Advanced Quantum Mechanics[1]

Fall 2023

Bernat Frangi Mahiques
Physics Engineering and Industrial
Technologies Engineering

[1]Based on *Quatum Mechanics (Vol 1)* by Cohen Tannoudji and *Quantum Mechanics Concepts and Applications* by Nouredine Zettili.

Contents

1 Mathematical Tools of Quantum Mechanics

1.1 Hilbert Spaces

Hilbert spaces, also known as **complex vector spaces**, are defined in **Definition 1**.

> **Definition 1.** A Hilbert space $\mathcal{H}$ consists of a set of vectors ψ, ψ, χ, ... and a set of scalars a, b, c, ... which satisfy the following four properties:
>
> 1. **$\mathcal{H}$ is a linear space**[a].
>
> 2. **$\mathcal{H}$ has a defined scalar product that is strictly positive**. The scalar product of an element ψ with another element ϕ is in general a complex number, denoted by (ψ, ϕ). The scalar product satisfies the following properties[b]:
> $$(\psi, \phi) = (\phi, \psi)^* \tag{1}$$
> $$(\phi, a\psi_1 + b\psi_2) = a(\phi, \psi_1) + b(\phi, \psi_2) \tag{2}$$
> $$(a\phi_1 + b\phi_2, \psi) = a^*(\phi_1, \psi) + b^*(\phi_2, \psi) \tag{3}$$
> $$(\psi, \psi) = ||\psi||^2 \geq 0 \text{ (the equality holds only for } \psi = 0) \tag{4}$$
>
> 3. **$\mathcal{H}$ is separable**.
>
> 4. **$\mathcal{H}$ is complete**.
>
> ---
> [a]See **Definition 31** for the definition of linear vector space.
> [b]**Note:** Watch out for the order! Since the scalar product is a complex number, the quantity $(\psi, \phi) = \psi^*\phi$ is generall not equal to $(\phi, \psi) = \phi^*\psi$.

We should note that in a scalar product (ϕ, ψ), the second factor, ψ, belongs to the Hilbert space $\mathcal{H}$, while the first factor, ϕ, belongs to its dual Hilbert space $\mathcal{H}^{*}$[2]. The distinction between $\mathcal{H}$ and $\mathcal{H}^*$ is due to the fact that, as mentioned above, the scalar product is not commutative: $(\psi, \phi) \neq (\phi, \psi)$; the order matters!

1.2 The dual space

Given any Hilbert space $\mathcal{H}$, one can construct another complex vector space $\mathcal{H}^*$, called the **dual vector space**. It contains all the linear functionals in $\mathcal{H}$, which are a special kind of operator that maps all elements of $\mathcal{H}$ onto complex numbers[3]. In general, for an abstract vector space $\mathcal{H}$:

> **Definition 2.** Given a Hilbert space $\mathcal{H}$, the dual space $\mathcal{H}^*$ is the vector space of all linear functionals in $\mathcal{H}$.

Therefore, all linear functionals $L : \mathcal{H} \to \mathbb{C}$ live in $\mathcal{H}^*$ ($L \in V^*$).

The reason that the dual space is so interesting for quantum mechanics is that our goal as quantum physicists is to build a mathematical model for the real world, and in the end we want to be able to extract useful values and predictions from this model. For example, we may want to know the probability of getting a certain energy; or the average position we expect in a certain state. All these are scalar values, that we need to extract from a quantum state $|\psi\rangle$, so we know we will need a linear functional someplace or other!

---
[2]More on the dual space in the next section.
[3]See Appendix **Section 9.2** for more on linear functionals.

This may all sound really abstract at first glance, but hopefully it will become a lot clearer in the next section when we look at the **Dirac notation**.

1.3   Dirac Notation

In quantum mechanics, we use the Dirac notation to represent wave functions:

- We call the elements of $\mathcal{H}$ "ket" vectors, and we represent them as $|\psi\rangle \in \mathcal{H}$.

- We call the elements of $\mathcal{H}^*$ "bra" vectors, and we represent them as $\langle\phi| \in \mathcal{H}^*$.

Bra vectors are operators that linearly map elements of $\mathcal{H}$ into complex numbers:

$$\begin{array}{rcl} \langle\phi| : & \mathcal{H} & \rightarrow & \mathbb{C} \\ \langle\phi| : & |\psi\rangle & \rightarrow & \langle\phi\rangle\,\psi \end{array} \tag{5}$$

1.3.1   Inner product and bra-ket notation

Notice that, when we put a bra and a ket together ($\langle\phi|\,|\psi\rangle$), they look suspiciously like an inner product in this notation: $\langle\phi\rangle\,\psi$. If we go back at how our L_x operator in $\mathbb{R}^2$ acts on a column vector:

$$L_x \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = 1 \cdot a + 0 \cdot b = a \tag{6}$$

Notice that its action is the same as if we were taking the dot product with the x unit vector:

$$x \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = 1 \cdot a + 0 \cdot b = a \tag{7}$$

In fact, when a linear functional in $\mathbb{R}^n$ acts on any vector, it can be written equivalently as a dot product with the corresponding column vector:

$$L_x \vec{v} = L_x^T \cdot \vec{v} \tag{8}$$

This is actually a very general mathematical fact, rooted within something called the **Riesz Representation Theorem**:

> **Theorem 1. (Riesz Representation Theorem)** For any linear functional L_ϕ, the action of L_ϕ is equivalent to taking the inner product with some unique vector $\vec{\phi}$.

In our example of L_x, we have that $\vec{\phi} = \vec{x} = [1\ 0]^T$:

$$L_x \vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \vec{v} \tag{9}$$

This is the reason for the suggestive notation for bra vectors: they are operators whose action on a ket is mathematically equivalent to taking the inner product with said ket:

$$\langle\phi|\,|\psi\rangle = \langle\phi\rangle\,\psi \tag{10}$$

That is the power of bra-ket notation: it has the Riesz Representation Theorem baked right into it. Whatever you do, breaking apart inner products and putting together bras and kets, you will always have

something that makes mathematical sense. Although bra and the inner product are two entities that are completely different mathematically, the bra-ket notation makes their connection completely seamless, thanks to the Riesz Representation Theorem.

1.3.2 Properties of bras and kets

Some properties that arise naturally from the Dirac notation:

$$\langle \psi \rangle \lambda_1 \phi_1 + \lambda_2 \phi_2 = \lambda_1 \langle \psi \rangle \phi_1 + \lambda_2 \langle \psi \rangle \phi_2 \tag{11}$$

$$\langle \lambda_1 \psi + \lambda_2 \psi_2 \rangle \phi = \lambda_1^* \langle \psi_1 \rangle \phi + \lambda_2^* \langle \psi_2 \rangle \phi \tag{12}$$

$$\langle \psi \rangle \phi = \langle \phi \rangle \psi^* \tag{13}$$

$$\langle \psi \rangle \psi \text{ is real, positive and only zero if } |\psi\rangle = 0 \tag{14}$$

1.4 Linear operators

A linear map is defined as:

> **Definition 3.** A linear map (or linear operator) is a mathematical entity A that associates a function with another function such that:
>
> $$A(\lambda_1 \psi_1 + \lambda_2 \psi_2) = \lambda_1 A \psi_1 + \lambda_2 A \psi_2 \tag{15}$$

In the quantum mechanical context, we can see them as entities that transform a ket into another ket. Some example linear operators are:

- **Commutator:** The commutator of two operators A and B is defined as:

$$[A, B] \equiv AB - BA \tag{16}$$

  Two operators are said to commute if their commutator is equal to zero, and hence $AB = BA$. See **Section 9.5** to see an interesting application of commutator algebra for finding the uncertainty products of two operators.

- **Anti-commutator:** The anti-commutator of two operators A and B is defined as:

$$\{A, B\} \equiv AB + BA \tag{17}$$

- **Projector:** $P_\phi = |\phi\rangle \langle \phi|$. The projector operator P_ϕ acting on a ket $|\psi\rangle$ gives a new ket that is proportional (parallel) to $|\phi\rangle$. It "projects" the vector $|\psi\rangle$ onto $|phi\rangle$. The coefficient of proportionality is the scalar product $\langle \phi \rangle \psi$.[4]

- **Inverse:** assuming it exists, the inverse operator A^{-1} of the operator A, when applied to A, gives the identity operator. Also, A is the inverse of A^{-1}, so that $AA^{-1} = A^{-1}A = \mathbb{1}$.

- **Hermitian conjugation:** the hermitian conjugate (or adjoint) $A^\dagger$ of an operator A is obtained by interchanging the columns of the operator by its rows, and taking the complex conjugate of all elements. For example:

$$A = \begin{bmatrix} i & 1 \\ 3-i & -i \end{bmatrix} \rightarrow A^\dagger = \begin{bmatrix} -i & 3+i \\ 1 & i \end{bmatrix} \tag{18}$$

  Some properties of the adjoint are:

---

[4]*Proof:* $P_\phi |\psi\rangle = |\phi\rangle \langle \phi| |\psi\rangle = |\phi\rangle \langle \phi \rangle \psi = \langle \phi \rangle \psi |\phi\rangle$. Note that any projector operator P_ϕ fulfils $P_\phi^2 = |\phi\rangle \langle \phi| |\phi\rangle \langle \phi| = |\phi\rangle \langle \phi| = P_\phi$. In fact, this property completely characterises projection operators.

a) $\left(A^\dagger\right)^\dagger = A$.

b) $(\lambda A)^\dagger = \lambda^* A^\dagger$.

c) $(A + B)^\dagger = A^\dagger + B^\dagger$.

d) $(AB)^\dagger = B^\dagger A^\dagger$.

e) $(|u\rangle \langle v|)^\dagger = |v\rangle \langle u|$.[5]

The adjoint of a bra is its ket, and the adjoint of a ket is its bra. To obtain the hermitian conjugate of an expression:

a) Replace constants with their complex conjugate: $\lambda \to \lambda^*$.

b) Replace operators with their Hermitian conjugates: $A \to A^\dagger$.

c) Replace kets with bras: $|\phi\rangle \to \langle\phi|$.

d) Replace bras with kets: $\langle\phi| \to |\phi\rangle$.

e) Reverse the order of factors: $A |\phi\rangle \to \langle\phi| A^\dagger$.

A special case of linear operators are **unitary operators**:

Definition 4. A linear operator U is said to be unitary if its inverse U^{-1} is equal to its adjoint $U^\dagger$, so that $U^{-1} = U^\dagger$ and $U^\dagger U = UU^\dagger = \mathbb{1}$.

Another special case are **Hermitian operators**:

Definition 5. An operator A is said to be Hermitian if $A^\dagger = A$.

and **anti-Hermitian operators**:

Definition 6. An operator A is said to be anti-Hermitian if $A^\dagger = -A$.

An example of a Hermitian operator is the projector operator, as $P_\phi^\dagger = (|\phi\rangle \langle\phi|)^\dagger = |\phi\rangle \langle\phi| = P_\phi$.

1.4.1   Expected value of an operator

In order to define the expected value of an operator, we first need to define the **matrix element**:

Definition 7. Let $|\psi\rangle$, $|\phi\rangle$ be two kets, we call the matrix element of an operator A between $|\psi\rangle$ and $|\phi\rangle$ the quantity $\langle\psi| (A |\phi\rangle)$.

Note that the matrix element of an operator A between $|\psi\rangle$ and $|\phi\rangle$ is a complex number, and it is equal to the scalar product of $|\psi\rangle$ with the ket $A |\phi\rangle$. If we now define the expected value of an operator:

Definition 8. The expected value $\langle A\rangle_\psi$ of A in the state $|\psi\rangle$ is defined as the matrix element of A between $|\psi\rangle$ and itself:

$$\langle A\rangle_\psi = \langle\psi|A|\psi\rangle \tag{19}$$

---

[5]*Proof:* $\langle\phi| (|u\rangle \langle v|)^\dagger |\psi\rangle = [\langle\psi| (|u\rangle \langle v|) |\phi\rangle]^* = \langle\psi| u^* \langle v| \phi^* = \langle u| \psi \langle\phi| v = \langle\phi| v \langle u| \psi = \langle\phi| (|v\rangle \langle u|) |\psi\rangle$

It is easy to see that, if ψ is chosen to be an eigenvector of A, then the expected value of A in the state $|\psi\rangle$ is equal to the eigenvalue λ of A corresponding to the eigenvector $|\psi\rangle$:

$$\langle\psi|A|\psi\rangle = \langle\psi|\,(A\,|\psi\rangle) = \langle\psi|\,(\lambda\,|\psi\rangle) = \lambda\,\langle\psi|\psi\rangle = \lambda \tag{20}$$

This means that, for an arbitrary vector ϕ expressed as a linear combination of eigenvectors ψ_i of A:

$$|\phi\rangle = \sum_i c_i\,|\psi_i\rangle \tag{21}$$

we have:

$$\langle\phi|A|\phi\rangle = \sum_i\sum_j c_i^* c_j\,\langle\psi_i|A|\psi_j\rangle = \sum_i\sum_j c_i^* c_j\,\langle\psi_i|\lambda_j|\psi_j\rangle = \sum_i\sum_j c_i^* c_j\lambda_j\,\langle\psi_i|\psi_j\rangle =$$

$$= \sum_i\sum_j c_i^* c_j\lambda_j\delta_{ij} = \sum_i c_i^* c_i\lambda_i = \sum_i |c_i|^2\lambda_i \tag{22}$$

1.5   Closure relation

For a set of vectors to form a basis of a Hilbert space $\mathcal{H}$, they must fulfil the **closure relation** (also known as the completeness relation). In simple terms, if the set of vectors fulfils the closure relation, it means that with those vectors you can reach all possible directions in $\mathcal{H}$, and any $|\psi\rangle \in \mathcal{H}$ is a linear combination of those basis vectors. In our general Hilbert space:

> **Definition 9.** A set of vectors $\{|A_1\rangle, |A_2\rangle, ...\}$ in a Hilbert space $\mathcal{H}$ form a basis for $\mathcal{H}$ if and only if they fulfil the closure relation:
> $$\mathbb{1} = \sum_i |A_i\rangle\langle A_i|. \tag{23}$$

Proof of this relation is given in **Section 9.4**.

1.6   Wave function space $\mathcal{F}$

The wave function in quantum mechanics is an object whose modulus squared is a probability density function. If we look back at **Definition 1**, we can see that, from a physical point of view, the set $\mathcal{H}$ is clearly too wide in scope for our purposes. We need to restrict it to a subset of $\mathcal{H}$ that is physically meaningful. This subset is called the **wave function space** $\mathcal{F}$, and it retains only the functions ψ of $\mathcal{H}$ which are everywhere defined, continuous, and infinitely differentiable. In addition, the functions of $\mathcal{F}$ must be normalizable by an arbitrary multiplicative constant in such a way that the area under the curve $|\psi|^2$ is exactly equal to 1.

1.7   Basis of the wave function space

If $\{|\psi_i\rangle\}$ is a basis for a Hilbert space $\mathcal{H}$ (in particular, the subset $\mathcal{F}$), we can expand every arbitrary vector $|\Psi\rangle$ according to this basis:

$$|\Psi\rangle = \sum_i c_i\,|\psi_i\rangle \tag{24}$$

so that we have[6]:

$$\langle\psi_i|\,\Psi = \sum_j c_i\,\langle\psi_i|\,\Psi = \sum_j c_i\delta_{ij} = c_i \tag{25}$$

---

[6]δ_{ij} is known as Kronecker's delta, and is equal to 1 if $i = j$ and 0 otherwise.

The choice of basis is arbitrary, and depending on the choice we make, we obtain different representations of state space. There are many different representations, which often have to do with physical properties of the system.

If you are familiar with basic quantum mechanics, you will probably have seen the position representation of the wave function, $\Psi(\vec{r})$. This representation of state space is particularly useful for working with position in a quantum system. However, it is not the *only* representation of state space that can have. What does this mean? Well, here, there is an important concept to understand, which is the difference between a vector and its representation: a vector is a mathematical entity that, once defined, is the same all the time, no matter where we look at it from or which basis we express it in. Its representation, however, may differ, depending on which basis we choose to represent it in. If the basis changes, the coordinates will also change, even though the vector is still the same.

For example, a vector $\vec{v}_B = (a, b)$ expressed in the basis $B = \{(1,0),\ (0,1)\}$ will change to $\vec{v}_{B'} = (b/2, a)$ $B' = \{(0,2),\ (1,0)\}$. The coordinates of the vector have changed, but we can see that both representations refer to the same vector:

$$\vec{v}_B = a \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + b \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{b}{2} \cdot \begin{pmatrix} 0 \\ 2 \end{pmatrix} + a \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \vec{v}_{B'} \tag{26}$$

Just as in this example, the position representation of the wave function vector space is only one of its many possible representations. The functions $\psi_i(\vec{r})$ form the basis for the position representation of state space. Other representations, like the momentum representation, can be useful in certain situations, as we will see later on.

1.8   Representations in state space

When studying quantum mechanical systems, we need a way to represent quantum states. We do that by choosing an orthonormal basis, either discrete or continuous, in the state space $\mathcal{F}$. Vectors and operators are then represented in this basis by numbers: components for the vectors and matrix elements for the operators.

As we mentioned before, the choice of a representation is, in principle, arbitrary. In fact, it depends on the particular problem being studied: in each case, one chooses the representation that leads to the simplest calculations.

Most useful bases come as eigenstates of some pertinent operator[7]. So far, we have mentioned the position and the momentum representation. These bases deal with the position and momentum operators, but you can think of many others. For example, the eigenstates of the Hamiltonian for some physical system are often used, especially when solving the Schrödinger equation. These might also be infinite dimensional but can be discrete, as opposed to the continuous bases of position and momentum.

1.8.1   General representation

For a general representation of state space (wave function space, $\mathcal{F}$), the elements of $\mathcal{F}$ are functions $\psi(\vec{\xi})$, $\phi(\vec{\xi})$, and the inner product in $\mathcal{F}$ is defined as:

$$\left(\phi(\vec{\xi}),\ \psi(\vec{\xi}) \right) \equiv \int \phi^*(\vec{\xi}) \psi(\vec{\xi}) d\vec{\xi} \tag{27}$$

The length (norm) of a vector is given by:

---

[7]More on this later add link, but the idea is that, if we represent the state space in the basis of eigenstates (analogous to eigenvectors) of an operator, then that operator will be expressed in that basis as a diagonal matrix, where the elements of the diagonals are the eigenvalues of the operator. This makes calculations very easy.

$$\text{length}\left(\psi(\vec{\xi})\right) = \sqrt{\left(\psi(\vec{\xi}),\ \psi(\vec{\xi})\right)} = \sqrt{\int \psi^*(\vec{\xi})\psi(\vec{\xi})d\vec{\xi}} = \sqrt{\int \left|\psi(\vec{\xi})\right|^2 d\vec{\xi}} \tag{28}$$

As the vectors of the basis $\{\phi_i\}$ of $\mathcal{F}$ are orthonormal, we have that:

$$\left(\phi_i(\vec{\xi}),\ \phi_j(\vec{\xi})\right) = \delta_{ij} \tag{29}$$

1.8.2   Discrete orthonormal bases

A **discrete orthonormal basis** is defined as:

Definition 10. A countable set of functions $\{u_i\}$ is called orthonormal if:

$$\left(u_i(\vec{\xi}),\ u_j(\vec{\xi})\right) = \delta_{ij} \tag{30}$$

And it constitutes a basis for $\mathcal{F}$ if every function in $\mathcal{F}$ can be written as a linear combination of the functions of the basis in exactly one way:

$$\psi(\vec{\xi}) = \sum_i c_i u_i(\vec{\xi}) \tag{31}$$

with the coefficients being:

$$c_i = \left(u_i(\vec{\xi}),\ \psi(\vec{\xi})\right) = \int u_i^*(\vec{\xi})\psi(\vec{\xi})d\vec{\xi} \tag{32}$$

Note that, in a discrete orthonormal basis, all basis vectors u_i are elements belonging to $\mathcal{F}$. In other words, every basis vector u_i is a valid physical state for the system in the space $\mathcal{F}$. As we will see, this will *not* be the case for continuous orthonormal "bases".

For a discrete orthonormal basis, we can express the **scalar product in terms of the components** as:

$$\left(\phi(\vec{\xi}),\ \psi(\vec{\xi})\right) = \left(\sum_i b_i u_i(\vec{\xi}),\ \sum_j c_j u_j(\vec{\xi})\right) = \sum_{i,\,j} b_i^* c_j \left(u_i(\vec{\xi}),\ u_j(\vec{\xi})\right) = \sum_{i,\,j} b_i^* c_i \delta_{ij} = \sum_i b_i^* c_i \tag{33}$$

With a similar proof as in **Section 9.4**, we can find the closure relation for a discrete orthonormal basis:

$$\psi(\vec{\xi}) = \sum_i c_i u_i(\vec{\xi}) = \sum_i \left(u_i(\vec{\xi}),\ \psi(\vec{\xi})\right) u_i(\vec{\xi}) = \sum_i \left(\int u_i^*(\vec{\xi}')\psi(\vec{\xi}')d\vec{\xi}'\right) u_i(\vec{\xi}) = $$

$$= \int \left(\sum_i u_i^*(\vec{\xi}')u_i(\vec{\xi})\right) \psi(\vec{\xi}')d\vec{\xi}' \tag{34}$$

Therefore, the term in the parenthesis must be equal to 1 for $\vec{\xi} = \vec{\xi}'$ and zero for every other case, so we obtain the **closure relation for a discrete orthonormal basis**[8]:

$$\sum_i u^*(\vec{\xi}')u_i(\vec{\xi}) = \delta(\vec{\xi} - \vec{\xi}') \tag{35}$$

---

[8] $\delta(\vec{\xi} - \vec{\xi}')$ is known as Dirac's delta function, and it is equal to 1 if $\vec{\xi} = \vec{\xi}'$ and 0 otherwise. It can also be expressed as the integral $\delta(\vec{\xi} - \vec{\xi}') = \frac{1}{(2\pi)^3} \int e^{i\vec{k}\cdot(\vec{\xi} - \vec{\xi}')} d^3 k$.

1.8.3 Continuous orthonormal bases

A **continuous orthonormal basis** is defined as:

> **Definition 11.** A continuous set of functions $\{\omega_\alpha\}$, labelled by a continuous index α, is called orthonormal if:
>
> $$\left(\omega_\alpha(\vec{\xi}),\ \omega_{\alpha'}(\vec{\xi})\right) = \delta(\alpha - \alpha') \tag{36}$$
>
> And it constitutes a basis for $\mathcal{F}$ if every function in $\mathcal{F}$ can be written as a linear combination of the functions of the basis in exactly one way:
>
> $$\psi(\vec{\xi}) = \int c(\alpha)\omega_\alpha(\vec{\xi})d\alpha \tag{37}$$
>
> with the continuous coefficient being:
>
> $$c(\alpha) = \left(\omega_\alpha(\vec{\xi}),\ \psi(\vec{\xi})\right) = \int \omega_\alpha^*(\vec{\xi})\psi(\vec{\xi})d\vec{\xi} \tag{38}$$
>
> Note that, $\langle\omega_\alpha(\vec{\xi})|\omega_{\alpha'}(\vec{\xi})\rangle = \delta(\alpha - \alpha')$ implies that the functions ω_α are not normaliseable[a], so these functions are *not* vectors in $\mathcal{F}$. Therefore, strictly speaking, they cannot be a basis for $\mathcal{F}$. In other words, basis vectors ω_α are *not* valid physical states for the system in the space $\mathcal{F}$. Rather, they are a mathematical tool that can help us to perform calculations in certain scenarios, and are formalised in what is known as a rigged Hilbert space[b]. However, we will still refer to them as "basis vectors" for simplicity.
>
> ---
>
> [a]As $\langle\omega_\alpha(\vec{\xi})|\omega_{\alpha'}(\vec{\xi})\rangle = \delta(\alpha - \alpha')$ would imply that $\|\omega_\alpha\|^2 = \langle\omega_\alpha(\vec{\xi})|\omega_\alpha(\vec{\xi})\rangle = \delta(\alpha - \alpha) = \infty \neq 1$. So, it turns out that any basis in $\mathcal{F}$ has to be a discrete basis with an orthogonality condition expressed in terms of a Kronecker delta instead of a Dirac delta.
> [b]See this and this post for more information.

For a continuous orthonormal basis, we can express the **scalar product in terms of the continuous coefficients** as:

$$\left(\phi(\vec{\xi}),\ \psi(\vec{\xi})\right) = \left(\int b(\alpha)\omega_\alpha(\vec{\xi})d\alpha,\ \int c(\alpha')\omega_{\alpha'}(\vec{\xi})d\alpha'\right) = \sum_{i,\,j} b_i^* c_j \left(u_i(\vec{\xi}),\ u_j(\vec{\xi})\right) =$$

$$= \int\left(\int b^*(\alpha)c(\alpha')(\omega_\alpha(\vec{\xi}),\omega_{\alpha'}(\vec{\xi}))d\alpha\right)d\alpha' = \int\left(\int b^*(\alpha)c(\alpha')\delta(\alpha - \alpha')d\alpha\right)d\alpha'$$

$$= \int b^*(\alpha)c(\alpha)d\alpha \tag{39}$$

With a similar proof as in **Section 9.4**, we can find the closure relation for a continuous orthonormal basis:

$$\psi(\vec{\xi}) = \int c(\alpha)\omega_\alpha(\vec{\xi})d\alpha = \int\left(\omega_\alpha(\vec{\xi}),\ \psi(\vec{\xi})\right)\omega_\alpha(\vec{\xi})d\alpha = \int\left(\int \omega_\alpha^*(\vec{\xi'})\psi(\vec{\xi'})d\vec{\xi'}\right)\omega_\alpha(\vec{\xi})d\alpha =$$

$$= \int\left(\int \omega_\alpha^*(\vec{\xi'})\omega_\alpha(\vec{\xi})d\alpha\right)\psi(\vec{\xi'})d\vec{\xi'} \tag{40}$$

Therefore, the term in the parenthesis must be equal to 1 for $\vec{\xi} = \vec{\xi'}$ and zero for every other case, so we obtain the **closure relation for a continuous orthonormal basis**:

$$\int \omega_\alpha^*(\vec{\xi'})\omega_\alpha(\vec{\xi})d\alpha = \left(\omega_\alpha(\vec{\xi'}),\ \omega_\alpha(\vec{\xi})\right) = \delta(\vec{\xi} - \vec{\xi'}) \tag{41}$$

Fourier transform

Take the inverse fourier transform of the position wave function, for example:

$$\psi(\vec{r}) = \frac{1}{\sqrt{2\pi\hbar}} \int \overline{\psi}(\vec{p}) e^{i\vec{p}\cdot\vec{r}/\hbar} d\vec{p} = \int \overline{\psi}(\vec{p}) v_{\vec{p}}(\vec{r}) d\vec{p}, \qquad v_{\vec{p}}(\vec{r}) = \frac{1}{\sqrt{2\pi\hbar}} e^{i\vec{p}\cdot\vec{r}/\hbar} \text{ (plane wave)} \qquad (42)$$

Notice that this is the same as:

$$\psi(\vec{\xi}) = \int c(\alpha) \omega_\alpha(\vec{\xi}) d\alpha \qquad (43)$$

where:

$$\alpha \to \vec{p} \text{ (continuous index)}, \quad c(\alpha) \to \overline{\psi}(\vec{p}), \quad \omega_\alpha(\vec{\xi}) \to v_{\vec{p}}(\vec{r}) \text{ (basis functions)} \qquad (44)$$

And the "continuous coefficient" function $\overline{\psi}$ can be found as:

$$\overline{\psi}(\vec{p}) = (v_{\vec{p}}, \psi) = \int v_{\vec{p}}^*(\vec{r}) \psi(\vec{r}) d\vec{r} \qquad (45)$$

It would seem that the set of uncountable $v_{\vec{p}}$ functions is a basis for $\mathcal{F}$, however, the integral of $\left|v_{\vec{p}}(\vec{r})\right|^2 = \frac{1}{2\pi\hbar}$ diverges, so $v_{\vec{p}}(\vec{r}) \notin \mathcal{F}$.

Delta function

In the same way, we can introduce the set of functions $\{\xi_{\vec{r}_0}(\vec{r})\}$ of $\vec{r}$, labelled by the continuous index $\vec{r}_0$ and defined as:

$$\xi_{\vec{r}_0}(\vec{r}) = \delta(\vec{r} - \vec{r}_0) \qquad (46)$$

$\{\xi_{\vec{r}_0}(\vec{r})\}$ represents the set of delta functions centered at each of the points $\vec{r}_0$ of space. Clearly, $\xi_{\vec{r}_0}(\vec{r})$ is not square integrable, so $\xi_{\vec{r}_0}(\vec{r}) \notin \mathcal{F}$. Then, consider the relations:

$$\psi(\vec{r}) = \int \psi(\vec{r}_0) \delta(\vec{r} - \vec{r}_0) \ d^3 r_0 = \int \psi(\vec{r}_0) \xi_{\vec{r}_0}(\vec{r}) \ d^3 r_0 \qquad (47)$$

$$\psi(\vec{r}_0) = (\xi_{\vec{r}_0}, \psi) = \int \delta(\vec{r}_0 - \vec{r}) \psi(\vec{r}) \ d^3 r = \int \xi_{\vec{r}_0}^*(\vec{r}) \psi(\vec{r}) \ d^3 r \qquad (48)$$

Equation 47 expresses the fact that every function $\psi(\vec{r}) \in \mathcal{F}$ can be expanded in terms of the $\xi_{\vec{r}_0}(\vec{r})$ functions in exactly one way. **Equation 48** shows that the value of the "continuous coefficient" corresponding to $\xi_{\vec{r}_0}(\vec{r})$ is precisely $\psi(\vec{r}_0)$.

The usefulness of the continuous bases that we have just introduced is revealed more clearly in what follows. However, we must not lose sight of the following point: a physical state must always correspond to a square-integrable wave function. In no case can $v_{\vec{p}}(\vec{r})$ or $\xi_{\vec{r}_0}(\vec{r})$ represent the state of a particle. These functions are nothing more than intermediaries, very useful in calculations involving operations on the wave functions $\psi(\vec{r})$ which are used to describe a physical state.[9]

1.9 Matrix formulation of quantum mechanics

1.9.1 Matrix representation in discrete bases

Recall that, for a basis $\{|\phi_n\rangle\}$ of $\mathcal{F}$, we can write any $\psi \in \mathcal{F}$ as a linear combination of the basis vectors:

$$|\psi\rangle = \sum_n c_n |\phi_n\rangle \qquad (49)$$

---

[9]aclarar

where $c_n = \langle\phi_n|\psi\rangle$ represents the projection of $|\psi\rangle$ onto $|\phi_n\rangle$. So, within the basis $\{|\phi_n\rangle\}$, the ket $|\psi\rangle$ is represented by the set of its components, $c_1, c_2, ...$, along $|\phi_1\rangle, |\phi_2\rangle, ...$, respectively. Hence, we can write the ket $|\psi\rangle$ as a column vector:

$$|\psi\rangle \rightarrow \begin{pmatrix} \langle\phi_1|\psi\rangle \\ \langle\phi_2|\psi\rangle \\ \vdots \\ \langle\phi_n|\psi\rangle \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \tag{50}$$

The bra $\langle\psi|$ is represented by the row vector:

$$\langle\psi| \rightarrow \begin{pmatrix} \langle\psi|\phi_1\rangle & \langle\psi|\phi_2\rangle & \dots & \langle\psi|\phi_n\rangle \end{pmatrix} = \begin{pmatrix} \langle\phi_1|\psi\rangle^* & \langle\phi_2|\psi\rangle^* & \dots & \langle\phi_n|\psi\rangle^* \end{pmatrix} = \begin{pmatrix} c_1^* & c_2^* & \dots & c_n^* \end{pmatrix} \tag{51}$$

Just as kets and bras are respresented by column and row vectors, respectively, operators are represented by square matrices.

1.10 Eigenvalues and eigenvectors of an operator

We can define the **eigenvectors of an operator** as:

Definition 12. A state vector $|\psi\rangle$ is said to be an eigenvector (also called eigenket or eigenvector) of an operator A if it is a solution of the eigenvalue equation:

$$A|\psi\rangle = a|\psi\rangle \tag{52}$$

where a is a complex number, called an eigenvalue of A.

Some theorems regarding eigenvectors and eigenvalues are:

Theorem 2. The eigenvalues of the inverse A^{-1} of an operator A are the inverse (with respect to the multiplication, $1/a$) of the eigenvalues a of A. *Proof:*

$$A|\psi\rangle = a|\psi\rangle \rightarrow A^{-1}A|\psi\rangle = aA^{-1}|\psi\rangle \rightarrow |\psi\rangle = aA^{-1}|\psi\rangle \rightarrow A^{-1}|\psi\rangle = \frac{1}{a}|\psi\rangle \tag{53}$$

Theorem 3. For a Hermitian operator A, all of its eigenvalues are real and the eigenvectors corresponding to different eigenvalues are orthogonal[a].

---
[a]Proof in **Section 9.8**.

Theorem 4. If two Hermitian operators, A and B, commute and if A has no degenerate eigenvalue, then each eigenvector of A is also an eigenvector of B. In addition, we can construct a common orthonormal basis that is made of the joint eigenvectors of A and B[a].

---
[a]Proof in **Section 9.9**.

1.11 Position representation of state space

In the position representation, the basis consists of an infinite set of vectors $\{|\vec{r}\rangle\}$, which are eigenkets of the position operator $\vec{R}$:

$$\vec{R}|\vec{r}\rangle = \vec{r}|\vec{r}\rangle \tag{54}$$

where $\vec{r}$, the position vector, is the eigenvalue of the position operator $\vec{R}$. The orthonormality and completeness relations are given by[10]:

$$\langle \vec{r}|\vec{r}\,'\rangle = \delta(\vec{r} - \vec{r}\,') \tag{55}$$

$$\int |\vec{r}\rangle \langle \vec{r}| \, d^3r = \mathbb{1} \tag{56}$$

And every state vector $|\psi\rangle$ can be expanded in terms of the position eigenkets $|\vec{r}\rangle$ as:

$$|\psi\rangle = \int \langle \vec{r}|\psi\rangle \, |\vec{r}\rangle \, d^3r = \int \psi(\vec{r}) \, |\vec{r}\rangle \, d^3r \tag{57}$$

where the **wave function** $\psi(\vec{r}) = \langle \vec{r}|\psi\rangle$ denotes the components of ψ in the $\{|\vec{r}\rangle\}$ basis. The quantity $|\langle \vec{r}|\psi\rangle|^2 d^3r = |\psi(\vec{r})|^2 d^3r$ represents the probability of finding the system inside the volume element d^3r.

The scalar product between two state vectors $|\psi\rangle$ and $|\phi\rangle$ can be written as:

$$\langle \phi|\psi\rangle = \langle \phi| \left(\int |\vec{r}\rangle \langle \vec{r}| \, d^3r \right) |\psi\rangle = \int \langle \vec{r}|\phi\rangle^* \langle \vec{r}|\psi\rangle \, d^3r = \int \phi^*(\vec{r})\psi(\vec{r}) d^3r \tag{58}$$

Since $\vec{R}\,|\vec{r}\rangle = \vec{r}\,|\vec{r}\rangle$, we can write:

$$\vec{R}^n \, |\vec{r}\rangle = \vec{r}^{\,n} \, |\vec{r}\rangle \rightarrow \langle \vec{r}\,'| \, \vec{R}^n \, |\vec{r}\rangle = \langle \vec{r}\,'| \, \vec{r}^{\,n} \, |\vec{r}\rangle \rightarrow \langle \vec{r}\,'| \, \vec{R}^n \, |\vec{r}\rangle = \vec{r}^{\,n} \delta(\vec{r} - \vec{r}\,') \tag{59}$$

Note that the operator $\vec{R}$ is Hermitian, since:

$$\langle \phi|\vec{R}\,|\psi\rangle = \int \langle \vec{r}|\phi\rangle^* \, \vec{r} \, \langle \vec{r}|\phi\rangle = \left[\int \langle \vec{r}|\phi\rangle \, \vec{r} \, \langle \vec{r}|\phi\rangle^* \right]^* = \langle \psi|\vec{R}\,|\phi\rangle^* \tag{60}$$

1.12   Momentum representation of state space

The basis $\{\vec{p}\}$ of the momentum representation is made of the eigenkets of the momentum operator $\vec{P}$:

$$\vec{P}\,|\vec{p}\rangle = \vec{p}\,|\vec{p}\rangle \tag{61}$$

where $\vec{p}$, the momentum vector, is the eigenvalue of the momentum operator $\vec{P}$. The orthonormality and completeness relations are given by[11]:

$$\langle \vec{p}|\vec{p}\,'\rangle = \delta(\vec{p} - \vec{p}\,') \tag{62}$$

$$\int |\vec{p}\rangle \langle \vec{p}| \, d^3p = \mathbb{1} \tag{63}$$

And every state vector $|\psi\rangle$ can be expanded in terms of the momentum eigenkets $|\vec{p}\rangle$ as:

$$|\psi\rangle = \int \langle \vec{p}|\psi\rangle \, |\vec{p}\rangle \, d^3p = \int \Psi(\vec{p}) \, |\vec{p}\rangle \, d^3p \tag{64}$$

where the expansion coefficient $\Psi(\vec{p}) = \langle \vec{p}|\psi\rangle$ is the **momentum space wave function**. The quantity $|\langle \vec{p}|\psi\rangle|^2 d^3p = |\Psi(\vec{p})|^2 d^3p$ represents the probability of finding the system's momentum inside the volume element d^3p located between $\vec{p}$ and $\vec{p} + d\vec{p}$.

The scalar product between two state vectors $|\psi\rangle$ and $|\phi\rangle$ can be written as:

$$\langle \phi|\psi\rangle = \langle \phi| \left(\int |\vec{p}\rangle \langle \vec{p}| \, d^3p \right) |\psi\rangle = \int \langle \vec{p}|\phi\rangle^* \langle \vec{p}|\psi\rangle \, d^3p = \int \Phi^*(\vec{p})\Psi(\vec{p}) d^3p \tag{65}$$

Since $\vec{P}\,|\vec{p}\rangle = \vec{p}\,|\vec{p}\rangle$, we can write:

$$\vec{P}^n \, |\vec{p}\rangle = \vec{p}^{\,n} \, |\vec{p}\rangle \rightarrow \langle \vec{p}\,'| \, \vec{P}^n \, |\vec{p}\rangle = \langle \vec{p}\,'| \, \vec{p}^{\,n} \, |\vec{p}\rangle \rightarrow \langle \vec{p}\,'| \, \vec{P}^n \, |\vec{p}\rangle = \vec{p}^{\,n} \delta(\vec{p} - \vec{p}\,') \tag{66}$$

[10]Remember that elements of an uncountable set of functions such as this *cannot* be elements of $\mathcal{F}$.

[11]Remember that, as in the position representation, elements of an uncountable set of functions such as this *cannot* be elements of $\mathcal{F}$.

1.13   Connecting position and momentum representations

When changing from the $\{\vec{r}\}$ basis to the $\{\vec{p}\}$ basis, we encounter a transformation function $\langle \vec{r}|\vec{p}\rangle$. To find the expression for this transformation, let us establish the following relations:

$$\langle \vec{r}|\psi\rangle = \langle \vec{r}| \left(\int |\vec{p}\rangle \langle \vec{p}| d^3p \right) |\psi\rangle = \int \langle \vec{r}|\vec{p}\rangle \langle \vec{p}|\psi\rangle \ d^3p = \int \langle \vec{r}|\vec{p}\rangle \Psi(\vec{p}) \ d^3p \tag{67}$$

which, as $\psi(\vec{r}) = \langle \vec{r}|\psi\rangle$, means:

$$\psi(\vec{r}) = \int \langle \vec{r}|\vec{p}\rangle \Psi(\vec{p}) \ d^3p \tag{68}$$

Similarly, we find:

$$\Psi(\vec{p}) = \langle \vec{p}|\psi\rangle = \langle \vec{p}| \left(\int |\vec{r}\rangle \langle \vec{r}| d^3r \right) |\psi\rangle = \int \langle \vec{p}|\vec{r}\rangle \langle \vec{r}|\psi\rangle d^3r = \int \langle \vec{p}|\vec{r}\rangle \psi(\vec{r}) \ d^3r \tag{69}$$

$$\Psi(\vec{p}) = \int \langle \vec{p}|\vec{r}\rangle \psi(\vec{r}) \ d^3r \tag{70}$$

Equation 68 and **Equation 70** imply that $\phi(\vec{r})$ and $\Phi(\vec{p})$ are Fourier transforms of each other. In quantum mechanics, the Fourier transform of a function $f(\vec{r})$ is given as:

$$f(\vec{r}) = \frac{1}{(2\pi\hbar)^{3/2}} \int e^{i\vec{p}\cdot\vec{r}/\hbar} g(\vec{p}) \ d^3p \tag{71}$$

Hence, the function $\langle \vec{r}|\vec{p}\rangle$ is given by:

$$\langle \vec{r}|\vec{p}\rangle = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\vec{p}\cdot\vec{r}/\hbar} \tag{72}$$

This function transforms from the momentum to the position representation. The function corresponding to the inverse transformation, $\langle \vec{p}|\vec{r}\rangle$, is given by:

$$\langle \vec{p}|\vec{r}\rangle = \langle \vec{r}|\vec{p}\rangle^* = \frac{1}{(2\pi\hbar)^{3/2}} e^{-i\vec{p}\cdot\vec{r}/\hbar} \tag{73}$$

The quantity $|\langle \vec{r}|\vec{p}\rangle|^2 = \frac{1}{(2\pi\hbar)^3}$ represents the probability of finding the system at the point $\vec{r}$ when its momentum is $\vec{p}$.

1.13.1   Momentum operator in the position representation

To determine the expression for the momentum operator $\vec{P}$ in the position representation, let us calculate:

$$\langle \vec{r}|\vec{P}|\psi\rangle = \int \langle \vec{r}|\vec{P}|\psi\rangle d^3p \stackrel{12}{=} \int \langle \vec{r}|\vec{P}|\vec{p}\rangle \langle \vec{p}|\psi\rangle d^3p \stackrel{13}{=} \int \vec{p}\langle \vec{r}|\vec{p}\rangle \langle \vec{p}|\psi\rangle d^3p \stackrel{14}{=} \frac{1}{(2\pi\hbar)^{3/2}} \int \vec{p} e^{i\vec{p}\cdot\vec{r}/\hbar} \Psi(\vec{p}) d^3p \quad (74)$$

Now, since $\vec{p}e^{i\vec{p}\cdot\vec{r}/\hbar} = -i\hbar\vec{\nabla} e^{i\vec{p}\cdot\vec{r}/\hbar}$, and using **Equation 72**, we can write:

$$\langle \vec{r}|\vec{P}|\psi\rangle = -i\hbar\vec{\nabla} \left(\frac{1}{(2\pi\hbar)^{3/2}} \int e^{i\vec{p}\cdot\vec{r}/\hbar} \Psi(\vec{p}) d^3p \right) = -i\hbar\vec{\nabla} \left(\int \langle \vec{r}|\vec{p}\rangle \langle \vec{p}|\psi\rangle d^3p \right) = -i\hbar\vec{\nabla} \langle \vec{r}|\psi\rangle \tag{75}$$

So, the momentum operator acting on the state $|\psi\rangle$ is expressed in the position representation as the differential operator:

$$\vec{P} = -i\hbar\vec{\nabla} \tag{76}$$

acting on the wave function $\psi(\vec{r}) = \langle \vec{r}|\psi\rangle$. Thus, the momentum operator is represented by the differential operator $-i\hbar\vec{\nabla}$ in the position representation.

---

[12]Here, we apply the closure relation.

[13]Here, we have used that $\vec{P}|\psi\rangle = \vec{p}|\vec{p}\rangle$.

[14]Here, we have used the definition of the transformation function from the momentum to the position representation, defined in **Equation 72**.

1.13.2 Canonical commutation relations

Calculating the separate terms of the x-component commutator of the operators $\vec{R}$ and $\vec{P}$ in the position representation, we find:

$$XP_x\psi(\vec{r}) = -i\hbar x\frac{\partial\psi(\vec{r})}{\partial x} \tag{77}$$

$$P_xX\psi(\vec{r}) = -i\hbar\frac{\partial}{\partial x}(x\psi(\vec{r})) = -i\hbar\psi(\vec{r}) - i\hbar x\frac{\partial\psi(\vec{r})}{\partial x} \tag{78}$$

so that:

$$[X,P_x]\psi(\vec{r}) = XP_x\psi(\vec{r}) - P_xX\psi(\vec{r}) = -i\hbar x\frac{\partial\psi(\vec{r})}{\partial x} + i\hbar\psi(\vec{r}) + i\hbar x\frac{\partial\psi(\vec{r})}{\partial x} = i\hbar\psi(\vec{r}) \tag{79}$$

Therefore:

$$[X,P_x] = i\hbar \tag{80}$$

Likewise, it can be shown for the y and z components:

$$[Y,P_y] = i\hbar, \qquad [Z,P_z] = i\hbar \tag{81}$$

We can also check that, crossing different components:

$$[X,P_y] = [X,P_z] = [Y,P_x] = [Y,P_z] = [Z,P_x] = [Z,P_y] = 0 \tag{82}$$

With this, and the fact that the three degrees of freedom are independent[15], we arrive at the **canonical commutation relations**[16]:

$$[X_j,P_k] = i\hbar\delta_{jk}, \quad [X_j,X_k] = [P_j,P_k] = 0, \quad j,k = x,y,z \tag{83}$$

It is important to note that, even though the particular expression of an operator in different representations may vary, the commutation relations for operators are representation independent[17].

1.14 Tensor product of two vector spaces

Up to this point, we have always talked about the wave functions of a single particle. However, even in that case, one can consider a one-dimensional, two-dimensional or even three-dimensional wave function. It is clear that the particularisation of the general state space $\mathcal{F}$ is not the same for functions of one variables (we will call the 1D state space ε_x) as for functions of three variables (we will call the 3D state space, $\varepsilon_{\vec{r}}$). $\varepsilon_{\vec{r}}$ and ε_x are therefore different spaces. Nevertheless, $\varepsilon_{\vec{r}}$ appears to be essentially a generalisation of ε_x. Does there exist a more precise relation between these two spaces?

In this section, we are going to define and study the operation of taking the tensor product of vector spaces, and apply it to state spaces. This will answer, in particular, the question we have just asked: $\varepsilon_{\vec{r}}$ can be constructed from ε_x and another two spaces, ε_y and ε_z, which are isomorphic to it[18].

We give the following definition of the tensor product of two vector spaces:

Definition 13. Let ε_1 and ε_2 be two[a] vector spaces, of (finite or infinite) dimension N_1 and N_2, respectively. Let vectors and operators belonging to ε_1 and ε_2 be denoted by an index (1) or (2), respectively. The vector space $\varepsilon = \varepsilon_1 \otimes \varepsilon_2$ is called the *tensor product of ε_1 and ε_2* if there is associated with each pair of vectors, $|\psi(1)\rangle$ belonging to ε_1 and $|\psi(2)\rangle$ belonging to ε_2, a vector belonging to ε, denoted by[b]:

$$|\psi(1)\rangle \otimes |\psi(2)\rangle = |\psi(1)\psi(2)\rangle \tag{84}$$

---

[15]See the **Comment on the three-dimensional state space** in **Section 1.17**.
[16]For simplicity of notation, we use $X_x = X$, $X_y = Y$ and $X_z = Z$.
[17]Add proof
[18]clarify what this means

which is called the tensor product of $|\psi(1)\rangle$ and $|\psi(2)\rangle$, and which satisfies the following conditions:

- It is *linear* with respect to multiplication by complex numbers:

$$\alpha\left[|\psi(1)\rangle \otimes |\psi(2)\rangle\right] = \left[\alpha|\psi(1)\rangle\right] \otimes |\psi(2)\rangle = |\psi(1)\rangle \otimes (\alpha|\psi(2)\rangle) \tag{85}$$

- It is *distributive* with respect to vector addition:

$$|\psi(1)\rangle \otimes \left[|\chi_1(2)\rangle + |\chi_2(2)\rangle\right] = |\psi(1)\rangle \otimes |\chi_1(2)\rangle + |\psi(1)\rangle \otimes |\chi_2(2)\rangle$$
$$\left[|\psi_1(1)\rangle + |\psi_2(1)\rangle\right] \otimes |\chi(2)\rangle = |\psi_1(1)\rangle \otimes |\chi(2)\rangle + |\psi_2(1)\rangle \otimes |\chi(2)\rangle \tag{86}$$

- When a basis has been chosen for ε_1, $\{u_i(1)\}$, and for ε_2, $\{u_j(2)\}$, the set of vectors $|u_i(1)\rangle \otimes |v_l(2)\rangle$ constitutes a basis in ε. If N_1 and N_2 are finite, the dimensions of ε is $N_1 N_2$.

---
[a]This definition can easily be extended to a finite number of vector spaces.
[b]This vector can be written as $|\psi(1)\rangle \otimes |\psi(2)\rangle$ or $|\psi(2)\rangle \otimes |\psi(1)\rangle$; the order is of no importance.

1.15 Vectors in the tensor product space

Let us consider a tensor product $\varepsilon = \varepsilon_1 \otimes \varepsilon_2$. Let us also consider a vector $|\psi(1)\rangle \otimes |\chi(2)\rangle$ in ε. Whatever $|\psi(1)\rangle$ and $|\chi(2)\rangle$ may be, they can be expressed in the basis of their respective spaces, $\{|u_i(1)\rangle\}$ and $\{|v_l(2)\rangle\}$, as:

$$|\psi(1)\rangle = \sum_i a_i |u_i(1)\rangle, \qquad |\chi(2)\rangle = \sum_l b_l |v_l(2)\rangle \tag{87}$$

The expansion of the vector $|\psi(1)\rangle \otimes |\chi(2)\rangle$ in the basis $\{|u_i(1)\rangle \otimes |v_l(2)\rangle\}$ of ε is then given by:

$$|\psi(1)\rangle \otimes |\chi(2)\rangle = \sum_{i,l} a_i b_l |u_i(1)\rangle \otimes |v_l(2)\rangle \tag{88}$$

Therefore, we obtain the following result:

Definition 14. The components of a tensor product vector $|\psi(1)\rangle \otimes |\chi(2)\rangle$ are the products of the components of the two vectors of the product $|\psi(1)\rangle$ and $|\chi(2)\rangle$.

Note that a general vector in the tensor product space ε can be written as:

$$|\psi\rangle = \sum_{i,l} c_{i,l} |u_i(1)\rangle \otimes |v_l(2)\rangle \tag{89}$$

Thus, by comparing this expression with **Equation 88**, we can see that there exist vectors in ε that cannot be written as a tensor product of two vectors belonging to ε_1 and ε_2. This is because, in general, there are many cases where $c_{i,j}$ cannot be written as a product of two numbers a_i and b_j. However, any arbitrary vector of ε can always be decomposed into a *linear combination* of tensor product vectors belonging to ε_1 and ε_2.

1.16 Scalar product in the tensor product space

The existence of the scalar product in ε_1 and ε_2 allows us to define one in ε too.

Definition 15. The scalar product of $|\psi(1)\chi(2)\rangle = |\psi(1)\rangle \otimes |\chi(2)\rangle$ and $|\psi'(1)\chi'(2)\rangle = |\psi'(1)\rangle \otimes |\chi'(2)\rangle$ is defined as:

$$\langle \psi'(1)\chi'(2)|\psi(1)\chi(2)\rangle = \langle \psi'(1)|\psi(1)\rangle \langle \chi'(2)|\chi(2)\rangle \tag{90}$$

Notice, in particular, that the basis $\{|u_i(1)v_l(2)\rangle = |u_i(1)\rangle \otimes |v_l(2)\rangle\}$ of ε is orthonormal if each of the bases of ε_1 and ε_2 is orthonormal:

$$\langle u_{i'}(1)v_{l'}(2)|u_i(1)v_l(2)\rangle = \langle u_{i'}(1)|u_i(1)\rangle \langle v_{l'}(2)|v_l(2)\rangle = \delta_{ii'}\delta_{ll'} \tag{91}$$

1.17 Tensor product of operators

We define operators inside the tensor product space in the following way:

Definition 16. Consider a linear operator $A(1)$ defined in ε_1. We associate it with a linear operator $\tilde{A}(1)$ defined in ε, which we call the extension of $A(1)$ in ε, and which is characterised in the following way: when $\tilde{A}(1)$ acts on a vector $|\psi(1)\rangle \otimes |\chi(2)\rangle$ of ε, one obtains, by definition:

$$\tilde{A}(1)\left[|\psi(1)\rangle \otimes |\chi(2)\rangle\right] = [A(1)|\psi(1)\rangle] \otimes |\chi(2)\rangle \tag{92}$$

We obtain an analogous definition for an operator $\tilde{B}(2)$ initially defined in ε_2:

$$\tilde{B}(2)\left[|\psi(1)\rangle \otimes |\chi(2)\rangle\right] = |\psi(1)\rangle \otimes [B(2)|\chi(2)\rangle] \tag{93}$$

Then, the action of $\tilde{A}(1)$ on an arbitrary vector of ε is given by:

$$\tilde{A}(1)|\psi\rangle = \tilde{A}(1)\left(\sum_{i,l} c_{i,l}|u_i(1)\rangle \otimes |v_l(2)\rangle\right) = \sum_{i,l} c_{i,l}[A(1)|u_i(1)\rangle] \otimes |v_l(2)\rangle \tag{94}$$

and the action of $\tilde{B}(2)$ is given by:

$$\tilde{B}(2)|\psi\rangle = \tilde{B}(2)\left(\sum_{i,l} c_{i,l}|u_i(1)\rangle \otimes |v_l(2)\rangle\right) = \sum_{i,l} c_{i,l}|u_i(1)\rangle \otimes [B(2)|v_l(2)\rangle] \tag{95}$$

Therefore, we obtain:

Definition 17. Let $A(1)$ and $B(2)$ be two linear operators acting, respectively, in ε_1 and ε_2. Their tensor product $A(1) \otimes B(2)$ is the linear operator acting in ε, defined by the following relation which describes its action on the tensor product vectors:

$$[A(1) \otimes B(2)]\left[|\psi(1)\rangle \otimes |\chi(2)\rangle\right] = [A(1)|\psi(1)\rangle] \otimes [B(2)|\chi(2)\rangle] \tag{96}$$

Comment on extension of operators

The extensions of operators are special cases of tensor products. If $\mathbb{1}(1)$ and $\mathbb{1}(2)$ are the identity operators in ε_1 and ε_2, respectively, then:

$$\begin{aligned} \tilde{A}(1) &= A(1) \otimes \mathbb{1}(2) \\ \tilde{B}(2) &= \mathbb{1}(1) \otimes B(2) \end{aligned} \tag{97}$$

By combining these expressions with **Definition 17**, it is now very easy to deduce the expressions in **Definition 16**.

It is also easy to show that any pair of such operators commute in ε, by looking at the action of $\tilde{A}(1)\tilde{B}(2)$ and $\tilde{B}(2)\tilde{A}(1)$ on an arbitrary vector of the basis of ε:

$$
\left.
\begin{aligned}
\tilde{A}(1)\tilde{B}(2) \,|u_i(1)\rangle \otimes |v_l(2)\rangle &= \tilde{A}(1) \left[|u_i(1)\rangle \otimes [B(2)\,|v_l(2)\rangle]\right] = \\
&= [A(1)\,|u_i(1)\rangle] \otimes [B(2)\,|v_l(2)\rangle] \\
\tilde{B}(2)\tilde{A}(1) \,|u_i(1)\rangle \otimes |v_l(2)\rangle &= \tilde{B}(2) \left[[A(1)\,|u_i(1)\rangle] \otimes |v_l(2)\rangle\right] = \\
&= [A(1)\,|u_i(1)\rangle] \otimes [B(2)\,|v_l(2)\rangle]
\end{aligned}
\right\} \longrightarrow \tilde{A}(1)\tilde{B}(2) = \tilde{B}(2)\tilde{A}(1) \quad (98)
$$

As the result of the action of both operator products on the basis vectors is the same, the operators commute:

$$[\tilde{A}(1), \tilde{B}(2)] = 0 \tag{99}$$

Comment on the three-dimensional state space

Going back to what we said earlier about the state space $\varepsilon_{\vec{r}}$ being a generalisation of ε_x, we can now see that this is true. In fact, we can write:

$$\varepsilon_{\vec{r}} = \varepsilon_x \otimes \varepsilon_y \otimes \varepsilon_z \tag{100}$$

which means that every vector in $\varepsilon_{\vec{r}}$ is the tensor product of three vectors: one in each of those three spaces. We can simplify the notation by writing:

$$|x, y, z\rangle = |x\rangle \otimes |y\rangle \otimes |z\rangle \tag{101}$$

Then, we can also see that the operators X, Y and Z can actually be extended in $\varepsilon_{\vec{r}}$ as:

$$\tilde{X} = X \otimes \mathbb{1} \otimes \mathbb{1}, \qquad \tilde{Y} = \mathbb{1} \otimes Y \otimes \mathbb{1}, \qquad \tilde{Z} = \mathbb{1} \otimes \mathbb{1} \otimes Z \tag{102}$$

We usually forget the tilde when we are working in $\varepsilon_{\vec{r}}$, and write directly X, Y and Z, for simplicity, but in order to understand fully what we are doing, we must remember the expressions in **Equation 102**. These give us new insight into the nature of these operators. X, Y and Z commute because, as we can clearly see from these results, they actually act independently on different state spaces.

In fact, as we proved in the previous *comment on extension of operators*, any pair of operators $\tilde{A}(1)$ and $\tilde{B}(2)$ in ε, that originally come from their respective operators $A(1)$ and $B(2)$ acting on different spaces ε_1 and ε_2, commute in $\varepsilon = \varepsilon_1 \otimes \varepsilon_2$. If we look back at **Theorem 4**, we can provide further proof of this commutation relation between $\tilde{A}(1)$ and $\tilde{B}(2)$ by showing that they always have common eigenvectors: let $|\psi_n(1)\rangle$ be an eigenvector of $A(1)$ in ε_1 and $|\chi_n(2)\rangle$ an eigenvector of $B(2)$ in ε_2, with eigenvalues $\lambda_n(1)$ and $\lambda_n(2)$ respectively. Then:

$$
\begin{aligned}
\tilde{A}(1) \left[|\psi_n(1)\rangle \otimes |\chi_n(2)\rangle\right] &= [A(1)\,|\psi_n(1)\rangle] \otimes |\chi_n(2)\rangle = \\
&= [\lambda_n(1)\,|\psi_n(1)\rangle] \otimes |\chi_n(2)\rangle = \\
&= \lambda_n(1) \left[|\psi_n(1)\rangle \otimes |\chi_n(2)\rangle\right]
\end{aligned}
\tag{103}
$$

so that $|\psi_n(1)\rangle \otimes |\chi_n(2)\rangle$ is an eigenvector of $\tilde{A}(1)$ with eigenvalue $\lambda_n(1)$. But we can also find the same relation for $\tilde{B}(2)$:

$$
\begin{aligned}
\tilde{B}(2) \left[|\psi_n(1)\rangle \otimes |\chi_n(2)\rangle\right] &= |\psi_n(1)\rangle \otimes [B(2)\,|\chi_n(2)\rangle] = \\
&= |\psi_n(1)\rangle \otimes [\lambda_n(2)\,|\chi_n(2)\rangle] = \\
&= \lambda_n(2) \left[|\psi_n(1)\rangle \otimes |\chi_n(2)\rangle\right]
\end{aligned}
\tag{104}
$$

so that $|\psi_n(1)\rangle \otimes |\chi_n(2)\rangle$ is also an eigenvector of $\tilde{B}(2)$ with eigenvalue $\lambda_n(2)$. Thus, $\tilde{A}(1)$ and $\tilde{B}(2)$ have common eigenvectors, and therefore commute.

Note: this does *not* mean that any pair of commuting operators, which have common eigenvectors, can be extended in a tensor product space. For example, as we will see in coming sections, the operators $\vec{L}^{\,2}$ and L_z commute, and have common eigenvectors, but they cannot be extended in a separate tensor product space, because they both act on the same space $\varepsilon_{\vec{r}}$.

2 Postulates of Quantum Mechanics

2.1 Introduction

Quantum mechanics is based on a number of postulates, which themselves are based on a range of experimental observations. **These postulates cannot be derived**, and they result exclusively from experiment. They are the minimal set of assumptions that one needs in order to build the theory of quantum mechanics. **The validity of the postulates of quantum mechanics can only be determined inferentially**: the theory works extremely well, so the postulates must also be valid. This represents its experimental justification.

The first four postulates concern the state of a system at a give time, whereas the last postulate concerns the time evolution of the system.

> **Postulate 1. State of a System:** The state of any physical system is specified, at each time t, by a state vector $|\psi(t)\rangle$ in a Hilbert space $\mathcal{H}$; $|\psi(t)\rangle$ contains (and serves as the basis to extract) all the needed information about the system. Any superposition of state vectors is also a state vector.

> **Postulate 2. Observables and Operators:** To every physical quantity a, called an observable or dynamical variable, there corresponds a linear Hermitian operator A whose eigenvectors form a complete basis.

> **Postulate 3. Measurements and Eigenvalues of Operators:** The measurement of an observable a may be represented formally by the action of its operator A on a state vector $|\psi(t)\rangle$. The only possible result of such a measurement is one of the eigenvalues a_n (which are real) of the operator A. If the result of a measurement of A on a state $|\psi(t)\rangle$ is a_n, the state of the system *immediately after* the measurement changes to $|\psi_n\rangle$.

> **Postulate 4. Probabilistic Outcome of Measurements:**
>
> - **Discrete Spectra:** When measuring an observable a of a system in a state $|\psi\rangle$, the probability of obtaining one of the non-degenerate eigenvalues a_n of the corresponding operator A is given by:
> $$P_n(a_n) = \frac{|\langle \psi_n|\psi\rangle|^2}{\langle \psi|\psi\rangle} = |a_n|^2 \tag{105}$$
> where $|\psi_n\rangle$ is the eigenstate of A with the eigenvalue a_n and we have assumed that the state is normalised. If the eigenvalue a_n is m-degenerate, P_n becomes:
> $$P_n(a_n) = \frac{\sum_{j=1}^{m}|\langle \psi_n^j|\psi\rangle|^2}{\langle \psi|\psi\rangle} = \sum_{j=1}^{m}|a_n^{(j)}|^2 \tag{106}$$
> where, again, we assume normalised states.
>
> The act of measurement changes the state of the system from $|\psi\rangle$ to $|\psi_n\rangle$. If the system is already in an eigenstate $|\psi_n\rangle$ of A, a measurement of a yields with certainty the corresponding eigenvalue a_n: $A|\psi_n\rangle = a_n|\psi_n\rangle$.

The mean value of an observable for a state $|\psi\rangle$ can be then calculated as:

$$\langle a \rangle_\psi = \sum_n a_n P_n(a_n) = \sum_n a_n \sum_{j=1}^m |a_n^j|^2 = \sum_n a_n \sum_{j=1}^m \langle\psi|\psi_n^j\rangle \langle\psi_n^j|\psi\rangle =$$

$$= \sum_n \sum_{j=1}^m \langle\psi|a_n|\psi_n^j\rangle \langle\psi_n^j|\psi\rangle = \sum_n \sum_{j=1}^m \langle\psi|A|\psi_n^j\rangle \langle\psi_n^j|\psi\rangle = \langle\psi| A \left[\sum_n \sum_{j=1}^m |\psi_n^j\rangle \langle\psi_n^j| \right] |\psi\rangle =$$

$$= \langle\psi| A |\psi\rangle \tag{107}$$

- **Continuous Spectra:** we can extend the previous relations to continuous spectra to determine the probability density that a measurement of A yields a value between a and $a+da$ on a system which is initially in a state $|\psi\rangle$:

$$\frac{dP(a)}{da} = \frac{|\langle\psi_a|\psi\rangle|^2}{\langle\psi|\psi\rangle} = |\psi(a)|^2 \tag{108}$$

where, again, we assume that the state $|\psi\rangle$ is normalised.

The mean value of an observable for a state $|\psi\rangle$ can be then calculated as:

$$\langle a \rangle_\psi = \int a\, dP(a) = \int a|\psi(a)|^2\, da = \int a\, |\langle\psi_a|\psi\rangle|^2\, da = \int a \langle\psi|\psi_a\rangle \langle\psi_a|\psi\rangle\, da =$$

$$= \int \langle\psi| a |\psi_a\rangle \langle\psi_a|\psi\rangle\, da = \int \langle\psi| A |\psi_a\rangle \langle\psi_a|\psi\rangle\, da = \langle\psi| A \left(\int |\psi_a\rangle \langle\psi_a|\, da \right) |\psi\rangle$$

$$= \langle\psi|A|\psi\rangle \tag{109}$$

Postulate 5. Time Evolution of State Vectors: The state vector $|\psi(t)\rangle$ of a system evolves in time according to the Schrödinger equation:

$$i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = H |\psi(t)\rangle \tag{110}$$

where H is the Hamiltonian operator corresponding to the total energy of the system.

2.2 Observables and operators

An observable is a dynamic variable that can be measured. According to **Postulate 2**, a Hermitian operator is associated with every physical observable. In previous sections, we have seen that the position representation of the linear momentum operator is given by:

$$\vec{P} = -i\hbar \vec{\nabla} \tag{111}$$

In general, any function $f(\vec{r}, \vec{p})$ can be "quantised" (made into a function of operators) by replacing $\vec{r}$ and $\vec{p}$ with the position and momentum operators $\vec{R}$ and $\vec{P}$, respectively:

$$f(\vec{r}, \vec{p}) \to F(\vec{R}, \vec{P}) \tag{112}$$

Some examples of operators are shown in **Table 2.2**.

Observable	Corresponding Operator
$\vec{r}$	$\vec{R}$
$\vec{p}$	$\vec{P} = -\hbar\vec{\nabla}$
$T = \frac{p^2}{2m}$	$T = -\frac{\hbar^2}{2m}\nabla^2$
$E = \frac{p^2}{2m} + V(\vec{r}, t)$	$H = -\frac{\hbar^2}{2m}\nabla^2 + V(\vec{R}, t)$
$\vec{L} = \vec{r} \times \vec{p}$	$\vec{L} = -i\hbar\vec{R} \times \vec{\nabla}$

Table 1: Some observables and their corresponding operators.

2.2.1 Compatibility of observables

Two observables a and b are said to be **compatible** if their corresponding operators A and B commute. This means that there exists a common set of eigenvectors for both operators. In other words, if $|\psi_n\rangle$ is an eigenvector of A with eigenvalue a_n, then it is also an eigenvector of B with eigenvalue b_n. It is said that such pair A and B of observables form a **complete set of commuting observables** (CSCO).

Two observables can only be measured simultaneously if they are compatible, and this is a fact that is tightly connected to **Postulate 3**: when we measure an observable (we can think of this as applying its operator to the state), we obtain one of the eigenvalues of the operator, and immediately the system changes to the eigenstate corresponding to that eigenvalue. In order to be able to measure two different observables at the same time, the system must be able to change to an eigenstate that is common to *both* operators. This is only possible if the operators commute, because for two operators to share the basis of eigenvectors they *must* commute.

Furthermore, for commuting operators, it does not matter in what order we measure their corresponding observables. Measurement of one does not affect the measurement of the other. On the other hand, for non-commuting (incompatible) operators, measurement of one produces loss of information obtained from the measurement of the other.

2.3 Conservative systems

When the Hamiltonian of a system does not depend explicitly on time, the system is said to be **conservative**. In classical mechanics, the most important consequence of such a situation is the conservation of energy over time[19]. In other words, **the energy of the system is a constant of motion**. This also means that the eigenvalues and eigenvectors of the Hamiltonian operator are constant over time. Then, any state can be expanded in terms of the eigenvectors of the Hamiltonian operator:

$$|\psi(t)\rangle = \sum_n c_n(t) |\psi_n\rangle, \qquad c_n(t) = \langle\psi_n|\psi(t)\rangle \tag{113}$$

In this case, the time evolution of the state vector $|\psi(t)\rangle$ is given by the time-dependent Schrödinger equation:

$$i\hbar\frac{\partial \langle\psi_n|\psi(t)\rangle}{\partial t} = \langle\psi_n|H|\psi(t)\rangle \overset{20}{\Longrightarrow} i\hbar\frac{\partial c_n(t)}{\partial t} = E_n c_n(t) \tag{114}$$

which has the solution:

$$c_n(t) = c_n(t_0)e^{-iE_n(t-t_0)/\hbar} \tag{115}$$

Therefore:

$$|\psi(t)\rangle = \sum_n c_n(t_0)e^{-iE_n(t-t_0)/\hbar} |\psi_n\rangle, \qquad c_n(t_0) = \langle\psi_n|\psi(t_0)\rangle \tag{116}$$

---

[19]Actually, eigenvectors *can* change over time, but only by a phase.

[20]Here, we use the eigenvalue equation: $\langle\psi_n|H = E_n \langle\psi_n|$ so that $\langle\psi_n|H|\psi(t)\rangle = E_n \langle\psi_n|\psi(t)\rangle = E_n c_n(t)$.

and, for a continuous spectrum:

$$|\psi(t)\rangle = \int c(E, t_0) e^{-iE(t-t_0)/\hbar} |\psi_E\rangle \, dE, \qquad c(E, t_0) = \langle \psi_E | \psi(t_0) \rangle \tag{117}$$

2.3.1 Stationary states

A **stationary state** $|\psi(t)\rangle$ is a state that does not present any *observable*[21] change over time. They exist only in conservative systems, although not all states in a conservative system are stationary. For a stationary state, it is verified that the probability density function does not change over time[22], so:

$$\| |\psi(t)\rangle \|^2 = \| |\psi(t_0)\rangle \|^2 \implies |\psi(t)\rangle = |\psi(t_0)\rangle \, e^{i\alpha(t)} \tag{118}$$

for a certain real function $\alpha(t)$. Using the time-dependent Schrödinger equation:

$$i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = H |\psi(t)\rangle \tag{119}$$

we obtain[23]:

$$i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = i\hbar \frac{\partial \left(|\psi(t_0)\rangle \, e^{i\alpha(t)} \right)}{\partial t} = i\hbar |\psi(t_0)\rangle \frac{\partial (e^{i\alpha(t)})}{\partial t} = i\hbar |\psi(t_0)\rangle \, i\alpha'(t) e^{i\alpha(t)} =$$
$$= -\hbar \alpha'(t) |\psi(t)\rangle = H |\psi(t)\rangle \tag{120}$$

Which means that $|\psi(t)\rangle$ is an eigenstate of the Hamiltonian operator with eigenvalue $-\hbar\alpha'(t)$. As the system is conservative, the eigenvectors evolve only by a phase and the eigenvalues of H are constant, we have $|\psi(t)\rangle \to |\psi_n\rangle \, e^{i\alpha(t)}$ and $\alpha'(t) \to \alpha'$:

$$-\hbar\alpha' |\psi_n\rangle \, e^{i\alpha(t)} = H |\psi_n\rangle \, e^{i\alpha(t)} \tag{121}$$

It is now easy to see, looking at **Equation 116**, that:

$$|\psi(t)\rangle = |\psi_n\rangle \, e^{-iE_n(t-t_0)/\hbar} \tag{122}$$

Therefore, we can define:

> **Definition 18.** Stationary states are eigenstates of the Hamiltonian of a conservative system. They evolve in time only by a phase, which is to say that they do not present any observable change over time.

2.3.2 Constants of motion

A **constant of motion** a is an observable that does not depend explicitly on time, and whose operator A commutes with H:

$$\frac{dA}{dt} = 0, \qquad [H, A] = 0 \tag{123}$$

Then, by **Theorem 4**, there is always a system of common eigenvectors for H and A. These eigenvectors are stationary states, and their eigenvalues are called **good quantum numbers**.

---

[21] Meaning that any measurement of any observable always yields the same value.

[22] Remember that, in general, the shape of the probability density function can change, but the state *must* remain normalised. It is in the particular case of stationary states that the shape is maintained also.

[23] Note that $|\psi(t_0)\rangle$ is a constant, and can be taken out of the derivative.

2.4   Superposition and interference

Superposition in quantum mechanics does not work in the way we would expect from classical mechanics. Take the double slit experiment as an example. According to the classical superposition principle, the light intensity distribution on the screen is given by the sum of the intensities of the two waves coming from the two slits. In quantum mechanics, however, the light intensity distribution on the screen is proportional to the *square* of the sum of the amplitudes of the two waves coming from the two slits:

$$I(\vec{r}, t) \propto |\psi_1(\vec{r}, t) + \psi_2(\vec{r}, t)|^2 = |\psi_1(\vec{r}, t)|^2 + |\psi_2(\vec{r}, t)|^2 + \boxed{\psi_1^*(\vec{r}, t)\psi_2(\vec{r}, t) + \psi_1(\vec{r}, t)\psi_2^*(\vec{r}, t)} \tag{124}$$

This extra term that appears as a result of the interference of the partial amplitudes is the origin of the **interference** phenomena in quantum mechanics.

3 Angular Momentum

3.1 Orbital angular momentum

3.1.1 Classical orbital angular momentum

In classical mechanics, the angular momentum of a particle relative to some axis is defined as:

$$\vec{L} = \vec{r} \times \vec{p} \tag{125}$$

where $\vec{r}$ is the position vector of the particle with respect to a point on the axis of rotation and $\vec{p}$ is its momentum.

The total angular momentum of a system of particles is the sum of angular momenta of the individual particles:

$$\vec{L}_{\text{total}} = \sum_i \vec{r}_i \times \vec{p}_i \tag{126}$$

Figure 1: Classical orbital angular momentum

The total angular momentum varies in time according to the net external torque, which we can obtain by differentiating the total angular momentum with respect to time:

$$\vec{\tau}_{\text{total}} = \sum_i \vec{\tau}_{\text{ext},i} = \frac{d\vec{L}_{\text{total}}}{dt} \tag{127}$$

It follows that the total angular momentum of a system is conserved if the resultant external torque acting on the system is zero.

3.1.2 Orbital angular momentum in quantum mechanics

As discussed in **Section 2.2**, to obtain the quantum mechanical operator for orbital angular momentum from its classical definition, we can apply quantisation to the classical expression provided in the previous section:

$$\vec{L} = \vec{R} \times \vec{P} = -i\hbar \vec{R} \times \vec{\nabla} \tag{128}$$

For a system of (spin-less) particles, the total angular momentum is defined as:

$$\vec{L}_{\text{total}} = \sum_i \vec{R}_i \times \vec{P}_i \tag{129}$$

The different cartesian components of the angular momentum operator are:

$$
\begin{aligned}
L_x &= Y P_z - Z P_y = -i\hbar \left(Y \frac{\partial}{\partial z} - Z \frac{\partial}{\partial y} \right) \\
L_y &= Z P_x - X P_z = -i\hbar \left(Z \frac{\partial}{\partial x} - X \frac{\partial}{\partial z} \right) \\
L_z &= X P_y - Y P_x = -i\hbar \left(X \frac{\partial}{\partial y} - Y \frac{\partial}{\partial x} \right)
\end{aligned}
\tag{130}
$$

We can also define the square of the angular momentum operator:

$$L^2 = L_x^2 + L_y^2 + L_z^2 \tag{131}$$

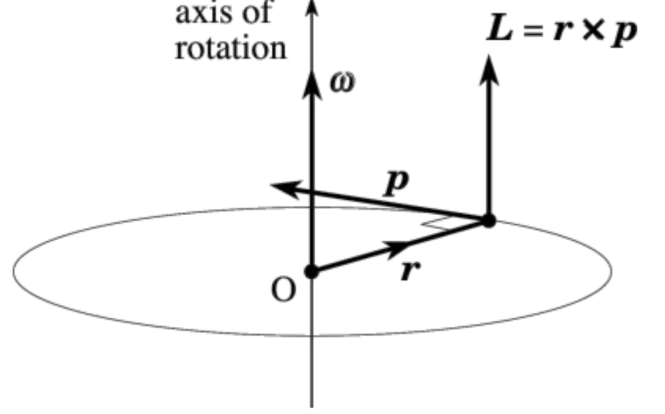As expected from operators corresponding to observables, all angular momentum operators are Hermitian.

3.1.3 Commutation relations

The commutation relations for the orbital angular momentum operators are:

$$
\begin{aligned}
[L_x, L_y] &= L_x L_y - L_y L_x = i\hbar L_z \\
[L_y, L_z] &= L_y L_z - L_z L_y = i\hbar L_x \\
[L_z, L_x] &= L_z L_x - L_x L_z = i\hbar L_y
\end{aligned}
\tag{132}
$$

3.2 General formalism of angular momentum

We have just defined the orbital angular momentum. However there exists a more general formalism of angular momentum, which is the *total* angular momentum. Its corresponding operator is $\vec{J}$. In the case that the only contribution to the *total* angular momentum is the *orbital* angular momentum, then $\vec{J} = \vec{L}$. Thus, all we will see in this section can be identically applied to orbital angular momentum[24].

$\vec{J}^2$ is defined by its three components that satisfy:

$$
\begin{aligned}
[J_x, J_y] &= i\hbar J_z \\
[J_y, J_z] &= i\hbar J_x \qquad \vec{J}^2 = J_x^2 + J_y^2 + J_z^2 \qquad [\vec{J}^2, J_k] = 0 \quad (k = x, y, z) \\
[J_z, J_x] &= i\hbar J_y
\end{aligned}
\tag{133}
$$

Thanks to these commutation relations, we know that we cannot measure the three components of the total angular momentum simultaneously. However, we *can* simultaneously measure the total angular momentum squared J^2 and one of its components. There is possibility to find simultaneous eigenstates of J^2 and any component of J. However, we can only choose one component of J to be measured simultaneously with J^2. By convention, we choose J_z, so that we work with a basis of eigenvectors that is common to J^2 and J_z in all our calculations[25].

Eigenstates of the total angular momentum operators

Let us now look for the joint eigenstates of J^2 and J_z and their corresponding eigenvalues. Denoting the joint eigenstates by $|\alpha, \beta\rangle$, and the corresponding eigenvalues of J^2 and J_z by $\hbar^2\alpha$ and $\hbar\beta$, respectively, we have:

$$
\begin{aligned}
J^2 |\alpha, \beta\rangle &= \hbar^2 \alpha |\alpha, \beta\rangle \\
J_z |\alpha, \beta\rangle &= \hbar\beta |\alpha, \beta\rangle
\end{aligned}
\tag{134}
$$

The factor $\hbar$ is introduced so that α and β are dimensionless. We also assume that the eigenstates are orthonormal.

Now we need to introduce *raising* and *lowering* operators J_+ and J_-, respectively, which are defined as:

$$
J_\pm = J_x \pm i J_y
\tag{135}
$$

This leads to:

$$
J_x = \frac{1}{2}(J_+ + J_-) \qquad J_y = \frac{1}{2i}(J_+ - J_-)
\tag{136}
$$

hence:

$$
J_x^2 = \frac{1}{4}(J_+^2 + J_+ J_- + J_- J_+ + J_-^2) \qquad J_y^2 = -\frac{1}{4}(J_+^2 - J_+ J_- - J_- J_+ + J_-^2)
\tag{137}
$$

Using the commutation relations from **Equation 133**, we can easily obtain:

$$
[\vec{J}^2, J_\pm] = 0 \qquad [J_+, J_-] = 2\hbar J_z \qquad [J_z, J_\pm] = \pm\hbar J_\pm
\tag{138}
$$

---

[24]In general, the total angular momentum is the sum of the *orbital* angular momentum and the *spin* angular momentum. The spin angular momentum is a purely quantum mechanical phenomenon, which we will discuss in **Section 3.4**.

[25]Note that this is just a convention. There is *nothing* special about the z direction compared to x and y.

And also:

$$J_+ J_- = J_x^2 + J_y^2 + \hbar J_z = \vec{J}^2 - J_z^2 + \hbar J_z$$
$$J_- J_+ = J_x^2 + J_y^2 - \hbar J_z = \vec{J}^2 - J_z^2 - \hbar J_z \tag{139}$$

These relations lead to:

$$\vec{J}^2 = J_\pm J_\mp + J_z^2 \mp \hbar J_z = \frac{1}{2}(J_+ J_- + J_- J_+) + J_z^2 \tag{140}$$

Since $J_\pm$ do not commute with J_z, the kets $|\alpha, \beta\rangle$ are not eigenstates of $J_\pm$. Using the expressions in **Equation 138**, we can obtain:

$$J_z(J_\pm |\alpha, \beta\rangle) = (J_\pm J_z \pm \hbar J_\pm) |\alpha, \beta\rangle = J_\pm J_z |\alpha, \beta\rangle \pm \hbar J_\pm |\alpha, \beta\rangle =$$
$$= \hbar \beta J_\pm |\alpha, \beta\rangle \pm \hbar J_\pm |\alpha, \beta\rangle = \hbar(\beta \pm 1)(J_\pm |\alpha, \beta\rangle) \tag{141}$$

hence the ket $J_\pm |\alpha, \beta\rangle$ is an eigenstate of J_z with eigenvalue $\hbar(\beta \pm 1)$. Since $\vec{J}^2$ commutes with J_z, $J_\pm |\alpha, \beta\rangle$ is also an eigenstate of $\vec{J}^2$. Using **Equation 138** again, which tells us that $\vec{J}^2$ commutes with $J_\pm$, we can determine the eigenvalue, which is $\hbar^2 \alpha$:

$$\vec{J}^2(J_\pm |\alpha, \beta\rangle) = J_\pm \vec{J}^2 |\alpha, \beta\rangle = J_\pm \hbar^2 \alpha |\alpha, \beta\rangle = \hbar^2 \alpha (J_\pm |\alpha, \beta\rangle) \tag{142}$$

If we rewrite **Equation 141** and **Equation 142** in terms of $|\alpha', \beta'\rangle = J_\pm |\alpha, \beta\rangle$:

$$J_z |\alpha', \beta'\rangle = \hbar(\beta \pm 1) |\alpha', \beta'\rangle$$
$$\vec{J}^2 |\alpha', \beta'\rangle = \hbar^2 \alpha |\alpha', \beta'\rangle \tag{143}$$

and if we compare with **Equation 134**, we can infer that $\alpha' = \alpha$ and $\beta' = \beta \pm 1$. In other words, the ket $J_\pm |\alpha, \beta\rangle$ is proportional to $|\alpha, \beta \pm 1\rangle$ (any eigenket multiplied by a constant is also an eigenket, although it may not be normalised), and we can write:

$$J_\pm |\alpha, \beta\rangle = C_{\alpha\beta}^\pm |\alpha, \beta \pm 1\rangle \tag{144}$$

So, when the operators $J_\pm$ act on a ket $|\alpha, \beta\rangle$, they do not change the first quantum number α, but they increase (or decrease) the second quantum number β by one unit. Hence the names *raising* and *lowering* operators.

Note that, for a given eigenvalue α of $\vec{J}^2$, there exists an upper limit for the *absolute value*[26] of the quantum number β. This is due to the fact that the operator $\vec{J}^2 - J_z^2$ is positive definite, as the matrix elements of $\vec{J}^2 - J_z^2 = J_x^2 + J_y^2 \geq 0$ are non-negative, so we can write:

$$\langle \alpha, \beta | \vec{J}^2 - J_z^2 | \alpha, \beta \rangle = \hbar^2(\alpha - \beta^2) \geq 0 \implies \alpha \geq \beta^2 \tag{145}$$

Since β has an upper limit, $\beta_{\max}$, there must exist a state $|\alpha, \beta_{\max}\rangle$ which cannot be raised further:

$$J_+ |\alpha, \beta_{\max}\rangle = 0 \tag{146}$$

Using this, along with **Equation 139**, we can obtain:

$$J_- J_+ |\alpha, \beta_{\max}\rangle = (\vec{J}^2 - J_z^2 - \hbar J_z) |\alpha, \beta_{\max}\rangle = \hbar^2(\alpha - \beta_{\max}^2 - \beta_{\max}) |\alpha, \beta_{\max}\rangle = 0 \tag{147}$$

hence:

$$\alpha = \beta_{\max}(\beta_{\max} + 1) \tag{148}$$

---

[26]That is to say, β is bounded from above *and* below.

Since β has a lower limit, $\beta_{\min}$, there must exist a state $|\alpha, \beta_{\min}\rangle$ which cannot be lowered further, which we reach after n successive applications of J_- on $|\alpha, \beta_{\max}\rangle$:

$$J_- |\alpha, \beta_{\min}\rangle = 0 \tag{149}$$

Using this, along with **Equation 139**, we can obtain:

$$J_+ J_- |\alpha, \beta_{\min}\rangle = (\vec{J}^2 - J_z^2 + \hbar J_z) |\alpha, \beta_{\min}\rangle = \hbar^2 (\alpha - \beta_{\min}^2 + \beta_{\min}) |\alpha, \beta_{\min}\rangle = 0 \tag{150}$$

hence:

$$\alpha = \beta_{\min}(\beta_{\min} - 1) \tag{151}$$

Comparing **Equation 149** and **Equation 151**, we can infer that:

$$\beta_{\max} = -\beta_{\min} \tag{152}$$

Since $\beta_{\min}$ was reached after n successive applications of J_- on $|\alpha, \beta_{\max}\rangle$, it follows that:

$$\beta_{\max} = \beta_{\min} + n, \qquad n \in \mathbb{N} \tag{153}$$

Combining the last two equations, we obtain:

$$\beta_{\min} = -\frac{n}{2} \qquad \beta_{\max} = \frac{n}{2} \tag{154}$$

Which means that $\beta_{\max}$ is either an integer or a half-odd-integer. We can now introduce the notation:

$$j = \beta_{\max} = \frac{n}{2} \qquad m = \beta \tag{155}$$

hence, we can express α as:

$$\alpha = j(j+1) \tag{156}$$

And we can infer that the values of m lie between $-j$ and j:

$$-j \leq m \leq j \tag{157}$$

We can now summarise the results we have obtained so far:

Definition 19. The eigenvalues of $\vec{J}^2$ and J_z corresponding to the joint eigenvectors $|j, m\rangle$ are given, respectively, by $\hbar^2 j(j+1)$ and $\hbar m$:

$$\vec{J}^2 |j, m\rangle = \hbar^2 j(j+1) |j, m\rangle \qquad J_z |j, m\rangle = \hbar m |j, m\rangle \tag{158}$$

where $j = 0, 1/2, 1, 3/2, ...$ and $m = -j, -(j-1), \ldots, j-1, j$. We see that the spectra of the angular momentum operators $\vec{J}^2$ and J_z are discrete. Since the eigenstates corresponding to different angular momenta are orthogonal, and since the angular momentum spectra are discrete, the orthonormality condition is:

$$\langle j', m' | j, m \rangle = \delta_{j',j} \delta_{m',m} \tag{159}$$

Let us now determine the normalisation constant $C_{\alpha\beta}^{\pm}$ in **Equation 144**, which we can rewrite in terms of j and m as:

$$J_\pm |j, m\rangle = C_{jm}^{\pm} |j, m \pm 1\rangle \tag{160}$$

Since $|j, m+1\rangle$ is normalised:

$$(J_+ |j, m\rangle)^{\dagger}(J_+ |j, m\rangle) = |C_{jm}^+|^2 \langle j, m+1 | j, m+1\rangle = |C_{jm}^+|^2 \tag{161}$$

Since $J_+ = J_x + iJ_y$ and the operators J_x and J_y are Hermitian, we have:

$$J_+^\dagger = (J_x + iJ_y)^\dagger = J_x^\dagger - iJ_y^\dagger = J_x - iJ_y = J_- \tag{162}$$

So we can also write:

$$\langle j, m | J_- J_+ | j, m \rangle = \langle j, m | J_+^\dagger J_+ | j, m \rangle = (J_+ | j, m \rangle)^\dagger (J_+ | j, m \rangle) = |C_{jm}^+|^2 \tag{163}$$

But since $J_- J_+ = \vec{J}^2 - J_z^2 - \hbar J_z$ and $|j, m\rangle$ is orthonormal, we can also write:

$$
\begin{aligned}
|C_{jm}^+|^2 &= \langle j, m | J_- J_+ | j, m \rangle = \langle j, m | \vec{J}^2 - J_z^2 - \hbar J_z | j, m \rangle = \\
&= \langle j, m | \vec{J}^2 | j, m \rangle - \langle j, m | J_z^2 | j, m \rangle - \hbar \langle j, m | J_z | j, m \rangle = \\
&= \langle j, m | \hbar^2 j(j+1) | j, m \rangle - \langle j, m | \hbar^2 m^2 | j, m \rangle - \hbar \langle j, m | \hbar m | j, m \rangle = \\
&= \hbar^2 j(j+1) \langle j, m | j, m \rangle - \hbar^2 m^2 \langle j, m | j, m \rangle - \hbar^2 m \langle j, m | j, m \rangle = \\
&= \hbar^2 j(j+1) - \hbar^2 m^2 - \hbar^2 m = \\
&= \hbar^2 (j(j+1) - m(m+1))
\end{aligned}
\tag{164}
$$

So we conclude that[27]:

$$C_{jm}^+ = \hbar \sqrt{j(j+1) - m(m+1)} \tag{165}$$

Similarly, we can obtain:

$$C_{jm}^- = \hbar \sqrt{j(j+1) - m(m-1)} \tag{166}$$

So, we now have the raising and lowering operators completely defined:

> **Definition 20.** The raising and lowering operators J_+ and J_- are given by:
>
> $$J_\pm = J_x \pm iJ_y \tag{167}$$
>
> Their action on a ket $|j, m\rangle$ raises or lowers the second quantum number m by one unit, respectively:
>
> $$J_\pm |j, m\rangle = \hbar \sqrt{j(j+1) - m(m \pm 1)} \, |j, m \pm 1\rangle \tag{168}$$

3.2.1 Matrix representation of angular momentum

The formalism we discussed in the previous section is completely general and independent on the choice of representation. There are many ways to represent the angular momentum operators and their eigenstates, and in this section we will discuss the matrix representation of angular momentum, where eigenkets and operators are represented by column and square matrices, respectively.

Since $\vec{J}^2$ and J_z commute, the set of their common eigenstates $\{|j, m\rangle\}$ can be chosen as a basis. This basis is discrete, orthonormal and complete. The orthonormality condition and the completeness relation can be expressed, respectively, as:

$$\langle j', m' | j, m \rangle = \delta_{j',j} \delta_{m',m} \qquad \sum_j \sum_{m=-j}^{j} |j, m\rangle \langle j, m| = \mathbb{1} \tag{169}$$

---

[27]Note that here we would actually have to add a phase factor $e^{i\theta}$ so that $C_{jm}^+ = \hbar \sqrt{j(j+1) - m(m+1)} e^{i\theta}$. This is beacuse C_{jm}^+ is, in general terms, a complex number. As we have the freedom of choice, we take C_{jm}^+ to be real and positive. The same argument applies to C_{jm}^-.

The matrix elements of the $\vec{J}^2$ and J_z operators in this basis are:

$$\langle j', m'| \, \vec{J}^2 \, |j, m\rangle = \hbar^2 j(j+1)\delta_{j',j}\delta_{m',m}$$
$$\langle j', m'| \, J_z \, |j, m\rangle = \hbar m\delta_{j',j}\delta_{m',m} \tag{170}$$

Clearly, $\vec{J}^2$ and J_z are diagonal in this basis, their diagonal elements being $\hbar^2 j(j+1)$ and $\hbar m$, respectively. However, as $J_\pm$ do not commute with $\vec{J}^2$ and J_z, they are not diagonal in this basis. Their matrix elements are:

$$\langle j', m'| \, J_\pm \, |j, m\rangle = \hbar\sqrt{j(j+1) - m(m\pm 1)}\delta_{j',j}\delta_{m',m\pm 1} \tag{171}$$

The matrices of J_x and J_y are obtained by adding and subtracting the matrices of J_+ and J_-, respectively, as seen in **Equation 136**:

$$J_x = \frac{1}{2}(J_+ + J_-) \qquad J_y = \frac{1}{2i}(J_+ - J_-) \tag{172}$$

3.2.2 Geometrical representation of angular momentum

The expected value of the angular momentum operator for a state $|j, m\rangle$, denoted as $\langle \vec{J} \rangle$, is given by:

$$\langle \vec{J} \rangle = \sqrt{\langle \vec{J}^2 \rangle} = \sqrt{\langle j, m|\vec{J}^2|j, m\rangle} = \hbar\sqrt{j(j+1)} \tag{173}$$

with z-component given by:

$$\langle J_z \rangle = \langle j, m|J_z|j, m\rangle = \hbar m \tag{174}$$

Recalling that $J_x^2 + J_y^2 = \vec{J}^2 - J_z^2$, we can also give the component of the angular momentum that lies in the xy-plane as:

$$\langle J_{xy} \rangle = \sqrt{\langle J_x^2 + J_y^2 \rangle} = \sqrt{\langle \vec{J}^2 - J_z^2 \rangle} = \hbar\sqrt{j(j+1) - m^2} \tag{175}$$

As expected, it is verified that:

$$\langle J_{xy} \rangle^2 + \langle J_z \rangle^2 = \left(\hbar\sqrt{j(j+1) - m^2}\right)^2 + (\hbar m)^2 = \hbar^2 j(j+1) = \langle \vec{J} \rangle^2 \tag{176}$$

Graphically, we can show this with a diagram like the one on **Figure 2**, which shows the graphical representation of the angular momentum $j = 2$ for the state $|j, m\rangle = |2, m\rangle$ with $m = -2, -1, 0, 1, 2$.

In the classical sense, we can think of $\vec{J}^2$ as a vector whose endpoint lies on a sphere of radius $\hbar\sqrt{j(j+1)}$, rotating about the z-axis along the surface of a cone, and whose projection on the z-axis has length $\hbar m$. The angle θ between the z-axis and the vector $\vec{J}^2$ is given by:

$$\cos\theta = \frac{\langle J_z \rangle}{\langle \vec{J} \rangle} = \frac{\hbar m}{\hbar\sqrt{j(j+1)}} = \frac{m}{\sqrt{j(j+1)}} \tag{177}$$

Notice that, as the values of the quantum number m are limited to $-j, -(j-1), ..., j-1, j$, the angle θ is also limited to a discrete set of $2j + 1$ values.

Since all the orientations of $\vec{J}^2$ on the surface of the cone are equaly likely, the projection of $\vec{J}^2$ on both the x-axes and y-axes averages out to zero. This can also be demostrated mathematically by calculating the expected values of J_x and J_y using **Equation 136**:

$$\langle J_x \rangle = \langle j, m|J_x|j, m\rangle = \frac{1}{2}\langle j, m|(J_+ + J_-)|j, m\rangle = \frac{1}{2}\langle j, m|J_+|j, m\rangle + \frac{1}{2}\langle j, m|J_-|j, m\rangle =$$
$$= \frac{\hbar}{2}\sqrt{j(j+1) - m(m+1)}\langle j, m|j, m+1\rangle + \frac{\hbar}{2}\sqrt{j(j+1) - m(m-1)}\langle j, m|j, m-1\rangle = 0 \tag{178}$$

$$\langle J_y \rangle = \langle j, m|J_y|j, m\rangle = \frac{1}{2i}\langle j, m|(J_+ - J_-)|j, m\rangle = \frac{1}{2i}\langle j, m|J_+|j, m\rangle - \frac{1}{2i}\langle j, m|J_-|j, m\rangle =$$
$$= \frac{\hbar}{2i}\sqrt{j(j+1) - m(m+1)}\langle j, m|j, m+1\rangle - \frac{\hbar}{2i}\sqrt{j(j+1) - m(m-1)}\langle j, m|j, m-1\rangle = 0 \tag{179}$$
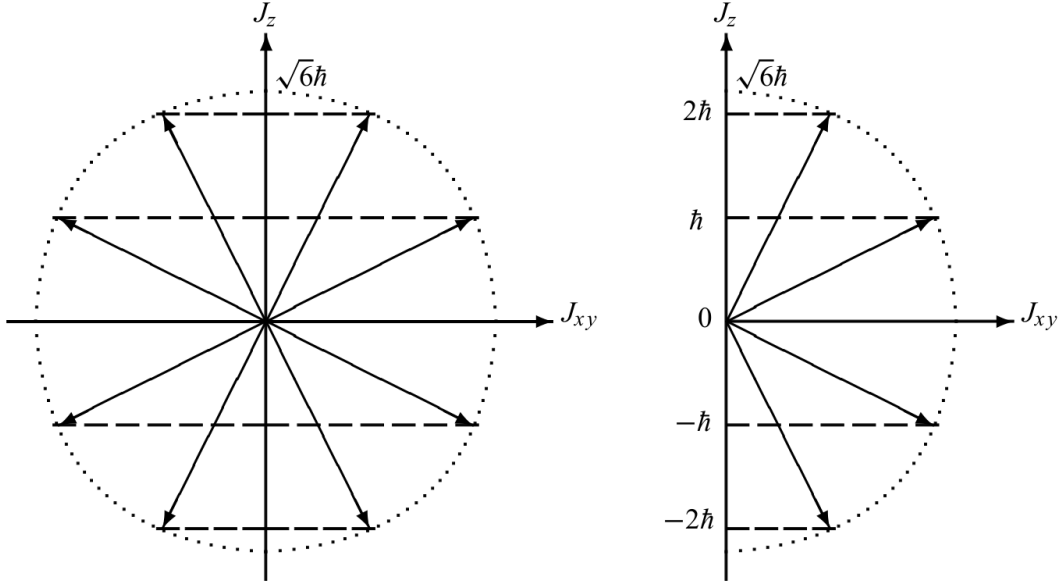
Figure 2: Graphical representation of the angular momentum $j = 2$ for the state $|j, m\rangle = |2, m\rangle$ with $m = -2, -1, 0, 1, 2$. The radius of the circle is $\hbar\sqrt{j(j+1)} = \hbar\sqrt{6}$.

3.3   Position representation of orbital angular momentum

As we already saw in a previous section, in the position representation, the orbital angular momentum operators are given in cartesian coordinates by **Equation 130**. However, it is more convenient to work in spherical (or polar) coordinates, since, as we shall see, the various angular momentum operators act only on the angular variables and and not on the r variable. We can rewrite the angular momentum operators in spherical coordinates as:

$$
\begin{aligned}
L_x &= i\hbar\left(\sin\phi\frac{\partial}{\partial\theta} + \frac{\cos\phi}{\tan\theta}\frac{\partial}{\partial\phi}\right) \\
L_y &= i\hbar\left(-\cos\phi\frac{\partial}{\partial\theta} + \frac{\sin\phi}{\tan\theta}\frac{\partial}{\partial\phi}\right) \\
L_z &= -i\hbar\frac{\partial}{\partial\phi}
\end{aligned}
\tag{180}
$$

We can also rewrite the $L_\pm$ and L^2 operators in spherical coordinates:

$$
\begin{aligned}
L_\pm &= \pm\hbar e^{\pm i\phi}\left(\frac{\partial}{\partial\theta} \pm i\cot\theta\frac{\partial}{\partial\phi}\right) \\
\vec{L}^{\,2} &= -\hbar^2\left[\frac{\partial^2}{\partial\theta^2} + \frac{1}{\tan\theta}\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]
\end{aligned}
\tag{181}
$$

Then, the original eigenproblem for a spinless particle can be written as:

$$
\begin{aligned}
\vec{L}^{\,2}|l, m\rangle &= \hbar^2 l(l+1)|l, m\rangle \\
L_z|l, m\rangle &= \hbar m|l, m\rangle
\end{aligned}
\tag{182}
$$

can be written in the r-representation as:

$$
\begin{aligned}
\vec{L}^{\,2}\psi_{l,m}(r, \theta, \phi) &= \hbar^2 l(l+1)\psi_{l,m}(r, \theta, \phi) \\
L_z\psi_{l,m}(r, \theta, \phi) &= \hbar m\psi_{l,m}(r, \theta, \phi)
\end{aligned}
\tag{183}
$$

where, as the operators act only on the angular variables, the eigenfunctions are separable:

$$\psi_{l,m}(r,\theta,\phi) = R_l(r)Y_l^m(\theta,\phi) \tag{184}$$

The solution to the angular part of this eigenproblem are **spherical harmonics**:

$$Y_l^m(\theta,\phi) = \varepsilon\sqrt{\frac{(2l+1)}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}}P_l^m(\cos\theta)e^{im\phi}$$

$$\varepsilon = \begin{cases} (-1)^m & \text{if } m \geq 0 \\ 1 & \text{if } m < 0 \end{cases} \tag{185}$$

where $P_l^m(\cos\theta)$ are the **associated Legendre polynomials**, which are given by:

$$P_l^m(x) = (1-x^2)^{|m|/2}\frac{d^{|m|}}{dx^{|m|}}P_l(x), \qquad P_l(x) = \frac{1}{2^l l!}\frac{d^l}{dx^l}(x^2-1)^l \tag{186}$$

Spherical harmonics are joint eigenfunctions of the operators $\vec{L}^2$ and L_z, as both these operators commute. We can therefore write:

$$\vec{L}^2 Y_l^m(\theta,\phi) = \hbar^2 l(l+1)Y_l^m(\theta,\phi)$$
$$L_z Y_l^m(\theta,\phi) = \hbar m Y_l^m(\theta,\phi) \tag{187}$$

They are also functions with a well-defined parity, as they are eigenfunctions of the parity operator P:

$$PY_l^m(\theta,\phi) = Y_l^m(\pi-\theta,\pi+\phi) = (-1)^l Y_l^m(\theta,\phi) \tag{188}$$

Also, the complex conjugate of a spherical harmonic is given by:

$$[Y_l^m(\theta,\phi)]^* = (-1)^m Y_l^{-m}(\theta,\phi) \tag{189}$$

3.3.1 Values of m and l

Using $L_z = -i\hbar\partial/\partial\phi$ (from **Equation 180**), we can rewrite **Equation 187** as:

$$-i\hbar\frac{\partial}{\partial\phi}Y_l^m(\theta,\phi) = \hbar m Y_l^m(\theta,\phi) \tag{190}$$

From this, we obtain that $Y_l^m(\theta,\phi) = F_l^m(\theta)e^{im\phi}$. However, as wave functions must be continuous at all points in space, we must impose that:

$$Y_l^m(\theta,\phi=0) = Y_l^m(\theta,\phi=2\pi) \longrightarrow 1 = e^{2\pi mi} \tag{191}$$

We already had from the general formalism of angular momentum that m had to be either integral or half-integral. Now, we have this additional condition that restricts m to being strictly integer, as half-integer values of m would result in $e^{2\pi mi} = -1$. But, if m is integral, then l must also be integral. Therefore, we can now define:

Definition 21. The eigenvalues of $\vec{L}^2$ and L_z corresponding to the joint eigenvectors $|l,m\rangle$ are given, respectively, by $\hbar^2 l(l+1)$ and $\hbar m$:

$$\vec{L}^2 |l,m\rangle = \hbar^2 l(l+1)|l,m\rangle \qquad L_z|l,m\rangle = \hbar m|l,m\rangle \tag{192}$$

where $l = 0, 1, 2, 3, \ldots$ and $m = -l, -(l-1), \ldots, l-1, l$. We see that the spectra of the orbital angular momentum operators $\vec{L}^2$ and L_z are discrete. Since the eigenstates corresponding to different angular momenta are orthogonal, and since the angular momentum spectra are discrete, the orthonormality

31

condition is:

$$\langle l', m'|l, m\rangle = \delta_{l',l}\delta_{m',m} \tag{193}$$

Note that the m we use here is different to the one used in the general angular momentum $\vec{J}$. In order to tell them apart, this one we are using here is often denoted as m_l.

3.4 Spin angular momentum

Spin angular momentum is one of the most fundamental properties of an electron and other elementary particles, despite the fact that it has no classical analogue, in contrast to orbital angular momentum. Its existence was discovered in 1922 by Stern and Gerlach, who observed that a beam of ground-state silver atoms, which have no orbital angular momentum, split in two when passing through a non-uniform magnetic field. This unexpected result, caused by the spin angular momentum of the $5s$ valence electron, led them to postulate the existence of a new angular momentum phenomenon different from orbital angular momentum, and they called it **spin**. The theory of spin is identical to the general theory of angular momentum and, by analogy with the vector $\vec{J}$, the spin is also represented by a vector operator $\vec{S}$.

From the classical theory of electromagnetism, an orbital magnetic dipole moment is generated with the orbital motion of a particle of charge q:

$$\vec{\mu}_L = \frac{q}{2m}\vec{L} \tag{194}$$

where $\vec{L}$ is the orbital angular momentum of the particle, m its mass, and c the speed of light. Similarly, if we follow a classical analysis, we can define the **spin magnetic dipole moment** of a particle as:

$$\vec{\mu}_S = \frac{q}{2m}\vec{S} \tag{195}$$

Although this expression cannot be derived classically, as there is no classical equivalent of spin, it can still be postulated by analogy to the orbital magnetic dipole moment. However, there is one correction that we need to make to **Equation 195**: it turns out that, for the electron, the orbital magnetic dipole moment is actually twice its classical expression. For other particles, this ratio, which we call gyromagnetic ratio (or Landé factor) g_s, can vary. Taking this effect into account, we obtain[28]:

$$\vec{\mu}_S = g_s\frac{q}{2m}\vec{S} \tag{196}$$

When the electron is placed in a magnetic field $\vec{B}$ and if the field is inhomogeneous, a force will be exerted on the electron's intrinsic dipole moment; the direction and magnitude of the force depend on the relative orientation of the field and the dipole. This force tends to align $\vec{\mu}_S$ along $\vec{b}$, producing a precessional motion of $\vec{\mu}_S$ around $\vec{b}$.

3.4.1 Quantum description of spin

Until now, we have studied the quantisation of *orbital variables*. With the position $\vec{r}$ and the momentum $\vec{p}$ of a particle such as the electron, we associated the observables $\vec{R}$ and $\vec{P}$ acting in the state space $\varepsilon_{\vec{r}}$, which is isomorphic to the space $\mathcal{F}$ of wave functions. All physical quantities are functions of the fundamental variables $\vec{r}$ and $\vec{p}$, and the quantisation rules enable us to associate with them observables acting in $\vec{r}$. We shall call $\varepsilon_{\vec{r}}$ the *orbital state space*.

To these *orbital variables* we have studied up until now, we will add the *spin variables*, which satisfy the following postulates:

---

[28]The ratio g_s can be obtained using Dirac's relativistic quantum theory. For the electron, $g_s = 2$.

Postulate 6. The spin operator: The spin operator $\vec{S}$ is an angular momentum vector operator whose components S_x, S_y and S_z satisfy the commutation relations of angular momentum operators:

$$[S_x, S_y] = i\hbar S_z, \qquad [S_y, S_z] = i\hbar S_x, \qquad [S_z, S_x] = i\hbar S_y \tag{197}$$

Postulate 7. Spin state space: The spin operators act in a new space, the "spin state space" ε_s, where $\vec{S}^2$ and S_z constitute a CSCO. The space ε_s is thus spanned by the set of eigenstates $|s, m_s\rangle$ common to $\vec{S}^2$ and S_z:

$$\vec{S}^2 |s, m_s\rangle = \hbar^2 s(s+1) |s, m_s\rangle, \qquad S_z |s, m_s\rangle = \hbar m_s |s, m_s\rangle \tag{198}$$

where $s = 0, 1/2, 1, 3/2, ...$ and $m_s = -s, -(s-1), ..., s-1, s$. A given particle is characterised by a *unique*, *specific* and *immutable* value[a] of s: this particle is said to have a spin s. The spin state space ε_s is therefore always of finite dimension $(2s+1)$, and all spin states are eigenvectors of $\vec{S}^2$ with the same eigenvalue $s(s+1)\hbar^2$.

Similarly, we have:

$$S_\pm |s, m_s\rangle = \hbar\sqrt{s(s+1) - m_s(m_s \pm 1)} |s, m_s \pm 1\rangle \tag{199}$$

where $S_\pm = S_x \pm iS_y$, and:

$$\langle S_x^2 \rangle = \langle S_y^2 \rangle = \frac{\hbar^2}{2} \left[s(s+1) - m_s^2 \right] \tag{200}$$

The spin states form an orthonormal and complete basis:

$$\langle s', m_s'|s, m_s\rangle = \delta_{s',s}\delta_{m_s',m_s} \qquad \sum_s \sum_{m_s=-s}^{s} |s, m_s\rangle \langle s, m_s| = \mathbb{1} \tag{201}$$

---

[a]For example, the electron has $s = 1/2$, the photon has $s = 1$, the proton has $s = 1/2$, the neutron has $s = 1/2$, etc. The value of s is called the **spin quantum number**.

Postulate 8. The state space: The state space ε of a particle is the tensor product of the orbital state space $\varepsilon_{\vec{r}}$ and the spin state space ε_s:

$$\varepsilon = \varepsilon_{\vec{r}} \otimes \varepsilon_s \tag{202}$$

Note that the spin angular momentum is an **intrinsic property of the particle**. It has nothing to do with motion of the particle in space. All angular momentum related to the spatial dependence of the wave function is included in $\vec{L}$. Therefore, spin angular momentum cannot be described by any function of the position variables. An important consequence of **Postulate 8** related to this is that all *orbital observables* commute with all *spin observables*, as they act on different state spaces. Spin operators act on the spin part and leave the spatial part unchanged, and vice versa. In other words, the spin of a particle and its orbital motion do not affect each other.

3.4.2 Spin 1/2

In the case of a particle of spin $s = 1/2$, the spin state ε_s is two-dimensional, as we have two possible values of $m_s = \pm 1/2$. In this space, we will consider the orthonormal basis of eigenstates common to $\vec{S}^2$ and S_z:

- For $s = 1/2$ and $m_s = 1/2$, we have $|1/2, 1/2\rangle = |\uparrow\rangle = |+\rangle$

- For $s = 1/2$ and $m_s = -1/2$, we have $|1/2, -1/2\rangle = |\downarrow\rangle = |-\rangle$

With this basis, we can write the eigenvaue equations from **Equation 198** as:

$$\vec{S}^2 |\pm\rangle = \frac{3}{4}\hbar^2 |\pm\rangle, \qquad S_z |\pm\rangle = \pm\frac{1}{2}\hbar |\pm\rangle \tag{203}$$

Matrix representation

Note from **Equation 203** that *all* kets of the basis of ε_s are eigenvectors of $\vec{S}^2$ with the same eigenvalue $3\hbar^2/4$. This causes $\vec{S}^2$ to be proportional to the unity operator $\mathbb{1}$ of ε_s:

$$\vec{S}^2 = \frac{3}{4}\hbar^2\mathbb{1} \tag{204}$$

We can also see this by taking the matrix elements of $\vec{S}^2$ explicitly:

$$\vec{S}^2 = \begin{pmatrix} \langle+|\vec{S}^2|+\rangle & \langle+|\vec{S}^2|-\rangle \\ \langle-|\vec{S}^2|+\rangle & \langle-|\vec{S}^2|-\rangle \end{pmatrix} = \begin{pmatrix} \frac{3}{4}\hbar^2 \langle+|+\rangle & \frac{3}{4}\hbar^2 \langle+|-\rangle \\ \frac{3}{4}\hbar^2 \langle-|+\rangle & \frac{3}{4}\hbar^2 \langle-|-\rangle \end{pmatrix} = \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{3}{4}\hbar^2\mathbb{1} \tag{205}$$

The matrix of S_z can also be obtained easily from the eigenvalue equation:

$$S_z = \begin{pmatrix} \langle+|S_z|+\rangle & \langle+|S_z|-\rangle \\ \langle-|S_z|+\rangle & \langle-|S_z|-\rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\hbar \langle+|+\rangle & -\frac{1}{2}\hbar \langle+|-\rangle \\ \frac{1}{2}\hbar \langle-|+\rangle & -\frac{1}{2}\hbar \langle-|-\rangle \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{206}$$

From **Equation 199**, we can obtain:

$$S_+ |+\rangle = 0$$
$$S_+ |-\rangle = \hbar\sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right) - \left(-\frac{1}{2}\right)\left(\left(-\frac{1}{2}\right)+1\right)} |+\rangle = \hbar |+\rangle$$
$$S_- |+\rangle = \hbar\sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right) - \frac{1}{2}\left(\frac{1}{2}-1\right)} |-\rangle = \hbar |-\rangle \tag{207}$$
$$S_- |-\rangle = 0$$

Therefore, the matrices of S_+ and S_- are:

$$S_+ = \begin{pmatrix} \langle+|S_+|+\rangle & \langle+|S_+|-\rangle \\ \langle-|S_+|+\rangle & \langle-|S_+|-\rangle \end{pmatrix} = \begin{pmatrix} 0 & \hbar\langle+|+\rangle \\ 0 & \hbar\langle-|+\rangle \end{pmatrix} = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{208}$$

$$S_- = \begin{pmatrix} \langle+|S_-|+\rangle & \langle+|S_-|-\rangle \\ \langle-|S_-|+\rangle & \langle-|S_-|-\rangle \end{pmatrix} = \begin{pmatrix} \hbar\langle+|-\rangle & 0 \\ \hbar\langle-|-\rangle & 0 \end{pmatrix} = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{209}$$

And since $S_x = \frac{1}{2}(S_+ + S_-)$ and $S_y = \frac{1}{2i}(S_+ - S_-)$, we can obtain the matrices of S_x and S_y:

$$S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad S_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{210}$$

We can express the joint eigenvectors of $\vec{S}^2$ and S_z in terms of two-element column matrices known as *spinnors*:

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{211}$$

Any vector in the state space ε_s can be written as a linear combination of the basis vectors $|\pm\rangle$:

$$|\psi\rangle = c_+ |+\rangle + c_- |-\rangle \tag{212}$$

where $c_{\pm}$ are complex numbers. We can also write $|\psi\rangle$ in the form:

$$|\psi\rangle = \begin{pmatrix} c_+ \\ c_- \end{pmatrix} \tag{213}$$

Pauli matrices

When $s = 1/2$, it is convenient to introduce the **Pauli matrices**, σ_x, σ_y and σ_z, which are defined as:

$$\vec{S} = \frac{\hbar}{2}\vec{\sigma} \tag{214}$$

So that:

$$\frac{\hbar}{2}\sqrt{S_x^2 + S_y^2 + S_z^2} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2} \longrightarrow \sqrt{\left(\frac{\hbar}{2}S_x\right)^2 + \left(\frac{\hbar}{2}S_y\right)^2 + \left(\frac{\hbar}{2}S_z\right)^2} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2} \rightarrow \tag{215}$$

Finally:

$$\sigma_x = \frac{\hbar}{2}S_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \frac{\hbar}{2}S_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \frac{\hbar}{2}S_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{216}$$

3.4.3 Spin-orbit coupling

Spin-orbit coupling in hydrogen is an effect that arises from the interaction between the electron's spin magnetic moment, $\vec{\mu}_S$, and the proton's orbtial magnetic field $\vec{B}$.

The origin of the magnetic field experienced by the electron moving at $\vec{v}$ in a circular orbit around the proton can be explained classically as follows. The electron, within its rest frame, sees the proton moving at $-\vec{v}$ in a circular orbit around it (**Figure 3**). From classical electrodynamics, the magnetic field experienced by the electron is:

$$\vec{B} = -\frac{\vec{v}}{c^2} \times \vec{E} = -\frac{\vec{p}}{m_e c^2} \times \vec{E} = \frac{\vec{E}}{m_e c^2} \times \vec{p} \tag{217}$$

where $\vec{p} = m_e \vec{v}$ is the linear momentum of the electron and $\vec{E}$ is the electric field generated by the proton, which is given by:

$$\vec{E}(\vec{r}) = \frac{e}{r^2}\frac{\vec{r}}{r} = \frac{e\vec{r}}{r^3} \tag{218}$$
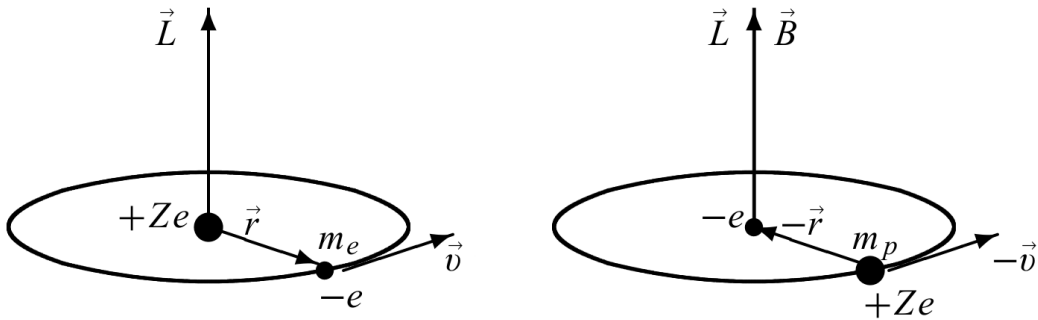
Figure 3: An electron moving in a circular orbit as seen by the nucleus (left) and the same motion as seen by the electron within its rest frame; the electron sees the nucleus moving in a circular orbit around it (right).

For a more general problem of hydrogen-like atoms (those with one valence electron outside a closed shell) were an electron moves in the (central) Coulomb potential of a nucleus $V(r) = -e\phi(r)$, the electric field is given by:

$$\vec{E}(\vec{r}) = -\nabla\phi(r) = \frac{1}{e}\vec{\nabla}V(r) = \frac{1}{e}\frac{\vec{r}}{r}\frac{dV}{dr} \tag{219}$$

So the magnetic field of the nucleus calculated in the rest frame of the electron is obtained by inserting **Equation 219** into **Equation 217**:

$$\vec{B} = \frac{\vec{E}}{m_e c^2} \times \vec{p} = \frac{1}{e m_e c^2} \frac{1}{r} \frac{dV}{dr} \vec{r} \times \vec{p} = \frac{1}{e m_e c^2} \frac{1}{r} \frac{dV}{dr} \vec{L} \tag{220}$$

where $\vec{L} = \vec{r} \times \vec{p}$ is the orbital angular momentum of the electron. The interaction of the electron's spin dipole moment $\vec{\mu}_S = -e\vec{S}/m_e$ with the orbital magnetic field $\vec{B}$ of the nucleus gives rise to the following interaction energy:

$$H_{SO} = -\vec{\mu}_S \cdot \vec{B} = \frac{e}{m_e} \vec{S} \cdot \vec{B} = \frac{1}{m_e^2 c^2} \frac{1}{r} \frac{dV}{dr} \vec{S} \cdot \vec{L} \tag{221}$$

This energy turns out to be twice the observed spin-orbit interaction energy. This is due to the fact that H_{SO} was calculated within the rest frame of the electron. This frame is not inertial, for the electron accelerates while moving in a circular orbit around the nucleus. For a correct treatment, we must transform to the rest frame of the nucleus (i.e., the lab frame). This transformation, which involves a relativistic transformation of velocities, gives rise to an additional motion resulting from the precession of $\vec{\mu}_S$; this is known as the *Thomas precession*. The precession of the electron's spin moment is a relativistic effect which occurs even in the absence of an external magnetic field. The transformation back to the rest frame of the nucleus leads to a reduction of the interaction energy computed before by a factor of 2:

$$H_{SO} = \frac{1}{2 m_e^2 c^2} \frac{1}{r} \frac{dV}{dr} \vec{S} \cdot \vec{L} \tag{222}$$

In order to obtain the quantum mechanical expression from this classically obtained result, one need only replace the classical quantities $\vec{S}$ and $\vec{L}$ by their quantum mechanical operators.

For a hydrogen's electron, $V(r) = -e^2/r$ and $dV/dr = e^2/r^2$, so that:

$$H_{SO} = \frac{e^2}{2 m_e^2 c^2} \frac{1}{r^3} \vec{S} \cdot \vec{L} \tag{223}$$

If we consider only this as a standalone Hamiltonian, we can find the energy eigenvalues by taking $\vec{S} \cdot \vec{L} = \frac{1}{2}(\vec{J}^2 - \vec{L}^2 - \vec{S}^2)$, which comes from expanding $\vec{J}^2 = (\vec{L} + \vec{S})^2$:

$$H_{SO} = \frac{e^2}{4 m_e^2 c^2} \frac{1}{r^3} (\vec{J}^2 - \vec{L}^2 - \vec{S}^2) \longrightarrow E_{SO} = \frac{\hbar^2}{4 m_e^2 c^2} \frac{e^2}{r^3} \left(j(j+1) - l(l+1) - s(s+1) \right) \tag{224}$$

In terms of the Bohr magneton $\mu_B = e\hbar/2m_e$:

$$E_{SO} = \frac{\mu_B^2}{c^2 r^3} \left[j(j+1) - s(s+1) - l(l+1) \right] \tag{225}$$

The scales of energies associated with the spin-orbit coupling are about two orders of magnitude smaller than the energies associated with the discrete atomic levels, so we can consider this energy H_{SO} as a *perturbation* to the Hamiltonian of the hydrogen atom and use *stationary perturbation theory*[29] to calculate corrections due to the spin-orbit coupling:

$$H = H_0 + H_{SO} = \frac{\vec{p}^2}{2 m_e} - \frac{e^2}{r} + \frac{e^2}{2 m_e^2 c^2} \frac{1}{r^3} \vec{S} \cdot \vec{L} \tag{226}$$

where H_0 is the unperturbed Hamiltonian and H_{SO} is the spin-orbit coupling perturbation.

---

[29]We will discuss this in more detail in the next chapter.

3.4.4 The Zeeman effect

As we have discussed in previous sections, the electron posesses both an orbital magnetic dipole moment, $\vec{\mu}_L$, and a spin magnetic dipole moment, $\vec{\mu}_S$. Both of these moments give rise to two energy terms when the electron is submitted to a magnetic field $\vec{B}$, $-\vec{\mu}_L \cdot \vec{B}$ and $-\vec{\mu}_S \cdot \vec{B}$, whose sum we call the Zeeman energy. Using **Equation 194** and **Equation 196**, we can write the Zeeman energy as[30]:

$$H_Z = -\vec{\mu}_L \cdot \vec{B} - \vec{\mu}_S \cdot \vec{B} = \frac{e}{2m_e} \vec{L} \cdot \vec{B} + \frac{e}{m_e} \vec{S} \cdot \vec{B} = \frac{e}{2m_e} \left(\vec{L} + 2\vec{S} \right) \cdot \vec{B} \tag{227}$$

We can write this in terms of Bohr's magneton $\mu_B = \frac{e\hbar}{2m_e}$ as:

$$H_Z = \frac{\mu_B}{\hbar} \left(\vec{L} + 2\vec{S} \right) \cdot \vec{B} \tag{228}$$

What we have just described is the energy corresponding to the so called **anomalous Zeeman effect**, which takes into account the spin of the electron. The **normal Zeeman effect** is the case where we only consider the orbital angular momentum of the electron, which yields the following expression for the energy:

$$H_Z = \frac{\mu_B}{\hbar} \vec{L} \cdot \vec{B} \tag{229}$$

The normal Zeeman effect

Let us first consider the normal Zeeman effect. For simplicity, we shall take $\vec{B} = B\vec{e}_z$, so that:

$$H_z = \frac{\mu_B B}{\hbar} L_z \tag{230}$$

Solving the eigenvalue equation for the total energy $H = H_0 + H_Z$, we find:

$$\begin{aligned} H \left| n, l, m_l \right\rangle = (H_0 + H_Z) \left| n, l, m_l \right\rangle = H_0 \left| n, l, m_l \right\rangle + \frac{\mu_B B}{\hbar} L_z \left| n, l, m_l \right\rangle = \\ = E_n^0 \left| n, l, m_l \right\rangle + \frac{\mu_B B}{\hbar} \hbar m_l \left| n, l, m_l \right\rangle = \left(E_n^0 + m_l \mu_B B \right) \left| n, l, m_l \right\rangle \end{aligned} \tag{231}$$

Therefore, the difference between two consecutive energy levels for each value of n is given by:

$$\Delta E = \mu_B B \tag{232}$$

Take the example of $n = 2$, where we have $l = 0, 1$:

- When $l = 0$, we will only have one value of m_l, which is $m_l = 0$. Therefore, we will only have a single energy level:

  $\rightarrow E_{0,0} = E_2^0$

- When $l = 1$, we will have three possible values of m_l, which are $m_l = -1, 0, 1$. Therefore, we will obtain three energy levels:

  $\rightarrow E_{1,-1} = E_2^0 - \mu_B B$
  $\rightarrow E_{1,0} = E_2^0$
  $\rightarrow E_{1,1} = E_2^0 + \mu_B B$

---

[30]Here, we have used the fact that $g_s = 2$ and $q = -e$ for the electron.

Overall, we have just three *unique* values of the energy. Experimentally, the way we would expect to see this is by observing the splitting of the spectral lines of the atom in the presence of a magnetic field. Instead of seeing *one* line, we would see *three* energy levels: one corresponding to the original level and two at the levels shifted by $\pm\mu_B B$. In general, we would see $2l+1$ lines, corresponding to the $2l+1$ possible values of m_l, so that we would always have an odd number of lines in any spectra.

The anomalous Zeeman effect

Unfortunately, in many cases the results of the normal Zeeman effect disagree with the experimental observations. For instance, spectral lines can sometimes split into an *even* number of levels, which could never happen according to the normal Zeeman effect. Why, then, does this happen? The answer, of course, is *spin*. We will now see the two different cases of anomalous Zeeman effect, depending on the strength of the magnetic field $\vec{B}$.

a) Strong $\vec{B}$

Let us again consider $\vec{B} = B\vec{e}_z$, so that:

$$H_Z = \frac{\mu_B B}{\hbar}\left(L_z + 2S_z\right) \tag{233}$$

In order to find a general result considering spin, it is also important to take into account the spin-orbit coupling effects. Therefore, we need to include the term H_{SO} in our Hamiltonian, so that the eigenvalue equation for the general anomalous Zeeman effect becomes:

$$
\begin{aligned}
H\left|n, l, m_l, m_s\right\rangle &= (H_0 + H_Z)\left|n, l, m_l, m_s\right\rangle = \\
&= H_0\left|n, l, m_l, m_s\right\rangle + \frac{\mu_B B}{\hbar}\left(L_z + 2S_z\right)\left|n, l, m_l, m_s\right\rangle = \\
&= E_n^0\left|n, l, m_l, m_s\right\rangle + \frac{\mu_B B}{\hbar}\left(\hbar m_l + 2\hbar m_s\right)\left|n, l, m_l, m_s\right\rangle = \\
&= \left[E_n^0 + (m_l + 2m_s)\mu_B B\right]\left|n, l, m_l, m_s\right\rangle
\end{aligned}
\tag{234}
$$

Therefore, the energy levels E_n^0 are shifted by:

$$\Delta E = (m_l + 2m_s)\mu_B B \tag{235}$$

Take the same example as before of $n = 2$, where we have $l = 0, 1$. The electron has spin $s = 1/2$, so $m_s = \pm 1/2$. Then:

- When $l = 0$, we will only have one value of m_l, which is $m_l = 0$. However, we now have to consider m_s, which is has two possible values, $\pm 1/2$. Therefore, we will obtain two different energy levels:

  $\rightarrow E_{0,0,\frac{1}{2}} = E_2^0 + \mu_B B$

  $\rightarrow E_{0,0,-\frac{1}{2}} = E_2^0 - \mu_B B$

- When $l = 1$, we will have three possible values of m_l, which are $m_l = -1, 0, 1$. For each of the values of m_l, we have two possible values of m_s, so we will obtain six different energy levels for $l = 1$:

  $\rightarrow E_{1,1,\frac{1}{2}} = E_2^0 + 2\mu_B B$

  $\rightarrow E_{1,0,\frac{1}{2}} = E_2^0 + \mu_B B$

  $\rightarrow E_{1,-1,\frac{1}{2}} = E_2^0$

  $\rightarrow E_{1,1,-\frac{1}{2}} = E_2^0$

  $\rightarrow E_{1,0,-\frac{1}{2}} = E_2^0 - \mu_B B$

  $\rightarrow E_{1,-1,-\frac{1}{2}} = E_2^0 - 2\mu_B B$

Overall, we have *five* unique values of the energy, which is two more than what we expected from the normal Zeeman effect. There is still an odd number of energy levels. However, if we computed the energy levels for $n = 1$, we will have only the case $l = 0$, which will give the same levels as the $l = 0$ case for $n = 2$ that we just computed, which were only two. That is, $n = 1$ will show an *even* number of energy levels. Both these cases ($n = 1$ and $n = 2$) are shown in **Figure 4**.
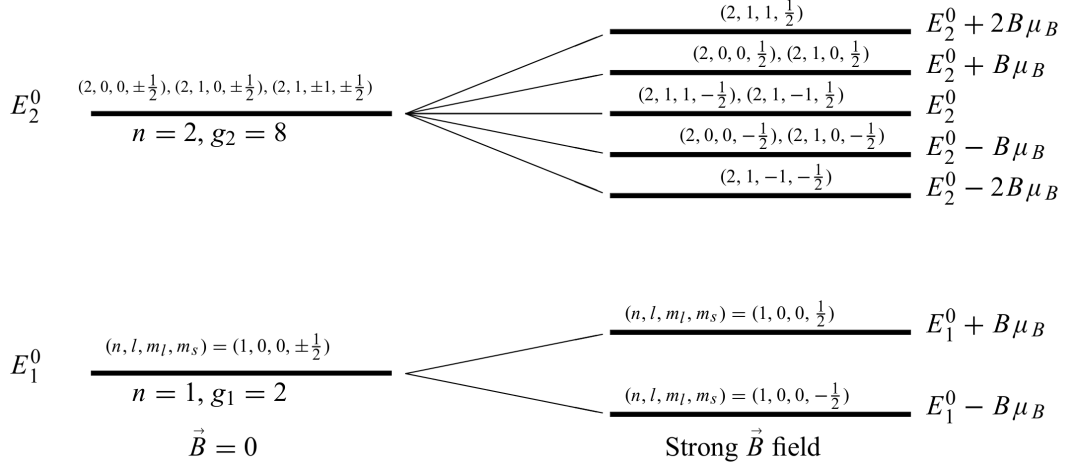

Figure 4: Energy level splitting in the absence of a magnetic field (left) and in the presence of a strong magnetic field $\vec{B}$ (right).

a) Weak $\vec{B}$

In the case where we have a weak magnetic field $\vec{B}$, we can no longer neglect the spin-orbit coupling effects[31]. Then, the best eigenstates to use are $|n, l, j, m_j\rangle$, as they are simultaneous eigenstates of H_0 and H_{SO}. It is therefore necessary to rewrite H_Z so that $\vec{S}$ and $\vec{L}$ only appear as square operators[32]. To do this, let us first consider:

$$\vec{J} = \vec{S} + \vec{L} \longrightarrow \vec{L} = \vec{J} - \vec{S} \tag{236}$$

Then:

$$\vec{L}^2 = \vec{J}^2 - \vec{S}^2 - \vec{J} \cdot \vec{S} - \vec{S} \cdot \vec{J} \overset{33}{=} \vec{J}^2 - \vec{S}^2 - 2\vec{J} \cdot \vec{S} \longrightarrow \vec{J} \cdot \vec{S} = \frac{\vec{J}^2 - \vec{S}^2 - \vec{L}^2}{2} \tag{237}$$

Therefore:

$$H_Z = \frac{\mu_B}{\hbar} \left(\vec{L} + 2\vec{S} \right) \cdot \vec{B} = \frac{\mu_B}{\hbar} \left(\vec{J} + \vec{S} \right) \cdot \vec{B} = \frac{\mu_B}{\hbar} \left(1 + \frac{\vec{S}}{\vec{J}} \right) \vec{J} \cdot \vec{B} =$$

$$= \frac{\mu_B}{\hbar} \left(1 + \frac{\vec{J} \cdot \vec{S}}{\vec{J}^2} \right) \vec{J} \cdot \vec{B} = \frac{\mu_B}{\hbar} \left(1 + \frac{\vec{J}^2 - \vec{S}^2 - \vec{L}^2}{2\vec{J}^2} \right) \vec{J} \cdot \vec{B} \tag{238}$$

---

[31]In reality, we should consider not only the spin-orbit coupling, but also the relativistic correction. Both of these together conform the *fine structure* of hydrogen. However, for simplicity, we will only consider the spin-orbit coupling effects in our analysis.

[32]Why?

[33]Remember that, as $\vec{L}$ and $\vec{S}$ commute (they act on different spaces $\varepsilon_{\vec{r}}$ and ε_s), then $\vec{J}$ and $\vec{S}$ also commute: $\vec{J} \cdot \vec{S} - \vec{S} \cdot \vec{J} = (\vec{L} + \vec{S}) \cdot \vec{S} - \vec{S} \cdot (\vec{L} + \vec{S}) = \vec{L} \cdot \vec{S} + \vec{S}^2 - \vec{S} \cdot \vec{L} - \vec{S}^2 = 0$.

Then, we write the eigenvalue equation, assuming that $\vec{J}||\vec{B}$[34]:

$$H\left|n,l,j,m_j\right\rangle = (H_0 + H_{SO} + H_Z)\left|n,l,j,m_j\right\rangle =$$

$$= E_n^0\left|n,l,j,m_j\right\rangle + E_{SO}\left|n,l,j,m_j\right\rangle + \frac{\mu_B}{\hbar}\left(1 + \frac{\vec{J}^2 - \vec{S}^2 - \vec{L}^2}{2\vec{J}^2}\right)\vec{J}\cdot\vec{B}\left|n,l,j,m_j\right\rangle =$$

$$= \left[E_n^0 + E_{SO} + \frac{\mu_B}{\hbar}\left(1 + \frac{j(j+1) - s(s+1) - l(l+1)}{2j(j+1)}\right)\hbar m_j B\right]\left|n,l,j,m_j\right\rangle =$$

$$= \left[E_n^0 + E_{SO} + g_j m_j \mu_B B\right]\left|n,l,j,m_j\right\rangle \tag{239}$$

Therefore, the energy levels are shifted by two terms: E_{SO}, corresponding to spin-orbit coupling and $E_Z = g_j m_j \mu_B B$, corresponding to the Zeeman effect, where g_j is called the Landé factor and is defined as:

$$g_j = 1 + \frac{j(j+1) - s(s+1) - l(l+1)}{2j(j+1)} \tag{240}$$

Recall from the discussion about spin-orbit coupling that:

$$E_{SO} = \frac{\mu_B^2}{c^2 r^3}\left[j(j+1) - s(s+1) - l(l+1)\right] \tag{241}$$

Therefore, the effect of the magnetic field on the atom is to split the energy levels first due to the spin-orbit cupling effect, and then to further split those lines due to the Zeeman effect by an amount equal to:

$$\Delta E_Z = B\mu_B \Delta m_j g_j \tag{242}$$

Take the simple example where $n = 1$, where we have $l = 0$. The electron has spin $s = 1/2$, so $m_s = \pm 1/2$. We also have two different values of j, which are $j = l + s = \frac{1}{2}$ and $j = l - s = -\frac{1}{2}$. Then[35]:

- When $j = \frac{1}{2}$, we will have a spin-orbit coupling energy of:

$$E_{SO,\frac{1}{2}} = \frac{\mu_B^2}{c^2 r^3}\left[\frac{1}{2}\left(\frac{1}{2}+1\right) - \frac{1}{2}\left(\frac{1}{2}+1\right) - 0\left(0+1\right)\right] = 0 \tag{243}$$

  and a Landé factor of:

$$g_{j=\frac{1}{2}} = 1 + \frac{\frac{1}{2}\left(\frac{1}{2}+1\right) - \frac{1}{2}\left(\frac{1}{2}+1\right) - 0\left(0+1\right)}{2\cdot\frac{1}{2}\left(\frac{1}{2}+1\right)} = 1 \tag{244}$$

  Therefore, for the two possible values of m_j, we will obtain:

  $\rightarrow E_{0,0,\frac{1}{2}} = E_n^0 + \frac{\mu_B B}{2}$

  $\rightarrow E_{0,0,-\frac{1}{2}} = E_n^0 - \frac{\mu_B B}{2}$

- When $j = -\frac{1}{2}$, we will have a spin-orbit coupling energy of:

$$E_{SO,\frac{1}{2}} = \frac{\mu_B^2}{c^2 r^3}\left[-\frac{1}{2}\left(-\frac{1}{2}+1\right) - \frac{1}{2}\left(\frac{1}{2}+1\right) - 0\left(0+1\right)\right] = -\frac{\mu_B^2}{c^2 r^3} \tag{245}$$

  and a Landé factor of:

$$g_{j=\frac{1}{2}} = 1 + \frac{-\frac{1}{2}\left(-\frac{1}{2}+1\right) - \frac{1}{2}\left(\frac{1}{2}+1\right) - 0\left(0+1\right)}{2\cdot\left(-\frac{1}{2}\right)\left(-\frac{1}{2}+1\right)} = 3 \tag{246}$$

  Therefore, for the two possible values of m_j, we will obtain:

---

[34]We are computing the energy when the electrons are exposed to $\vec{B}$, which causes their spin angular momentum to align with the magnetc field. Why do we assume the whole angular momentum is aligned?

[35]Check all this

$$\rightarrow E_{0,0,\frac{1}{2}} = E_n^0 - \frac{\mu_B^2}{c^2 r^3} + \frac{3}{2}\mu_B B$$

$$\rightarrow E_{0,0,-\frac{1}{2}} = E_n^0 - \frac{\mu_B^2}{c^2 r^3} - \frac{3}{2}\mu_B B$$

Overall, just for $n = 1$, we have *four* unique values of the energy, due to the spin-orbit coupling effect.

4 Approximation Methods for Stationary States in Quantum Mechanics

Solving a general quantum mechanical problem involves finding the Hamiltonian of a system and solving the Schrödinger Equation to find energies and states. However, most problems encountered in quantum mechanics cannot be solved exactly. Exact solutions of the Schrödinger equation exist only for a few idealised systems[36]. To solve general problems, one must resort to approximation methods.

In this chapter we consider approximation methods that deal with *stationary states* corresponding to *time-independent* Hamiltonians. To study problems of stationary states, we focus on two approximation methods: **time-independent perturbation theory** and the **variational method**.

4.1 Introduction to perturbation theory

Perturbation theory provides a method to approximate energy eigenvalues and eigenstates of Hamiltonians of the form:

$$H = H_0 + H_p \tag{247}$$

where H_0 is the Hamiltonian of an exactly solvable system (unperturbed system) and H_p is a small time-independent perturbation. In other words, perturbation theory enables us to obtain solutions for systems that are "close" to exactly solvable systems.

For a general (possibly degenerate[37]) case, the discrete eigenvalues and eigenstates of the unperturbed Hamiltonian can be defined as:

$$H_0 |\psi_n^{i\,(0)}\rangle = E_n^{(0)} |\psi_n^{i\,(0)}\rangle \tag{248}$$

where the states are assumed to be orthonormal. We can also write the eigenvalue equation for the perturbed system:

$$H |\psi_n^i\rangle = E_n |\psi_n^i\rangle \tag{249}$$

The idea is to determine the small corrections required so that $|\psi_n^{i\,(0)}\rangle \to |\psi_n^i\rangle$ and $E_n^{(0)} \to E_n$. Of course, this can only be accomplished if H_p is "small enough" (we will quantify this later) so that $|\psi_n^{i\,(0)}\rangle$ are close to $|\psi_n^i\rangle$. It is common to write this using a dimensionless parameter $\lambda \ll 1$ so that:

$$H_p = \lambda W \tag{250}$$

Then, we can rewrite the perturbed Hamiltonian as:

$$H = H_0 + \lambda W \tag{251}$$

Depending on whether the unperturbed system has any degenerate eigenvalues or not, we have two different ways of treating the problem. We shall discuss both cases separately.

4.2 Non-degenerate perturbation theory

Let us assume that we have a system with an unperturbed and non-degenerate Hamiltonian H_0, the eigenvalue problem for which is solved and the spectrum of which is discrete:

$$H_0 |\psi_p^{(0)}\rangle = E_p^0 |\psi_p^{(0)}\rangle \tag{252}$$

---

[36]For example, for the *quantum harmonic oscillator*, the *quantum anharmonic oscillator*, the *infinite square well*, the *free particle*, and the *hydrogen atom* can be solved exactly. However, the *spin-orbital interactions*, the *Zeeman effect* or the *hyperfine split* cannot.

[37]Degeneracy is indexed by i.

The eigenstates $|\psi_p^{(0)}\rangle$ form a complete orthonormal basis:

$$\langle\psi_n^{(0)}|\psi_p^{(0)}\rangle = \delta_{np} \qquad \sum_p |\psi_p^{(0)}\rangle\langle\psi_p^{(0)}| = \mathbb{1} \tag{253}$$

We can also write the perturbed eigenproblem for the Hamiltonian $H(\lambda) = H_0 + \lambda W$ as:

$$H(\lambda)|\psi_n(\lambda)\rangle = E_n(\lambda)|\psi_n(\lambda)\rangle \tag{254}$$

The main idea of perturbation theory consists in now assuming that the perturbed eigenvalues and eigenstates can both be expanded in power series in the parameter λ:

$$E_n(\lambda) = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots = \sum_{k=0}^{\infty} E_n^{(k)} \lambda^k$$
$$|\psi_n(\lambda)\rangle = |\psi_n^{(0)}\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle + \cdots = \sum_{k=0}^{\infty} |\psi_n^{(k)}\rangle \lambda^k \tag{255}$$

The parameters $E_n^{(k)}$ and $|\psi_n^{(k)}\rangle$ are the k-th corrections to the eigenenergies and the eigenstates, respectively. We need to make two remarks here. First, one might think that, whenever the perturbation is sufficiently weak, the expansions in **Equation 255** always exist. Unfortunately, this is not always the case. In some cases where the perturbation is small, $E_n(\lambda)$ and $|\psi_n(\lambda)\rangle$ are not expandable in powers of λ. Second, these series are frequently not convergent. However, when λ is small, the first few terms do provide a reliable description of the system. So in practice, we keep only one or two terms in these expansions; hence the problem of nonconvergence of these series is avoided. Note that when $\lambda = 0$, the expressions in **Equation 255** yield the unperturbed solutions: $E_n = En^{(0)}$ and $|\psi_n\rangle = |\psi_n^{(0)}\rangle$.

The job of perturbation theory then reduces to the calculation of $E_n^{(1)}$, $E_n^{(2)}$, $\cdots$ and $|\psi_n^{(1)}\rangle$, $|\psi_n^{(2)}\rangle$, $\cdots$. We shall only be concerned in this section with the determination of the first and second corrections to the energy, as well as the first correction to the eigenkets.

As we assumed that the unperturbed states were non-degenerate, we can write:

$$E_n(\lambda)|\psi_n(\lambda)\rangle = \left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots\right)\left(|\psi_n^{(0)}\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle + \cdots\right) =$$
$$= E_n^{(0)}|\psi_n^{(0)}\rangle + \left[E_n^{(0)}|\psi_n^{(1)}\rangle + E_n^{(1)}|\psi_n^{(0)}\rangle\right]\lambda +$$
$$+ \left[E_n^{(0)}|\psi_n^{(2)}\rangle + E_n^{(1)}|\psi_n^{(1)}\rangle + E_n^{(2)}|\psi_n^{(0)}\rangle\right]\lambda^2 + \cdots = \tag{256}$$
$$= E_n^{(0)}|\psi_n^{(0)}\rangle + \sum_{a=1}^{\infty}\left(\lambda^a \sum_{i=0}^{a}\sum_{j=0}^{a} E_n^{(i)}|\psi_n^{(j)}\rangle\,\delta_{i+j,a}\right)$$

$$H(\lambda)|\psi_n(\lambda)\rangle = (H_0 + \lambda W)\left(|\psi_n^{(0)}\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle + \cdots\right) =$$
$$= H_0|\psi_n^{(0)}\rangle + \left[H_0|\psi_n^{(1)}\rangle + W|\psi_n^{(0)}\rangle\right]\lambda + \left[H_0|\psi_n^{(2)}\rangle + W|\psi_n^{(1)}\rangle\right]\lambda^2 + \cdots = \tag{257}$$
$$= H_0|\psi_n^{(0)}\rangle + \sum_{a=1}^{\infty}\left(\lambda^a\left[H_0|\psi_n^{(a)}\rangle + W|\psi_n^{(a-1)}\rangle\right]\right)$$

Therefore, as $E_n(\lambda)|\psi_n(\lambda)\rangle = H(\lambda)|\psi_n(\lambda)\rangle$, we have:

$$E_n^{(0)}|\psi_n^{(0)}\rangle + \sum_{a=1}^{\infty}\left(\lambda^a\sum_{i=0}^{a}\sum_{j=0}^{a}E_n^{(i)}|\psi_n^{(j)}\rangle\,\delta_{i+j,a}\right) = H_0|\psi_n^{(0)}\rangle + \sum_{a=1}^{\infty}\left(\lambda^a\left[H_0|\psi_n^{(a)}\rangle + W|\psi_n^{(a-1)}\rangle\right]\right) \tag{258}$$

This gives us all the equations to solve in order to find the coefficients of the expansion:

- Zero order[38] in λ:

$$H_0 \, |\psi_n^{(0)}\rangle = E_n^{(0)} \, |\psi_n^{(0)}\rangle \tag{259}$$

- First order in λ:

$$H_0 \, |\psi_n^{(1)}\rangle + W \, |\psi_n^{(0)}\rangle = E_n^{(0)} \, |\psi_n^{(1)}\rangle + E_n^{(1)} \, |\psi_n^{(0)}\rangle \tag{260}$$

- Second order in λ:

$$H_0 \, |\psi_n^{(2)}\rangle + W \, |\psi_n^{(1)}\rangle = E_n^{(0)} \, |\psi_n^{(2)}\rangle + E_n^{(1)} \, |\psi_n^{(1)}\rangle + E_n^{(2)} \, |\psi_n^{(0)}\rangle \tag{261}$$

$$\vdots$$

- a-th order in λ:

$$H_0 \, |\psi_n^{(a)}\rangle + W \, |\psi_n^{(a-1)}\rangle = \sum_{i=0}^{a} \sum_{j=0}^{a} E_n^{(i)} \, |\psi_n^{(j)}\rangle \, \delta_{i+j,a} \tag{262}$$

Note that, since $|\psi_n^{(0)}\rangle \simeq |\psi_n(\lambda)\rangle$, we can normalise $|\psi_n(\lambda)\rangle$ so that $\langle \psi_n^{(0)}|\psi_n(\lambda)\rangle = 1$. Then:

$$1 = \langle \psi_n^{(0)}|\psi_n(\lambda)\rangle = \langle \psi_n^{(0)}|\psi_n^{(0)}\rangle + \lambda \, \langle \psi_n^{(0)}|\psi_n^{(1)}\rangle + \lambda^2 \, \langle \psi_n^{(0)}|\psi_n^{(2)}\rangle + \cdots = 1 + \sum_{k=1}^{\infty} \langle \psi_n^{(0)}|\psi_n^{(k)}\rangle \, \lambda^k \tag{263}$$

Therefore:

$$\sum_{k=1}^{\infty} \langle \psi_n^{(0)}|\psi_n^{(k)}\rangle \, \lambda^k = 0 \quad \longrightarrow \quad \langle \psi_n^{(0)}|\psi_n^{(k)}\rangle = 0, \quad \forall k = 1, 2, 3, ... \tag{264}$$

4.2.1 First order correction

If we multiply both sides of **Equation 260** by $\langle \psi_n^{(0)}|$, we find[39]:

$$\langle \psi_n^{(0)}|H_0|\psi_n^{(1)}\rangle + \langle \psi_n^{(0)}|W|\psi_n^{(0)}\rangle = \langle \psi_n^{(0)}|E_n^{(0)}|\psi_n^{(1)}\rangle + \langle \psi_n^{(0)}|E_n^{(1)}|\psi_n^{(0)}\rangle$$

$$\longrightarrow E_n^{(0)} \underbrace{\langle \psi_n^{(0)}|\psi_n^{(1)}\rangle}_{0} + \langle \psi_n^{(0)}|W|\psi_n^{(0)}\rangle = E_n^{(0)} \underbrace{\langle \psi_n^{(0)}|\psi_n^{(1)}\rangle}_{0} + E_n^{(1)} \underbrace{\langle \psi_n^{(0)}|\psi_n^{(0)}\rangle}_{1} \tag{265}$$

$$\longrightarrow \langle \psi_n^{(0)}|W|\psi_n^{(0)}\rangle = E_n^{(1)}$$

Therefore:

$$E_n^{(1)} = \langle \psi_n^{(0)}|W|\psi_n^{(0)}\rangle \tag{266}$$

This energy shift must be small compared to the level spacing in the unperturbed system. Note that, for some systems, the first order correction $E_n^{(1)}$ vanishes exactly. In such cases, one needs to consider higher order terms. We will later discuss the second order correction.

In order to determine the first order correction to the eigenstates, we can take advantage of the fact that the unperturbed states form a complete orthonormal basis:

$$|\psi_n^{(1)}\rangle = \left(\sum_m |\psi_m^{(0)}\rangle \langle \psi_m^{(0)}| \right) |\psi_n^{(1)}\rangle = \sum_m \langle \psi_m^{(0)}|\psi_n^{(1)}\rangle \, |\psi_m^{(0)}\rangle = \sum_{m \neq n} \langle \psi_m^{(0)}|\psi_n^{(1)}\rangle \, |\psi_m^{(0)}\rangle \tag{267}$$

Note that the term $m = n$ does not contribute as, from **Equation 264**, we know that $\langle \psi_n^{(0)}|\psi_n^{(1)}\rangle = 0$. Multiplying **Equation 260** on both sides by $|\psi_m^{(0)}\rangle$, we can find the coefficient $\langle \psi_m^{(0)}|\psi_n^{(1)}\rangle$:

$$\langle \psi_m^{(0)}|H_0|\psi_n^{(1)}\rangle + \langle \psi_m^{(0)}|W|\psi_n^{(0)}\rangle = \langle \psi_m^{(0)}|E_n^{(0)}|\psi_n^{(1)}\rangle + \langle \psi_m^{(0)}|E_n^{(1)}|\psi_n^{(0)}\rangle$$

$$\longrightarrow E_m^{(0)} \, \langle \psi_m^{(0)}|\psi_n^{(1)}\rangle + \langle \psi_m^{(0)}|W|\psi_n^{(0)}\rangle = E_n^{(0)} \, \langle \psi_m^{(0)}|\psi_n^{(1)}\rangle + E_n^{(1)} \underbrace{\langle \psi_m^{(0)}|\psi_n^{(0)}\rangle}_{0} \tag{268}$$

$$\longrightarrow \langle \psi_m^{(0)}|W|\psi_n^{(0)}\rangle = \left(E_n^{(0)} - E_m^{(0)} \right) \langle \psi_m^{(0)}|\psi_n^{(1)}\rangle$$

---

[38]The zero order correction is trivial, as it is just the eigenvalue equation for the unperturbed system.

[39]Remember that H_0 is Hermitian, so it can be applied to its eigenstate $\psi_n^{(0)}$ on the left.

Therefore:

$$\langle \psi_m^{(0)} | \psi_n^{(1)} \rangle = \frac{\langle \psi_m^{(0)} | W | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \tag{269}$$

And, finally:

$$|\psi_n^{(1)}\rangle = \sum_{m \neq n} \frac{\langle \psi_m^{(0)} | W | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} |\psi_m^{(0)}\rangle \tag{270}$$

Unlike the formula for the energy shift, the correction to the eigenstates involves in general an infinite sum mixing all the states of the unperturbed system. We speak of the perturbation mixing the unperturbed eigenstates.

We can now completely define the first order correction:

> **Definition 22.** The first order correction to the energy is defined as:
>
> $$E_n^{(1)} = \langle \psi_n^{(0)} | W | \psi_n^{(0)} \rangle \tag{271}$$
>
> where $W = \lambda^{-1} H_p$, so that the total energy corrected to the first order is:
>
> $$E_n = E_n^{(0)} + \langle \psi_n^{(0)} | H_p | \psi_n^{(0)} \rangle \tag{272}$$
>
> Furthermore, the first order correction to the eigenstate is given by:
>
> $$|\psi_n^{(1)}\rangle = \sum_{m \neq n} \frac{\langle \psi_m^{(0)} | W | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} |\psi_m^{(0)}\rangle \tag{273}$$
>
> where again we have $W = \lambda^{-1} H_p$, so that the total corrected eigenket is:
>
> $$|\psi_n\rangle = |\psi_n^{(0)}\rangle + \sum_{m \neq n} \frac{\langle \psi_m^{(0)} | H_p | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} |\psi_m^{(0)}\rangle \tag{274}$$

4.2.2 Second order correction

If we multiply both sides of **Equation 261** by $\langle \psi_n^{(0)} |$, we find[40]:

$$\langle \psi_n^{(0)} | H_0 | \psi_n^{(2)} \rangle + \langle \psi_n^{(0)} | W | \psi_n^{(1)} \rangle = \langle \psi_n^{(0)} | E_n^{(0)} | \psi_n^{(2)} \rangle + \langle \psi_n^{(0)} | E_n^{(1)} | \psi_n^{(1)} \rangle + \langle \psi_n^{(0)} | E_n^{(2)} | \psi_n^{(0)} \rangle$$

$$\longrightarrow E_n^{(0)} \underbrace{\langle \psi_n^{(0)} | \psi_n^{(2)} \rangle}_{0} + \langle \psi_n^{(0)} | W | \psi_n^{(1)} \rangle = E_n^{(0)} \underbrace{\langle \psi_n^{(0)} | \psi_n^{(2)} \rangle}_{0} + E_n^{(1)} \underbrace{\langle \psi_n^{(0)} | \psi_n^{(1)} \rangle}_{0} + E_n^{(2)} \underbrace{\langle \psi_n^{(0)} | \psi_n^{(0)} \rangle}_{1}$$

$$\longrightarrow \langle \psi_n^{(0)} | W | \psi_n^{(1)} \rangle = E_n^{(2)} \tag{275}$$

Therefore:

$$E_n^{(2)} = \langle \psi_n^{(0)} | W | \psi_n^{(1)} \rangle \tag{276}$$

Combining this expression with **Equation 273** (and using that W is Hermitian), we can obtain:

$$E_n^{(2)} = \langle \psi_n^{(0)} | W \left(\sum_{m \neq n} \frac{\langle \psi_m^{(0)} | W | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} |\psi_m^{(0)}\rangle \right) = \sum_{m \neq n} \frac{\langle \psi_m^{(0)} | W | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \langle \psi_n^{(0)} | W | \psi_m^{(0)} \rangle =$$

$$= \sum_{m \neq n} \frac{\langle \psi_m^{(0)} | W | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \langle \psi_m^{(0)} | W | \psi_n^{(0)} \rangle^* = \sum_{m \neq n} \frac{\left| \langle \psi_m^{(0)} | W | \psi_n^{(0)} \rangle \right|^2}{E_n^{(0)} - E_m^{(0)}} \tag{277}$$

---

[40]Remember that H_0 is Hermitian, so it can be applied to its eigenstate $\phi_n^{(0)}$ on the left.

Finally, we can define:

4.3 Degenerate perturbation theory

The method we described before is valid for cases in which the eigenenergies are non-degenerate. However, for perturbing degenerate eigenenergies and their corresponding eigenstates, we must apply a different theory[41]. As we will see, the effect of a perturbation on a degenerate level is to remove some or all of the degeneracy. We say that the degeneracy is *lifted*. In other words, the perturbation can (it may not) split the level $E_p^{(0)}$ into distinct sublevels.

Let us consider a g_p-degenerate eigenvalue, $E_p^{(0)}$, of the unperturbed Hamiltonian, H_0, with eigenstates $|\psi_p^{i\,(0)}\rangle$, where $i = 1, 2, 3, \ldots, g_p$ is the index identifying each of the g_p degenerate states[42]. The eigenvalue equation for the degenerate eigenenergy is written as:

$$H_0 |\psi_p^{i\,(0)}\rangle = E_p^{(0)} |\psi_p^{i\,(0)}\rangle \tag{280}$$

The eigenstates $|\psi_p^{i\,(0)}\rangle$ corresponding to $E_p^{(0)}$ form a degenerate g_p-dimensional subpace of the full state space, in which they conform a complete orthonormal basis:

$$\langle \psi_p^{i\,(0)} | \psi_p^{j\,(0)} \rangle = \delta_{ij} \qquad \sum_{i=1}^{g_p} |\psi_p^{i\,(0)}\rangle \langle \psi_p^{i\,(0)}| = \mathbb{1} \tag{281}$$

Note that any set of g_p linearly independent vectors in the degenerate subspace will conform a basis for the degenerate subspace. The choice is arbitrary until we subject the wave functions to the requirement that the change in them under the action of the applied perturbation should be small. Therefore, the zero order approximation consists in finding the correct linear combination of the initial vectors $\{|\psi_n^{i\,(0)}\rangle\}$ such that this condition is met. For a general zero order correction of the eigenvectors of the degenerate subspace, we can write[43]:

$$|\psi_n^{(0)}\rangle = \sum_{k=1}^{g_p} a_k |\psi_p^{k\,(0)}\rangle \tag{282}$$

From this point to the end of this section, we will use the index p to denote the initial eigenstates $\{|\psi_p^{i\,(0)}\rangle\}$ corresponding to the degenerate eigenvalue $E_p^{(0)}$ that were obtained directly from solving the unperturbed

---

[41]Note that, if we have only *some* degenerate eigenvalues, then we apply degenerate perturbation theory only to those degenerate eigenenergies and their states, and we apply normal perturbation theory to the remaining non-degenerate eigenvalues and eigenstates. Check

[42]The i stands for one or more quanutm numbers that the energy is independent of.

[43]This index-less $|\psi_n^{(0)}\rangle$ represents the *general "shape"* of the zero order correction to the eigenstates. We will later find that, as a result of our calculations, this separates into g_p distinct definitions for each of the g_p eigenstates of the zero correction. For now, however, it makes no sense to carry an index labeling each of those, as we cannot yet distinguish between them.

system. We will use the index n to denote the zero order approximation of the eigenstates $\{|\psi_n^{i\,(0)}\rangle\}$ of the perturbed Hamiltonian, which are just a linear combination of $\{|\psi_p^{i\,(0)}\rangle\}$. By definition, $\{|\psi_n^{i\,(0)}\rangle\}$ are eigenstates of the unperturbed system and almost or complete? eigenstates of the perturbed system.

Considering orthonormality, we can determine the following relation concerning the coefficients:

$$1 = \langle \psi_n^{k\,(0)}|\psi_n^{k\,(0)}\rangle = \sum_{k=1}^{g_p} |a_k|^2 \tag{283}$$

Let us now rewrite the eigenvalue equation for the perturbed Hamiltonian using the expression from **Equation 282** Why do we introduce directly $|\psi_n^{(0)}\rangle$? This is related to the above question:

$$H|\psi_n^{(0)}\rangle = E_n|\psi_n^{(0)}\rangle \longrightarrow (H_0 + H_p)\sum_{k=1}^{g_p} a_k|\psi_p^{k\,(0)}\rangle = E_n \sum_{k=1}^{g_p} a_k|\psi_p^{k\,(0)}\rangle$$

$$\longrightarrow \sum_{k=1}^{g_p} a_k\left[H_0|\psi_p^{k\,(0)}\rangle + H_p|\psi_p^{k\,(0)}\rangle\right] = E_n \sum_{k=1}^{g_p} a_k|\psi_p^{k\,(0)}\rangle \tag{284}$$

$$\longrightarrow \sum_{k=1}^{g_p} a_k\left[E_n^{(0)}|\psi_p^{k\,(0)}\rangle + H_p|\psi_p^{k\,(0)}\rangle\right] = E_n \sum_{k=1}^{g_p} a_k|\psi_p^{k\,(0)}\rangle$$

Multiplying both sides of this equation by $\langle \psi_p^{j\,(0)}|$ yields:

$$\sum_{k=1}^{g_p} a_k\left[E_n^{(0)}\langle\psi_p^{j\,(0)}|\psi_p^{k\,(0)}\rangle + \langle\psi_p^{j\,(0)}|H_p|\psi_p^{k\,(0)}\rangle\right] = E_n \sum_{k=1}^{g_p} a_k\langle\psi_p^{j\,(0)}|\psi_p^{k\,(0)}\rangle$$

$$\longrightarrow \sum_{k=1}^{g_p} a_k\left[E_n^{(0)}\delta_{k,j} + \langle\psi_p^{j\,(0)}|H_p|\psi_p^{k\,(0)}\rangle\right] = E_n \sum_{k=1}^{g_p} a_k\delta_{k,j} \tag{285}$$

$$\longrightarrow a_j E_n^{(0)} + \sum_{k=1}^{g_p} a_k\langle\psi_p^{j\,(0)}|H_p|\psi_p^{k\,(0)}\rangle = a_j E_n$$

If we denote $H_{p_{jk}} = \langle\psi_p^{j\,(0)}|H_p|\psi_p^{k\,(0)}\rangle$ and $E_n^{(1)} = E_n - E_n^{(0)}$, we can write this as:

$$\sum_{k=1}^{g_p} a_k\left(H_{p_{jk}} - \delta_{k,j}E_n^{(1)}\right) = 0 \tag{286}$$

which, considering $j = 1, 2, 3, \ldots, g_p$, gives a system of g_p linear equations for the coefficients a_k. These coefficients are non-vanishing only when the determinant $|H_{p_{jk}} - \delta_{k,j}E_n^{(1)}|$ is zero:

$$\begin{vmatrix} H_{p_{11}} - E_n^{(1)} & H_{p_{12}} & \cdots & H_{p_{1g_p}} \\ H_{p_{21}} & H_{p_{22}} - E_n^{(1)} & \cdots & H_{p_{2g_p}} \\ \vdots & \vdots & \ddots & \vdots \\ H_{p_{g_p 1}} & H_{p_{g_p 2}} & \cdots & H_{p_{g_p g_p}} - E_n^{(1)} \end{vmatrix} = 0 \tag{287}$$

This determinant is a g_p-th degree equation in $E_n^{(1)}$, which in general has g_p different roots $E_n^{i\,(1)}$ (some may not be different), each of which is the first order correction to the energy corresponding to the i-th eigenvector of the degenerate eigenvalue $E_n^{(0)}$ of the unperturbed system. Similarly to when we find the characteristic polynomial $P(\lambda)$ of a matrix A and solve $P(\lambda) = 0$ to find the eigenvectors, once we have

the first order corrections of the energy we substitute the different values into the system of equations:

$$
\begin{bmatrix}
H_{p_{11}} - E_n^{i\,(1)} & H_{p_{12}} & \cdots & H_{p_{1g_p}} \\
H_{p_{21}} & H_{p_{22}} - E_n^{i\,(1)} & \cdots & H_{p_{2g_p}} \\
\vdots & \vdots & \ddots & \vdots \\
H_{p_{g_p 1}} & H_{p_{g_p 2}} & \cdots & H_{p_{g_p g_p}} - E_n^{i\,(1)}
\end{bmatrix}
\cdot
\begin{bmatrix}
a_1^i \\
a_2^i \\
\vdots \\
a_{g_p}^i
\end{bmatrix}
=
\begin{bmatrix}
0 \\
0 \\
\vdots \\
0
\end{bmatrix}
\tag{288}
$$

To find the coefficients a_k. The i-th root $E_n^{i\,(1)}$ allows us to find the coefficients a_k^i for the zero order correction of the eigenstate $|\psi_n^{i\,(0)}\rangle$, such that:

$$
|\psi_n^{i\,(0)}\rangle = \sum_{k=1}^{g_p} a_k^i \, |\psi_p^{k\,(0)}\rangle, \qquad i = 1, 2, 3, \ldots, g_p
\tag{289}
$$

In general, the roots $E_n^{i\,(0)}$ are different, so the g_p-degenerate level $E_n^{(0)}$ is split into g_p different levels $E_n^i = E_n^{(0)} + E_n^{i\,(1)}$. In this way, the perturbation lifts the degeneracy, depending on whether all the roots or only some of them are different.

In summary, the degenerate perturbation method consists of the following steps:

1. For each of the g_p-fold degenerate levels, determine the $g_p \times g_p$ matrix of the perturbation H_p:

$$
H_p = \begin{pmatrix}
H_{p_{11}} & H_{p_{12}} & \cdots & H_{p_{1g_p}} \\
H_{p_{21}} & H_{p_{22}} & \cdots & H_{p_{2g_p}} \\
\vdots & \vdots & \ddots & \vdots \\
H_{p_{g_p 1}} & H_{p_{g_p 2}} & \cdots & H_{p_{g_p g_p}}
\end{pmatrix}
\tag{290}
$$

where $H_{p_{jk}} = \langle \psi_p^{j\,(0)} | H_p | \psi_p^{k\,(0)} \rangle$.

2. Diagonalise this matrix and find the g_p eigenvalues $E_n^{i\,(1)}$ $(i = 1, 2, 3, \ldots, g_p)$ and their corresponding eigenvectors:

$$
\vec{a}^{\,i} = \begin{pmatrix}
a_1^i \\
a_2^i \\
\vdots \\
a_{g_p}^i
\end{pmatrix}, \qquad (i = 1, 2, 3, \ldots, g_p)
\tag{291}
$$

3. Finally, the first order corrections to the energy for the g_p eigenstates corresponding to the original degenerate level are given by:

$$
E_n^i = E_n^{(0)} + E_n^{i\,(1)}, \qquad (i = 1, 2, 3, \ldots, g_p)
\tag{292}
$$

And the corresponding eigenvectors are given to order zero as:

$$
|\psi_n^{i\,(0)}\rangle = \sum_{k=1}^{g_p} a_k^i \, |\psi_p^{k\,(0)}\rangle, \qquad i = 1, 2, 3, \ldots, g_p
\tag{293}
$$

Just as for non-degenerate levels, the method described in this section is valid only if the matrix elements of the perturbation are much smaller than the differences between the energy of the level under study and those of the other levels (this conclusion would have been evident if we had calculated higher-order corrections).

4.3.1 The Stark effect

The Stark effect is an example where degenerate perturbation can be applied to solve a quantum mechanical problem. It is related to the interaction of the electric dipole of the hydrogen atom with an external electric field.

In the absence of an external magnetic field, the first excited state of the hydrogen atom, $n = 2$, is fourfold degenerate: the states $|n, l, m\rangle = |200\rangle, |210\rangle, |211\rangle, |21-1\rangle$ have the same energy $E_2 = -R_y/4$, where $R_y = \mu e^4/(2\hbar^2) = 13.6 eV$ is the Rydberg constant.

Consider now a hydrogen atom placed in a uniform static electric field $\vec{\mathcal{E}}$ parallel to the z axis. To the Hamiltonian of the Hydrogen atom, we must add the Stark Hamiltonian W_S, which describes the interaction energy of the electric dipole of the atom $\vec{d} = -e\vec{r}$ with the electric field $\vec{\mathcal{E}} = \mathcal{E}\vec{e}_z$. We can write W_S as:

$$W_S = -\vec{d} \cdot \vec{\mathcal{E}} = e\vec{r} \cdot \vec{\mathcal{E}} = e\mathcal{E}Z \tag{294}$$

To calculate the first order correction to the energies, we need to determine and then diagonalise the 4×4 matrix elements $W_{s_{l'm'lm}} = \langle 2l'm'|W_s|2lm\rangle$. For calculating the matrix elements[44]:

$$\langle 2l'm'|W_s|2lm\rangle = e\mathcal{E}\,\langle 2l'm'|Z|2lm\rangle =$$
$$= e\mathcal{E} \int_0^\infty dr \int_0^\pi d\theta \int_0^{2\pi} d\phi \cdot r^2 \sin\theta \cdot \left[\psi_{l',m'}(r,\theta,\phi)\right]^* \cdot r\cos\theta \cdot \psi_{l,m}(r,\theta,\phi) = \tag{295}$$
$$= e\mathcal{E} \int_0^\infty dr \cdot r^3 \int_0^\pi d\theta \cdot \sin\theta\cos\theta \int_0^{2\pi} d\phi \cdot \left[\psi_{l',m'}(r,\theta,\phi)\right]^* \cdot \psi_{l,m}(r,\theta,\phi)$$

Using that $\psi_{l,m}(r,\theta,\phi) = R_l(r)Y_l^m(\theta,\phi)$, we obtain:

$$\langle 2l'm'|W_s|2lm\rangle = e\mathcal{E} \int_0^\infty dr \cdot r^3 \int_0^\pi d\theta \cdot \sin\theta\cos\theta \int_0^{2\pi} d\phi \cdot \left[R_{l'}(r)Y_{l'}^{m'}(\theta,\phi)\right]^* \cdot R_l(r)Y_l^m(\theta,\phi) =$$
$$= e\mathcal{E} \int_0^\infty dr \cdot r^3 R_{l'}^*(r)R_l(r) \int_0^\pi d\theta \cdot \sin\theta\cos\theta \int_0^{2\pi} d\phi \left[Y_{l'}^{m'}(\theta,\phi)\right]^* Y_l^m(\theta,\phi) \tag{296}$$

From here, we can easily see that the only non-zero matrix elements are those for which $m' = m$[45]. Therefore, the only non-vanishing matrix elements will be those between any combination of the states $|200\rangle$ and $|210\rangle$. Furthermore, taking into account the parity of spherical harmonics, from **Equation 188**, and the fact that the operator Z is odd, we can see that the only non-zero matrix elements are those for which $l' = l \pm 1$. Therefore, the only matrix element that does not vanish (its complex conjugate matrix element does not vanish either, of course), is:

$$\langle 200|W_s|210\rangle = e\mathcal{E} \int_0^\infty dr \cdot r^3 R_0^*(r)R_1(r) \int_0^\pi d\theta \cdot \sin\theta\cos\theta \int_0^{2\pi} d\phi \left[Y_0^0(\theta,\phi)\right]^* Y_1^0(\theta,\phi) = -3e\mathcal{E}a_0 \tag{297}$$

where $a_0 = \frac{\hbar^2}{\mu e^2}$ is the Bohr radius. Therefore, the matrix of the perturbation is:

$$W_s = -3e\mathcal{E}a_0 \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{298}$$

Diagonalizing this matrix leads to the following eigenvalues:

$$E_2^{1(1)} = -3e\mathcal{E}a_0, \qquad E_2^{2(1)} = E_2^{3(1)} = 0, \qquad E_2^{4(1)} = 3e\mathcal{E}a_0 \tag{299}$$

---

[44]We use that $z = r\cos\theta$.

[45]Clarify this, which is stated in Zetilli.

The corrections eigenvectors to the zeroth order are:

$$|\psi_2^{1(0)}\rangle = \frac{1}{\sqrt{2}}\left(|200\rangle + |210\rangle\right), \qquad |\psi_2^{2(0)}\rangle = |211\rangle,$$

$$|\psi_2^{3(0)}\rangle = |21-1\rangle, \qquad |\psi_2^{4(0)}\rangle = \frac{1}{\sqrt{2}}\left(|200\rangle - |210\rangle\right),$$

$$(300)$$

As we can see, this perturbation has only partially removed the degeneracy of the $n = 2$ level; the states $|211\rangle$ and $|21-1\rangle$ still have the same energy.

4.4   The variational method

The variational method is used for finding approximate solutions to problems that cannot be solved exactly. It is based upon the fact that the energy of any trial wave function ψ_T is an upper bound to the true ground state energy of the system. The variational method consists in choosing a trial wave function ψ_T that depends on one or more parameters, and then minimizing the expectation value of the energy with respect to these parameters. The minimum value of the energy is an upper bound to the true ground state energy of the system. The closer the trial wave function is to the true ground state wave function, the closer the upper bound is to the true ground state energy.

Consider a system described by a Hamiltonian H, with a complete set of eigenstates $\{|n\rangle\}$ and corresponding eigenvalues $\{E_n\}$ ordered in increasing value:

$$E_1 < E_2 < E_3 < \dots \tag{301}$$

An arbitrary state $|n\rangle$ can be expanded as a linear superposition

$$|\psi\rangle = \sum_n c_n |n\rangle, \text{ with } c_n = \langle n|\psi\rangle \tag{302}$$

Then, we can compute the expectation value of the energy as:

$$\langle E \rangle = \frac{\langle \psi|H|\psi\rangle}{\langle \psi|\psi\rangle} = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2} \geq \frac{\sum_n |c_n|^2 E_1}{\sum_n |c_n|^2} = E_1 \tag{303}$$

where we have used the fact that $E_n \geq E_1$ for $n > 1$. Therefore, the expectation value of the energy for any state is an upper bound to the ground state energy E_1. This is the basis of the variational method, which gives way to the following procedure:

1. Choose a trial state $|\psi_T\rangle$ that depends on one or more parameters $\alpha_1, \alpha_2, \cdots, \alpha_r$.

2. Compute:

$$E(\alpha_1, \alpha_2, \cdots, \alpha_r) = \frac{\langle \psi_T|H|\psi_T\rangle}{\langle \psi_T|\psi_T\rangle} \tag{304}$$

3. Minimise $E(\alpha_1, \alpha_2, \cdots, \alpha_r)$ with respect to the variational parameters $\alpha_1, \alpha_2, \cdots, \alpha_r$ by solving the equations:

$$\frac{\partial E}{\partial \alpha_i} = 0, \qquad i = 1, 2, \cdots, r \tag{305}$$

4. Evaluate the averge energy of the system using the values of the variational parameters that minimise $E(\alpha_1, \alpha_2, \cdots, \alpha_r)$.

5. The minimum value of the energy that we obtain is an upper bound to the true ground state energy of the system.

A good trial function incorporates as many of the features of the exact wave function as possible, some of which can be deduced from symmetry arguments. Note that the variational method is aimed at optimising the estimate of the *energy* of the energy. It may not necessarily give a good approximation to the true *wave function*, and could give poor results if used to calculate things other than the energy.

5 Introduction to Light Matter Interaction and Time Dependent Perturbation Theory

So far, We have dealt with Hamiltonians that do not depend explicitly on time. In nature, however, most quantum phenomena are governed by time-dependent Hamiltonians. In this chapter we are going to consider approximation methods treating Hamiltonians that depend explicitly on time.

To study the structure of molecular and atomic systems, we need to know how electromagnetic radiation interacts with these systems. Molecular and atomic spectroscopy deals in essence with the absorption and emission of electromagnetic radiation by molecules and atoms. As a system absorbs or emits radiation, it undergoes transitions from one state to another.

Time-dependent perturbation theory is most useful for studying processes of absorption and emission of radiation by atoms or, more generally, for treating the transitions of quantum systems from one energy level to another.

5.1 Time dependent perturbation theory

For systems with time-dependent Hamiltonians the energy is not conserved and thus there are no stationary states. Solving the Schrödinger equations is then very difficult. However, in the case where the time-dependent part is small, progress can be made applying **time dependent perturbation theory**. One example of this is the interaction of a quantum system, like an atom, with electromagnetic radiation. The objective of this method is to compute how the system transitions in time between states in an initially prescribed basis (usually the eigenstates of the time-independent free Hamiltonian).

Let us assume that we have a system with an unperturbed and time-independent Hamiltonian H_0. Its spectrum is discrete and non-degenerate:

$$H_0 \left| \phi_n \right\rangle = E_n \left| \phi_n \right\rangle, \tag{306}$$

and the eigenstates form a complete orthonormal basis:

$$\left\langle \phi_n | \phi_p \right\rangle = \delta_{np} \qquad \sum_n \left| \phi_n \right\rangle \left\langle \phi_n \right| = \mathbb{1} \tag{307}$$

Recall also that the complete stationary states can be written as:

$$\left| \Phi_n(t) \right\rangle = e^{-itH_0/\hbar} \left| \phi_n \right\rangle = e^{-iE_n t/\hbar} \left| \phi_n \right\rangle \tag{308}$$

During the time interval $0 < t < \tau$, a small perturbation is introduced so that the new Hamiltonian for the duration of this interval is:

$$H(t) = H_0 + H_p(t) = H_0 + \lambda W(t) \tag{309}$$

where $H_p(t) = \lambda W(t)$ for $\lambda \ll 1$. When the system interacts with $H_p(t)$, it either absorbs or emits energy, which inevitably causes the system to undergo transitions from one unperturbed eigenstate to another. At time $t < 0$, the system is in the stationary state $\left| \phi_i \right\rangle$. However, at time $t > 0$, the system evolves and can be found in another state. The question that arises naturally is: what is the probability $\mathcal{P}_{if}$ of finding the system at time t in another eigenstate $\left| \phi_f \right\rangle$ of the unperturbed Hamiltonian of H_0?

The evolution of the system is described by the Schrödinger equation:

$$i\hbar \frac{d \left| \psi(t) \right\rangle}{dt} = (H_0 + \lambda W(t)) \left| \psi(t) \right\rangle \tag{310}$$

with the initial condition given by:

$$\left| \psi(0) \right\rangle = \left| \phi_i \right\rangle \tag{311}$$

At time t, we can obtain the probability of finding the system in a state $|\phi_f\rangle$, which we call $\mathcal{P}_{if}$, in the usual manner; by finding the magnitude square of the projection of the state $|\psi(t)\rangle$ onto the state $|\phi_f\rangle$:

$$\mathcal{P}_{if} = |\langle\phi_f|\psi(t)\rangle|^2 \tag{312}$$

As $|\phi_f\rangle$ is an eigenvector of H_0, this formula is basically giving us the time dependent coefficient $c_f(t) = \langle\phi_f|\psi(t)\rangle$ corresponding $|\phi_f\rangle$ for the expansion of $|\psi(t)\rangle$ in the basis of the eigenstates of H_0, at time t:

$$|\psi(t)\rangle = \sum_n c_n(t)\,|\phi_n\rangle \tag{313}$$

Using this expansion together with the closure relation in the Schrödinger equation, we obtain:

$$
\begin{aligned}
i\hbar\frac{d}{dt}\left(\sum_n c_n(t)\,|\phi_n\rangle\right) &= \left(\sum_n |\phi_n\rangle\langle\phi_n|\right)(H_0+\lambda W(t))\,|\psi(t)\rangle \\
i\hbar\sum_n \frac{dc_n(t)}{dt}\,|\phi_n\rangle &= \left(\sum_n |\phi_n\rangle\langle\phi_n|\,(H_0+\lambda W(t))\,|\psi(t)\rangle\right) \\
i\hbar\sum_n \frac{dc_n(t)}{dt}\,|\phi_n\rangle &= \left(\sum_n |\phi_n\rangle\langle\phi_n|\,(H_0+\lambda W(t))\left(\sum_k c_k(t)\,|\phi_k\rangle\right)\right) \\
i\hbar\sum_n \frac{dc_n(t)}{dt}\,|\phi_n\rangle &= \sum_{n,k} |\phi_n\rangle\langle\phi_n|H_0|\phi_k\rangle\,c_k(t) + \lambda\sum_{n,k} |\phi_n\rangle\langle\phi_n|W(t)|\phi_k\rangle\,c_k(t)
\end{aligned}
\tag{314}
$$

Using now the fact that $|\phi_k\rangle$ are eigenstates of the unperturbed Hamiltonian, we can write:

$$
\begin{aligned}
i\hbar\sum_n \frac{dc_n(t)}{dt}\,|\phi_n\rangle &= \sum_{n,k} |\phi_n\rangle\,E_k\langle\phi_n|\phi_k\rangle\,c_k(t) + \lambda\sum_{n,k} |\phi_n\rangle\langle\phi_n|W(t)|\phi_k\rangle\,c_k(t) \\
i\hbar\sum_n \frac{dc_n(t)}{dt}\,|\phi_n\rangle &= \sum_{n,k} |\phi_n\rangle\,[E_k\langle\phi_n|\phi_k\rangle + \lambda\langle\phi_n|W(t)|\phi_k\rangle]\,c_k(t) \\
i\hbar\sum_n \frac{dc_n(t)}{dt}\,|\phi_n\rangle &= \sum_{n,k} |\phi_n\rangle\,[E_k\delta_{nk} + \lambda\langle\phi_n|W(t)|\phi_k\rangle]\,c_k(t)
\end{aligned}
\tag{315}
$$

Therefore, the first in the sum appears only for $n = k$ and we can write:

$$i\hbar\sum_n \frac{dc_n(t)}{dt}\,|\phi_n\rangle = \sum_n |\phi_n\rangle\left(E_n c_n(t) + \lambda\sum_k \langle\phi_n|W(t)|\phi_k\rangle\,c_k(t)\right) \tag{316}$$

We can now split this into different equations for each n:

$$i\hbar\frac{dc_n(t)}{dt} = E_n c_n(t) + \lambda\sum_k \langle\phi_n|W(t)|\phi_k\rangle\,c_k(t) \tag{317}$$

Defining the matrix elements of $W(t)$ as:

$$W_{nk}(t) \equiv \langle\phi_n|W(t)|\phi_k\rangle \tag{318}$$

we can write:

$$i\hbar\frac{dc_n(t)}{dt} = E_n c_n(t) + \lambda\sum_k W_{nk}(t)c_k(t) \tag{319}$$

These equations, written for the various n, constitute a set of coupled linear differential equations of first order in t, which enables us, in theory to determine the components $c_n(t)$.

When the perturbation $\lambda W(t)$ is zero, the equations are no longer coupled, and we can easily find the solution:

$$c_n(t) = b_0\,e^{-iE_n t/\hbar} \tag{320}$$

With a non-zero perturbation, but much smaller than H_0, we expect solutions close to the previous one. Thus, we can look for the solution in the form:

$$c_n(t) = b_n(t) \, e^{-iE_n t/\hbar} \tag{321}$$

where the functions $b_n(t)$ will be slowly varying functions of time. Substituting this into **Equation 319**, we obtain:

$$i\hbar \frac{d}{dt}\left(b_n(t) \, e^{-iE_n t/\hbar}\right) = E_n b_n(t) \, e^{-iE_n t/\hbar} + \lambda \sum_k W_{nk}(t)b_k(t) \, e^{-iE_k t/\hbar}$$

$$i\hbar \frac{db_n(t)}{dt}e^{-iE_n t/\hbar} + i\hbar \frac{d\left(e^{-iE_n t/\hbar}\right)}{dt}b_n(t) = E_n b_n(t) e^{-iE_n t/\hbar} + \lambda \sum_k W_{nk}(t)b_k(t) \, e^{-iE_k t/\hbar}$$

$$i\hbar \frac{db_n(t)}{dt}e^{-iE_n t/\hbar} = \lambda \sum_k W_{nk}(t)b_k(t) \, e^{-iE_k t/\hbar} \tag{322}$$

$$i\hbar \frac{db_n(t)}{dt} = \lambda \sum_k e^{i(E_n - E_k)t/\hbar}W_{nk}(t)b_k(t)$$

Finally, if we define:

$$\omega_{nk} \equiv \frac{E_n - E_k}{\hbar}, \tag{323}$$

we can write:

$$i\hbar \frac{db_n(t)}{dt} = \lambda \sum_k e^{i\omega_{nk}t}W_{nk}(t)b_k(t) \tag{324}$$

This equation is equivalent to the Schrödinger equation for the problem we are trying to solve. Note, however, that this is an approximate equation, since we have assumed the shape of the solution for small perturbations.

We will now apply the same method as in time independent perturbation methods to find the q-th order approximation of the coefficients $b_n(t)$ using the power expansion:

$$b_n(t) = b_n^{(0)}(t) + \lambda b_n^{(1)}(t) + \lambda^2 b_n^{(2)}(t) + \cdots = \sum_{q=0}^{\infty} \lambda^q b_n^{(q)}(t) \tag{325}$$

Plugging this into **Equation 324**, we obtain:

$$i\hbar \frac{d}{dt}\left(\sum_{q=0}^{\infty} \lambda^q b_n^{(q)}(t)\right) = \lambda \sum_k e^{i\omega_{nk}t}W_{nk}(t)\sum_{q=0}^{\infty} \lambda^q b_k^{(q)}(t)$$

$$i\hbar \sum_{q=0}^{\infty} \lambda^q \frac{db_n^{(q)}(t)}{dt} = \sum_{q=0}^{\infty}\sum_k \lambda^{q+1}e^{i\omega_{nk}t}W_{nk}(t)b_k^{(q)}(t) \tag{326}$$

Now, matching the terms of equal power in λ, we can obtain a set of equations that give us the corrections of $b_n(t)$ at the different orders. For example, for the zeroth order (terms multiplying λ^0), we have:

$$i\hbar \frac{db_n^{(0)}(t)}{dt} = 0, \tag{327}$$

which means that $b_n^{(0)}(t)$ are constant functions:

$$b_n^{(0)} \equiv \text{const.} \tag{328}$$

As we are initially in the state $|\phi_i\rangle$, these coefficients at time $t = 0$, and for any t (as they are constant), must be equal to δ_{ni}.

$$b_n^{(0)} = \delta_{ni} \tag{329}$$

This completely determines the zeroth order solution:

$$|\psi^{(0)}(t)\rangle = \sum_n c_n(t)\,|\phi_n\rangle = \sum_n b_n(t)\,e^{-iE_nt/\hbar}\,|\phi_n\rangle = \sum_n b_n^{(0)}\,e^{-iE_nt/\hbar}\,|\phi_n\rangle =$$
$$= \sum_n \delta_{ni}\,e^{-iE_nt/\hbar}\,|\phi_n\rangle = \ e^{-iE_it/\hbar}\,|\phi_i\rangle\,, \tag{330}$$

which is simply the initial stationary state we are starting from.

For higher orders, we need to apply this recursive formula, which is found from **Equation 326** by taking the q-th order on the left hand side and the $(q-1)$-th order on the right hand side, to match the powers of λ:

$$i\hbar\frac{db_n^{(q)}(t)}{dt} = \sum_k e^{i\omega_{nk}t}W_{nk}(t)b_k^{(q-1)}(t) \tag{331}$$

Therefore, from the zeroth order solution, we can recursively restore solutions for the higher orders.

We can then write:

Definition 24. Consider a system with an unperturbed and time-independent Hamiltonian H_0 whose spectrum is discrete and non-degenerate:

$$H_0\,|\phi_n\rangle = E_n\,|\phi_n\rangle\,, \tag{332}$$

and which is initially in the state $|\psi(t=0)\rangle = |\phi_i\rangle$. During the time interval $0 < t < \tau$, it is exposed to a small time dependent perturbation $H_p(t) = \lambda W(t)$. The probability $\mathcal{P}_{if}$ of finding the system in a state $|\phi_f\rangle$ at time t is given by:

$$\mathcal{P}_{if} = |\langle\phi_f|\psi(t)\rangle|^2 = |c_f(t)|^2 \tag{333}$$

where the coefficients $c_n(t)$ are of the form:

$$c_n(t) = b_n(t)\,e^{-iE_nt/\hbar} \tag{334}$$

and the functions $b_n(t)$ are given at q-th order by the recursive formula:

$$i\hbar\frac{db_n^{(q)}(t)}{dt} = \sum_k e^{i\omega_{nk}t}W_{nk}(t)b_k^{(q-1)}(t) \tag{335}$$

starting from the constant zeroth order term given by:

$$b_n^{(0)} = \delta_{ni} \tag{336}$$

5.1.1   First order solution

The initial state of the system is $|\phi_i\rangle$. Therefore, in the initial instant when the perturbation is applied the only term in the expansion of the $b_n(t)$ coefficients that is non-zero, for any n-eigenstate and q-order, is the zeroth order coefficient for $n = i$. In other words:

$$\begin{cases} b_n^{(0)}(0) = \ \delta_{ni} \\ b_n^{(q)}(0) = \ 0, \quad \forall q > 0 \end{cases} \tag{337}$$

Then, using **Equation 335**, we obtain:

$$i\hbar\frac{db_n^{(1)}(t)}{dt} = \sum_k e^{i\omega_{nk}t}W_{nk}(t)b_k^{(0)}(t) = \sum_k e^{i\omega_{nk}t}W_{nk}(t)\delta_{ki} = e^{i\omega_{ni}t}W_{ni}(t) \tag{338}$$

Finally, we obtain:

$$b_n^{(1)}(t) = \frac{1}{i\hbar} \int_0^t e^{i\omega_{ni}t'} W_{ni}(t')dt' \tag{339}$$

Recalling the expression for the probability we are trying to find:

$$\mathcal{P}_{if}(t) = |\langle\phi_f|\psi(t)\rangle|^2 = |c_f(t)|^2 = |b_f(t)|^2 = \left|b_f^{(0)}(t) + \lambda b_f^{(1)}(t) + \cdots\right|^2 \tag{340}$$

Since we found that $b_f^{(0)}(t) = \delta_{fi}$, when we consider $i \neq f$ the expression for the probability $\mathcal{P}_{if}(t)$ to the first order is reduced to:

$$\mathcal{P}_{if}^{(1)}(t) = \lambda^2 \left|b_f^{(1)}(t)\right|^2 \tag{341}$$

We can then write:

Definition 25. The first order term in the time dependent perturbation method is given by:

$$b_n^{(1)}(t) = \frac{1}{i\hbar} \int_0^t e^{i\omega_{ni}t'} W_{ni}(t')dt' \tag{342}$$

And the corresponding probability to the first order is given by:

$$\mathcal{P}_{if}^{(1)}(t) = \lambda^2 \left|b_f^{(1)}(t)\right|^2 = \frac{1}{\hbar^2}\left|\int_0^t e^{i\omega_{fi}t'} H_{pfi}(t')dt'\right|^2 \tag{343}$$

where we recall that $H_p(t) = \lambda W(t)$ and $H_{pfi} = \langle\phi_f|H_p|\phi_i\rangle$.

Note that the probability $\mathcal{P}_{if}^{(1)}$ is none other than the square of the modulus of the Fourier transform of the perturbation matrix element.

5.1.2 An important special case: the sinusoidal perturbation

Let us assume that the perturbation is given by:

$$H_p(t) = H_p \sin(\omega t), \tag{344}$$

where $\omega \geq 0$. In this case, we can write the matrix elements as:

$$H_{pfi}(t) = H_{pfi}\sin(\omega t) = \frac{H_{pfi}}{2i}\left(e^{i\omega t} - e^{-i\omega t}\right), \tag{345}$$

where $H_{pfi} = \langle\phi_f|H_p|\phi_i\rangle$. Now, we can calculate the probability:

$$\mathcal{P}_{if}(t;\omega) \approx \mathcal{P}_{if}^{(1)}(t;\omega) = \frac{1}{\hbar^2}\left|\int_0^t e^{i\omega_{fi}t'} H_{pfi}(t')dt'\right|^2 = \frac{\left|H_{pfi}\right|^2}{4\hbar^2}\left|\int_0^t e^{i\omega_{fi}t'}\left(e^{i\omega t'} - e^{-i\omega t'}\right)dt'\right|^2 =$$

$$= \frac{\left|H_{pfi}\right|^2}{4\hbar^2}\left|\int_0^t \left(e^{i(\omega_{fi}+\omega)t'} - e^{i(\omega_{fi}-\omega)t'}\right)dt'\right|^2$$

$$= \frac{\left|H_{pfi}\right|^2}{4\hbar^2}\left|\left(\frac{e^{i(\omega_{fi}+\omega)t'}}{i(\omega_{fi}+\omega)} - \frac{e^{i(\omega_{fi}-\omega)t'}}{i(\omega_{fi}-\omega)}\right)_0^t\right|^2 = \tag{346}$$

$$= \frac{\left|H_{pfi}\right|^2}{4\hbar^2}\left|\frac{-1}{i}\left(\frac{1-e^{i(\omega_{fi}+\omega)t}}{\omega_{fi}+\omega} - \frac{1-e^{i(\omega_{fi}-\omega)t}}{\omega_{fi}-\omega}\right)\right|^2$$

Finally, the probability is:

$$\mathcal{P}_{if}(t;\omega) \approx \frac{\left|H_{pfi}\right|^2}{4\hbar^2}\left|\frac{1 - e^{i(\omega_{fi}+\omega)t}}{\omega_{fi}+\omega} - \frac{1 - e^{i(\omega_{fi}-\omega)t}}{\omega_{fi}-\omega}\right|^2 \tag{347}$$

Here, we can make an important observation related to the concept of resonance. Consider the case where $\omega \to \omega_{fi}$. In this scenario, the second term of the subtraction in **Equation 347** (called the "resonant term") is much larger than the first, so we can write:

$$\mathcal{P}_{if}^{\omega \to \omega_{fi}}(t;\omega) \approx \frac{\left|H_{pfi}\right|^2}{4\hbar^2}\left|\frac{1 - e^{i(\omega_{fi}-\omega)t}}{\omega_{fi}-\omega}\right|^2 \tag{348}$$

Taking $e^{i(\omega_{fi}-\omega)t/2}$ as common factor in the numerator, we can write:

$$\begin{aligned}
\mathcal{P}_{if}^{\omega \to \omega_{fi}}(t;\omega) &\approx \frac{\left|H_{pfi}\right|^2}{4\hbar^2}\left|\frac{e^{i(\omega_{fi}-\omega)t/2}\left(e^{-i(\omega_{fi}-\omega)t/2} - e^{i(\omega_{fi}-\omega)t/2}\right)}{\omega_{fi}-\omega}\right|^2 = \\
&= \frac{\left|H_{pfi}\right|^2}{4\hbar^2}\left|\frac{e^{i(\omega_{fi}-\omega)t/2}\left(2\sin\frac{(\omega_{fi}-\omega)t}{2}\right)}{\omega_{fi}-\omega}\right|^2
\end{aligned} \tag{349}$$

Finally, we obtain:

$$\mathcal{P}_{if}^{\omega \to \omega_{fi}}(t;\omega) \approx \frac{\left|H_{pfi}\right|^2}{4\hbar^2}\left(\frac{\sin\frac{(\omega_{fi}-\omega)t}{2}}{\frac{\omega_{fi}-\omega}{2}}\right)^2 \tag{350}$$

Similarly, for $\omega \to -\omega_{fi}$, only the first term in the subtraction of **Equation 347** (called the "anti-resonant" term) survives, so we obtain:

$$\mathcal{P}_{if}^{\omega \to -\omega_{fi}}(t;\omega) \approx \frac{\left|H_{pfi}\right|^2}{4\hbar^2}\left(\frac{\sin\frac{(\omega_{fi}+\omega)t}{2}}{\frac{\omega_{fi}+\omega}{2}}\right)^2 \tag{351}$$

If we recall the definition of ω_{fi}:

$$\omega_{fi} = \frac{E_f - E_i}{\hbar}, \tag{352}$$

We can see that, for negative ω_{fi}, we have $E_f < E_i$, and for positive ω_{fi}, we have $E_f > E_i$. In **Figure 5**, we can see the effect of each of these cases for the energy transition.
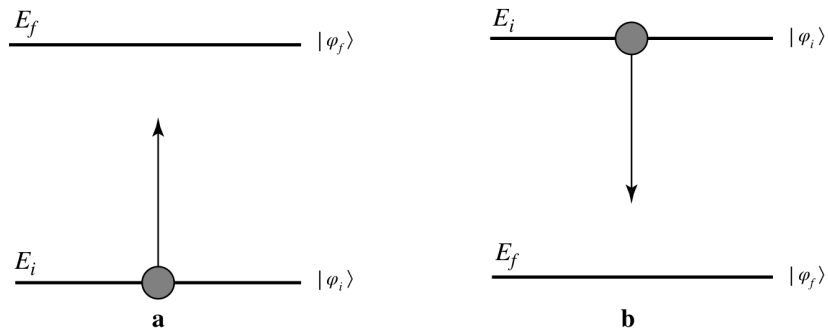

Figure 5: Relative disposition of the energies E_i and E_f associated with the states $|\phi_i\rangle$ and $|\phi_f\rangle$. If $E_i < E_f$, (fig. a), the $|\phi_i\rangle \to |\phi_f\rangle$ transition occurs through absorption of an energy quantum $\hbar\omega$. If, on the other hand $E_i > E_f$ (fig. b), the transition occurs through induced emission of an energy quantum $\hbar\omega$.

From now, we will consider the case $\omega \to \omega_{fi}$ only, however the analysis for $\omega \to -\omega_{fi}$ can be made analogously.

Plotting $\mathcal{P}_{if}^{\omega \to \omega_{fi}}(t;\omega)$ as a function of ω for a specific time yields the curve shown in **Figure 6**. We can see that when $\omega \simeq \omega_{fi}$, a resonance appears whose intensity is proportional to t^2 and whose width is inversely proportional to t. The probability of transition is greatest when the driving frequency ω is closest to the natural frequency ω_{if}. It is also larger the longer the system is exposed to the perturbation.
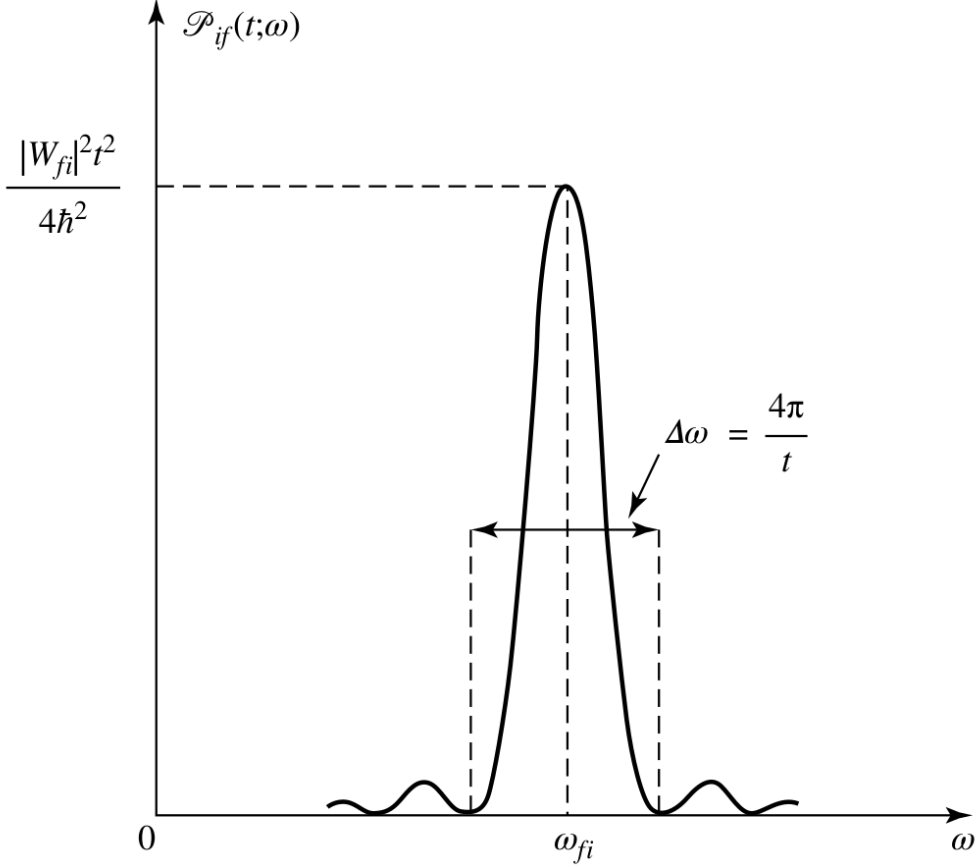

Figure 6: Variation of the first order transition probability $\mathcal{P}_{if}(t;\omega)$ for $\omega \to \omega_{fi}$ as a function of ω for a fixed time t.

Note that, as the limit when $t \to \infty$, we have:

$$\lim_{t \to \infty} \left(\frac{\sin(at)}{a} \right)^2 = \pi t \delta(a) \tag{353}$$

Therefore, in our case:

$$\lim_{t \to \infty} \mathcal{P}_{if}(t;\omega) = \frac{\left| H_{pfi} \right|^2}{4\hbar^2} \lim_{t \to \infty} \left(\frac{\sin \frac{(\omega_{fi}-\omega)t}{2}}{\frac{\omega_{fi}-\omega}{2}} \right)^2 = \frac{\left| H_{pfi} \right|^2}{4\hbar^2} \pi t \, \delta\left(\frac{\omega_{fi}-\omega}{2} \right) \tag{354}$$

5.1.3   Limits to the first order calculation

If we look at the plot in **Figure 6**, it is clear that there must be some limits to the application of the first order approximation. For instance, if we take a look at the expression of the maximum probability, we can see that it depends on time:

$$\max \left(\mathcal{P}_{if}(t;\omega) \right) = \frac{\left| H_{pfi} \right|^2 t^2}{4\hbar^2}, \tag{355}$$

57

and that it goes to infinity at $t \to \infty$. This is clearly absurd, as probability can never go above 1. In practice, for the first order approximation to be valid at resonance, it must be smaller than 1. That is:

$$t \ll \frac{\hbar}{\left| H_{p_{fi}} \right|} \tag{356}$$

Furthermore, resonance will not take place unless the perturbation has had enough time to appear to the system as sinusoidal (in other words, it needs to perform a few oscillations for the system to perceive the perturbation as sinusoidal). This condition can be expressed as:

$$2|\omega_{fi}| \gg \Delta\omega, \tag{357}$$

where $\Delta\omega = \frac{4\pi}{t}$ is the width of the resonant part in **Figure 6**. The condition then becomes:

$$\cancel{2}|\omega_{fi}| \gg \frac{\cancel{4}\pi}{t} \quad \longrightarrow \quad t \gg \frac{1}{\omega_{fi}} \tag{358}$$

Therefore, combining the conditions from **Equation 356** and **Equation 358**, we obtain:

$$\frac{1}{\omega_{fi}} \ll \frac{1}{\left| H_{p_{fi}} \right|} \tag{359}$$

5.1.4 Coupling between discrete states: Rabi's formula

In this section, we will study the oscillations of a system between two discrete states when they are exposed to a sinusoidal resonant perturbation for a long time, for which the previous limiting conditions are not met.

For these cases, as the first order solution is insufficient, in principle it would be possible to simply find higher order terms to obtain a better expression. However, such a method would lead to unnecessarily long calculations. In this section, we shall see that it is possible to solve the problem in a quicker and more elegant way by considering the fact that the resonance condition implies that only the two discrete states $|\phi_i\rangle$ and $|\phi_f\rangle$ are effetively coupled by the perturbation. In other words, the only possible transitions (with non-negligible probability) are between these two states.

Back in **Section 5.1**, we replaced all the coefficients $b_k(t)$ of **Equation 324** by their expansion in λ:

$$b_k(t) = \sum_{q=0}^{\infty} \lambda^q b_k^{(q)}(t), \tag{360}$$

In this case, we will simply neglect all terms $k \neq i, f$ and leave $b_i(t)$ and $b_f(t)$ as they are:

$$
\begin{aligned}
i\hbar \frac{db_i(t)}{dt} &= \lambda \sum_k e^{i\omega_{ik}t} W_{ik}(t) b_k(t) = \\
&= \lambda \sum_{k \neq i,f} \cancel{e^{i\omega_{ik}t} W_{ik}(t) b_k(t)} + e^{i\omega_{ii}t} H_{p_{ii}}(t) b_i(t) + e^{i\omega_{if}t} H_{p_{if}}(t) b_f(t),
\end{aligned} \tag{361}
$$

where we have used that $H_p(t) = \lambda W(t)$. Since the perturbation is sinusoidal, we have $H_p(t) = H_p \sin \omega t$. Therefore, using this and the fact that $\omega_{if} = -\omega_{fi}$:

$$
\begin{aligned}
i\hbar \frac{db_i(t)}{dt} &= \cancel{e^{i\omega_{ii}t}}^{1} H_{p_{ii}} \sin(\omega t) b_i(t) + e^{-i\omega_{fi}t} H_{p_{if}} \sin(\omega t) b_f(t) = \\
&= H_{p_{ii}} \left(\frac{e^{i\omega t} - e^{-i\omega t}}{2i} \right) b_i(t) + e^{-i\omega_{fi}t} H_{p_{if}} \left(\frac{e^{i\omega t} - e^{-i\omega t}}{2i} \right) b_f(t) = \\
&= \frac{1}{2i} \left\{ H_{p_{ii}} \left(e^{i\omega t} - e^{-i\omega t} \right) b_i(t) + e^{-i\omega_{fi}t} H_{p_{if}} \left(e^{i\omega t} - e^{-i\omega t} \right) b_f(t) \right\} = \\
&= \frac{1}{2i} \left\{ H_{p_{ii}} b_i(t) \left(e^{i\omega t} - e^{-i\omega t} \right) + H_{p_{if}} b_f(t) \left(e^{i(\omega - \omega_{fi})t} - e^{-i(\omega + \omega_{fi})t} \right) \right\}
\end{aligned} \tag{362}
$$

58

Finally, if we do the same for $b_f(t)$, we obtain the following system of differential equations:

$$\begin{cases} i\hbar\dfrac{db_i(t)}{dt} = \dfrac{1}{2i}\left\{ H_{p_{ii}}b_i(t)\left(e^{i\omega t}-e^{-i\omega t}\right) + H_{p_{if}}b_f(t)\left(e^{i(\omega-\omega_{fi})t}-e^{-i(\omega+\omega_{fi})t}\right)\right\} \\ i\hbar\dfrac{db_f(t)}{dt} = \dfrac{1}{2i}\left\{ H_{p_{fi}}b_i(t)\left(e^{i(\omega+\omega_{if})t}-e^{-i(\omega-\omega_{if})t}\right) + H_{p_{ff}}b_f(t)\left(e^{i\omega t}-e^{-i\omega t}\right)\right\} \end{cases} \quad (363)$$

On the right-hand side of these equations, certain coefficients of $b_i(t)$ and $b_f(t)$ are proportional to $e^{\pm i(\omega-\omega_{fi})t}$, so they oscillate slowly in time when $\omega \simeq \omega_{fi}$. On the other hand, the coefficients proportional to either $e^{\pm i\omega t}$ or $e^{\pm i(\omega+\omega_{fi})t}$ oscillate much more rapidly. Here, we will use the so called "secular approximation", which consists in neglecting the second type of terms, which average to zero when integrated over a large number of periods. The remaining ones (called "secular terms") are then those whose coefficients are reduced to constants for $\omega = \omega_{fi}$.

Note that, for the preceding argument to be valid, it is necessary for the temporal variation of a term $e^{i\omega t}b_{i,f}(t)$ to be due principally to the exponential, and not to the component $b_{i,f}(t)$. Since ω is very close to ω_{fi}, this means that $b_{i,f}(t)$ must not vary significantly over a time interval of the order of $1/|\omega_{fi}|$. This is indeed true with the assumptions we have made, that is, with $H_p \ll H_0$. The variations of $b_{i,f}(t)$ (which are constants if $H_p = 0$) are due to the presence of the perturbation H_p, and are appreciable for times of the order of $\hbar/|H_{p_{if}}|$. Since, by hypothesis $H_{p_{if}} \ll \hbar|\omega_{fi}|$, this time is much greater $1/|\omega_{fi}|$.

In conclusion, the secular approximation leads to the following system of equations:

$$\begin{cases} \dfrac{db_i(t)}{dt} = -\dfrac{1}{2\hbar}\, e^{i(\omega-\omega_{fi})t}H_{p_{if}}b_f(t) \\ \dfrac{db_f(t)}{dt} = \dfrac{1}{2\hbar}\, e^{-i(\omega-\omega_{if})t}H_{p_{fi}}b_i(t) \end{cases}, \quad (364)$$

whose solution gives **Rabi's formula**, defined below:

Definition 26. Rabi's formula gives the probability of a system in an initial state $|\phi_i\rangle$ transitioning to a state $|\phi_f\rangle$ when submitted to a resonant perturbation $H_p(t) = H_p\sin(\omega t)$, where $\omega \simeq \omega_{fi}$. It is given by the expression:

$$\mathcal{P}_{if}(t;\omega) = \frac{|H_{p_{if}}|^2}{|H_{p_{if}}|^2 + \hbar^2(\omega-\omega_{fi})^2}\sin\left[\frac{t}{2}\sqrt{\frac{|H_{p_{fi}}|^2}{\hbar} + (\omega-\omega_{fi})^2}\right] \quad (365)$$

Rabi's formula measures the magnitude of the resonance phenomenon. Note that, when $\omega = \omega_{fi}$, however small the perturbation is, it can cause the system to move completely to the final state $|\phi_f\rangle$, irrespective of the magnitude of H_p. The only effect of the magnitude of H_p is to make the transition faster (for larger H_p) or slower (for smaller H_p).

Note also that the probability is an oscillating function of time. For certain periodic time intervals, the probability of transition is zero, and the system has returned to the initial state $|\phi_i\rangle$

Finally, we can see that this expression is valid both for $t \to \infty$ and $t \to 0$, and it remains always bounded between 0 and 1.

5.2   Interaction of an atom with an electromagnetic wave

Consider a plane electromagnetic wave[46], of wave vector $\vec{k}$ parallel to the y-axis and angular frequency ω. In this configuration, as shown on **Figure 7**, the electric field of the plane wave is parallel to the z-axis, and its magnetic field is parallel to the x-axis.

---

[46]This result can be generalised for any electromagnetic field.
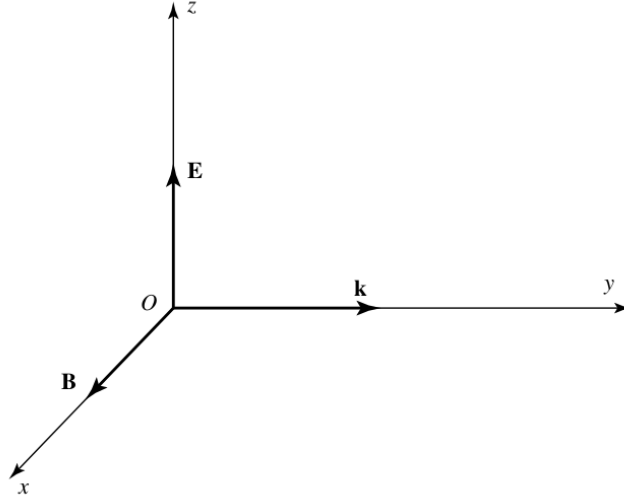
Figure 7: The electric field $\vec{E}$ and the magnetic field $\vec{B}$ of a plane wave of wave vector $\vec{k}$.

From electromagnetism, we have that, for a specific choice of gauge, consisting of a vector potential $\vec{A}(\vec{r}, t)$ and a scalar potential $U(\vec{r}, t)$, we have:

$$\vec{B}(\vec{r}, t) = \vec{\nabla} \times \vec{A}(\vec{r}, t)$$
$$\vec{E}(\vec{r}, t) = -\nabla U(\vec{r}, t) - \frac{\partial}{\partial t} A(\vec{r}, t) \tag{366}$$

It is always possible to make the scalar potential zero, so that the vector potential is given by:

$$\vec{A}(\vec{r}, t) = \mathcal{A}_0 \vec{e_z} e^{i(ky - \omega t)} + \mathcal{A}_0^* \vec{e_z} e^{-i(ky - \omega t)} \tag{367}$$

where $\mathcal{A}_0$ is a complex constant whose argument depends on the choice of the time origin. Then, using **Equation 366**, we obtain:

Finish!!

6 Identical Particles and Many Electron Atoms

6.1 Identical particles

In quantum mechanics, two particles are said to be identcal if they have the same intrinsic properties, such as mass, spin, and charge. Identical particles cannot be distinguished from each other in any experiment. For example, all electrons in the universe are considered identical, as are all protons and hydrogen atoms. However, particles with the same mass and spin, but different charges (like an electron and a positron) are not identical.

An important implication of identical particles is that when a physical system contains two identical particles, swapping their roles does not alter the system's properties or evolution. This definition is independent of experimental conditions, emphasizing the fundamental nature of identical particles in physics[47].

In *classical mechanics*, having identical particles poses no problem. If we have a system with two completely interchangeable particles, we can still label one of them with the index "1", and the other one with the index "2", and treat them as if they were completely different. If each particle follows a trajectory, we can always follow each one and distinguish it from the other one. The Lagrangian and the Hamiltonian of the system are invariant under the exchange of the indices, but one need only choose one of the possible descriptions at the initial moment and simply ignore the other; no ambiguity introduced.

In quantum mechanics, however, the situation is radically different, as particles no longer have well defined trajectories. Even if, at time t_0, the wave functions of two identical particles are completely separated, their subsequent time evolution may mix them. It is therefore easy to "lose track" of particles. When we detect a particle in a region of space in which another identical particle also has non-zero position probability, we have no way of knowing which of the two particles we really measured. The numbering of particles then easily becomes ambiguous when their positions are measured, since there exist several different paths by which the system can end up in the same measured state.

6.2 Permutation operators for 2 particles

In this section, we shall introduce permutation operators to simplify the calculations and reasoning in what follows.

Consider a system composed of two particles with the same spin s. Here it is not necessary for these two particles to be identical; it is sufficient that their individual state spaces be isomorphic[48]. We denote the particles with the indices "1" (for example, a proton) and "2" (for example, an electron).

We choose a basis, $\{|u_i\rangle\}$, in the state space $\mathcal{E}(1)$ of the first particle. Since the state space $\mathcal{E}(2)$ of the second particle is isomorphic to $\mathcal{E}(1)$ (both particles have the same spin $s = \frac{1}{2}$), it can be spanned by the same basis. Then, we can construct the tensor product space $\mathcal{E} = \mathcal{E}(1) \otimes \mathcal{E}(2)$ of the system, with basis:

$$\{|1 : u_i; \ 2 : u_j\rangle\} \tag{368}$$

Note that the order of the vectors in a tensor product is irrelevant, so we have:

$$|1 : u_i; \ 2 : u_j\rangle \equiv |2 : u_j; \ 1 : u_i\rangle \tag{369}$$

However:

$$|1 : u_i; \ 2 : u_j\rangle \neq |1 : u_j; \ 2 : u_i\rangle \tag{370}$$

The permutation operator P_{21} is then defined as the linear operator whose action on the basis vectors is given by:

$$P_{21} |1 : u_i; \ 2 : u_j\rangle = |2 : u_i; \ 1 : u_j\rangle = |1 : u_j; \ 2 : u_i\rangle \tag{371}$$

---

[47]This is to say that, even if, in a given experiment, the charges of the particles are not measured, an electron and a positron can never be treated like identical particles

[48]In other words, their individual state spaces should have the same dimensions. They should have the same structure, so that we can identify each element of the state space of particle "1" with exactly one element from the state space of particle "2". Both state spaces are the same after a simple *renaming* of the elements.

Clearly, the permutation operator is *its own inverse*:

$$P_{21}\left(P_{21}\left|1:u_i;\ 2:u_j\right\rangle\right)=P_{21}\left|1:u_j;\ 2:u_i\right\rangle=\left|1:u_i;\ 2:u_j\right\rangle\longrightarrow\left(P_{21}\right)^2=1 \tag{372}$$

Also, it is *Hermitian*:

$$\begin{cases}\left\langle1:u_{i'};\ 2:u_{j'}\right|P_{21}\left|1:u_i;\ 2:u_j\right\rangle=\left\langle1:u_{i'};\ 2:u_{j'}\right|1:u_j;\ 2:u_i\right\rangle=\delta_{i'j}\delta_{ij'}\\\left\langle1:u_{i'};\ 2:u_{j'}\right|P_{21}^\dagger\left|1:u_i;\ 2:u_j\right\rangle=\left(\left\langle1:u_i;\ 2:u_j\right|P_{21}\left|1:u_{i'};\ 2:u_{j'}\right\rangle\right)^*=\delta_{i'j}\delta_{ij'}\end{cases}\longrightarrow\quad P_{21}^\dagger=P_{21}, \tag{373}$$

which means it is also *unitary*:

$$P_{21}^\dagger P_{21}=P_{21}P_{21}^\dagger=1 \tag{374}$$

The fact that permutation operators are unitary means that their eigenvalues are either 1 or -1. Eigenvectors of the permutation operators are classified according to the eigenvalue they correspond to:

- **Symmetric eigenkets** are those that correspond to the eigenvalue $+1$.

$$P_{21}\left|\psi_S\right\rangle=\left|\psi_S\right\rangle \tag{375}$$

- **Antisymmetric eigenkets** are those that correspond to the eigenvalue -1.

$$P_{21}\left|\psi_A\right\rangle=-\left|\psi_A\right\rangle \tag{376}$$

At this point, it is useful to introduce the following two projector operators[49]:

- The **symmetriser** projects onto the subspace of symmetric kets:

$$S\equiv\frac{1+P_{12}}{2} \tag{377}$$

- The **antisymmetriser** projects onto the subspace of antisymmetric kets:

$$A\equiv\frac{1-P_{12}}{2} \tag{378}$$

Since the projector operator is Hermitian, so are A and B. Furthermore, we can check that they project onto orthogonal subspaces[50] by checking that $AS=SA=0$:

$$SA=\frac{1+P_{21}}{2}\frac{1-P_{21}}{2}=\frac{1+P_{21}-P_{21}-P_{21}P_{21}}{2}=0 \tag{379}$$

$$AS=\frac{1-P_{21}}{2}\frac{1+P_{21}}{2}=\frac{1+P_{21}-P_{21}-P_{21}P_{21}}{2}=0 \tag{380}$$

And it is also easy to check that the subspaces of symmetric kets and antisymmetric kets are supplementary, as $A+S=\mathbb{1}$.

It is now easy to see that, for an arbitrary ket $\left|\psi\right\rangle$ of state space, $A\left|\psi\right\rangle$ is antisymmetric and $S\left|\psi\right\rangle$ is symmetric. Then:

$$P_{21}A\left|\psi\right\rangle=-A\left|\psi\right\rangle,\qquad P_{21}S\left|\psi\right\rangle=S\left|\psi\right\rangle \tag{381}$$

Since the subspaces of symmetric and antisymmetric kets are orthogonal, we can always write this arbitrary ket $\left|\psi\right\rangle$ as a combination of a symmetric ket and an antisymmetric ket[51]:

$$\left|\psi\right\rangle=\mathbb{1}\left|\psi\right\rangle=(A+S)\left|\psi\right\rangle=A\left|\psi\right\rangle+S\left|\psi\right\rangle \tag{382}$$

---

[49]Remember from **Section 1.4** that the definition of a projector operator A is that $A^2=A$ (it is idempotent).

[50]The subspace of symmetric kets is orthogonal to that of antisymmetric kets.

[51]This is a nice mathematical result, but as we will see later in **Postulate 9**, physical systems *must* be either completely symmetric or completely antisymmetric. There are no asymmetric physical kets.

6.3 Permutation operators for N particles

In the state space of a system composed of N identical particles, $N!$ permutation operators can be defined (one of which is the identity operator). If N is greater than 2, the properties of these operators are more complex than those of P_{21}.

For example, consider a system with $N = 3$ particles with the same spin s, as in the case we considered with 2 particles. In this case, we will have the basis:

$$\{|1 : u_i;\ 2 : u_j;\ 3 : u_k\rangle\} \tag{383}$$

We will then have $3! = 6$ permutation operators:

$$P_{123} = \mathbb{1}, \quad P_{312}, \quad P_{231}, \quad P_{132}, \quad P_{213}, \quad P_{321}, \tag{384}$$

where:

$$P_{npq} |1 : u_i;\ 2 : u_j;\ 3 : u_k\rangle \equiv |n : u_i;\ p : u_j\ q : u_k\rangle \tag{385}$$

We can also define *transposition operators*, T_{npq}, which are permutations that only exchange two particles. These are the basis for the definition of symmetric and antisymmetric kets in systems of N identical particles:

Definition 27. Let T_α be an arbitrary transposition in a system of N identical particles in a state space $\mathcal{E}$. We define *symmetric* and *antisymmetric* eigenkets, $|\psi_S\rangle$ and $|\psi_A\rangle$, as follows:

$$T_\alpha |\psi_S\rangle = |\psi_S\rangle \qquad P_\alpha |\psi_A\rangle = - |\psi_A\rangle \tag{386}$$

The space of symmetric kets $\mathcal{E}_S$ and the space of antisymmetric kets $\mathcal{E}_A$ are subspaces of the whole state space $\mathcal{E}$. They are orthogonal to each other and supplementary, since $\mathcal{E} = \mathcal{E}_S \oplus \mathcal{E}_A$.

Permutation (and, thus, transposition) operators have the following properties:

- The product of any pair of permutation operators is another permutation operator. For example:

$$P_{312}P_{132} |1 : u_i;\ 2 : u_j;\ 3 : u_k\rangle = P_{312} |1 : u_i;\ 3 : u_j;\ 2 : u_k\rangle = P_{312} |1 : u_i;\ 2 : u_k;\ 3 : u_j\rangle =$$
$$= |3 : u_i;\ 2 : u_j;\ 1 : u_k\rangle = P_{321} |1 : u_i;\ 2 : u_j;\ 3 : u_k\rangle \rightarrow \tag{387}$$
$$\rightarrow P_{312}P_{132} = P_{321}$$

- Each permutation operator has an inverse, which is also a permutation operator.

$$
\begin{aligned}
P_{123}^{-1} = P_{123}; \qquad & P_{312}^{-1} = P_{231}; \qquad && P_{231}^{-1} = P_{312} \\
P_{132}^{-1} = P_{132}; \qquad & P_{213}^{-1} = P_{213}; \qquad && P_{321}^{-1} = P_{321}
\end{aligned} \tag{388}
$$

- Any permutation operator can be broken down as a product of transposition operators:

$$P_{312} = P_{132}P_{213} = P_{321}P_{132} = \cdots \tag{389}$$

This decomposition is not unique. However, for a given permutation, it can be shown that the parity of the number of transpositions into which it can be broken down is always the same. This is called the *parity of the permutation*:

$$\text{Even}: P_{123}, P_{312}, P_{231} \qquad \text{Odd}: P_{132}, P_{213}, P_{321}$$

For any given number of particles N, there are always the same number of even and odd permutations.

For an arbitrary number of particles N, we can define the eigenkets of permutation operators as:

Definition 28. Let P_α be an arbitrary permutation of N identical particles in a state space $\mathcal{E}$. We define *symmetric* and *antisymmetric* eigenkets, $|\psi_S\rangle$ and $|\psi_A\rangle$, as follows:

$$P_\alpha |\psi_S\rangle = |\psi_S\rangle \qquad P_\alpha |\psi_A\rangle = \varepsilon_\alpha |\psi_A\rangle \tag{390}$$

where[a]:

$$\varepsilon_\alpha = \begin{cases} +1; & P_\alpha \text{ is even} \\ -1; & P_\alpha \text{ is odd} \end{cases} \tag{391}$$

The space of symmetric kets $\mathcal{E}_S$ and the space of antisymmetric kets $\mathcal{E}_A$ are subspaces of the whole state space $\mathcal{E}$. They are orthogonal to each other and supplementary, since $\mathcal{E} = \mathcal{E}_S \oplus \mathcal{E}_A$.

---

[a]Note that this ε_α arises from the fact that *all transpositions on an antisymmetric ket give an eigenvalue of* -1. If the parity of P_α is even, it means P_α is equivalent to applying an even number of transpositions, which means its eigenvalue will be $(-1)^{2k} = 1$, for some arbitrary $k \in \mathbb{N}$. On the other hand, if the parity of P_α is odd, it means P_α is equivalent to applying an odd number of transpositions, which means its eigenvalue will be $(-1)^{2k+1} = -1$, for some arbitrary $k \in \mathbb{N}$.

We can also generalise the definition of the symmetriser and the antisymmetriser:

Definition 29. For a system of N particles, we define the **symmetriser**, which projects any ket onto $\mathcal{E}_S$, as:

$$S \equiv \frac{1}{N!} \sum_{\alpha=1}^{N!} P_\alpha \tag{392}$$

and the **antisymmetriser**, which projects any ket onto $\mathcal{E}_A$:

$$A \equiv \frac{1}{N!} \sum_{\alpha=1}^{N!} \varepsilon_\alpha P_\alpha \tag{393}$$

These verify:

$$P_\alpha S |\psi\rangle = S |\psi\rangle, \qquad P_\alpha A |\psi\rangle = \varepsilon_\alpha A |\psi\rangle \tag{394}$$

Since P_α are Hermitian, so are S and A:

$$A^\dagger = A, \qquad S^\dagger = S \tag{395}$$

We can see that S and A are projection operators:

$$S^2 = \frac{1}{N!} \sum_{\alpha=1}^{N!} P_\alpha S = \frac{S}{N!} \sum_{\alpha=1}^{N!} 1 = S \tag{396}$$

$$A^2 = \frac{1}{N!} \sum_{\alpha=1}^{N!} \varepsilon_\alpha P_\alpha A = \frac{A}{N!} \sum_{\alpha=1}^{N!} \varepsilon_\alpha^2 = A \tag{397}$$

And that $\mathcal{E}_S$ and $\mathcal{E}_A$ are orthogonal:

$$AS = \frac{1}{N!} \sum_{\alpha=1}^{N!} \varepsilon_\alpha P_\alpha S = \frac{1}{N!} \sum_{\alpha=1}^{N!} \varepsilon_\alpha S = \frac{S}{N!} \sum_{\alpha=1}^{N!} \varepsilon_\alpha \overset{52}{=} 0 \tag{398}$$

---

[52]Remember that there are the same number of even $\varepsilon_\alpha = 1$ and odd $\varepsilon_\alpha = -1$ permutations.

7 Symmetrisation Postulate

All these theoretical concepts around permutations are the basis for postulating the **symmetrisation postulate**. We have seen that the ket of a system of N identical particles is either symmetric or antisymmetric under the interchange of any pair of particles (transposition) T_α:

$$T_\alpha \ket{\psi} = \pm \ket{\psi} \tag{399}$$

This is the very basis for the symmetrisation postulate, which states the following:

> **Postulate 9. Symmetrisation Postulate:**
>
> - In nature, the states of systems containing N identical particles are either totally symmetric or totally antisymmetric under the interchange of any pair of particles. States with mixed symmetry do not exist — they are not physical kets[a].
>
> - Particles with **integral spins**, or **bosons**, have **symmetric** states.
>
> - Particles with **half-odd-integral spins**, or **fermions**, have **antisymmetric states**.
>
> _____
> [a]add intuition behind this regarding $|\psi(a,b)|^2 = |\psi(b,a)|^2 \rightarrow \psi(a,b) = \pm\psi(b,a)$

This new postulate which we have introduced considerably restricts the class of mathematical kets able to describe a physical state: these kets must be symmetric for bosons, or antisymmetric for fermions.

For composite particles (particles composed of other particles), the symmetrisation postulate also applies. All we have to do is add the spin of all the constituents. If the sum is an integer number, the composite particle will be a boson; while if the sum is a half-odd-integer, the composite particle will be a fermion.

8 Construction of physical kets

Based on the preceding discussion one can formulate the rule for constructing a unique physical ket corresponding to a given physical state of a system of N identical particles. The process is simple:

1. Number the N particles arbitrarily and construct a ket corresponding to a given physical state and to the numbers given to the particles.

2. Apply S or A to this ket, depending on whether the particles are bosons or fermions.

3. Normalize the resulting ket.

8.1 System of 2 particles

Let us take the example of a system of two identical particles in two different states $\ket{\varphi} \neq \ket{\chi}$. Following the process we defined in the previous section, we construct a ket matching this state:

$$\ket{u} = \ket{1 : \varphi; \ 2 : \chi} \tag{400}$$

- If the particle is a **boson**, we symmetrise this ket applying S:

$$\ket{\varphi; \ \chi} = S\ket{u} = \frac{1 + P_{21}}{2}\ket{u} = \frac{\ket{1 : \varphi; \ 2 : \chi} + \ket{1 : \chi; \ 2 : \varphi}}{2} \tag{401}$$

Finally we normalise (assuming $\ket{u}$ is normalised):

$$\ket{\varphi; \ \chi}_{\text{boson}} = \frac{\ket{1 : \varphi; \ 2 : \chi} + \ket{1 : \chi; \ 2 : \varphi}}{\sqrt{2}} \tag{402}$$

- If the particle is a **fermion**, we antisymmetrise this ket applying A:

$$|\varphi; \chi\rangle = S|u\rangle = \frac{1 - P_{21}}{2}|u\rangle = \frac{|1:\varphi;\ 2:\chi\rangle - |1:\chi;\ 2:\varphi\rangle}{2} \tag{403}$$

Finally we normalise (assuming $|u\rangle$ is normalised):

$$|\varphi; \chi\rangle_{\text{fermion}} = \frac{|1:\varphi;\ 2:\chi\rangle - |1:\chi;\ 2:\varphi\rangle}{\sqrt{2}} \tag{404}$$

Note that, in the case that the two identical particles are in the same state, $|\varphi\rangle = |\chi\rangle$, we have:

$$|\varphi; \varphi\rangle_{\text{boson}} = \frac{|1:\varphi;\ 2:\varphi\rangle + |1:\varphi;\ 2:\varphi\rangle}{\sqrt{2}} = |1:\varphi;\ 2:\varphi\rangle \tag{405}$$

$$|\varphi; \varphi\rangle_{\text{fermion}} = \frac{|1:\varphi;\ 2:\varphi\rangle - |1:\varphi;\ 2:\varphi\rangle}{\sqrt{2}} = 0 \tag{406}$$

What does this zero mean? Well, we have just obtained the **Pauli Exclusion Principle**: there exists no ket in which two fermions are in the same individual state.

8.2   System of 3 particles

Let us now consider the example of a system of *three* identical particles in *three* different states $|\varphi\rangle \neq |\chi\rangle \neq |\psi\rangle$. Following the same process as before, we construct a ket matching this state:

$$|u\rangle = |1:\varphi;\ 2:\chi;\ 3:\psi\rangle \tag{407}$$

- If the particle is a **boson**, we symmetrise this ket applying S:

$$
\begin{aligned}
|\varphi; \chi;\ \psi\rangle = S|u\rangle &= \frac{1}{N!}\sum_{\alpha=1}^{N!} P_\alpha |u\rangle = \\
&= \frac{1}{6}\big(|1:\varphi;\ 2:\chi;\ 3:\psi\rangle + |1:\psi;\ 2:\varphi;\ 3:\chi\rangle + |1:\chi;\ 2:\psi;\ 3:\varphi\rangle + \\
&\quad + |1:\varphi;\ 2:\psi;\ 3:\chi\rangle + |1:\chi;\ 2:\varphi;\ 3:\psi\rangle + |1:\psi;\ 2:\chi;\ 3:\varphi\rangle\big)
\end{aligned} \tag{408}
$$

Finally we normalise (assuming $|u\rangle$ is normalised):

$$
\begin{aligned}
|\varphi; \chi;\ \psi\rangle = \frac{1}{\sqrt{6}}\big(&|1:\varphi;\ 2:\chi;\ 3:\psi\rangle + |1:\psi;\ 2:\varphi;\ 3:\chi\rangle + |1:\chi;\ 2:\psi;\ 3:\varphi\rangle + \\
&+ |1:\varphi;\ 2:\psi;\ 3:\chi\rangle + |1:\chi;\ 2:\varphi;\ 3:\psi\rangle + |1:\psi;\ 2:\chi;\ 3:\varphi\rangle\big)
\end{aligned} \tag{409}
$$

If two of the three states coincide ($|\varphi\rangle = |\chi\rangle$), then:

$$|\varphi; \varphi;\ \psi\rangle = \frac{1}{\sqrt{3}}\big(|1:\varphi;\ 2:\varphi;\ 3:\psi\rangle + |1:\psi;\ 2:\varphi;\ 3:\varphi\rangle + |1:\varphi;\ 2:\psi;\ 3:\varphi\rangle\big) \tag{410}$$

If all three states coincide ($|\varphi\rangle = |\chi\rangle = |\psi\rangle$), then:

$$|\varphi; \varphi;\ \varphi\rangle = |1:\varphi;\ 2:\varphi;\ 3:\varphi\rangle \tag{411}$$

- If the particle is a **fermion**, we antisymmetrise this ket applying A:

$$|\varphi;\,\chi;\,\psi\rangle = S\,|u\rangle = \frac{1}{N!}\sum_{\alpha=1}^{N!}\varepsilon_\alpha P_\alpha\,|u\rangle =$$
$$= \frac{1}{6}\left(|1:\varphi;\,2:\chi;\,3:\psi\rangle + |1:\psi;\,2:\varphi;\,3:\chi\rangle + |1:\chi;\,2:\psi;\,3:\varphi\rangle -\right.$$
$$\left.- |1:\varphi;\,2:\psi;\,3:\chi\rangle - |1:\chi;\,2:\varphi;\,3:\psi\rangle - |1:\psi;\,2:\chi;\,3:\varphi\rangle\right) \quad (412)$$

Finally we normalise (assuming $|u\rangle$ is normalised):

$$|\varphi;\,\chi;\,\psi\rangle = \frac{1}{\sqrt{6}}\left(|1:\varphi;\,2:\chi;\,3:\psi\rangle + |1:\psi;\,2:\varphi;\,3:\chi\rangle + |1:\chi;\,2:\psi;\,3:\varphi\rangle -\right.$$
$$\left.- |1:\varphi;\,2:\psi;\,3:\chi\rangle - |1:\chi;\,2:\varphi;\,3:\psi\rangle - |1:\psi;\,2:\chi;\,3:\varphi\rangle\right) \quad (413)$$

It is easy to check that, if any two states coincide, the wave function becomes zero.

It is evident that, the larger the number of particles, the longer and more unwieldy that the expressions become. For **fermions**, there exists a simple notation that simplifies all this for us by using determinants. If we look at **Equation 413**, we see that the combinations that appear are exactly the same as those that are present in a determinant, signs included. Understanding the product here as a tensor product, we can then write:

$$|\varphi;\,\chi;\,\psi\rangle = \frac{1}{\sqrt{6}}\begin{vmatrix} |1:\varphi\rangle & |1:\chi\rangle & |1:\psi\rangle \\ |2:\varphi\rangle & |2:\chi\rangle & |2:\psi\rangle \\ |3:\varphi\rangle & |3:\chi\rangle & |3:\psi\rangle \end{vmatrix} \quad (414)$$

It is now even clearer to see that, if any two states or more are equal, we will have as many equal columns as equal states. This will immediately make the determinant go to zero.

It is also possible to extend the same notation to **bosons**, but forgetting all the minus signs in the determinant.

8.3 Systems of N particles

We can generalise our discussion in the previous two sections for a system of N identical particles:

> **Definition 30.** For a system of N identical particles, we can construct a physical ket $|\phi_1, \phi_2, ..., \phi_N\rangle$ for a certain physical state in the following way:
>
> - If the particles are **fermions**, we use the *Slater Determinant* to construct an **antisymmetric ket**:
> $$|\phi_1;\phi_2;...;\phi_N\rangle = \frac{1}{\sqrt{N!}}\begin{vmatrix} |1:\phi_1\rangle & |1:\phi_2\rangle & \cdots & |1:\phi_N\rangle \\ |2:\phi_1\rangle & |2:\phi_2\rangle & \cdots & |2:\phi_N\rangle \\ \vdots & \vdots & \ddots & \vdots \\ |N:\phi_1\rangle & |N:\phi_2\rangle & \cdots & |N:\phi_N\rangle \end{vmatrix} \quad (415)$$
>
> - If the particles are **bosons**, we use the *Slater Determinant* to construct an **symmetric ket** by turning all the minus signs into positive signs.

8.4 Systems with spin

We have seen that when the Schrödinger equation involves the spin, the wave function of a single particle $|\Psi\rangle$ is equal to the product of the spatial part $|\psi\rangle$ and the spin part $|\chi\rangle$: $|\Psi\rangle = |\psi\rangle \otimes |\chi\rangle$. The wave function

of a system of N particles, which have spins, is the product of the spatial part and the spin part:

$$|\Psi_1; \Psi_2; ...; \Psi_N\rangle = |\psi_1; \psi_2; ...; \psi_N\rangle \otimes |\chi_1; \chi_2; ...; \chi_N\rangle \tag{416}$$

This wave function must satisfy the appropriate symmetry requirements when the N particles are identical. In the case of a system of N identical **bosons**, the wave function must be symmetric; hence the spatial and spin parts must have the same parity:

$$|\Psi_1; \Psi_2; ...; \Psi_N\rangle_S = \begin{cases} |\psi_1; \psi_2; ...; \psi_N\rangle_S \otimes |\chi_1; \chi_2; ...; \chi_N\rangle_S \\ |\psi_1; \psi_2; ...; \psi_N\rangle_A \otimes |\chi_1; \chi_2; ...; \chi_N\rangle_A \end{cases} \tag{417}$$

In the case of a system of N identical **fermions**, however, the space and spin parts must have different parities, leading to an overall wave function that is antisymmetric:

$$|\Psi_1; \Psi_2; ...; \Psi_N\rangle_A = \begin{cases} |\psi_1; \psi_2; ...; \psi_N\rangle_A \otimes |\chi_1; \chi_2; ...; \chi_N\rangle_S \\ |\psi_1; \psi_2; ...; \psi_N\rangle_S \otimes |\chi_1; \chi_2; ...; \chi_N\rangle_A \end{cases} \tag{418}$$

9 Appendix

9.1 Linear vector space

The definition of linear vector space is as follows:

> **Definition 31.** A linear vector space consists of two sets of elements and two algebraic rules:
>
> - A set of vectors ψ, ϕ, χ, ... and a set of scalars a, b, c, ...
> - A rule for adding vectors and a rule for multiplying vectors by scalars.
>
>   a) **Addition:**
>   - If ψ and ϕ are vectors (elements) of a space, their sum, $\psi + \phi$, is also a vector of the same space.
>   - Commutativity: $\psi + \phi = \phi + \psi$.
>   - Associativity: $(\psi + \phi) + \chi = \psi + (\phi + \chi)$.
>   - Existence of a zero or neutral vector: for each vector ψ, there must exist a zero vector O such that: $\psi + O = O + \psi = \psi$.
>   - Existence of a symmetric or inverse vector: each vector ψ must have a symmetric vector $(-\psi)$ such that $\psi + (-\psi) = (-\psi) + \psi = O$.
>
>   b) **Multiplication:** The multiplication of vectors by scalars (scalars can be real or complex numbers) has these properties:
>   - The product of a scalar with a vector gives another vector. In general, if ψ and ϕ are two vectors of the space, any linear combination $a\psi + b\phi$ is also a vector of the space, a and b being scalars.
>   - Distributivity with respect to addition: a $(\psi + \phi) = a\psi + a\phi$, and $(a + b)\psi = a\psi + b\psi$.
>   - Associativity with respect to multiplication of scalars: $a(b\psi) = (ab)\psi$
>   - For each element ψ there must exist a unitary scalar I and a zero scalar "0" such that $I\psi = \psi I = \psi$ and $\psi 0 = 0\psi = O$. (2.3)

9.2 Linear functionals

In order to understand the mathematical background of the dual space, it is interesting to know the definitions of **linear maps** and **linear functionals**. We already defined linear operators in **Definition 3**. As for linear functionals:

> **Definition 32.** A linear functional is a linear map L that associates a function with a scalar value, which may be real or complex.

An example of a linear functional could be the linear map $L_x : \mathbb{R}^2 \to \mathbb{R}$ that returns the x-coordinate of the vector it is given. For example:

$$L_x \begin{bmatrix} a \\ b \end{bmatrix} = a \tag{419}$$

In this case, L_x takes us from $\mathbb{R}^2$ to $\mathbb{R}^1$, so its matrix will be 1 by 2 in dimension:

$$L_x = \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{420}$$

so that:

$$L_x \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = a \tag{421}$$

Taking a step back, we know that all linear functionals in $\mathbb{R}^2$ take us from $\mathbb{R}^2$ to $\mathbb{R}^1$. Therefore, by definition, all linear functionals in $\mathbb{R}^2$ are represented by 1×2 matrices. In other words, the set of all linear functionals in $\mathbb{R}^2$ consists of the set of all row matrices. More generally, the set of all linear functionals in $\mathbb{R}^n$ consists of the set of all $1 \times n$ row matrices. In fact, the set of all row matrices, much like column matrices, form their own vector space, which is known as the dual space[53].

9.3 Matrix element of an operator

The matrix element of an operator A is:

Definition 33. The matrix element of an operator A expressed in the basis $B = \{u_1, u_2, ...\}$ is defined as:

$$A_{ij} = \langle u_i | A | u_j \rangle \tag{422}$$

where u_i and u_j are the i-th and j-th vectors of the basis, respectively.

9.4 Proof of the Closure relation

The closure relation states that, for a set of vectors $\{|A_1\rangle, |A_2\rangle, ...\}$ to form a basis of a Hilbert space $\mathcal{H}$, they must fulfil the following relation:

If $B = \{\psi_1, \psi_2, ...\}$ is a basis for a Hilbert space $\mathcal{H}$, we can write any vector $|\psi\rangle \in \mathcal{H}$ as:

$$|\psi\rangle = \sum_i c_i |A_i\rangle \tag{423}$$

Where the coefficients are $c_i = \langle A_i \rangle \psi$[54]. If we substitute this expression:

$$|\psi\rangle = \sum_i \langle A_i \rangle \psi |A_i\rangle \tag{424}$$

As the coefficient is a complex number, we can move it to the end of the expression and separate the inner product:

$$|\psi\rangle = \sum_i |A_i\rangle \langle A_i | |\psi\rangle \tag{425}$$

As $|\psi\rangle$ is the same for every element of the sum, we can pull it out of the sum:

$$|\psi\rangle = \left(\sum_i |A_i\rangle \langle A_i| \right) |\psi\rangle \tag{426}$$

And it is now easy to see that this relation will hold if and only if the closure relation holds:

$$\sum_i |A_i\rangle \langle A_i| = \mathbb{1} \tag{427}$$

The inverse argument is easy to follow.

---

[53]See **Section 1.2** for more on the dual space

[54]*Proof:* As A_i form an orthonormal basis: $\langle A_i \rangle \psi = \langle A_i | \sum_j c_i |A_j\rangle = \sum_j c_i \langle A_i \rangle A_j = \sum_j c_i \delta_{ij} = c_i$

9.5 Uncertainty relation between two operators

An interesting application of commutator algebra is to derive a general relation giving the uncertainty product of two Hermitian operators, A and B. Let $\langle A \rangle$ and $\langle B \rangle$ be the expected values of the operators A and B, respectively, with respect to the normalised state vector $|\psi\rangle$. Introducing the operators ΔA and ΔB as:

$$\Delta A = A - \langle A \rangle, \qquad \Delta B = B - \langle B \rangle, \tag{428}$$

we have $(\Delta A)^2 = A^2 - 2A\langle A \rangle + \langle A \rangle^2$ and $(\Delta B)^2 = B^2 - 2B\langle B \rangle + \langle B \rangle^2$, and hence:

$$\langle \psi | (\Delta A)^2 | \psi \rangle = \langle (\Delta A)^2 \rangle = \left\langle A^2 - 2A\langle A \rangle + \langle A \rangle^2 \right\rangle = \langle A^2 \rangle - 2\langle A \rangle\langle A \rangle + \langle A \rangle^2 = \langle A^2 \rangle - \langle A \rangle^2 \tag{429}$$

$$\langle \psi | (\Delta B)^2 | \psi \rangle = \langle (\Delta B)^2 \rangle = \left\langle B^2 - 2B\langle B \rangle + \langle B \rangle^2 \right\rangle = \langle B^2 \rangle - 2\langle B \rangle\langle B \rangle + \langle B \rangle^2 = \langle B^2 \rangle - \langle B \rangle^2 \tag{430}$$

Then, we can write the uncertainties as:

$$\Delta A = \sqrt{\langle (\Delta A)^2 \rangle} = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}, \qquad \Delta B = \sqrt{\langle (\Delta B)^2 \rangle} = \sqrt{\langle B^2 \rangle - \langle B \rangle^2} \tag{431}$$

Writing the action of the operators in **Equation 428** on the state vector $|\psi\rangle$, we have:

$$|\chi\rangle = \Delta A |\psi\rangle = (A - \langle A \rangle)|\psi\rangle, \qquad |\phi\rangle = \Delta B |\psi\rangle = (B - \langle B \rangle)|\psi\rangle \tag{432}$$

The Cauchy-Schwarz inequality[55] gives us:

$$|\langle \chi | \phi \rangle|^2 \leq \langle \chi | \chi \rangle \langle \phi | \phi \rangle \tag{433}$$

Since A and B are Hermitian, ΔA and ΔB must also be Hermitian, so that: $(\Delta A)^\dagger = \Delta A$ and $(\Delta B)^\dagger = \Delta B$:

$$\langle \chi | \chi \rangle = \left\langle \psi \left| (\Delta A)^\dagger (\Delta A) \right| \psi \right\rangle = \langle \psi | (\Delta A)^2 | \psi \rangle = \langle (\Delta A)^2 \rangle \tag{434}$$

$$\langle \phi | \phi \rangle = \left\langle \psi \left| (\Delta B)^\dagger (\Delta B) \right| \psi \right\rangle = \langle \psi | (\Delta B)^2 | \psi \rangle = \langle (\Delta B)^2 \rangle \tag{435}$$

$$\langle \chi | \phi \rangle = \left\langle \psi \left| (\Delta A)^\dagger (\Delta B) \right| \psi \right\rangle = \langle \psi | \Delta A \Delta B | \psi \rangle = \langle \Delta A \Delta B \rangle \tag{436}$$

From this, we obtain a new expression for the Cauchy-Schwarz inequality:

$$|\langle \Delta A \Delta B \rangle|^2 \leq \langle (\Delta A)^2 \rangle \langle (\Delta B)^2 \rangle \tag{437}$$

We can write the $\Delta A \Delta B$ terms of this equation as:

$$\Delta A \Delta B = 2 \cdot \frac{1}{2} \Delta A \Delta B + \frac{1}{2} \Delta B \Delta A - \frac{1}{2} \Delta B \Delta A = \frac{1}{2}[\Delta A, \Delta B] + \frac{1}{2}\{\Delta A, \Delta B\} \tag{438}$$

Since $[\Delta A, \Delta B] = [A, B]$[56]:

$$\Delta A \Delta B = \frac{1}{2}[A, B] + \frac{1}{2}\{\Delta A, \Delta B\} \tag{439}$$

Since $[A, B]$ is anti-Hermitian[57] and $[\Delta A, \Delta B]$ is Hermitian[58]; and since the expectation value of a Hermitian operator is real[59] and the expectation value of an anti-Hermitian operator is imaginary[60], $\langle \Delta A \Delta B \rangle$ becomes equal to a real part $\frac{1}{2}\langle\{\Delta A, \Delta B\}\rangle$ plus an imaginary part $\frac{1}{2}\langle[A, B]\rangle$. Then:

$$|\langle \Delta A \Delta B \rangle|^2 = \frac{1}{4}|\langle[A, B]\rangle|^2 + \frac{1}{4}|\langle\{\Delta A, \Delta B\}\rangle|^2 \geq \frac{1}{4}|\langle[A, B]\rangle|^2 \tag{440}$$

---

[55]See for more on this inequality in this article.

[56]Check this

[57]*Proof:* As A and B are Hermitian, then $([A, B])^\dagger = (AB - BA)^\dagger = (AB)^\dagger - (BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = BA - AB = -[A, B]$. Then, $[A, B]$ is anti-Hermitian.

[58]*Proof:* As A and B are Hermitian, then $([\Delta A, \Delta B])^\dagger = ((A - \langle A \rangle)(B - \langle B \rangle) - (B - \langle B \rangle)(A - \langle A \rangle))^\dagger = ((A - \langle A \rangle)(B - \langle B \rangle))^\dagger - ((B - \langle B \rangle)(A - \langle A \rangle))^\dagger = (B - \langle B \rangle)^\dagger(A - \langle A \rangle)^\dagger - (A - \langle A \rangle)^\dagger(B - \langle B \rangle)^\dagger = (B - \langle B \rangle)(A - \langle A \rangle) - (A - \langle A \rangle)(B - \langle B \rangle) = [\Delta A, \Delta B]$. Then, $[\Delta A, \Delta B]$ is Hermitian.

[59]See **Section 9.6** for proof.

[60]See **Section 9.7** for proof.

Plugging this into **Equation 437**, we obtain:

$$\frac{1}{4} |\langle [A, B] \rangle|^2 \leq \langle (\Delta A)^2 \rangle \langle (\Delta B)^2 \rangle \tag{441}$$

Now, taking the square root, we obtain the uncertainty relation:

$$\Delta A \Delta B \geq \frac{1}{2} |\langle [A, B] \rangle|^2 \tag{442}$$

Definition 34. The uncertainty relation between two Hermitian operators A and B is defined as:

$$\Delta A \Delta B \geq \frac{1}{2} |\langle [A, B] \rangle|^2 \tag{443}$$

9.6 Expected value of a Hermitian operator

From **Equation 22**, we know that the expected value of an operator A in a state $|\phi\rangle$ is $\langle \phi | A | \phi \rangle = \sum_i |c_i|^2 \lambda_i$, where c_i are the components of the vector $|\phi\rangle$ in the basis of eigenvectors ψ_i of A, and λ_i are the corresponding eigenvalues. If A is Hermitian ($A^\dagger = A$):

$$\langle A\psi_i | \psi_i \rangle = \lambda^* ||\psi_i||^2 \tag{444}$$

$$\langle \psi_i | A^\dagger \psi_i \rangle = \langle \psi_i | A\psi_i \rangle = \lambda ||\psi_i||^2 \tag{445}$$

As $\langle A\psi_i | \psi_i \rangle = \langle \psi_i | A^\dagger \psi_i \rangle$, this means that $\lambda^* = \lambda$, so λ must be pure real (or zero). Therefore, the expected value of A must also be pure real.

9.7 Expected value of an anti-Hermitian operator

From **Equation 22**, we know that the expected value of an operator A in a state $|\phi\rangle$ is $\langle \phi | A | \phi \rangle = \sum_i |c_i|^2 \lambda_i$, where c_i are the components of the vector $|\phi\rangle$ in the basis of eigenvectors ψ_i of A, and λ_i are the corresponding eigenvalues. If A is anti-Hermitian ($A^\dagger = -A$):

$$\langle A\psi_i | \psi_i \rangle = \lambda^* ||\psi_i||^2 \tag{446}$$

$$\langle \psi_i | A^\dagger \psi_i \rangle = -\langle \psi_i | A\psi_i \rangle = -\lambda ||\psi_i||^2 \tag{447}$$

As $\langle A\psi_i | \psi_i \rangle = \langle \psi_i | A^\dagger \psi_i \rangle$, this means that $\lambda^* = -\lambda$, so λ must be pure imaginary (or zero). Therefore, the expected value of A must also be pure imaginary.

9.8 Eigenvalues of a Hermitian operator are real and eigenvectors are orthogonal

For a Hermitian operator, all of its eigenvalues are real and the eigenvectors corresponding to different eigenvalues are orthogonal. *Proof:*

Note that:

$$A |\psi_n\rangle = a_n |\psi_n\rangle \rightarrow \langle \psi_m | A |\psi_n\rangle = a_n \langle \psi_m | \psi_n \rangle \tag{448}$$

and:

$$\langle \psi_m | A^\dagger = (A |\psi_m\rangle)^\dagger = (a_m |\psi_m\rangle)^\dagger = a_m^* \langle \psi_m | \rightarrow \langle \psi_m | A^\dagger |\psi_n\rangle = a_m^* \langle \psi_m | \psi_n \rangle \tag{449}$$

Subtracting **Equation 449** from **Equation 448** and using the fact that A is Hermitian ($A^\dagger = A$):

$$\langle \psi_m | A |\psi_n\rangle - \langle \psi_m | A^\dagger |\psi_n\rangle = (a_n - a_m) \langle \psi_m | \psi_n \rangle = (a_n - a_m^*) \langle \psi_m | \psi_n \rangle = 0 \tag{450}$$

From this:

- When $m = n$, we obtain $a_n = a_n^*$, so a_n is real.

- Since, in general, $a_n - a_m^* \neq 0$ when $n \neq m$, we have that $\langle \psi_m | \psi_n \rangle = 0$, so the eigenvectors are orthogonal.

9.9 Common eigenvector basis of two commuting operators

This section gives the proof of **Theorem 4**:

Since A has no degenerate eigenvalues[61], to each eigenvalue of A there corresponds only one eigenvector. We can write the eigenvalue equation as:

$$A |\phi_n\rangle = a_n |\phi_n\rangle \tag{451}$$

And since A commutes with B:

$$AB |\phi_n\rangle = BA |\phi_n\rangle = Ba_n |\phi_n\rangle = a_n B |\phi_n\rangle \rightarrow A(B |\phi_n\rangle) = a_n(B |\phi_n\rangle) \tag{452}$$

This means that $B |\phi_n\rangle$ is an eigenvector of A with eigenvalue a_n. But since A has no degenerate eigenvalues, $B |\phi_n\rangle$ must be proportional to $|\phi_n\rangle$, so that $B |\phi_n\rangle = b_n |\phi_n\rangle$. Therefore, $|\phi_n\rangle$ is also an eigenvector of B, with eigenvalue b_n.

Since each eigenvector of A is also an eigenvector of B (and vice versa), both of these operators must have a common basis. This basis is unique; it is made of the eigenvectors of A, which are the same as the eigenvectors of B (we say they are joint eigenvectors of A and B). This theorem also holds for any number of mutually commuting Hermitian operators.

---

[61]Degenerate eigenvalues are those that correspond to more than one eigenvector.