

Introduction to Quantum and Statistical Physics

Tanguy Marsault

May 29, 2024

Preface

"I think I can safely say that nobody understands quantum mechanics."

-Richard Feynman

These words often evoke fear in the minds of uninitiated readers. They should not constitute a psychological barrier to the study of quantum theory, but rather fuel curiosity and give newcomers to the subject the desire to uncover the secrets of this still poorly understood theory. Although this quote is now old, it remains unfortunately (or fortunately!) no less true. Quantum theory, which has proven itself in many fundamental experiments for the history of science, is far from a closed subject. The problems that persist in it probably constitute a gateway to a more fundamental understanding of the universe, but also to the development of ever more advanced technologies.

With this idea in mind, I decided to write this introduction to quantum mechanics. Like any good science author, I am obliged, out of pedantry or honest modesty, to include the word "Introduction" in the title of the text. And then I understand that it seems foolish to spend a few hours of such precious time reading fifty pages, just to understand the "introduction" of a subject. But let me at least try to prove you wrong. The word "introduction" here does not mean that the subject matter is simple, but rather that it only deals with the foundations of quantum theory. It omits, and sometimes admits ignorance of, all the underlying theories that follow from it. However, having read and understood these 50 pages already constitutes a solid foundation for understanding the basics of quantum technologies, but also, for readers interested in theoretical physics, to delve into more fundamental theories such as quantum field theory.

So this note does not pretend to lead anywhere or to bring revolutionary ideas. All the material presented here can be found in other works of literature. They are not paraphrases, however, as I have always adapted the text by emphasizing the points that are generally poorly understood (and that I may have misunderstood during my own learning of the subject). Thus, there is no particular reason why the format will suit you, as it was written to suit *me*. Moreover, I hope that it will be useful to you, and that you will be able to find explanations and interpretations that you can hold onto.

This note was originally written for first-year students at CentraleSupélec and therefore focuses on the concepts they encounter in their study of quantum mechanics. It presents, as clearly and precisely as I have managed to write, all the basic concepts of quantum mechanics, giving proof for most of them and detailing the calculations as much as possible. For each subject addressed, many properties are explained mathematically and physically so that the reader can best understand the phenomenon in question. Most of the demonstrations represent more a way of manipulating concepts than a real mathematical-physical interest. Thus, they should be taken as application exercises rather than demonstrations that one would skip due to intellectual laziness. In addition, many small exercises that seem important to me are scattered throughout the text. They will help the reader in understanding the subject matter, but also in their mathematical understanding of the material, which at first glance can be intimidating.

The text is organized into two parts: quantum physics and statistical physics. Only the last section concerns the latter field. The proportion of space used for statistical physics does not reflect its importance. Statistical physics is one of the most fruitful theories for

studying macroscopic phenomena. If I am addressing students from CentraleSupélec more specifically, it is not to be overlooked, and it should be kept in mind that its applications are generally much simpler than those of quantum mechanics. This note starts from simple mathematical concepts that allow for the introduction of formal notations. Then, the postulates of quantum mechanics are exposed, and then the reader gets into the heart of the matter by studying many different Hamiltonians (harmonic oscillator, potential well, hydrogen atom, ...). The amount of information presented may seem important, but it is more a question of understanding than learning. Good understanding will allow most of the presented problems to be redone with ease.

Finally, it seems important to me to thank Constance de Lestrangé and Thomas Soupizet, who kindly lent me the notes they took during a lecture I gave on the subject. This allowed me to identify the conceptually and computationally misunderstood points, which I was able to detail here. By the way, thanks also to Alexandre Sajus whose numerous questions allowed me to orient my speech in the right direction. Thanks also to the Club Scientifique of CentraleSupélec for the feedback it gave me to develop this note. Special mention to Tom de Coninck, whose proofreading was so precise that one could speak of co-writing. And thank you in advance to all those who will contribute to the improvement of this work.

In case of noticed errors or misunderstandings, do not hesitate to contact me by email at tanguy.marsault02@gmail.com.

Enjoy your reading (and good luck if you're not reading for pleasure), I hope to spark your curiosity for a few moments.

Contents

1	Mathematical formalism and Dirac's notations	1
2	Postulates of quantum mechanics	3
2.1	Postulates	3
2.2	Time evolution of a state	4
3	Operators	5
3.1	Commutators	5
3.2	Mean, variance, and standard deviation of an operator in a state	6
3.3	Continuous basis and wave functions	7
3.4	Position and momentum operators	10
4	Quantum wells	12
4.1	Schrödinger equation and regularity of the wave function	12
4.2	Infinite well	14
4.3	A two-dimensional case	15
5	Harmonic oscillator	17
5.1	Solution, eigenstates, eigenenergies	17
5.2	2D and 3D Harmonic Oscillators	20
5.3	Application exercises	21
6	Basics of perturbation theory	23
6.1	Principle	23
6.2	Application exercise	24
7	Angular momentum and spin	26
7.1	Angular momentum operators	26
7.2	Spin	27
7.3	Study of electron spin	28
7.4	Coupling to magnetic field	28
8	Hydrogenoid atom	30
8.1	Solution of the two-body problem	30
8.2	Some application exercises	31
8.3	More Electrons and the Variational Method	32
9	Statistical physics	34
9.1	Foundations	34
9.2	Microcanonical Ensemble	35
9.3	Canonical ensemble	37
9.4	Applications	41
9.4.1	Spin orientation	41
9.4.2	Magnetization	42
9.4.3	Properties of the monatomic ideal gas	43

1 Mathematical formalism and Dirac's notations

"To those who do not know mathematics it is difficult to get across a real feeling as to the beauty, the deepest beauty, of nature ... If you want to learn about nature, to appreciate nature, it is necessary to understand the language that she speaks in."

-Richard Feynman

In this first part, the mathematical formalism necessary for quantum physics is introduced. Some definitions about Hilbert spaces are recalled, and Dirac notations are formally introduced.

Definition 1.0.1 (Scalar product). Let E be a complex vector space equipped with $B : E \times E \rightarrow \mathbb{C}$ such that:

- B is linear on the right
- B is anti-linear on the left

$$\forall \lambda, \mu \in \mathbb{C}, \forall u, v, w \in E, B(\lambda u + \mu v, w) = \lambda^* B(u, w) + \mu^* B(v, w)$$

- B is Hermitian symmetric $B(u, v) = B(v, u)^*$

Then B is said to be sesquilinear on E . Furthermore,

- B is said to be positive if $B(u, u) \geq 0$
- B is said to be definite if $B(u, u) = 0 \Leftrightarrow u = 0$

A positive definite sesquilinear form on E is called a scalar or Hermitian product on E .

Definition 1.0.2. If moreover the vector space E is complete, then it is a Hilbert space. This is the type of space that will be used for all of quantum physics.

Notation 1.0.3 (Dirac). In quantum physics, quite particular notations are used to designate the vectors of the considered Hilbert space (see Postulate 1 of the following section). Let \mathcal{H} be the considered Hilbert space, which is associated with its natural scalar product denoted $\langle \cdot | \cdot \rangle$. For a vector $v \in \mathcal{H}$, it will now be denoted $|v\rangle$. In addition, its transconjugate, defined as the transpose of the vector with the conjugate components, will be denoted by $(v^t)^* = \langle v| = (|v\rangle)^\dagger$. This definition is clear in finite dimension, although it remains true in infinite dimension, it can be conceptually disturbing; we can then rather define $\langle u| : \mathcal{H} \rightarrow \mathbb{R}$ as the unique linear form such that $\langle u| (|v\rangle) \equiv \langle u|v\rangle$ is the scalar product of u by v . With these notations, it follows that for $|u\rangle, |v\rangle \in \mathcal{H}$, $\langle u|v\rangle = \langle u| \cdot |v\rangle$ (scalar product of u and v). For any linear operator $A : \mathcal{H} \rightarrow \mathcal{H}$, this operator will be denoted \hat{A} and its action on a ket $|v\rangle$, $\hat{A}|v\rangle$.

We then notice that if $|u\rangle, |v\rangle \in \mathcal{H}$, the quantity $|v\rangle \langle u|$ defines an operator $\mathcal{H} \rightarrow \mathcal{H}$ such that

$$(|v\rangle \langle u|) |w\rangle = |v\rangle \langle u|w\rangle = \langle u|w\rangle |v\rangle$$

We note that the major advantage of Dirac's notation is its associativity, which allows us to manipulate it in an intuitive way and in the order we prefer. Let us also note once and for all that:

$$|v\rangle \langle u| \neq \langle u|v\rangle$$

Indeed, the left-hand side is an operator and the right-hand side is a scalar!

Now that the notation is introduced, we can give some relations that will be useful to us later.

Proposition 1.0.4 (Closure). Let $|e_n\rangle_{n \in \mathbb{N}}$ be an orthonormal basis¹ of \mathcal{H} (we will study the generalization to a non-countable basis later), then we have

$$\hat{\mathbb{I}} = \sum_{n=0}^{\infty} |e_n\rangle \langle e_n| \quad (1.1)$$

where the identity operator is denoted by $\hat{\mathbb{I}}$ such that $\hat{\mathbb{I}}|v\rangle = |v\rangle$

Proof. First, note that $|v\rangle = \sum_{k=0}^{\infty} \langle e_k|v\rangle |e_k\rangle$, hence

$$\begin{aligned} \left(\sum_{n=0}^{\infty} |e_n\rangle \langle e_n| \right) |v\rangle &= \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \langle e_k|v\rangle |e_n\rangle \langle e_n|e_k\rangle \quad \text{since } \langle e_k|v\rangle \text{ is a scalar} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \langle e_k|v\rangle |e_n\rangle \delta_{n,k} \quad \text{because } |e_n\rangle \text{ are orthonormal} \\ &= \sum_{k=0}^{\infty} \langle e_k|v\rangle |e_k\rangle \\ &= |v\rangle \end{aligned}$$

□

Proposition 1.0.5. With the previous notations, an operator $\hat{A} : \mathcal{H} \rightarrow \mathcal{H}$ can be seen as a matrix A (of potentially infinite size) with coefficients given by

$$A_{ij} = \langle e_i | \hat{A} | e_j \rangle \quad (1.2)$$

called matrix elements of \hat{A} .

Proof.

$$\begin{aligned} \hat{A} &= \hat{\mathbb{I}} \hat{A} \hat{\mathbb{I}} \\ &= \sum_{i=0}^{\infty} |e_i\rangle \langle e_i| \hat{A} \sum_{j=0}^{\infty} |e_j\rangle \langle e_j| \\ &= \sum_{i,j} |e_i\rangle \langle e_i| \hat{A} |e_j\rangle \langle e_j| \\ &= \sum_{i,j} \langle e_i | \hat{A} | e_j \rangle |e_i\rangle \langle e_j| \end{aligned}$$

Which then proves that the family $\{|e_i\rangle \langle e_j|\}_{i,j \in \mathbb{N}}$ is a base for the operators. In this base, an operator \hat{A} should be regarded as a matrix with coefficients $\langle e_i | \hat{A} | e_j \rangle$ □

¹In mathematics we usually call it a *complete basis*.

2 Postulates of quantum mechanics

"I do not like it, and I am sorry I ever had anything to do with it"

-Erwin Schrödinger

In this section, the postulates that form the foundations of quantum mechanics are first introduced, followed by an extremely important application.

2.1 Postulates

Postulate 1. Every physical system S is described (at every instant) by a normalized ket $|\psi\rangle$ belonging to a Hilbert space \mathcal{H} . Furthermore, $|\psi\rangle$ is of norm 1. $|\psi\rangle$ is called the state of the system, and \mathcal{H} is the space of states.

Postulate 2. Any physical quantity \mathcal{A} or "observable" is represented by a Hermitian operator \hat{A} (i.e. $\hat{A}^\dagger = \hat{A}$).

Postulate 3. A measurement of the observable \hat{A} of a physical system can only give as the observed value an eigenvalue of \hat{A} . (Note that since \hat{A} is Hermitian, its eigenvalues are real, which we quickly prove below.)

Proof. Let X be an eigenvector of A (a Hermitian operator) associated with a value $\lambda \in \mathbb{C}$. Then we have $AX = \lambda X$, and taking the transconjugate, $X^\dagger A^\dagger = \lambda^* X^\dagger$. Using $A^\dagger = A$ and multiplying by X , we get $\lambda X^\dagger X = \lambda^* X^\dagger X$. Noting that $X^\dagger X = \|X\|^2 \neq 0$, we have $\lambda = \lambda^*$, which means that $\lambda \in \mathbb{R}$. \square

Postulate 4 (Born's rule). The system is in the state $|\psi\rangle$ and a measurement of the observable \hat{A} is performed. Let $\{a_n\}_{n \in \mathbb{N}}$ be the eigenvalues of \hat{A} associated with the orthonormal basis (which exists because \hat{A} is Hermitian and thus diagonalizable in an orthonormal basis, by the complex spectral theorem) $\{e_n^i\}_{n \in \mathbb{N}, 1 \leq i \leq g_n}$ (where g_n is the degeneracy order of a_n , the dimension of its eigenspace, the number of independent $|v\rangle$ associated with the eigenvalue a_n , also called multiplicity in linear algebra) such that $\hat{A}|e_n^i\rangle = a_n|e_n^i\rangle$ with $1 \leq i \leq g_n$. Then the probability of obtaining the measurement result a_n is

$$P(a_n) = \sum_{i=1}^{g_n} |\langle e_n^i | \psi \rangle|^2 \quad (2.1)$$

In particular, if a_n is non-degenerate,

$$P(a_n) = |\langle e_n | \psi \rangle|^2 \quad (2.2)$$

Postulate 5 (Wavefunction collapse). Using the notation of postulate 4, the measured value is a_n . We are then certain of the state $|\psi'\rangle$ immediately after the measurement:

$$|\psi'\rangle = \frac{1}{\sqrt{P(a_n)}} \sum_{i=1}^{g_n} \langle e_n^i | \psi \rangle |e_n^i\rangle \quad (2.3)$$

Once again, if the state is non-degenerate,

$$|\psi'\rangle = |e_n\rangle \quad (2.4)$$

Postulate 6 (Schrödinger equation). Every system possesses a particular physical quantity called energy. The operator associated with it is the Hamiltonian, denoted \hat{H} . The time evolution of $|\psi(t)\rangle$ is given by the Schrödinger equation:

$$i\hbar \frac{d}{dt} |\psi\rangle = \hat{H} |\psi\rangle \quad (2.5)$$

2.2 Time evolution of a state

Since \hat{H} is Hermitian and we assume that it does not depend on time, it has an orthonormal basis that diagonalizes it, which is denoted $\{|\phi_n\rangle\}_{n \in \mathbb{N}}$, which also does not depend on time, so that $\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle$, where E_n are the eigenenergies (eigenvalues of \hat{H}).

At time $t = 0$, the state is $|\psi(0)\rangle = \sum_{n=0}^{\infty} a_n(0) |\phi_n\rangle$. We seek the expression for $|\psi(t)\rangle$. At any time, it is possible to write:

$$|\psi(t)\rangle = \sum_{n=0}^{\infty} a_n(t) |\phi_n\rangle$$

Injecting this expression into the Schrödinger equation, we have:

$$\begin{aligned} i\hbar \frac{d}{dt} |\psi\rangle &= \hat{H} |\psi\rangle \\ i\hbar \sum_n \frac{da_n(t)}{dt} |\phi_n\rangle &= \sum_n a_n(t) E_n |\phi_n\rangle \end{aligned}$$

Since $|\phi_n\rangle$ is a basis, for all $n \in \mathbb{N}$:

$$i\hbar \frac{da_n(t)}{dt} = a_n(t) E_n \quad (2.6)$$

which has solution:

$$a_n(t) = a_n(0) e^{-\frac{iE_n t}{\hbar}} \quad (2.7)$$

Therefore, we have the important relation:

$$\boxed{|\psi(t)\rangle = \sum_{n=0}^{\infty} a_n(0) e^{-\frac{iE_n t}{\hbar}} |\phi_n\rangle} \quad (2.8)$$

3 Operators

"I was taught at school never to start a sentence without knowing the end of it."

-Paul Dirac

After stating the postulates, we obtained an important result, namely the way the kets evolve in time. However, to continue our study of quantum physics, we need to examine operators more deeply.

3.1 Commutators

Throughout this section, any quantity not explicitly defined and carrying a hat will be an operator.

In quantum physics, the postulates indicate that measuring a physical observable \hat{A} amounts to transforming the state of our physical system $|\psi\rangle$ into the state ² $|\psi'\rangle = \hat{A}|\psi\rangle$. Thus, successive measurement of physical observables \hat{A} and \hat{B} should not yield the same result depending on whether we apply \hat{A} or \hat{B} first. In fact, for this to be the case, it would generally be necessary for $\hat{A}\hat{B} = \hat{B}\hat{A}$. That is why we define one of the most important objects in quantum physics:

Definition 3.1.1 (Commutator). Let \hat{A} and \hat{B} be two operators $\mathcal{H} \rightarrow \mathcal{H}$. Their commutator is defined by

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \quad (3.1)$$

The following lemma is one of the most important lemmas for most of the proofs that we will carry.

Lemma 3.1.2. If $[\hat{A}, \hat{B}] = 0$, then \hat{A} and \hat{B} have a common basis of eigenvectors.

Proof. This follows from mathematical codiagonalization. □

In quantum mechanics, all physical observables are necessarily represented by Hermitian operators, which then have an orthonormal basis of eigenvectors. We even know that if several of these observables commute, then they have a common basis of eigenvectors. The legitimate question is then, for a physical observable, can we find eigenvectors whose eigenvalue determines our physical state uniquely? The answer is obviously *no* as soon as this observable is degenerate. And that is what leads us to the following definition.

Definition 3.1.3 (Complete Set of Commuting Observables). A Complete Set of Commuting Observables (CSCO) is called a set of observables $E = \{\hat{A}, \hat{B}, \dots, \hat{Z}\}$ whose commutators are all zero (the observables commute pairwise) and whose n -tuple (a, b, \dots, z) of respective eigenvalues of each operator defines a unique quantum state that will then be denoted by $|a, b, \dots, z\rangle$ and such that

$$\begin{cases} \hat{A}|a, b, \dots, z\rangle = a|a, b, \dots, z\rangle \\ \hat{B}|a, b, \dots, z\rangle = b|a, b, \dots, z\rangle \\ \vdots \\ \hat{Z}|a, b, \dots, z\rangle = z|a, b, \dots, z\rangle \end{cases}$$

²What we say here is actually not quite true. We'd better say that the information concerning the state after the measure of the observable \mathcal{A} is contained in the ket $|\psi'\rangle$

The definition as it stands is rather mathematical and somewhat obscure, but the idea is quite simple. We wish to find enough observables that commute, meaning their values can be determined at the same time, to determine the state of our system uniquely. We will see many examples of such observables, called E.C.O.C., later on.

3.2 Mean, variance, and standard deviation of an operator in a state

Consider a physical system in a state $|\psi\rangle$ and a physical observable \hat{A} . We want to determine the average value of \hat{A} in this state.

Statistically, we define the mean of \hat{A} as the sum of the values that \hat{A} can take times their probability (which depends on the state $|\psi\rangle$, as shown by the Born rule). The postulates tell us that \hat{A} can only take its eigenvalues a_n , associated with eigenvectors $|\phi_n\rangle$ forming a basis of \mathcal{H} , so we have:

$$\begin{aligned}
 \langle A \rangle_\psi &= \sum_n p_n a_n \\
 &= \sum_n |\langle \psi | \phi_n \rangle|^2 a_n \\
 &= \sum_n (\langle \psi | \phi_n \rangle)^* \langle \psi | \phi_n \rangle a_n \\
 &= \sum_n \langle \phi_n | \psi \rangle \langle \psi | \phi_n \rangle a_n \\
 &= \sum_n \langle \psi | a_n | \phi_n \rangle \langle \phi_n | \psi \rangle \quad \text{since } a_n \text{ is a scalar and thus commutes with bras and kets} \\
 &= \sum_n \langle \psi | \hat{A} | \phi_n \rangle \langle \phi_n | \psi \rangle \quad \text{since } \hat{A} | \phi_n \rangle = a_n | \phi_n \rangle \\
 &= \langle \psi | \hat{A} \left(\sum_n | \phi_n \rangle \langle \phi_n | \right) | \psi \rangle \quad \text{using the closure formula} \\
 \langle A \rangle_\psi &= \langle \psi | \hat{A} | \psi \rangle
 \end{aligned}$$

This leads naturally to the definition:

Definition 3.2.1 (Mean). The mean value of \hat{A} in the state $|\psi\rangle$ is defined as

$$\langle A \rangle_\psi = \langle \psi | \hat{A} | \psi \rangle \quad (3.2)$$

In the same way as in statistics, we can define the standard deviation:

Definition 3.2.2 (Standard deviation). We define the squared standard deviation (variance) of \hat{A} in the state $|\psi\rangle$ as

$$\sigma_\psi^2(\hat{A}) = \langle A^2 \rangle_\psi - \langle A \rangle_\psi^2 = \langle \psi | \hat{A}^2 | \psi \rangle - \left(\langle \psi | \hat{A} | \psi \rangle \right)^2 \quad (3.3)$$

This leads to a well-known and fundamental property that will be studied again for two particular operators later on:

Proposition 3.2.3 (Heisenberg's inequality). Let \hat{A}, \hat{B} be two operators and $|\psi\rangle \in \mathcal{H}$, then we have

$$\sigma_\psi(\hat{A}) \sigma_\psi(\hat{B}) \geq \frac{1}{2} \left| \langle [\hat{A}, \hat{B}] \rangle_\psi \right| \quad (3.4)$$

Proof. The proof is actually quite easy and is based solely on the study of a 2nd order polynomial. It can be found almost everywhere but will not be detailed here because of the little physical interest it represents. \square

Another property can be determined quite easily, however, it is rather to be taken as an exercise.

Proposition 3.2.4 (Ehrenfest's theorem).

$$\frac{d}{dt} \langle \hat{A} \rangle_\psi = \left\langle \frac{\partial \hat{A}}{\partial t} \right\rangle_\psi + \frac{1}{i\hbar} \langle [\hat{A}, \hat{H}] \rangle_\psi \quad (3.5)$$

Proof. It suffices to write $\langle A \rangle_\psi = \langle \psi | \hat{A} | \psi \rangle$, to differentiate (Dirac's notations can be differentiated as a product or a multilinear form) and to use the fact that $i\hbar \frac{d}{dt} |\psi\rangle = \hat{H} |\psi\rangle \implies -i\hbar \frac{d}{dt} \langle \psi| = \langle \psi| \hat{H}^\dagger = \langle \psi| \hat{H}$. This is a good exercise to practice with the manipulation of mean values and the Schrödinger equation in bra form. \square

These last two properties allow us to deduce a fundamental inequality for quantum physics and more generally for high energy physics. Indeed, let us consider an observable \hat{A} , not explicitly depending on time, Ehrenfest's theorem then becomes

$$\frac{d}{dt} \langle \hat{A} \rangle_\psi = \frac{1}{i\hbar} \langle [\hat{A}, \hat{H}] \rangle_\psi \quad (3.6)$$

We assume that the average of the observable \hat{A} varies over a characteristic time Δt , that is

$$\frac{d}{dt} \langle \hat{A} \rangle_\psi \simeq \frac{\sigma_\psi(\hat{A})}{\Delta t}$$

Using the Heisenberg inequality in the Ehrenfest theorem, we obtain

$$\left| \frac{\sigma_\psi(\hat{A})}{\Delta t} \right| \simeq \left| \langle [\hat{A}, \hat{H}] \rangle_\psi \right| \leq \frac{2}{\hbar} \sigma_\psi(\hat{A}) \sigma_\psi(\hat{H})$$

which is often (incorrectly) rewritten as

$$\Delta E \Delta t \geq \frac{\hbar}{2} \quad (3.7)$$

In this last equation, the two Δ symbols do not have the same value but express that the energy cannot deviate from its mean value ($\sigma_\psi(\hat{H})$) for a short duration Δt which depends on the process studied represented by \hat{A} . Since time t is not a physical observable in quantum physics, it is not possible to deduce a real Heisenberg inequality as previously stated.

3.3 Continuous basis and wave functions

Let us consider a typical experiment such as Young's double-slit or diffraction. We wish to measure the position of an electron at the output of the optical device considered. We then take space and divide it as shown in the Figure 1.

When the position of an electron is measured in a small interval i , we note that it is in the state $|X_i\rangle$ and we have $\langle X_i | X_j \rangle = \delta_{i,j}$ (the probability of being measured as i and j at the same time is zero for $i \neq j$). We can reduce the size of these intervals more and more until they become infinitely precise and obtain a basis $|x\rangle$ for $x \in \mathbb{R}$. However, the properties of this basis are no longer the same.

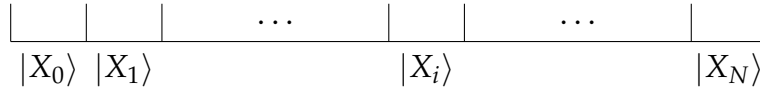


Figure 1: Measure of the position of the electron

Definition 3.3.1 (Basis $|x\rangle$). We define the basis $|x\rangle$ for measuring the position of a particle as an *orthogonal* basis satisfying $\langle y|x\rangle = \delta(y - x)$. δ is the Dirac distribution, viewed by physicists as a function that is zero everywhere except at 0 where it is infinite and whose integral is equal to 1. We note that this basis is no longer normalizable because $\langle x|x\rangle = \infty$! Thus, this basis generates a larger space than the Hilbert space of physical states.

If its properties are surprising, we will understand the usefulness of this basis for properly defining the wave function. Naturally, since the basis is continuous, we have the relation $\sum_n \rightarrow \int dx$. Let us take a physical state $|\psi\rangle \in \mathcal{H}$. It is then possible to write it in the form

$$|\psi\rangle = \int dx \psi(x) |x\rangle \quad (3.8)$$

where we have simply decomposed $|\psi\rangle$ on the basis $|x\rangle$ and where the function $\psi : \mathbb{R} \rightarrow \mathbb{C}$ just represents the coefficients of this decomposition. It is these coefficients that will form the classical and well-known wave function. Note that we must always have $\langle \psi|\psi\rangle = \int dx |\psi(x)|^2 = 1$.

Definition 3.3.2 (Wave function). For any ket $|\psi\rangle \in \mathcal{H}$, we can define a function $\psi : \mathbb{R} \rightarrow \mathbb{C}$ called the wave function, which corresponds to the representation on the basis $|x\rangle$ of $|\psi\rangle$. We then have in particular:

$$\psi(x) = \langle x|\psi\rangle \quad (3.9)$$

Proposition 3.3.3 (Inner product). Assuming the wave functions are in $L^2(\mathbb{C})$, for $|\psi\rangle, |\phi\rangle \in \mathcal{H}$, we have

$$\langle \phi|\psi\rangle = (\phi, \psi)_{L^2(\mathbb{C})} = \int dx \phi(x)^* \psi(x) \quad (3.10)$$

Proof.

$$\begin{aligned} \langle \phi|\psi\rangle &= \left(\int dy \phi(y)^* \langle y| \right) \left(\int dx \psi(x) |x\rangle \right) \\ &= \iint dx dy \phi(y)^* \psi(x) \langle y|x\rangle \\ &= \iint dx dy \phi(y)^* \psi(x) \delta(y - x) \\ &= \int dx \phi(x)^* \psi(x) \langle \phi|\psi\rangle \\ &= (\phi, \psi)_{L^2(\mathbb{C})} \end{aligned}$$

□

Definition 3.3.4 (Probability density). To relate with postulates and probability density,

$$\rho(x) = |\psi(x)|^2 = |\langle x|\psi\rangle|^2 \quad (3.11)$$

defines a valid probability density.

Proof. ρ is obviously positive, and we have

$$\begin{aligned}\int dx \rho(x) &= \int dx \langle x | \psi \rangle^* \langle x | \psi \rangle \\ &= \int dx \psi(x)^* \psi(x) \\ &= (\psi, \psi)_{L^2(\mathbb{C})} \\ &= \langle \psi | \psi \rangle \\ \int dx \rho(x) &= 1\end{aligned}$$

□

One last property needs to be generalized from countable basis to continuous basis:

Proposition 3.3.5 (Closure).

$$\hat{\mathbb{I}} = \int dx |x\rangle \langle x| \quad (3.12)$$

So far we have worked by seeking to know very precisely the position of a particle. We could also seek to know very precisely its momentum.

Definition 3.3.6. We define the $|p\rangle$ basis as

$$|p\rangle = \int dx \frac{e^{\frac{ipx}{\hbar}}}{\sqrt{2\pi\hbar}} |x\rangle \quad (3.13)$$

The normalization chosen with $\frac{1}{\sqrt{2\pi\hbar}}$ is not the only one possible, and in the literature (especially in QFT), it is possible to find others that are generally chosen to have a normalization different from the relation (demonstrated in the following property) $\langle p' | p \rangle = \delta(p - p')$.

Proposition 3.3.7. We have the relation

$$\langle x | p \rangle = \frac{e^{\frac{ipx}{\hbar}}}{\sqrt{2\pi\hbar}} \quad (3.14)$$

which then gives us the normalization of the kets $|p\rangle$

$$\langle p' | p \rangle = \delta(p - p') \quad (3.15)$$

and for any $|\psi\rangle \in \mathcal{H}$ we can define a wave function in the p representation by

$$\psi(p) = \langle p | \psi \rangle \quad (3.16)$$

and we have

$$|\psi\rangle = \int dp \psi(p) |p\rangle \quad (3.17)$$

Proof. The only property whose demonstration is interesting is $\langle p' | p \rangle$. It is done very easily, we just have to remember that $\int dx e^{ikx} = 2\pi\delta(k)$. □

Proposition 3.3.8. We then have

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int dp \psi(p) e^{\frac{ipx}{\hbar}} \quad (3.18)$$

The variables x and p are conjugated by a Fourier transform of the wave function.

Proof.

$$\begin{aligned}
 |\psi\rangle &= \int dp \psi(p) |p\rangle \\
 &= \int dp \psi(p) \hat{\mathbb{I}} |p\rangle \\
 &= \int dp \psi(p) \left(\int dx |x\rangle \langle x| \right) |p\rangle \\
 &= \iint dp dx \psi(p) |x\rangle \langle x|p\rangle \\
 |\psi\rangle &= \int dx \left(\int dp \frac{e^{\frac{ipx}{\hbar}}}{\sqrt{2\pi\hbar}} \psi(p) \right) |x\rangle
 \end{aligned}$$

Hence the result. □

One last closure relation can then be written:

Proposition 3.3.9.

$$\hat{\mathbb{I}} = \int dp |p\rangle \langle p| \quad (3.19)$$

3.4 Position and momentum operators

Now that we have defined states that designate the position and momentum of a particle, we seek to find operators that allow us to measure them.

Definition 3.4.1 (Position operator). We define the position operator $\hat{x} : \mathcal{H} \rightarrow \mathcal{H}$ such that $\hat{x}|\psi\rangle = |\psi'\rangle$ with

$$\langle y|\psi'\rangle = y\psi(y) \quad (3.20)$$

The position operator has a simple multiplication effect on the wave function:

$$\psi(y) \xrightarrow{\hat{x}} y\psi(y)$$

We will often write $|\psi'\rangle = |x\psi\rangle$, which is an abuse of notation that should be used very carefully and explained very clearly if used.

Definition 3.4.2 (Momentum operator). We define the momentum operator $\hat{p} : \mathcal{H} \rightarrow \mathcal{H}$ such that $\hat{p}|\psi\rangle = |\psi'\rangle$ with

$$\langle y|\psi'\rangle = -i\hbar \frac{\partial \psi(y)}{\partial y} \quad (3.21)$$

The momentum operator has a differentiation effect on the wave function:

$$\psi(y) \xrightarrow{\hat{p}} -i\hbar \partial_y \psi(y)$$

We will also often write $|\psi'\rangle = -i\hbar |\partial_x \psi\rangle$, which is also an abuse of notation.

Proposition 3.4.3 (Commutator).

$$[\hat{x}, \hat{p}] = i\hbar \hat{\mathbb{I}} \quad (3.22)$$

Proof. Let us calculate this commutator for any $|\psi\rangle \in \mathcal{H}$:

$$\begin{aligned}
 [\hat{x}, \hat{p}] |\psi\rangle &= \hat{x}\hat{p} |\psi\rangle - \hat{p}\hat{x} |\psi\rangle \\
 &= -i\hbar \left(\hat{x} |\partial_x \psi\rangle - \partial_x |x\psi\rangle \right) \\
 &= -i\hbar (|x\partial_x \psi\rangle - |\partial_x(x\psi)\rangle) \\
 &= -i\hbar (|x\partial_x \psi\rangle - |x\partial_x \psi\rangle - |\psi\rangle) \\
 &= i\hbar |\psi\rangle
 \end{aligned}$$

Thus, we have

$$[\hat{x}, \hat{p}] = i\hbar \hat{\mathbb{I}}$$

□

Remark 3.4.4. In the previous proof, we used a ket $|\psi\rangle$ to arrive at the result, sometimes it is not necessary and we can use purely algebraic methods such as the formula:

$$[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B} \quad (3.23)$$

A direct consequence can be obtained from the Heisenberg inequality:

Proposition 3.4.5.

$$\sigma_\psi(\hat{x})\sigma_\psi(\hat{p}) \geq \frac{\hbar}{2} \quad (3.24)$$

Proof. Recall the Heisenberg inequality:

$$\sigma_\psi(\hat{A})\sigma_\psi(\hat{B}) \geq \frac{1}{2} \left| \langle [\hat{A}, \hat{B}] \rangle \right|$$

which we apply directly using

$$[\hat{x}, \hat{p}] = i\hbar \hat{\mathbb{I}}$$

and the result follows. □

Everything we have done in this section has been done in 1D, but it is fairly easy to generalize to 3D. The main result in 3D is then the following:

Proposition 3.4.6.

$$[\hat{x}_i, \hat{p}_j] = i\hbar \hat{\mathbb{I}} \delta_{i,j} \quad (3.25)$$

where the indices i, j refer to orthogonal directions (often x, y, z).

In particular, observables that don't *look* in the same direction commute.

4 Quantum wells

Now that all the mathematical arsenal has been developed and the postulates of quantum mechanics have been established, we can tackle concrete problems. As suggested by the Schrödinger equation, there is a particular operator that interests us, the Hamiltonian. Generally, the Hamiltonian represents energy, so it is relatively easy to obtain:

- We take the classical expression of the energy of the problem under consideration.
- We *put hats* on all the quantities of interest which become operators in quantum mechanics.

Here, we study a first case of Hamiltonian with a potential energy (referred to only as potential in quantum mechanics) \hat{V} . By applying our quantization process to the classical energy $E = \frac{p^2}{2m} + V$, the general Hamiltonian is written as:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \hat{V} \quad (4.1)$$

In a few moment, we will make many assumptions about \hat{V} .

4.1 Schrödinger equation and regularity of the wave function

Let's first consider a one-dimensional situation. The goal is to establish a link between the Schrödinger equation on kets and the differential equation on the wave function. We then assume that the potential \hat{V} depends only on the position x and we denote $\widehat{V(x)}$ as follows:

$$\widehat{V(x)} |\psi\rangle = |\psi'\rangle \quad \text{with} \quad \langle x | \psi' \rangle = V(x) \psi(x) \quad (4.2)$$

Let's consider a state $|\psi(t)\rangle$. The end of the section on postulates helps us understand that finding the eigenvectors for the Hamiltonian completely solves the problem. The reasoning at the end of II indicates that if $|\psi\rangle$ is an eigenstate for \hat{H} associated with the proper energy E , then

$$|\psi(t)\rangle = e^{-\frac{iEt}{\hbar}} |\psi(0)\rangle$$

Thus, only $|\psi(0)\rangle$ depends on space and will be denoted more easily by $|\psi\rangle$. By injecting this form into the Schrödinger equation, we obtain the so-called *stationary* equation.

$$\begin{aligned} \hat{H} |\psi\rangle &= \left(\frac{\hat{p}^2}{2m} + \hat{V} \right) |\psi\rangle \\ &= E |\psi\rangle \end{aligned}$$

Proof. We can detail this calculation by inserting a wave function of the form $\phi(x, t) = f(t)g(x)$ into the Schrödinger equation. We then obtain:

$$i\hbar g(x) \frac{\partial f}{\partial t} = f(t) H g(x)$$

where we have used the fact that the Hamiltonian does not depend on time to bring $f(t)$ in front of H . It follows that:

$$\frac{1}{f(t)} i\hbar \frac{\partial f}{\partial t} = E = \frac{1}{g(x)} H g(x)$$

where E is a constant that does not depend on x or t and represents the energy of the particle (since $Hg(x) = Eg(x)$). Then we have:

$$f(t) = f(0)e^{-\frac{iEt}{\hbar}} \quad (4.3)$$

And the equation for the spatial part is $Hg(x) = Eg(x)$ which, when converted back to vector form, becomes:

$$\hat{H}|\psi\rangle = E|\psi\rangle$$

where $|\psi\rangle$ is a vector that does not depend on time. \square

(In general, we can remember that if the Hamiltonian does not depend on time, we can look for the eigenfunctions $\phi(x, t)$ as $f(t)g(x)$ with f as a complex exponential.)

Since $|\psi\rangle$ does not depend on time, we can decompose it onto the basis of $|y\rangle$:

$$|\psi\rangle = \int dy \psi(y) |y\rangle$$

We have $\hat{p}^2 |\psi\rangle = |\psi'\rangle$ with:

$$\langle x|\psi'\rangle = -\hbar^2 \frac{\partial^2 \psi(x)}{\partial x^2} \quad (4.4)$$

and $\widehat{V(x)} |\psi\rangle = |\psi''\rangle$ with

$$\langle x|\psi''\rangle = V(x)\psi(x) \quad (4.5)$$

Thus, the stationary equation becomes:

$$\frac{|\psi'\rangle}{2m} + |\psi''\rangle = E|\psi\rangle \quad (4.6)$$

Then by projecting this relation onto the ket $|x\rangle$, we obtain:

Proposition 4.1.1 (Stationary equation). The wave function and eigenenergy of \hat{H} are defined by the equation:

$$\frac{-\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x) + V(x)\psi(x) = E\psi(x) \quad (4.7)$$

We can now obtain some general properties on the regularity of the wave function without specifying $V(x)$ further:

Proposition 4.1.2 (Regularity of ψ). The previous equation is to be considered in the sense of distributions. The following properties hold for ψ :

- ψ is a function in $L^2(\mathbb{C})$
- ψ is continuous on \mathbb{R} regardless of the form of V
- The derivative of ψ is continuous at any point where V has only finite discontinuities (*finite jumps*).

Proof. The first point follows from the probabilistic description forcing $\int dx |\psi(x)|^2 = \langle \psi|\psi \rangle = 1$.

The second point is proved by contradiction. If there exists a point a where ψ is not continuous, in the sense of distributions the derivative of ψ will contain a $\delta_a(x)$, which is not differentiable in the sense of distributions so $\frac{\partial^2}{\partial x^2} \psi(x)$ would not make sense. (The argument is actually a bit fallacious, and it is possible to rigorously show continuity for a continuous

potential V with a method similar to the one developed for ψ' .) For the third point, let $a \in \mathbb{R}$ such that $V(a^+) - V(a^-) < \infty$.

$$\begin{aligned}\psi'(a + \epsilon) - \psi'(a) &= \int_a^{a+\epsilon} dx \frac{\partial^2}{\partial x^2} \psi(x) = -\frac{2m}{\hbar^2} \int_a^{a+\epsilon} dx (E - V(x)) \psi(x) \\ &= -\frac{2m}{\hbar^2} \left(E \int_a^{a+\epsilon} dx \psi(x) - \int_a^{a+\epsilon} dx V(x) \psi(x) \right)\end{aligned}$$

The first integral goes to 0 because ψ is continuous at a . The second integral goes to 0 by noticing that we can find an upper bound of $V(x)$ on $[a + \epsilon, a]$ because the discontinuity is not infinite, hence:

$$\psi'(a + \epsilon) - \psi'(a) \xrightarrow{\epsilon \rightarrow 0} 0$$

Hence the continuity of ψ' at a . □

These continuity properties are very useful, especially in cases where the potential $V(x)$ is piecewise-defined, where different equations are obtained depending on the intervals studied, and where we want to find the boundary conditions for each of these intervals.

4.2 Infinite well

Let us now consider a less general case:

$$\begin{cases} V(x) = \infty & \text{for } x < 0 \\ V(x) = 0 & \text{for } 0 < x < a \\ V(x) = \infty & \text{for } a < x \end{cases}$$

The stationary Schrödinger equation already tells us that $\psi(x) = 0$ for $x \in [-\infty, 0] \cup [a, +\infty]$ (indeed, to satisfy this equation, we must have $\psi = 0$ to overcome the infinity of the potential).

We note that we have $\psi(0) = \psi(a) = 0$ by continuity of ψ . Also note that during the resolution, we will not be surprised not to have the continuity of ψ' at 0 and a because of the infinite potential jumps.

Let us write and solve the Schrödinger equation on $[0, a]$:

$$\frac{-\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x) = E \psi(x) \quad (4.8)$$

This is a second-order differential equation:

$$\frac{\partial^2}{\partial x^2} \psi(x) + \frac{2mE}{\hbar^2} \psi(x) = 0 \quad (4.9)$$

We are interested in states of positive energy (higher than the potential), otherwise it seems natural that we cannot propagate particles. (Negative energies are the equivalent of evanescent waves in electromagnetism, the equations can also be solved for negative energies.)

We then note $k = \sqrt{\frac{2mE}{\hbar^2}}$ and we have

$$\psi(x) = A e^{ikx} + B e^{-ikx} \quad (4.10)$$

Let us use the continuity of ψ at 0:

$$\psi(0) = A + B = 0 \quad (4.11)$$

Hence,

$$\psi(x) = C \sin(kx) \quad (4.12)$$

where $C \in \mathbb{C}$. Now, the continuity at a gives us

$$\psi(a) = C \sin(ka) = 0 \quad (4.13)$$

We obviously want $A \neq 0$, otherwise, nothing happens! Therefore, we must have $ka = n\pi$ with $n \in \mathbb{N}^*$ (once again, $n = 0$ is excluded, otherwise, $\psi = 0$ everywhere and nothing happens).

Proposition 4.2.1. The possible energies of a particle in an infinite potential well are quantized and given by

$$E_n = n^2 \frac{\hbar^2 \pi^2}{2ma^2} \quad \text{for } n \in \mathbb{N}^* \quad (4.14)$$

They correspond to non-degenerate eigenstates

$$\phi_n(x) \propto \sin\left(\frac{n\pi}{a}x\right) \quad \text{on } [0, a] \quad (4.15)$$

and are zero outside of $[0, a]$.

We can also determine C by remembering that ϕ_n must be normalized:

$$\begin{aligned} 1 &= \int_0^a dx |\phi_n(x)|^2 \\ &= \int_0^a dx |C|^2 \sin^2\left(\frac{n\pi}{a}x\right) \\ &= |C|^2 \int_0^a dx \left(\frac{1 - \cos\left(2\frac{n\pi}{a}x\right)}{2} \right) \\ &= |C|^2 \left(\frac{a}{2} - \frac{a}{2n\pi} \left[\sin\left(2\frac{n\pi}{a}x\right) \right]_0^a \right) \\ &= \frac{a}{2} |C|^2 \end{aligned}$$

We can then take $C = \sqrt{\frac{2}{a}}$, but we notice, and this is a general property, that we can only define the wave function up to a phase constant $e^{i\theta}$, which appears in the calculation of C .

4.3 A two-dimensional case

We can now consider confining a particle in a two-dimensional potential of the form

$$\begin{cases} V(x, y) = \infty & \text{for } x < 0; \text{ or } y < 0 \\ V(x, y) = 0 & \text{for } 0 < x < a; \text{ and } 0 < y < b \\ V(x, y) = \infty & \text{for } a < x; \text{ or } b < y \end{cases}$$

Finally, we can write $V(x, y) = V_x(x) + V_y(y)$ with

$$\begin{cases} V_x(x) = \infty & \text{for } x < 0 \\ V_x(x) = 0 & \text{for } 0 < x < a \\ V_x(x) = \infty & \text{for } a < x \end{cases} \quad \text{and} \quad \begin{cases} V_y(y) = \infty & \text{for } y < 0 \\ V_y(y) = 0 & \text{for } 0 < y < b \\ V_y(y) = \infty & \text{for } a < y \end{cases}$$

The Hamiltonian of the problem is then written as:

$$\widehat{H} = \frac{\widehat{p}_x^2}{2m} + \widehat{V}_x + \frac{\widehat{p}_y^2}{2m} + \widehat{V}_y = \widehat{H}_x + \widehat{H}_y \quad (4.16)$$

where \widehat{H}_x acts only on x and \widehat{H}_y acts only on y . We have $[\widehat{H}, \widehat{H}_x] = [\widehat{H}, \widehat{H}_y] = 0$ and we can use the important **Lemma (3.1.2)**. The eigenfunctions for \widehat{H} must also be eigenfunctions for \widehat{H}_x and \widehat{H}_y , hence:

Proposition 4.3.1. The eigenfunctions for the Hamiltonian of the two-dimensional infinite square well are indexed by two integers n_x and n_y , and are given by:

$$\phi_{n_x, n_y}(x, y) = \phi_{n_x}(x)\phi_{n_y}(y) \text{ for } (n_x, n_y) \neq (0, 0) \quad (4.17)$$

They are associated with the eigenenergies:

$$E_{n_x, n_y} = n_x^2 \frac{\hbar^2 \pi^2}{2ma^2} + n_y^2 \frac{\hbar^2 \pi^2}{2mb^2} \text{ for } (n_x, n_y) \neq (0, 0) \quad (4.18)$$

Proof. The eigenstates come from the previous reasoning. To obtain the eigenenergies, it is sufficient to substitute the found eigenfunction and to use the one-dimensional results telling us that $H_x \phi_{n_x}(x) = E_{n_x} \phi_{n_x}(x)$ while remembering that H_x has no action on $\phi_{n_y}(y)$ (it is a constant for H_x). \square

The same reasoning will be carried out for the 2D harmonic oscillator, so if the reasoning is not perfectly understood here, don't worry, it will be done again in another case which will probably help the reader's understanding.

5 Harmonic oscillator

"The career of a young theoretical physicist consists of treating the harmonic oscillator in ever-increasing levels of abstraction."

-Sidney Coleman

As you may have understood, in this section we will thoroughly explore the quantum harmonic oscillator. Let's apply our quantization process to obtain the Hamiltonian of the problem

$$E = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \longrightarrow \hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2 \hat{x}^2$$

In this energy, the first part is the kinetic energy and the second part is a Hooke potential energy, like that of a spring. In this section, we will find the eigenenergies and eigenstates of this Hamiltonian.

$$\boxed{\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2 \hat{x}^2} \quad (5.1)$$

5.1 Solution, eigenstates, eigenenergies

To solve this problem, two paths are possible. We could write equations as for the potential well and solve them analytically (which is very computationally heavy and clever) or use a purely algebraic method that we adopt here.

First of all, as in any good physics problem, let's make the quantities considered dimensionless.

We then define: $\tilde{x} = \sqrt{\frac{m\omega}{\hbar}} \hat{x}$ and $\tilde{p} = \sqrt{\frac{1}{m\omega\hbar}} \hat{p}$ so that

$$\hat{H} = \frac{1}{2}\hbar\omega(\tilde{p}^2 + \tilde{x}^2)$$

We then study the new dimensionless Hamiltonian

$$\tilde{H} = \frac{1}{2}(\tilde{p}^2 + \tilde{x}^2)$$

In this form, we cannot solve this Hamiltonian, so let's introduce two new operators that will be very important, not only for the resolution.

$$\hat{a} = \frac{\tilde{x} + i\tilde{p}}{\sqrt{2}} \quad \text{and then} \quad \hat{a}^\dagger = \frac{\tilde{x} - i\tilde{p}}{\sqrt{2}} \quad (5.2)$$

(In the calculation of \hat{a}^\dagger , we used the fact that the momentum and position operators are Hermitian because they are physical observables). Note that these two operators are not Hermitian and therefore do not represent anything physical!

We can then transform the Hamiltonian again into

$$\tilde{H} = \hat{a}^\dagger \hat{a} + \hat{\mathbb{I}}/2 = \hat{N} + \hat{\mathbb{I}}/2$$

where we have defined $\hat{N} = \hat{a}^\dagger \hat{a}$.

Proof. First, let's calculate $[\tilde{x}, \tilde{p}]$:

$$\begin{aligned} [\tilde{x}, \tilde{p}] &= \sqrt{\frac{m\omega}{\hbar}} \sqrt{\frac{1}{m\omega\hbar}} [\hat{x}, \hat{p}] \\ &= i \frac{1}{\hbar} \hbar \hat{\mathbb{I}} \\ &= i \hat{\mathbb{I}} \end{aligned}$$

Then we have

$$\begin{aligned} \hat{N} &= \hat{a}^\dagger \hat{a} \\ &= \frac{1}{2} (\tilde{x}^2 + \tilde{p}^2 - i\tilde{p}\tilde{x} + i\tilde{x}\tilde{p}) \\ &= \frac{1}{2} (\tilde{x}^2 + \tilde{p}^2) + \frac{i}{2} [\tilde{x}, \tilde{p}] \\ &= \tilde{H} - \frac{\hat{\mathbb{I}}}{2} \end{aligned}$$

□

We will use **Lemma (3.1.2)**. Obviously, $[\tilde{H}, \hat{N}] = 0$, so it suffices to find the eigenvectors of \hat{N} to obtain those of \tilde{H} , and the eigenenergies will be

$$\text{sp}(\hat{H}) = \hbar\omega(\text{sp}(\hat{N}) + \frac{1}{2}) \quad (5.3)$$

To determine the eigenvalues of \hat{N} , a significant purely mathematical algebraic work must be done, which will not be detailed here, and we simply give the results:

$$\text{sp}(\hat{N}) = \mathbb{N}$$

And we show that these eigenvalues are non-degenerate, which allows us to note the eigenket associated with the eigenvalue n , $|n\rangle$. These kets are orthonormal.

Proposition 5.1.1. The eigenkets of \hat{H} are then the $|n\rangle$ (they form an orthonormal basis) $\langle n|m\rangle = \delta_{n,m}$ for $n \in \mathbb{N}$ and the eigenvalues of \hat{H} (the energies that the system can take, see postulates) are given by

$$\hbar\omega(n + 1/2) \quad (5.4)$$

Note that even at the lowest level, there is a fundamental energy of $\hbar\omega/2$.

During the resolution of this problem, we obtain the three most important relationships:

$$\begin{cases} \hat{N}|n\rangle &= n|n\rangle \\ \hat{a}^\dagger|n\rangle &= \sqrt{n+1}|n+1\rangle \\ \hat{a}|n\rangle &= \sqrt{n}|n-1\rangle \end{cases}$$

Proof. The proof will be incomplete and assumes that $\hat{N}|n\rangle = n|n\rangle$ and that the eigenvalues are non-degenerate. However, it constitutes a good exercise in manipulating these operators.

Firstly, a simple calculation gives $[\hat{a}, \hat{a}^\dagger] = \hat{\mathbb{I}}$, then by noticing that

$$[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] - [\hat{C}, \hat{A}]\hat{B}$$

It follows that $[\hat{N}, \hat{a}^\dagger] = \hat{a}^\dagger$ and $[\hat{N}, \hat{a}] = -\hat{a}$.

Therefore, $\hat{N}(\hat{a}^\dagger |n\rangle) = \hat{a}^\dagger \hat{N} |n\rangle + \hat{a}^\dagger |n\rangle = (n+1)\hat{a}^\dagger |n\rangle$. Hence, $\hat{a}^\dagger |n\rangle \propto |n+1\rangle$.

Similarly, $\hat{N}(\hat{a} |n\rangle) = \hat{a} \hat{N} |n\rangle - \hat{a} |n\rangle = (n-1)\hat{a} |n\rangle$. Hence, $\hat{a} |n\rangle \propto |n-1\rangle$. Let us write $\hat{a}^\dagger |n\rangle = \beta |n+1\rangle$ and $\hat{a} |n\rangle = \alpha |n-1\rangle$. Then we have

$$\begin{aligned} n &= \langle n | \hat{N} | n \rangle \\ &= \langle n | \hat{a}^\dagger \hat{a} | n \rangle \\ &= (\hat{a} | n \rangle)^\dagger (\hat{a} | n \rangle) \\ &= \| \hat{a} | n \rangle \|^2 \\ &= \alpha^2 \end{aligned}$$

So $\alpha = \sqrt{n}$. Similarly:

$$\begin{aligned} n+1 &= \langle n | (\hat{N} + \hat{I}) | n \rangle \\ &= \langle n | (\hat{a}^\dagger \hat{a} + \hat{I}) | n \rangle \\ &= \langle n | \hat{a} \hat{a}^\dagger | n \rangle \\ &= (\hat{a}^\dagger | n \rangle)^\dagger (\hat{a}^\dagger | n \rangle) \\ &= \| \hat{a}^\dagger | n \rangle \|^2 \\ &= \beta^2 \end{aligned}$$

So $\beta = \sqrt{n+1}$. □

These relations lead us to call \hat{a}^\dagger the creation operator and \hat{a} the annihilation operator.

The relation on \hat{a}^\dagger is the most interesting because it allows us to quickly find all the eigenvectors:

Proposition 5.1.2.

$$|n\rangle = \frac{1}{\sqrt{n!}} (\hat{a}^\dagger)^n |0\rangle \quad (5.5)$$

Proof. We have

$$\begin{aligned} |n\rangle &= \frac{1}{\sqrt{n}} \hat{a}^\dagger |n-1\rangle \\ |n\rangle &= \frac{1}{\sqrt{n}} \hat{a}^\dagger \left(\frac{1}{\sqrt{n-1}} \hat{a}^\dagger |n-2\rangle \right) \\ &\vdots \\ |n\rangle &= \frac{1}{\sqrt{n!}} (\hat{a}^\dagger)^n |0\rangle \end{aligned}$$

□

Remark 5.1.3. One could also act \hat{a} and \hat{a}^\dagger on wave functions by noting that

$$a = \frac{1}{\sqrt{2}} \left(\sqrt{\frac{m\omega}{\hbar}} x + \hbar \frac{1}{\sqrt{m\hbar\omega}} \frac{\partial}{\partial x} \right) \quad \text{and} \quad a^\dagger = \frac{1}{\sqrt{2}} \left(\sqrt{\frac{m\omega}{\hbar}} x - \hbar \frac{1}{\sqrt{m\hbar\omega}} \frac{\partial}{\partial x} \right) \quad (5.6)$$

These expressions can be obtained using

$$\hat{a} = \frac{\tilde{x} + i\tilde{p}}{\sqrt{2}} \quad \hat{a}^\dagger = \frac{\tilde{x} - i\tilde{p}}{\sqrt{2}} \quad \tilde{x} = \sqrt{\frac{m\omega}{\hbar}} \hat{x} \quad \tilde{p} = \sqrt{\frac{1}{m\omega\hbar}} \hat{p} \quad \hat{p} = -i\hbar \hat{\partial}_x$$

It is then sufficient to find $|0\rangle$ (or equivalently $\phi_0(x) = \langle x|0\rangle$) to have (theoretical) access to all the eigenstates. To do this, the easiest way is to write that $a|0\rangle = 0$, which translates into the wave function as

$$\frac{1}{\sqrt{2}} \left(\sqrt{\frac{m\omega}{\hbar}} x + \hbar \frac{1}{\sqrt{m\hbar\omega}} \frac{\partial}{\partial x} \right) \phi_0(x) = 0$$

Proposition 5.1.4. The ket $|0\rangle$ is defined by

$$\phi_0(x) = \langle x|0\rangle = \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} \quad (5.7)$$

Proof. Solving the preceding first-order equation. \square

5.2 2D and 3D Harmonic Oscillators

We have considered a system that can only oscillate in a single direction. However, we can imagine a system that oscillates in two directions. (generalizing to 3 directions will be very easy). The Hamiltonian of the problem is then

$$\hat{H} = \frac{\hat{p}_x^2}{2m} + \frac{\hat{p}_y^2}{2m} + \frac{1}{2}m\omega^2(\hat{x}^2 + \hat{y}^2) = \hat{H}_x + \hat{H}_y \quad (5.8)$$

where \hat{H}_x and \hat{H}_y commute (a consequence of the last property on position and momentum operators). Thus, we have $[\hat{H}, \hat{H}_x] = [\hat{H}, \hat{H}_y] = 0$, which means by the **Lemma (3.1.2)** that there exists a basis of common eigenvectors for \hat{H} , \hat{H}_x , and \hat{H}_y . Therefore, if $\psi(x, y)$ is an eigenfunction of \hat{H} , then there exist $(n_x, n_y) \in \mathbb{N}^2$ such that

$$\begin{aligned} \psi(x, y) &\propto \phi_{n_x}(x) \\ &\propto \phi_{n_y}(y) \end{aligned}$$

Hence, $\psi(x, y) \propto \phi_{n_x}(x)\phi_{n_y}(y)$ and since the wave functions are normalized, we have $\psi(x, y) = \phi_{n_x}(x)\phi_{n_y}(y)$. It follows that:

Proposition 5.2.1 (Eigenstates of \hat{H}). The eigenstates of \hat{H} are denoted $|n_x, n_y\rangle$ with $(n_x, n_y) \in \mathbb{N}^2$, and the associated wave function is given by:

$$\phi_{n_x, n_y}(x, y) = \langle x, y | n_x, n_y \rangle = \langle x | n_x \rangle \langle y | n_y \rangle = \phi_{n_x}(x)\phi_{n_y}(y) \quad (5.9)$$

The corresponding eigenenergies are:

$$E_{n_x, n_y} = \hbar\omega \left(n_x + \frac{1}{2} \right) + \hbar\omega \left(n_y + \frac{1}{2} \right)$$

which can be rewritten as:

$$E_n = \hbar\omega(n + 1) \quad (5.10)$$

Proof. The eigenstates come from the previous reasoning. For the eigenenergies, it suffices to reinsert the wave function into the equation $\hat{H}\phi_{n_x, n_y}(x, y) = E_{n_x, n_y}\phi_{n_x, n_y}(x, y)$ and to use the fact that ϕ_{n_x} is constant for \hat{H}_y and is an eigenfunction for \hat{H}_x associated with the eigenenergy $\hbar\omega \left(n_x + \frac{1}{2} \right)$ and vice versa. By denoting $n_x + n_y = n$, we obtain the last relation. \square

Remark 5.2.2. Note that we have considered here that both harmonic oscillators have the same angular frequency, but there is no reason for this to be the case. If we now consider two angular frequencies ω_x and ω_y , redoing the calculation gives the new eigenenergies

$$E_{n_x, n_y} = \hbar\omega_x \left(n_x + \frac{1}{2}\right) + \hbar\omega_y \left(n_y + \frac{1}{2}\right) \quad (5.11)$$

As an exercise, it may be interesting to repeat the reasoning for a 3D harmonic oscillator. The reasoning is the same and if the frequencies along the three dimensions are the same, the eigenenergies will be:

$$E_n = \hbar\omega \left(n + \frac{3}{2}\right) \quad (5.12)$$

5.3 Application exercises

In this section, we show the importance of creation and annihilation operators through exercises.

Exercises:

1. Calculate $\langle \hat{x} \rangle_n$ and $\langle \hat{p} \rangle_n$
2. Calculate $\langle \hat{x}^2 \rangle_n$ and $\langle \hat{p}^2 \rangle_n$
3. For 2D and 3D harmonic oscillators, calculate the degeneracy of each energy.

Solutions: Firstly, recall that $\langle A \rangle_\psi = \langle \psi | \hat{A} | \psi \rangle$. Thus, it is not easy to directly calculate the requested quantities since the $|n\rangle$ do not have direct relations with the position and momentum operators. However, we can express these two operators in terms of creation and annihilation operators whose action on the $|n\rangle$ is known.

$$\hat{x} = \sqrt{\frac{\hbar}{m\omega}} \frac{\hat{a} + \hat{a}^\dagger}{\sqrt{2}} \quad \text{and} \quad \hat{p} = \sqrt{m\omega\hbar} \frac{\hat{a} - \hat{a}^\dagger}{i\sqrt{2}} \quad (5.13)$$

1.

$$\begin{aligned} \langle x \rangle_n &= \langle n | \hat{x} | n \rangle \\ &= \sqrt{\frac{\hbar}{2m\omega}} \langle n | (\hat{a} + \hat{a}^\dagger) | n \rangle \\ &= \sqrt{\frac{\hbar}{2m\omega}} \left(\langle n | \hat{a} | n \rangle + \langle n | \hat{a}^\dagger | n \rangle \right) \\ &= \sqrt{\frac{\hbar}{2m\omega}} \left(\sqrt{n} \langle n | n-1 \rangle + \sqrt{n+1} \langle n | n+1 \rangle \right) \\ &= 0 \end{aligned}$$

Similarly, we could show that $\langle \hat{p} \rangle_n = 0$. These results seem logical since we have chosen a symmetric potential centered at $x = 0$.

2. The method is the same, but we must keep in mind that \hat{a} and \hat{a}^\dagger do not commute and we cannot directly apply a notable identity. For example:

$$\begin{aligned}\langle \hat{p}^2 \rangle_n &= -\frac{m\omega\hbar}{2} \langle n | \left(\hat{a}^2 + (\hat{a}^\dagger)^2 - \hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a} \right) | n \rangle \\ &= \frac{m\omega\hbar}{2} \left(\langle n | \hat{a}\hat{a}^\dagger | n \rangle + \langle n | \hat{a}^\dagger\hat{a} | n \rangle \right) \\ &= \frac{2n+1}{2} m\hbar\omega\end{aligned}$$

Similarly, we can calculate $\langle \hat{x}^2 \rangle_n = \frac{2n+1}{2} \frac{\hbar}{m\omega}$.

3. Let's start with the 2D case.

First, let's understand that there is degeneracy. Indeed, the energies are written as $E_n = \hbar\omega(n+1) = \hbar\omega(n_x + n_y + 1)$. Thus, all pairs (n_x, n_y) that satisfy $n_x + n_y = n$ correspond to the same energy. It is then sufficient to count these pairs. When we choose n_x , n_y is automatically fixed by the condition $n_x + n_y = n$. As there are $n+1$ possible choices for n_x (between 0 and n), the degeneracy is $g_n = n+1$.

For the 3D case, the counting is slightly more complicated, with the condition becoming $n_x + n_y + n_z = n$. With n_x fixed, n_y can vary between 0 and $n - n_x$, and with n_x and n_y fixed, there is only one possible value for n_z . We then have

$$g_n = \sum_{n_x=0}^n \sum_{n_y=0}^{n-n_x} 1 \quad (5.14)$$

$$= \sum_{n_x=0}^n (n - n_x + 1) \quad (5.15)$$

$$= (n+1)(n+1) - \frac{n(n+1)}{2} \quad (5.16)$$

$$= \frac{1}{2}(n+2)(n+1) \quad (5.17)$$

6 Basics of perturbation theory

In this chapter, we do not introduce a new Hamiltonian, but rather a way to find the eigenstates and eigenenergies when a small perturbation is added to a known Hamiltonian. Let us consider a potential $V(x)$, symmetric in x and that can be developed in a power series. The Hamiltonian of the system will then be of the form

$$\hat{H} = \frac{\hat{p}^2}{2m} + \alpha \hat{x}^2 + \beta \hat{x}^4 + \dots \quad (6.1)$$

By assuming that terms of order higher than 2 are very weak, we *almost* recover the Hamiltonian of the harmonic oscillator that we know how to solve exactly. The goal is then to find the new eigenenergies and eigenstates due to this perturbation.

6.1 Principle

We consider a more general case than what was presented in the previous introduction and write

$$\hat{H} = \hat{H}_0 + \hat{W} \quad (6.2)$$

where \hat{W} represents a perturbation that is very weak (in eigenvalues) compared to \hat{H}_0 . Let us assume that we know exactly the eigenstates and eigenenergies of \hat{H}_0 .

Let us first consider that they are non-degenerate, denoted by $|\phi_n^{(0)}\rangle$ and $\epsilon_n^{(0)}$. We then have the following properties, which we will not prove here:

Proposition 6.1.1. The new eigenenergies are given by

$$\epsilon_j \simeq \epsilon_j^{(0)} + \langle \phi_j^{(0)} | \hat{W} | \phi_j^{(0)} \rangle + \sum_{k \neq j} \frac{|\langle \phi_j^{(0)} | \hat{W} | \phi_k^{(0)} \rangle|^2}{\epsilon_j^{(0)} - \epsilon_k^{(0)}} \quad (6.3)$$

and are associated with the eigenstates

$$|\phi_j\rangle \simeq |\phi_j^{(0)}\rangle + \sum_{k \neq j} \frac{\langle \phi_k^{(0)} | \hat{W} | \phi_j^{(0)} \rangle}{\epsilon_j^{(0)} - \epsilon_k^{(0)}} |\phi_k^{(0)}\rangle \quad (6.4)$$

at the 2nd order in energy and at the 1st order in state (the 2nd order in state is much too complicated to be used).

Remark 6.1.2. Two important remarks following this first point:

1. \hat{H} remains a Hamiltonian, which implies that $\hat{W}^\dagger = \hat{W}$ (this relation can sometimes be useful!)
2. When we are in finite dimension and the operators \hat{H}_0 and \hat{W} are put in matrix form in an eigenvector basis for \hat{H}_0 , then \hat{H}_0 is diagonal and $\langle \phi_j^{(0)} | \hat{W} | \phi_k^{(0)} \rangle$ corresponds to the entry in row j and column k of the matrix representing \hat{W}

The degenerate case is treated differently but is no more difficult and we will only express how to calculate the new eigenenergies. If the following property is not understood, don't panic, the exercises will allow you to understand the method to be used.

Proposition 6.1.3. In the case where an eigenenergy is degenerate, we restrict the operators to the degeneracy space by

$$\widehat{H}_g = \widehat{H}_{0_g} + \widehat{W}_g$$

and a good approximation of the new eigenvalues is to compute the eigenvalues of this new operator.

A precise application of these formulas will be made in the exercises.

6.2 Application exercise

Consider a system such that the dimension of \mathcal{H} is 3 (3 possible eigenstates for \widehat{H}_0). We can then write $\widehat{H}_0 = \begin{pmatrix} \epsilon_0 & 0 & 0 \\ 0 & \epsilon_0 & 0 \\ 0 & 0 & \epsilon_1 \end{pmatrix}$ and add a perturbation of the form: $\widehat{W} = \begin{pmatrix} 0 & W_{12} & W_{13} \\ W_{21} & 0 & W_{23} \\ W_{31} & W_{32} & W_{33} \end{pmatrix}$. Assume furthermore that $\epsilon_0 < \epsilon_1$.

Questions:

1. Give the new energy ϵ'_1 to first order in the theory of perturbations.
2. Note that the unperturbed energy state ϵ_0 is degenerate and write the Hamiltonian in the degeneracy space.
3. Note that \widehat{W} is a physical observable and deduce a condition on the coefficients of the operator \widehat{W}_g restricted to the degenerate subspace.
4. Conclude on the new perturbed energy ϵ'_0 . Show that the ground state is no longer degenerate. We say that there is lifting of degeneracy.

Solution:

1. Let's notice that the eigenvector associated with ϵ_1 is $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. By the first-order perturbation theory, we have:

$$\epsilon'_1 \simeq \epsilon_1 + (0 \ 0 \ 1) \widehat{W} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \epsilon_1 + W_{33}$$

2. The energy ϵ_0 is degenerate and has eigenvectors $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. The associated Hamiltonian in the degenerate space is then:

$$\widehat{H}_g = \widehat{H}_{0_g} + \widehat{W}_g = \begin{pmatrix} \epsilon_0 & W_{12} \\ W_{21} & \epsilon_0 \end{pmatrix} \quad (6.5)$$

3. Since \widehat{W} is a physical observable, it is Hermitian, so in particular, \widehat{W}_g is also Hermitian. By detailing the condition $\widehat{W}_g^\dagger = \widehat{W}_g$, we obtain

$$W_{12}^* = W_{21}$$

4. Perturbation theory indicates that the new eigenenergies are the eigenvalues of the Hamiltonian \widehat{H}_g . Its characteristic polynomial is

$$\begin{aligned}
 \det(X\mathbb{I}_2 - \widehat{H}_g) &= (X - \epsilon_0)^2 - W_{21}W_{12} \\
 &= (X - \epsilon_0)^2 - W_{12}^*W_{12} \\
 &= (X - \epsilon_0)^2 - |W_{12}|^2 \\
 &= (X - \epsilon_0 - |W_{12}|)(X - \epsilon_0 + |W_{12}|)
 \end{aligned}$$

The new eigenenergies are then

$$\epsilon_{0\pm} = \epsilon_0 \pm |W_{12}|$$

The new ground state has an energy of $\epsilon_0 - |W_{12}|$ and is no longer degenerate.

7 Angular momentum and spin

In classical mechanics, we encounter purely rotational systems whose energy is given by $E = \frac{L^2}{2I}$, where L is the angular momentum vector and I is the moment of inertia. After the quantization process, the associated Hamiltonian becomes

$$\boxed{\hat{H} = \frac{\hat{L}^2}{2I}} \quad (7.1)$$

Therefore, it is important to study the angular momentum.

7.1 Angular momentum operators

Classically, the orbital angular momentum is defined by

$$\vec{L} = \vec{r} \times \vec{p}$$

It is therefore natural to define the angular momentum operators as follows:

Definition 7.1.1 (Classical angular momentum). The classical angular momentum operators of a particle are defined as

$$\begin{cases} \hat{L}_x &= \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \\ \hat{L}_y &= \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \\ \hat{L}_z &= \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \end{cases}$$

and we define the square of the angular momentum as

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \quad (7.2)$$

These different operators do not all commute:

Proposition 7.1.2. We have

$$[\hat{L}_x, \hat{L}_y] = i\hbar \hat{L}_z \quad (7.3)$$

and this relation remains valid under cyclic permutation $x \rightarrow y \rightarrow z$, and we also have

$$[\hat{L}^2, \hat{L}_i] = 0 \quad (7.4)$$

for $i \in x, y, z$.

These relations show in particular that \hat{L}^2 and \hat{L}_z commute and thus have a common set of eigenvectors. Algebraic manipulations similar to those performed for the harmonic oscillator, which we will not detail here, give us the property:

Proposition 7.1.3. \hat{L}^2, \hat{L}_z form a complete set of commuting observables (CSCO) whose eigenkets will be denoted $|l, m_l\rangle$ with $l \in \mathbb{N}$ and $m_l \in -l, -l+1, \dots, l-1, l$ and satisfy the relations:

$$\begin{cases} \hat{L}^2 |l, m_l\rangle &= l(l+1)\hbar^2 |l, m_l\rangle \\ \hat{L}_z |l, m_l\rangle &= m_l\hbar |l, m_l\rangle \end{cases}$$

l quantifies the norm of the angular momentum and m_l its projection along the z -axis.

As classically, the angular momentum, when written in spherical basis, affects only the coordinates (θ, φ) , we can define a sort of wave function as for the harmonic oscillator:

Definition 7.1.4. We define the wave functions associated with the kets $|l, m_l\rangle$ by:

$$Y_{lm_l}(\theta, \varphi) = \langle \theta, \varphi | l, m_l \rangle \quad (7.5)$$

These are called spherical harmonics and are tabulated.

This definition is less clean because we have not defined the basis $|\theta, \varphi\rangle$. In reality, it will not be used and only the notation $Y_{lm_l}(\theta, \varphi)$ will be reused for the hydrogen atom, which is the only reason it is introduced here.

7.2 Spin

Experiments (such as the Stern-Gerlach experiment) and quantum field theory predict an intrinsic angular momentum for particles. It resembles classical angular momentum in almost every respect, except for its coupling with the magnetic field and the eigenvalues associated with its squared norm.

Definition 7.2.1 (Spin angular momentum). As for classical angular momentum, we define 3 spin operators $\widehat{S}_x, \widehat{S}_y, \widehat{S}_z$ and the operator for the squared norm of the spin:

$$\widehat{S}^2 = \widehat{S}_x^2 + \widehat{S}_y^2 + \widehat{S}_z^2 \quad (7.6)$$

Just as for classical angular momenta, we can find the eigenvalues of these operators. Here, we consider a system that is not rotating and only the spin can change.

Definition 7.2.2. $\widehat{S}^2, \widehat{S}_z$ form an E.C.O.C whose eigenvectors are denoted by $|s, m_s\rangle$ with $s \in \frac{1}{2}\mathbb{N}$ and $m_s \in -s, -s+1, \dots, s-1, s$, and satisfy the following relations:

$$\begin{cases} \widehat{S}^2 |s, m_s\rangle &= s(s+1)\hbar^2 |s, m_s\rangle \\ \widehat{S}_z |s, m_s\rangle &= m_s \hbar |s, m_s\rangle \end{cases}$$

In reality, for elementary particles, the spin s is fixed, and the states are only labeled by $|m_s\rangle$ (see the case of the electron below). Two groups of particles are then defined:

- If s is a half-integer, the particle is a fermion.
- If s is an integer, the particle is a boson.

Fermions and bosons have different properties for their wave function, which give rise to Pauli's exclusion principle.

Proposition 7.2.3 (Spin-statistics theorem). Consider a set of n particles. For an arbitrary Hamiltonian, the state of the system is defined by a wave function

$$\psi(\vec{r}_1, \dots, \vec{r}_n)$$

- If the N particles are bosons, then the wave function is totally symmetric, meaning that

$$\psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_j, \dots, \vec{r}_n) = \psi(\vec{r}_1, \dots, \vec{r}_j, \dots, \vec{r}_i, \dots, \vec{r}_n) \quad (7.7)$$

- If the N particles are fermions, then the wave function is totally antisymmetric, meaning that

$$\psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_j, \dots, \vec{r}_n) = -\psi(\vec{r}_1, \dots, \vec{r}_j, \dots, \vec{r}_i, \dots, \vec{r}_n) \quad (7.8)$$

Proof. The proof of that, here?! You're crazy, way too complicated. \square

This last property of fermions gives rise to Pauli's exclusion principle.

Proposition 7.2.4 (Pauli's exclusion principle). In a system of N fermions, two fermions cannot occupy the same quantum state.

Proof. Suppose that two fermions are in the same quantum state, then

$$\psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_i, \dots, \vec{r}_n) = -\psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_i, \dots, \vec{r}_n)$$

Hence the wave function is identically zero, which is not possible. \square

7.3 Study of electron spin

Often, the spin of a particle is actually fixed (this is not always the case when combining several particles as we will see in an application exercise). Thus, the quantum number s is not of real interest because it is fixed and only m_s has a degree of freedom. Here, we focus on the electron whose spin s is fixed and equals $s = 1/2$ (in particular, the electron is a fermion). If we consider only the spin of the electron, there are only two states corresponding to $m_s = \pm \frac{1}{2}$, which we will denote $|\uparrow\rangle$ (spin up, $m_s = 1/2$) and $|\downarrow\rangle$ (spin down, $m_s = -1/2$). The Hilbert space \mathcal{H} considered is then of dimension 2 and can be identified with \mathbb{R}^2 , we can then write:

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Under this representation, the spin operators take on the form of matrices:

$$\widehat{S}_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \widehat{S}_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \widehat{S}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (7.9)$$

The three matrices are called Pauli matrices, so that the commutation relations are well verified and the kets up and down are proper for \widehat{S}_z .

We note that a classical rotation can be easily added to this electron simply by noting that $\widehat{L}^2, \widehat{L}_z, \widehat{S}_z$ forms an E.C.O.C whose proper kets are denoted $|l, m_l, m_s\rangle$. However, in this case, the previous matrix notations are obviously no longer valid because $\dim \mathcal{H} \neq 2$.

7.4 Coupling to magnetic field

An important property of angular momentum is its coupling with magnetic field. Indeed, a charged particle with angular momentum will generate a magnetic moment that can couple with the present magnetic field \vec{B} . We can derive this result by a small calculation (whose rigor will not be the greatest). Consider, for example, a hydrogen atom, where we would describe the electron classically as spinning around the nucleus (see Figure 2). The electron, with charge $-e$, spinning in a loop generates a current which generates a magnetic moment $\vec{\mu}$. The excess energy generated by the coupling between the magnetic field and the magnetic moment is $-\vec{\mu} \cdot \vec{B}$. It remains to determine $\vec{\mu}$. Suppose that the electron spins on a circular orbit of radius r , then the current of the loop is $I = \frac{-ev}{2\pi r}$ with v the velocity of the electron. The magnetic moment generated is then $\mu = IS = \frac{-ev}{2\pi r} \times \pi r^2$. By noting that $L = mrv$, we obtain the famous equation:

$$-\vec{\mu} \cdot \vec{B} = -\gamma \vec{L} \cdot \vec{B}$$

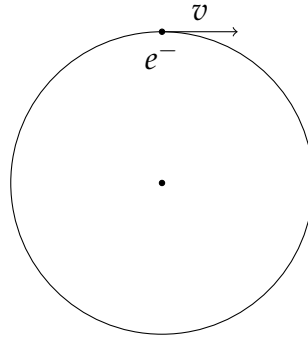


Figure 2: Modèle classique de l'atome d'hydrogène

(we note that the calculation is not rigorous because it alternates vectors and their norms, it is mainly a way to quickly find the value of γ).

Then, simply adding hats on the operators, we obtain:

$$\boxed{\hat{H} = -\gamma \hat{\vec{L}} \cdot \vec{B}} \quad (7.10)$$

with $\hat{\vec{L}} = \hat{L}_x \hat{e}_x + \hat{L}_y \hat{e}_y + \hat{L}_z \hat{e}_z$ (a way to condense the writing) and where $\gamma = \frac{q}{2m}$ is the gyromagnetic ratio of the particle considered. The operator $\hat{\vec{M}} = \gamma \hat{\vec{L}}$ is the magnetic moment operator.

For example, consider $\vec{B} = B \hat{e}_z$ with $B > 0$. The Hamiltonian becomes

$$\hat{H} = -\gamma B \hat{L}_z$$

Knowing the eigenvalues of \hat{L}_z , we can then find the associated energies.

However, we have omitted to consider the intrinsic spin angular momentum, which can also couple to the magnetic field, but quantum field theory shows that it couples slightly differently and then the Hamiltonian takes the form

$$\boxed{\hat{H} = -g\gamma \hat{\vec{S}} \cdot \vec{B}} \quad (7.11)$$

With g , the Landé factor, a constant that depends on the particle considered. We have $g \simeq 2$ for an electron.

To understand the behavior of spin in the presence of a magnetic field, consider a stationary electron in a uniform magnetic field $\vec{B} = B \hat{e}_z$ with $B > 0$. The Hamiltonian of the problem is then:

$$\hat{H} = \frac{e}{2m} B g \hat{S}_z \quad (7.12)$$

Thus, to minimize their energy, electrons tend to anti-align their spin with the magnetic field ($m_s = -1/2$ is energetically favorable). We will study these interactions more precisely in statistical mechanics.

Exercise: For a spin $3/2$ particle, a uniform magnetic field $\vec{B} = B \hat{e}_z$, what are the possible energies depending on γ, g, m ?

Solution: We have $s = 3/2$ then $m_s \in -3/2, -1/2, 1/2, 3/2$ hence the possible energies $E = -\gamma B g \hbar m_s$.

8 Hydrogenoid atom

"When it comes to atoms, language can be used only as in poetry. The poet, too, is not nearly so concerned with describing facts as with creating images."

-Niels Bohr

Let us now consider a more concrete and global problem. A single electron of charge $-e$ is in orbit around a punctual nucleus of charge $+Ze$. The Hamiltonian of the problem is then given by:

$$\hat{H} = \frac{\widehat{P}_Z^2}{2M_Z} + \frac{\widehat{p}_e^2}{2m_e} - \frac{Ze^2}{4\pi\epsilon_0|\widehat{r}_e - \widehat{R}_Z|} \quad (8.1)$$

The first two terms correspond to the kinetic energies of the nucleus and the electron, respectively, and the last term corresponds to the Coulomb interaction between the two.

We will solve this problem as a classical two-body problem.

8.1 Solution of the two-body problem

Let us introduce new position and momentum variables corresponding to the relative coordinates and the coordinates of the center of mass:

$$\begin{cases} \widehat{r} &= \widehat{r}_e - \widehat{R}_Z \\ \widehat{R} &= \frac{m_e\widehat{r}_e + M_Z\widehat{R}_Z}{m_e + M_Z} \\ \widehat{P} &= \widehat{p}_e + \widehat{P}_Z \\ \widehat{p} &= \frac{M_Z\widehat{p}_e - m_e\widehat{P}_Z}{m_e + M_Z} \end{cases}$$

We also introduce the total and reduced masses:

$$\begin{cases} m &= \frac{m_e M_Z}{m_e + M_Z} \\ M &= m_e + M_Z \end{cases}$$

These new variables allow us to decouple the problem into the study of the overall motion of the system and the relative motion. We then reduce to the case of the motion of a single particle. A series of calculations shows that

$$\hat{H} = \frac{\widehat{P}^2}{2M} + \frac{\widehat{p}^2}{2m} - \frac{Ze^2}{4\pi\epsilon_0\widehat{r}} \quad (8.2)$$

In reality, \widehat{p} and \widehat{r} are very close to the real momentum and position of the electron because the nucleus is much heavier. Thus, we can assume that they represent the electron's motion well. By placing ourselves in the center-of-mass reference frame, we have $\widehat{P} = 0$, and we then look for the eigenfunctions of the form $\psi_e(\vec{r})$ associated with the eigenenergies ϵ_e for the electron. Noting that

$$\widehat{p}^2 = -\frac{\hbar^2}{r} \frac{\partial^2}{\partial r^2} r + \frac{\widehat{L}^2}{r^2} \quad (8.3)$$

which comes from the expression of the spherical gradient, we can rewrite the stationary wave equation for the electron's wave function as

$$\frac{1}{2m} \left(-\frac{\hbar^2}{r} \frac{\partial^2}{\partial r^2} r + \frac{\widehat{L}^2}{r^2} \right) \psi_e(\vec{r}) - \frac{Ze^2}{4\pi\epsilon_0\widehat{r}} \psi_e(\vec{r}) = \epsilon_e \psi_e(\vec{r}) \quad (8.4)$$

Hence the Hamiltonian considered for this equation:

$$\widehat{H}_e = \frac{1}{2m} \left(-\frac{\hbar^2}{r} \widehat{\partial}_r^2 r + \frac{\widehat{L}^2}{r^2} \right) - \frac{Ze^2}{4\pi\epsilon_0 r} \quad (8.5)$$

We then have $[\widehat{H}, \widehat{L}^2] = 0$ (indeed, \widehat{L}^2 only acts on the angles (θ, φ)). Thus, using the most important lemma again, we have $\psi_e(\vec{r}) \propto Y_{lm_l}(\theta, \varphi)$ (where the Y_{lm_l} were defined in (7.5) as the eigenfunctions for \widehat{L}^2) for $l \in \mathbb{N}$ and $|m_l| \leq l$, and then

$$\psi_e(\vec{r}) = R(r) Y_{lm_l}(\theta, \varphi) \quad (8.6)$$

Then, by reintroducing into the equation using the action of \widehat{L}^2 (not forgetting that \widehat{r} and $\widehat{\partial}_r^2$ have no influence on Y_{lm_l}), we obtain the new equation:

$$-\frac{\hbar^2}{2m} \frac{1}{r} \frac{\partial^2}{\partial r^2} (rR(r)) + V_l(r)R(r) = \epsilon_e R(r) \quad (8.7)$$

where we have introduced the effective potential $V_l(r) = -\frac{Ze^2}{4\pi\epsilon_0 r} + \frac{\hbar^2 l(l+1)}{2mr^2}$. By making the substitution $P(r) = rR(r)$, we could solve this equation, but we will not do so here. Upon solving, for fixed l, m_l , we find several R functions that satisfy the equation, which we index using an integer n . The solution also tells us that n must be greater than l . More generally, we obtain the following result:

Proposition 8.1.1 (Hydrogen atom). $\widehat{H}, \widehat{L}^2, \widehat{L}_z, \widehat{S}_z$ form an E.C.O.C for the problem of the hydrogen atom. Thus, the eigenkets can be denoted as follows:

$$|n, l, m_l, m_s\rangle \text{ with } n \in \mathbb{N}^*, l \leq n-1, m_l \in -l, -l+1, \dots, l-1, l, m_s = \pm 1/2 \quad (8.8)$$

Moreover,

$$\langle n, l, m_l, m_s | \widehat{H} | n, l, m_l, m_s \rangle = -R \left(\frac{Z}{n} \right)^2 = -13.6 \left(\frac{Z}{n} \right)^2 \text{ eV} \quad (8.9)$$

where $R = \frac{m_e e^4}{8\epsilon_0^2 \hbar^2}$ is the Rydberg constant given in eV.

Note that $E_H = -13.6 \text{ eV}$ is the fundamental energy for the electron in the hydrogen atom.

One last property to keep in mind regarding the possible position of an electron is:

Proposition 8.1.2. The ground state of an electron (without considering its spin) is $|1, 0, 0\rangle$, whose wave function is defined as:

$$\phi_0(\vec{r}) = \langle r, \theta, \varphi | 1, 0, 0 \rangle = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0} \right)^{3/2} e^{-\frac{Zr}{a_0}} \quad (8.10)$$

where $a_0 = \frac{4\pi\epsilon_0 \hbar^2}{m_e e^2}$ is the Bohr radius.

8.2 Some application exercises

Exercise 1: How many quantum states are there for a fixed value of n ?

Exercise 2: Consider an electron/positron system. The previous demonstration applied to this system shows that it reduces to a single particle problem with the same characteristics as a single electron. However, this system can take on spin values of $s = 0$ and $s = 1$. How many quantum states are there for a fixed value of n ?

Exercise 3: The electron's angular momentum is measured to be $|L| = 2\hbar\sqrt{3}$. How many compatible states are there?

Solution 1: For each n , there are $n - 1$ possible values of l , and for each l , there are $2l + 1$ possible values of m_l . Taking into account the spin, the total number of states is then:

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

Solution 2: Without considering the spin, the system can be in n^2 states (as in the previous exercise). If we take into account the spin, in the $s = 0$ state, we have only one possible state with $m_s = 0$, and in the $s = 1$ state, we have $m_s \in -1, 0, 1$, which corresponds to three states. In total, there are then $4n^2$ possible states.

Solution 3: We know that during a measurement, $L^2 = l(l + 1)\hbar^2$ is a given. Thus:

$$2\hbar\sqrt{3} = \hbar\sqrt{l(l + 1)} \implies l = 3$$

When $l = 3$, m_l can take on $2l + 1 = 7$ values. Not forgetting a factor of 2 for the electron's spin, there are 14 possible states.

8.3 More Electrons and the Variational Method

In this final section on quantum physics, we detail an approach to approximate the energy of the ground state and the wave function of the ground state. Let us consider any Hamiltonian \hat{H} . Let $|\phi_n\rangle$ and ϵ_n be its kets and energies arranged in increasing order. Recall that the eigenkets form an orthonormal basis of \mathcal{H} . Note that we do not have access to these quantities and their introduction is only for the demonstration of the following property.

Having a good foundation in quantum mechanics, and in view of \hat{H} , our intuition tells us that the ground state will belong to a class of kets $|\psi_s\rangle_s$ (we will understand how to find this class in the example). In general, this method is applied numerically and works very well even with a very large class of functions to start with. How do we know which $|\psi_s\rangle$ to choose? We will rely on the following property.

Proposition 8.3.1. For any ket $|\psi_s\rangle$, we have

$$\langle \hat{H} \rangle_{\psi_s} = \langle \psi_s | \hat{H} | \psi_s \rangle \geq \epsilon_0 \quad (8.11)$$

Proof. The proof is quite simple. First, let us write

$$|\psi_s\rangle = \sum_n a_n |\phi_n\rangle$$

Note that $\langle \psi_s | \psi_s \rangle = 1$ implies that $\sum_n |a_n|^2 = 1$. Then, we calculate

$$\begin{aligned}
 \langle \psi_s | \hat{H} | \psi_s \rangle &= \sum_{k,n} a_k^* a_n \langle \phi_k | \hat{H} | \phi_n \rangle \\
 &= \sum_{k,n} a_k^* a_n \langle \phi_k | \epsilon_n | \phi_n \rangle \\
 &= \sum_{k,n} a_k^* a_n \epsilon_n \langle \phi_k | \phi_n \rangle \\
 &= \sum_{k,n} a_k^* a_n \epsilon_n \delta_{n,k} \\
 &= \sum_n |a_n|^2 \epsilon_n \\
 &\geq \epsilon_0 \sum_n |a_n|^2 = \epsilon_0
 \end{aligned}$$

□

Thus, we need to choose $|\psi_s\rangle$ that minimizes $\langle \psi_s | \hat{H} | \psi_s \rangle$ and we will then obtain an approximation of the ground state energy and its associated ket.

Application Exercise: Consider a helium atom with $2e^-$. Write the Hamiltonian of the problem (assuming the nucleus is fixed). Propose a class of functions for the wave function of the ground state. Justify your choice. Explain how to solve the problem.

Solution: The Hamiltonian for the fixed nucleus is given by:

$$\hat{H} = \frac{\hat{p}_1^2}{2m} - \frac{Ze^2}{4\pi\epsilon_0\hat{r}_1} + \frac{\hat{p}_2^2}{2m} - \frac{Ze^2}{4\pi\epsilon_0\hat{r}_2} + \frac{e^2}{4\pi\epsilon_0|\hat{r}_2 - \hat{r}_1|} \quad (8.12)$$

Where indices 1 and 2 refer to the 2 electrons and the last term is the interaction term between the two electrons. For a helium atom, we have $Z = 2$. We would like to write the wave function as

$$\phi(\vec{r}_1, \vec{r}_2) = \phi_0(\vec{r}_1)\phi_0(\vec{r}_2) \quad (8.13)$$

where ϕ_0 is the ground state found for the hydrogen-like atom. However, this is only true if there is no interaction between the electrons. By considering the physical situation, we realize that from the point of view of a single electron, the other electron only reduces the charge Z of the nucleus (called screening) to a charge Z' that we can try to determine. Let us choose the class of functions

$$\phi(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{\pi}} \left(\frac{Z'}{a_0} \right)^{3/2} e^{-\frac{Z'r_1}{a_0}} \frac{1}{\sqrt{\pi}} \left(\frac{Z'}{a_0} \right)^{3/2} e^{-\frac{Z'r_2}{a_0}} = \frac{1}{\pi} \left(\frac{Z'}{a_0} \right)^3 e^{-\frac{Z'(r_1+r_2)}{a_0}} \quad (8.14)$$

(We recall that the wave function should be anti-symmetric, which is generally achieved through the spin, but we will not detail this phenomenon here). It is then sufficient to calculate $\langle \phi | \hat{H} | \phi \rangle$, differentiate to find the Z' minimizing the energy, and thus obtain the ground state energy and associated wave function. The calculation is not trivial, however, as it involves integrals of the form

$$\int d^3\vec{r}_1 d^3\vec{r}_2 \phi(\vec{r}_1, \vec{r}_2)^* \hat{H} \phi(\vec{r}_1, \vec{r}_2)$$

which can prove to be non-trivial.

9 Statistical physicss

"It is the mark of a truly intelligent person to be moved by statistics"

-George Bernard Shaw

9.1 Foundations

So far, we have exclusively studied microscopic phenomena with a small number of particles. However, it can be interesting to understand the macroscopic behavior of a large number of particles N . In reality, this is not an easy task. Indeed, if we look at the phenomenon from a classical point of view, the state of the system at a time t is described by a vector of dimension $6N$:

$$\vec{\Gamma}(t) = (\vec{r}_1, \dots, \vec{r}_N, \vec{p}_1, \dots, \vec{p}_N) \quad (9.1)$$

And it becomes totally impossible to solve the equations of mechanics on these N particles! Such a state is called the microstate of the system and we will remember:

$$\boxed{\text{Microstate} \longleftrightarrow \vec{\Gamma}} \quad (9.2)$$

Here we used states defined classically, but quantum mechanics teaches us that particles are not localized and it is not always relevant to define a particle by its position and its momentum. The state can also be defined quantumly by a ket, for example by considering a system of N electrons whose spin we only study (we will come back to this example in the applications), a state can be:

$$|\uparrow, \uparrow, \downarrow, \dots, \downarrow\rangle \quad (9.3)$$

In quantum mechanics:

$$\boxed{\text{Microstate} \longleftrightarrow |\rangle \in \mathcal{H}} \quad (9.4)$$

Note that the number of microstates is finite here and is equal to 2^N . As argued earlier, what really interests us in a system with a large number of particles are its macroscopic characteristics (energy, temperature, pressure, ...). Let \mathcal{A} be a quantity (we make the connection with the physical observable of quantum mechanics), macroscopic characteristic of the system. The value of \mathcal{A} depends on the vector $\vec{\Gamma}$. There are two ways to consider the average of \mathcal{A} . Linking with quantum mechanics, one can calculate its statistical average over the states that the system can occupy:

$$A^* = \sum_l P_l \langle \hat{\mathcal{A}} \rangle_l \quad (9.5)$$

$$= \int d^3\vec{\Gamma} \rho(\vec{\Gamma}) \mathcal{A}(\vec{\Gamma}) \quad (9.6)$$

(the first equality being considered in the quantum case and the second in the classical case). P_l denotes the probability of being in the state l , and $\rho(\vec{\Gamma})$ is the probability density of the phase space.

Alternatively, we can calculate its time average:

$$\bar{A}_{\mathcal{T}} = \frac{1}{\mathcal{T}} \int_t^{t+\mathcal{T}} dt' \mathcal{A}(\vec{\Gamma}(t')) \quad (9.7)$$

In practice, and this can only be demonstrated in simple cases, we make the ergodic hypothesis, which assumes that the processes are so chaotic that over an interval of time of size \mathcal{T}

tending to infinity, the system will pass through all the states $\vec{\Gamma}$ and will remain in each state for a duration $\tau_{\vec{\Gamma}} = \mathcal{T} \rho(\vec{\Gamma})$. The ergodic hypothesis thus leads us to write:

$$\lim_{\mathcal{T} \rightarrow \infty} \bar{A}_{\mathcal{T}} \simeq A^* \quad (9.8)$$

This is why we will only study A^* . Therefore, for any quantity \mathcal{A} that we want to study, we need to know $\rho(\vec{\Gamma})$ (or $p(\phi)$), which then defines what we will call a macrostate:

$$\boxed{\text{Macrostate} \longleftrightarrow \rho(\vec{\Gamma})} \quad (9.9)$$

As for microstates, in quantum mechanics we will have rather:

$$\boxed{\text{Macrostate} \longleftrightarrow \{P_l, l \in \mathcal{H}\}} \quad (9.10)$$

From a notation point of view, we will no longer distinguish between the classical and quantum descriptions (except when explicitly stated), and we will denote l as a microstate. It should be noted that:

$$\sum_l P_l \longleftrightarrow \int d^3\vec{\Gamma} \rho(\vec{\Gamma}) \quad (9.11)$$

where, given the probability distribution P_l , we wish to determine a measure of information about our system. This measure is obtained through entropy (which is actually linked to thermodynamic entropy), defined as:

Definition 9.1.1.

$$S = -k_b \sum_l P_l \ln(P_l) \quad (9.12)$$

where the sum is only taken over accessible states (otherwise $P_l = 0$).

Now that all the relevant quantities are well-defined, we can study the macroscopic behavior of physical systems.

9.2 Microcanonical Ensemble

First, consider a physical system \mathcal{S} whose energy is fixed at E . We assume that there are a finite number of possible states $\Omega(E)$. Using the principle of maximum entropy (i.e., taking the distribution that maximizes entropy subject to the given constraints), we obtain:

Proposition 9.2.1 (Microcanonical Probabilities). In the microcanonical case where the energy of the system is fixed at E , all states l with energy E are equally probable:

$$P_l = \frac{1}{\Omega(E)} \quad (9.13)$$

Proof. Let u be the uniform distribution and p be an arbitrary distribution. For the proof, we take $k_b = 1$.

$$\begin{aligned} S(p) - S(u) &= - \sum_x p(x) \ln p(x) + \ln u \\ &= - \sum_x p(x) \ln p(x) + \sum_x p(x) \ln u(x) \\ &= \sum_x p(x) \ln \frac{u(x)}{p(x)} \\ &\leq \sum_x p(x) \left(1 - \frac{u(x)}{p(x)} \right) \\ &= 0 \end{aligned}$$

where we used $u(x) = u = \frac{1}{\Omega}$, $S(u) = -\sum_x u(x) \ln u(x) = -\ln u$, and the concavity of the logarithm. \square

This result is generalized to a uniform distribution in the classical case. The associated entropy is then:

Proposition 9.2.2 (Microcanonical Entropy).

$$S = k_b \ln \Omega(E) \quad (9.14)$$

Proof. Substitute P_i into the definition of entropy. \square

The statistical temperature is defined by analogy with thermodynamics, which in most cases corresponds to the real temperature:

Definition 9.2.3 (Statistical temperature and parameter β). The statistical temperature T is defined by:

$$\frac{1}{T} \equiv \frac{\partial S}{\partial E} \quad (9.15)$$

and the parameter β that we will continue to use is defined by:

$$\beta \equiv \frac{1}{k_b T} \quad (9.16)$$

Depending on the problem, by analogy with thermodynamics, if we denote V as the volume and N as the number of particles, we can also define pressure and chemical potential:

Definition 9.2.4.

$$p \equiv T \frac{\partial S}{\partial V} \quad (9.17)$$

$$\mu \equiv -T \frac{\partial S}{\partial N} \quad (9.18)$$

$$(9.19)$$

All these relationships have been defined with respect to classical thermodynamic identities, for example:

$$dE = TdS - PdV + \mu dN \quad (9.20)$$

Now let's consider several subsystems \mathcal{S}_i . For example, let's study the case with two subsystems and the complete system $\mathcal{S} = \mathcal{S}_1 \otimes \mathcal{S}_2$. If the two systems interact very little, the total energy of the system is

$$E_{\mathcal{S}} = E_1 + E_2 + E_{1 \leftrightarrow 2} \simeq E_1 + E_2$$

And the number of states of the system is

$$\Omega(E = E_1 + E_2) = \Omega(E_1)\Omega(E_2)$$

This implies the additivity of entropy:

$$\begin{aligned} S &= k_b \ln \Omega(E = E_1 + E_2) \\ &= k_b \ln (\Omega(E_1)\Omega(E_2)) \\ &= k_b (\ln \Omega(E_1) + \ln \Omega(E_2)) \\ S &= S_1 + S_2 \end{aligned}$$

9.3 Canonical ensemble

We now consider a system $\mathcal{S} = \mathcal{S}_1 \otimes \mathcal{S}_2$. The goal is to study \mathcal{S}_1 for which the fixed energy constraint of the microcanonical study has been relaxed. The fixed total energy is then written as

$$E = E_1 + E_2$$

And let us consider the subsystem \mathcal{S}_2 as an energy reservoir such that $E_1 \ll E_2 \simeq E$. The system \mathcal{S} , considered as isolated, can be treated using microcanonical methods but not the system \mathcal{S}_1 . Let $l_1 \otimes l_2$ be the microstates of \mathcal{S} , then we have:

$$P_{l_1 \otimes l_2} = \frac{1}{\Omega(E)}$$

We want to determine $P_{l_1}(E_1)$, the probability that system 1 is in microstate l_1 given that its energy is E_1 . The law of total probability gives:

$$P_{l_1}(E_1) = \sum_{\substack{l_2 \\ E_2 = E - E_1}} P(l_1 \cap l_2) = \sum_{\substack{l_2 \\ E_2 = E - E_1}} \frac{1}{\Omega(E)} = \frac{\Omega_2(E_2 = E - E_1)}{\Omega(E)}$$

Thus,

$$P_{l_1}(E_1) = \frac{\Omega_2(E_2 = E - E_1)}{\Omega(E)} \quad (9.21)$$

Since $E_1 \ll E$, we can expand $\ln \Omega_2(E - E_1)$ in a Taylor series:

$$\ln \Omega_2(E - E_1) = \ln \Omega_2(E) - \left. \frac{\partial \ln \Omega_2}{\partial E_2} \right|_E E_1 + \dots$$

Then, using the definition of β from the previous section (in fact, we show that the equality of temperatures between the two systems is a necessary equilibrium condition, which allows us to make this definition):

$$P_{l_1}(E_1) = \frac{e^{\ln \Omega_2(E - E_1)}}{\Omega(E)} \propto e^{-\beta E_1}$$

Since $E_{l_1} = E_1$, we have:

$$P_{l_1} \propto e^{-\beta E_{l_1}} \quad (9.22)$$

We then define the canonical partition function:

Definition 9.3.1 (Canonical partition function).

$$Z = \sum_l e^{-\beta E_l} \quad (9.23)$$

which is called the canonical partition function, and since the sum of probabilities is 1, we have:

$$P_l = \frac{e^{-\beta E_l}}{Z} \quad (9.24)$$

The probability of system 1 being in microstate l .

Remark 9.3.2. It is very important to understand the summation index in the partition function. The sum is taken over the microstates l and not over the energies E_l . To sum over the energies E_l , one must take into account the degeneracy of each energy, which is denoted by g_E . Thus, we have

$$\sum_l \longleftrightarrow \sum_E g_E \quad (9.25)$$

It turns out that the partition function contains all the information necessary to obtain all the characteristics of the system.

Proposition 9.3.3 (Average Energy).

$$\bar{E} = -\frac{\partial \ln Z}{\partial \beta} \quad (9.26)$$

Proof.

$$\begin{aligned} \bar{E} &= \sum_l P_l E_l \\ &= \frac{1}{Z} \sum_l e^{-\beta E_l} E_l \\ &= -\frac{1}{Z} \frac{\partial Z}{\partial \beta} \\ &= -\frac{\partial \ln Z}{\partial \beta} \end{aligned}$$

□

It is essential to be able to do this calculation without hesitation, as it is the most important calculation in statistical physics. We can also define the canonical entropy by the calculation:

$$\begin{aligned} S &= -k_b \sum_l P_l \ln(P_l) \\ &= -k_b \left(\sum_l -P_l \beta E_l - \sum_l P_l \ln Z \right) \\ &= k_b \ln Z + \frac{\bar{E}}{T} \end{aligned}$$

Proposition 9.3.4 (Canonical Entropy).

$$S = k_b \ln Z + \frac{\bar{E}}{T} \quad (9.27)$$

This relation echoes the free energy $F = E - TS$, from which we have:

Definition 9.3.5 (Free Energy).

$$F = -k_b T \ln Z \quad (9.28)$$

Then, as in thermodynamics, we have:

Proposition 9.3.6 (Chemical Potential).

$$\mu = \frac{\partial F}{\partial N} = -k_b T \frac{\partial \ln Z}{\partial N} \quad (9.29)$$

One last important property is:

Proposition 9.3.7 (Heat Capacity).

$$C_v = \frac{\partial E}{\partial T} = \frac{(\sigma_E)^2}{k_b T^2} \quad (9.30)$$

Proof. The proof is a good exercise and allows us to manipulate the equality

$$T \frac{\partial}{\partial T} = -\beta \frac{\partial}{\partial \beta}$$

(One should be able to retrieve this operator equality by explicating β ; it simplifies many calculations greatly) \square

We can now consider independent systems. We consider N independent subsystems, meaning that for a microstate l of the global system, it is possible to write $l = (\lambda_1, \dots, \lambda_N)$ with λ_i characterizing the state of subsystem i . The total energy is then written as

$$E = \sum_i E_{\lambda_i}$$

and we have

$$\sum_l \longrightarrow \sum_{\lambda_1} \sum_{\lambda_2} \dots \sum_{\lambda_N} \quad (9.31)$$

Then we have:

Proposition 9.3.8. For N independent subsystems, the global partition function is written as

$$Z = \prod_i z_i \quad (9.32)$$

where z_i is the partition function of subsystem i .

Proof.

$$\begin{aligned} Z &= \sum_l e^{-\beta E_l} \\ &= \sum_{\lambda_1} \sum_{\lambda_2} \dots \sum_{\lambda_N} e^{-\beta \sum_i E_{\lambda_i}} \\ &= \left(\sum_{\lambda_1} e^{-\beta E_{\lambda_1}} \right) \left(\sum_{\lambda_2} e^{-\beta E_{\lambda_2}} \right) \dots \left(\sum_{\lambda_N} e^{-\beta E_{\lambda_N}} \right) \\ Z &= \prod_i z_i \end{aligned}$$

\square

When the subsystems are identical, this equation reduces to $Z = z^N$. However, when the subsystems are identical and indistinguishable (indiscernible case), we cannot differentiate the ways to number the subsystems as $1, 2, \dots, N$. Thus, we must divide by $N!$:

Proposition 9.3.9 (Identical indistinguishable subsystems). For N identical indistinguishable subsystems:

$$Z = \frac{z^N}{N!} \quad (9.33)$$

We will study the case of the ideal gas where the particles are indistinguishable and the case of a crystal where the particles are identical but distinguishable as they are attached to fixed nodes of the crystal.

Classical approximation and ideal gas Let's place a particle in a box of volume $V = L_x L_y L_z$. Let's choose periodic boundary conditions, i.e., $\psi(\dots, x_i + L_{x_i}, \dots) = \psi(\dots, x_i, \dots)$. Quantum mechanics tells us that the particle can take wave vectors of the form

$$\vec{k} = 2\pi \left(\frac{n_x}{L_x}, \frac{n_y}{L_y}, \frac{n_z}{L_z} \right), \quad (9.34)$$

where n_x , n_y , and n_z are integers. When the box dimensions become large, the different states in wave vector space become increasingly close together (separation of $1/L^3$) and a continuum of states appears in wave vector space that can be expressed in terms of a density of states $\rho(\vec{k})$. Multiplying this density by $d^3\vec{k}$ gives the number of states between \vec{k} and $\vec{k} + d^3\vec{k}$.

The form of the wave vectors shows that each wave vector occupies an elementary volume $\frac{(2\pi)^3}{V}$ in wave vector space. Hence, by cross-multiplication, $\rho(\vec{k}) = \frac{V}{(2\pi)^3} d^3\vec{k}$. This yields

$$\sum_l \rightarrow \int d^3\vec{k} \frac{V}{(2\pi)^3} = \int d^3\vec{r} \int d^3\vec{p} \frac{1}{h^3}, \quad (9.35)$$

where we have used $\vec{p} = \hbar\vec{k}$.

A classical application of this result is the calculation of the partition function of an ideal gas of N particles in a volume V . Since there are N indistinguishable particles, we can use the result

$$Z = \frac{z^N}{N!},$$

where z is the partition function of a single particle in the gas. The energy of a particle is $\frac{\vec{p}^2}{2m}$ (ideal gas, so no interactions), so that

$$\begin{aligned} z &= \int d^3\vec{r} \int d^3\vec{p} \frac{1}{h^3} e^{-\beta \frac{\vec{p}^2}{2m}} \\ &= \frac{V}{h^3} \int dp_x dp_y dp_z e^{-\beta \frac{p_x^2 + p_y^2 + p_z^2}{2m}} \\ &= \frac{V}{h^3} \left(\int dp e^{-\beta \frac{p^2}{2m}} \right)^3 \\ z &= \frac{V}{\Lambda_T^3} \end{aligned}$$

where we have defined $\Lambda_T = \frac{h}{\sqrt{2\pi m k_B T}}$, which is called the de Broglie wavelength.

We will determine more characteristics of the ideal gas in applications.

Proposition 9.3.10. (Boltzmann's equipartition theorem) At thermal equilibrium, each independent squared degree of freedom in the energy is associated with an average energy quantum of $\frac{1}{2}k_B T$.

Proof. In general, we can write the energy in the form

$$E = H(p_i, q_j)$$

We focus on the particular case where the energy takes the following form,

$$H(p_i, q_j) = \alpha p_k^2 + \mathcal{H}(p_{i \neq k}, q_j)$$

Where the second term of the *rhs* of this equation doesn't not depend on p_k .
Let us calculate $\langle \alpha p_k^2 \rangle$:

$$\begin{aligned}
 \langle \alpha p_k^2 \rangle &= \frac{\int \frac{dp_1}{2\pi\hbar} \dots dq_1 \dots \alpha p_k^2 e^{-\beta E}}{\int \frac{dp_1}{2\pi\hbar} \dots dq_1 \dots e^{-\beta E}} \\
 &= \frac{\left(\int dq_k \prod_{i \neq k} dp_i dq_i e^{-\beta \mathcal{H}(p_{i \neq k}, q_j)} \right) \left(\int dp_k \alpha p_k^2 e^{-\beta \alpha p_k^2} \right)}{\left(\int dq_k \prod_{i \neq k} dp_i dq_i e^{-\beta \mathcal{H}(p_{i \neq k}, q_j)} \right) \left(\int dp_k e^{-\beta \alpha p_k^2} \right)} \\
 &= - \frac{1}{\int dp_k e^{-\beta \alpha p_k^2}} \frac{\partial}{\partial \beta} \left(\int dp_k e^{-\beta \alpha p_k^2} \right) \\
 &= - \sqrt{\frac{\beta \alpha}{\pi}} \frac{\partial}{\partial \beta} \left(\sqrt{\frac{\pi}{\beta \alpha}} \right) = - \sqrt{\beta} \frac{\partial \beta^{-1/2}}{\partial \beta} \\
 &= \frac{1}{2} \sqrt{\beta} \beta^{-3/2} = \frac{1}{2\beta} = \frac{1}{2} k_b T
 \end{aligned}$$

where we used the important relation :

$$\int_{-\infty}^{+\infty} dx e^{-ax^2} = \sqrt{\frac{\pi}{a}}$$

□

9.4 Applications

9.4.1 Spin orientation

Let us consider a crystal with N atoms. The i -th atom carries a magnetic moment $\vec{m}_i = \gamma \widehat{\vec{S}}^{(i)}$. To simplify, we assume that each atom behaves like an electron with a fixed total spin (s) and a spin projection (m_s) on the z axis $\left(\widehat{S}_z^{(i)} \right)$ that can take the values $\pm \hbar/2$. Thus, the i -th spin can be in the state $|\uparrow\rangle$ or $|\downarrow\rangle$.

Any microstate l of the system can then be written as

$$|\uparrow\uparrow\downarrow \dots \downarrow\rangle$$

Let's place the system in a magnetic field $\vec{B} = B\hat{e}_z$ and assume that the spins do not interact with each other. Consider the Hamiltonian (the energy) of the system which then becomes:

$$\hat{H} = -\vec{B} \cdot \vec{M} = -B\gamma \sum_i \widehat{S}_z^{(i)} \quad (9.36)$$

where $\vec{M} = \sum_i \vec{m}_i$ is the total magnetization of the system. Each spin can therefore take 2 energy values: $\pm B\gamma \frac{\hbar}{2} = \pm \epsilon_B$. Let's assume that the system is isolated and therefore its energy is fixed at E (this is really the mantra to remember: if the energy is fixed, the approach is microcanonical; if the temperature is fixed, the approach is canonical). Thus, the number

of atoms in the state $|\uparrow\rangle$, which we denote as n_+ , and the number of atoms in the state $|\downarrow\rangle$, which we denote as n_- , are fixed. We then have

$$N = n_+ + n_- \quad (9.37)$$

$$E = (-n_+ + n_-)\epsilon_B \quad (9.38)$$

These relationships can be inverted and we find:

$$n_{\pm} = \frac{1}{2} (N \mp E/\epsilon_B) \quad (9.39)$$

The number of microstates that are accessible to achieve the energy E is easily counted (we must select n_+ spins up out of N spins) and is given by

$$\Omega(E) = \binom{N}{n_+} = \frac{N!}{n_+!n_-!} \quad (9.40)$$

We can then calculate $S(E, N) = k_b \ln \Omega(E, N)$ using the Stirling's approximation $\ln(N!) \simeq N(\ln N - 1)$, leading to the formula:

$$\begin{aligned} S(E, N) &= k_b \ln \Omega(E, N) \\ &= k_b [\ln(N!) - \ln(n_+!) - \ln(n_-!)] \\ &\simeq k_b [N \ln N - N - n_+ \ln n_+ + n_+ - n_- \ln n_- + n_-] \\ &= k_b [N \ln N - n_+ \ln n_+ - n_- \ln n_-] \\ &= Nk_b \left[\ln 2 - \frac{(1 - \xi) \ln(1 - \xi) + (1 + \xi) \ln(1 + \xi)}{2} \right] \end{aligned}$$

where we have defined $\xi = E/N\epsilon_B \in [-1, +1]$.

9.4.2 Magnetization

We can also study the previous problem in another way by assuming the temperature to be T and relaxing the constraint on energy E . We now work in the canonical ensemble. There are N identical atoms that are distinguishable because they occupy fixed sites. The canonical partition function is then given by:

$$Z = z^N \quad (9.41)$$

where z is the partition function associated with a single atom. We can explicitly calculate z since there are only two possible states per atom with energies $\pm\epsilon_B$:

$$\begin{aligned} z &= e^{-\beta\epsilon_B} + e^{\beta\epsilon_B} \\ &= 2 \cosh(\beta\epsilon_B) \end{aligned}$$

We can now calculate the interesting characteristics of the system, starting with the energy:

$$\bar{E} = -\frac{\partial \ln Z}{\partial \beta} \quad (9.42)$$

$$= -N \frac{\partial \ln z}{\partial \beta} \quad (9.43)$$

$$= -N\epsilon_B \frac{\sinh(\beta\epsilon_B)}{\cosh(\beta\epsilon_B)} \quad (9.44)$$

$$\bar{E} = -N\epsilon_B \tanh(\beta\epsilon_B) \quad (9.45)$$

We are happy to find an extensive result that shows that the average energy of the system is the sum of the average energies per atom (which reassures us because we assumed that the atoms did not interact with each other). We could have anticipated this result by noting that all the quantities of interest depend on $\ln Z = N \ln z$, so we could calculate each quantity on a per-atom basis and sum them. Since the calculation with Z is not much more complicated, this is the approach we will keep. We can also calculate the entropy:

$$\begin{aligned} S(N) &= k_b \ln Z + \frac{\bar{E}}{T} \\ &= k_b N [\ln(2 \cosh(\beta \epsilon_B)) - \beta \epsilon_B \tanh(\beta \epsilon_B)] \end{aligned}$$

The last quantity of interest that we will calculate is the average magnetization \bar{M}_z . It was said earlier that

$$M_z = \sum_i m_z^{(i)}$$

Hence,

$$\bar{M}_z = \sum_i \bar{m}_z^{(i)} = N \bar{m}_z$$

Let's calculate \bar{m}_z . There are two possible states for m_z . If the considered spin is up, then $m_z = \gamma \hbar / 2$, if it is down $m_z = -\gamma \hbar / 2$. The probability of spin up is $e^{\beta \epsilon_B} / z$ (the $+\epsilon_B$ comes from $e^{-\beta(-\epsilon_B)}$), the probability of spin down is $e^{-\beta \epsilon_B} / z$, then we have

$$\begin{aligned} \bar{m}_z &= \frac{1}{z} \left(\frac{\gamma \hbar}{2} e^{\beta \epsilon_B} - \frac{\gamma \hbar}{2} e^{-\beta \epsilon_B} \right) \\ &= \frac{1}{z} \gamma \frac{\hbar}{2 \epsilon_B} \frac{\partial}{\partial \beta} (e^{\beta \epsilon_B} + e^{-\beta \epsilon_B}) \\ &= \frac{1}{B} \frac{\partial \ln z}{\partial \beta} \\ &= -\frac{\bar{\epsilon}}{B} \end{aligned}$$

where $\bar{\epsilon}$ is the average energy per atom.

9.4.3 Properties of the monatomic ideal gas

To conclude this section on statistical physics, let's go back to the model of the ideal gas and try to determine its other characteristics, in particular, to find the ideal gas equation. We recall the result

$$z = \frac{V}{\Lambda_T^3} \quad \text{with} \quad \Lambda_T = \frac{h}{\sqrt{2\pi m k_b T}} \quad (9.46)$$

and $Z = \frac{z^N}{N!}$.

Let's first calculate the average energy:

$$\begin{aligned}\bar{E} &= -\frac{\partial \ln Z}{\partial \beta} \\ &= -N \frac{\partial \ln z}{\partial \beta} \\ &= -N \frac{\partial \ln(\beta^{-3/2})}{\partial \beta} \\ &= \frac{3}{2}N \\ \bar{E} &= \frac{3}{2}Nk_bT\end{aligned}$$

It then naturally follows that

$$C_v = \frac{\partial E}{\partial T} = \frac{3}{2}Nk_b \quad (9.47)$$

which corresponds to the classical thermodynamics result for a monoatomic ideal gas (for a diatomic gas, the energy would need to include degrees of freedom associated with rotation). Next, let's calculate the pressure using the thermodynamic identity

$$dF = -SdT - pdV \quad (9.48)$$

which tells us that

$$p = -\frac{\partial F}{\partial V}$$

Remembering that $F = -k_bT \ln Z$, we have

$$F = -Nk_bT \ln(V) + \text{const.}(N, T, \dots) \quad (9.49)$$

where the second term is independent of V . Thus, we have

$$\begin{aligned}p &= -\frac{\partial F}{\partial V} \\ &= Nk_bT \frac{\partial \ln V}{\partial V} \\ &= Nk_bT \frac{1}{V} \\ &= nk_bT\end{aligned}$$

which is indeed the ideal gas law. We could calculate the other quantities, but as you may have gathered by now, it is not particularly exciting and will be intentionally omitted here.

Conclusion

Congratulations to all of you who have made it to the end of these fifty pages. I hope you have enjoyed discovering quantum mechanics. If my work has been well done, you should no longer perceive quantum theory as a forbidden black science. If some of you have even developed a taste for it, I am delighted.

I would like to take this opportunity to provide a few words for those who wish to go further. For those interested in applied physics, the next step is to perfect your knowledge in the field with works such as *Modern Quantum Mechanics* by J.J. Sakurai and then dive into condensed matter physics. For theorists, you will need to roll up your sleeves even more and dive headfirst into the most advanced theory to date, quantum field theory. To do this, I suggest taking a course on the basics of analytical mechanics (variational calculus) and then reading Tom Lancaster's magnificent work *Quantum Field Theory for the Gifted Amateur*. David Tong's lecture notes are also a real gift to humanity on the subject.

Before I let you go, I want to express my gratitude. Many thanks to Juliette Kalflèche, Bilel El Yaagoubi, and Quentin Berlioux for their valuable feedback, which allowed me to better understand the points that needed clarification. And more generally, thank you to ViaRézo for letting me squat in their flat and their space.