g\frac{q}{2m}\frac{\hbar}{2} \approx (-2)\frac{-1e}{2m}\frac{\hbar}{2} = \frac{e\hbar}{2m_e} = \mu_B$.

All in all, the Hamiltonian reads as the sum of a time-independent term, proportional to B_0, and a time-dependent contribution, proportional to B_1,

$$\hat{H} = \hat{H}_0 + \hat{H}_1(t) = -\mu_0 B_0\hat{\sigma}_z - \mu_0 B_1\cos\omega t\hat{\sigma}_x - \mu_0 B_1\sin\omega t\hat{\sigma}_y\,. \quad (6.1)$$

One can define frequencies associated to the magnetic moment and

the two magnetic field components,

$$\frac{\hbar\omega_0}{2} = -\mu_0 B_0, \quad \frac{\hbar\omega_1}{2} = -\mu_0 B_1 .$$

and reexpress the hamiltonian as

$$\hat{H} = \hat{H}_0 + \hat{H}_1(t) = \frac{\hbar\omega_0}{2}\hat{\sigma}_z + \frac{\hbar\omega_1}{2}\cos\omega t\hat{\sigma}_x + \frac{\hbar\omega_1}{2}\sin\omega t\hat{\sigma}_y . \quad (6.2)$$

Importantly, the following solution applies for any Hamiltonian of this form. If, for instance, the oscillations were to be generated with time-dependent electric fields, this may yield a different definition for $\hbar\omega_1$, but would not change the hamiltonian itself.

6.2 Exact solution

The time-dependent evolution of the wavefunction is described by two time-dependent coefficients $a_\pm(t)$, such that

$$|\psi(t)\rangle = a_+(t)|+\rangle + a_-(t)|-\rangle .$$

To find the time evolution of these coefficients, we use the time-dependent Schrödinger equation:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = i\hbar\dot{a}_+(t)|+\rangle + i\hbar\dot{a}_-(t)|-\rangle$$
$$= \left[\hat{H}_0 + \hat{H}_1(t)\right]|\psi(t)\rangle = \left[\hat{H}_0 + \hat{H}_1(t)\right](a_+(t)|+\rangle + a_-(t)|-\rangle)$$

To find the action of the hamiltonian on the states in the second line, we use the relations

$$\sigma_x|\pm\rangle = |\mp\rangle, \quad \sigma_y|\pm\rangle = \pm i|\mp\rangle, \quad \sigma_z|\pm\rangle = \pm|\pm\rangle,$$

and obtain

$$\left[\hat{H}_0 + \hat{H}_1(t)\right]|+\rangle = \frac{\hbar\omega_0}{2}|+\rangle + \frac{\hbar\omega_1}{2}\cos\omega t|-\rangle + \frac{\hbar\omega_1}{2}i\sin\omega t|-\rangle$$
$$= \frac{\hbar\omega_0}{2}|+\rangle + \frac{\hbar\omega_1}{2}e^{i\omega t}|-\rangle ,$$
$$\left[\hat{H}_0 + \hat{H}_1(t)\right]|-\rangle = \frac{\hbar\omega_0}{2}|-\rangle + \frac{\hbar\omega_1}{2}\cos\omega t|+\rangle - \frac{\hbar\omega_1}{2}i\sin\omega t|+\rangle$$
$$= \frac{\hbar\omega_0}{2}|-\rangle + \frac{\hbar\omega_1}{2}e^{-i\omega t}|+\rangle .$$

Projecting on the bras $\langle+|$ and $\langle-|$, we find the dynamical equations for $a_\pm$:

$$i\dot{a}_+(t) = \frac{\omega_0}{2}a_+(t) + \frac{\omega_1}{2}e^{-i\omega t}a_-(t) , \quad (6.3)$$
$$i\dot{a}_-(t) = \frac{\omega_1}{2}e^{i\omega t}a_+(t) + \frac{\omega_1}{2}a_-(t) . \quad (6.4)$$

This is a coupled set of differential equations. The system can be decoupled by using a change of functions $b_\pm(t) = e^{\pm i\frac{\omega t}{2}}a_\pm(t)$. With the relations $\dot{a}_\pm(t) = e^{\mp i\frac{\omega t}{2}}\left[\dot{b}_\pm(t) \mp \frac{i\omega}{2}b_\pm(t)\right]$, we find

$$i\dot{b}_+(t) = \frac{\omega_0 - \omega}{2}b_+(t) + \frac{\omega_1}{2}b_-(t) , \quad (6.5)$$
$$i\dot{b}_-(t) = \frac{\omega_1}{2}b_+(t) - \frac{\omega_0 - \omega}{2}b_-(t) . \quad (6.6)$$

The change from $a_\pm$ to $b_\pm$ can also be seen as a change of reference frame. In this case, we go from the laboratory frame, where $a_\pm$ are defined, to a rotating frame that co-moves with the magnetic field B_1 around the $z-$axis. This is known as a *rotating wave approximation*, whichis useful in magnetic systems of this kind. Note that here the approximation in fact leads to the exact result.

While these may not look simpler, the set of equations above can be simplified by looking not at the first, but rather at the second time derivatives. In this case, one finds the common equation

$$\ddot{b}_{\pm}(t) + \left(\frac{\Omega}{2}\right)^2 b_{\pm}(t) = 0\,, \tag{6.7}$$

where the frequency Ω is given by

$$\Omega^2 = (\omega - \omega_0)^2 + \omega_1^2\,. \tag{6.8}$$

These equations are two simple, independent harmonic equations with solutions,

$$b_{\pm}(t) = A_{\pm} e^{-i\frac{\Omega}{2}t} + B_{\pm} e^{+i\frac{\Omega}{2}t}, \tag{6.9}$$

which need to fulfill the normalization condition $|b_+(t)|^2 + |b_-(t)|^2 = |a_+(t)|^2 + |a_-(t)|^2 = 1$.

6.3 *Transition probability*

We have found the most general solution for the coupled set of equations that describe the dynamics of the system. Let us consider now one specific example. Suppose we perform a measurement of the electron spin before we switch the time dependence of the hamiltonian (eg at $t < 0$). Let us assume that the measurement provided an initial state $|+\rangle$. The boundary conditions associated to this physical situation correspond to so that $b_-(t = 0) = 0$ and $b_+(t = 0) = 1$. This immediately implies the relations $A_- = -B_-$ and $B_+ = 1 - A_+$, which lead to

$$b_-(t) = A_- \left(e^{-i\frac{\Omega}{2}t} - e^{+i\frac{\Omega}{2}t}\right) = 2iA_- \sin\frac{\Omega}{2}t$$

On the one hand, taking a time derivative and looking at $t = 0$, this yields

$$\dot{b}_-(0) = 2iA_- \frac{\Omega}{2}\,.$$

On the other hand, we can now use Eq. (6.6) at $t = 0$ to find

$$i\dot{b}_-(0) = \frac{\omega_1}{2} b_+(0) - \frac{\omega_0 - \omega}{2} b_-(0) = \frac{\omega_1}{2} \Rightarrow \dot{b}_-(0) = -i\frac{\omega_1}{2}$$

Using both equations, we find that $A_- = -\frac{\omega_1}{2\Omega}$ and hence the solution

$$b_-(t) = -\frac{i\omega_1}{\Omega} \sin\left(\frac{\Omega t}{2}\right)\,. \tag{6.10}$$

For reference, the solution for b_+ reads $b_+(t) = \cos\left(\frac{\Omega t}{2}\right) + i\frac{\omega-\omega_0}{\Omega}\sin\left(\frac{\Omega t}{2}\right)$.

In this set-up, we started with the system in a $|+\rangle$ state, and the transition probability to the $|-\rangle$ state is then given by:

$$\begin{aligned}\mathcal{P}_{+\to-}(t) &= |\langle -|\psi(t)\rangle|^2 = |a_-(t)|^2 = |b_-(t)|^2 \\ &= \frac{\omega_1^2}{(\omega - \omega_0)^2 + \omega_1^2} \sin^2\left(\frac{\Omega t}{2}\right)\,.\end{aligned} \tag{6.11}$$

This formula exhibits a resonance phenomenon. In other words, depending on the relation between ω, ω_0 and ω_1, the transition probability has a very different behaviour. We can identify two clearly different regimes.

When $|\omega - \omega_0| \gg \omega_1$, the denominator in Eq. (6.11) is large and the transition probability is small. In this regime, the probability oscillates with frequency $\Omega \approx |\omega - \omega_0|$. The probability of measuring the $|-\rangle$ state is also low *at all times*. An example of this behaviour is shown in the solid lines of the top panel of Fig. 6.1.

A very different regime occurs when $\omega = \omega_1$. Here, the transition probability reduces to $\mathcal{P}_{+\to-}(t) = \sin^2\left(\frac{\omega_1 t}{2}\right)$, and oscillates with frequency $\Omega = \omega_1$. The probability of a spin flip is *exactly one* at times $t_n = (2n+1)\frac{\pi}{\omega_1}$. This happens independently of the amplitude of the field B_1. In other words, even if B_1 is very small, the spin population will flip. An example of this behaviour is shown in the bottom panel of Fig. 6.1.

Rabi oscillations and magnetic resonances are extremely important quantum phenomena. Rabi was awarded the Nobel prize for the experimental realization of this method, which allowed for unprecedentedly accurate measurements of the proton magnetic moment. This experiment is also a precursor of nuclear magnetic resonance (NMR) techniques of widespread use in biophysics and medicine.

In quantum computing, one works with a two-state qubit $|0\rangle$ and $|1\rangle$. If a Rabi oscillation setting can be engineered at resonance, one can use the spin-flip transition to transform a $|0\rangle$ qubit into a $|1\rangle$ qubit. This can be obtained by letting a B_1 perturbation act for a period $t_0 = \frac{\pi}{\omega_1}$, which is a so-called $\pi-$pulse.

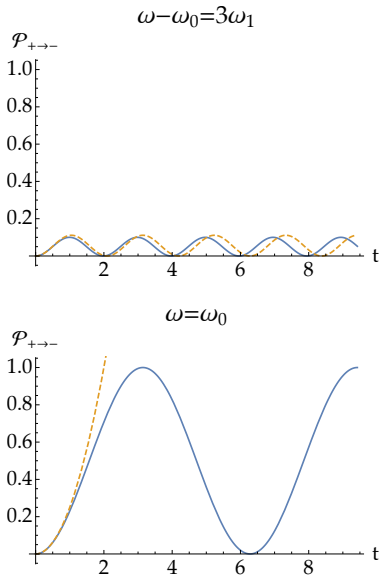

Figure 6.1: Transition probability for a Rabi oscillation as a function of time for two different cases: far away from resonance, $\omega - \omega_0 = 3\omega_1$ (top panel), and at resonance $\omega = \omega_0$ (bottom panel). ω_1 is arbitrarily set to one for illustration purposes. The dashed line correspond to the first-order perturbation theory results of Eqs. (7.54) and (7.56).

7 Approximation Methods

Up until now, we have solved analytical models that provide exact solutions of the systems at hand. In real applications of quantum mechanics, it is rare to have such exact solutions. In addition to useful numerical methods, we can try to find approximate solutions by analytical methods. This chapter is concerned with methods that allow us to find such approximate solutions - in static, time-independent configurations, but also in dynamic, time-dependent situations.

7.1 Time-independent perturbation theory

We discuss first the more standard case of perturbation theory for a time-independent setting. The starting point is a separation of the full Hamiltonian of the problem, $\hat{H}$, into two sub-pieces,

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}_1 \,. \tag{7.1}$$

Typically, the term $\hat{H}_0$ is solvable, meaning that we have access to the eigenstates and eigenvalues

$$\hat{H}_0 |n\rangle = \epsilon_n |n\rangle \,. \tag{7.2}$$

We call this the "unperturbed" part of the Hamiltonian[1]. The term $\lambda \hat{H}_1$ is then treated as a *perturbation* on top of $\hat{H}_0$, meaning that it is assumed to be small in some sense that we shall clarify later on. In some cases, λ is a variable parameter (the strength of an external field) which can be made small at will. In others, it may be a fixed parameter, governed by some physical mechanism (the spin-orbit interactions). And yet in other occasions it may introduced for mathematical convenience to switch from the $\lambda = 0$ unperturbed case to the fully perturbed $\lambda = 1$ case.

[1] Sometimes this is also referred to as the reference Hamiltonian.

We denote the solutions of the full Hamiltonian in Eq. (7.1) as

$$\hat{H} |\phi_n\rangle = E_n |\phi_n\rangle \,. \tag{7.3}$$

In perturbation theory, the unknown quantities E_n and ϕ_n are expanded in powers of the small parameter λ

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots \tag{7.4}$$

$$|\phi_n\rangle = |\phi_n^{(0)}\rangle + \lambda |\phi_n^{(1)}\rangle + \lambda^2 |\phi_n^{(2)}\rangle + \cdots \tag{7.5}$$

and explicit expressions for the different coefficients $E_n^{(j)}$ and $\phi_n^{(j)}$ at $j-$th order are obtained. Before discussing how these coefficients can be obtained, we look at the simplest non-trivial case: a two-dimensional Hilbert space. This helps us set the scene for the more complete developments that follow.

7.1.1 Two-level system: exact results

Consider the case in which the unperturbed Hamiltonian has only two known eigenstates,

$$\hat{H}_0|1\rangle = \epsilon_1|1\rangle\,, \quad \hat{H}_0|2\rangle = \epsilon_2|2\rangle\,. \tag{7.6}$$

When the perturbation $\lambda\hat{H}_1$ is added, the dimensionality of the system does not change. There are two solutions to the Schrödinger equation with the full Hamiltonian, $|\phi_1\rangle$ and $|\phi_2\rangle$. These solutions can be expanded in the states of the reference Hamiltonian,

$$|\phi_i\rangle = c_{i1}|1\rangle + c_{i2}|2\rangle\,. \tag{7.7}$$

If we find c_{ij}, we have solved the problem.

In principle, i runs over two indexes that, in this subsection, we shall name $+$ and $-$. We can substitute the full solution back into the Schrödinger equation,

$$\begin{aligned}(\hat{H}_0 + \lambda\hat{H}_1)|\phi_\pm\rangle &= c_{\pm1}(\epsilon_1 + \lambda\hat{H}_1)|1\rangle + c_{\pm2}(\epsilon_2 + \lambda\hat{H}_1)|2\rangle &(7.8)\\ &= c_{\pm1}E|1\rangle + c_{\pm2}E|2\rangle\,. &(7.9)\end{aligned}$$

Taking the product with the two bras $\langle 1|$ and $\langle 2|$, we get the two coupled equations:

$$\begin{aligned}c_{\pm1}(\epsilon_1 + \lambda H_{11}) + c_{\pm2}\lambda H_{12} &= c_{\pm1}E\\ c_{\pm1}\lambda H_{21} + c_{\pm2}(\epsilon_2 + \lambda H_{22}) &= c_{\pm2}E\,,\end{aligned}$$

where we have introduced the matrix elements of $\hat{H}_1$ on unperturbed states, $H_{mn} = \langle m|\hat{H}_1|n\rangle$. These equations can be easily recast in terms of a matrix representation

$$\begin{pmatrix}\epsilon_1 + \lambda H_{11} & H_{12}\\ H_{21} & \epsilon_2 + \lambda H_{22}\end{pmatrix}\begin{pmatrix}c_{\pm1}\\ c_{\pm2}\end{pmatrix} = E\begin{pmatrix}c_{\pm1}\\ c_{\pm2}\end{pmatrix}.$$

This is a simple eigenvalue problem on the energy E with two solutions,

$$\begin{aligned}E_\pm =& \frac{1}{2}\Big(\epsilon_1 + \epsilon_2 + \lambda H_{11} + \lambda H_{22}\\ &\pm \sqrt{(\epsilon_1 - \epsilon_2 + \lambda H_{11} - \lambda H_{22})^2 + 4\lambda^2 H_{12}H_{21}}\Big).\end{aligned} \tag{7.10}$$

It is important to note that this solves the problem *exactly*, and provides the complete eigenvalues, $E_\pm$, starting from known information on the unperturbed system (namely ϵ_i and the matrix elements). Note that we could also find the eigenvectors $\phi_\pm$ if we wanted to.

For an $N-$level system, the exact solution requires the solution of an $N \times N$ eigenvalue problem, a process that easily becomes difficult to treat analytically. It is convenient instead to try and find approximate answers. Let us look for instance at the case where λ is small and expand the exact expression on Eq. (7.10) up to second order. We expand first the square root term and find

$$\begin{aligned}
&\sqrt{(\epsilon_1 - \epsilon_2 + \lambda H_{11} - \lambda H_{22})^2 + 4\lambda^2 H_{12} H_{21}} \\
&= \sqrt{(\epsilon_1 - \epsilon_2)^2 + 2\lambda\,(\epsilon_1 - \epsilon_2)\,(H_{11} - H_{22}) + \lambda^2 \left[(H_{11} - H_{22})^2 + 4H_{12}H_{21}\right]} \\
&= (\epsilon_1 - \epsilon_2)\sqrt{1 + 2\lambda \frac{H_{11} - H_{22}}{\epsilon_1 - \epsilon_2} + \lambda^2 \frac{(H_{11} - H_{22})^2 + 4H_{12}H_{21}}{(\epsilon_1 - \epsilon_2)^2}} \\
&\approx (\epsilon_1 - \epsilon_2)\left[1 + \lambda \frac{H_{11} - H_{22}}{\epsilon_1 - \epsilon_2} + 2\lambda^2 \frac{H_{12}H_{21}}{(\epsilon_1 - \epsilon_2)^2}\right] \\
&= (\epsilon_1 - \epsilon_2) + \lambda\,(H_{11} - H_{22}) + 2\lambda^2 \frac{H_{12}H_{21}}{(\epsilon_1 - \epsilon_2)}\,.
\end{aligned}$$

In the square root expansion, one has to expand to second order

$$\sqrt{1 + a\lambda + b\lambda^2} \approx 1 + \frac{a}{2}\lambda + \frac{1}{2}\left(b - \frac{a^2}{2}\right)\lambda^2.$$

Noting that $a = 2\frac{H_{11}-H_{22}}{\epsilon_1-\epsilon_2}$ and $b = \frac{(H_{11}-H_{22})^2+4H_{12}H_{21}}{(\epsilon_1-\epsilon_2)^2}$, one finds that the prefactor in the second order term simplifies to

$$\frac{1}{2}\left(b - \frac{a^2}{2}\right) = 2\frac{H_{12}H_{21}}{(\epsilon_1 - \epsilon_2)^2}.$$

One can then use this result in Eq. (7.10) to get the two *approximate* solutions:

$$E_+ = \epsilon_1 + \lambda H_{11} + \lambda^2 \frac{H_{12}H_{21}}{(\epsilon_1 - \epsilon_2)}\,, \tag{7.11}$$

$$E_- = \epsilon_2 + \lambda H_{22} - \lambda^2 \frac{H_{12}H_{21}}{(\epsilon_1 - \epsilon_2)}\,. \tag{7.12}$$

Clearly, when $\lambda = 0$, one has $E_+ = \epsilon_1$ and $E_- = \epsilon_2$. When λ raises from that value, the previous equation gives the shift in levels as a function of λ. At first order, the shift is linear so long as $H_{nn} \neq 0$. The quadratic term only involves off-diagonal terms of the perturbation, $H_{m,n\neq m}$, but it also depends on the inverse of the unperturbed energy spacing, $\epsilon_1 - \epsilon_2$. Importantly, these require the calculation of the explicit matrix elements H_{mn}. In other words, after the $2 \times 2 = 4$ matrix elements are evaluated, the approximate solution can be found without the need for diagonalization. Let us now try to derive similar expressions explicitly from perturbation theory, where similar results are found for systems with Hilbert space dimensions larger than 2.

7.1.2 *Non-degenerate perturbation theory*

We start with the simplest possible case, in which the unperturbed spectrum is non-degenerate, i.e. $\epsilon_n \neq \epsilon_m$ if $n \neq m$. In perturbation theory, we are interested in finding the expansion coefficients for states and eigenvalues in Eqs. (7.5) and (7.4). We know that the 0-th order terms should coincide with the $\lambda = 0$ case, so that

$$|\phi_n^{(0)}\rangle = |n\rangle\,, \tag{7.13}$$

$$E_n^{(0)} = \epsilon_n\,. \tag{7.14}$$

The key difference with respect to the exact case is that we expand

the k-th order coefficients in Eqs. (7.5) in terms of the states $|n\rangle$:

$$|\phi_n^{(k)}\rangle = \sum_{m\neq n} a_m^{(k)} |m\rangle\,. \tag{7.15}$$

The sum on m avoids the n term. This implies the following orthogonality condition:

$$\langle n|\phi_n^{(k)}\rangle = 0,\, k>0\,. \tag{7.16}$$

In the $\lambda = 0$ case the norm of the state $\langle n|\phi_n\rangle = \langle n|n\rangle = 1$. If we want to keep this normalization for finite λ, one can see that the sum above must skip the n^{th} term. To see this, we use the expansion of Eq. (7.5),

$$\langle n|\phi_n\rangle = \underbrace{\langle n|\phi_n^{(0)}\rangle}_{=1} + \lambda\,\langle n|\phi_n^{(1)}\rangle + \lambda^2\,\langle n|\phi_n^{(2)}\rangle + \cdots \tag{7.17}$$

and find that the normalization immediately requires $\langle n|\phi_n^{(k)}\rangle = 0$. In consequence, the sum in Eq. (7.16) should not have an $m = n$ term[2].

[2] This comes at the expense of having to renormalize differently the wavefunction at every order k in the perturbative expansion.

We can use the expansions of Eqs. (7.5) and (7.4) in Eq. (7.3):

$$\begin{aligned}\left(\hat{H}_0 + \lambda\hat{H}_1\right)&\left(|\phi_n^{(0)}\rangle + \lambda|\phi_n^{(1)}\rangle + \lambda^2|\phi_n^{(2)}\rangle + \cdots\right)\\ &= \left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots\right)\left(|\phi_n^{(0)}\rangle + \lambda|\phi_n^{(1)}\rangle + \lambda^2|\phi_n^{(2)}\rangle + \cdots\right)\end{aligned}$$

Collecting terms on both sides of the equation order by order in λ, one finds:

$$\begin{aligned}
&(0): \left(\hat{H}_0 - E_n^{(0)}\right)|\phi_n^{(0)}\rangle = 0,\\
&(1): \left(\hat{H}_0 - E_n^{(0)}\right)|\phi_n^{(1)}\rangle = \left(E_n^{(1)} - \hat{H}_1\right)|\phi_n^{(0)}\rangle,\\
&(2): \left(\hat{H}_0 - E_n^{(0)}\right)|\phi_n^{(2)}\rangle = \left(E_n^{(1)} - \hat{H}_1\right)|\phi_n^{(1)}\rangle + E_n^{(2)}|\phi_n^{(0)}\rangle,\\
&\qquad\vdots\\
&(k): \left(\hat{H}_0 - E_n^{(0)}\right)|\phi_n^{(k)}\rangle = \left(E_n^{(1)} - \hat{H}_1\right)|\phi_n^{(k-1)}\rangle + E_n^{(2)}|\phi_n^{(k-2)}\rangle + \cdots\\
&\qquad\qquad + E_n^{(k)}|\phi_n^{(0)}\rangle
\end{aligned} \tag{7.18}$$

To solve for the first order term, we introduce the expansion of Eq. (7.15) at order $k = 1$, $|\phi_n^{(1)}\rangle = \sum_{m\neq n} a_m^{(1)}|m\rangle$, in the equation (1) for the first order terms. Further, taking the product with an arbitrary bra $\langle l|$, we find

$$\sum_{m\neq n}\langle l|\left(\hat{H}_0 - E_n^{(0)}\right)a_m^{(1)}|m\rangle = \langle l|\left(E_n^{(1)} - \hat{H}_1\right)|n\rangle.$$

The sum on the left-hand-side is simplified by acting with $\hat{H}_0$ on the bra, $\langle l|\hat{H}_0 = \langle l|\hat{e}_l$, and by the orthogonality condition $\langle l|m\rangle = \delta_{lm}$, so that

$$(\epsilon_l - \epsilon_n)\,a_l^{(1)} = E_n^{(1)}\delta_{ln} - \hat{H}_{ln}. \tag{7.19}$$

The first order equation takes different forms depending on whether $l = n$ or $l \neq n$. In the first instance, when $l = n$, we immediately find the first-order correction to the energy

$$E_n^{(1)} = \hat{H}_{nn} = \langle n|\hat{H}_1|n\rangle, \tag{7.20}$$

which is provided by the expectation value of $\hat{H}_1$ on the unperturbed basis. In contrast, the case $l \neq n$ allows us to find the expansion coefficient for the wave functions:

$$a_l^{(1)} = \frac{H_{ln}}{\epsilon_n - \epsilon_l}, \tag{7.21}$$

so that the complete first-order correction to the state reads

$$\lambda|\phi_n^{(1)}\rangle = \sum_{m\neq n} \frac{\lambda H_{mn}}{\epsilon_n - \epsilon_m}|m\rangle. \tag{7.22}$$

Once again, this suggests that the ratio $\frac{H_{mn}}{\epsilon_n - \epsilon_m}$ is indeed the small parameter in the perturbation theory expansion.

To find the higher-order terms in Eq. (7.18), one could mechanically continue through the sequence. It is however more useful to find the k-th term from the $(k-1)$-th one. To this end, we take the order k term in Eq. (7.18) and multiply with the unperturbed bra $\langle n|$:

$$\langle n|\left(\hat{H}_0 - \epsilon_n\right)|\phi_n^{(k)}\rangle = \langle n|\left(E_n^{(1)} - \hat{H}_1\right)|\phi_n^{(k-1)}\rangle + E_n^{(2)}\underbrace{\langle n|\phi_n^{(k-2)}\rangle}_{=0}$$
$$+ \cdots + E_n^{(k)}\underbrace{\langle n|\phi_n^{(0)}\rangle}_{=1}.$$

The left-hand side vanishes by construction, and all but the second and the last term of the right-hand side remain. This provides the following expression for the energy,

$$E_n^{(k)} = \langle n|\hat{H}_1|\phi_n^{(k-1)}\rangle\,.$$

This indicates that the energy at order k can be directly obtained from the wave function at order $k-1$. Using this expression for $k = 2$, one finds

$$E_n^{(2)} = \langle n|\hat{H}_1|\phi_n^{(1)}\rangle = \sum_{m\neq n} \frac{H_{nm}H_{mn}}{\epsilon_n - \epsilon_m}\,.$$

We note that the second term is dictated by the all the possible off-diagonal couplings between the state of interest, n, and all the allowed m internal states. The difference in energies $\epsilon_n - \epsilon_m$ also plays a crucial role. Nearly degenerate reference states can potentially provide very large contributions here. This suggests that the parameter in the expansion for perturbation theory need not be λ itself, but rather the ratio $\lambda\frac{H_{mn}}{(\epsilon_n - \epsilon_m)}$. Importantly, these equations only require the calculation of the $N \times N$ matrix elements H_{mn}, rather than their diagonalization.

7.1.3 *Two-level system: perturbation theory*

We can easily apply the equations above to the two-level system of Sec. 7.1.1. In that case, we have two levels, $n = 1, 2$, and the first order correction to the energies are just $E_1^{(1)} = H_{11}$ and $E_2^{(1)} = H_{22}$. The second order corrections to the energy can also be obtained. The

sum would normally run over the 2 states, but the $m \neq n$ restriction immediately reduces the sum to one single term for each eigenvalue:

$$E_1^{(2)} = \frac{H_{12}H_{21}}{\epsilon_1 - \epsilon_2}, \quad E_2^{(2)} = \frac{H_{21}H_{12}}{\epsilon_2 - \epsilon_1} = -E_1^{(2)}.$$

With these two corrections, the energies up to second order read:

$$E_1 \approx \epsilon_1 + \lambda H_{11} + \lambda^2 \frac{H_{12}H_{21}}{\epsilon_1 - \epsilon_2},$$
$$E_2 \approx \epsilon_2 + \lambda H_{22} - \lambda^2 \frac{H_{12}H_{21}}{\epsilon_1 - \epsilon_2}.$$

These expressions coincide with the corresponding expansion for the two-level system, Eqs. (7.11) and (7.12).

7.1.4 *Harmonic oscillator with harmonic perturbation*

Another somewhat trivial, but possibly useful example of perturbation theory arises by considering an unperturbed harmonic oscillator,

$$\hat{H}_0 = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2, \tag{7.23}$$

with an harmonic perturbation,

$$\hat{H}_1 = \frac{1}{2}m\omega^2\hat{x}^2. \tag{7.24}$$

The full hamiltonian is then

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 + \lambda\frac{1}{2}m\omega^2\hat{x}^2 = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega'^2\hat{x}^2, \tag{7.25}$$

where the redefined harmonic oscillator frequency $\omega' = \sqrt{1+\lambda}\omega$. Clearly, this is an harmonic oscillator of frequency ω' and we can solve the problem analytically. In particular, we can write the spectrum of the "perturbed" harmonic oscillator, E_n, in terms of the "unperturbed" spectrum, ϵ_n,

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega' = \left(n + \frac{1}{2}\right)\hbar\omega\sqrt{1+\lambda} = \epsilon_n\sqrt{1+\lambda}. \tag{7.26}$$

For small values of the coupling λ, this yields the expansion

$$E_n \approx \epsilon_n\left(1 + \frac{1}{2}\lambda - \frac{1}{8}\lambda^2 + \cdots\right). \tag{7.27}$$

We can now use perturbation theory to check if we recover this result. For simplicity, we will work with the ground $n = 0$ state, but the approach works for other states too. To find the corrections to the spectrum we need the matrix elements H_{m0}. In this case, the perturbation is quadratic in position,

$$\langle m|\hat{H}_1|n = 0\rangle = \frac{1}{2}m\omega^2\langle m|\hat{x}^2|0\rangle. \tag{7.28}$$

Using $\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}\left(\hat{a}+\hat{a}^\dagger\right)$, and $\left(\hat{a}^\dagger\right)^2|0\rangle = \sqrt{2}|2\rangle$, the matrix elements can be easily computed,

$$\begin{aligned}\langle m|\hat{H}_1|0\rangle &= \frac{\hbar\omega}{4}\langle m|\hat{a}^\dagger\hat{a}^\dagger + \hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger + \hat{a}\hat{a}|n\rangle \\ &= \frac{\hbar\omega}{4}\left(\delta_{m,0} + \sqrt{2}\delta_{m,2}\right) . \end{aligned} \tag{7.29}$$

With this, the first-order correction to the energy is simply:

$$E_0^{(1)} = \langle 0|\hat{H}_1|0\rangle = \frac{\hbar\omega}{4} = \frac{\epsilon_0}{2} . \tag{7.30}$$

In the second order term, the sum reduces to a single state, $m = 2 \neq 0$, and the second-order correction is

$$E_0^{(2)} = \frac{H_{02}H_{20}}{\epsilon_0 - \epsilon_2} = -\frac{\frac{2}{4^2}(\hbar\omega)^2}{2\hbar\omega} = -\frac{1}{16}\hbar\omega = -\frac{\epsilon_0}{8} . \tag{7.31}$$

Clearly, using perturbation theory, we manage to reproduce the results from the expansion of the "exact" case, Eq. (7.27).

7.1.5 *Degenerate perturbation theory*

The examples and derivations above have assumed that there is no degeneracy in the unperturbed spectrum. While *a priori* one may think that formulae like Eqs. (7.22) or (7.23) cannot be applied in this case, we shall see now that this is not indeed the case. The extension for the degenerate case is of interest particularly in the atomic physics context, for the so-called Stark effect[3].

[3] L. E. Ballentine. *Quantum Mechanics: A Modern Development*. World Scientific, 1998; and W. Greiner. *Quantum Mechanics: an Introduction*. Springer-Verlag, 1989

Consider a hydrogen atom that is subject to an electric field, **E**. The Stark effect is the shift in spectral lines associated to this field. It is similar in principle to the Zeeman effect associated to changes in the magnetic field. Under the assumption that the field points in the $z-$direction and that we work in spherical coordinates (as we should for the hydrogen atom), the perturbing hamiltonian becomes

$$\hat{H}_1 = qEz = qEr\cos\theta . \tag{7.32}$$

We can apply perturbation theory as described above, and in this case we shall assume $\lambda = 1$. The strength of the perturbation is then given by the size of the controllable external parameter E.

The eigenstates of the hydrogen atom are of the form $|nlm\rangle$, with n the radial quantum number; l, the angular momentum; and m the component along the $z-$axis. Using the wave function for the ground state,

$$\psi_{100}(r) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0} , \tag{7.33}$$

one immediately finds that the matrix elements of the perturbation are zero, $\langle 100|\hat{H}_1|100\rangle = 0$. Mathematically, this result arises from the integral of the $\cos\theta$ term in spherical coordinates, which yields 0. Physically, it can also be seen to arise from parity arguments[4]. The

[4] L. E. Ballentine. *Quantum Mechanics: A Modern Development*. World Scientific, 1998

important point here is that, as a consequence, the shift in energy,

$$E_{100} \approx \epsilon_{100} + \langle 100|\hat{H}_1|100\rangle + O(E^2) = \epsilon_{100} + O(E^2)\,. \tag{7.34}$$

The shift in the spectral level of the ground state due to the electric field is not linear. It must be (and indeed it is verified experimentally to be) quadratic in E.

The situation is different for the $n = 2$ state, where the spectral lines are verified experimentally to behave linearly with respect to E. The reason underlying this very different behaviour has to do with degeneracy. The levels in a hydrogen atom with quantum number n are n^2-degenerate, so in the case $n = 2$ we have 4 states that belong to the same energy ϵ_2. These 4 states correspond to the quantum numbers $|200\rangle$, $|21-1\rangle$, $|210\rangle$ and $|211\rangle$.

How can we go about and include such degenerate reference states in perturbation theory? We start by writing explicitly the eigenvalue equation of the unperturbed Hamiltonian

$$\hat{H}_0|n^r\rangle = \epsilon_n|n^r\rangle\,. \tag{7.35}$$

including a label r that runs over the degenerate states, $r = 1,\ldots,f_n$, where f_n is the degree of degeneracy of reference state n. Because of this degeneracy, we assume that at $0-$th order the approximate eigenstates are a linear combination of all the degenerate states,

A doubly-degenerate state has a degree of degeneracy $f_n = 2$. For the $n = 2$ state of the hydrogen atom, $f_n = 4$

$$|\phi_n^{r,(0)}\rangle = \sum_{r'} c_{rr'}|n^r\rangle\,. \tag{7.36}$$

We can now plug this equation into the order 1 equality of Eq. (7.18), carefully adding all the degeneracy labels in the different states:

$$\left(\hat{H}_0 - \epsilon_n\right)|\phi_n^{r,(1)}\rangle = \left(E_n^{(1)} - \hat{H}_1\right)|\phi_n^{r,(0)}\rangle = \left(E_n^{(1)} - \hat{H}_1\right)\sum_{r'} c_{rr'}|n^r\rangle$$

We can now act with a bra on the left of this equation. Starting with a bra on the state n corresponding to a degenerate state s (which need not be equal to r), we find

$$\langle n^s|(\epsilon_n - \epsilon_n)\,|\phi_n^{r,(1)}\rangle = 0 = E_n^{(1)}c_{rs} - \sum_{r'}\langle n^s|\hat{H}_1|n^r\rangle c_{rr'}\,, \tag{7.37}$$

which leads to the eigenvalue problem

$$\sum_{r'}\langle n^s|\hat{H}_1|n^r\rangle c_{rr'} = E_n^{(1)}c_{rs}\,. \tag{7.38}$$

This is a set of f_n linear equations for the coefficients $c_{rr'}$. In other words, to determine the coefficients $c_{rr'}$ we need to diagonalize the perturbing hamiltonian in the subspace spanned by the degenerate eigenvectors of energy ϵ_n. Importantly, solving the eigenvalue problem above also yields f_n (generally different) energies $E_n^{(1)}$. In other words, in general, the degeneracy of the levels is lifted by the perturbation. Originally, one had f_n unperturbed degenerate levels, which split up into f_n distinct, different levels when the perturbation is active.

For the Stark effect in the $n = 2$ hydrogen atom states, we can write the $0-$th order state as a linear combination of the 4 degenerate states,

$$|\phi_n^0\rangle = c_1|200\rangle + c_2|21-1\rangle + c_3|210\rangle + c_4|211\rangle\,.$$

We also need the off-diagonal matrix elements of the perturbing hamiltonian over these states. Luckily, the dipolar electric field component, $\langle nlj|\hat{H}_1|n'l'j'\rangle$, is only non-zero if the conditions $m = m'$ and $l = l' \pm 1$ are met[5]. The only non-zero matrix elements are $\langle 200|\hat{H}_1|210\rangle = \langle 210|\hat{H}_1|200\rangle^* = -3qEa_0$, where a_0 is the Bohr length. In matrix form, and considering the 4 relevant states, this perturbation reads

$$H_1 = \begin{pmatrix} 0 & 0 & -3qEa_0 & 0 \\ 0 & 0 & 0 & 0 \\ -3qEa_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

To find the eigenvalues, we use the condition

$$\begin{vmatrix} E^{(1)} & 0 & -3qEa_0 & 0 \\ 0 & E^{(1)} & 0 & 0 \\ -3qEa_0 & 0 & E^{(1)} & 0 \\ 0 & 0 & 0 & E^{(1)} \end{vmatrix} = 0\,,$$

which is met if either $E^{(1)} = 0$ or $E^{(1)} = \pm 3qEa_0$. Each one of these energies is associated to different eigenvectors,

$$E^{(1)} = 0 \to \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, E^{(1)} = 0 \to \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, E^{(1)} = \pm 3qEa_0 \to \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ \mp 1 \\ 0 \end{pmatrix}.$$

To summarise, the originally four degenerate states with energy ϵ_2 have now transformed into one state with energy $\epsilon_2 + 3qEa_0$, another one with energy $\epsilon_2 - 3qEa_0$ and two states that remain unchanged. Importantly, the effect of the energy shift is now linear in E - rather than quadratic as was the case for non-degenerate states. The evolution of the spectral lines as a function of E is sketched in Fig. 7.1.

[5] These conditions can be argued on the grounds that $\hat{H}_1$ does not depend on ϕ and hence can't change m; and requiring that parity be conserved, which implies the condition on l.

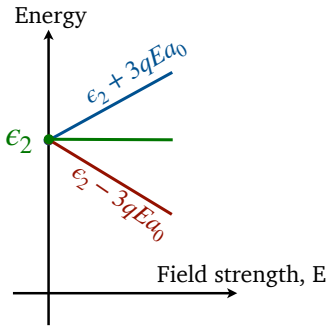

Figure 7.1: The four degenerate $n = 2$ hydrogen states split up by the Stark effect into three different states. The states in the middle remain at energy ϵ_2 and are doubly degenerate.

7.2 *Time-dependent perturbation theory*

We can apply similar ideas to perturbation theory in static systems to approximately solve for the dynamics of quantum systems. In this case, we require a time-dependent perturbation theory. To this end, it is useful to consider a time-dependent perturbation, $\lambda\hat{H}_1(t)$, on top of a static, unperturbed hamiltonian $\hat{H}_0$. The unperturbed hamiltonian defines a stationary basis $\hat{H}_0|n\rangle = \epsilon_n|n\rangle$. The wave function of the system evolves in time according to the Schrödinger equation,

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \left[\hat{H}_0 + \lambda\hat{H}_1(t)\right]|\psi(t)\rangle\,. \tag{7.39}$$

The Hamiltonian is time-dependent, so $\hat{H}$ may not longer correspond to a conserved energy[6]. Just as we did in solving the Schrödinger equation in the context of the 6th postulate (Chapter 1.2), we can expand the time-evolved state in terms of the static unperturbed basis, $\{|n\rangle\}$. But, because the hamiltonian depends on time, the expansion coefficients are in general time-dependent. We choose to expand the time-dependent ket using an ansatz that includes the phases associated to the unperturbed hamiltonian evolution,

$$|\psi(t)\rangle = \sum_n a_n(t) e^{-i\frac{\epsilon_n t}{\hbar}} |n\rangle \,. \tag{7.40}$$

Note that, in comparison to Eq. (1.31) in our initial solution of the time-dependent Schrödinger equation, the amplitudes $a_n(t)$ here are time-dependent. This is necessary to accommodate the time dependence of the perturbing hamiltonian, $\hat{H}_1(t)$. You can think about this in the following way. When $\lambda = 0$, the system is unperturbed and its dynamics is trivial - given by changing phases and constant (in time) coefficients $a_n^{(0)}$,

$$|\psi(t)\rangle = \sum_n a_n^0(t) |n\rangle \,. \tag{7.41}$$

When $\lambda \neq 0$ but it is small (in some sense to be determined), we expect that the time dependence of the coefficients $a_n(t)$ is mild and can be captured by a perturbative expansion.

Let's see how this works. The Schrödinger equation is in this case an initial value problem. We assume we know the initial state, $|\psi(t = t_0)\rangle$, and evolve it in time. As a consequence, the projection of the initial states on the unperturbed kets, $\{a_n(t = t_0)\}$, are also assumed to be known. Using the ansatz of Eq. (7.40) in the Schrödinger equation, Eq. (7.39), we find

$$\begin{aligned} i\hbar \frac{d}{dt}|\psi(t)\rangle &= \sum_n \left[i\hbar \dot{a}_n(t) + \epsilon_n a_n(t)\right] e^{-i\frac{\epsilon_n t}{\hbar}} |n\rangle \\ \left(\hat{H}_0 + \lambda \hat{H}_1(t)\right)|\psi(t)\rangle &= \sum_n \left[\epsilon_n a_n(t) + \lambda \hat{H}_1 a_n(t)\right] e^{-i\frac{\epsilon_n t}{\hbar}} |n\rangle \,, \end{aligned}$$

Clearly, the terms proportional to ϵ_n cancel on both sides. Projecting on a bra $\langle m|$, we end up with an equation that determines the time dependence of the coefficients:

$$i\hbar \dot{a}_m(t) = \lambda \sum_n \langle m|\hat{H}_1|n\rangle e^{-i\frac{(\epsilon_n - \epsilon_m)t}{\hbar}} a_n(t) \,.$$

By introducing the following notation for the matrix elements, $H_{mn}(t) = \langle m|\hat{H}_1(t)|n\rangle$, and the so-called *Bohr frequencies* $\omega_{mn} = \frac{\epsilon_m - \epsilon_n}{\hbar}$, this equation simplifies to

$$i\hbar \dot{a}_m(t) = \lambda \sum_n H_{mn}(t) e^{i\omega_{mn} t} a_n(t) \,. \tag{7.42}$$

So far, we have not performed any simplifications. In fact, this expression is still equivalent to the full Schrödinger equation.

[6] In the context of the Ehrenfest theorem derivation of Section 1.2.3, one sees that the expectation value of the hamiltonian is no longer conserved, $\frac{d\langle h\rangle}{dt} = \langle\psi|\frac{dH_1(t)}{dt}|\psi\rangle \neq 0$.

Importantly, this equation illustrates transparently that the time dependence of the $a_m(t)$ coefficients is entirely due to the time-dependent perturbation matrix elements, H_{mn}. The phase factors associated to the Bohr frequencies absorb the time dependence of $\hat{H}_0$. We can transform the previous equation to integral form,

$$a_m(t) = a_m(t = t_0) + \frac{\lambda}{i\hbar}\int_{t_0}^{t} d\bar{t} \sum_n H_{mn}(\bar{t}) e^{i\omega_{mn}\bar{t}} a_n(t)\,, \tag{7.43}$$

with an expression that explicitly incorporates the initial conditions $a_m(t = t_0)$.

It is at this point that we can introduce some approximations. First, we introduce a perturbation expansion of the coefficients in terms of λ,

$$a_n(t) = a_n^{(0)}(t) + \lambda a_n^{(1)}(t) + \lambda^2 a_n^{(2)}(t) + \cdots$$

Using this expansion in the integral equation above, and matching order by order in λ the left and the right hand side, one finds

$$\begin{aligned}
&(0) : a_m^{(0)}(t) = a_m(t = t_0),\\
&(1) : a_m^{(1)}(t) = \frac{\lambda}{i\hbar}\int_{t_0}^{t} d\bar{t} \sum_n H_{mn}(\bar{t}) e^{i\omega_{mn}\bar{t}} a_n^{(0)},\\
&\qquad\vdots\\
&(k) : a_m^{(k+1)}(t) = \frac{\lambda}{i\hbar}\int_{t_0}^{t} d\bar{t} \sum_n H_{mn}(\bar{t}) e^{i\omega_{mn}\bar{t}} a_n^{(k)}(\bar{t}).
\end{aligned} \tag{7.44}$$

This is a complicated, nested set of equations for the contributions at different orders of $a_n(t)$. We focus on the first two equations to attempt and solve the problem at first order.

Clearly, the 0th order term is just equal to the initial conditions - which are known. The first-order contribution is particularly simple because the coefficients $a_m^{(0)}$ are time-independent. To further simplify the equations at play, let us consider a specific problem and consider the set-up illustrated in Fig. 7.2. We assume that the perturbation acts for a *finite* amount of time, say between t_1 and t_1' as shown in the Figure. Far in the past before the perturbation acts, at $t_0 \to \infty$, the system is in a well-defined unperturbed eigenstate $|i\rangle$. With this, $a_n(t \to -\infty) = \delta_{ni}$, which collapses the sum in the first-order coefficient to single term. Furthermore, we consider that we are looking into a unique final state, $|f\rangle \neq |i\rangle$ after the perturbation has been switched off, $t > t_1'$. Because the perturnation has been switched off, we can choose a state belonging to the reference Hamiltonian without ambiguities.

With these simplications, the first-order coefficient becomes

$$a_f^{(1)}(t) = \frac{\lambda}{i\hbar}\int_{-\infty}^{t} d\bar{t}\, H_{fi}(\bar{t}) e^{i\omega_{fi}\bar{t}}, \tag{7.45}$$

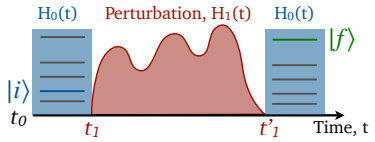

Figure 7.2: The usual set-up in time-dependent perturbation theory is illustrated here. At a faraway initial time, t_0, the system is in a well-defined eigenstate of the unperturbed hamiltonian, $|i\rangle$. The perturbation acts over a finite period of time, say between t_1 and t_1'. We then look at a well-defined final state $|f\rangle$.

and the corresponding transition probability is simply

$$\mathcal{P}_{i\to f}(t) = \frac{\lambda^2}{\hbar^2}\left|\int_{-\infty}^{t} d\bar{t}\, H_{fi}(\bar{t})e^{i\omega_{fi}\bar{t}}\right|^2 . \tag{7.46}$$

Moreover, we have assumed that $t > t_1'$, and so the perturbation H_1 has already ceased to act at time t. This allows us to take the limit $t \to \infty$ and find the transition probability,

$$\mathcal{P}_{i\to f}(t) = \frac{\lambda^2}{\hbar^2}\left|\int_{-\infty}^{\infty} d\bar{t}\, H_{fi}(\bar{t})e^{i\omega_{fi}\bar{t}}\right|^2 . \tag{7.47}$$

Note that this is just a Fourier harmonic of the perturbation at a frequency that is equal to the transition Bohr frequency.

Physically, we can interpret this formula in terms of *virtual* states. The transition can happen at any time $\bar{t} < t$ and the phase jumps once with the Bohr frequency, but we integrate over all possible times to account for all the allowed intermediate virtual jumps.

It is also important to understand that a transition means absorbing an energy difference $\hbar\omega_{fi} = \epsilon_f - \epsilon_i$ from the perturbing field. The system goes from a well-defined initial state at t_0 to one of many populated final states at the end of the perturbation as illustrated in Fig. 7.3. This does not mean that $|\psi(t_f)\rangle = |f\rangle$. $|\psi(t_f)\rangle$ is instead a superposition of several reference states, $|\psi(t_f)\rangle = \sum_n a_n(t_f)e^{-i\frac{\epsilon_n t_f}{\hbar}}|n\rangle$.

Finally, an implicit assumption in this derivation is that the transitions $f \to i$ with $f \neq i$ are small. One can easily see that the transition probability is of order λ^2. This is as opposed to the transition $i \to i$, which is instead of order 1 and hence large compared to the others.

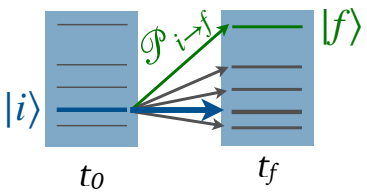

Figure 7.3: When we discuss transitions, we do not mean to say that the final state $|f\rangle$ is the only eigenstate at the end of the time evolution. The final state is fragmented among several different states, and the transition probability $\mathcal{P}_{i\to f}$ quantifies how likely it is to go from $|i\rangle$ into $|f\rangle$. This is in contrast to the initial state which is prepared in a well-defined eigenstate $|i\rangle$. Note that the transition $\mathcal{P}_{i\to i}$ is assumed to be dominant in perturbation theory.

7.2.1 *Harmonic perturbation*

We have so far not specified the explicit type of time dependence of the perturbation hamiltonian. A very common case is an harmonic perturbation[7], in which the time dependence has a well-defined frequency ω. We specify the time dependence in terms of the hamiltonian

$$\lambda\hat{H}_1(t) = H'e^{-i\omega t} + H'^{\dagger}e^{i\omega t},\ 0 \le t < T. \tag{7.48}$$

We use two terms in this expression that are Hermitian conjugate to each other. This guarantees the hermiticity of $\hat{H}$. We assume that this harmonic perturbation starts at a time $t = 0$ and ends at $t = T$. Using the expression in Eq. (7.45), we find:

$$\begin{aligned} a_f^{(1)}(T) &= \frac{1}{i\hbar}\int_0^T d\bar{t}\,\left(H'_{fi}e^{-i\omega\bar{t}} + H'^{\dagger}_{fi}e^{i\omega\bar{t}}\right)e^{i\omega_{fi}\bar{t}} \\ &= \frac{H'_{fi}}{i\hbar}\int_0^T d\bar{t}\, e^{-i(\omega-\omega_{fi})\bar{t}} + \frac{H'^{\dagger}_{fi}}{i\hbar}\int_0^T d\bar{t}\, e^{-i(\omega+\omega_{fi})\bar{t}} \\ &= H'_{fi}\frac{1-e^{-i(\omega-\omega_{fi})T}}{\hbar\left(\omega_{fi}-\omega\right)} + H'^{\dagger}_{fi}\frac{1-e^{-i(\omega+\omega_{fi})T}}{\hbar\left(\omega_{fi}+\omega\right)}. \end{aligned} \tag{7.49}$$

[7] This is the case of a monochromatic electromagnetic wave, for instance.

Note that we define the perturbation matrix elements $\hat{H}'$ in a way that they incorporate λ, which we will ignore from now on.

For a weak perturbation, this amplitude will be small whenever ω is substantially different from $\pm\omega_{fi}$. The transition probability becomes large only in the vicinity of $\omega \approx \pm\omega_{fi}$, a condition which is called *resonance*. In fact, assuming $\epsilon_f > \epsilon_i$ and $\omega > 0$, the condition $\hbar\omega \approx \epsilon_f - \epsilon_i$ is associated to *resonant absorption* processes. The system absorbs energy from the perturbation with large probability when this condition is fulfilled. Similarly, when $\omega < 0$, the condition $\hbar\omega \approx \epsilon_i - \epsilon_f$ is indicative of *resonant emission*, and the system looses energy to the perturbation with a large probability.

Let us consider the first case, $\hbar\omega \approx \epsilon_f - \epsilon_i$. The first term in Eq. (7.49) dominates and the corresponding transition probability becomes

$$\begin{aligned}\mathcal{P}_{i\to f}(T) &\approx \frac{\left|H'_{fi}\right|^2}{\hbar^2}\frac{\left|1-e^{-i\left(\omega-\omega_{fi}\right)T}\right|^2}{\left(\omega_{fi}-\omega\right)^2} \\ &= \frac{\left|H'_{fi}\right|^2}{\hbar^2}\frac{\left|e^{-\frac{i\left(\omega-\omega_{fi}\right)T}{2}}\left[e^{\frac{i\left(\omega-\omega_{fi}\right)T}{2}}-e^{-\frac{i\left(\omega-\omega_{fi}\right)T}{2}}\right]\right|^2}{\left(\omega_{fi}-\omega\right)^2} \\ &= \frac{\left|H'_{fi}\right|^2}{\hbar^2}\frac{\sin^2\left(\frac{\omega_{fi}-\omega}{2}T\right)}{\left(\frac{\omega_{fi}-\omega}{2}\right)^2}\end{aligned} \tag{7.50}$$

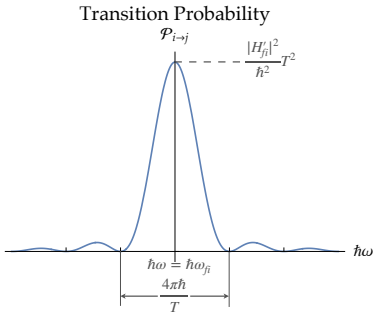

Figure 7.4: The resonance part of the transition probability at finite time as a function of the perturbation frequency energy, $\hbar\omega$. The height of the probability peak increases with T^2, but its width decreases as $\frac{1}{T}$.

The probability that we obtain is the product of two terms. The first term concerns the matrix elements (squared), $\left|H'_{fi}\right|^2$. This is associated to the strength of the perturbation - or to how likely it is that the interaction connects the states $i \to f$. The second term encodes all the time dependence. The transition probability is shown in Fig. 7.4 as a function of the energy $\hbar\omega$. This is the frequency ω of the external field, which is usually a tunable parameter. Because of the quadratic nature of the function, the dependence on $\hbar\omega_i$ or $\hbar\omega_f$ is exactly the same. The figure clearly indicates that at resonance, $\hbar\omega = \hbar\omega_{fi}$, the probability shows a strong peak. By Taylor expanding the numerator of Eq. (7.50) around $\frac{\omega_{fi}-\omega}{2}T \approx 0$, one sees that at resonance

$$\mathcal{P}_{i\to f}(T) \xrightarrow[\omega=\omega_{fi}]{} \frac{\left|H'_{fi}\right|^2 T^2}{\hbar^2} \tag{7.51}$$

or, in other words, the peak is proportional to T^2 independently of the energy.

Further, upon inspection of Fig. 7.4, one sees that the probability only has a significant value within an energy interval of width $\Delta\omega \approx \frac{4\pi\hbar}{T}$ around the resonance. In other words, as T increases, the maximum of the peak grows with T^2, but the width of the peak decreases like T^{-1}. Because the area under the peak remains finite, we shall see in the next chapter that, in the limit $T \to \infty$, the probability is essentially a δ function.

7.3 *Rabi oscillations: perturbation theory*

The quadratic growth of the maximum of $\mathcal{P}_{i\to f}$ with time, Eq. (7.51), is somewhat problematic. It indicates that the probability may grow arbitrarily with time, in contrast to the arguments of time-dependent perturbation theory which indicate that $\mathcal{P}_{i\to f}$ should be small for all times. One may however wonder whether this prediction from first-order perturbation theory is even valid since, naively, this theory should only be applicable at short times. In order to assess this situation, we look back at Rabi oscillations. This will allow us to compare a transition probability from exact dynamics to the expression in Eq. (7.50), establishing some conditions of validity for perturbation theory. As a reminder, in the Rabi oscillation case, we found that the exact transition probability from a $|+\rangle$ state at the origin of time to a $|-\rangle$ state a given time t later is given by the expression:

$$\mathcal{P}_{+\to-}(t) = \frac{\omega_1^2}{(\omega-\omega_0)^2+\omega_1^2}\sin^2\left(\frac{\Omega t}{2}\right). \tag{7.52}$$

One can look at Rabi oscillations from the perspective of time-dependent perturbation theory. In that case, the unperturbed time-independent term of the Rabi hamiltonian Eq. (6.1), is $\hat{H}_0 = \frac{\hbar\omega_0}{2}\sigma_z$. This Hamiltonian generates the unperturbed basis $|+\rangle$ and $|-\rangle$ with energies $\epsilon_\pm = \pm\frac{\hbar\omega_0}{2}$.

The harmonic time-dependent perturbation, in contrast, corresponds to the term

$$\hat{H}_1 = \frac{\hbar\omega_1}{2}\cos\omega t\hat{\sigma}_x + \frac{\hbar\omega_1}{2}\sin\omega t\hat{\sigma}_y = \frac{\hbar\omega_1}{2}\left[e^{i\omega t}\hat{\sigma}_- + e^{-i\omega t}\hat{\sigma}+\right], \tag{7.53}$$

with $\hat{\sigma}_\pm = \frac{1}{2}(\hat{\sigma}_x \mp i\hat{\sigma}_y)$ such that $\sigma_\pm|\mp\rangle = |\pm\rangle$. The Bohr frequencies in this case correspond to unperturbed energy differences, $\omega_{-+} = -\omega_0$. We thus need to look for Eq. (7.50) in the region where $\omega < 0$. The dominant term is resonant absorption, and we find

$$\mathcal{P}_{+\to-}(t) \approx \frac{|H'_{-+}|^2}{\hbar^2}\frac{\sin^2\left(\frac{\omega_{-+}+\omega}{2}T\right)}{\left(\frac{\omega_{-+}+\omega}{2}\right)^2}. \tag{7.54}$$

The matrix elements can be easily computed

$$H'_{-+} = \langle-|\hat{H}_1(t)|+\rangle = \frac{\hbar\omega_1}{2}e^{i\omega t} \Rightarrow |H'_{-+}|^2 = \frac{(\hbar\omega_1)^2}{4},$$

so that the transition probability becomes

$$\mathcal{P}_{+\to-}(t) = \frac{\omega_1}{(\omega-\omega_0)^2}\sin^2\left(\frac{\omega-\omega_0}{2}T\right). \tag{7.55}$$

In principle, we may have anticipated that this equation would only hold for short times near the resonance $\omega \approx |\omega_0|$. Closeness to the resonance is one of the criteria we used to derive Eq. (7.50), and

first-order perturbation theory should only work for short times. We indeed find

$$\mathcal{P}_{+\to-}(t) \approx \left(\frac{\omega_1 t}{2}\right)^2, \tag{7.56}$$

which is exactly the same result one would obtain by expanding the exact formula in Eq. (6.11) for small t. This expression is shown in dashed lines in the bottom panel of Fig. 6.1.

Surprisingly, however, the formula is also valid in a different, off-resonance regime. When $|\omega - \omega_0| \gg \omega_1$, one finds that $\Omega \approx \omega - \omega_0$ and the two expressions Eq. (6.11) and Eq. (7.54) are close to each other not only at small times but, in principle, for *all times*. An example of this behaviour is given in the top panel of Fig. 6.1, which shows in a dashed line the perturbation theory result. One can see that this follows relatively closely the exact result for a relatively long period of time. In this regime, the perturbation is small and time-dependent perturbation theory works excellently.

The lesson here is that perturbation theory works at resonance for small times (as expected), but it can also work well off resonance across all times. In this second situation, the time dependence of the transition probability is relatively mild and well-captured by the perturbative expansion at first order, independently of the time evolution.

7.4 *Fermi Golden Rule*

Fermi's Golden Rule allows us to compute transition probabilities per unit of time (or simply transition rates) between two states. To derive the rule, we start from Eq. (7.50), which provides the transition probability between two different states, $i \neq f$, when the external perturbation is monochromatic with a given frequency $\omega \approx \omega_{fi}$. We consider, instead of an individual final state $|f\rangle$, a set of states around ϵ_f, which we characterise with a typical width $\Delta\epsilon_f$. The transition rate is given by a sum of terms in Eq. (7.50) that differ in the value of ϵ_f. If we multiply and divide by T^2 and sum over the set of final states between $\epsilon_f - \Delta\epsilon_f$ and $\epsilon_f + \Delta\epsilon_f$, we find

$$\mathcal{P}_{i\to\Delta\epsilon_f}(T) = \frac{T^2}{\hbar^2} \sum_{\epsilon_f-\Delta\epsilon_f}^{\epsilon_f+\Delta\epsilon_f} \left|H'_{fi}\right|^2 \frac{\sin^2\left(\frac{\epsilon_f-\epsilon_i-\hbar\omega}{2\hbar}T\right)}{\left(\frac{\epsilon_f-\epsilon_i-\hbar\omega}{2\hbar}T\right)^2}. \tag{7.57}$$

Note that matrix elements $\left|H'_{fi}\right|^2$ depend on the final state f through the first index so, in principle, they may change when changing ϵ_f. A sketch of the physical picture at hand is provided in Fig. 7.5.

Importantly, because of the dependence on the energy ω through the $\frac{\sin^2(\dots)}{(\dots)^2}$, factor, the available subset of states in the final state is increasingly small with time. In other words, the width of probability distribution becomes sharper and the area with a

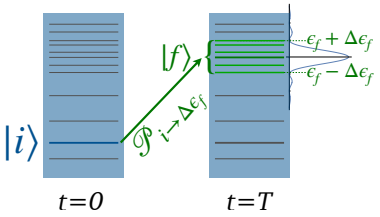

Figure 7.5: We consider a transition from a well-defined initial state, $|i\rangle$, at $t = 0$ to a set of final states in the region $(\epsilon_f - \Delta\epsilon_f, \epsilon_f + \Delta\epsilon_f)$ around ϵ_f at $t = T$. As time evolves, the probability distribution becomes sharper and the area with a significant transition probability becomes differentially small, $\Delta\epsilon_f \to d\epsilon_f$.

significant transition probability becomes differentially small, $\Delta\epsilon_f \to d\epsilon_f$. In these circumstances (eg for long enough times T), we can replace the sum over discrete states by an integral over a continuum of states. To do so, we need to count how many states, ν, occur around a given energy, ϵ_f. This is precisely the definition of the density of states $\frac{d\nu}{d\varepsilon_f}$. With this, the integral version of Eq. (7.57) becomes:

$$\mathcal{P}_{i\to d\epsilon_f}(T) = \frac{T^2}{\hbar^2}\int_{\epsilon_f-\Delta\epsilon_f}^{\epsilon_f+\Delta\epsilon_f} \mathrm{d}\bar{\varepsilon}_f \left(\frac{d\nu}{d\varepsilon_f}\right)_{\epsilon_f=\bar{\varepsilon}_f} \left|H'_{\bar{f}i}\right|^2 \frac{\sin^2\left(\frac{\bar{\varepsilon}_f-\epsilon_i-\hbar\omega}{2\hbar}T\right)}{\left(\frac{\bar{\varepsilon}_f-\epsilon_i-\hbar\omega}{2\hbar}T\right)^2}. \tag{7.58}$$

As the time T evolves, the $\sin^2(\cdots)/\cdots^2$ factor in the integrand becomes more and more peaked around $\epsilon_f = \epsilon_i + \hbar\omega$. In doing so, however, the term keeps a constant area and hence behaves like a $\delta-$function. Let us explicitly look at the area under the curve associated to this term, by integrating across the spectrum and performing a change of variables $x = \frac{\bar{\varepsilon}_f-\epsilon_i-\hbar\omega}{2\hbar}T$:

$$\int_{-\infty}^{+\infty} \mathrm{d}\bar{\varepsilon}_f \frac{\sin^2\left(\frac{\bar{\varepsilon}_f-\epsilon_i-\hbar\omega}{2\hbar}T\right)}{\left(\frac{\bar{\varepsilon}_f-\epsilon_i-\hbar\omega}{2\hbar}T\right)^2} = \frac{2\hbar}{T}\int_{-\infty}^{+\infty} \mathrm{d}x \frac{\sin^2(x)}{(x)^2} = \frac{2\pi\hbar}{T}.$$

In other words, for times that are long enough, we can assume that the $\sin^2(\cdots)/\cdots^2$ factor behaves like a $\delta-$function with the normalization above,

$$\frac{\sin^2\left(\frac{\epsilon_f-\epsilon_i-\hbar\omega}{2\hbar}T\right)}{\left(\frac{\epsilon_f-\epsilon_i-\hbar\omega}{2\hbar}T\right)^2} \xrightarrow[T\gg\frac{2\pi\hbar}{\Delta E}]{} \frac{2\pi\hbar}{T}\delta\left(\epsilon_f-\epsilon_i-\hbar\omega\right). \tag{7.59}$$

Strictly speaking, for this approximation to work, times must be large compared to the associated prefactor $\frac{\epsilon_f-\epsilon_i-\hbar\omega}{2\hbar} = \frac{\Delta E}{2\hbar}$. In other words, the time must be long enough to account for several oscillations in the sinusoidal factor, so $T \gg \frac{2\pi\hbar}{\Delta E}$.

With this, the transition probability for long enough times, $T \gg \frac{2\pi\hbar}{\Delta E}$, becomes:

$$\mathcal{P}_{i\to d\epsilon_f} = \frac{2\pi}{\hbar}T\left|H'_{fi}\right|^2\left(\frac{d\nu}{d\varepsilon_f}\right)_{\epsilon_f=\epsilon_i+\hbar\omega}. \tag{7.60}$$

Taking the time derivative of this object, we find a time-independent quantity. This is the *rate* of the transition,

$$\mathcal{W}_{i\to d\epsilon_f} = \frac{d\mathcal{P}_{i\to d\epsilon_f}}{dT} = \frac{2\pi}{\hbar}\left|H'_{fi}\right|^2\left(\frac{d\nu}{d\varepsilon_f}\right)_{\epsilon_f=\epsilon_i+\hbar\omega}, \tag{7.61}$$

which provides the change of probability per unit of time.

Fermi Golden Rule

The transition rate induced by an harmonic perturbation of energy $\hbar\omega$ from a well-defined initial state with energy ϵ_i to a (differentially small) set of states around a final energy ϵ_f is given by

$$\mathcal{W}_{i\to d\epsilon_f} = \frac{2\pi}{\hbar}\left|H'_{fi}\right|^2\left(\frac{d\nu}{d\varepsilon_f dx_1\cdots dx_n}\right)_{\epsilon_f=\epsilon_i+\hbar\omega}. \tag{7.62}$$

The rate is proportional to the squared perturbation matrix elements, $\left|H'_{fi}\right|^2$, time the density of states, $\left(\frac{d\nu}{d\varepsilon_f}dx_1\cdots dx_n\right)_{\epsilon_f=\epsilon_i+\hbar\omega}$. In this expression, conservation of energy must always be assumed, so that $\epsilon_f = \epsilon_i + \hbar\omega$. $x_1,\cdots,x_n$ represent variables in the final state that not entirely fixed by energy conservation.

We have derived Fermi's golden rule in the context of first-order perturbation theory, and under several approximations such as an harmonic perturbation with energy close to a resonance. To set up energy conservation precisely, we have assumed that sufficiently long times occur, so $T \gg \frac{2\pi\hbar}{\Delta E}$. One can show, however, that the rule holds in general terms, and it is in fact the starting point of several different treatments - including scattering, which we will discuss in the next chapter.

It is also important to keep in mind the exact meaning of the $i \to d\epsilon_f$ notation. We are considering here the rate to a given set of states around ϵ_f - in a differential sense. There are instances where the energies themselves do not have a one-to-one correspondence with final states, $|f\rangle$. Consider for instance a three-dimensional setting with spherical symmetry. In that case, the final energies do not depend on the (solid) angle Ω. The density of states in this situation is such that counts the number of states per unit of energy and for a differential angular change, $\left(\frac{d\nu}{d\varepsilon_f d\Omega}\right)$. Assuming there are $x_1,\cdots,x_n$ variables that are not fixed by energy conservation, the rate given by the golden rule is then a (multiple) differential quantity,

$$d\mathcal{W}_{i\to d\epsilon_f} = \frac{2\pi}{\hbar}\left|H'_{fi}\right|^2\left(\frac{d\nu}{d\varepsilon_f dx_1\cdots dx_n}\right)dx_1\cdots dx_n\,, \tag{7.63}$$

depending on the set of differentially small final angles. Further integration over all the $x_1,\cdots,x_n$ variables gives a global rate that is independent of the angle. In other words, if one is looking into the rate for a process such that the final state cannot be fully constrained by the energy, the rate should account for any potentially small changes of the unconstrained variables around ϵ_f. It is important to keep in mind that, whenever such degeneracies occur, the rates need to be multiplied by differentially small changes of *all* the variables that the density of states cannot constrain. In this spirit, for instance, if the final state includes not just a single particle, but more than one, the rates need to include the density of states for each one of these final particles.

7.5 Schrödinger, Heisenberg and Interaction Pictures

The description of quantum systems can be undertaken from different perspectives, which we call *pictures*. This is due to the fact that we have the freedom to use unitary transformations to "move"

from one perspective to another while keeping the same observable expectation values. We will now discuss some examples of such pictures, which have to do with time.

7.5.1 *Schrödinger Picture*

We have so far employed the *Schrödinger Picture*. In this picture, the states evolve according to the Schrödinger equation,

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle \,. \tag{7.64}$$

We shall for the time being assume that $\hat{H}$ is time-independent. The expectation value of any given (time-independent) operator $\hat{A}$ is given by

$$\langle a\rangle(t) = \langle\psi(t)|\hat{A}|\psi(t)\rangle \,. \tag{7.65}$$

We have explicitly written the time dependence of the expectation value which, as the right-hand-side indicates, is due only to the time dependence of the ket. As we discussed in Section 1.2.3, the time evolution of this expectation value is given by Ehrenfest's law

$$\frac{d}{dt}\langle a\rangle(t) = \frac{1}{i\hbar}\langle\psi|[\hat{A},\hat{H}]|\psi\rangle + \langle\psi|\frac{d\hat{A}}{dt}|\psi\rangle \,,$$

and the second term is dropped whenever operators do not depend on time (which is a common situation).

7.5.2 *Time Evolution Operator*

As we discussed in Section 1.2.1, one can always perform a unitary transformation of the basis and the operators. As a reminder, a unitary transformation can be expressed in terms of an operator $\hat{U}$, such that $\hat{U}\hat{U}^\dagger = \hat{I}$. The states transform according to $|\psi\rangle \to |\psi'\rangle = \hat{U}|\psi\rangle$. Such transformations keep the expectation values of operators (and hence, the observables) unchanged.

In the case of the Heisenberg picture, the unitary transformation is related to the time evolution operator

$$\hat{U}(t,t_0) = e^{-i\frac{\hat{H}(t-t_0)}{\hbar}} \,. \tag{7.66}$$

This operator can be obtained as a formal solution of the time-dependent Schrödinger equation, Eq. (7.64), and it is such that

$$|\psi(t)\rangle = \hat{U}(t,t_0)|\psi(t_0)\rangle \,.$$

In other words, $\hat{U}(t,t_0)$ translates a state in time from an initial time t_0 to another time t. This unitary transformation fulfils the standard relation $\hat{U}(t,t_0)\hat{U}^\dagger(t,t_0) = \hat{I}$. It is relatively easy to see that the inverse of $\hat{U}(t,t_0)$ is $\hat{U}^{-1}(t,t_0) = \hat{U}^\dagger(t,t_0) = \hat{U}(t_0,t)$. Or, in other words, that the inverse of the time evolution in going from t_0 to t is the *time reversed* time evolution, going from t to t_0.

As a unitary transformation that depends on a continuous parameter t, the transformation also fulfills a *group property*, $\hat{U}(t_2,t_0) = \hat{U}(t_2,t_1)\hat{U}(t_1,t_0)$. The later can be used in the case $t_2 = t_1 + \epsilon$ to find the time derivative of this operator,

$$\begin{aligned}\frac{d}{dt}\hat{U}(t,t_0) &= \lim_{\epsilon\to 0}\frac{\hat{U}(t+\epsilon,t_0)-\hat{U}(t,t_0)}{\epsilon} = \lim_{\epsilon\to 0}\frac{\left(\hat{U}(\epsilon,t_0)-\hat{I}\right)\hat{U}(t,t_0)}{\epsilon}\\ &= \lim_{\epsilon\to 0}\frac{\left(-i\frac{\hat{H}(\epsilon)}{\hbar}\right)\hat{U}(t,t_0)}{\epsilon} = \frac{-i}{\hbar}\hat{H}\hat{U}(t,t_0)\,,\end{aligned}$$

where we have expanded Eq. (7.66) up to linear terms in time. In other words, the equation of motion for $\hat{U}$ is

$$i\hbar\frac{d}{dt}\hat{U}(t,t_0) = \hat{H}\hat{U}(t,t_0)\,. \tag{7.67}$$

Note that this is a differential equation for an operator, not a function. The solution of such equations requires specific methods.

7.5.3 *Heisenberg Picture*

So far, we have discussed the time evolution of states within the Schrödinger picture. The Heisenberg picture differs from this because one chooses to remove the time dependence of the states by acting with the *inverse* of $\hat{U}(t,t_0)$ on Schrödinger kets:

$$|\psi_H\rangle = \hat{U}^{-1}(t,t_0)|\psi_S(t)\rangle = e^{i\hat{H}(t-t_0)}|\psi_S(t)\rangle = |\psi_S(t_0)\rangle\,. \tag{7.68}$$

We use the notation $|\psi_H\rangle$ and $\hat{A}_H$ for kets and operators in the Heisenberg picture. We temporarily use $|\psi_S(t)\rangle$ for the Schrödinger kets for clarity.

Looking at the expectation value of a given operator $\hat{A}$ and introducing this time evolution, we find

$$\begin{aligned}\langle a\rangle(t) &= \langle\psi_S(t)|\hat{A}_S|\psi_S(t)\rangle = \langle\psi_H|\hat{e}^{i\hat{H}(t-t_0)}\hat{A}\hat{e}^{-i\hat{H}(t-t_0)}|\psi_H\rangle\\ &\equiv \langle\psi_H|\hat{A}_H(t)|\psi_H\rangle\,.\end{aligned} \tag{7.69}$$

The last equality defines the Heisenberg picture of an operator,

$$\hat{A}_H(t) = e^{i\frac{\hat{H}(t-t_0)}{\hbar}}\hat{A}_S e^{-i\frac{\hat{H}(t-t_0)}{\hbar}}\,, \tag{7.70}$$

as a sandwich of exponentials of the Hamiltonian with respect to the Schrödinger picture operator. Clearly, this operator now depends on time and is therefore a time-evolving object.

Just as we did for $\hat{U}$, we can now find the equation of motion for this operator through direct differentiation

$$i\hbar\frac{d}{dt}\hat{A}_H(t) = \left[\hat{A}(t),\hat{H}\right]_H + i\hbar\left(\frac{d\hat{A}_S}{dt}\right)_H\,. \tag{7.71}$$

The first term corresponds to a commutator of Schrödinger-picture operators computed in the Heisenberg picture. but it could equally be computed in the Heisenberg picture altogether

$$\begin{aligned}\left[\hat{A},\hat{H}\right]_H &= e^{i\frac{\hat{H}(t-t_0)}{\hbar}}\left[\hat{A},\hat{H}\right]e^{-i\frac{\hat{H}(t-t_0)}{\hbar}}\\ &= e^{i\frac{\hat{H}(t-t_0)}{\hbar}}\left(\hat{A}\hat{H}-\hat{H}\hat{A}\right)e^{-i\frac{\hat{H}(t-t_0)}{\hbar}}\\ &= e^{i\frac{\hat{H}(t-t_0)}{\hbar}}\left(\hat{A}e^{i\frac{-\hat{H}(t-t_0)}{\hbar}}e^{i\frac{\hat{H}(t-t_0)}{\hbar}}\hat{H}-\hat{H}e^{i\frac{-\hat{H}(t-t_0)}{\hbar}}e^{i\frac{\hat{H}(t-t_0)}{\hbar}}\hat{A}\right)e^{-i\frac{\hat{H}(t-t_0)}{\hbar}}\\ &= \left[e^{i\frac{\hat{H}(t-t_0)}{\hbar}}\hat{A}e^{-i\frac{\hat{H}(t-t_0)}{\hbar}}, e^{i\frac{\hat{H}(t-t_0)}{\hbar}}\hat{H}e^{-i\frac{\hat{H}(t-t_0)}{\hbar}}\right] = \left[\hat{A}_H,\hat{H}_H\right].\end{aligned}$$

Note, moreover, that $\hat{H}_H = \hat{H}$. The second term accounts for any possible (if rare) explicit time dependent of the operator $\hat{A}$ in the Schrödinger picture. The expression should be interpreted as

$$\left(\frac{d\hat{A}_S}{dt}\right)_H = e^{i\frac{\hat{H}(t-t_0)}{\hbar}} \frac{d\hat{A}_S(t)}{dt} e^{-i\frac{\hat{H}(t-t_0)}{\hbar}} \tag{7.72}$$

To summarise, in the Heisenberg picture we have replaced all the time dependence of the states by time dependence of the operators. This picture is useful for theoretical derivation, and is also the closest to classical mechanics in mathematical terms. This is, however, not the only available choice when it comes to time evolution, as we shall shortly see.

7.5.4 *Interaction Picture*

In the interaction picture, we work in a setting that is reminiscent to time-dependent perturbation theory. We split the Hamiltonian in the Schrödinger equation in two terms,

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \left[\hat{H}_0 + \hat{V}(t)\right]|\psi(t)\rangle\,. \tag{7.73}$$

$\hat{H}_0$ is again a time-independent, unperturbed Hamiltonian (eg a non-interacting term), whereas $\hat{V}(t)$ is a possibly time-dependent interaction term. The definition of the interaction picture state is

We employ $\hat{V}$ rather than $\lambda\hat{H}_1$ in the following for notational simplicity.

$$|\psi_I(t)\rangle = e^{i\frac{\hat{H}_0(t-t_0)}{\hbar}}|\psi_S(t)\rangle\,. \tag{7.74}$$

By multiplying with the inverse of the non-interacting part of the time evolution operator, we effectively remove the time dependence from $\hat{H}_0$ on the interaction picture ket $|\psi_I(t)\rangle$. Clearly, at the origin of time the states coincides with its Schrödinger picture counterpart, $|\psi_I(t = t_0)\rangle = |\psi_S(t_0)\rangle$. As time evolves, however, the interaction picture ket and the Schrödinger picture ket differ. One can see that this is due to the fact that the dynamics of the kets is due entirely to the interaction, $\hat{V}(t)$, rather than to $\hat{H}_0$. We see this by looking at the equation of motion of the interaction-picture ket:

$$\begin{aligned}
i\hbar\frac{d}{dt}|\psi_I(t)\rangle &= i\hbar\frac{d}{dt}\left[e^{i\frac{\hat{H}(t-t_0)}{\hbar}}|\psi_S(t)\rangle\right] \\
&= -e^{i\frac{\hat{H}_0(t-t_0)}{\hbar}}\hat{H}_0|\psi_S(t)\rangle + e^{i\frac{\hat{H}_0(t-t_0)}{\hbar}}\left(\hat{H}_0 + \hat{V}(t)\right)|\psi_S(t)\rangle \\
&= \underbrace{e^{i\frac{\hat{H}_0(t-t_0)}{\hbar}}\hat{V}(t)e^{-i\frac{\hat{H}_0(t-t_0)}{\hbar}}}_{\hat{V}_I(t)}\underbrace{e^{i\frac{\hat{H}_0(t-t_0)}{\hbar}}|\psi_S(t)\rangle}_{|\psi_I(t)\rangle} = \hat{V}_I(t)|\psi_I(t)\rangle\,,
\end{aligned}$$

which is just

$$i\hbar\frac{d}{dt}|\psi_I(t)\rangle = \hat{V}_I(t)|\psi_I(t)\rangle \tag{7.75}$$

This clearly indicates that the dynamics of $|\psi_I(t)\rangle$ is entirely governed by the interaction (in the interaction picture representation), $\hat{V}_I(t)$.

The derivation of previous equation also provides the transformation of operators from the Schrödinger to the interaction picture,

$$\hat{A}_I = e^{i\frac{\hat{H}_0(t-t_0)}{\hbar}} \hat{A}_S e^{-i\frac{\hat{H}_0(t-t_0)}{\hbar}} . \tag{7.76}$$

The corresponding equation of motion of the operator can easily be found following the same derivation as the Heisenberg picture case and we find

$$i\hbar \frac{d}{dt}\hat{A}_I(t) = \left[\hat{A}_I(t), \hat{H}_0\right] + i\hbar \left(\frac{d\hat{A}}{dt}\right)_I . \tag{7.77}$$

Note that in the commutator only $\hat{H}_0$ is active. One interprets this as an indication that the dynamics of the operator is governed by the non-interacting term of the hamiltonian, $\hat{H}_0$, rather than the interaction. The second term is to be interpreted analogously to Eq. (7.72) with $\hat{H} \to \hat{H}_0$.

The differences between the time evolution of states and operators in the three pictures discussed so far are summarised in Table 7.1. It is important to keep in mind that all observables remain the same in either picture. For the interaction picture, for instance, one finds:

$$\begin{aligned}\langle a\rangle(t) &= \langle\psi_I(t)|\hat{A}_I(t)|\psi_I(t)\rangle \\ &= \langle\psi_S(t)|e^{-i\frac{\hat{H}_0(t-t_0)}{\hbar}} e^{i\frac{\hat{H}_0(t-t_0)}{\hbar}} \hat{A}_S e^{-i\frac{\hat{H}_0(t-t_0)}{\hbar}} e^{i\frac{\hat{H}_0(t-t_0)}{\hbar}} |\psi_S(t)\rangle \\ &= \langle\psi_S(t)|\hat{A}_S|\psi_S(t)\rangle .\end{aligned}$$

Picture	Operators	States
Schrödinger	static	evolves ($\hat{H}$)
Heisenberg	evolves ($\hat{H}$)	static
Interaction	evolves ($\hat{H}_0$)	evolves ($\hat{V}$)

Table 7.1: A short summary of the different pictures in quantum mechanics. We stress that in Schrödinger not all operators are time-independent.

We can now ask ourselves how a ket in the interaction picture at time t_0 is connected to a ket at a later time t. We postulate the existence of a unitary time evolution operator, $\hat{U}_I$, analogous to Eq. (7.66) but in the interaction picture,

$$|\psi_I(t)\rangle = \hat{U}_I(t,t_0)|\psi_I(t_0)\rangle .$$

The definition above clearly implies the boundary condition $\hat{U}_I(t_0,t_0) = \hat{I}$. Using the equation of motion of $|\psi_I(t)\rangle$, Eq. (7.75), and keeping in mind that t_0 (and hence $|\psi_I(t_0)\rangle$ are arbitrary), one finds the equation of motion for $\hat{U}_I$:

$$i\hbar \frac{d}{dt}\hat{U}_I(t,t_0) = \hat{V}_I(t)\hat{U}_I(t,t_0) .$$

Note that because the interaction $\hat{V}$ is explicitly time dependent, and because generally it does not commute with $\hat{H}_0$, the operator $\hat{U}$ is not simply the exponential of a hamiltonian term, $\hat{U}_I(t,t_0) \neq e^{-i\frac{\hat{H}(t-t_0)}{\hbar}}$. Here we cannot use the same derivation as in Eq. (7.67) because, again, the U_I is not a simple exponential.

We can now consider the integral version of this equation. Together with the boundary condition $\hat{U}_I(t_0,t_0) = \hat{I}$, the integral version reads

$$\hat{U}_I(t,t_0) = \hat{I} - \frac{i}{\hbar}\int_{t_0}^{t} d\bar{t}\, \hat{V}_I(\bar{t})\hat{U}_I(\bar{t},t_0) .$$

This is an integral equation for the operator, involving a time evolution dictated only by $\hat{V}_I$, as expected. Note however that the operator appears both on the left and the right-hand-side, so there is no easy solution for this. A different, possibly more useful

expression can be obtained by iterating the previous equation. One can start by setting $\hat{U}_I(\bar{t}, t_0) = \hat{I}$ on the right hand side. The approximated $\hat{U}_I^{(1)}$ can then be used again in the integrand. Doing this repeatedly, we obtain a series expansion that does not involve $\hat{U}_I$ in the right-hand-side:

$$\hat{U}_I(t, t_0) = \hat{I} - \frac{i}{\hbar}\int_{t_0}^{t} d\bar{t}\, \hat{V}_I(\bar{t}) + \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} d\bar{t} \int_{t_0}^{\bar{t}} d\bar{\bar{t}}\; \hat{V}_I(\bar{t})\hat{V}_I(\bar{\bar{t}}) + \cdots . \tag{7.78}$$

It turns out that approximate solutions of perturbation theory can be obtained by truncating this series at different orders. In particular, suppose we start from a well-defined eigenstate of $\hat{H}_0$, $|i\rangle$, at $t = t_0$. The transition amplitude in the interaction picture to a final state $|f\rangle$ (another eigenstate of $\hat{H}_0$) is given by the expression:

$$\begin{aligned} a_f(t) &= \langle f|\hat{U}_I(t, t_0)|i\rangle \\ &= \delta_{fi} - \frac{i}{\hbar}\int_{t_0}^{t} d\bar{t}\, \langle f|\hat{V}_I(\bar{t})|i\rangle + \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} d\bar{t} \int_{t_0}^{\bar{t}} d\bar{\bar{t}}\; \langle f|\hat{V}_I(\bar{t})\hat{V}_I(\bar{\bar{t}})|i\rangle + \cdots . \end{aligned} \tag{7.79}$$

Keeping in mind the definition of the interaction picture interaction, we find that the matrix elements can be written as

$$\langle f|\hat{V}_I(t)|i\rangle = \langle f|e^{i\hat{H}_0(t-t_0)}\hat{V}(t)e^{-i\frac{\hat{H}(t-t_0)}{\hbar}}|i\rangle = e^{i\omega_{fi}t}\underbrace{\langle f|\hat{V}(t)|i\rangle}_{V_{fi}(t)},$$

where ω_{fi} are the corresponding Bohr frequencies of the system. With this, the first order term in the expansion of Eq. (7.79) reads

$$a_f^{(1)}(t) = -\frac{i}{\hbar}\int_{t_0}^{t} d\bar{t}\, V_{fi}(\bar{t})e^{i\omega_{fi}\bar{t}}.$$

in complete analogy with the first-order time-dependent perturbation theory result of Eq. (7.45), under the replacement $\hat{V} \to \lambda\hat{H}_1$. The second order term can be easily worked out by adding a resolution to the identity in between the two interactions to find

$$a_f^{(2)}(t) = -\frac{1}{\hbar^2}\int_{t_0}^{t} d\bar{t} \int_{t_0}^{\bar{t}} d\bar{\bar{t}}\; V_{fn}(\bar{t})e^{i\omega_{fn}\bar{t}}V_{ni}(\bar{\bar{t}})e^{i\omega_{ni}\bar{\bar{t}}}.$$

One could continue this expansion and obtain higher order terms if necessary.

We can, however, also find a closed formal expression for the operator $\hat{U}_I$ starting from Eq. (7.78). Let us look at the two time integrals in the second term,

$$\int_{t_0}^{t} d\bar{t} \int_{t_0}^{\bar{t}} d\bar{\bar{t}}\; \hat{V}_I(\bar{t})\hat{V}_I(\bar{\bar{t}}).$$

These time integrals cover the gray area of the $\bar{t} - \bar{\bar{t}}$ plane in Fig. 7.6. The area is bounded by above by the line $\bar{t} = \bar{\bar{t}}$. For a fixed value of $\bar{t}$, we integrate $\bar{\bar{t}}$ vertically up until $\bar{t} = \bar{\bar{t}}$. This is illustrated by the red lines in the top panel of Fig. 7.6. We then integrate $\bar{t}$ from t_0 to t. We can change the order of integration to cover the same area, as

illustrated in the bottom panel in of Fig. 7.6. Here, we integrate first at fixed $\bar{\bar{t}}$ from $\bar{t}$ to t and then integrate over all $\bar{\bar{t}}$'s:

$$\int_{t_0}^{t} d\bar{\bar{t}} \int_{\bar{t}}^{t} d\bar{t}\ \hat{V}_I(\bar{t})\hat{V}_I(\bar{\bar{t}})\,.$$

This expression amounts to a permutation of integration variables, together with the integration limits. In this second expression, we can use the change of dummy integration variables $\bar{t} \leftrightarrow \bar{\bar{t}}$ to find that the term can be equivalently written as

$$\int_{t_0}^{t} d\bar{t} \int_{\bar{t}}^{t} d\bar{\bar{t}}\ \hat{V}_I(\bar{\bar{t}})\hat{V}_I(\bar{t})\,.$$

Because this covers the same area and hence gives the same result, we can replace the original integral by the half-sum of the two integrals

$$\frac{1}{2}\left[\int_{t_0}^{t} d\bar{t} \int_{t_0}^{\bar{t}} d\bar{\bar{t}}\ \hat{V}_I(\bar{t})\hat{V}_I(\bar{\bar{t}}) + \int_{t_0}^{t} d\bar{t} \int_{\bar{t}}^{t} d\bar{\bar{t}}\ \hat{V}_I(\bar{\bar{t}})\hat{V}_I(\bar{t})\right]\,.$$

Note that the second term has now two operators in a different time ordering. Operators at different times do not generally commute, which is why we cannot simplify this term straight away. In fact, looking at the integral, one finds that the expression is such that the operator product is written in a chronological order. The operators are ordered according to their time variables, with operators occurring earlier to the right of those occurring later. We can therefore rewrite the second-order term as

$$\left(-\frac{i}{\hbar}\right)^2 \frac{1}{2} \int_{t_0}^{t} d\bar{t} \int_{t_0}^{t} d\bar{\bar{t}}\ \mathcal{T}\left[\hat{V}_I(\bar{t})\hat{V}_I(\bar{\bar{t}})\right]\,,$$

with a time-ordering operator such that

$$\mathcal{T}\left[\hat{V}_I(t_1)\hat{V}_I(t_2)\right] = \begin{cases} \hat{V}_I(t_1)\hat{V}_I(t_2), & t_1 > t_2, \\ \hat{V}_I(t_2)\hat{V}_I(t_1), & t_1 < t_2. \end{cases}$$

This expression is useful because one now has the *same* integration limits in the two variables t_1 and t_2, which allows for a much more symmetric notation.

One can employ a similar procedure for the higher order terms. At nth order, there are n $\hat{V}_I(t_i)$ operators placed within the integral,

$$a_f^{(n)}(t) = \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n\ \hat{V}_I(t_1)\hat{V}_I(t_2)\cdots\hat{V}_I(t_n)\,.$$

The chronological order of this term is fixed by construction. We can explore all the possible $n!$ equivalent permutations of the integration times. Relabeling the time variables as we did for the second-order case, one finds that each of these permutations provides the same result. As a consequence, the original expression can be replaced by $1/n!$ times the time-ordered contribution,

$$a_f^{(n)}(t) = \left(-\frac{i}{\hbar}\right)^n \frac{1}{n!} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \cdots \int_{t_0}^{t} dt_n\ \mathcal{T}\left[\hat{V}_I(t_1)\hat{V}_I(t_2)\cdots\hat{V}_I(t_n)\right]\,.$$

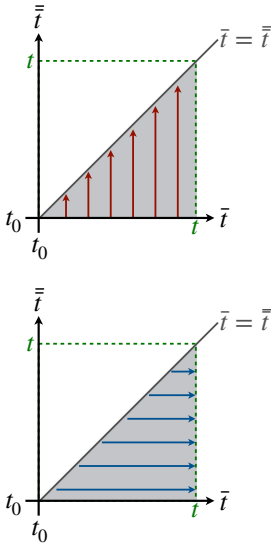

Figure 7.6: The time integral over the $\bar{t} - \bar{\bar{t}}$ half-plane can be performed in two analogous ways described in this figure.

The time-ordered product can be generalised to cases involving more than two products of operators. One always puts operators at earlier times to the right of operators at later times.

The full series now resembles an exponential,

$$\hat{U}(t,t_0) = \hat{I} + \sum_{n=1} \left(-\frac{i}{\hbar}\right)^n \frac{1}{n!} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \cdots \int_{t_0}^{t} dt_n \; \mathcal{T}\left[\hat{V}_I(t_1)\hat{V}_I(t_2)\cdots\hat{V}_I(t_n)\right] . \tag{7.80}$$

This type of time-ordered expansion is usually calles a *Dyson series*, and is in fact sometimes written formally and compactly as

$$\hat{U}(t,t_0) = \mathcal{T}\left[e^{-\frac{i}{\hbar}\int_{t_0}^{t} d\bar{t}\, \hat{V}_I(\bar{t})}\right] . \tag{7.81}$$

Sandwiching the previous expression over an initial and a final state, one can also find generic expressions for the transitions probabilities,

$$a_f(t) = \langle f|\hat{U}(t,t_0)|i\rangle = U_{fi}(t,t_0) = \mathcal{T}\left[e^{-\frac{i}{\hbar}\int_{t_0}^{t} d\bar{t}\, V_I(\bar{t})}\right]_{fi} . \tag{7.82}$$

8 Scattering

8.1 Born approximation scattering

The formalism of the previous Chapter, particularly the Fermi golden rule, allows us to describe scattering problems. Time-dependent perturbation theory is necessary to account for the time evolution of the scattering, but the Fermi golden rule already accounts for it. This makes the problem particularly simple - and apparently time-independent!

The standard set-up for a scattering experiment looks like the diagram in Fig. 8.1. We consider a moving projectile of mass m that scatters off a target placed at the origin of coordinates. In the remote past, far away from the target, this projectile is a freely propagating particle, an eigenstate of the hamiltonian $\hat{H}_0 = \frac{\hat{p}^2}{2m}$. The projectile thus has a well-defined momentum $\mathbf{p}$, which we choose to lie in the $z-$axis. Over a certain period of time, the projectile interacts with the target through a perturbation potential, $\hat{H}_1 = V(\mathbf{r}\{t\})$, that depends on the distance between projectile and target, $\mathbf{r}\{t\}$. This distance in turn depends on time, and implicitly builds a time dependence on the hamiltonian. After the scattering process, the particle ends up with a momentum $\mathbf{p}'$, pointing in a different direction characterised by a polar angle θ. A detector with a solid angle aperture $\Delta\Omega$ is placed far away from the target ($r \gg 1$), and detects outgoing particles at a rate of $N(\Omega, \Delta\Omega)$ particles per unit of time. In a typical experiment, one can generally tune in the energy (or momentum) of the projectile beam as well as its intensity. The intensity is often characterised by the number of particles crossing a unit area in a direction perpendicular to the beam (the $x-y$ plane in the case of Fig. 8.1). This is the so-called beam flux, j_{b}. One also knows a few properties of the target, like its density, area and volume. With this, one can characterise how many projectiles interact with target particles over an overlap interaction volume V, giving the effective density $n_{\text{t+b}} = \frac{N_{\text{t+b}}}{V}$.

We assume that the energies are relatively high, so multi-step interactions between the projectile and the target can be neglected. At large times, we expect that the total energy, sum of the projectile, ϵ, and target, E, energies, to be conserved throughout the process. Hence, $\epsilon_i + E_i = \epsilon_f + E_f$. Further, if the target is structureless and inert, we cannot excite it through the scattering process and $E_i = E_f$. This immediately implies that the only possible scattering channel is

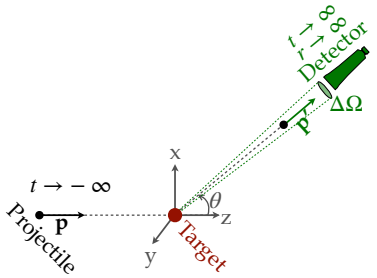

Figure 8.1: The standard setup of a scattering experiment.

In general, the projectile scatters at a solid angle $\Omega = (\theta, \phi)$. For a given fixed detector position, though, we can always rotate the coordinates along the $z-$axis so the detector lies in the $x-z$ plane. This immediately makes the azimutal angle zero, $\phi = 0$. For central potentials, as we shall see, one can altogether ignore the ϕ dependence of the scattering cross sections.

This is the relevant density for the scattering process, since we are counting how many interactions take place per unit volume , rather than how many particles are in a given volume.

elastic scattering, where $\epsilon_i = \epsilon_f$. For eigenstates of $\hat{H}_0$, this immediately implies $\epsilon_i = \frac{p^2}{2m} = \epsilon_f = \frac{p'^2}{2m}$ and hence the magnitudes of the incoming and outgoing momenta are the same, $p = p'$.

8.1.1 Density of states

In order to apply Fermi's Golden Rule in Eq. (7.61), we need to understand how to work with the density of states, $\left(\frac{d\nu}{d\epsilon_f}\right)_{\epsilon_f=\epsilon_i+\hbar\omega}$, for states in the continuum. In Section 4.5, we have argued the the number ν_p of properly quantized states between a momentum p and $p + dp$ is given by

The results of scattering theory are independent of the normalization, but the density of states is easier to compute with specific normalization choices.

$$d\nu_p = \left(\frac{L}{2\pi\hbar}\right)^3 d^{(3)}\mathbf{p}. \tag{8.1}$$

We can separate the solid angle and the modulus dependence of this differential,

$$d\nu_p = \left(\frac{L}{2\pi\hbar}\right)^3 p^2 dp d\Omega. \tag{8.2}$$

Each plane wave of momentum $\mathbf{p}$ carries an energy $\epsilon_p = \frac{p^2}{2m}$. This energy only depends on the modulus $p = |\mathbf{p}|$; it does not depend on the angular properties of $\mathbf{p}$. Moreover, a small change in momentum modulus, dp, brings about a change in energy of size $d\epsilon_p = \frac{p}{m}dp$. This allows us to write the number of states per unit of energy from Eq. (8.2),

$$d\nu_\epsilon = \left(\frac{L}{2\pi\hbar}\right)^3 pm d\epsilon_p d\Omega \Rightarrow \frac{d\nu}{d\epsilon d\Omega} = \left(\frac{L}{2\pi\hbar}\right)^3 pm\,. \tag{8.3}$$

This expression indicates that the density of states in the continuum can be computed up to a differential of solid angle $d\Omega$. Using Eq. (7.63), we can express the rate of scattering into a final state formed of a single non-relativistic particle of mass m and momentum $p' = p$ as

$$\begin{aligned}\frac{d\mathcal{W}_{i\to f}}{d\Omega} &= \frac{2\pi}{\hbar}\left|H'_{fi}\right|^2\left(\frac{d\nu}{d\epsilon d\Omega}\right) = \frac{2\pi}{\hbar}\left|H'_{fi}\right|^2\left(\frac{L}{2\pi\hbar}\right)^3 p'm \\ &= \frac{L^3 pm}{4\pi^2\hbar^4}\left|H'_{fi}\right|^2.\end{aligned} \tag{8.4}$$

In this expression, the energy is assumed to be conserved. L is a physical size, assumed to be much larger than the scattering volume.

8.1.2 Cross sections and rates

We can connect the properties of the incoming and outgoing wave functions to the scattering cross section. In a scattering experiment, a detector at angle Ω of small aperture $\Delta\Omega$ detects $N(\Omega, \Delta\Omega)$ particles per unit of time. If the aperture is small, $N(\Omega, \Delta\Omega)$ should be proportional to $\Delta\Omega$. Assuming multiple scatting is neglected, $N(\Omega, \Delta\Omega)$ should also be proportional to the number of target

particles in the region overlapping with the beam, $n_{\text{t+b}}$. If beam particles do not interact, the count rate $N(\Omega, \Delta\Omega)$ should also be proportional to the incident flux j_{b}. The constant of proportionality between $N(\Omega, \Delta\Omega)$ and these quantities is the differential cross section,

$$N(\Omega, \Delta\Omega) = \Delta\Omega \times n_{\text{t+b}} \times j_{\text{b}} \times \frac{d\sigma}{d\Omega},$$

which is thus defined as

$$\frac{d\sigma}{d\Omega} = \frac{N(\Omega, \Delta\Omega)}{\Delta\Omega\, n_{\text{t+b}}\, j_{\text{b}}}. \tag{8.5}$$

This expression provides a clear procedure to compute the cross section from experimental data.

In a theoretical set-up like the one described in Fig. 8.1, we can also compute N and j_{b} from the properties of the wave functions. Given a wavefunction $\psi(\mathbf{r})$, the probability flux is given by,

$$\mathbf{j}(\mathbf{r}) = -i\frac{\hbar}{2m}\left[\psi^*(\mathbf{r})\nabla\psi(\mathbf{r}) - \psi(\mathbf{r})\nabla\psi^*(\mathbf{r})\right]. \tag{8.6}$$

The quantum mechanical flux $\mathbf{j}$ is equivalent to the flux of fluid, or the flux of the electromagnetic field. For instance, it fulfills the continuity equation

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{j} = 0,$$

where ρ is the probability density given by $\rho(\mathbf{r}, t) = \psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)$. In time-independent situations, the flux is conserved, $\nabla\cdot\mathbf{j} = 0$.

We compute first the beam flux, j_b. This is the number of beam particles going through a unit surface in the xy plane, but can also be obtained as the projection of the flux vector of the wave function along the $z-$axis

$$\begin{aligned} j_{\text{b}} = \mathbf{k}\cdot\mathbf{j}_{\text{in}}(\mathbf{r}) &= -i\frac{\hbar}{2m}\left[e^{-i\mathbf{k}\cdot\mathbf{z}}\frac{d}{dz}e^{i\mathbf{k}\cdot\mathbf{z}} - e^{i\mathbf{k}\cdot\mathbf{z}}\frac{d}{dz}e^{-i\mathbf{k}\cdot\mathbf{z}}\right] \\ &= -i\frac{\hbar}{2m}2i\underset{=k}{k_z} = \frac{\hbar k}{m} = \frac{p}{m} \end{aligned} \tag{8.7}$$

Note that the flux is just the velocity of a single particle of mass m and momentum k, $v = \frac{p}{m}$.

A flux of beam particles hits the target. $n_{\text{t+b}}$ is the number of scattering centers per unit of time and unit of volume in the initial state. Out of all these initial states, only a few will transition into the final state. How many of these transition to the final state is dictated by the transition rate, so the rate of scattered particles that is read by the detector is simply

$$N(\Omega, \Delta\Omega) = N_{\text{t+b}} \times d\mathcal{W}_{i\to f} = n_{\text{t+b}} \times L^3 \times d\mathcal{W}_{i\to f}.$$

Note that N detects the total number of particles (as opposed to a particle density n), which explains the presence of the factor L^3. With these definitions, the cross section reads

$$\frac{d\sigma}{d\Omega} = \frac{d\mathcal{W}_{i\to f}}{d\Omega}\frac{n_{\text{t+b}}L^3}{n_{\text{t+b}}\frac{p}{m}} = \frac{mL^3}{p}\frac{d\mathcal{W}_{i\to f}}{d\Omega} = L^6\frac{p'}{p}\frac{m^2}{4\pi^2\hbar^4}\left|H'_{fi}\right|^2. \tag{8.8}$$

where we work in the $\Delta\Omega \to d\Omega$ limit and we have used the Fermi golden rule expression, Eq. (8.4).

8.1.3 *Central potentials*

The previous expression allows us to compute the differential scattering cross section in terms of the initial and final momenta, p and p'; the mass of the scatterer, m; and the matrix elements H'_{fi}. We need to compute such matrix elements to find the cross section. Clearly, these matrix elements depend on the interaction between the projectile and the target. We model this perturbation in terms of a potential that depends on the distance between the projectile and the target, $\hat{H} = V(\mathbf{r})$. We assume that the initial and final states are plane waves as given by Eq. (4.76), with momenta $\mathbf{p_i} = \mathbf{p}$ and $\mathbf{p_f} = \mathbf{p}'$. In elastic scattering, we must have equal initial and final momentum magnitudes, $p = p'$. The matrix elements are given by the integral

$$H'_{fi} = \langle f|\hat{H}'|i\rangle = \int \mathrm{d}^{(3)}\mathbf{r}\, \psi^*_{\mathbf{p_f}}(\mathbf{r}) V(\mathbf{r}) \psi_{\mathbf{p_i}}(\mathbf{r}) = \frac{1}{L^3}\int \mathrm{d}^{(3)}\mathbf{r}\, e^{-i\frac{(\mathbf{p_i}-\mathbf{p_f})\cdot\mathbf{r}}{\hbar}} V(\mathbf{r}) \tag{8.9}$$

We can introduce the momentum transfer $\mathbf{q} = \frac{\mathbf{p_f}-\mathbf{p_i}}{\hbar}$. It is also useful to define an interaction form factor

$$V(\mathbf{q}) = \int \mathrm{d}^{(3)}\mathbf{r}\, e^{-i\mathbf{q}\cdot\mathbf{r}} V(\mathbf{r})\,, \tag{8.10}$$

which is nothing but the Fourier transform of the interaction. Using the matrix elements above in the cross section formula Eq. (8.8), we see that the normalization volumes disappear and the final expression for the Born approximation cross section is

$$\frac{d\sigma}{d\Omega} = \left(\frac{m}{2\pi\hbar^2}\right)^2 |V(\mathbf{q})|^2\,. \tag{8.11}$$

This remarkably simple formula indicates that the scattering cross section of elastic scattering is only a function of the transfer momentum $\mathbf{q}$, and depends on the Fourier structure of the interaction. Note that the integral we perform to compute the form factor is independent of the scattering process. We are therefore free to choose the frame of reference for the dummy integration variable $\mathbf{r}$. Typically, one chooses $\mathbf{r}$ along the $z-$axis, so the dot product of the exponential term is just $\mathbf{q}\cdot\mathbf{r} = qr\cos\theta$, where θ is the angle between $\mathbf{r}$ and $\mathbf{q}$. This integration angular variable should not be confused with the angle θ of the scattering process of Fig. 8.1.

Further simplifications arise in the case of a scattering potential that is central, i.e. that depend only on the modulus of $\mathbf{r}$, $V(\mathbf{r}) = V(r)$. In this situation, the form factor integral,

$$V(\mathbf{q}) = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \mathrm{d}\theta \sin\theta \int_0^{\infty} \mathrm{d}r\, r^2 e^{-iqr\cos\theta} V(r)\,, \tag{8.12}$$

can be simplified because the angular integrals are analytical. The ϕ integral can be performed straightforwardly. We can also use a change of variables $x = \cos\theta$ to find the integral over θ,

$$\begin{aligned} V(\mathbf{q}) &= 2\pi \int_0^{\infty} \mathrm{d}r\, r^2 V(r) \int_{-1}^{1} \mathrm{d}x\, e^{-iqrx} = \frac{2\pi}{iq}\int_0^{\infty} \mathrm{d}r\, rV(r)\left[e^{iqr} - e^{-iqr}\right] \\ &= \frac{4\pi}{q}\int_0^{\infty} \mathrm{d}r\, rV(r)\sin(qr)\,. \end{aligned} \tag{8.13}$$

Note that for a central potential the form factor is a real function of the modulus of the momentum transfer, $V(\mathbf{q}) = V(q)$. Further, in elastic scattering, the momentum transfer is given by

$$q^2 = \frac{1}{\hbar^2}\left(\mathbf{p}-\mathbf{p}'\right)^2 = \frac{1}{\hbar^2}\left(p^2+p'^2-2pp'\cos\theta\right) = \frac{2p^2}{\hbar^2}\left(1-\cos\theta\right)$$
$$= \frac{4p^2}{\hbar^2}\sin^2\frac{\theta}{2} \Rightarrow q = \frac{2p}{\hbar}\sin\frac{\theta}{2}\,. \tag{8.14}$$

In other words, q is entirely determined by the modulus of the momentum, p, and the angle between the initial and final momenta, θ.

8.1.4 *Yukawa potential*

We are now in a position to work out a specific example for the cross section. Consider an interaction between projectile and target which is of the Yukawa type,

$$V(r) = \frac{V_0}{r}e^{-\frac{r}{\mu}}\,. \tag{8.15}$$

This interaction diverges at short distances, $r \to 0$, but goes to zero exponentially when $r \gg \mu$, as shown in Fig. 8.2. You can generate this plot with the Jupyter Notebook Scattering_Born_Yukawa.ipynb available at the Master's github repository shorturl.at/fEXZ5. μ is therefore a proxy for the range of the interaction. In particular, in the limit $\mu \to \infty$, the Yukawa interaction cannot be distinguished from a $1/r$ potential, like the Coulomb force. Note that V_0 is a constant with units of energy times length.

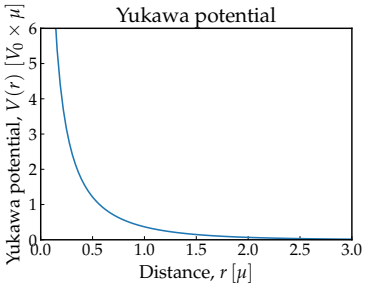

Figure 8.2: The Yukawa potential as a function of distance. For this plot, we choose $V_0 = 1$ and $\mu = 1$ (or, alternatively, we plot $V(r)$ in units of V_0/μ and distances in units of μ).

To compute the Yukawa interaction form factor, we can start from the last expression in the first line of Eq. (8.13), and multiply the exponentials

$$V(q) = \frac{2\pi}{iq}\int_0^\infty \mathrm{d}r\, r\frac{V_0}{r}\left[e^{-\frac{r}{\mu}}e^{iqr} - e^{-\frac{r}{\mu}}e^{-iqr}\right]$$
$$= \frac{2\pi V_0}{iq}\left[\frac{1}{-\frac{1}{\mu}+iq}\,e^{-\left(\frac{1}{\mu}-iq\right)r}\Big|_0^\infty + \frac{1}{\frac{1}{\mu}+iq}\,e^{-\left(\frac{1}{\mu}+iq\right)r}\Big|_0^\infty\right]$$
$$= \frac{2\pi V_0}{iq}\left[\frac{1}{\frac{1}{\mu}-iq} - \frac{1}{\frac{1}{\mu}+iq}\right] = \frac{4\pi V_0}{\frac{1}{\mu^2}+q^2} = \frac{4\pi V_0\mu^2}{1+\mu^2q^2}.$$

Introducing this form factor in the cross section formula of Eq. (8.11), we find

$$\frac{d\sigma}{d\Omega} = \left(\frac{m}{2\pi\hbar^2}\right)^2\left|\frac{4\pi V_0\mu^2}{1+\mu^2q^2}\right|^2 = \left(\frac{2mV_0\mu^2}{\hbar^2}\right)^2\frac{1}{\left[1+\frac{8mE\mu^2}{\hbar^2}\sin^2\frac{\theta}{2}\right]^2},$$

where we have used the definition of the momentum transfer, Eq. (8.14), and $p^2 = 2mE$. Introducing a characteristic energy scale $\varepsilon_\mu = \frac{\hbar^2}{2m\mu^2}$, the previous expression simplifies to

$$\frac{d\sigma}{d\Omega} = \left(\frac{V_0}{\varepsilon_\mu}\right)^2\frac{1}{\left[1+\frac{4E}{\varepsilon_\mu}\sin^2\frac{\theta}{2}\right]^2}\,. \tag{8.16}$$

The differential cross section as a function of angle is shown in Fig. 8.3 for three values of the energy. At low energies, like the case $E = 0.1\varepsilon_\mu$ shown in solid lines, the prefactor in the denominator is small and the angular dependence is mild.

For higher energies, the angular dependence of the cross section is more pronounced. More importantly, the cross section tends to a constant value $\frac{d\sigma}{d\Omega} \to \left(\frac{2mV_0\mu^2}{\hbar^2}\right)^2$ at forward angles, $\theta \approx 0$, independently of the energy. For an unknown potential, this may allow us to infer a value for the combination of parameters $V_0\mu^2$. You can generate this plot with the Jupyter Notebook Scattering_Born_Yukawa.ipynb available at the Master's github repository shorturl.at/fEXZ5.

We can also find a closed expression in the case $\mu \to \infty$ which, as we argued earlier, would be the equivalent of a Coulomb potential. In that case, starting from Eq. (8.16), we find

$$\frac{d\sigma}{d\Omega} \approx \left(\frac{2mV_0\mu^2}{\hbar^2}\right)^2 \frac{1}{\left[\frac{8mE\mu^2}{\hbar^2}\sin^2\frac{\theta}{2}\right]^2} = \left(\frac{V_0}{4E}\right)^2 \frac{1}{\sin^4\frac{\theta}{2}}\,. \tag{8.17}$$

This expression corresponds to the Rutherford cross section of Coulomb scattering. The very same expression is obtained from classical physics. The cross section as a function of angle is shown for different energies in Fig. 8.4. Importantly, the cross section has a universal angular dependence and diverges at forward angles.

Finally, we can perform an angular integration to obtain the *total cross section* for the Yukawa potential,

$$\begin{aligned}
\sigma = \int \mathrm{d}\Omega \frac{d\sigma}{d\Omega} &= \int_0^{2\pi} \mathrm{d}\phi \int_{-1}^{1} \mathrm{d}x \left(\frac{V_0}{\varepsilon_\mu}\right)^2 \frac{1}{\left[1+\frac{2E}{\varepsilon_\mu}-\frac{2E}{\varepsilon_\mu}x\right]^2} \\
&= 2\pi\left(\frac{V_0}{\varepsilon_\mu}\right)^2 \frac{\varepsilon_\mu}{2E} \left.\frac{1}{\left[1+\frac{2E}{\varepsilon_\mu}-\frac{2E}{\varepsilon_\mu}x\right]}\right|_{-1}^{1} \\
&= 2\pi\left(\frac{V_0}{\varepsilon_\mu}\right)^2 \frac{\varepsilon_\mu}{2E}\left(1-\frac{1}{1+\frac{4E}{\varepsilon_\mu}}\right) \\
&= 4\pi\left(\frac{V_0}{\varepsilon_\mu}\right)^2 \frac{1}{1+\frac{4E}{\varepsilon_\mu}}\,.
\end{aligned}$$

A plot of this function is shown in Fig. 8.5. At high energies, the projectile is too energetic to interact with the target and the cross section decreases like $\sigma \approx E^{-1}$. In contrast, at low energies the cross section saturates to a value $\sigma \to 4\pi\left(\frac{V_0}{\varepsilon_\mu}\right)^2$. This formula is equivalent to the surface of a sphere with effective radius $R = \frac{V_0}{\varepsilon_\mu}$. This plot can be generated with the Jupyter Notebook Scattering_Born_Yukawa.ipynb available at the Master's github repository shorturl.at/fEXZ5.

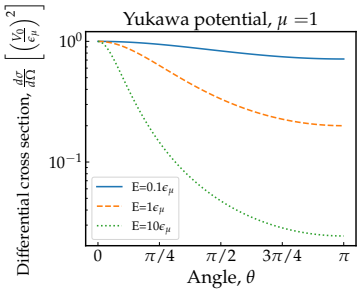

Figure 8.3: The Yukawa potential differential cross section as a function of the angle. This is shown in units of $\left(\frac{V_0}{\varepsilon_\mu}\right)^2$ as a function of the angle for three energies: $E = 0.1\varepsilon_\mu$ (solid line), $E = \varepsilon_\mu$ (dashed line) and $E = 10\varepsilon_\mu$ (dotted line).

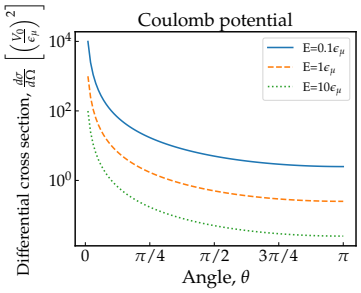

Figure 8.4: The Rutherford scattering differential cross section as a function of the angle. For this plot, we choose $V_0 = 1$ and energies $E = 0.1V_0$ (solid line), $E = V_0$ (dashed line) and $E = 10V_0$ (dotted line).

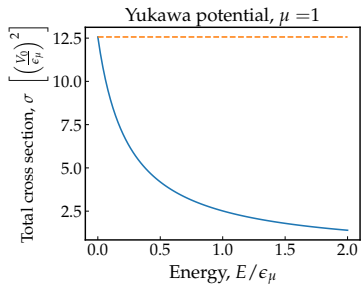

Figure 8.5: The Yukawa potential total cross section as a function of the energy in units of the typical energy scale ε_μ. The cross section has units of $\left(\frac{V_0}{\varepsilon_\mu}\right)^2$. The dashed line corresponds to the zero energy result, $\sigma \to 4\pi\left(\frac{V_0}{\varepsilon_\mu}\right)^2$.

8.2 *Formal theory of scattering*

The discussion of scattering in the preceding subsection was based on the Fermi golden rule in the setting described by time-dependent perturbation theory. Several approximations were required (eg no multi-step processes) and the formal connection with time-dependent theory is somewhat tenuous. Scattering theory can be reformulated from the ground up by providing a more formal approach as we shall see now.

We assume that we can use an initial state which is an eigenstate of an unperturbed Hamiltonian $\hat{H}_0$ in the remote past, $|i(t=-\infty)\rangle$. In most situations, this is just the non-interacting hamiltonian, $\hat{H}_0 = \frac{\hat{p}^2}{2m}$. The eigenstate wavepacket then propagates in time and evolves through a region where it interacts with a target. This is described by a total Hamiltonian $\hat{H} = \hat{H}_0 + \hat{V}(t)$ which describes the projectile + target system. In the interaction representation, the initial state, which is not an eigenstate of $\hat{H}$, will evolve in time through the equation

$$|i(t)\rangle = \hat{U}_I(t,-\infty)|i(t=-\infty)\rangle\,, \tag{8.18}$$

where $\hat{U}_I$ is the time evolution operator in the interaction picture.

In a scattering experiment, one typically measures a final state some time after the reaction process has occurred, $|f(t\to\infty)\rangle$. The transition amplitude between the two states is thus given by the overlap of the two states (measured at one and the same time, here chosen to be $t=\infty$),

$$a_{i\to f} = \langle f(t=\infty)|i(t=\infty)\rangle = \langle f(t=\infty)|\underbrace{\hat{U}_I(\infty,-\infty)}_{\hat{S}}|i(t=-\infty)\rangle\,.$$

We define the scattering or $S-$*matrix* as the operator connecting the two remote past and future states,

$$\hat{S} = \hat{U}_I(\infty,-\infty)\,. \tag{8.19}$$

Transition amplitudes are just matrix elements of the $S-$matrix,

$$a_{i\to f} = \langle f|\hat{S}|i\rangle = S_{fi}\,. \tag{8.20}$$

We use the interaction representation in this setting because it allows us to get a closed expression and an expansion for these matrix elements. Considering Eq. (7.82) for $t_0\to-\infty$ and $t\to\infty$, we immediately find

$$S_{fi} = a_{i\to f} = \mathcal{T}\left[-\frac{i}{\hbar}\int_{-\infty}^{\infty} d\bar{t}\, V_I(\bar{t})\right]_{fi}\,. \tag{8.21}$$

Further, using the Dyson expansion in Eq. (7.80), one finds the perturbative expansion

$$S_{fi} = \delta_{fi} + \sum_{n=1}\left(-\frac{i}{\hbar}\right)^n \frac{1}{n!}\int_{-\infty}^{\infty}dt_1\int_{-\infty}^{\infty}dt_2\cdots\int_{-\infty}^{\infty}dt_n\;\mathcal{T}\left[\hat{V}_I(t_1)\hat{V}_I(t_2)\cdots\hat{V}_I(t_n)\right]_{fi}\,. \tag{8.22}$$

The transition probabilities are then given by the expression $\mathcal{P}_{i\to f} = |S_{fi}|^2$. In other words, the modulus squared of the $S-$matrix are the transition probabilities.

The $S-$matrix inherits some useful properties from the time evolution operator, $\hat{U}$. First and foremost, it is unitary, $\hat{S}\hat{S}^\dagger = \hat{S}^\dagger\hat{S} = \hat{I}$. From this, one can derive a completeness relation:

$$\delta_{fi} = \langle f|\hat{S}^\dagger\hat{S}|i\rangle = \sum_n \langle f|\hat{S}^\dagger|n\rangle\langle n|\hat{S}|i\rangle = \sum_n S^\dagger_{fn}S_{ni} = \sum_n S^*_{nf}S_{ni}\,.$$

More specifically, for $f = i$, we find the conservation of probability law

$$\sum_n |S_{ni}|^2 = \sum_n \mathcal{P}_{i\to n} = 1\,.$$

The $S-$matrix is somewhat more fundamental than the Hamiltonian, in that it describes the full evolution of quantum systems. Some hamiltonian-independent formulations of quantum mechanics can be formulated in terms of the $S-$matrix alone. Moreover, the $S-$matrix must know about all the conservation laws in the system. A state i belonging to a given constant of motion (or a given symmetry class) can never be connected via the time evolution to a state f belonging to a different constant of motion (or a different symmetry class). Stated differently, the $S-$matrix must be diagonal in the quantum numbers of commuting conserved operators. For instance, if the Hamiltonian is rotationally invariant, it conserves angular momentum. The $S-$matrix is then diagonal in the angular momentum quantum number, S_l. This provides a substantial simplification and is generally helpful in determining the $S-$matrix matrix elements.

Separating explicitly the $i \to i$ transitions and taking energy conservation into account, one usually defines the so-called $T-$matrix implicitly as,

$$S_{fi} = \delta_{fi} - 2\pi i\delta(\epsilon_f - \epsilon_i)T_{fi}\,. \tag{8.23}$$

The $T-$matrix in a sense corresponds to the "interaction" or "perturbation" part of the transition $i \to f$. In fact, by taking the square of this expression[1] carefully, one can see that the "exact" transition rate is given by the expression

[1] V. Zelevinsky. *Quantum Physics*, volume 2. Wiley-VCH, 2011

$$\mathcal{W}_{i\to f} = \frac{2\pi}{\hbar}|T_{fi}|^2\delta(\epsilon_f - \epsilon_i)\,, \tag{8.24}$$

where we have for brevity ignored the subtleties associated to continuous final states. This is a more general and formally correct expression for Fermi's golden rule. We can see the connection between this approach and our previous derivation by looking at the expansion of Eq. (8.22). Considering the first-order term,

$$S^{(1)}_{fi} = \left(-\frac{i}{\hbar}\right)\int_{-\infty}^{\infty} dt_1 \ \left[\hat{V}_I(t_1)\right]_{fi} = -\frac{i}{\hbar}\int_{-\infty}^{\infty} dt_1 \langle f|e^{i\frac{\hat{H}_0 t_1}{\hbar}}\hat{V}e^{-i\frac{\hat{H}_0 t_1}{\hbar}}|i\rangle$$
$$= -\frac{i}{\hbar}\langle f|\hat{V}|i\rangle\int_{-\infty}^{\infty} dt_1 e^{i(\omega_f-\omega_i)t_1} = -\frac{2\pi i}{\hbar}V_{fi}\delta(\omega_f-\omega_i) = -2\pi i V_{fi}\delta(\epsilon_f-\epsilon_i)$$

where we have assumed a time-independent perturbation $\hat{V}$. Comparing this equation to the definition of Eq. (8.23), it is clear that at lowest order in perturbation theory $T_{fi} \approx V_{fi}$. When computing the Fermi golden rule, one should indeed use the $T-$matrix rather than the interaction in the matrix elements. Higher order perturbation theory corrections change the expression for the $T-$matrix by allowing multi-step processes involving not just a single interaction, V, but other terms proportional to V^2, V^3, etc.

8.2.1 *Scattering amplitude and the cross section*

How do these formal developments translate into useful properties for scattering calculations? We consider the same set-up like the one described in Fig. 8.1. The beam travels along the z axis with momentum k. A particle in this beam has an incident wavefunction

$$\psi_{\text{in}}(r) = e^{i\mathbf{k}\cdot\mathbf{z}}$$

We use a different normalization to what was discussed in Sec. 4.5. The final cross section should not be affected by this choice.

A detector is placed at a distance $r \gg 1$ from the target, forming an angle (θ, ϕ) with respect to the beam's direction. This detector registers elastically scattered particles in a detector area of solid angle $\Delta\Omega$ in the direction $\mathbf{r} = r\boldsymbol{e_r}(\theta, \phi)$. We assume that, after interacting with the target, the particles move freely away, and thus propagate radially outwards, $\mathbf{k}' = k'\boldsymbol{e_r}$.

In the asymptotic region $r \gg 1$, the wave function is still a superposition of solutions of the Schrödinger equation for free motion for the incident and the scattered wave functions. We choose to parametrize the scattered wave function as

$$\psi_{\text{sc}}(r) = f(\mathbf{k}', \mathbf{k})\frac{e^{ik'r}}{r}\,.$$

The term $e^{ik'r}/r$ is analogous to the radial solution of a non-interacting Schrödinger equation. The quantity $f(\mathbf{k}', \mathbf{k})$ goes by the name of *scattering amplitude* and has units of length. In elastic scattering, since $k = k'$, the scattering amplitude is only a function of the momentum value k and the angle $\Omega = (\theta, \phi)$, between $\mathbf{k}$ and $\mathbf{k}$. It is typically written as $f(\Omega)$, and the dependence in energy is assumed implicitly. The scattering amplitude encodes all the information about the scattering process and can be directly related to the $S-$matrix. At any one time, the wave function of the target+beam system is the superposition

$$\psi(r) = \psi_{\text{in}}(r) + \psi_{\text{sc}}(r) = \psi_{\text{in}}(r) + f(\Omega)\frac{e^{ik'r}}{r} \tag{8.25}$$

How can we connect this generic description to a measurable quantity like the cross section? We can compute $\frac{d\sigma}{d\Omega}$ in analogy to what we did in Section 8.1.2. The key is to find expressions for the flux of incoming and scattered particles. For the incoming beam, the incident flux is obtained as the projection of the flux associated along the $z-$axis, and the result is $j_{\text{b}} = \frac{\hbar k}{m} = v$ as explained in the context of Eq. (8.7).

The count rate $N(\Omega, \Delta\Omega)$ is proportional to the flux of scattered current, $\mathbf{j}_{\rm sc}$, through the detector surface. The vector normal to this surface points along the radial direction, and so the (finite) differential of surface vector is $\Delta\mathbf{S} = \Delta S \boldsymbol{e}_r = \Delta\Omega r^2 \boldsymbol{e}_r$. The count rate must be related to the projection of the flux onto this surface,

$$N(\Omega, \Delta\Omega) = \mathbf{j}_{\rm sc} \cdot \Delta\mathbf{S} = \Delta\Omega\, r^2 \boldsymbol{e}_r \cdot \mathbf{j}_{\rm sc}.$$

The projection along the radial direction of the flux singles out the radial component of the gradient, $\boldsymbol{e}_r \cdot \nabla = \frac{\partial}{\partial r}$, so one finds

$$\begin{aligned}\boldsymbol{e}_r \cdot \mathbf{j}_{\rm sc} &= -i\frac{\hbar}{2m}\left[f^*(\Omega)\frac{e^{-ikr}}{r}\frac{\partial}{\partial r}\left(f(\Omega)\frac{e^{ikr}}{r}\right) - f(\Omega)\frac{e^{ikr}}{r}\frac{\partial}{\partial r}\left(f^*(\Omega)\frac{e^{-ikr}}{r}\right)\right] \\ &= -i\frac{\hbar}{2m}|f(\Omega)|^2\left(\frac{2ik}{r^2}\right) = \frac{\hbar k}{m}\frac{|f(\Omega)|^2}{r^2} = v\frac{|f(\Omega)|^2}{r^2}\,.\end{aligned}$$

With this, the count rate becomes

$$N(\Omega, \Delta\Omega) = v|f(\Omega)|^2\Delta\Omega,$$

and the cross section from Eq. (8.5) is

$$\frac{d\sigma}{d\Omega} = \frac{v|f(\Omega)|^2\Delta\Omega}{\Delta\Omega\, v} = |f(\Omega)|^2\,. \tag{8.26}$$

You will notice that we take $n_{\rm t+b} = 1$ here. Indeed, we assume that a single beam particle interacts with a single target particle per unit of volume. More complex configurations do not change this result.

We stress that the result for the cross section is independent of the wave function normalization because it is the ratio of two fluxes, and normalization constants cancel in this ratio.

8.3 *Partial waves*

As we discussed earlier in Section 8.1.3, central interactions are those that do not depend on the direction of the relative position between beam and target particles, $V(\mathbf{r}) = V(r)$. This simplifies things substantially. Because the interaction does not depend on angles, the relative angular momentum between projectile and target commutes with the Hamiltonian of the system. The angular momentum, $\hat{L}$, between the projectile and the target is thus a conserved quantity that cannot change across the collision. A consequence of this fact is that we can treat **separately** the scattering of individual partial waves in the wave function.

To start with, the $S-$matrix reflects this fact. Because different values of l cannot be connected between the initial and final states, S must be diagonal in l, $S_{ll'} = S_l\delta_{ll'}$. Moreover, because S is unitary, we have that

$$\hat{S}\hat{S}^\dagger = \hat{I} \Rightarrow \sum_n S^*_{nf}S_{ni} = \sum_n S_f\delta_{nf}S^*_i\delta_{ni} = |S_i|^2\delta_{ij} = \delta_{ij}$$

which shows that $|S_l|^2 = 1$. In other words, S_l is a complex number of unit modulus, only given by a phase which we choose to represent as

$$S_l = e^{i2\delta_l}\,.$$

The parameter $\delta_l = \delta_l(k)$ is the so-called phase-shift. This, again, depends implicitly on energy. One can think of the phase shift as the additional phase that a scattered wave collects with respect to an incoming one.

The general strategy to compute the phaseshift is as follows. First, one looks at the properties of an incident non-interacting wave, $\psi_{\text{in}}(r)$. This can be expressed as a sum of several partial waves, which we can take as a reference for the scattering problem. In this general scattering scenario, we can also compute the total wave function in terms of partial waves and find how partial waves are modified by the scattering amplitude, $f(\Omega)$. Second, one solves the Schrödinger equation for the interaction potential between the beam and the target. This is a three-dimensional problem, and we know how to express it in partial waves as described in Section 5.2. Far away enough from the origin, the corresponding wavefunction is shifted from the non-interacting case by an angle δ_l. Matching the solutions of the first and the second terms, we can find the corresponding phase shifts. We now describe the mathematical developments that are necessary to perform this matching procedure. In the next section, you can also find specific (simple) examples for a square potential well and barrier in spherical coordinates.

8.3.1 *Scattering wavefunction partial wave expansion*

Let us see how we arrive at this result by using the spherical harmonics and the partial wave expansion of Section 5.2. As we discussed there, the solution of the three-dimensional Schrödinger equation is a linear combination of simultaneous eigenfunctions of $\hat{H}, \hat{L}^2, \hat{L}_z$, so

$$\psi_{k,l,m}(r,\theta,\phi) = R_{kl}(r)Y_{lm}(\theta,\phi)\,. \tag{8.27}$$

Y_{lm} are spherical harmonics and $R_{kl}(r)$ are the solutions of the radial equation,

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right) - \frac{l(l+1)}{r^2}\right]R_{kl}(r) + V(r)R_{kl}(r) = ER_l(r)\,. \tag{8.28}$$

We use the wavenumber $k = \sqrt{2mE}/\hbar$ to characterise the collision energy. Each term of Eq. (8.27) is a *partial wave* from which the solution to the scattering problem can be built. We rewrite this equation by introducing a new variable $\rho = kr$ and reshuffling terms to get

$$\left[\frac{d^2}{d\rho^2} + \frac{2}{\rho}\frac{d}{d\rho} + \left(1 - \frac{l(l+1)}{\rho^2}\right)\right]R_{kl}(\rho) = U(\rho)R_{kl}(\rho)\,, \tag{8.29}$$

with $U(\rho) = V(\rho/k)/E$. One can also use the reduced wavefunction $u_{kl}(\rho) = rR_{kl}(\rho) = \frac{\rho}{k}R_{kl}(\rho)$, so the differential equation simplifies

further

$$\left[\frac{d^2}{d\rho^2} + \left(1 - \frac{l(l+1)}{\rho^2}\right)\right] u_{kl}(\rho) = U(\rho) u_{kl}(\rho)\,. \tag{8.30}$$

We start by considering the non-interacting case, $U = 0$. Equation (8.29) is the spherical Bessel equation[2] and its solution can be expressed as linear combinations of the pair of real solutions $j_l(\rho)$ (spherical Bessel functions) and $n_l(\rho)$ (Neumann functions). The former are regular at the origin, and the latter diverge as $\rho \to 0$. For $l = 0$ and $l = 1$ these are:

[2] G. B. Arfken, H. J. Weber, and F. E. Harris. *Mathematical Methods for Physicists*. Elsevier, 2005; and K. F. Riley and M. P. Hobson. *Essential Mathematical Methods for the Physical Sciences*. Cambridge University Press, 2011

$$j_0(\rho) = \frac{\sin\rho}{\rho}, \quad j_1(\rho) = \frac{\sin\rho}{\rho^2} - \frac{\cos\rho}{\rho}\,, \tag{8.31}$$

$$n_0(\rho) = \frac{\cos\rho}{\rho}, \quad n_1(\rho) = \frac{\cos\rho}{\rho^2} + \frac{\sin\rho}{\rho}\,. \tag{8.32}$$

The lth functions may be obtained from successive derivations,

$$j_l(\rho) = (-\rho)^l \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^l j_0(\rho)\,,$$
$$n_l(\rho) = (-\rho)^l \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^l n_0(\rho).$$

We shall need the limiting values of these functions in the asymptotic region, where $\rho \gg 1$,

$$j_l(\rho \to \infty) \approx \frac{\sin\left(\rho - \frac{l\pi}{2}\right)}{\rho}\,, \tag{8.33}$$

$$n_l(\rho \to \infty) \approx \frac{\cos\left(\rho - \frac{l\pi}{2}\right)}{\rho}\,. \tag{8.34}$$

If the potential has a finite range, R, there is always an asymptotic region, $r \gg R$, where the solution of Eq. (8.29) is similar to the free, $U = 0$ case. In that region, we can expand the incident wave function in terms of partial waves using the Bauer partial-wave expansion

Here, we make use of the mathematical identity called the *addition theorem*, $4\pi \sum_m Y^*_{lm}(\boldsymbol{e_r}) Y_{lm}(\boldsymbol{e_k}) = (2l+1) P_l(\boldsymbol{e_r} \cdot \boldsymbol{e_k})$ with $\cos\theta$ incorporating the angle between $\mathbf{e_r}$ and $\mathbf{e_k}$.

$$\begin{aligned}\psi_{\text{in}}(r) = e^{ikz} = e^{ikr\cos\theta} = e^{i\rho\cos\theta} &= \sum_{lm} 4\pi Y^*_{lm}(\boldsymbol{e_r}) Y_{lm}(\boldsymbol{e_k}) i^l j_l(\rho) \\ &= \sum_l (2l+1)\, P_l(\cos\theta) i^l j_l(\rho)\,,\end{aligned} \tag{8.35}$$

This indicates that the plane wave e^{ikz} can be understood as a sum of partial waves, each one with orbital angular momentum $\sqrt{l(l+1)}\hbar$. The terms $P_l(\cos\theta)$ and $j_l(\rho)$ specify separately the angular, θ, and radial, ρ, dependence of the each partial wave l in the sum.

Using the asymptotic expansion of Eq. (8.33), we expect that the

incoming plane wave reads in the asymptotic region:

$$\psi_{\rm in}(r\to\infty) = \sum_l (2l+1)\, P_l(\cos\theta) i^l \frac{\sin\left(\rho - \frac{l\pi}{2}\right)}{\rho} \tag{8.36}$$
$$= \frac{1}{2i}\sum_l (2l+1)\, i^l P_l(\cos\theta)\left[\frac{e^{i\left(kr-\frac{l\pi}{2}\right)} - e^{-i\left(kr-\frac{l\pi}{2}\right)}}{kr}\right] =$$
$$= \frac{e^{-ikr}}{r}\sum_l P_l(\cos\theta)\left[-\frac{2l+1}{2ik} i^l e^{i\frac{l\pi}{2}}\right]$$
$$+ \frac{e^{ikr}}{r}\sum_l P_l(\cos\theta)\left[\frac{2l+1}{2ik} i^l e^{-i\frac{l\pi}{2}}\right]$$

Note that the radial dependence of the wave function is all encoded in the exponential factors, whereas the angular dependence is captured by the sums over l, which we denote

These factors have an intrinsic dependence in $\cos\theta$ and k, which we ignore for simplicity.

$$X^{\rm in}_{\pm} = \mp\sum_l (2l+1)\, P_l(\cos\theta)\left[\frac{2l+1}{2ik} i^l e^{\mp i\frac{l\pi}{2}}\right]$$
$$= \mp\sum_l (2l+1)\,(\mp 1)^l \left[\frac{2l+1}{2ik}\right] P_l(\cos\theta),$$

where we have used $i^l e^{\pm i\frac{l\pi}{2}} = (\mp 1)^l$. With this, we find an expression for the asymptotic expansion of the $z-$directed incident scattering wave,

$$\psi_{\rm in}(r\to\infty) = \frac{e^{-ikr}}{r} X^{\rm in}_- + \frac{e^{ikr}}{r} X^{\rm in}_+ , \tag{8.37}$$

which separates into incoming and outgoing terms (with a $-$ and a $+$ sign, respectively).

This expression can easily be incorporated into the total scattering wave function of Eq. (8.25), which includes both the incoming and scattered components. With $\psi_{\rm sc}(r) = f(\Omega)\frac{e^{ik'r}}{r}$, we write

$$\psi_{\rm sc}(r\to\infty) = \frac{e^{-ikr}}{r} X^{\rm sc}_- + \frac{e^{ikr}}{r} X^{\rm sc}_+ , \tag{8.38}$$

where the angular-dependent factors are now

$$X^{\rm sc}_- = -\sum_l P_l(\cos\theta)\left[\frac{2l+1}{2ik}(-1)^l\right] \tag{8.39}$$

$$X^{\rm sc}_+ = f(\theta,\phi) + \sum_l P_l(\cos\theta)\left[\frac{2l+1}{2ik}\right] . \tag{8.40}$$

With a central potential and the beam in the $z-$direction, the scattering problem should be independent of the angle ϕ. As a consequence, the scattering amplitude is such that $f(\theta,\phi) = f(\theta)$. One can define a partial wave scattering amplitudes, f_l, so that

$$f(\theta) = \frac{1}{k}\sum_l (2l+1) P_l(\cos\theta) f_l, \tag{8.42}$$

which in turn translates into the following expansion for the angular dependence of the $+$ component,

$$X^{\rm sc}_+ = \frac{1}{k}\sum_l (2l+1) P_l(\cos\theta)\left[f_l - \frac{i}{2}\right] . \tag{8.43}$$

The Legendre polynomials $P_l(x)$ are complete and orthogonal in the interval $[-1,1]$, and fulfill the relations

$$\int_{-1}^{1} dx\, P_l(x) P_{l'}(x) = \frac{2}{2l+1}\delta_{l,l'}. \tag{8.41}$$

It follows that any function of $x = \cos\theta$ can be expanded as

$$f(x) = \sum_l a_l P_l(x) ,$$

which is precisely what we do in Eq. (8.42).

8.3.2 *Schrödinger wavefunction partial wave expansion*

On the other hand, we know that the scattering solution to the Schrödinger equation can be found by solving Eq. (8.28). The solutions to that equation are linear cominations of the solutions shown in Eq. (8.27),

$$\psi_{\text{Sch}}(\mathbf{r}) = \sum_{lm} D_{lm} R_{kl}(r) Y_{lm}(\theta, \phi), \tag{8.44}$$

where the D_{lm} factors would be obtained upon solving the equation. We can now introduce the reduced wavefunction $u_{kl} = rR_{kl}$ and use the fact that $\phi-$independent solutions of the spherical harmonics, Y_{lm}, are essentially Legendre polynomials, P_l. With this, the previous equation becomes

$$\psi_{\text{Sch}}(\mathbf{r}) = \sum_{l} C_l \frac{u_l(kr)}{kr} P_l(\cos\theta)\,. \tag{8.45}$$

where again the C_l factors are obtained from a solution to the equation.

Far away from the scattering point at the origin (or near the detector, placed away at $r \to \infty$) the asymptotic expansion must be the sum of an incoming term and an outgoing term. The incoming waves have not yet met the target, so they are not modified by scattering. They must be similar to the non-interacting solutions of Eq. (8.37). The outgoing partial waves scatter independently. By unitarity of the $S-$matrix, they can only change by a phase. Taking into account that the outgoing term is thus distorted by the $S-$matrix, the Schrödinger wavefunction in the asymptotic region must be

$$\begin{aligned}\psi_{\text{Sch}}(r \to \infty) &= \frac{1}{2i} \sum_l C_l P_l(\cos\theta) \left[\frac{S_l e^{i\left(kr - \frac{l\pi}{2}\right)} - e^{-i\left(kr - \frac{l\pi}{2}\right)}}{kr} \right] \\ &= \frac{e^{-ikr}}{r} X_-^{\text{Sch}} + \frac{e^{ikr}}{r} X_+^{\text{Sch}}\,. \end{aligned} \tag{8.46}$$

with

$$X_-^{\text{Sch}} = \sum_l P_l(\cos\theta) \left[-\frac{C_l}{2ik} e^{i\frac{l\pi}{2}} \right] = -\sum_l P_l(\cos\theta) \left[\frac{C_l}{2ik} i^l \right] \tag{8.47}$$

$$X_+^{\text{Sch}} = \sum_l P_l(\cos\theta) \left[\frac{C_l}{2ik} i^{-l} S_l \right] \tag{8.48}$$

In the asymptotic region, Eqs. (8.38) and (8.46) must yield the same results With the $+$ and $-$ components being linearly independent, this immediately implies the following equalities between the angular coefficients of the two equations:

$$X_-^{\text{Sch}} = X_-^{\text{sc}}\,, \quad X_+^{\text{Sch}} = X_+^{\text{sc}}. \tag{8.49}$$

For the $-$ components, the equation allows us to find $C_l = (2l + 1)i^l$. For the $+$ components, the equality provides a relation between

the phase-shifts, δ_l, and the scattering amplitude partial wave f_l,

$$\left.\begin{aligned} X_+^{\text{Sch}} &= \sum_l P_l(\cos\theta)(2l+1)\left[\frac{1}{2ik}S_l\right] \\ X_+^{\text{sc}} &= \frac{1}{k}\sum_l P_l(\cos\theta)(2l+1)\left[f_l - \frac{i}{2}\right] \end{aligned}\right\} \Rightarrow \frac{1}{2ik}S_l = \frac{1}{k}\left[f_l - \frac{i}{2}\right]$$

In turn, we can isolate for the partial wave scattering amplitude in terms of the $S-$matrix elements,

$$f_l = \frac{1}{2i}\left[S_l - 1\right] . \tag{8.50}$$

With $S_l = e^{2i\delta_l}$, one can also write the scattering amplitude in terms of the phase shifts,

$$f_l = e^{i\delta_l}\sin\delta_l . \tag{8.51}$$

The partial wave expansion of the scattering amplitude is, in turn,

$$f(\theta) = \frac{1}{k}\sum_l (2l+1)P_l(\cos\theta)e^{i\delta_l}\sin\delta_l \tag{8.52}$$

There are several equivalent ways in which one can write the previous expressions. Consider a case in which the $l = 0$ is the dominant contribution in the previous expansion. Because $P_0(x) = 1$, one can rewrite the full scattering amplitude as the following angular independent contribution

$$f(\theta) \approx \bar{f} = \frac{\sin\delta_0}{k}e^{i\delta_0} = \frac{\sin\delta_0}{ke^{-i\delta_0}} = \frac{\sin\delta_0}{k(\cos\delta_0 - i\sin\delta_0)} = \frac{1}{k\cot\delta_0 - ik} \tag{8.53}$$

This expression for the scattering amplitude is often used whenever the scattering energy is low. An example of such a situation is the scattering of particles in low density gases.

8.3.3 *Partial waves and scattering cross sections*

So far, we have used a matching procedure to find the scattering amplitude in terms of the scattering phase shift. The first quantity is a useful parametrization of scattering properties. The second quantities, in contrast, can be found by solving the Schrödinger equation in the asymptotic region. How can we relate these properties to obtain scattering observables, like the cross section?

We start by using Eq. (8.26) for the scattering cross section. Introducing the partial wave expansion of the scattering cross section, Eq. (8.42), we find:

$$\begin{aligned} \frac{d\sigma}{d\Omega} = |f(\Omega)|^2 &= \frac{1}{k^2}\sum_{ll'}(2l+1)(2l'+1)P_l(\cos\theta)P_{l'}(\cos\theta)f_l^* f_{l'} = \\ &= \frac{1}{k^2}\sum_{ll'}(2l+1)(2l'+1)P_l(\cos\theta)P_{l'}(\cos\theta)\sin^2\delta_l . \end{aligned} \tag{8.54}$$

In other words, if the scattering phase shifts for all the values of l are known, one can compute the differential cross section. Further,

using the orthogonality property of Eq. (8.41), one can find the total scattering cross section

$$\sigma = \int d\Omega \frac{d\sigma}{d\Omega} = \frac{2\pi}{k^2} \times \sum_l (2l+1)^2 \times \frac{2}{2l+1} \times \sin^2 \delta_l$$
$$= \frac{4\pi}{k^2} \sum_l (2l+1) \sin^2 \delta_l \,. \tag{8.55}$$

This expression provides the partial wave expansion for the total cross section in terms of the phase shifts. For a given collision energy, each partial contributes a term of size $\sigma_l = \frac{4\pi}{k^2}(2l+1)\sin^2\delta_l$ to the wave function. The minimum contribution to the cross sections occurs when $\delta_l = 0$ or $\delta_l = n\pi$, in which case the contribution is exactly zero and the partial wave is invisible to scattering. The maximum contribution, in contrast, happens for $\delta_l = (2n+1)\frac{\pi}{2}$, for which

$$\sigma_l^{\max} = \frac{4\pi}{k^2}(2l+1) \,. \tag{8.56}$$

We shall see some examples of both extremes below.

Equation (8.55) clearly shows that the phase-shifts are key objects in the calculation of cross sections. We shall see below how these are computed in some particularly simple examples. In the most general case, one typically uses rewrites Eq. (8.46) as

$$\psi_{\mathbf{k}}^{\text{Sch}}(r) = \sum_l (2l+1) i^l e^{i\delta_l} P_l(\cos\theta) \frac{\sin(kr - \frac{1}{2}l\pi + \delta_l)}{kr} \,. \tag{8.57}$$

This is the wave function at asymptotically large distances, $r \gg R$, well beyond the range of the potential R. Let us assume that one can solve the Schrödinger equation for a region $r < R$, providing a wave function $\psi^{(1)}$ that depends on the energy k. This solution, and the asymptotic one, must coincide at some pointe where $r \approx R$. This matching procedure allows one to compute δ_l from the known, interior solution of the Schrödinger equation.

8.4 *Examples*

Let us now consider a few examples of scattering from simple potentials. These solutions will be useful in characterising the general properties of the scattering phase shifts. We shall start by considering the scattering off a repulsive and attractive square well. You can find more details about these in the Jupyter Notebooks Scattering_spherical.ipynb available at the Master's github repository shorturl.at/fEXZ5.

8.4.1 *Repulsive square well: scattering below the barrier*

Let us consider the case of a repulsive ($V_0 > 0$) square well potential,

$$V(r) = \begin{cases} V_0, \, r \leq R, \\ 0, \, r > R \,. \end{cases} \tag{8.58}$$

This potential is schematically depicted in Fig. 8.6. The Schrödinger equation for the radial component of the wavefunction, Eq. (8.28), is substantially simplified if we consider $R_l(k,r) = \frac{u_l(k,r)}{r}$. With this change, the equation becomes:

$$-\frac{\hbar^2}{2m}\frac{1}{r}\frac{d^2}{dr^2}u_l(k,r) + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^3}u_l(k,r) + V(r)\frac{u_l(k,r)}{r} = E\frac{u_l(k,r)}{r}\,.$$

Multiplying by r and $-2m/\hbar^2$, we get:

$$\frac{d^2}{dr^2}u_l(k,r) - \frac{l(l+1)}{r^2}u_l(k,r) - \frac{2mV(r)}{\hbar^2}u_l(k,r) = -\frac{2mE}{\hbar^2}u_l(k,r)\,. \tag{8.59}$$

We shall only consider for simplicity the $S-$wave scattering case, for which $l = 0$. In this case, the centrifugal term drops and the equation is a second-order homogeneous differential equation:

$$\frac{d^2}{dr^2}u_0(k,r) + \frac{2m[E - V(r)]}{\hbar^2}u_0(k,r) = 0\,. \tag{8.60}$$

For the repulsive square well, we have two regions of space with different equations depending on whether $r < R$ or $r > R$,

$$\frac{d^2u_0}{dr^2} + \frac{2m[E - V_0]}{\hbar^2}u_0 = 0\,, \quad r \leq R, \tag{8.61}$$

$$\frac{d^2u_0}{dr^2} + \frac{2mE}{\hbar^2}u_0 = 0\,, \quad r > R. \tag{8.62}$$

Introducing the wave-vector $k^2 = \frac{2mE}{\hbar^2}$, the second equation is just an harmonic equation. The first equation is also harmonic, but with a different wave-vector, $\bar{k}^2 = k^2 - k_0^2$. Here, we have defined the momentum scale associated to the square well as $k_0 = \sqrt{\frac{2mV_0}{\hbar^2}}$. With this, the two equations are

$$\frac{d^2u_0}{dr^2} + \bar{k}^2u_0 = 0\,, \quad r \leq R, \tag{8.63}$$

$$\frac{d^2u_0}{dr^2} + k^2u_0 = 0\,, \quad r > R. \tag{8.64}$$

The solution to the first equation is different depending on whether $\bar{k}^2$ is positive or negative. We start with the negative case, which corresponds to $k^2 < k_0^2$ and, hence, to $E < V_0$. Physically, this represents scattering at energies below the square well barrier. Mathematically, this means that $\bar{k}$ is a pure imaginary number, $\bar{k} = i\kappa$, with $\kappa = \sqrt{k_0^2 - k^2} > 0$.

The solution of the two equations in the two spatial regions are essentially exponentials of imaginary or real numbers times r. The solutions in the interior region, $r < R$, need to be regular at the origin, so $u_0(r \to 0) \to 0$. In contrast, in the exterior region, $r > R$, the waves are asymptotically free but modified by the phase shift $u_0(r \to \infty) \to \sin(kr + \delta_0)$ as dictated by Eq. (8.36) with $l = 0$. With this, we can immediately guess the following two solutions:

$$\frac{d^2u_0}{dr^2} + \bar{k}^2u_0 = 0\,, r \leq R \Rightarrow u_0(r) = A\sinh\kappa r \tag{8.65}$$

$$\frac{d^2u_0}{dr^2} + k^2u_0 = 0\,, r > R \Rightarrow u_0(r) = B\sin\left(kr + \delta_0\right). \tag{8.66}$$

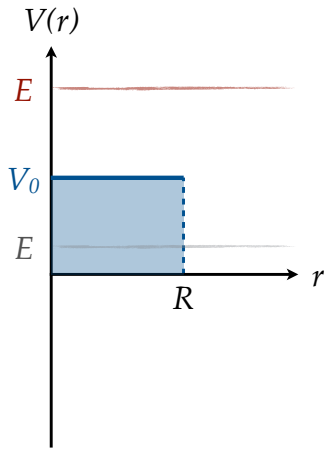

Figure 8.6: The repulsive square well is a positive constant potential, $V_0 > 0$, between $r = 0$ and $r = R$. The equations for the case in which the particle scattering energy is $E < V_0$ and $E > V_0$ are slightly different, and discussed in the text.

To determine δ_0, we look at the wave-functions at values $R_- = R - \epsilon$, right below R, and at $R_+ = R + \epsilon$, right above R with $\epsilon \to 0$. A physically valid matching can only be achieved if the wavefunctions fulfil $u_0(R_-) = u_0(R_+)$ and if their derivatives fulfill $\left.\frac{du_0}{dr}\right|_{r=R_-} = \left.\frac{du_0}{dr}\right|_{r=R_+}$. An example for this matching procedure is shown in Fig. 8.7.

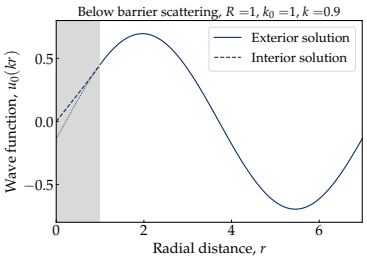

Figure 8.7: The wavefunctions in the interior region, $r < R$ (grey area), and the exterior region, $r > R$, must match at $r = R$. For this to happen, the phaseshift must be given by Eq. (8.68). The dotted line in the interior region represents the continuation of the exterior wavefunction into the interior region.

A useful way to compute the phase shift straight away is to consider the logarithmic derivative,

$$\left.\frac{1}{u_0}\frac{du_0}{dr}\right|_{R_-} = \left.\frac{1}{u_0}\frac{du_0}{dr}\right|_{R_+} . \tag{8.67}$$

This allows us to compute the phaseshifts directly without having to worry about the normalization constants A and B. Taking the logarithmic derivatives for the solutions in the interior region, Eq. (8.65), and in the exterior region, Eq. (8.66), we find:

$$\frac{1}{A\sinh\kappa R}A\kappa\cosh\kappa R = \frac{1}{B\sin(kR+\delta_0)}B\sin(kR+\delta_0)$$
$$\frac{1}{\tanh\kappa R} = \frac{1}{\tan(kR+\delta_0)}$$

This is an equation for the phaseshift, which can be isolated to find:

$$\begin{aligned}\delta_0 &= \arctan\left[\frac{k}{\kappa}\tanh(\kappa R)\right] - kR \\ &= \arctan\left[\frac{k}{\sqrt{k_0^2-k^2}}\tanh\left(\sqrt{k_0^2-k^2}R\right)\right] - kR\,,\end{aligned} \tag{8.68}$$

The argument of the arctan function may diverge as the energy increases and $k \to k_0$, but the phaseshift remains finite because the arctan function itself is bound between -1 and 1.

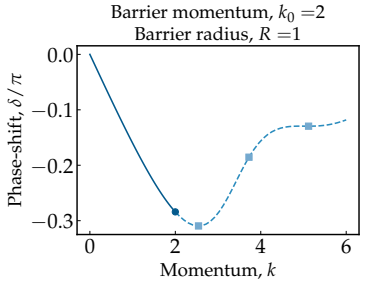

Figure 8.8: The $S-$wave scattering phase shift for a repulsive square well with $R = 1$ and $k_0 = 2$. The solid line corresponds to Eq. (8.68), for values $k < k_0$. The dashed line corresponds to Eq. (8.77), for values $k > k_0$. The filled circle marks the point at which $k = k_0$. The filled square correspond to the first three "resonance" values of Eq. (8.78).

The phaseshift as a function of k is shown in Fig. 8.8 for an arbitrary value of $k_0 = 0$. The grey solid line corresponds to the case $k < k_0$ that we have just discussed. The expression is valid from $k = 0$ to $k = k_0$. The behaviour at low values of momenta, $k \ll k_0$, appears to be linear. This can be justified from a low-momentum expansion in Eq. (8.68), which indicates

$$\begin{aligned}\delta_0 &\approx \arctan\left[\frac{k}{k_0}\tanh k_0R\right] - kR \approx \left[\frac{k}{k_0}\tanh k_0R\right] - kR \\ &= -kR\left[1 - \frac{\tanh k_0R}{k_0R}\right] \equiv -ka_s\end{aligned} \tag{8.69}$$

The (negative of the) constant of proportionality between δ_0 and k at low energies is the so-called *scattering length*, a_s. This only depends on the properties of the potential and for a repulsive square well it is given by the expression

$$a_s = R\left[1 - \frac{\tanh k_0R}{k_0R}\right] \tag{8.70}$$

Because $\tanh x < x$ regardless of x, the scattering length of a repulsive potential well is always positive, $a_s > 0$.

How can we relate these low-momentum properties to the cross section? The $l = 0$ component of the total cross section is, according to Eq. (8.55),

$$\sigma_0 = \frac{4\pi}{k^2} \sin^2 \delta_0 \,. \tag{8.71}$$

The cross section is shown as a function of the momentum in Fig. 8.9. In the low energy region, one finds that

$$\sigma_0 \xrightarrow[k\to 0]{} \frac{4\pi}{k^2} \sin^2(-ka_s) \approx 4\pi a_s^2 \tag{8.72}$$

This corresponds to the cross section of a sphere of effective radius a_s, shown as a horizontal dotted line in the figure. This indicates that, at low energies, the scattering wave "sees" only a portion of the potential characterised by the scattering length a_s, which is in general smaller than the radius, $a_s < R$.

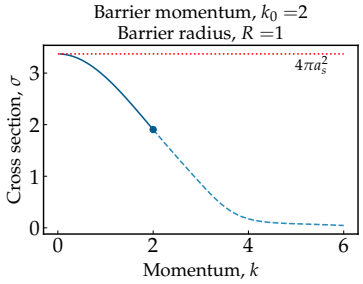

Figure 8.9: The $S-$wave scattering cross section for a repulsive square well with $R = 1$ and $k_0 = 2$. The solid and dashed lines and symbols have the same meanings as in Fig. 8.8. The dashed line corresponds to the low-energy limit $\sigma \to 4\pi a_s^2$.

8.4.2 *Repulsive square well: scattering above the barrier*

The phase-shift solution of Eq. (8.69) is only valid in the case of low scattering energies, $k < k_0$. For $k > k_0$, the solution is mathematically different. The momentum $\bar{k} = \sqrt{k^2 - k_0^2}$ in Eq. (8.63) is not imaginary anymore, but a real number. As a consequence, the solution to the first equation (in the interior region) is different and becomes a $\sin(x)$ instead of a $\sinh(x)$:

$$\bar{k}^2 = k^2 - k_0^2 \tag{8.73}$$

$$\frac{d^2u_0}{dr^2} + \bar{k}^2 u_0 = 0\,, r \le R \Rightarrow u_0(r) = A \sin \bar{k}r \tag{8.74}$$

$$\frac{d^2u_0}{dr^2} + k^2 u_0 = 0\,, r > R \Rightarrow u_0(r) = B \sin\left(kr + \delta_0\right). \tag{8.75}$$

We can still find the phase-shift δ_0 by imposing the continuity of logarithmic derivatives, Eq. (8.67). The matching procedure of the interior and exterior wavefunctions is illustrated in Fig. 8.10.

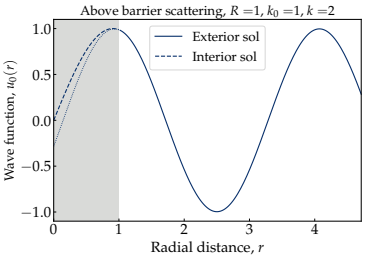

Figure 8.10: For energies above the barrier, the scattering wavefunctions in the interior region, $r < R$ (grey area), and the exterior region, $r > R$, must match at $r = R$. For this to happen, the phaseshift must be given by Eq. (8.77). The dotted line in the interior region represents the continuation of the exterior wavefunction into the interior region.

In this case, the matching of logarithmic derivatives at $r = R$ leads to the following equation,

$$\frac{\bar{k}}{\tan \bar{k}R} = \frac{k}{\tan\left(kR + \delta_0\right)}. \tag{8.76}$$

Isolating the phaseshift, we find:

$$\delta_0 = \arctan\left[\frac{k}{\sqrt{k^2 - k_0^2}} \tan\left(\sqrt{k^2 - k_0^2}R\right)\right] - kR\,, \tag{8.77}$$

This corresponds to the dashed line in Fig. 8.8, which shows that the phase-shift typically peaks at a given value, and progressively becomes smaller in size. Importantly, this expression matches perfectly well with Eq. (8.69) in the $k \to k_0$ limit even though the arguments of the arctan function diverge in the two expressions.

This is not the only divergence "hidden" in Eq. (8.77). Whenever the argument of the tan function in the first term becomes a multiple of $\pi/2$, the function itself will diverge. This happens whenever the condition

$$\sqrt{k_n^2 - k_0^2}R = \frac{\pi}{2}n \Rightarrow k_n = \sqrt{k_0^2 + \left(\frac{\pi}{2}\frac{n}{R}\right)^2} \tag{8.78}$$

is met. The divergence is not apparent in the phase shift because the arctan function itself does not diverge for infinite arguments. The position of k_1 is marked with a filled square in Fig. 8.8. It is clear that this coincides with a minimum of the phase-shift. The condition for k_n corresponds also to the points in energy where the real part of the logarithmic derivative in the inner region, Eq. (8.76), vanishes. This is in general a *resonance* condition, which is known to give rise to "bumps" in the cross section. As the energy (or k) increases, the phase shift decreases to zero - indicating that the scattering becomes less and less likely. The decay is not smooth, though, and every time $k \approx k_n$, structures appear in the phase shift, as shown in Fig. 8.8.

See the tutorial problem looking at this in the Week 8 sheet.

8.4.3 *Attractive square well*

We now consider a similar, yet physically different, problem. This is now an attractive square well,

$$V(r) = \begin{cases} -V_0, \, r \leq R, \\ 0, \, r > R\,, \end{cases} \tag{8.79}$$

as shown in Fig. 8.11. Mathematically, the problem is very similar to the cases above. The differential equations in the two regions of space are now:

$$\frac{d^2u_0}{dr^2} + \bar{\bar{k}}^2 u_0 = 0\,, r \leq R \Rightarrow u_0(r) = A\sin\bar{\bar{k}}r \tag{8.80}$$

$$\frac{d^2u_0}{dr^2} + k^2 u_0 = 0\,, r > R \Rightarrow u_0(r) = B\sin\left(kr + \delta_0\right). \tag{8.81}$$

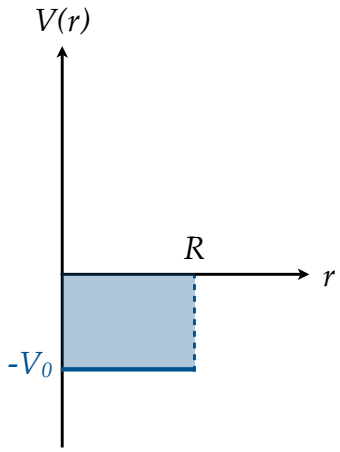

Figure 8.11: The attractive square well is a negative constant potential, V_0, between $r = 0$ and $r = R$.

where now the relevant momentum is real

$$\bar{\bar{k}} = \sqrt{k^2 + k_0^2}\,. \tag{8.82}$$

The phase-shift δ_0 is found again by matching logarithmic derivatives, Eq. (8.67), so that

$$\frac{\bar{\bar{k}}}{\tan\bar{\bar{k}}R} = \frac{k}{\tan\left(kR + \delta_0\right)}\,,$$

which yields

$$\delta_0 = \arctan\left[\frac{k}{\sqrt{k_0^2 + k^2}}\tan\left(\sqrt{k_0^2 + k^2}R\right)\right] - kR\,. \tag{8.83}$$

The result for the phase shift as a function of k equation is shown in Fig. 8.12 for the case $k_0 = 0.5$. Clearly, at low energies the phase shift

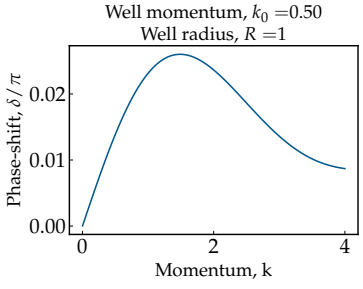

Figure 8.12: The $S-$wave scattering phase shift from an attractive well with $R = 1$ and $k_0 = 0.5$.

is linear in k. We can find the constant of proportionality by looking at Eq. (8.83) at low momentum,

$$\delta_0 \approx \frac{k}{k_0}\tan(k_0 R) - kR = -k\underbrace{\left[R - \frac{\tan k_0 R}{k_0}\right]}_{=a_s}, \tag{8.84}$$

which indicates that the scattering length in this case is

$$a_s = R\left[1 - \frac{\tan k_0 R}{k_0 R}\right]. \tag{8.85}$$

Because $\tan x > x$, the scattering length of this potential is *negative*, in contrast to the repulsive case. The phase shift thus grows linearly with k at low energies.

The procedure we have just describe only works for sufficiently shallow wells or, in other words, for small values of k_0. Imagine a situation in which we increase k_0 slowly from zero - thus progressively increasing the depth of the well. Eq. (8.85) indicates that the scattering length should become increasingly large and, in fact, when $k_0 R = \pi/2$, it is (in principle) infinite!

The scattering phase shift for the cases $k_0 = \frac{\pi}{2R} - 0.1$ and $k_0 = \frac{\pi}{2R} + 0.1$ are shown in the top and bottom panels of Fig. 8.13, respectively. A markedly different behaviour is observed. When the well has a depth that is slightly more repulsive than the critical value $V_b = \frac{\hbar^2\pi^2}{4R^2}$, the zero energy phase shift is zero and increases quickly with k. It saturates quickly at a value close to $\pi/2$, and then decays as a function of energy. The maximum indicates again the presence of a resonance at the point where $\sqrt{k_0^2 + k^2}R = \pi/2$. In contrast, when V_0 is just a bit more attractive than V_b, the phase shift at zero energy is π, and as the energy increases δ_0 decreases steeply with energy.

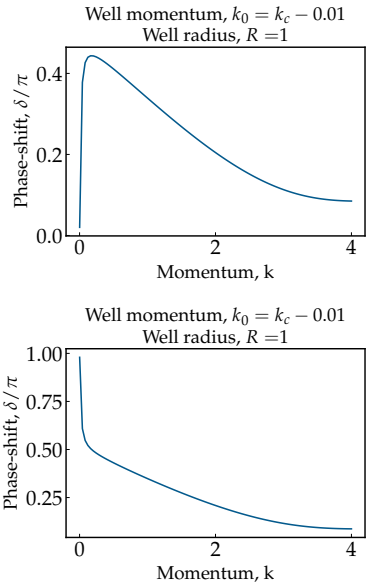

Figure 8.13: The $S-$wave scattering phase shift from an attractive well with $R = 1$ and $k_0 = \frac{\pi}{2R} - 0.01$ (top panel) and $k_0 = \frac{\pi}{2R} + 0.01$ (bottom panel).

These markedly different behaviours are due to the presence of bound states in the square well. One can prove (see for instance Refs. [3]) that a spherical attractive square well has one single bound state in the $S-$wave when the condition $|V_b| < |V_0| < 4|V_b|$ is met. We thus see that the scattering phase shift captures information of the bound states of the beam+target system. In fact, a general result due to Levinson stipulates that for a potential with N bound states, the scattering phase shift at zero energies is $N\pi$. This is remarkable because we are seeing the properties of the bound state (a negative energy state!) from the scattering properties (which occur at positive, real energies!).

[3] C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum Mechanics*, volume 1. Wiley-VCH, 1977; and V. Zelevinsky. *Quantum Physics*, volume 1. Wiley-VCH, 2011

8.4.4 *Hard spheres*

Hard spheres are a particularly useful example of the scattering problem. These correspond to the repulsive square well, where the barrier height is extremely large, $V_0 \to \infty$. The hard sphere model is useful because it allows us to look at what happens for the scattering of partial waves with $l > 0$. You can work with these in the Jupyter

Notebook Scattering_hard_spheres.ipynb available at the Master's github repository shorturl.at/fEXZ5.

The infinitely repulsive well forces the wavefunctions to vanish at $r = R$, so that $R_l(r < R) = 0$. At $r > R$, the Schrödinger equation is nothing but the spherical Bessel equation, Eq. (**??**). We know the solutions to this equation must be linear combinations of the independent functions $j_l(kr)$ and $n_l(kr)$ with at most a phase difference, so

$$R_l(k,r) = A\left[j_l(kr)\cos\delta_l(k) + n_l(kr)\sin\delta_l(k)\right], \tag{8.86}$$

The boundary condition imposed by the hard sphere at $r = R$ is such that

$$R_l(k, r = R) = 0 \Rightarrow j_l(kR)\cos\delta_l = -n_l(kR)\sin\delta_l, \tag{8.87}$$

and so

$$\delta_l = \arctan\left(-\frac{j_l(kR)}{n_l(kR)}\right). \tag{8.88}$$

At $l = 0$, using Eqs. (8.31) and (8.32), one finds $\delta_0 = -kR$ exactly.

The hard sphere wavefunctions (multiplied by r) are shown as the solid lines in Fig. 8.14, for the case $R = 1$ and $k = 1$. The three panels correspond to the partial waves $l = 0$ (top), 1 (central) and 2 (bottom panel). As expected, these wavefunctions vanish at $r = R$. The corresponding non-interacting wavefunction is simply a spherical Bessel function, $R_l = j_l$ (eg see Eq. (8.86) for $\delta_l = 0$). One clearly finds that the $l = 0$ wavefunction is a sinusoidal wave, displaced from the origin by a constant value δ_0. For hard spheres, and using Eq. (8.66), we indeed see that the scattering wave function behaves like

$$u_0(r) \approx B\sin(kr - ka_s). \tag{8.89}$$

At low energies, we therefore expect the wavefunction to have a node at $r = a_s$. For the $l = 1$ wave (central panel), the displacement is more subtle and is more visible at lower energies. When we get to $l = 2$ (bottom panel), the modifications due to the interaction are negligible. This indicates that the expansion in terms of partial waves may be terminated at relatively low values when the incident momentum, k, is low.

For $l \neq 0$, we can extract useful information in the limit of very low energies, $kR \ll 1$. Using the expansions in Eq. (8.33) and (8.34), one finds that the low energy phaseshifts become

$$\delta_l \approx \arctan\left(-\frac{(kR)^{2l+1}}{(2l+1)!!(2l-1)!!}\right) \approx -\frac{(kR)^{2l+1}}{(2l+1)!!(2l-1)!!}. \tag{8.90}$$

For a fixed value of k and R, the phase-shift becomes smaller as l increases. This behaviour is illustrated in the top panel of Fig. 8.15. Whereas the $l = 0$ phase-shift is active for all momenta, the contribution of the $l = 1$ phase-shift only becomes finite beyond

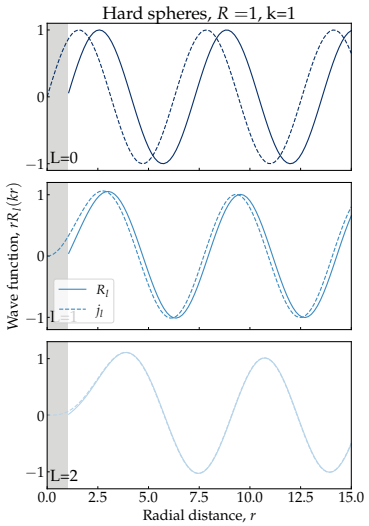

Figure 8.14: The wave functions $rR_l(kr)$ as a function of radial distance for hard-sphere scattering with radius $R = 1$ and momentum $k = 1$. The top, central and bottom panels correspond to $l = 0$, $l = 1$ and $l = 2$, respectively. The dashed line is the non-interacting wavefunction.

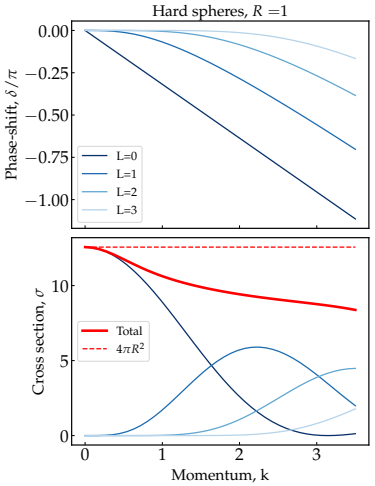

Figure 8.15: Top panel: phase-shifts δ_l for hard spheres with $R = 1$ as a function of momentum for partial waves from $l = 0$ to $l = 3$. At low energies, only the phase-shift with lower l matter. Bottom panel: cross section σ, and contributions σ_l, as a function of momentum. The low-momentum value $\sigma = 4\pi R^2$ is shown with a horizontal dashed line. wavefunction.

$k \approx R^{-1}$, as expected from the previous equation. Higher l modes are pushed away from the low energy region by the centrifugal barrier, and they only contribute at successively higher values of momentum k.

We can find the scattering length of hard spheres by taking the limit $V_0 \to \infty$ in Eq. (8.70). Interestingly, this quantity remains finite in the limit of a very large (infinite!) repulsive wall. Indeed, for $k_0 R \gg 1$, one has $\tanh x \to 1$ so that

$$a_s \approx R \left[1 - \frac{1}{k_0 R} \right] \to R \,. \tag{8.91}$$

and the scattering length is the radius of the barrier $a_s = R$.

In this case, for $l = 0$, the cross section corresponds exactly to the surface of a hard sphere, $\sigma = 4\pi R^2$. We show this with a dashed line in the bottom panel of Fig. 8.15. The remaining contributions to the wavefunctions, given by $\sigma_l = \frac{4\pi}{k^2}(2l+1)\sin^2\delta_l$, are also shown in the figure for completeness. Clearly, the $l = 0$ wave is the only one that contributes at low energies. As k increases, this wave decreases in value and the $l = 1$ dominates for momenta $k \approx 2$. After that, it is the $l = 2$ (and, further on, the $l = 3$ wave) that become relevant for the total cross section. This indicates that the partial wave expansion requires more and more terms as the momentum increases.

8.5 *Scattering length*

We have just seen in the hard spheres example that only the $l = 0$ waves contribute to scattering at low energies. As we discussed earlier in the context of Eq. (8.53), in this situation one can write the full scattering amplitude as

$$f = \frac{1}{k \cot \delta_0 - ik} \,. \tag{8.92}$$

Importantly, δ_0 remains finite even in the case of a very strong interaction, even in the case of a hard sphere with a formally infinite repulsive potential.

Let us summarize what we have learned about the scattering properties at low energies. We have discussed the $S-$wave phase shift at low energies and we have characterised them in terms of the scattering length, a_s, through the relation $\delta_0 \approx -k a_s$. We find that for a *repulsive* square well of range R, the scattering length is *positive* and fulfils $0 < a_s < R$ [see Eq. (8.70)]. In contrast, Eq. (8.85) indicates that for an *attractive* potential well holding no bound states, the scattering length is *negative*. If the square well is such that it holds a bound state, the scattering length is instead large and positive, as seen in the bottom panel of Fig. 8.13.

The fact that δ_0 behaves linearly with k at low energies is a generic result of scattering theory. The notions relating the sign of the scattering length to the attractive or repulsive nature of the potential also hold independently of the shape of $V(r)$. For generic potentials,

one defines the scattering length as the following limit:

$$a_s = -\lim_{k\to 0} \frac{\tan\delta_0}{k}\,, \tag{8.93}$$

so that $\delta_0 = -ka_s$ in the $k \to 0$ limit. With this, the scattering amplitude of Eq. (8.92) becomes

$$\bar{f} = \frac{1}{-\frac{1}{a_s} - ik}\,. \tag{8.94}$$

One can show that repulsive potentials lead to positive scattering lengths, whereas attractive potentials with no bound states lead to negative values of a_s.

The generic treatment of the low-momentum scattering properties of potentials can be extended beyond the lowest order in k. By considering an expansion of f as a function of momentum to the second order in k, one then finds that any scattering amplitude at sufficiently low energies is given by the general expression

$$\bar{f} = \frac{1}{-\frac{1}{a_s} + \frac{1}{2}k^2 r_{\text{eff}} - ik}\,. \tag{8.95}$$

The parameter r_{eff} is the so-called *effective range* and it is, just as a_s, a generic property of the interaction.

-