Quantum Mechanics

Università degli studi di Roma "La Sapienza"
Physics and Astrophysics BSc

MATTEO CHERI

NOTES ON QUANTUM MECHANICS

MAY 29, 2023

VERSION 0.9β

Quantum Mechanics

Notes on Quantum Mechanics

Written by

Matteo Cheri

Università degli Studi di Roma "La Sapienza"
Physics and Astrophysics BSc

$\LaTeX 2_\varepsilon$ inside, VıM powered.

May 29, 2023

Version 1

Preface

The idea of writing my notes this particular way came from attending the classes my third year of my BSc course in Physics, especially *Meccanica Quantistica e Statistica* with Prof.Presilla and *Struttura della Materia* with Prof.Postorino. Both these beautiful subjects were explained flawlessly during the lectures, but with a huge problem: there are way too many books from where to get your really needed study, and most aren't really books I'd say I like, they're either too long, don't dwelve with enough precision on important topics and most, unfortunately, get lost in a sea of words that just messes with the mind of the reader and helps to make such a beautiful subject even harder than it already is. The reader I'm talking about is obviously me, and my lazy mind thought that it would be better to write this set of notes using the power of typesetting with $\LaTeX 2_\varepsilon$ finally creating a "huge" collection of notes, enlarged using various books I found really helpful in my time passed by studying.

These notes are, obviously, only a student's notes, and might have errors, phrases or whole paragraphs that make no sense and stains of coffee on some equations[1]. I hope that these "little" problems won't hurt the reader but rather encourage him to write a detailed mail to me at this address cheri.1686219@studenti.uniroma1.it, so that I can actually fix the error and send out a new, updated, version.

I'd like to thank *AISF Local Committee Roma Sapienza* and everyone that contributed even with a simple "thanks" or a "keep going". One day when I'll be rich I'll pay you back with a coffee and a hug, thank you.

Rome, May 29, 2023

Sincerely yours, Matteo Cheri

---

[1]

Contents

Part I

Quantum Mechanics

1 The Failure of Classical Physics

The failure of classical physics starts in the first years of the 1900s, when the first experimental measurements on the world of the very small begun. The first discrepancies found, after Planck's quantization of energy "trick" for avoiding the UV catastrophe, were in the experimental results given from the measurements of the wavelength of the emission of Hydrogen, Bremsstrahlung radiation and the famous photoelectric effect.

The first approaches for a correct theoretical modelization of a Hydrogen atom were put forward by Thomson, where the atom itself is considered as a charged sphere, in which there are inside positive and negative charges.

For Hydrogen we will have a sphere of radius a with charge $|q| = e = 1.6 \cdot 10^{-19}$ C. Using Gauss' theorem, we know that the flux of the electric field $\underline{E}$ will be given by the following piecewise function

$$\Phi_E(r) = \begin{cases} 4\pi e \dfrac{r^3}{a^3} & 0 \leq r \leq a \\ 4\pi e & r \geq a \end{cases} \tag{1.1a}$$

Since we are in a spherically simmetrical system, the flux of the $\underline{E}$ field will simply be $4\pi r^2 \underline{E}(r)$, and our $\underline{E}$ field will be

$$\underline{E}(r) = \begin{cases} \dfrac{er}{a}\hat{\mathbf{r}} & 0 \leq r \leq a \\ \dfrac{e}{r^2}\hat{\mathbf{r}} & r \leq a \end{cases} \tag{1.1b}$$

Since $\underline{E}$ is conservative, we can define a scalar potential ϕ such that $\nabla\phi = \underline{E}$. This function is easily determined by the solution of a 1st order ODE

$$\begin{cases} \dfrac{\mathrm{d}\phi}{\mathrm{d}r} = \begin{cases} -\dfrac{er}{a^3} & 0 \leq r \leq a \\ -\dfrac{e}{r^2} & r \geq a \end{cases} \\ \lim\limits_{r\to\infty} (\phi(r)) = 0 \end{cases} \tag{1.2a}$$

The ODE is a separable differential equation with the following solution

$$\phi(r) = \begin{cases} \dfrac{3e}{2a} - \dfrac{er^2}{2a^3} & 0 \leq r \leq a \\ -\dfrac{e^2}{r} & r \geq a \end{cases} \tag{1.2b}$$

5

Since the total charge of the system is $q = -e$, and the potential energy of the system will be given from the scalar potential times the total charge, we have that $V(r) = -e\phi(r)$. Since the ionization energy is defined as $E_I = V(0)$, we get for Hydrogen,

$$E_I = -\frac{3e^2}{2a} \approx -13.6 \text{ eV}$$

For evaluating the emission frequency of this system, we write the Hamiltonian for a harmonic oscillator with the mass of an electron, m_e, coupled to a Hamiltonian with the potential $V(r)$.
The Hamiltonian will then be defined piecewise as such

$$\mathcal{H}(p,r) = \begin{cases} \dfrac{p^2}{2m_e} + \dfrac{1}{2}m_e\omega^2 r^2 \\ \dfrac{p^2}{2m_e} + \dfrac{e^2 r^2}{2a^3} \end{cases} \tag{1.3}$$

Solving the system for ω we get

$$\omega = \sqrt{\frac{e^2}{m_e a^3}}$$

Plugging the measured values of the electron mass and the radius of the charged sphere, we get a frequency $\nu \approx 1.2 \cdot 10^{15}$ Hz and a corresponding wavelength of $\lambda \approx 3 \cdot 10^3$. Although the values obtained from this classical model of the atom core are consistent, the whole idea has been disproven by Geiger and Madsen, whom have demonstrated that if a Gold sheet is irradiated with α particles, most of them will pass through without scattering, some get slightly deflected after interacting with the Gold atoms and another part gets deflected with angles that are completely incompatible with Thomson's model.
A different model which explains the Geiger-Madsen experiment is the Rutherford model of the atom, where the system is evaluated as two opposed charges, the nucleus positively charged, and the orbiting electrons with negative charge.
The total energy of the Rutherford atom can be calculated using the classical Virial theorem, which gives

$$E = \frac{1}{2}m_e v^2 - \frac{e^2}{a} = -\frac{e^2}{2a} \tag{1.4}$$

Forcing in the measured valued for E_I, we get for Hydrogen $a \approx 0.5$ and $\lambda \approx 455$.
Although these values are consistent with observations, the model isn't physically accurate, since the electrons in these orbits have a nonzero acceleration, which brings them to emit sincrotron radiation and consequently lose angular momentum, falling towards the nucleus. If one explicits all the values in game with such system, will get that a random atom will have a mean life of 10^{-8} s, which is completely in contrast with observation, since atoms exist, and didn't get annihilated 10^{-8} s after their formation in the early universe.
The final blow to this huge crisis in classical physics has been given by Einstein and his discovery of the photoelectric effect, where it's shown that, if a metal is irradiated from a source with an energy $E > W$, with W a work function, there is an emission of electrons

linearly proportional to the frequency of the radiation, with coupling constant being exactly Planck's constant, $h = 6.6 \cdot 10^{-34}$ Js. The inverted experiment gives instead a "stopping radiation", Bremsstrahlung in German, where at given frequencies the blackbody radiation of the source gets "stopped". This kind of behavior is not explainable with classical physics, which gave rise to the formulation of quantum mechanics.

The photoelectric effect, with its astonishing results, gave rise to the idea that radiation is quantized, hence it behaves as a particle, and at the same time, due to the certainity behind classical optics, it has the behavior of a wave.

§ 1.1 Old Quantum Mechanics

After the discovery of Bremsstrahlung and the photoelectric effect, there has been an attempt to formalize this new mechanics of quanta by Bohr, through 3 hypotheses

Hypothesis 1 (Bound States). For any atom, only states with discrete energies E_n are allowed, where E_n is a monotonically increasing succession of values, called Energy Levels. The set of these discrete states is called the set of Bound States, the minimum value of this succession is E_0 and is commonly referred to as Ground State.

Hypothesis 2 (Transition Between Levels). When the system is in a bound state radiates only in transitions between levels.

Taking a level E_n and a level E_m where $m > n$, the frequency of the radiation is

$$\nu_{nm} = \frac{|E_m - E_n|}{h}$$

Where h is Planck's Constant.

Corollary 1.1.1 (Ritz Combination Principle). The emission spectrum of an atom, with this hypothesis, is then given by evaluating all the energy differences of the absorption spectrum's difference, hence

$$|\nu_{0n} - \nu_{0m}| = \left| \frac{E_n - E_0}{h} - \frac{E_m - E_0}{h} \right| = \frac{|E_n - E_m|}{h} = \nu_{nm}$$

Hypothesis 3 (Bohr-Sommerfield Quantization). Defining $\hbar$ as $h/2\pi$ as the reduced Planck Constant, we get that the only permitted orbits are those where $L \propto n\hbar$.

For an orbit γ then holds that

$$L = \oint_\gamma p \, \mathrm{d}q = n\hbar$$

For a circular orbit, $L = \mu v r$, hence $\mu v r = n\hbar$

For an Hydrogen atom we then get the following results.

Considering that the system is virialized, we can write the energy as such

$$E = \frac{1}{2}V = -\frac{Ze^2}{2r} \tag{1.5a}$$

The third Bohr hypothesis requires that $\mu v r = n\hbar$, with μ being the reduced mass of the system nucleus-electron.

Squaring the previous relation and inserting it in the expression of energy, we get

$$\frac{1}{2}\mu v^2 = \frac{n^2\hbar^2}{2\mu r^2} = \frac{Ze^2}{2r} \tag{1.5b}$$

From this we get that the only possible orbits are at a radius r_n, where

$$r_n = \frac{n^2\hbar^2}{\mu Ze^2} = \frac{n^2 m_e a_B}{Z\mu}$$

With $a_B = \hbar^2/m_e e^2$ a constant with dimensions of length, called Bohr radius.

Inserting what we have just derived in (1.5a) we get the succession of quantized levels of energy, with the following relation

$$E_n = -\frac{Z^2 e^2 m_e}{2 a_B n^2} \tag{1.5c}$$

Inserting $Z = 1$ for restricting these levels to hydrogen, we get, with the approximation $\mu \approx m_e$ that

$$r_n = n^2 a_B$$

$$a_B = \frac{\hbar^2}{m_e e^2} = 0.53 \tag{1.6}$$

$$E_n = -\frac{e^2}{2 a_B n^2}$$

The values obtained are in accord with experimental results, but the new "theory" of quantum mechanics had yet to be formalized with a set of fundamental principles.

§ 1.2 Wave-Particle Duality

The photoelectric effect, as said before, gave rise to the discovery of the particle-like behavior of light. L. de Broglie, put forward the following problem:

Taking as true the wave like behavior of the electromagnetic radiation, and at the same time taking as true the particle behavior of the same, there must be a particular wavelength for any particle, for which can be then defined a *matter wave*, which has the same ondulatory behavior of classical waves, while still having all the properties of matter.

Taking as an Ansatz the Bohr-Sommerfield quantization hypothesis, we get that

$$pL = nh$$

Dividing by p we finally get that

$$L = n\frac{h}{p}$$

Since for a photon $\lambda = h/p$, we get the following hypothesis

Hypothesis 4 (de Broglie Hypothesis)**.** A wave is associated with each particle, and its wavelength is given by the relation $\lambda = h/p$. All the allowed orbits are those which contain an integer number of wavelengths.

    We can then define the *de Broglie wavelength* of matter as such:
Since $p = \sqrt{2mE}$, we get

$$\lambda_{DB} = \frac{h}{\sqrt{2mE}} \tag{1.7}$$

If preferred, defining a "reduced" de Broglie wavelenght as $\lambdabar = \lambda_{DB}/2\pi$, we get, in terms of $\hbar$

$$\lambdabar = \frac{\hbar}{\sqrt{2mE}}$$

§ 1.3 Experimental Verifications of Quantum Mechanics

§§ 1.3.1 Double Slit Experiment and Quantum Measurement

In order to verify de Broglie's hypothesis of matter waves, many experiments were carried on, but the one that is really worth of notice is the *double slit experiment*.
Let a beam of monochromatic light hit a completely opaque screen, on which there are two parallel slits. Due to the wave nature of electromagnetic radiation, the light passing through the two waves interferes with itself, and if a detector is put on front of such screen, a diffraction pattern can be observed and measured, as expected from classical electromagnetism and optics. In terms of photons, this diffraction pattern indicates where more photons hit the detector and where less did, and the diffraction can be seen as mere interaction with photons that got through different slits.
The real hassle comes when the same experiment is repeated without a continuous beam of light, but with a single photon. In this case an interference pattern «can't» be explained with interacting photons, as there is a single photon in the whole system.
According to the corpuscular model, the photon must pass through one single slit, but the interference pattern wouldn't be observed in this case, giving a further confirmation of the wave-particle duality.
Getting back again to the first case, we can also try to interpret the interference pattern as a sum of the interference patterns of the photons passing only through one of the two slits. Introducing the idea of *state*, we can identify with $|A\rangle, |B\rangle$ the diffraction patterns of the photons passing through a single slit, and with $|C\rangle$ the interference pattern identified in the doble slit experiment.
It's easy to see that the final state isn't exactly a sum of the first two.
Leaving the world of classical physics, we can interpret the final state $|C\rangle$ as being a "mixture" of the state $|A\rangle$ and $|B\rangle$. taking a, b constants, then we can write that

$$|C\rangle = a\,|A\rangle + b\,|B\rangle$$

This result has no similar results in classical physics, since, if evaluated for a single photon, it explicitly indicates that the photon is passing through both slits at the same time!

A similar experiment uses a Mach-Zehnder interferometer, where a beam of neutrons gets shot through two different paths, where at the end of both a detector is carefully placed. The result for a single neutron experiment are basically the same of the double slit experiment, because it is measured that the single neutron passes through both paths, and gets then measured from both detectors.

Modifying the experiment, and putting a detector on both slits, removes this indecision on which slit the particle went through (or path in the interferometer), and would let a detailed description of the construction of the interference pattern to be brought up from data.

What is observed though is extremely different from what is expected, in fact, measuring from which slit the particle goes through completely destroys the interference pattern, and what is measured is a corpuscular behavior of the particles.

This result, tells us that measuring a state, changes it. In fact, it changes the state in such a way that we can either have an interference pattern of a wave-like behavior or a particle-like behavior.

In fact, wave-particle duality still holds, but now gets a whole different meaning: *particles aren't waves nor corpuscules, but both at the same time*. This affirmation has no meaning whatsoever in the classical world that we perceive, and it's the key factor in making quantum mechanics not understandable with physical intuition.

Although not humanly understandable, quantum mechanics can be formalized in a full fledged physical theory with a formal mathematical background, which lets us "understand" quantum mechanics through mathematics. In fact, we can grasp the mathematics of it, even without grasping the physical reality behind it.

2 The Fundamentals

§ 2.1 Dirac Notation

In order to fully grasp this and the future chapters is useful to really know how Dirac notation works in a mathematical framework. In Dirac notation, a vector of a complex Hilbert space $\mathbb{H}$ is indicated with the following notation: $|\cdot\rangle$, called a *ket*, where in place of the dot there is usually a label or an index. Hence, if we have a vector $\underline{a} \in \mathbb{H}$, this vector can be indicated as $|a\rangle \in \mathbb{H}$. Defining the dual space of $\mathbb{H}$ as the space of all linear functionals, i.e. all elements $\underline{b} \in \mathbb{H}^\star$ such that $\underline{b} \cdot \underline{a} = \lambda \in \mathbb{C}$, we define such vectors (or covectors) with the notation $\langle b|$, called *bra*. With this notation, the scalar product simply becomes $\langle b|a\rangle = \lambda$. It's also useful to remember that there exist an isomophism $\iota : \mathbb{H} \to \mathbb{H}^\star$, where it is defined as follows

$$\iota(\underline{a}) = \overline{\underline{a}}^\mathsf{T} = \overline{\underline{a}^\mathsf{T}} \in \mathbb{H}^\star \quad \forall \underline{a} \in \mathbb{H}$$

In Dirac notation, we will then have the following

$$\iota(|a\rangle) = \overline{|a\rangle}^\mathsf{T} = \overline{|a\rangle^\mathsf{T}} = \langle a| \quad \forall |a\rangle \in \mathbb{H}$$

§§ 2.1.1 Generalized Vectors and Tensors in Dirac Notation

Generalizing to tensor spaces, we get the following notation equivalences in the case of rank 2 tensors.
Let T_{ij}, G_i^j, H^{ij} be tensors in the spaces, respectively $\mathbb{H}^\star \otimes \mathbb{H}^\star, \mathbb{H}^\star \otimes \mathbb{H}, \mathbb{H} \otimes \mathbb{H}$
In Dirac notation, since they represent the direct product of two vectors (or covectors), they can be indicated as such

$$\begin{aligned} T_{ij} &\longrightarrow \langle i| \otimes \langle j| = \langle i| \langle j| = \langle ij| \\ G_j^i &\longrightarrow |i\rangle \otimes \langle j| = |i\rangle \langle j| \\ H^{ij} &\longrightarrow |i\rangle \otimes |j\rangle = |i\rangle |j\rangle = |ij\rangle \end{aligned} \tag{2.1}$$

Remembering that the isomorphism between the space and its dual is given by a trasposition and a complex conjugation, we have that the operation of index raising and index lowering

(also known as musical isomorphisms), in Dirac notation will be indicated as follows:

$$T^{ij} = \overline{T_{ij}^{\mathsf{T}}} = \overline{T_{ji}} \longrightarrow |i\rangle\,|j\rangle = \langle j|\,\langle i| \tag{2.2}$$

This is especially useful when treating the contraction of all indexes of a tensor.
Let $R_{ij} = a_i b_j$ and $S^{ij} = c^i d^j$. The product $R_{ij}S^{ij}$ will then be, using the Einstein convention for repeated indexes and the rules defined before for raising and lowering indexes,

$$R_{ij}S^{ij} = a_i b_j c^i d^j = \overline{b^j a^i} c^i d^j \to \langle b|\,\langle a|c\rangle\,|d\rangle = \langle a|c\rangle\,\langle b|d\rangle \tag{2.3}$$

Since a tensor product of two vector spaces is a vector space itself, it's common to directly use superindexes or labels in the labels of the kets and bras, such that a tensor $|ij\rangle = |I\rangle$, where I is the chosen label. It's not hard to then generalize the notation to rank-n tensors. There is only one exception, irreducible tensors. Since they can't be factorized as the direct product of two vectors, they're usually directly indicated with a simple label as before.

§§ 2.1.2 Function Spaces and Linear Operators

Since quantum mechanics can be also represented with elements of function spaces, Dirac notation can be extended to indicate even functions and operators.
Let $\mathcal{F}$ be a function space, and let $\psi \in \mathcal{F}$ be an element of such.
We can write the function $\psi(x) : \mathbb{C} \to \mathbb{R}$ as the projection of the element of $\mathcal{F}$ to the field of complex numbers $\mathbb{C}$. In Dirac notation this becomes

$$\psi(x) \leftrightarrow \langle x|\psi\rangle \tag{2.4}$$

An operator, is a linear endomorphism of $\mathbb{H}$. If we have an operator $\hat{\eta}$ bounded, and an operator $\hat{\sigma}$ unbounded, their definition spaces will be the following

$$\begin{aligned} \hat{\eta} &: \mathbb{H} \to \mathbb{H} \\ \hat{\sigma} &: \mathcal{D} \subset \mathbb{H} \to \mathcal{D} \subset H \end{aligned} \tag{2.5}$$

Since they're linear, if we define a new operator $\hat{\imath}_i$, where $\hat{\imath}_1 = \hat{\eta}$ and $\hat{\imath}_2 = \sigma$, we have, that $\forall \alpha, \beta \in \mathbb{C}$ and $\forall\,|A\rangle, |B\rangle \in \mathcal{D}$, ($\in \mathbb{H}$ if we're considering a bounded operator)

$$\hat{\imath}_i\,(\alpha\,|A\rangle + \beta\,|B\rangle) = \alpha\hat{\imath}_i\,|A\rangle + \beta\hat{\imath}_i\,|B\rangle \tag{2.6}$$

Let's now define an operator $\hat{R}$ such that $\hat{R}\psi(x) = e^{i\pi}\psi(x)$. In Dirac notation it's simply $\hat{R}\,|\psi\rangle = e^{i\pi}\,|\psi\rangle$, nothing changes much.
Although, since operators acting on functions can be seen as tensors (or matrices) acting on vectors, we can write the (infinite) matrix representation of such objects, as such:
Let $\underline{e}_i, \underline{e}_j$ be two basis elements of $\mathcal{F}$. We then will have the following relation, if the scalar product is indicated as $\langle \cdot, \cdot \rangle$

$$R_{ij} = \langle \underline{e}_i, \hat{R}\underline{e}_j \rangle$$

In Dirac notation, its formulation it's similar in shape, and it's written as such

$$R_{ij} = \langle i | \hat{R} | j \rangle$$

If now we define the *adjoint* operation (indicated with (†) as the composition of complex conjugation and transposition ($\square^\dagger = \overline{\square} \circ \square^\mathsf{T}$), we have

$$R_{ij}^\dagger = \langle \underline{e}_i, \hat{R}\underline{e}_j \rangle^\dagger = \langle \underline{e}_j \hat{R}^\dagger, \underline{e}_i \rangle$$

Without writing the equivalent operation in Dirac notation, and just confronting how the adjoint operator works, we get immediately get how it will behave with kets and bras. The operator will act on the right, and its adjoint on the left.
Let $|a\rangle$ be an eigenvector (or eigenket) of $\hat{R}$, with eigenvalue α. We then can express the fact that the operator acts on the right, and its adjoint on the left using $\hat{R}$'s secular equation

$$\hat{R} |a\rangle = \alpha |a\rangle$$
$$\langle a | \hat{R}^\dagger = \langle a | \overline{\alpha}$$

But since applying the double adjoint means applying the identity transformation, we get that

$$\left[\left(\hat{R} |a\rangle \right)^\dagger \right]^\dagger = \left(\langle a | \hat{R}^\dagger \right)^\dagger = (\langle a | \overline{\alpha})^\dagger = \alpha |a\rangle = \hat{R} |a\rangle$$

Hence, an adjoint of an operator acts on the left, and the adjoint of the adjoint operator is the operator itself, and if a ket is a solution to the secular equation, its equivalent bra will be the solution to the adjoint secular equation.
Let's remember that in general, $\hat{R}^\dagger \neq \hat{R}$, hence generally $\hat{R} |A\rangle \neq \langle A | \hat{R}$

§ 2.2 Axioms of Quantum Mechanics

The first postulate of quantum mechanics used to formalize mathematically its structure is the *superposition principle*

Postulate 1 (Superposition Principle). The state of a quantum system is a vector of a separable Hilbert space ($\mathbb{H}$), and if two different vectors are proportional to eachother, they represent the same state. This space must be endowed with a Hermitian scalar product and its dimension is usually infinite.

Since this Hilbert space is a linear vector space, it's algebraically closed, which means that, for two different states $|a\rangle$ and $|b\rangle$, with $\alpha \in \mathbb{C}$, we have

$$\alpha \left(|a\rangle + |b\rangle \right) = \alpha |a\rangle + \alpha |b\rangle = |c\rangle \in \mathbb{H} \tag{2.7}$$

Postulate 2 (Observables). Every quantity that can be observed, hence measured, is called observable, and it's mathematically represented by a self-adjoint operator.
The eigenvalues of this operator are the possible results of the measurement.

   With such definition, the eigenvectors of the self-adjoint operator will be all the states invariant to the action of the operator.

If to every eigenvalue of the observable corresponds one and only one eigenvector, the observable is said to be nondegenerate.

Postulate 3 (Postulate of von Neumann). If a measurement of an observable $\hat{\alpha}$ on a state $|a\rangle$ gives a degenerate eigenvalue α, the eigenstate of the observable will then be given by the projection of $|a\rangle$ onto the eigenspace of $\hat{\alpha}$, corresponding to the subspace generated by all vectors that satisfy the equation

$$\hat{\alpha}\,|r_i\rangle = \alpha\,|r_i\rangle$$

The searched state will then be

$$|s\rangle = \sum_i p_i\,|r_i\rangle$$

Postulate 4 (Transition Probabilities). Let $\hat{\beta}$ be an observable, hence $\hat{\beta} = \hat{\beta}^\dagger$. Consider now a state $|d\rangle$. If $\hat{\beta}\,|d\rangle = p\,|f\rangle$, we can define the probability of the transition $|d\rangle \to |f\rangle$ as follows:

$$P\left(|d\rangle \to |f\rangle\right) = \frac{|\langle f|d\rangle|^2}{\langle d|d\rangle\,\langle f|f\rangle} \leq 1$$

   There are two main cases for the study of transition probabilities, one for nondegenerate states and one for degenerate states. Since the case for nondegenerate states is a particular case of the degenerate case, we can study directly the latter.

Let $\hat{S}$ be a degenerate observable. For a state $|A\rangle$ (with $\langle A|A\rangle = 1$), we will then have, after a projection to an orthonormal basis of eigenstates $|s_i\rangle$, as indicated in the von Neumann postulate

$$|A\rangle = \sum_i a_i\,|s_i\rangle$$

Let's define a state $\hat{S}\,|A\rangle = |k\rangle = \sum_i a_i\,|s_i\rangle$ The problem is straightforward. What's the transition probability from a state $|A\rangle$ to a state $|s_i\rangle$ (with i fixed)?

From the last postulate we can write

$$p_i = P\left(|A\rangle \to |k\rangle\right) = \frac{|\langle k|A\rangle|^2}{\langle k|k\rangle}$$

Due to the orthonormality of $|s_i\rangle$ ($\langle s_i|s_j\rangle = \delta_{ij}$) we will have

$$p_i = \sum_i |a_i|^2$$

Which is a direct consequence of von Neumann's postulate.

Hence, in order to treat a general state $|B\rangle$ which is not an eigenstate of a given observable $\hat{b}$, we can apply a projection, as stated from von Neumann. But how does this projection

work? What's its mathematical form?

The answers to these questions are quite easy to think what shape would they take. As the projection of a state can be interpreted as a Fourier series, we get that the projection will be given by an operator that sends $|B\rangle$ to a set of eigenstates $|b_i\rangle$, multiplied by some constant which is exactly given by the scalar product $\langle B|b_i\rangle$. Summarizing, if we indicate the projected state with $|b\rangle$

$$|b\rangle = \hat{\pi}_i |B\rangle = \sum_i \langle b_i|B\rangle |b_i\rangle$$

$$\hat{\pi}_i = |b_i\rangle \langle b_i|$$

(2.8)

Where $\hat{\pi}_i^\dagger = \hat{\pi}_i$ and $\hat{\pi}^2 = \hat{\pi}$, due to the definition of the projector operator.

Moreover, we can even define a theorem from this definition, where, if $|b_i\rangle$ is a complete orthonormal basis, then

$$|b_i\rangle \langle b_i| = \hat{\mathbb{1}}$$

(2.9)

Which indicates the completeness relation for a basis.

We subsequently define the expectation value of an observable, its variance, commutators and anticommutators

Definition 2.2.1 (Mean Value). Given an observable $\hat{\eta}$, we can define its mean value on a state $|a\rangle$ as follows

$$\langle \hat{\eta} \rangle_{|a\rangle} = \langle a| \hat{\eta} |a\rangle = \sum_i \langle a| \hat{\pi}_i \hat{\eta} \hat{\pi}_i |a\rangle = \sum_i p_i \eta_i, \quad p_i = \frac{N_i}{N}$$

(2.10)

Due to the arbitrary definition of $\hat{\eta}$ and $|a\rangle$, what has been written is valid generally

Definition 2.2.2 (Variance). If we define the variance as $Var(x) = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$, we can define the variation of an observable on a state $|a\rangle$ as such

$$Var(\hat{\eta}) = \sqrt{\langle a| \hat{\eta}\hat{\eta} |a\rangle - \langle a| \hat{\eta} |a\rangle \langle a| \hat{\eta} |a\rangle} = \sqrt{\langle \hat{\eta}^2 \rangle_{|a\rangle} - \langle \hat{\eta} \rangle_{|a\rangle}^2}$$

(2.11)

Definition 2.2.3 (Commutators and Anticommutators). The commutator is an operator of operators defined for $\hat{\eta}, \hat{\gamma}$ as such:

$$[\hat{\eta}, \hat{\gamma}] = \hat{\eta}\hat{\gamma} - \hat{\gamma}\hat{\eta}$$

(2.12)

The anticommutator is defined as such

$$\{\hat{\eta}, \hat{\gamma}\} = \hat{\eta}\hat{\gamma} + \hat{\gamma}\hat{\eta}$$

(2.13)

In general, the commutator is antihermitian and bilinear, whereas an anticommutator is hermitian and bilinear.

Theorem 2.1 (Compatibility). *Two operators $\hat{a}$ and $\hat{b}$ are said compatible, if there exists a common basis of eigenstates and their commutator is equal to 0*

Proof. If $\hat{a}$ and $\hat{b}$ are compatible, then there exists a common basis of eigenstates. Let $|A\rangle \neq 0$ be such basis, then, such statement is true

$$\begin{cases} \hat{a}\,|A\rangle = \alpha\,|a\rangle \\ \hat{b}\,|A\rangle = \beta\,|a\rangle \end{cases} \tag{2.14}$$

Then,

$$\left[\hat{a}, \hat{b}\right]|A\rangle = (\hat{a}\hat{b} - \hat{b}\hat{a})\,|A\rangle =$$
$$= \hat{a}\hat{b}\,|A\rangle - \hat{b}\hat{a}\,|A\rangle = \beta\hat{a}\,|A\rangle - \alpha\hat{b}\,|A\rangle = \tag{2.15}$$
$$= \beta\alpha\,|A\rangle - \alpha\beta\,|A\rangle = 0$$

But, if $\left[\hat{a}, \hat{b}\right]|A\rangle = 0$

$$\hat{a}\,|A\rangle = \alpha\,|A\rangle$$
$$\hat{b}\alpha\,|A\rangle = \alpha\beta\,|A\rangle \tag{2.16}$$
$$(\hat{a}\hat{b} - \hat{b}\hat{a})\,|A\rangle = 0$$

As expected from the statement of the theorem. $\qquad\square$

Postulate 5 (Canonical Quantization). There exists a strict relation between commutators and Poisson brackets. If we define the Poisson brackets as $[\cdot,\cdot]_{PB}$, we can quantize Poisson brackets with the following relation

$$[\cdot,\cdot] = i\hbar[\cdot,\cdot]_{PB}$$

Since $[q_i, p_j]_{PB} = \delta_{ij}$, in quantum mechanics holds this fundamental relation

$$[\hat{q}, \hat{p}] = i\hbar\hat{\mathbb{1}} \tag{2.17}$$

This is known as canonical quantization, where position and momentum are represented by observables.

Postulate 6 (Heisenberg Uncertainity). In quantum mechanics there is an intrinsic uncertainty on measurements.

Taking two observables $\hat{a}$ and $\hat{b}$, we define an operator $\alpha = \hat{a} + ix\hat{b}$ with $x \in \mathbb{R}$. Its expectation value on a state $|s\rangle$ will be the following

$$\langle s|\,\hat{\alpha}^\dagger\hat{\alpha}\,|s\rangle \leq \langle s|\,\hat{\alpha}^\dagger\,|s\rangle\,\langle s|\,\hat{\alpha}\,|s\rangle$$

Since $\hat{\alpha}^\dagger\hat{\alpha} = \hat{a}^2 + x^2\hat{b}^2 + ix\left[\hat{a}, \hat{b}\right]$, we get that

$$\langle\hat{a}^2\rangle + x^2\left\langle\hat{b}^2\right\rangle + ix\left\langle\left[\hat{a}, \hat{b}\right]\right\rangle \geq \langle\hat{a}\rangle^2 + x^2\left\langle\hat{b}\right\rangle^2$$

Hence

$$\langle \hat{a}^2 \rangle - \langle \hat{a} \rangle^2 + x^2 \left(\langle \hat{b}^2 \rangle - \langle \hat{b} \rangle^2 \right) - ix \left\langle \left[\hat{a}, \hat{b} \right] \right\rangle \geq 0$$

Rewriting $\langle c^2 \rangle - \langle c \rangle^2 = \sigma_c^2$, with c arbitrary and σ its standard deviation, we get

$$\sigma_a^2 - x^2 \sigma_b^2 + ix \left\langle \left[\hat{a}, \hat{b} \right] \right\rangle \geq 0$$

In order for this equation to be true, the discriminant of the quadratic equation must be greater than 0, hence

$$- \left\langle \left[\hat{a}, \hat{b} \right] \right\rangle^2 + 4\sigma_a^2 \sigma_b^2 \geq 0$$

And finally we get

$$\sigma_a \sigma_b \geq \frac{1}{2} \left\langle \left[\hat{a}, \hat{b} \right] \right\rangle$$

Taking $\hat{a} = \hat{q}$ and $\hat{b} = \hat{p}$, and following canonical quantization rules, we get that momentum and position can't be measured simultaneously due to a fundamental uncertainity, given by Heisenberg's uncertainity principle

$$\sigma_q \sigma_p \geq \frac{1}{2} \hbar \tag{2.18}$$

§ 2.3 Representation Theory

§§ 2.3.1 Heisenberg Representation

In quantum mechanics, we can define two main representations: Heisenberg representation or Schrödinger representation.

Heisenberg representation is given by matrix elements of operators, through an isomorphism between l_2 and the Hilbert space of quantum configurations $\mathbb{H}$. Defining a CON basis (complete orthonormal) $|e_i\rangle$ on $\mathbb{H}$, we get that, $\forall |A\rangle \in \mathbb{H}$

$$|A\rangle = \hat{\mathbb{1}} |A\rangle = \sum_i |e_i\rangle \langle e_i|A\rangle = \sum_i a_i |A\rangle \tag{2.19}$$

Where the operator $\hat{\pi} = |e_i\rangle \langle e_i|$ is defined as $\hat{\pi} : \mathbb{H} \to l_2$, which it's simply a projection, as defined in (2.8).

Hence, we can define the operation of a general operator $\hat{\rho}$ in l_2 as a combination of the projection $\hat{\pi}$ and $\hat{\rho}$. So, if $\rho |A\rangle = |B\rangle$, we have that

$$b_n = \langle e_n| \hat{\rho} \sum_i |e_i\rangle \langle e_i|A\rangle \to b_n = \rho_{ni} a_i \in l_2 \tag{2.20}$$

Where ρ_{ni} is the matricial representation in l_2 of the operator $\hat{\rho}$.

Theorem 2.2 (Block Representation for Degenerate Observables)*. Let $\hat{a}$ and $\hat{b}$ be two observables. If they're compatible, and the basis of eigenvectors of $|a\rangle$ is degenerate, then $\hat{b}$ can be represented as a block matrix*

Proof. Let $|e_i\rangle$ be such degenerate base, then we can write b_{ij} as such

$$b_{ij} = \langle e_i | \hat{b} | e_j \rangle$$

Since $\langle e_i | e_j \rangle = 1$ for all the degenerate i, j, and it's 0 elsewhere, b_{ij} will be a block matrix. $\square$

§§§ 2.3.1.1 Unitary Transformations

If $|e_i\rangle$ and $|r_i\rangle$ are two ON bases in $\mathbb{H}$, we can define an unitary operator of base change $\hat{U}$ with the following equation.

$$\hat{U} |e_i\rangle = |r_i\rangle \tag{2.21}$$

Since the two bases are ON, this operator will be defined as the projection $|r_i\rangle \langle r_i|$.
It follows, in order for the equation (2.21) to be true, $\langle r_i | e_j \rangle = \delta_{ij}$, and for $\hat{U}$, it must hold that

$$\hat{U}\hat{U}^{\dagger} = \hat{U}^{\dagger}\hat{U} = \hat{\mathbb{1}} \tag{2.22}$$

The last relation can be reduced to the fact that $\hat{U}$ must be left unitary and right unitary.

Theorem 2.3 (Von Neumann Theorem)*. Quantization is invariant to unitary transformations*

Proof. Let $\hat{U}$ be an unitary operator, for which $\tilde{\hat{q}} = \hat{U}\hat{q}\hat{U}^{\dagger}$ and $\tilde{\hat{p}} = \hat{U}\hat{p}\hat{U}^{\dagger}$.
Then, since $\hat{U}^{\dagger}\hat{U} = \hat{\mathbb{1}}$

$$\begin{cases} \left[\tilde{\hat{q}}, \tilde{\hat{p}}\right] = \hat{U}\hat{q}\hat{U}^{\dagger}\hat{U}\hat{p}\hat{U}^{\dagger} - \hat{U}\hat{p}\hat{U}^{\dagger}\hat{U}\hat{q}\hat{U}^{\dagger} = \\ = \hat{U}\hat{q}\hat{\mathbb{1}}\hat{p}\hat{U}^{\dagger} - \hat{U}\hat{p}\hat{\mathbb{1}}\hat{q}\hat{U}^{\dagger} = \hat{U}[\hat{q}, \hat{p}]\hat{U}^{\dagger} = i\hbar\hat{\mathbb{1}} \end{cases} \tag{2.23}$$

The same holds for every commutator and anticommutator. $\square$

§§ 2.3.2 Schrödinger Representation

Schrödinger representatio is obtained through an isomorphism between $\mathbb{H}$ and $L^2(\mathbb{R})$.
Let's consider a space $\mathbb{H}$ formed by the direct product of n Hilbert spaces, then $\mathbb{H} \simeq L^2(\mathbb{R}^n)$.
The isomorphism between these two spaces is given through a projection to the space $\mathbb{R}^n$, indicated as such.
Let $|A\rangle \in \mathbb{H}$, then $\exists \hat{S} : \mathbb{H} \to L^2(\mathbb{R}^n)$ defined as such

$$\hat{S} |A\rangle = \langle x_i | A \rangle = \psi_A(x_i) \in L^2(\mathbb{R}^n) \tag{2.24}$$

Where the function ψ_A is called *wavefunction*.
In this space, the scalar product is defined through the following integral

$$\langle A | B \rangle \to \int_{\mathbb{R}^n} \overline{\psi_B(x_i)} \psi_A(x_j) \, \mathrm{d}^n x \tag{2.25}$$

In order to be a valid representation of quantum mechanics, canonical quantization must hold, and the momentum and position operators will be defined as such, in position space

$$\begin{cases} p_i \to -i\hbar\nabla_x \\ q_i \to \hat{x}_i \end{cases} \tag{2.26a}$$

And in momentum space as such

$$\begin{cases} p_i \to \hat{k}_i \\ q_i \to -i\hbar\nabla_p \end{cases} \tag{2.26b}$$

The canonical quantization rules will then become, for position space

$$[\hat{q}, \hat{p}] \, |A\rangle \to -i\hbar \left(x_i \nabla_j - \nabla_j x_i \right) \psi_A = i\hbar \delta_{ij} \psi_A \tag{2.27}$$

Since a quantum state must be normalizable, in order to have a proper probability of transition, we need that even the wavefunction must be normalizable.
The normalization equation will then be

$$N \langle A|A \rangle = 1 \to N \int_{\mathbb{R}^n} \overline{\psi_A} \psi_A \, \mathrm{d}^n x = N \int_{\mathbb{R}^n} |\psi_A|^2 \, \mathrm{d}^n x = 1 \tag{2.28}$$

Since $\psi_A \in L^2$, the integral converges, and N^{-1} is the searched normalization constant.
Due to the definition of probability density function, we have that the absolute value squared of the wavefunction is the probability density of finding the particle in a certain position (in position representation) or in a certain momentum (momentum representation). Mean values and superior statistical models are then calculated as
In order to switch between representation, we can use Fourier transforms. Defining the integral operator $\hat{\mathcal{F}}$ as the Fourier transform operator, we get, if $\phi(k_i)$ is the momentum wavefunction and $\psi(x_i)$ the position wavefunction

$$\phi(k_i) = \hat{\mathcal{F}}_k \psi(x_i) \tag{2.29}$$

And, defining an inverse Fourier transform operator $\hat{\mathcal{F}}^{-1}$

$$\psi(x_i) = \hat{\mathcal{F}}_x^{-1} \phi(k) \tag{2.30}$$

§§ 2.3.3 Hamiltonian Operators

In order to define a function for a whole system, a Lagrangian can be written. In non relativistic quantum mechanics although, its Legendre tranform is used, the Hamiltonian. We then have, for a system with a general potential, that the Hamiltonian will be given in this general form

$$\mathcal{H} = \frac{p^2}{2m} + V(x_i) \tag{2.31}$$

Quantizing, in Heisenberg representation, we will get an observable, a self adjoint operator:

$$\hat{\mathcal{H}} = \frac{\hat{p}^2}{2m} + V(x_i) \tag{2.32}$$

This operator is used, since its eigenvalues are the energy levels of the system.
Hence, given an eigenstate $|E\rangle$, we will have the following equation

$$\hat{\mathcal{H}} |E\rangle = E |E\rangle \tag{2.33}$$

Which is the secular equation of the Hamiltonian, with eigenket $|E\rangle$ and eigenvalue E. Remembering (2.26a), we will get the Schrödinger representation of the Hamiltonian operator

$$\hat{\mathcal{H}} = -\frac{\hbar^2}{2m} \nabla^2 + V(x_i) \tag{2.34}$$

The secular equation will then become a eigenfunction equation for the differential operator $\hat{\mathcal{H}}$, and the energy levels will be the spectrum of the operator

$$\hat{\mathcal{H}}\psi = -\frac{\hbar^2}{2m} \nabla^2 \psi + V(x_i)\psi = E\psi(x_i) \tag{2.35}$$

The solution to this equation, commonly called *Schrödinger equation* is the wavefunction of the system, with associated energy eigenvalue E. Due to what said before, Schrödinger and Heisenberg representations are equivalent, and a problem can be solved either as an eigenstate problem or an eigenfunction problem. The most common example of problem that can be solved with both methods is the Quantum Harmonic Oscillator, treated in the next chapter.

3 Quantum Dynamics in 1D

§ 3.1 Free Particle

The motion of a free particle is described by the following Hamiltonian:

$$\mathcal{H} = \frac{p^2}{2m} \tag{3.1}$$

Quantizing in Heisenberg representation (3.1), we get the following operatorial representation

$$\hat{\mathcal{H}} = \frac{\hat{p}^2}{2m} \tag{3.2}$$

Since $\left[\hat{\mathcal{H}}, \hat{p}\right] = 0$, there exists a common basis of eigenvectors that we will indicate with $|p\rangle$, for which

$$\hat{p} |p\rangle = p |p\rangle$$

Hence, the secular equation for the Hamiltonian will be

$$\hat{\mathcal{H}} |p\rangle = \frac{1}{2m} \hat{p}^2 |p\rangle = \frac{p^2}{2m} |p\rangle \tag{3.3}$$

The spectrum of the Hamiltonian operator is twice degenerate, since momentum can be either negative or positive, and the final energy state will then be given by the linear combination of the two states with positive and negative momentum, with energies $\pm \frac{1}{2m} p^2$

$$\hat{\mathcal{H}} |E\rangle = \frac{p^2}{2m} |p\rangle - \frac{p^2}{2m} |-p\rangle \tag{3.4}$$

It's obvious, from this equation, that the particle is delocalized, since the momentum is positive and negative at the same time, meaning that the particle is "going" both forward and backward.
Representing the Hamiltonian in the Schrödinger form we will get instead

$$\hat{\mathcal{H}} = -\frac{\hbar^2}{2m} \frac{\mathrm{d}^2}{\mathrm{d}x^2} \tag{3.5}$$

The Schrödinger equation will then be

$$\hat{\mathcal{H}}\psi = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} - E\psi = 0 \tag{3.6}$$

Solving this simple 2nd order ODE, we get the following wavefunction

$$\psi(x) = Ae^{i\sqrt{\frac{2mE}{\hbar^2}}x} + Be^{-i\sqrt{\frac{2mE}{\hbar^2}}x} \tag{3.7}$$

Where its twofold degeneracy is obvious

Theorem 3.1 (Degeneration Theorem). *Let $\hat{p}, \hat{\mathcal{H}}, \hat{I}$ be three observables. If*

$$\begin{cases} \left[\hat{p}, \hat{I}\right] \neq 0 \\ \left[\hat{p}, \hat{\mathcal{H}}\right] = \left[\hat{p}, \hat{H}\right] = 0 \end{cases} \tag{3.8}$$

Then $\hat{I}$ is 2-degenerate

Proof. Considering the case of a free particle, introducing operator $\hat{I}$ as a spatial inversion operator, we have that

$$\begin{cases} \left[\hat{p}, \hat{\mathcal{H}}\right] = \left[\hat{I}, \hat{\mathcal{H}}\right] = 0 \\ \left[\hat{p}, \hat{I}\right] \neq 0 \end{cases}$$

Where $\hat{I}|p\rangle = \pm|-p\rangle$ It's obvious that $\hat{I}^2 = \hat{\mathbb{1}}$ and that $\left\{\hat{I}, \hat{p}\right\} = \left\{\hat{I}, \hat{q}\right\} = 0$.
Due to this fact, we can write a new state which is the sum of the simmetrization and antisimmetrization of a the first state found.
Since $\hat{I}|\pm p\rangle = \pm|\pm p\rangle$ we get this state, indicating the antisimmetric one with $|p\rangle_a$ and the simmetric one as $|p\rangle_s$

$$|p\rangle = \frac{1}{2}\left(|p\rangle_s + |-p\rangle_s\right) + \frac{1}{2}\left(|p\rangle_a - |-p\rangle_a\right) \tag{3.9}$$

Which is also an eigenstate of $\hat{\mathcal{H}}$, since it commutes with $\hat{I}$ $\qquad\square$

§ 3.2 Quantum Harmonic Oscillator

§§ 3.2.1 Dirac Formulation

The classical harmonic oscillator, is described by the following Hamiltonian

$$\mathcal{H} = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2 q^2 \tag{3.10}$$

Quantizing

$$\hat{\mathcal{H}} = \frac{1}{2m}\hat{p}^2 + \frac{1}{2}m\omega^2\hat{q}^2 \tag{3.11}$$

It's evident that the Hamiltonian is an observable, hence its eigenvalues will be real.
A solution in operatorial representation can be given, defining two operators, called *creation*
and *annihilation* operators, defined as follows

$$\begin{cases} \hat{\eta} = \dfrac{1}{\sqrt{2m\hbar\omega}}(\hat{p} - im\omega\hat{q}) \\[2mm] \hat{\eta}^\dagger = \dfrac{1}{\sqrt{2m\hbar\omega}}(\hat{p} + im\omega\hat{q}) \end{cases} \tag{3.12}$$

Inverting the relations and expressing position and momentum in terms of creation and
distruction operators, we get

$$\begin{cases} \hat{p} = \sqrt{\dfrac{m\hbar\omega}{2}}(\hat{\eta}^\dagger + \hat{\eta}) \\[3mm] \hat{q} = -i\sqrt{\dfrac{\hbar}{2m\omega}}(\hat{\eta}^\dagger - \hat{\eta}) \end{cases} \tag{3.13}$$

In order to evaluate the commutators of these two new operators $\hat{\eta}$, we evaluate the product
of the two, from the left and right, obtaining the following result

$$\hat{\eta}^\dagger\hat{\eta} = \frac{1}{2m\hbar\omega}(\hat{p}^2 + m^2\omega^2\hat{q}^2 - im\omega[\hat{q},\hat{p}]) = \frac{1}{\hbar\omega}\left(\hat{\mathcal{H}} - \frac{1}{2}\hbar\omega\hat{\mathbb{1}}\right) \tag{3.14a}$$

And

$$\hat{\eta}\hat{\eta}^\dagger = \frac{1}{2m\hbar\omega}(\hat{p}^2 + m^2\omega^2\hat{q}^2 + im\omega[\hat{q},\hat{p}]) = \frac{1}{\hbar\omega}\left(\hat{\mathcal{H}} + \frac{1}{2}\hbar\omega\hat{\mathbb{1}}\right) \tag{3.14b}$$

Having evaluated this, we therefore get

$$\begin{aligned} \left[\hat{\eta},\hat{\eta}^\dagger\right] &= \frac{1}{2m\hbar\omega}(-2im\omega[\hat{q},\hat{p}]) = \hat{\mathbb{1}} \\[2mm] \left[\hat{\eta}^\dagger,\hat{\eta}\right] &= \frac{1}{2m\hbar\omega}(2im\omega[\hat{q},\hat{p}]) = -\hat{\mathbb{1}} \end{aligned} \tag{3.15}$$

And therefore, inverting the previous relation where the Hamiltonian appears on the right
hand side, and utilizing the commutators in order to switch the position of the multiplication
of the two operators, we get these two new expressions for the Hamiltonian

$$\hat{\mathcal{H}} = \hbar\omega\left(\hat{\eta}^\dagger\hat{\eta} + \frac{1}{2}\hat{\mathbb{1}}\right) \tag{3.16a}$$

$$\hat{\mathcal{H}} = \hbar\omega\left(\hat{\eta}\hat{\eta}^\dagger - \frac{1}{2}\hat{\mathbb{1}}\right) \tag{3.16b}$$

The commutators between the Hamiltonian and the annihilation/creation operators is
calculated easily, utilizing the latter form of the Hamiltonian and the properties of the
commutator (remembering that operators in general «do not» commute)

$$\begin{aligned} \left[\hat{\mathcal{H}},\hat{\eta}\right] &= \hbar\omega\left(\left[\hat{\eta}^\dagger\hat{\eta},\hat{\eta}\right] + \frac{1}{2}[\hat{\mathbb{1}},\hat{\eta}]\right) = \hbar\omega\left(\hat{\eta}^\dagger[\hat{\eta},\hat{\eta}] + \left[\hat{\eta}^\dagger,\hat{\eta}\right]\hat{\eta}\right) = -\hbar\omega\hat{\eta} \\[2mm] \left[\hat{\mathcal{H}},\hat{\eta}^\dagger\right] &= \hbar\omega\left(\hat{\eta}^\dagger\left[\hat{\eta},\hat{\eta}^\dagger\right] + \left[\hat{\eta}^\dagger,\hat{\eta}^\dagger\right]\hat{\eta}\right) = \hbar\omega\hat{\eta} \end{aligned} \tag{3.17}$$

Consequently, writing the secular equation of the system we get the following results

$$\hat{\mathcal{H}}\hat{\eta}\ket{E} = \left(\hat{\eta}\hat{\mathcal{H}} + \left[\hat{\mathcal{H}},\hat{\eta}\right]\right)\ket{E} = \left(\hat{\eta}\hat{\mathcal{H}} - \hbar\omega\hat{\eta}\right)\ket{E} = (E - \hbar\omega)\,\hat{\eta}\ket{E}$$
$$\hat{\mathcal{H}}\hat{\eta}^\dagger\ket{E} = \left(\hat{\eta}^\dagger\hat{\mathcal{H}} + \left[\hat{\mathcal{H}},\hat{\eta}^\dagger\right]\right)\ket{E} = \left(\hat{\eta}^\dagger\hat{\mathcal{H}} + \hbar\omega\hat{\eta}^\dagger\right)\ket{E} = (E + \hbar\omega)\,\hat{\eta}^\dagger\ket{E} \tag{3.18}$$

So, the action of these operators actually creates and annihilates quantas of energy, with a spacing of $\hbar\omega$, creating a ladder. This characteristics gives them also the name of *ladder operators*.

Since the energies of an harmonic oscillator can't obtain negative values, there must exist a state for which the application of the annihilation operators "annihilates" the system, returning back zero. Indicating the energy eigenvalues with $\ket{n}$, we identify such state with $\ket{0}$ (watch out, this is «not» the null vector of the Hilbert space, 0 is simply a label), such state is usually called *vacuum* state.

Its energy can be determinated without much hassle, noting that the square norm of 0 is still 0 (obviously). Rewriting 0 as $\hat{\eta}\ket{0}$, we get, in Dirac notation, and remembering (3.14a)

$$\bra{0}\hat{\eta}^\dagger\hat{\eta}\ket{0} = \frac{1}{\hbar\omega}\bra{0}\left(\hat{\mathcal{H}} - \frac{1}{2}\hbar\omega\hat{\mathbb{1}}\right)\ket{0} = E_0 - \frac{1}{2}\hbar\omega = 0 \tag{3.19}$$

It's clear then, that the minimum possible value of energy is $E_0 = \frac{1}{2}\hbar\omega$, usually called *zero point energy* In order to properly see how the annihilation and creation operators act on the energy eigenstates it's useful to normalize their succession.

We know, as for (3.18), that creation operators increase by one the number of quants in the system, hence, knowing that there is a ground state, indicated with $\ket{0}$, we can write the n-th state as follows

$$\hat{\eta}^n\ket{0} = \nu\ket{n}$$

With $\nu \in \mathbb{C}$. Since the normalization condition asks that the square braket of the vector must be equal to one, we simply get the following expression

$$\bra{0}\hat{\eta}^n\hat{\eta}^{\dagger n}\ket{0} = |\nu|^2\braket{n|n} \tag{3.20}$$

Since the former operator multiplication can be written as $\hat{\eta}^{n-1}\left[\hat{\eta},\hat{\eta}^{\dagger n}\right]$. It can be easily demonstrated that the previous commutator gives the following value

$$\left[\hat{\eta},\hat{\eta}^{\dagger n}\right] = n\hat{\eta}^{\dagger(n-1)}$$

We can then write (3.20) as

$$n\bra{0}\hat{\eta}^{n-1}\hat{\eta}^{\dagger(n-1)}\ket{0} = \nu^2\braket{n-1|n-1}$$

Iterating, we get the following condition

$$\nu^2 = \frac{1}{n!} \tag{3.21}$$

Henceforth, we get $\nu = \frac{1}{\sqrt{n!}}$, without any loss of generality.

The normalization of the n-th state, can then be defined recursively starting from $|0\rangle$

$$|n\rangle = \frac{1}{\sqrt{n!}} \hat{\eta}^{\dagger n} |0\rangle \tag{3.22}$$

Since all eigenstates must be normalized, we get that the action of the creation and annihilation operators will be, respectively

$$\begin{aligned} \hat{\eta}^{\dagger} |n\rangle &= \sqrt{n+1} |n+1\rangle \\ \hat{\eta} |n\rangle &= \sqrt{n} |n-1\rangle \end{aligned} \tag{3.23}$$

Defining a *particle number operator* as follows

$$\hat{N} = \hat{\eta}^{\dagger} \hat{\eta} \tag{3.24}$$

We get the following commutation relations

$$\begin{cases} \left[\hat{N}, \hat{\eta}^{\dagger}\right] = \hat{\eta}^{\dagger} \\ \left[\hat{N}, \hat{\eta}\right] = -\hat{\eta} \end{cases} \tag{3.25}$$

And the new Hamiltonian in terms of $\hat{N}$ becomes, evidently

$$\hat{\mathcal{H}} = \hbar\omega \left(\hat{N} + \frac{1}{2}\hat{\mathbb{1}}\right) \tag{3.26}$$

Calculating the commutator $\left[\hat{\mathcal{H}}, \hat{N}\right]$ explicitly we get

$$\begin{aligned} \left[\hat{\mathcal{H}}, \hat{N}\right] &= \left[\hbar\omega \left(\hat{N} + \frac{1}{2}\hat{\mathbb{1}}\right), \hat{N}\right] = \hbar\omega \left[\hat{N} + \frac{1}{2}\hat{\mathbb{1}}, \hat{N}\right] = \\ &= \hbar\omega \left(\left[\hat{N}, \hat{N}\right] + \frac{1}{2}\left[\hat{\mathbb{1}}, \hat{N}\right]\right) = 0 \end{aligned} \tag{3.27}$$

Since $\left[\hat{\mathcal{H}}, \hat{N}\right] = 0$, we get from Theorem 2.1 on the compatibility of operators, that there exist a common ON basis between $\hat{N}$ and $\hat{\mathcal{H}}$. This basis will obviously be the energy eigenstates of the Hamiltonian, and remembering the action of $\hat{\eta}$ and $\hat{\eta}^{\dagger}$ on the energy eigenkets, we get that the number operator will act on these states as such

$$\hat{N} |n\rangle = n |n\rangle \tag{3.28}$$

It's now obvious why such operator is called *particle number* operator, since applying it to an energy eigenstate, it will give as an eigenvalue the "number of particles" present.

Since these eigenstates are eigenstates for both the particle number operator and the

Hamiltonian, we can write the secular equation of the Hamiltonian (3.26), which gives immediately the following result

$$\hat{\mathcal{H}} |n\rangle = \hbar\omega \left(n + \frac{1}{2} \right) |n\rangle \tag{3.29}$$

Since $\hat{\mathcal{H}} |n\rangle = E_n |n\rangle$, we get that energy is quantized, and has the following expression

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right) \tag{3.30}$$

From this definition, we can define the ground state of the system as $|0\rangle$, since, applying the annihilation operator on $|0\rangle$ the state gets "annihilated" and its action gives 0, we get our ground state energy (also known as *zero point energy*)

$$\hat{\mathcal{H}} |0\rangle = \frac{1}{2} \hbar\omega |0\rangle \tag{3.31}$$

Switching to a Schrödinger representation for the definition (3.22), we can also find the eigenfunctions of the Hamiltonian, $\langle x|n\rangle = \psi_n(x)$. Since the operator $\hat{\eta}^\dagger$ and $\hat{\eta}$ are defined in (3.12), we can then write the normalized eigenfunctions as the solutions for the following differential equation. For the ground state, we will have, applying the annihilation operator

$$\hat{\eta}\psi_0(x) = -\frac{1}{\sqrt{2m\hbar\omega}} \left(i\hbar\nabla + im\omega x \right) \psi_0(x) = 0 \tag{3.32}$$

Its solution is simply given by the following exponential

$$\psi_0(x) = A e^{-\frac{m\omega}{2\hbar} x^2}$$

The normalization condition will be given by the fact that $\psi_0 \in L^2(\mathbb{R})$, hence it must be square integrable. Using Gauss' identity, we get

$$|A|^2 \int_{-\infty}^{\infty} e^{-\frac{m\omega}{\hbar} x^2} \, \mathrm{d}x = \sqrt{\frac{\hbar\pi}{m\omega}}$$

The constant A is easily determined, and the normalized wavefunction for the ground state is

$$\psi_0(x) = \sqrt{\sqrt{\frac{m\omega}{\hbar\pi}}} e^{-\frac{m\omega}{2\hbar} x^2} \tag{3.33}$$

Now, applying the operator $\hat{\eta}^\dagger$ multiple times we can get the n-th wavefunction. In formulae, we get that

$$\frac{1}{\sqrt{n!}} \hat{\eta}^{\dagger n} \psi_0(x) = \frac{1}{(2m\hbar\omega)^{\frac{n}{2}}} \left(im\omega x - i\hbar\nabla \right)^n \psi_0(x) = \psi_n(x)$$

Substituting the function we found for ψ_0, we get the following differential equation of the n-th order

$$\psi_n(x) = \frac{1}{\sqrt[n]{2m\hbar\omega}} \sqrt{\frac{1}{n!} \sqrt{\frac{m\omega}{\hbar\pi}}} \left(im\omega x - i\hbar \frac{\mathrm{d}}{\mathrm{d}x} \right)^n e^{-\frac{m\omega}{2\hbar}x^2} \tag{3.34}$$

Using the substitution $\xi = \sqrt{\frac{m\omega}{\hbar}} x$ we can write the solution in terms of *Hermite polynomials*, where they're defined through *Rodrigues formula*

$$H_n(\xi) = (-1)^n e^{\xi^2} \frac{\mathrm{d}^n}{\mathrm{d}\xi^n} e^{-\xi^2} \tag{3.35}$$

The normalization constant for Hermite polynomials can be calculated to be exactly $(2^n n!)^{-1/2}$, and the eigenfunction succession of the quantum harmonic oscillator will be the following

$$\psi_n(x) = \sqrt{\frac{1}{2^n n!} \sqrt{\frac{m\omega}{\hbar\pi}}} H_n\left(\sqrt{\frac{m\omega}{\hbar}} x \right) e^{-\frac{m\omega}{\hbar}x^2} \tag{3.36}$$

§§ 3.2.2 Coherent States of the Quantum Harmonic Oscillator

The coherent states of a quantum harmonic oscillator, are those states defined as the eigenvalues of the annihilation operator $\hat{\eta}$. Hence, we are finding all states $|h\rangle$ such that $\hat{\eta}|h\rangle = h|h\rangle$. Since we're in the Hilbert space of the quantum harmonic oscillator (obviously) we can use Von Neumann's principle in order to Fourier transform our state $|h\rangle$ to an eigenstate of the Hamiltonian. We then get, applying a projection $\hat{\pi}_h^{(i)} = |h\rangle\langle h|$, the following Fourier series

$$|h\rangle = \sum_{n=0}^{\infty} \langle h|n\rangle |h\rangle \tag{3.37}$$

Applying $\hat{\eta}$ to its eigenstate $|h\rangle$, we get, knowing its action to the energy eigenstates

$$\hat{\eta}|h\rangle = \sum_{n=0}^{\infty} \langle h|n\rangle \hat{\eta}|n\rangle = \sum_{n=0}^{\infty} \langle h|n\rangle \sqrt{n} |n-1\rangle = h|h\rangle \tag{3.38a}$$

Changing the indexes of our sum to $k = n - 1$ we get

$$\sum_{k=0}^{\infty} \langle h|k+1\rangle \sqrt{k+1} |k\rangle = h \sum_{k=0}^{\infty} \langle h|k\rangle |k\rangle \tag{3.38b}$$

We henceforth get the following relation between $\langle h|k+1\rangle$ and $\langle h|k\rangle$

$$\langle h|k+1\rangle = \frac{h}{\sqrt{k+1}} \langle h|k\rangle \tag{3.38c}$$

Through induction we get, after substituting again the index,

$$\langle h|n\rangle = \frac{h^n}{\sqrt{n!}}\langle h|0\rangle$$

$$|h\rangle = \langle h|0\rangle \sum_{n=0}^{\infty} \frac{h^n}{\sqrt{n!}}|n\rangle \tag{3.38d}$$

In order to find $\langle h|0\rangle$ we normalize the state $\langle h|h\rangle$

$$\langle h|h\rangle = |\langle h|0\rangle|^2 \sum_{n=0}^{\infty} \frac{|h|^{2n}}{n!} = 1 \tag{3.38e}$$

We finally get $\langle h|0\rangle = e^{-\frac{|h|^2}{2}}$, hence, our coherent state will be

$$|h\rangle = e^{-\frac{|h|^2}{2}} \sum_{n=0}^{\infty} \frac{h^n}{\sqrt{n!}}|n\rangle \tag{3.38f}$$

Writing $|n\rangle$ as in (3.22) we get a new form of this state, in terms of $\hat{\eta}^\dagger$

$$|h\rangle = e^{\frac{|h|^2}{2}} \sum_{n=0}^{\infty} \frac{h^n}{n!}\hat{\eta}^{\dagger n}|0\rangle \tag{3.39}$$

Summing, we then find a new representation for this coherent state, in terms of both annihilation and creation operators

$$|h\rangle = e^{-\frac{\bar{h}h}{2}+h\hat{\eta}}|0\rangle = e^{h\hat{\eta}^\dagger - \bar{h}\hat{\eta}}|0\rangle \tag{3.40}$$

§§ 3.2.3 Schrödinger Formulation

The Schrödinger equation for the quantum harmonic oscillator is simply given converting momentum and position operator to their representation in $L^2(\mathbb{R})$. Taking (3.11) and converting, we get the following differential equation for $\psi_n(x)$ as follows

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_n}{\mathrm{d}x^2} + \frac{1}{2}m\omega^2 x^2 \psi_n(x) = E_n \psi_n(x) \tag{3.41}$$

In order to ease the equation, we utilize the following change of variables

$$\xi = \sqrt{\frac{m\omega}{\hbar}}x$$

Through this change of variables we get that the second derivative of ψ_n will become

$$\frac{\mathrm{d}^2\psi_n}{\mathrm{d}x^2} = \frac{m\omega}{\hbar}\frac{\mathrm{d}^2\psi_n}{\mathrm{d}\xi^2}$$

Hence, the Schrödinger equation will become

$$-\frac{\hbar\omega}{2}\frac{d^2\psi_n}{d\xi^2} + \left(\frac{\hbar\omega}{2}\xi^2 - E_n\right)\psi_n(\xi)$$

Rearranging everything, we get the following "easier" to tackle equation

$$\frac{d^2\psi_n}{d\xi^2} - \left(\xi^2 - \frac{2E_n}{\hbar\omega}\right)\psi_n(\xi) = 0 \tag{3.42}$$

Considering the limit where $\xi >> \frac{2E_n}{\hbar\omega}$, we can define an asymptotic differential equation for ψ_n

$$\psi_a(\xi) - \xi^2\psi_a(\xi) = 0$$

Its solution will be a linear combination of esponentials

$$\psi_a(\xi) = Ae^{\frac{\xi^2}{2}} + Be^{-\frac{\xi^2}{2}}$$

Due to normalization problems, we choose $A = 0$.
Through the definition of ξ, we can easily find the normalization constant B, which it is, simply

$$B = \sqrt{\sqrt{\frac{m\omega}{\hbar\pi}}}$$

Henceforth, the asymptotic solution will simply be the following

$$\psi_a(x) = \sqrt{\sqrt{\frac{m\omega}{\hbar\pi}}}e^{-\frac{m\omega}{\hbar}x^2} \tag{3.43}$$

The complete solution, will then be the product of a function $h(\xi)$ with the asymptotic solution

$$\psi_n(\xi) = h(\xi)\psi_a(\xi) \tag{3.44}$$

Where, $h(x)$ is a power series, defined as follows

$$h(\xi) = \sum_{j=0}^{\infty} a_j\xi^j \tag{3.45}$$

Deriving the new definition of $\psi_n(\xi)$, we get the following relation

$$\frac{d^2\psi_n}{d\xi^2} = \psi_a(\xi)\frac{d^2h}{d\xi^2} + 2\frac{dh}{d\xi}\frac{d\psi_a}{d\xi} + h(\xi)\frac{d^2\psi_a}{d\xi^2}$$

Due to ψ_a being known, we end up with the following equation

$$\frac{d^2\psi_n}{d\xi^2} = \sqrt{\sqrt{\frac{m\omega}{\hbar\pi}}}\frac{d^2h}{d\xi^2}e^{-\frac{\xi^2}{2}} - 2\xi\sqrt{\sqrt{\frac{m\omega}{\hbar\pi}}}\frac{dh}{d\xi}e^{-\frac{\xi^2}{2}} + (\xi^2 - 1)\sqrt{\sqrt{\frac{m\omega}{\hbar\pi}}}h(\xi)e^{-\frac{\xi^2}{2}}$$

The Schrödinger equation will, finally, become the following, after cleaning up multiplicative constants and nonvanishing exponentials

$$\frac{\mathrm{d}^2 h}{\mathrm{d}\xi^2} - 2\xi \frac{\mathrm{d}h}{\mathrm{d}\xi} + \left(\frac{2E_n}{\hbar\omega} - 1\right) h(\xi) = 0 \tag{3.46}$$

The derivatives of $h(\xi)$ are easy to calculate, and they give the following result

$$\frac{\mathrm{d}h}{\mathrm{d}\xi} = \sum_{j=1}^{\infty} j a_j \xi^{j-1} = \sum_{n=0}^{\infty} \alpha a_\alpha \xi^{\alpha-1}$$

$$\frac{\mathrm{d}^2 h}{\mathrm{d}\xi^2} = \sum_{j=2}^{\infty} j(j-1) a_j \xi^{j-2} = \sum_{\alpha=0}^{\infty} (\alpha+1)(\alpha+2) a_{\alpha+2} \xi^{\alpha}$$

Inserting in our Schrödinger equation, we get the following

$$\sum_{\alpha=0}^{\infty} \left[(\alpha+1)(\alpha+2) a_{\alpha+2} + \left(\frac{2E_n}{\hbar\omega} - 2\alpha - 1\right) a_\alpha \right] \xi^{\alpha} = 0$$

For which, the only non trivial solutions will be given if the summand is zero, thing that'll be true only and only if persists a recursive relation

$$a_{\alpha+2} = \frac{2\alpha + 1 - \frac{2E_n}{\hbar\omega}}{(\alpha+1)(\alpha+2)} a_\alpha \tag{3.47}$$

Since we want the sum to converge, we suppose that $\exists n \in \mathbb{N} : \forall \alpha > n \; a_\alpha = 0$, hence, using the recursive relation we get

$$\frac{2E_n}{\hbar\omega} = 2n + 1 \tag{3.48}$$

Solving for E_n we get the quantization of energy

$$E_n = \hbar\omega \left(n + \frac{1}{2}\right) \tag{3.49}$$

Getting a closer look on (3.46), we get that a general solution for this differential equation is known, and it's the Hermite polynomials $H_n(\xi)$. The final product solution will then be our eigenfunctions of the Hamiltonian

$$\psi_n(x) = \sqrt{\frac{1}{2^n n!}} \sqrt{\frac{m\omega}{\hbar\pi}} H_n\left(\sqrt{\frac{m\omega}{\hbar}} x\right) e^{-\frac{m\omega}{2\hbar} x^2} \tag{3.50}$$

The factor $(2^n n!)^{-1/2}$ is given by the normalization of Hermite polynomials.

§ 3.3 Infinite Square Well

Let's consider now a massive particle inside an infinite square well, i.e., where the potential is defined as follows

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

Since this problem is simple enough with the Schrödinger approach, we immediately write the Schrödinger equation for the problem

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = -k^2\psi(x) \tag{3.51}$$

Where we define k as $k^2 = 2mE/\hbar^2$.
The general solution is easily computed as being

$$\psi(x) = A\sin(kx) + B\cos(kx) \tag{3.52}$$

Since ψ must be square integrable in all space, we impose its continuity at the borders of the well, hence we must have $\psi(0) = \psi(a) = 0$, which gives us $B = 0$ and $A\sin(ka) = 0$. Since $A = 0$ would give a trivial solution, we impose $\sin(ka) = 0$, and we get a restriction on the possible values of k.

$$ka = n\pi \longrightarrow k_n = \frac{n\pi}{a} \tag{3.53}$$

Due to the definition of k, we get that energy must be quantized, with the following succession

$$E_n = \frac{n^2\pi^2\hbar^2}{2ma^2} \tag{3.54}$$

Integrating over all space the square modulus of what we have defined, finally, lets us determine the normalization factor A

$$|A|^2 \int_0^a \sin^2(k_n x)\,\mathrm{d}x = |A|^2\frac{a}{2} = 1 \tag{3.55}$$

The complete solution will be, finally

$$\psi_n(x) = \sqrt{\frac{2}{a}}\sin\left(\frac{n\pi}{a}x\right)$$
$$E_n = \frac{n^2\pi^2\hbar^2}{2ma^2} \tag{3.56}$$

§ 3.4 Infinite Wall

In order to treat the idea of an infinite wall in quantum mechanics, we have to first define two particular states, *scattering states* and *bound states*

Definition 3.4.1 (Bound State). We define a bound state, as the set of configurations of the system where $E < V(x)$, and the particle is henceforth "trapped"

Definition 3.4.2 (Scattering State). A scattering state is defined as the set of configurations of the system for which $E > V(x)$, hence, the particle coming from $-\infty$ will simply interact with the potential without getting trapped by it.

Getting back to our problem, we define an infinite wall as a system whose potential is described by a Dirac delta function $\delta(x)$. Using a potential $V(x) = -\alpha\delta(x)$, we get that the Hamiltonian will be

$$\hat{\mathcal{H}} = \frac{\hat{p}^2}{2m} - \alpha\delta(x) \tag{3.57}$$

The associated Schrödinger equation will be

$$\hat{\mathcal{H}}\psi(x) = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} - \alpha\delta(x)\psi(x) = E\psi(x) \tag{3.58}$$

It's obvious that we can both have bound states and scattering states. Considering first the bound states, in the region $x < 0$, $V(x) = 0$, we get the following equation, where $\kappa = (-2mE)^{1/2}/\hbar$ ($E < 0$ by assumption, since we're considering bound states only).

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = \kappa^2\psi(x)$$

Its solution will simply be $\psi(x) = A\exp(-\kappa x) + B\exp(\kappa x)$, that for normalization reasons, in the region $x < 0$, becomes simply

$$\psi(x) = Be^{\kappa x} \tag{3.59}$$

In the region $x > 0$, instead, we get the following solution

$$\psi(x) = Ae^{-\kappa x} \tag{3.60}$$

Since $\psi(x)$ must be «always» continuous, we get that the solution for our bound states must have $A = B$, and our general solution becomes

$$\psi(x) = \begin{cases} Be^{\kappa x} & x \leq 0 \\ Be^{-\kappa x} & x \geq 0 \end{cases} \tag{3.61}$$

Since $\psi(x)$ must also be a square integrable function, we need that its derivative must be continuous too, hence for $x = 0$, we have to check another few things.
The first idea that comes up to mind is to utilize the properties of the delta function and integrate in a infinitesimal interval around 0, henceforth, we get

$$-\frac{\hbar^2}{2m}\int_{-\epsilon}^{\epsilon}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}\,\mathrm{d}x - \alpha\int_{-\epsilon}^{\epsilon}\delta(x)\psi(x)\,\mathrm{d}x = E\int_{-\epsilon}^{\epsilon}\psi(x)\,\mathrm{d}x \tag{3.62}$$

Which becomes the following general relation

$$\lim_{x\to 0^{\pm}}\Delta\psi'(x) = -\frac{2m\alpha}{\hbar^2}\psi(0) \tag{3.63}$$

Calculating the derivatives from the left and from the right of our solution, and imposing what has been found previously, we get

$$\frac{\mathrm{d}\psi}{\mathrm{d}x} = \begin{cases} -B\kappa & x \to 0^+ \\ B\kappa & x \to 0^- \end{cases} \tag{3.64}$$

From this, we get $\Delta\psi'(x) = -2B\kappa$ and, then, since $\kappa = (-2mE)^{1/2}/\hbar$, we get from (3.63)

$$E = -\frac{m\alpha^2}{2\hbar^2} \tag{3.65}$$

Normalizing ψ, we get

$$2|B|^2 \int_{\mathbb{R}} e^{-2\kappa x} \, \mathrm{d}x = \frac{|B|^2}{\kappa} = 1 \tag{3.66}$$

The final solution for the wavefunction of the bound states, will then be

$$\psi(x) = \frac{\sqrt{m\alpha}}{\hbar} e^{-\frac{m\alpha}{\hbar^2}|x|}$$
$$E = -\frac{m\alpha^2}{2\hbar^2} \tag{3.67}$$

It's obvious that there is only «one» bound state.
For scattering states, we define $k = (2mE)^{1/2}/\hbar$, and the Schrödinger equation becomes the following

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = -k^2\psi(x) \tag{3.68}$$

The solution will be formed by the superposition of two complex esponentials, since neither of the two blows up for $x \to \pm\infty$. Hence, we will get

$$\psi(x) = \begin{cases} Ae^{ikx} + Be^{-ikx} & x < 0 \\ Fe^{ikx} + Ge^{-ikx} & x > 0 \end{cases} \tag{3.69}$$

Considering again their derivatives coming from left and right, we get the following

$$\frac{\mathrm{d}\psi}{\mathrm{d}x} = \begin{cases} ik(A - B) & x \to 0^- \\ ik(F - G) & x \to 0^+ \end{cases} \tag{3.70}$$

Since $\Delta\psi'(x) = ik(F - G - A + B)$ and $\psi(0) = A + B$, we get from (3.63) the following

$$ik\,(F - G - A + B) = -\frac{2m\alpha}{\hbar^2}\,(A + B)$$

Defining $\beta = m\alpha/\hbar^2 k$, we can write the previous relation in a much more compact way

$$F - G = A(1 + 2i\beta) - B(1 - 2i\beta) \tag{3.71}$$

In this case normalizating $\psi(x)$ won't help, since this state is absolutely non normalizable. In order to give a viable solution we reason on how the particle would scatter in this potential. Imagine shooting this quantum particle from $-\infty$, since a negative complex exponential describes a wave coming from $+\infty$ to $-\infty$ we can easily set $G = 0$, and get the following result:

$$\begin{cases} B = \dfrac{i\beta}{1 - i\beta} A \\ F = \dfrac{1}{1 - i\beta} A \end{cases}$$

(3.72)

Reasoning in a physical way, we can deduce then that A is the amplitude of the incident wave, B of the reflected wave and F of the trasmitted wave. Since the probability in quantum mechanics is given by the square modulus of the wavefunction, we get that the probability of having a particle reflected back or trasmitted forward, will be given by two coefficients, respectively R and T, where they're defined as follows

$$R = \frac{|B|^2}{|A|^2} = \frac{\beta^2}{1 + \beta^2}$$

$$T = \frac{|F|^2}{|A|^2} = \frac{1}{1 + \beta^2}$$

(3.73a)

Since their sum must be 1, being probabilities themselves, we get

$$R + T = \frac{|B|^2 + |F|^2}{|A|^2} = 1$$

(3.73b)

And substituting β with its full expression,

$$R = \frac{1}{1 + \frac{2\hbar^2 E}{m\alpha^2}}$$

$$T = \frac{1}{1 + \frac{m\alpha^2}{2\hbar^2 E}}$$

(3.73c)

A fun thing to do, with this potential, is to change the sign of α and reason on what is happening really.

First of all, since $E \propto \alpha$, we get that the only bound state we found gets brutally killed since $E \not< 0$ everywhere, but since $R, T \propto \alpha^2$, they stay unchanged.

Naively evaluating this as a classical problem, we get that this "particle" is thrown towards an infinitely strong wall n times and passes through nT times, and bounces back nR times. This is obviously impossible in the classical world, but instead in the quantum world is much more than possible. This effect is known commonly as *quantum tunneling*. This phenomenon is not restricted to infinite potential walls, let's take a finite potential wall V, for which exists a maximum V_{max}. If the energy E of the particle is $E < V_{max}$ we get that it might pass through with a nonzero probability T, and for $E > V_{max}$ there is still a probability of it bouncing back, expressed with R.

§ 3.5 Finite Square Well

After having considered an infinite square well and the difference between bound and scattering states, we can describe properly the finite square well problem, where the potential is defined as such

$$V(x) = \begin{cases} -V_0 & -a \leq x \leq a \\ 0 & |x| > a \end{cases}$$

It's obvious how this potential admits both scattering and bound states.
Let's consider first the region $x < -a$, where the potential is 0, from the previous problem on the infinite wall, we can write directly the solution

$$\psi(x) = Be^{\kappa x} \quad \kappa = \frac{\sqrt{-2mE}}{\hbar}, \ x < -a \tag{3.74}$$

Inside the well, the problem is similar, but it's useful to directly write Schrödinger's equation

$$\hat{\mathcal{H}}\psi(x) = \frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} - \frac{2m}{\hbar^2}(V_0 + E)\psi(x) = 0 \tag{3.75}$$

Replacing $[2m(E + V_0)^{-1/2}]/\hbar$ with l it reduces, yet again, to the following equation

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} = -l^2\psi(x)$$

Since $E > V_{min}$, the solution must be real and positive, hence, we can write it as follows

$$\psi(x) = C\sin(lx) + D\cos(lx) \quad -a < x < a \tag{3.76}$$

In the outer region, for $x > a$, we have again an exponential solution, this time decreasing

$$\psi(x) = Fe^{\kappa x} \quad x > a \tag{3.77}$$

Now, we need to impose the boundary conditions for ψ and ψ', that due to the potential being odd, can only be of two kinds, either odd or even.
Choosing the even solution, we get immediately that $C = 0$, and $\psi(x)$ becomes

$$\psi(x) = \begin{cases} Fe^{-\kappa x} & x > a \\ D\cos(lx) & 0 < x < a \\ \psi(-x) & x < 0 \end{cases} \tag{3.78}$$

For the continuity of ψ, ψ' at $x = a$, we get the following system

$$\begin{cases} Fe^{-\kappa a} = D\cos(la) \\ -\kappa Fe^{-\kappa a} = -lD\sin(la) \end{cases} \tag{3.79}$$

Dividing the two, we get $\kappa = l \tan(la)$. This is clearly a restriction on energies, due to the definition of κ, but it's a trascendental equation, that can't be solved directly. Applying the transformation $z = la$ and $z_0 = a(2mV_0)^{1/2}/\hbar$ we end with the following trascendental equation

$$\tan(z) = \sqrt{\left(\frac{z_0}{z}\right)^2 - 1} \tag{3.80}$$

We can also approximate energy eigenvalues for the case of a deep well or a shallow well. In the first case we get that the solutions to the aforementioned equation will be at $z_n = n\pi/2$ with n odd, and follows that

$$E_n + V_0 \approx \frac{n^2\pi^2\hbar^2}{8ma^2}$$

In the second case instead, the intersections between the tangent and the square root get fewer and fewer, until, for $z_0 < \pi/2$ we end up with a single eigenstate, no matter how shallow is the well.

The only thing we miss to properly evaluate this problem (other than compute the normalization constant for $\psi(x)$) is to check the scattering states of the system.

We assume an ian incoming wave from the right, with a wavefunction $\psi(x) = Fe^{ikx}$, with $k^2 = 2mE/\hbar^2$. Putting, as before, A as the incident amplitude, B as the reflected amplitude and F the transmitted amplitude, for continuity of ψ and its derivative, we get that the following systems must hold. At $x = -a$ we have

$$\begin{cases} Ae^{-ika} + Be^{ika} = D\cos(la) - C\sin(la) \\ ik\left[Ae^{-ika} - Be^{ika}\right] = l\left[C\cos(la) + D\sin(la)\right] \end{cases} \tag{3.81}$$

At $x = a$, instead we get the following system

$$\begin{cases} C\sin(la) + D\cos(la) = Fe^{ika} \\ l\left[C\cos(la) - D\sin(la)\right] = ikFe^{ika} \end{cases} \tag{3.82}$$

Eliminating C and D and solving for B and F, we get the following

$$B = i\frac{\sin(2la)}{2kl}\left(l^2 - k^2\right)F$$

$$F = \frac{e^{-2ika}}{\cos(2la) - i\frac{k^2+l^2}{2kl}\sin(2la)}A \tag{3.83}$$

Substituting everything back to the original variables, we get that the transmission coefficient T is

$$\frac{1}{T} = 1 + \frac{V_0^2}{4E(E + V_0)}\sin^2\left(\frac{2a}{\hbar}\sqrt{2m(E + V_0)}\right) \tag{3.84}$$

The reflection coefficient can be calculated knowing that $R = 1 - T$.

§§ 3.5.1 Scattering and Transfer Matrices

In order to treat efficiently scattering problems for general potentials, we first of all, suppose that the wavefunction will be the following:

For an incoming particle, since it's a free particle, we get that $\psi(x)$ will be a superposition of complex exponentials

$$\psi(x) = Ae^{ikx} + Be^{-ikx} \tag{3.85}$$

The same goes for the post-scattering region, but with changed coefficients due to the interaction with the potential

$$\psi(x) = Fe^{ikx} + Ge^{-ikx} \tag{3.86}$$

In the interaction region, we can still suppose that the wavefunction will be a superposition of two functions $f(x), g(x)$, albeit both will remain unknown, until a specific potential is given.

$$\psi(x) = Cf(x) + Dg(x) \tag{3.87}$$

Using this systems, we can write a linear system of equations for B and F (the reflected and trasmitted amplitudes), as follows

$$\begin{aligned} B &= S_{11}A + S_{12}G \\ F &= S_{21}A + S_{22}G \end{aligned} \tag{3.88}$$

Reuniting everything in a *scattering matrix*, also known as *S-matrix*, we get the following matrix equation

$$\begin{pmatrix} B \\ F \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} A \\ G \end{pmatrix} \tag{3.89}$$

In this formalism, the transmission and reflection coefficients will be:

For a particle coming from the left

$$\begin{aligned} R_l &= \frac{|B|^2}{|A|^2} = |S_{11}|^2 \\ T_l &= \frac{|F|^2}{|A|^2} = |S_{21}|^2 \end{aligned} \tag{3.90}$$

For a scattering from the right instead, we get

$$\begin{aligned} R_r &= \frac{|F|^2}{|G|^2} = |S_{22}|^2 \\ T_r &= \frac{|B|^2}{|G|^2} = |S_{12}|^2 \end{aligned} \tag{3.91}$$

The scattering matrix, hence, gives us the outgoing amplitudes in terms of the incoming amplitudes.

In case we'd like to have the amplitudes on the right of the potential in terms of the amplitudes on the left of the potential, we can "build" a second matrix, called *transfer matrix* or *M-matrix*

$$\begin{pmatrix} F \\ G \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} \tag{3.92}$$

The usefulness of this matrix, is that if the potential is formed by two separated pieces, the complete M-matrix of the system, will be given as the product of the two M-matrices of the single pieces of the potential, hence $\underline{M} = \underline{M_2 M_1}$. It's easy then to generalize this product to multiple pieces of the potential.

§ 3.6 One Dimensional Motion in Generic Potentials

What has been studied with the one dimensional quantum system that have been solved before, can be generalized to a one-dimensional Hamiltonian with a general potential $V(x)$. Without passing through the operatorial representation, we write directly the Schrödinger equation for a general system.
We will have

$$\hat{\mathcal{H}}\psi(x) = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + V(x)\psi(x) = E\psi(x)$$

Rewriting the equation in its normal form, we will have the following

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} - \frac{2m}{\hbar^2}\left(V(x) - E\right)\psi(x) = 0 \tag{3.93}$$

We can immediately bring up three conclusions

1. If $\psi(x)$ is normalizable, there exists a *discrete spectrum* of eigenvalues of $\hat{\mathcal{H}}$, E_n

2. If $\psi(x)$ is «not» normalizable, there exists a *continuous spectrum* of $\hat{\mathcal{H}}$, $\sigma(E)$

3. Since $\psi : \mathbb{R} \to \mathbb{C}$, from the theory of differential equations, we will have that $\mathfrak{Re}\psi(x)$ will be a solution.

Again, if we consider what we derived for scattering and bound states, we can derive from (3.93) two considerations on $\psi(x)$ and its second derivative, that we will indicate with $\psi''(x)$ in order to avoid heavy notations.

1. If we are evaluating a bound state, then $\psi''(x)/\psi < 0$

2. If we are evaluating a scattering state, then $\psi''(x)/\psi > 0$

3. In passing from a bound state to a scattering state there is an inversion point, where $\psi''(x)/\psi(x) = 0$, $\psi(x) \neq 0$ and $E = V(x)$

It can be demonstrated that, if $E < \min(V)$ there won't be any eigenvalues E, for $\min(V) < E < 0$ there will be at least one eigenvalue E, if there is more than one in this region, they will be discrete and nondegenerate. For $0 < E < \max(V)$ we will have nondegenerate continue eigenvalues $\sigma(E)$, and for $E > \max(V)$ they will be continuous and twice degenerate. Utilizing this general potential Schrödinger equation, we can also define a quite useful theorem

Theorem 3.2 (Theorem of the Oscillations). *Let $\hat{D}_s$ be a differential operator that acts in the following way:*

$$\hat{D}_s = \frac{\mathrm{d}^2}{\mathrm{d}x^2} - \frac{2m}{\hbar^2}\left(V(x) - E\right)$$

Suppose then that there exist two functions $\psi(x)$ and $\phi(x)$, for which $\hat{D}_s\psi(x) = 0$ and $\hat{D}_s\phi(x) = 0$, then ψ and ϕ are linearly dependent with n zeroes.

Proof. We begin writing directly the action of the operator $\hat{D}_s$ on the two functions $\psi(x)$ and $\phi(x)$

$$\begin{cases} \hat{D}_s\psi(x) = \dfrac{\mathrm{d}^2\psi}{\mathrm{d}x^2} - \dfrac{2m}{\hbar^2}\left(V(x) - E\right)\psi(x) = 0 \\[2mm] \hat{D}_s\phi(x) = \dfrac{\mathrm{d}^2\phi}{\mathrm{d}x^2} - \dfrac{2m}{\hbar^2}\left(V(x) - E\right)\phi(x) = 0 \end{cases}$$

We solve this linear system multiplying the first equation by $\phi(x)$ and the second by $\psi(x)$, and then subtracting the two rows, obtaining the following relation

$$\phi(x)\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} - \psi(x)\frac{\mathrm{d}^2\phi}{\mathrm{d}x^2} = 0$$

It's easy to see that this it's the derivative of $\phi(x)\psi'(x) - \psi(x)\phi'(x)$, and since this derivative must be zero, we know that it must be a constant k

$$\phi\frac{\mathrm{d}\psi}{\mathrm{d}x} - \psi(x)\frac{\mathrm{d}\phi}{\mathrm{d}x} = k$$

Since both functions must be normalizable, we must have $k = 0$, hence we can write the equation in the following way

$$\frac{\psi'(x)}{\psi(x)} = \frac{\phi'(x)}{\phi(x)}$$

Integrating once, we get that $\psi(x) = \lambda\phi(x)$, hence they're linearly independent, and hence essentially the same. $\square$

Corollary 3.6.1. Since $\phi(x), \psi(x)$ are eigenfunctions of the Hamiltonian, we know that there is a direct relation between the energy eigenvalue E_n and the eigenfunction $\psi_n(x)$, hence, analyzing everything, we get that it must have n zeroes.

Another interesting feature of quantum dynamics, is given by the quantum version of the virial theorem, which states the following

Theorem 3.3. *Let $\hat{T}$ be the kinetic energy operator, and $\hat{V}$ the potential energy operator, then the expectation value of the kinetic energy on an energy eigenstate is equal to the expectation value of $qV'(q)$.*

Proof. In order to demonstrate this theorem we first calculate the commutator between the Hamiltonian operator and the product operator $\hat{q}\hat{p}$.
The Hamiltonian in question is

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

Then the desired commutator will be the following

$$\left[\hat{\mathcal{H}}, \hat{q}\hat{p}\right] = \hat{q}\left[\hat{\mathcal{H}}, \hat{p}\right] + \left[\hat{\mathcal{H}}, \hat{q}\right]\hat{p}$$

In order to evaluate the previous commutators, we write the Poisson brackets of each, and then quantize, deforming the brackets through mltiplication by $i\hbar$

$$[\mathcal{H}, p]_{PB} = \frac{\partial \mathcal{H}}{\partial q}\frac{\partial p}{\partial p} - \frac{\partial \mathcal{H}}{\partial p}\frac{\partial p}{\partial q} = \frac{\partial V}{\partial q}$$

$$[\mathcal{H}, q]_{PB} = \frac{\partial \mathcal{H}}{\partial q}\frac{\partial q}{\partial p} - \frac{\partial \mathcal{H}}{\partial p}\frac{\partial q}{\partial q} = -\frac{p}{m}$$

Quantizing, we get the following

$$\left[\hat{\mathcal{H}}, \hat{p}\right] = i\hbar\frac{\partial V}{\partial q}$$

$$\left[\hat{\mathcal{H}}, \hat{q}\right] = -i\hbar\frac{\hat{p}}{m}$$

And the searched commutator will be

$$\hat{q}\left[\hat{\mathcal{H}}, \hat{p}\right] + \left[\hat{\mathcal{H}}, \hat{q}\right]\hat{p} = i\hbar\hat{q}\frac{\partial V}{\partial q} - i\hbar\frac{\hat{p}^2}{m}$$

Rearranging, and remembering how $\hat{T}$ is defined, we get the following

$$\left[\hat{\mathcal{H}}, \hat{q}\hat{p}\right] = i\hbar\left(\hat{q}\frac{\partial V}{\partial q} - 2\hat{T}\right)$$

We calculate the expectation value of this commutator in the energy eigenstate

$$\left\langle\left[\hat{\mathcal{H}}, \hat{q}\hat{p}\right]\right\rangle = i\hbar\left\langle\hat{q}\frac{\partial V}{\partial q} - 2\hat{T}\right\rangle$$

We first calculate the left side, and we get the following

$$\left\langle\hat{q}\hat{\mathcal{H}}\hat{p} - \hat{q}\hat{p}\hat{\mathcal{H}} + \hat{\mathcal{H}}\hat{q}\hat{p} - \hat{q}\hat{\mathcal{H}}\hat{p}\right\rangle = \left\langle\hat{\mathcal{H}}\hat{q}\hat{p} - \hat{q}\hat{p}\hat{\mathcal{H}}\right\rangle = E\left\langle\hat{q}\hat{p} - \hat{q}\hat{p}\right\rangle = 0$$

Hence, using the linearity of the $\langle \cdot \rangle$ operator, we get the virial theorem

$$i\hbar \left\langle \hat{q}\frac{\partial V}{\partial q} \right\rangle - 2i\hbar \left\langle \hat{T} \right\rangle = 0$$

$$2\left\langle \hat{T} \right\rangle = \left\langle \hat{q}\frac{\partial V}{\partial q} \right\rangle$$

$\square$

Corollary 3.6.2. Considering the particular case of potentials of the form $V(q) = c\hat{q}^\alpha$, we get the following

$$\left\langle \hat{T} \right\rangle = \frac{\alpha}{2} \left\langle \hat{V} \right\rangle$$

Proof. We use the previous identity used in the general case, and we get that $\hat{q}V'(q) = c\alpha\hat{q}^\alpha = \alpha\hat{V}$, hence

$$2\left\langle \hat{T} \right\rangle = \alpha \left\langle \hat{V} \right\rangle$$

$\square$

The next theorem that will be stated, will be directly generalized to n-dimensions.

Theorem 3.4 (Probability Conservation). *Defining $J_i(x_j)$ as the probability current density vector, for a time-independent system the following continuity equation holds*

$$\frac{\partial J_i}{\partial x_i} = 0$$

If $J_i(x_j)$ is defined as follows

$$J_i(x_j) = -\frac{i\hbar}{2m} \left(\overline{\psi}(x_j)\frac{\partial \psi}{\partial x_i} - \psi(x_j)\frac{\partial \overline{\psi}}{\partial x_i} \right)$$

Or, in operatorial form

$$\hat{J}_i = \frac{1}{2m} \left(\langle \psi | \hat{p}_i | \psi \rangle \, \hat{\mathbb{1}} - | \psi \rangle \, \hat{p}_i \, \langle \psi | \right)$$

Proof. We will begin by writing the Schrödinger equation for both ψ and $\overline{\psi}$. We will then have the following system

$$\begin{cases} -\dfrac{\hbar^2}{2m}\dfrac{\partial^2 \psi}{\partial x_i \partial x_i} + V(x_i)\psi(x_j) = E\psi(x_j) \\ -\dfrac{\hbar^2}{2m}\dfrac{\partial^2 \overline{\psi}}{\partial x_i \partial x_i} + V(x_i)\overline{\psi}(x_j) = E\overline{\psi}(x_j) \end{cases}$$

Multiplying the first by $\overline{\psi}_i(x_j)$ and the second by $\psi_i(x_j)$ and adding the second to the first we get the following

$$-\frac{\hbar^2}{2m}\left(\overline{\psi}(x_j)\frac{\partial^2\psi}{\partial x_i \partial x_i} - \psi(x_j)\frac{\partial^2\overline{\psi}}{\partial x_i \partial x_i}\right) + V(x_i)\left(|\psi(x_j)|^2 - |\psi(x_j)|^2\right) = E\left(|\psi|^2 - |\psi|^2\right)$$

Simplifying everything, we get

$$-\frac{\hbar^2}{2m}\left(\overline{\psi}(x_j)\frac{\partial^2\psi}{\partial x_i \partial x_i} - \psi(x_j)\frac{\partial^2\psi}{\partial x_i \partial x_i}\right) = 0$$

Bringing outside the gradient operator we get

$$-\frac{\hbar^2}{2m}\frac{\partial}{\partial x_i}\left(\overline{\psi}(x_j)\frac{\partial\psi}{\partial x_i} - \psi(x_j)\frac{\partial\psi}{\partial x_i}\right) = \frac{\partial J_i}{\partial x_i} = 0$$

This theorem, in one dimension, reduces simply to the equation

$$\mathfrak{Im}(\overline{\psi}(x)\psi'(x)) = \mathfrak{Im}(\psi(x)\overline{\psi}'(x))$$

$\square$

§ 3.7 Time Evolution of Quantum Systems

§§ 3.7.1 General Remarks and Schrödinger's Picture

The problems we studied in the previous sections never found how the eigenstate of the system evolves in time.

Time evolution of a system can be seen as a traslation, hence taking the eigenstate of the time independent problem and "moving" it to a time $t \neq 0$. It's not hard to imagine then how time evolution can be seen as the action of an unitary group with a single parameter, which will be t, our time.

The first question that comes up to the mind is actually how this group is defined in quantum mechanics.

Here, the following theorem comes to our rescue

Theorem 3.5 (Stone Theorem). *Let U_t be a strongly continuous one-parameter unitary group. Then*

$$\exists! \hat{A} : D_a \subset \mathbb{H} \to \mathbb{H}, \ A = A^\dagger \text{ in } D_a \text{ then } \hat{\mathcal{U}}(t) = e^{it\hat{A}} \in U_t$$

Where holds the following relation $\hat{\mathcal{U}}(t_1 + t_2) = \hat{\mathcal{U}}(t_1)\hat{\mathcal{U}}(t_2)$ where $\hat{\mathcal{U}}(t_1 + t_2) \in U_{t_1}U_{t_2}$ and $\hat{\mathcal{U}}(t_1) \in U_{t_1}$, $\hat{\mathcal{U}}(t_2) \in U_{t_2}$ (i.e., $\hat{\mathcal{U}}$ is an homomorphism)
The set D_a is defined as follows

$$D_a := \left\{|\psi\rangle \in \mathbb{H} \ \Big| \ \exists \lim_{\epsilon \to 0} \frac{-i}{\epsilon}\left(\hat{\mathcal{U}}(\epsilon)|\psi\rangle - \hat{\mathbb{1}}|\psi\rangle\right)\right\}$$

$\hat{A}$ is called the infinitesimal generator of the unitary group, and can be computed as follows

$$\hat{A}\ket{\psi} = -i \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(\hat{\mathcal{U}}(\epsilon)\ket{\psi} - \hat{\mathbb{1}}\ket{\psi} \right)$$

As a spoiler, we can say that this infinitesimal generator is the Hamiltonian of the system, and the operator $\hat{\mathcal{U}}$ has the form $\hat{\mathcal{U}} = e^{-\frac{i}{\hbar}\hat{\mathcal{H}}t}$. In order to see how this works out, informally, we write the time-dependent Schrödinger equation, which can be derived considering energy as a differential operator $E \to i\hbar\partial_t$. Substituting, we get the following equation

$$\hat{E}\ket{\psi} = \hat{\mathcal{H}}\ket{\psi} \to i\hbar\frac{\partial\psi}{\partial t} = \hat{\mathcal{H}}\psi(x,t) \tag{3.94a}$$

In the case that $\hat{\mathcal{H}}$ is time-independent, we get that the solution to the differential equation in time will be the following

$$\psi(x,t) = c(x)e^{-\frac{i}{\hbar}\hat{\mathcal{H}}t} \tag{3.94b}$$

It's obvious then that the time-evolved state will be the time-independent state at which gets applied the time evolution operator.

$$\psi(x,t) = \hat{\mathcal{U}}(t)\psi(x) \tag{3.94c}$$

In case the Hamiltonian is time dependent, the operator easily becomes the following

$$\hat{\mathcal{U}}(t) = e^{-\frac{i}{\hbar}\int_0^t \hat{\mathcal{H}}\,\mathrm{d}t} \tag{3.94d}$$

In order to see how this operator actually acts on the states, it's quite useful to utilize the fact that all complex exponentials are holomorphic, and can be written as a power series. This holds even when we talk about operator exponentials, hence $\hat{\mathcal{U}}(t)$ becomes:

$$\hat{\mathcal{U}}(t) = \sum_{n=0}^{\infty} \frac{1}{n!}\left(-\frac{i}{\hbar}\hat{\mathcal{H}}t\right)^n \tag{3.95a}$$

Applying this to an eigenstate of the Hamiltonian $\ket{\psi}$, we get the following

$$\hat{\mathcal{U}}(t)\ket{\psi} = \sum_{n=0}^{\infty} \frac{1}{n!}\left(-\frac{i}{\hbar}\hat{\mathcal{H}}t\right)\ket{\psi}$$

$$\hat{\mathcal{U}}(t) = \sum_{n=0}^{\infty} \frac{1}{n!}\left(-\frac{i}{\hbar}\hat{\mathcal{H}}t\right)^{n-1}\left(-\frac{i}{\hbar}\hat{\mathcal{H}}t\right)\ket{\psi} \tag{3.95b}$$

Using Schrödinger's equation and iterating, we finally get, retransforming the series to an exponential, the following result for eigenstates

$$\hat{\mathcal{U}}(t)\ket{\psi} = e^{-\frac{i}{\hbar}Et}\ket{\psi} \tag{3.95c}$$

For a general state $|s\rangle$, we then get using Von Neumann's principle, the following result

$$\hat{\mathcal{U}}(t)\left|s\right\rangle = e^{-\frac{i}{\hbar}\hat{\mathcal{H}}t}\sum_n \langle\psi|s\rangle\left|\psi\right\rangle = \sum_n c_n e^{-\frac{i}{\hbar}Et}\left|\psi\right\rangle \qquad (3.95\text{d})$$

Since all the problems that we discussed were with a time-independent Hamiltonian, they're quite easy to generalize to a time-evolved problem without solving a time dependent Schrödinger equation or even redoing any calculus. This stress on time evolution on the state, is commonly called *Schrödinger picture* of time evolution.

Having now described time evolution in quantum mechanics, the next step is redescribing the probability conservation equation. Since now we have that the wavefunction is time-dependent, we have now that the time dependent probability amplitude will be defined as follows

$$\langle s(t)|s(t)\rangle = \langle s|\hat{\mathcal{U}}^\dagger\hat{\mathcal{U}}\left|s\right\rangle \rightarrow \int_{-\infty}^\infty \overline{\psi}_s(x,t)\psi_s(x,t)\,\mathrm{d}x = \rho(t)$$

Having defined this time dependent probability amplitude with $\rho(t)$. We now have an advanced version of the theorem

Theorem 3.6 (Time Dependent Probability Conservation)**.** *The probability amplitude of a wavefunction that solves the Schrödinger's time-dependent equation must always solve the following equation*

$$\frac{\partial \rho}{\partial t} + \frac{\partial J_i}{\partial x_i} = 0$$

Where

$$\rho(x_j,t) = \overline{\psi}(x_j,t)\psi(x_j,t)$$

$$J_i(x_j,t) = -\frac{i\hbar}{2m}\left(\overline{\psi}(x_j,t)\frac{\partial\psi}{\partial x_i} - \psi(x_j,t)\frac{\partial\overline{\psi}}{\partial x_i}\right)$$

Proof. The first thing to write is the time dependent Schrödinger equation for ψ and its complex conjugate, since both must solve it as a proposition of the theorem.

$$\begin{cases} i\hbar\dfrac{\partial\psi}{\partial t} = -\dfrac{\hbar^2}{2m}\dfrac{\partial^2\psi}{\partial x_i\partial x_i} + V(x_j)\psi(x_j,t) \\[2mm] -i\hbar\dfrac{\partial\overline{\psi}}{\partial t} = -\dfrac{\hbar^2}{2m}\dfrac{\partial^2\overline{\psi}}{\partial x_i\partial x_i} + V(x_j)\psi(x_j,t) \end{cases}$$

We then multiply the first line by $\overline{\psi}$ and the second by ψ

$$\begin{cases} i\hbar\overline{\psi}(x_j,t)\dfrac{\partial\psi}{\partial t} = -\dfrac{\hbar^2}{2m}\overline{\psi}(x_j,t)\dfrac{\partial^2\psi}{\partial x_i\partial x_i} + \overline{\psi}(x_j,t)V(x_j)\psi(x_j,t) \\[2mm] -i\hbar\psi(x_j,t)\dfrac{\partial\overline{\psi}}{\partial t} = -\dfrac{\hbar^2}{2m}\psi(x_j,t)\dfrac{\partial^2\overline{\psi}}{\partial x_i\partial x_i} + \psi(x_j,t)V(x_j)\overline{\psi}(x_j,t) \end{cases}$$

We then subtract the first line to the second, and we obtain the following equation

$$-i\hbar\left(\overline{\psi}(x_j,t)\frac{\partial\psi}{\partial t} + \psi(x_j,t)\frac{\partial\overline{\psi}}{\partial t}\right) = \frac{\hbar^2}{2m}\left(\overline{\psi}(x_j,t)\frac{\partial^2\psi}{\partial x_i\partial x_i} - \psi(x_j,t)\frac{\partial^2\overline{\psi}}{\partial x_i\partial x_i}\right)$$

We recognize immediately the left part being exactly equal to $\partial_t \rho(x_j, t)$, hence we substitute it in the equation

$$-i\hbar \frac{\partial \rho}{\partial t} = \frac{\hbar^2}{2m} \left(\overline{\psi}(x_j, t) \frac{\partial^2 \psi}{\partial x_i \partial x_i} - \psi(x_j, t) \frac{\partial^2 \overline{\psi}}{\partial x_i \partial x_i} \right)$$

Bringing out a gradient operator (∂_i), we recognize the part on the right being the probability density current J_i, to which has been applied a gradient operator ∂_i (Einstein summation convention is implied, hence it's a divergence)

$$-i\hbar \frac{\partial \rho}{\partial t} = i\hbar \frac{\partial J_i}{\partial x_i}$$

We simplify everything and bring the right hand side on the left of the equation and we finally have demonstrated the theorem

$$\frac{\partial \rho}{\partial t} + \frac{\partial J_i}{\partial x_i} = 0$$

$\square$

§§ 3.7.2 Heisenberg Picture and Constants of Motion

In the Heisenberg picture, the stress on time dependence is given only on operators. It may not be immediately clear how this is equivalent to Schrödinger's picture, where the time dependence is always on the state.

In order to clearly see this equivalence, we calculate the expectation state of an operator $\hat{A}$ on a time evolved state $|s(t)\rangle$. We get the following relation

$$\left\langle \hat{A} \right\rangle_t = \langle s(t)| \hat{A} |s(t)\rangle = \langle s| \hat{\mathcal{U}}^\dagger(t) \hat{A} \hat{\mathcal{U}}(t) |s\rangle \tag{3.96}$$

It's immediate to write the previous relation as $\left\langle \hat{A}(t) \right\rangle_s$, where the time evolved operator is defined as $\hat{A}(t) = \hat{\mathcal{U}}^\dagger(t) \hat{A} \hat{\mathcal{U}}(t)$. Using this definition, we can also calculate the derivative of an operator, in the following way

$$\begin{aligned}
\frac{\mathrm{d}\hat{A}}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t} \left(\hat{\mathcal{U}}^\dagger(t) \hat{A} \hat{\mathcal{U}}(t) \right) \\
&= \frac{\mathrm{d}\hat{\mathcal{U}}^\dagger}{\mathrm{d}t} \hat{A} \hat{\mathcal{U}}(t) + \hat{\mathcal{U}}^\dagger \hat{A} \frac{\mathrm{d}\hat{\mathcal{U}}}{\mathrm{d}t} = \\
&= \frac{i}{\hbar} \left(\hat{\mathcal{H}} \hat{\mathcal{U}}^\dagger \hat{A} \hat{\mathcal{U}} - \hat{\mathcal{U}}^\dagger \hat{A} \hat{\mathcal{H}} \hat{\mathcal{U}} \right) = \\
&= \frac{i}{\hbar} \hat{\mathcal{U}}^\dagger \left[\hat{\mathcal{H}}, \hat{A} \right] \hat{\mathcal{U}}
\end{aligned} \tag{3.97}$$

Or, using the full fledged Heisenberg picture, we can write the last line as follows, for a time-independent Hamiltonian

$$\frac{\mathrm{d}\hat{A}}{\mathrm{d}t} = \frac{i}{\hbar} \left[\hat{\mathcal{H}}, \hat{A}(t) \right] \tag{3.98}$$

Remembering the connection to the Poisson brackets, we already know that if the $[\mathcal{H}, A]_{PB} = 0$, then A is a constant of motion. In quantum mechanics, this will be represented as follows

$$\frac{\mathrm{d}\hat{A}}{\mathrm{d}t} = 0 \longrightarrow \left[\hat{\mathcal{H}}, \hat{A}\right] = 0 \tag{3.99}$$

The operator $\hat{A}$, is then called a *constant of motion* in quantum mechanics.

4 Angular Momentum

§ 4.1 Rotations

Taking as granted the knowledge that in classical physics rotations around the same axis commute and around different axes do not (it's easy to prove, even mentally by yourself). A classical rotation around a certain axis can be writted in tensor notation as follows. Let v_i be a vector in the fixed system and v_i' the same vector in the rotated system. We can then write the following relation (where $i, j = x, y, z$ or $i, j = 1, 2, 3$)

$$v_i' = R_{ij} v_j$$

Whereas

$$R_{ij} R_{ji} = R_{ji} R_{ij} = \delta_{ij} \longrightarrow R^{\mathsf{T}} = R^{-1}$$

The second equation is obvious, since applying an identical but opposed rotation on the same axis is exactly like letting act the identity rotation onto the vector (basically doing nothing to it).

Due to this matrix being then, orthogonal, we know from linear algebra that the following relation holds, and is automatically satisfied.

$$\sqrt{v_1^2 + v_2^2 + v_3^2} = \sqrt{v_1^{2'} + v_2^{2'} + v_3^{2'}}$$

From this relation, we can define pretty easily, that for a finite rotation of an angle ϕ, the matrix $R_{ij}^{(z)}(\phi)$, will have the following general form

$$R_{ij}^{(z)}(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.1}$$

The infinitesimal form all the rotation matrices, in order $\mathcal{O}(\varepsilon^3)$, will be the following, using the taylor expansions of the trigonometric functions

$$R_{ij}^{(z)}(\varepsilon) = \begin{pmatrix} 1 - \frac{\varepsilon^2}{2} & -\varepsilon & 0 \\ \varepsilon & 1 - \frac{\varepsilon^2}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.2a}$$

$$R_{ij}^{(x)}(\varepsilon) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - \frac{\varepsilon^2}{2} & -\varepsilon \\ 0 & \varepsilon & 1 - \frac{\varepsilon^2}{2} \end{pmatrix} \tag{4.2b}$$

$$R_{ij}^{(y)}(\varepsilon) = \begin{pmatrix} 1 - \frac{\varepsilon^2}{2} & 0 & \varepsilon \\ 0 & 1 & 0 \\ -\varepsilon & 0 & 1 - \frac{\varepsilon^2}{2} \end{pmatrix} \tag{4.2c}$$

It's obvious, that the rotations commute only for $\mathcal{O}(\varepsilon)$, and that they follow a cyclic relation. We might write this as follows for rotations of order $\mathcal{O}(\varepsilon^3)$

$$[R_x(\varepsilon), R_y(\varepsilon)] = R_x(\varepsilon^2) - R_{(i)}(0) \tag{4.3}$$

Where $R_{(i)}(0)$ is a null rotation along a generic axis, which is exactly an identity matrix, for every axis we choose.

§§ 4.1.1 Infinitesimal Rotations in Quantum Mechanics

In order to "quantize" the previous section, we assign to the rotation matrix, a rotation operator $\hat{\mathcal{D}}(R)$, with D as in *Drehung*, which means rotation in German.
Calling $|\cdot\rangle_r$ the rotated ket, we will have that this operator will act in the following way

$$|s\rangle_r = \hat{\mathcal{D}}(R)|s\rangle \tag{4.4}$$

Using an analogy with classical mechanics, and knowing that angular momentum is the generator for infinitesimal rotations, we can imagine that in quantum mechanics, this generator $\hat{G}$ will simply be the following $\hat{\underline{J}}/\hbar$, where we indicated with $\hat{\underline{J}}$ a generic angular momentum operator, and an $\hbar$ added for dimensionality reasons.
In first approximation, we can see the Drehung operator to the first order, for a rotation of $\mathrm{d}\phi$ degrees, as follows

$$\hat{\mathcal{D}}(\hat{\mathbf{n}}, \mathrm{d}\phi) = 1 - i\left(\frac{\hat{\underline{J}} \cdot \hat{\mathbf{n}}}{\hbar}\right)\mathrm{d}\phi$$

Repeating this rotation N times and sending this N to infinity, we can see this as a finite rotation, and due to the way it's written, we can immediately see how the Drehung operator for a finite rotation is shaped. Taking without loss of generality a rotation around the z axis, we get the following result for a rotation of ϕ degrees.

$$\hat{\mathcal{D}}(\phi) = \lim_{N \to \infty}\left(1 - \frac{i\hat{J}_z\phi}{N\hbar}\right)^N = e^{-\frac{i\hat{J}_z\phi}{\hbar}} \tag{4.5}$$

From this simple result, we can immediately see how the Drehung operator is tied to classical rotation, as it has the same group properties of the classical rotation matrices, which are the following

1. Identity: $\hat{\mathcal{D}}(R) \cdot \hat{\mathbb{1}} = \hat{\mathcal{D}}(R)$

2. Closure: $\hat{\mathcal{D}}(R_1)\hat{\mathcal{D}}(R_2) = \hat{\mathcal{D}}(R_3)$

3. Invertibility: $\hat{\mathcal{D}}^{-1}(R)\hat{\mathcal{D}}(R) = \hat{\mathcal{D}}(R)\hat{\mathcal{D}}^{-1}(R) = \hat{\mathbb{1}}$

4. Associativity: $\hat{\mathcal{D}}(R_1)\left(\hat{\mathcal{D}}(R_2)\hat{\mathcal{D}}(R_3)\right) = \left(\hat{\mathcal{D}}(R_1)\hat{\mathcal{D}}(R_2)\right)\hat{\mathcal{D}}(R_3) = \hat{\mathcal{D}}(R_1)\hat{\mathcal{D}}(R_2)\hat{\mathcal{D}}(R_3)$

Now it's quick to ask how do these operators commute, then we use the analogy with the R matrices and use the commutation relations (4.3). We obviously use the approimation to the second order of $\hat{\mathcal{D}}(R)$

$$\left(1 - \frac{i\hat{J}_x\varepsilon}{\hbar} - \frac{\hat{J}_x^2\varepsilon^2}{2\hbar^2}\right)\left(1 - \frac{i\hat{J}_y\varepsilon}{\hbar} - \frac{\hat{J}_y^2\varepsilon^2}{2\hbar^2}\right) - \left(1 - \frac{i\hat{J}_y\varepsilon}{\hbar} - \frac{\hat{J}_y^2\varepsilon^2}{2\hbar^2}\right)\left(1 - \frac{i\hat{J}_x\varepsilon}{\hbar} - \frac{\hat{J}_x^2\varepsilon^2}{2\hbar^2}\right) =$$

$$= 1 - \frac{i\hat{J}_z\varepsilon^2}{\hbar} - 1$$

Equating the terms of order $\mathcal{O}(\varepsilon^2)$, we get that $\left[\hat{J}_x, \hat{J}_y\right] = i\hbar\hat{J}_z$, and through cyclic permutations (using the ϵ_{ijk} tensor) we get the following result

$$\left[\hat{J}_i, \hat{J}_j\right] = i\hbar\epsilon_{ijk}\hat{J}_k \tag{4.6}$$

These are the fundamental commutation relations of angular momentum, and they are said to generate a non-Abelian group of rotations, since two different operators of the same group do not commute.
Since we have defined angular momentum from rotations, we can easily say that these relations hold for «every kind» of angular momentum we find.

§ 4.2 Eigenvalues and Eigenstates of Angular Momentum

We have already seen how angular momentum operators between different axes do not commute, hence, in order to find a suitable eigenbasis we build a new operator starting from $\underline{\hat{J}}$. From (4.6) we can immediately imagine that the simplest operator we can find is $\hat{J}^2$, defined as follows

$$\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 = \underline{\hat{J}} \cdot \underline{\hat{J}}$$

This operator commutes with every $\hat{J}_i$, and hence, we have another commutation relations

$$\left[\hat{J}^2, \hat{J}_i\right] = 0 \tag{4.7}$$

We know already that there exists simultaneously an eigenbasis for any operator $\hat{J}_i$ and $\hat{J}^2$. As a convention we take $\hat{J}_3 = \hat{J}_z$ as our main direction, and we will call the eigenkets as $|a, b\rangle$, for which holds the already known secular equation

$$\hat{J}^2 |a, b\rangle = a |a, b\rangle$$
$$\hat{J}_z |a, b\rangle = b |a, b\rangle$$

In order to work out the results, we define two non hermitian operators, that we will call $\hat{J}_+$ and $\hat{J}_+^\dagger = \hat{J}_-$. These operators are defined as follows

$$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y \tag{4.8}$$

These operators satisfy the commutation relations

$$\left[\hat{J}_z, \hat{J}_\pm\right] = \pm\hbar\hat{J}_\pm$$
$$\left[\hat{J}_\pm, \hat{J}_\mp\right] = \pm 2\hbar\hat{J}_z \tag{4.9}$$
$$\left[\hat{J}^2, \hat{J}_\pm\right] = 0$$

Defined as such, these are the ladder operators of angular momentum, but why are they called ladder operators? It's easy to see why if we let them act on an eigenstate and utilize the previous commutation relations. We then get

$$\hat{J}_z\hat{J}_\pm |a, b\rangle = \left(\left[\hat{J}_z, \hat{J}_\pm\right] + \hat{J}_\pm\hat{J}_z\right)|a, b\rangle = (b \pm \hbar)\,\hat{J}_\pm |a, b\rangle$$

In an analogy with the quantum harmonic oscillator, we see then really why these operators are called "ladder" operators, since they move up or down of one step ($\hbar$ long) the "measured" value.

We now redo the same calculations with $\hat{J}^2$, and using (4.9), we get

$$\hat{J}^2\hat{J}_\pm |a, b\rangle = \hat{J}_\pm\hat{J}^2 |a, b\rangle = a\hat{J}_\pm |a, b\rangle \tag{4.10}$$

Remembering that $|a, b\rangle$ are simultaneous eigenstates for both $\hat{J}_z$ and $\hat{J}^2$, we can summarize everything in the following way

$$\hat{J}_\pm |a, b\rangle = c_\pm |a, b \pm \hbar\rangle \tag{4.11}$$

Now, let's use what we found with this ladder operator machinery for finding the actual eigenstates and eigenvalues of angular momentum. We give the following Ansatz:

$$a \geq b^2$$

But why should the eigenvalue of $\hat{J}_z$ be limited? Let's write a new operator, made through a combination of $\hat{J}^2$ and $\hat{J}_z$

$$\hat{J}^2 - \hat{J}_z^2 = \frac{1}{2}\left(\hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+\right) = \frac{1}{2}\left(\hat{J}_+\hat{J}_+^\dagger + \hat{J}_+^\dagger\hat{J}_+\right) \tag{4.12}$$

Now, $\hat{J}_+^\dagger\hat{J}_+$ and $\hat{J}_+\hat{J}_+^\dagger$ must be positive definite, hence it follows that

$$\langle a, b|\left(\hat{J}^2 - \hat{J}_z^2\right)|a, b\rangle \geq 0$$

Which implies the previous Ansatz. It follows, then, that there exists a value b_{max} and b_{min} for the following relations hold

$$\hat{J}_+ |a, b_{max}\rangle = 0$$
$$\hat{J}_- |a, b_{min}\rangle = 0$$

(4.13)

Studying the first case, we know that it also implies $\hat{J}_- \hat{J}_+ |a, b_{max}\rangle = 0$, but analyzing further and utilizing the definition of the ladder operators, we get that

$$\hat{J}_- \hat{J}_+ = \hat{J}_x^2 + \hat{J}_y^2 - i\left[\hat{J}_y, \hat{J}_x\right] = \hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z$$

Hence, we get the following important result

$$\left(\hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z\right) |a, b_{max}\rangle = 0$$

(4.14)

Now, since $|a, b_{max}\rangle$ is an eigenket of every operator acting on it, and for such can't be a null ket, it must hold that

$$a - b_{max}^2 - \hbar b_{max} = 0$$
$$a = b_{max}(b_{max} + \hbar)$$

(4.15)

Similarly, letting $\hat{J}_-$ on the ket $|a, b_{min}\rangle$, we obtain that a must also be equal to

$$a = b_{min}(b_{min} - \hbar)$$

Hence, it's obvious that $b_{max} = -b_{min}$, and hence $-b_{max} \leq b \leq b_{max}$, with $b_{max} \geq 0$. Due to how $\hat{J}_+$ acts on the eigenkets, it must definitely hold that $b_{max} = b_{min} + n\hbar$, with $n \in \mathbb{N}$.
Then, it must hold definitely that

$$b_{max} = \frac{n\hbar}{2}$$

Due to convention it's usual to define directly the eigenvalue j as $n/2$, hence, we get that

$$a = \hbar^2 j(j+1)$$

(4.16)

Defining another integer eigenvalue m, such that we have

$$b = m\hbar$$

(4.17)

Due to what we found before, j is an half-integer, hence all m must be half-integers too, and they can only take the following values

$$m = \underbrace{-j, -j+1, \cdots, j-1, j}_{2j+1 \text{ times}}$$

(4.18)

For a little recap, after renaming a, b with j, m we have that the simultaneous eigenkets of $\hat{J}^2$ and $\hat{J}_z$ are the following

$$\hat{J}^2 |j, m\rangle = \hbar^2 j(j+1) |j, m\rangle$$
$$\hat{J}_z |j, m\rangle = \hbar m |j, m\rangle$$

(4.19)

§§ 4.2.1 Matrix Elements of the Ladder Operators of Angular Momentum

Now that we have defined the eigenvalues of the $\hat{J}^2$ and $\hat{J}_z$ angular momentum operators, is immediately seen that their tensor representation will be the following:

$$J^2_{\bar{j},\bar{m},j,m} = \langle \bar{j}, \bar{m} | \hat{J}^2 | j, m \rangle = \hbar^2 j(j+1)\delta_{\bar{j}j}\delta_{\bar{m}m}$$
$$J^{(z)}_{\bar{j},\bar{m},j,m} = \langle \bar{j}, \bar{m} | \hat{J}_z | j, m \rangle = \hbar m \delta_{\bar{j}j}\delta_{\bar{m}m}$$

(4.20)

The first question that might pop up is, then, what are the matrix elements of the ladder operators? Hence, what's their action on the eigenkets of angular momentum?
We use the definition in (4.12) in order to find this out. From that equation, we can write the two following useful relations

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

(4.21)

Let's now calculate the matrix elements for these operators, bearing in mind that $\hat{J}_- = \hat{J}_+^\dagger$

$$\langle j.m | \hat{J}_+^\dagger \hat{J}_+ | j, m \rangle = \langle j, m | \left(\hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z\right) | j, m \rangle = \hbar^2 j(j+1) - \hbar^2 m^2 - \hbar^2 m$$
$$\langle j, m | \hat{J}_-^\dagger \hat{J}_- | j, m \rangle = \langle j, m | \left(\hat{J}^2 - \hat{J}_z^2 + \hbar\hat{J}_z\right) | j, m \rangle = \hbar^2 j(j+1) - \hbar^2 m^2 + \hbar^2 m$$

(4.22)

But, we can also write the following relations

$$\hat{J}_\pm | j, m \rangle = c_\pm | j, m \pm 1 \rangle$$

And, in order to find $c_\pm$, we impose that $\left\| \hat{J}_\pm | j, m \rangle \right\| = 1$, and confronting with (4.22), we can conclude that

$$|c_\pm|^2 = \hbar^2 \left(j(j+1) - m(m \pm 1)\right) \longrightarrow c_\pm = \hbar\sqrt{j(j+1) - m(m \pm 1)}$$

(4.23)

And we can conclude, definitely, that the action of the ladder operators will be the following

$$\hat{J}_\pm | j, m \rangle = \hbar\sqrt{j(j+1) - m(m \pm 1)} | j, m \pm 1 \rangle$$

(4.24)

The tensor representations of the operators will then be

$$J^{(+)}_{\bar{j},\bar{m},j,m} = \hbar\sqrt{j(j+1) - m(m-1)}\delta_{\bar{m}m+1}\delta_{\bar{j}j}$$
$$J^{(-)}_{\bar{j},\bar{m},j,m} = \hbar\sqrt{j(j+1) - m(m+1)}\delta_{\bar{m}m-1}\delta_{\bar{j}j}$$

(4.25)

§ 4.3 Orbital Angular Momentum

We have seen how a general angular momentum, seen as a generator of rotations can be quantized, and be represented as an operator. To delve deeper on the physical meaning of this, we then quantize the classical angular momentum, where it can be written quantum mechanically as follows

$$\hat{\underline{L}} = \hat{\underline{q}} \wedge \hat{\underline{p}} = \epsilon_{ijk} \hat{q}_j \hat{p}_k \qquad (4.26)$$

It is immediately seen that it satisfies the main commutation relation of angular momentum, and hence it must be a generator of rotations. In fact, it's easy to demonstrate that

$$\left[\hat{L}_i, \hat{L}_j \right] = i\hbar\epsilon_{ijk} \hat{L}_k \qquad (4.27)$$

Since this must be a generator of rotations, we can write at the first order the Drehung operator as follows

$$\hat{\mathcal{D}}_z(\delta\phi) = 1 - \left(\frac{i\delta\phi}{\hbar} \right) \hat{L}_z = 1 - \left(\frac{i\delta\phi}{\hbar} \right) (\hat{x}\hat{p}_y - \hat{y}\hat{p}_x) \qquad (4.28)$$

It's already clear from the beginning, that letting this act on a ket $|x, y, z\rangle$ is uncomfortable, hence we choose a spherical coordinate set $|r, \theta, \phi\rangle$, and we see that the action, to the first order, of the Drehung operator will be the following, considering a toy wavefunction $|\alpha\rangle$

$$\langle r, \theta, \phi| \left(1 - \left(\frac{i\delta\phi}{\hbar} \right) \hat{L}_z \right) |\alpha\rangle = \langle r, \theta, \phi - \delta\phi|\alpha\rangle = \langle r, \theta, \phi|\alpha\rangle - \frac{\partial}{\partial\phi} \langle r, \theta, \phi|\alpha\rangle \, \delta\phi \quad (4.29)$$

Since $|r, \theta, \phi\rangle$ is arbitrary, we easily identify the following representation of $\hat{L}_z$

$$\langle r_i| \hat{L}_z |\alpha\rangle = -i\hbar \frac{\partial}{\partial\phi} \langle r_i|\alpha\rangle \qquad (4.30)$$

Identically, doing the same evaluation with $\hat{L}_x$ and $\hat{L}_y$, we get the following representations

$$\langle r_i| \hat{L}_x |\alpha\rangle = -i\hbar \left(-\sin\phi \frac{\partial}{\partial\theta} - \cot\theta\cos\phi \frac{\partial}{\partial\phi} \right) \langle r_i|\alpha\rangle$$
$$\langle r_i| \hat{L}_y |\alpha\rangle = -i\hbar \left(\cos\phi \frac{\partial}{\partial\theta} - \cot\theta\sin\phi \frac{\partial}{\partial\phi} \right) \langle r_i|\alpha\rangle \qquad (4.31)$$

Now, defining two new ladder operators for orbital angular momentum as $\hat{L}_\pm$, we have, combining the representations (4.31), the following result

$$\langle r_i| \hat{L}_\pm |\alpha\rangle = -i\hbar e^{\pm i\phi} \left(\pm i\frac{\partial}{\partial\theta} - \cot\theta \frac{\partial}{\partial\phi} \right) \langle r_i|\alpha\rangle \qquad (4.32)$$

As for $\hat{J}^2$, we can write $\hat{L}^2$ as follows

$$\hat{L}^2 = \hat{L}_z^2 + \frac{1}{2} \left(\hat{L}_+\hat{L}_- + \hat{L}_-\hat{L}_+ \right) \qquad (4.33)$$

Applying to our wavefunction $|\alpha\rangle$ we, hence get

$$\langle r_i| \hat{L}^2 |\alpha\rangle = -\hbar^2 \left(\csc^2\theta \frac{\partial^2}{\partial\phi^2} + \csc\theta \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) \right) \langle r_i|\alpha\rangle \tag{4.34}$$

Which is only the angular part of the Laplacian in spherical coordinates.
Using the property (A.2d) of the tensor ϵ_{ijk}, we can also write $\hat{L}^2$ in operatorial form as follows

$$\hat{L}^2 = \hat{q}^2\hat{p}^2 - \left(\underline{\hat{q}} \cdot \underline{\hat{p}} \right)^2 + i\hbar \underline{\hat{q}} \cdot \underline{\hat{p}} \tag{4.35}$$

Using this last expression, we manage to get the following results

$$\langle r_i| \hat{q}^j\hat{p}_j |\alpha\rangle = \hat{q}^j \left(-i\hbar \frac{\partial}{\partial x_j} \langle r_i|\alpha\rangle \right)$$

$$\langle r_i| \left(\hat{q}^j\hat{p}_j \right)^2 |\alpha\rangle = -\hbar^2 r \frac{\partial}{\partial r} \left(r \frac{\partial}{\partial r} \langle r_i|\alpha\rangle \right)$$

Thus, we have

$$\langle r_i| \hat{L}^2 |\alpha\rangle = r^2 \langle r_i| \hat{p}^2 |\alpha\rangle + \hbar^2 \left(r^2 \frac{\partial^2}{\partial r^2} + 2r \frac{\partial}{\partial r} \right) \langle r_i|\alpha\rangle \tag{4.36}$$

Now, having $\hat{p}^2$ in the definition of $\hat{L}^2$, we can write the kinetic energy as follows

$$\hat{T} \langle r_i|\alpha\rangle = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} \langle r_i|\alpha\rangle + \frac{2}{r} \frac{\partial}{\partial r} \langle r_i|\alpha\rangle - \frac{\langle r_i| \hat{L}^2 |\alpha\rangle}{\hbar^2 r^2} \right) \tag{4.37}$$

Let's now write the actual eigenfunctions $\langle r_i|l, m\rangle^1$ of $\hat{L}^2, \hat{L}_z$. Solving for both $\hat{L}_z$ and $\hat{L}^2$, it's easy to see that they are the Spherical Harmonics $Y_l^m(\theta, \phi)$, hence $\langle r_i|l, m\rangle = Y_l^m(\theta, \phi)$, and

$$\hat{L}_z Y_l^m(\theta, \phi) = \hbar m Y_l^m(\theta, \phi)$$
$$\hat{L}^2 Y_l^m(\theta, \phi) = \hbar^2 l(l + 1) Y_l^m(\theta, \phi) \tag{4.38}$$

The treatment of the spherical harmonics and its derivation is given in the appendices.

§ 4.4 Spin

So far we managed to define how the algebra of angular momentum works, and how the quantization of orbital angular momentum follows it. A main question arises: due to how orbital angular momentum is defined, it can only have integer eigenvalues, but as we've seen in the general picture, the eigenvalues can also be half-integers, which cannot be explained by orbital angular momentum. This problem arises, where there is this kind of

---

[1]The two integers l and m, are usually identified in literature as *principal quantum number* and *magnetic quantum number*, respectively, due to experimental reasons

intrinsic angular momentum, which cannot be defined through the quantization of classical objects! This new angular momentum, is called *Spin Angular Momentum* or simply *Spin*, and it's represented through a vector operator $\underline{\hat{S}}$.

Since it's an angular momentum, it follows every single commutation rule of a generator of rotations, hence it satisfies the following rules

$$
\begin{aligned}
\left[\hat{S}_i, \hat{S}_j\right] &= i\hbar\epsilon_{ijk}S_k \\
\left[\hat{S}^2, \hat{S}_i\right] &= 0 \\
\hat{S}^2 \left|s, m\right\rangle &= \hbar^2 s(s+1) \left|s, m\right\rangle \\
\hat{S}_z \left|s, m\right\rangle &= \hbar m_s \left|s, m\right\rangle
\end{aligned}
\tag{4.39}
$$

Due to the previous statements, we already know that s takes half integer values, but using the algebraic definition of angular momentum, it's obvious that it can also take integer values.

The simplest case that comes to mind, is the case where $s = 1/2$. In this case we have, since $-s \leq m \leq s$, that

$$
\begin{aligned}
\hat{S}_z \left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle &= \pm\frac{\hbar}{2} \left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle \\
\hat{S}^2 \left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle &= \frac{3\hbar^2}{4} \left|\frac{1}{2}, \pm\frac{1}{2}\right\rangle
\end{aligned}
\tag{4.40}
$$

Due to the two different possible values of m_s, we can split the ket, in order to have the following result

$$
\begin{aligned}
\hat{S}_z \left|\frac{1}{2}, \frac{1}{2}\right\rangle &= \frac{\hbar}{2} \left|\frac{1}{2}, \frac{1}{2}\right\rangle \\
\hat{S}_z \left|\frac{1}{2}, -\frac{1}{2}\right\rangle &= -\frac{\hbar}{2} \left|\frac{1}{2}, -\frac{1}{2}\right\rangle
\end{aligned}
\tag{4.41}
$$

Immediately, it's easy to define then, two possible states, either spin "up" or spin "down". This new notation follows the previous rules, although adding some day to day physical intuition that this argument completely lacks:

$$
\begin{aligned}
\hat{S}_z \left|\uparrow\right\rangle &= \frac{\hbar}{2} \left|\uparrow\right\rangle \\
\hat{S}_z \left|\downarrow\right\rangle &= -\frac{\hbar}{2} \left|\downarrow\right\rangle \\
\hat{S}^2 \left|\uparrow\right\rangle = \hat{S}^2 \left|\downarrow\right\rangle &= \frac{3\hbar^2}{4} \left|\uparrow\right\rangle = \frac{3\hbar^2}{4} \left|\downarrow\right\rangle
\end{aligned}
\tag{4.42}
$$

The $\hat{S}_z$ is diagonal in this basis, hence we can immediately write it in matrix form as follows

$$
\hat{S}_z \rightarrow \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
\tag{4.43}
$$

And $\hat{S}^2$, as

$$\hat{S}^2 \rightarrow \frac{3\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{4.44}$$

It jumps immediately to the eye that we're now working on a 2D Hilbert space. This space is spanned by the kets $|\uparrow\rangle$ and $|\downarrow\rangle$, which, due to their nature as spin eigenkets, are called *spinors*.

Immediately, a question comes to mind: how do the other spin components act on these spinors? Firstly, let's define our spin ladder operators $\hat{S}_\pm$ as usual. Since the only possible values of m are $m = -1/2, 1/2$ they'll be represented by the following matrices

$$\hat{S}_+ \rightarrow \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$\hat{S}_- \rightarrow \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{4.45}$$

The $\hbar$ pops out remembering the following rule of angular momentum ladder operators

$$\hat{S}_\pm |s, m\rangle = \hbar\sqrt{s(s+1) - m(m \pm 1)} |s, m \pm 1\rangle \tag{4.46}$$

Now it's easy to answer our first question! We remember how the ladder operators are defined, and we invert them in order to find how we can define $\hat{S}_x$ and $\hat{S}_y$ in terms of these operators, their calculation is quite straightforward, and we get

$$\hat{S}_x = \frac{1}{2} \left(\hat{S}_+ + \hat{S}_- \right)$$

$$\hat{S}_y = \frac{1}{2i} \left(\hat{S}_+ - \hat{S}_- \right) \tag{4.47}$$

Their representation is then easy to find

$$\hat{S}_x \rightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\hat{S}_y \rightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{4.48}$$

Putting it all together, the $\underline{\hat{S}}$ is then represented as follows

$$\hat{S}_x \rightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \frac{\hbar}{2}\hat{\sigma}_x$$

$$\hat{S}_y \rightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \frac{\hbar}{2}\hat{\sigma}_y \tag{4.49}$$

$$\hat{S}_z \rightarrow \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \frac{\hbar}{2}\hat{\sigma}_z$$

Using index notation, we get that $\hat{S}_i = \frac{\hbar}{2}\hat{\sigma}_i$, where $\hat{\sigma}_i$ is the i-th *Pauli matrix* for a spin $1/2$ system.

These matrices have the following properties, inherited from the properties of Spin

$$[\hat{\sigma}_i, \hat{\sigma}_j] = 2i\epsilon_{ijk}\hat{\sigma}_k$$
$$\{\hat{\sigma}_i, \hat{\sigma}_j\} = 2\delta_{ij}\hat{\mathbb{1}}$$

(4.50)

From these, it's easy to derive some additional properties

$$\hat{\sigma}_i\hat{\sigma}_j = \delta_{ij}\hat{\mathbb{1}} + \epsilon_{ijk}\hat{\sigma}_k$$
$$A_{ij} = c_0\delta_{ij} + c_k\hat{\sigma}^k$$
$$\hat{\sigma}^2 = \delta_{ij}$$
$$\hat{\sigma}_i\hat{\sigma}^i = 3\delta_{ij}$$

§ 4.5 Addition of Angular Momenta

In order to add up two different angular momenta, we need, first of all, to understand how the underlying maths works.

The two operators act on two different Hilbert spaces, and the *total* angular momentum must act on both. This constraints bring us to the definition of the total angular momentum Hilbert space, which «must» be given by the tensor product of the two, as follows

$$\mathbb{H} = \mathbb{H}_1 \otimes \mathbb{H}_2$$

(4.51)

The operator $\hat{\underline{J}}$ will then be defined as

$$\hat{\underline{J}} = \hat{\underline{J}}_1 \otimes \hat{\mathbb{1}} + \hat{\mathbb{1}} \otimes \hat{\underline{J}}_2$$

(4.52)

From this definition, the two following commutation relations are then obvious

$$\left[\hat{J}_2^2, \hat{J}^2\right] = 0$$
$$\left[\hat{J}^2, \hat{J}_1^2\right] = 0$$

(4.53)

Although, it follows that $\hat{J}^2$ and $\hat{J}_1^z, \hat{J}_2^z$ do not commute.

$$\left[\hat{J}_2^z, \hat{J}^2\right] = \left[\hat{J}^2, \hat{J}_1^z\right] = 2i\hbar\left(\hat{J}_1^x\hat{J}_2^y - \hat{J}_1^y\hat{J}_2^x\right) \neq 0$$

(4.54)

Due to the last relation a common basis of eigenvalues between $\hat{J}^2$ and $\hat{J}_1^z, \hat{J}_2^z$ cannot be formed, but two choices are possible:

1. Since $\hat{J}_1^2$ commutes with $\hat{J}_2^2, \hat{J}_1^z, \hat{J}_2^z$ and each of them commutes, we can utilize a basis $B := |j_1, m_1\rangle \otimes |j_2, m_2\rangle$

2. Since $\hat{J}^2$ commutes only with $\hat{J}_1^2, \hat{J}_2^2, \hat{J}_z$, we can also define a second basis $C :=$ $|j_1, j_2\rangle \otimes |j, m\rangle$

We will then have the following eigenvalues (where we avoid writing the tensor product between states in order to ease the notation)

$$\hat{J}^2 |j_1, j_2, j, m\rangle = \hbar^2 j(j+1) |j_1, j_2, j, m\rangle$$
$$\hat{J}_z |j_1, j_2, j, m\rangle = \hbar m |j_1, j_2, j, m\rangle$$
$$\hat{J}_1^2 |j_1, j_2, j, m\rangle = \hbar^2 j_1(j_1+1) |j_1, j_2, j, m\rangle$$
$$\hat{J}_2^2 |l.s.j.m\rangle = \hbar^2 j_2(j_2+1) |j_1, j_2, j.m\rangle \tag{4.55}$$
$$\hat{J}_2^z |j_1, j_2, m_1, m_2\rangle = \hbar m_2 |j_1, j_2, m_1, m_2\rangle$$
$$\hat{J}_1^z |j_1, j_2, m_1, m_2\rangle = \hbar m_1 |j_1, j_2, m_1, m_2\rangle$$

We can obviously define an unitary transformation in $\mathbb{H}$ that lets us switch between the two basi

$$|j_1, j_2, j, m\rangle = \hat{\mathcal{U}} |j_1, j_2, m_1, m_2\rangle \tag{4.56}$$

The shape of $\hat{\mathcal{U}}$ will obviously be that of a projection, and hence we will have the following result

$$|j_1, j_2, j, m\rangle = \sum_{m_1, m_2} |j_1, j_2, m_1, m_2\rangle \langle j_1, j_2, m_1, m_2|j_1, j_2, j, m\rangle \tag{4.57}$$

Where here we utilized the completeness relation

$$\sum_{m_1, m_2} |j_1, j_2, m_1, m_2\rangle \langle j_1, j_2, m_1, m_2| = \hat{\mathbb{1}}$$

Looking at the braket on the right of (4.57), we immediately see the property of these coefficients, called *Clebsch-Gordan coefficients*. These coefficients vanish, unless we have that $m = m_1 + m_2$, but how?

Proof. We know that the following assertion is true

$$\left(\hat{J}_z - \hat{J}_1^z - \hat{J}_2^z\right) |j_1, j_2, j, m\rangle = 0$$

We simply multiply the relation with $\langle j_1, j_2, m_1, m_2|$ and we get our proof

$$(m - m_1 - m_2) \langle j_1, j_2, m_1, m_2|j_1, j_2, j, m\rangle$$

Where the last braket are our Clebsch-Gordan coefficients   □

Similarly, we have that the Clebsch-Gordan coefficients are nonzero also if only the $\hat{J}^2$ eigenvalue j holds the following values

$$|j_1 - j_2| \leq j \leq j_1 + j_2 \tag{4.58}$$

One might now ask how many states we can count, after adding the two angular momentums.

We can already count $2j_1 + 1$ possible values of m_1 and $2j_2 + 1$ values of m_2. Summing it all up we get that there will be N states, counted as follows

$$N = \sum_{j=j_1-j_2}^{j_1+j_2} (2j+1) = \frac{1}{2}\left(2(j_1 - j_2) + 1 + 2(j_1 + j_2) + 1\right)(2j_2 + 1)$$
$$= (2j_1 + 1)(2j_2 + 1)$$

(4.59)

The Clebsch-Gordan coefficients are taken to be real by covention, hence the inverse coefficients are equal to the coefficient themselves, since it's a unitary transformation.

Another way that these coefficients can be written, is with the *Wigner 3-j symbols*, through this relation

$$\langle j_1, j_2, m_1, m_2 | j_1, j_2, j, m \rangle = (-1)^{j_1 - j_2 + m}\sqrt{2j+1}\begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & -m \end{pmatrix}$$

(4.60)

The values that the "matrix" can take are tabulated, and therefore, even the Clebsch-Gordan coefficients.

5 Quantum Dynamics in 3D

In this chapter, we will treat problems defined by Hamiltonians of the following kind

$$\hat{\mathcal{H}} = \frac{\hat{p}^2}{2m} + V(r) \tag{5.1}$$

Where

$$\begin{cases} \hat{p}^2 = p_i p^i \\ \hat{r} = \sqrt{\hat{q}_i \hat{q}^i} \end{cases} \quad i = 1, 2, 3$$

Due to this definition, it's evident that this Hamiltonian is *spherically symmetrical*, because it's easy to show that angular momentum is conserved. In fact

$$\left[\hat{L}_i, \hat{p}^2 \right] = \left[\hat{L}_i, \hat{q}^2 \right] = 0$$

Due to how the Hamiltonian is defined, we have

$$\left[\hat{L}_i, \hat{\mathcal{H}} \right] = \left[\hat{L}^2, \hat{\mathcal{H}} \right] = 0 \tag{5.2}$$

Which, indicates that angular momentum is conserved. The fact that the commutator between every element of the angular momentum and the Hamiltonian brings our problem to the search of common eigenstates of angular momentum and energy, that for simplicity we will call $|Elm\rangle$.

The secular equations that we need to solve simultaneously are the following three

$$\hat{\mathcal{H}} |Elm\rangle = E |Elm\rangle$$
$$\hat{L}^2 |Elm\rangle = \hbar^2 l(l+1) |Elm\rangle \tag{5.3}$$
$$\hat{L}_z |Elm\rangle = \hbar m |Elm\rangle$$

Using what we found for angular momentum, we write $\langle x_i| \hat{p}^2 |Elm\rangle$ in spherical coordinates, and, we then get

$$\langle x_i| \hat{\mathcal{H}} |Elm\rangle = -\frac{\hbar^2}{2mr^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) \psi_{Elm}(x_i) +$$
$$+ \frac{\hbar^2 l(l+1)}{2mr^2} \psi_{Elm}(x_i) + V(r)\psi_{Elm}(x_i) = E\psi_{Elm}(x_i) \tag{5.4}$$

As we seen for angular momentum, $\hat{L}^2$ takes the angular dependence, hence this equation can be reduced to a single variable differential equation on r through separation of variable, utilizing the fact that $\langle \theta, \phi | lm \rangle \neq 0$ for all θ, ϕ. As a convention, we will call the radial part $R_{El}(r)$

We simplify the equation (the *radial* equation), introducing a new radial function, and an effective potential

$$\begin{cases} R_{El} = \dfrac{u_{El}(r)}{r} \\ V_{eff}(r) = \dfrac{\hbar^2 l(l+1)}{2mr^2} + V(r) \end{cases} \tag{5.5}$$

The new equation then becomes the following

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2 u_{El}}{\mathrm{d}r^2} + V_{eff}(r)u_{El}(r) = Eu_{El}(r) \tag{5.6}$$

Considering the behavior in an infinitesimal ball around $r = 0$, and supposing that our potential is regular enough to have $\lim_{r\to 0} V(r) = 0$, we have that our equation becomes the following

$$\frac{\mathrm{d}^2 u_{El}}{\mathrm{d}r^2} = \frac{l(l+1)}{r^2}u_{El}(r) \tag{5.7}$$

Which has the following general solution

$$u(r) = Ar^{l+1} + Br^{-l}$$

For our needs, we set $B = 0$, since r^{-l} blows up at $r \to 0$. We have now a restriction for R_{El}, which imposes that $R_{El}(r) \to r^l$ for $r \to 0$.

For $r \to \infty$ we get a new equation, of simple solution

$$\frac{\mathrm{d}^2 u_E}{\mathrm{d}r^2} = \kappa^2 u_E(r) \quad \kappa^2 = -\frac{2mE}{\hbar^2} > 0 \tag{5.8}$$

It's solution is a decaying exponential $u_E(r) \propto e^{-\kappa r}$

Introducing a new dimensionless variable $\rho = \kappa r$, we can write our radial function as a product of the asymptotic behaviors and a new unknown function $w(\rho)$

$$u_{El} = \rho^{l+1}e^{-\rho}w(\rho)$$

The function $w(\rho)$ is well behaved, and satisfies the following differential equation

$$\rho\frac{\mathrm{d}^2 w}{\mathrm{d}\rho^2} + 2\left(l+1-\rho\right)\frac{\mathrm{d}w}{\mathrm{d}\rho} + \left(\frac{V(\rho/\kappa)}{E}\rho - 2(l+1)\right)w(\rho) = 0 \tag{5.9}$$

This equation depends on the potential $V(r)$, hence a formal solution can't yet be defined.

§ 5.1 Free Particles and Infinite Wells Revisited

Starting from the radial equation, we define as before $\rho = kr$ and $E = \hbar^2 k^2/2m$. Plugging it all into the radial equation, we then get the following

$$\frac{\mathrm{d}^2 R}{\mathrm{d}\rho^2} + \frac{2}{\rho}\frac{\mathrm{d}R}{\mathrm{d}\rho} + \left(1 - \frac{l(l+1)}{\rho^2}\right) R(\rho) = 0 \tag{5.10}$$

This differential equation is solved immediately including the *spherical Bessel functions* j_l, n_l, defined as follows

$$j_l(\rho) = (-1)^l \rho^l \left(\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}\right)^l \left(\frac{\sin\rho}{\rho}\right)$$
$$n_l(\rho) = (-1)^{l+1} \rho^l \left(\frac{1}{\rho}\frac{\mathrm{d}}{\mathrm{d}\rho}\right)^l \left(\frac{\cos\rho}{\rho}\right) \tag{5.11}$$

This result can be applied directly to the problem of a particle confined inside an infinite spherical well ($V(\rho) = 0$ in $r < a$) through the imposition of the condition $j_l(ka) = 0$. At $l = 0, 1, 2$ the levels are non degenerate, and for $l = 0$ we then have the following energy values

$$E_{nl} = E_{n0} = \frac{\hbar^2(n\pi)^2}{2ma} \tag{5.12}$$

§ 5.2 Isotropic Harmonic Oscillator

The Hamiltonian for a 3D isotropic oscillator can be written as follows

$$\hat{\mathcal{H}} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{r}^2 \tag{5.13}$$

We introduce immediately the two following dimensionless variables and introduce them in the radial equation

$$E = \frac{1}{2}\hbar\omega\lambda$$
$$r = \sqrt{\frac{\hbar}{m\omega}}\rho$$

We obtain the following equation

$$\frac{\mathrm{d}^2 u}{\mathrm{d}\rho^2} - \frac{l(l+1)}{\rho^2}u(\rho) + (\lambda - \rho)^2 u(\rho) = 0 \tag{5.14}$$

We separate immediately the asymptotic behavior of $u(\rho)$, and since the potential diverges for $r \to \infty$, we write

$$u(\rho) = \rho^{l+1} e^{-\frac{\rho^2}{2}} f(\rho) \tag{5.15}$$

Inserting it back into (5.14), we get an equation on $f(\rho)$

$$\rho \frac{\mathrm{d}^2 f}{\mathrm{d}\rho^2} + 2\left(l + 1 - \rho^2\right) \frac{\mathrm{d}f}{\mathrm{d}\rho} + (\lambda - 2l + 3)\rho f(\rho) = 0 \tag{5.16}$$

This equation is solvable by writing $f(\rho)$ as a power series, and after manipulating it analogously to how it's done for a simple linear quantum harmonic oscillator, we get, after plugging it back into (5.16)

$$\sum_{n=2}^{\infty} \left[(n+2)(n+1)a_{n+2} + 2(l+1)(n+2)a_{n+2} - 2na_n + (\lambda - 2l + 3)a_n\right] \rho^{l+1} = 0 \tag{5.17}$$

Which gives the following recursion relation

$$a_{n+2} = \frac{2n + 2l + 3 - \lambda}{(n+2)(n+2l+3)} a_n \tag{5.18}$$

It's immediate to see that $\lim_{n\to\infty} a_{n+2}/a_n = 2/n = q^{-1}$, therefore, we get that $\lim_{\rho\to\infty} f(\rho) \propto e^{\rho^2}$, which gives us a non normalizable wavefunction. Due to this, the series must terminate for some even $q = 2n$, which gives the quantization of energy

$$\lambda = 2n + 2l + 3 \tag{5.19}$$

Which, gives finally

$$E_{ql} = \hbar\omega \left(2q + l + \frac{3}{2}\right) = \hbar\omega \left(N + \frac{3}{2}\right) \tag{5.20}$$

It's easy so see how energy is degenerate in l, and how for even (or odd) N we can only have even (or odd) values for l.

Another way to solve this problem is by seeing how the Hamiltonian is simply the sum of three Hamiltonians, one for each coordinate

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_x + \hat{\mathcal{H}}_y + \hat{\mathcal{H}}_y$$

And write each in terms of creation and destruction operators as follows

$$\hat{\mathcal{H}}_i = \hbar\omega \left(\hat{\eta}_i^\dagger \hat{\eta}_i + \frac{1}{2}\hat{\mathbb{1}}\right)$$

Labeling the eigenstates as $|n_x\rangle \otimes |n_y\rangle \otimes |n_z\rangle = |n_x, n_y, n_z\rangle$, we get then the simple result

$$\hat{\mathcal{H}} |n_x, n_y, n_z\rangle = \hbar\omega \left(n_x + \frac{1}{2} + n_y + \frac{1}{2} + n_z + \frac{1}{2}\right) |n_x, n_y, n_z\rangle$$

It's obvious that in this basis, the degeneration is the same of the previous, and it can be seen changing the basis using the following unitary transformation matrix $\langle n_x, n_y, n_z | qlm\rangle$

§ 5.3 Particle in a Coulomb Potential, Hydrogen Atom

We basically treated most of the common problems in 3 dimensions, but there is one potential that we didn't treat in a single dimension that will pop up various times in this book, especially when we'll start touching particular themes such as atomic physics and quantum chemistry: **The Coulomb potential**.
We write our potential in Gaussian units as follows

$$V(r) = -\frac{Ze^2}{r} \tag{5.21}$$

We already know the shape of this potential and what do the constants really mean, since we already treated it in the old quantum theory, so we immediately write our Schrödinger equation, remembering how on the general case, the assumption of $V(r) \propto r^{-1}$ brought the equation (5.9).
Considering that bound states happen only for $E < 0$ and defining a ρ_0 as

$$\rho_0 = \sqrt{-\frac{2m}{E}}\frac{Ze^2}{\hbar} = \sqrt{-\frac{2mc^2}{E}}Z\alpha$$

Where $\alpha = e^2/\hbar c \approx 137^{-1}$ is the fine structure constant. Inserting everything in (5.9) we get a particular differential equation

$$\rho\frac{\mathrm{d}^2w}{\mathrm{d}\rho^2} + 2(l+1-\rho)\frac{\mathrm{d}w}{\mathrm{d}\rho} + (\rho_0 - 2l + 2)w(\rho) = 0 \tag{5.22}$$

This equation is immediately solved by a confluent hypergeometric function[1], with parameters $\alpha = l + 1 - \rho$, $\gamma = 2l + 2$ and $z = 2\rho$, so

$$w(\rho) = F(2l + 2 - \rho_0, 2l + 2, 2\rho) \tag{5.23}$$

Approximating w with a power series for large N we get that

$$w(\rho) = \sum_{N=0}^{\infty}\frac{(2l+2-2\rho_0)_N}{(2l+2)_N}\frac{(2\rho)^N}{N!} \approx \sum_{N=0}^{\infty}\frac{(N/2)^N(2\rho)^N}{N^N N!} \approx \sum_{N=0}^{\infty}\frac{\rho^N}{N!} \approx e^\rho$$

Hence, this series must terminate, for some $n \geq \tilde{N}$, defined as $n = N + l + 1$. This number is called the *principal quantum number*[2].
Since we defined ρ_0 as $2N + 2l + 2$, we can write the energy eigenvalues as follows

$$\rho_0 = \sqrt{-\frac{E}{2mc^2}} = 2n$$

---

[1]See appendix B.2
[2]So far we found 3 quantum numbers, n the principal quantum number, l the angular quantum number and m the magnetic quantum number.

And therefore, solving for E, we get the energy quantization rule

$$E_n = -\frac{1}{2}mc^2\frac{Z^2\alpha^2}{n^2} \tag{5.24}$$

Now we can define properly our eigenfunction $\langle r, \theta, \phi | nlm \rangle$. As we know already the symmetries of the system, we know it will be composed by a radial part and a spherical part, multiplied together tensorially. Therefore, we have our wavefunction as follows

$$\psi_{nlm}(r, \theta, \phi) = \frac{1}{(2l+1)!}\left(\frac{2Zr}{na_0}\right)^l \\ \sqrt{\left(\frac{2Z}{na_0}\right)^3 \left(\frac{(n+l)!}{2n(n-l-1)!}\right)} F\left(-n+l+1, 2l+2, \frac{2Zr}{na_0}\right) Y_l^m(\theta, \phi) \tag{5.25}$$

Where $a_0 = \hbar/mc\alpha$ is Bohr's radius. The appearance of Bohr's radius in this equation is not casual, since the solution of the Schrödinger equation for a Coulomb potential is «identical» to the direct solution of the equation for a Hydrogen atom (non-relativistic). This finally closes at least partially all the questions that the old quantum theory left, and gave a proper solution to the main problem of atomic physics: the Hydrogen atom.

6 Approximation Methods

§ 6.1 Perturbation Theory

Since most problems in quantum mechanics can't be solved directly, there are various methods in order to approximate the results. In perturbation theory, we have two main problems and two cases: time-independent and time-dependent perturbations with or without degeneration. As a first approach, we will consider nondegenerate and time independent perturbations.

§§ 6.1.1 Rayleigh-Schrödinger Perturbation Theory, Nondegenerate Case

Let $\hat{\mathcal{H}}$ be our non directly solvable Hamiltonian, divisible in a sum of a solvable Hamiltonian $\hat{\mathcal{H}}_0$ and a perturbation $\hat{V}$.

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{V} \tag{6.1}$$

For $\hat{\mathcal{H}}_0$ we already know the solution of the secular equation, and we label them as follows

$$\hat{\mathcal{H}}_0 \ket{n_0} = E_n^{(0)} \ket{n_0} \tag{6.2}$$

Where the 0 should be seen strictly as a label, otherwise told.
The full secular equation will then be the following

$$\hat{\mathcal{H}} \ket{n} = \left(\hat{\mathcal{H}}_0 + \hat{V} \right) \ket{n} = E_n \ket{n} \tag{6.3}$$

It is customary to have $\hat{V}' = \lambda \hat{V}$, where $0 \leq |\lambda| \leq 1$ is a parameter that can be manipulated in order to see the effect of the perturbation.
A nice example on how this works is given by a two state system, defined as follows

$$\begin{aligned} \hat{\mathcal{H}}_0 &= E_1^{(0)} \ket{1_0}\bra{1_0} + E_2^{(0)} \ket{2_0}\bra{2_0} \\ \hat{V} &= \lambda V_{12} \ket{1_0}\bra{2_0} + \lambda V_{21} \ket{2_0}\bra{1_0} \end{aligned} \tag{6.4}$$

In matrix representation, our Hamiltonian will then be the following

$$\mathcal{H}_{ij} = \begin{pmatrix} E_1^{(0)} & \lambda V_{12} \\ \lambda V_{21} & E_2^{(0)} \end{pmatrix} \tag{6.5}$$

Since it must be (obviously) an Hermitian operator, we have $V_{12} = V_{21}$ Calculating the eigenvalues is simple, and we get

$$E_{1,2} = \frac{E_1^{(0)} + E_2^{(0)}}{2} \pm \sqrt{\frac{(E_1^{(0)} - E_2^{(0)})^2}{4} + \lambda^2 V_{12}^2} \tag{6.6}$$

Since we are considering perturbations tied to a parameter, if $\lambda|V_{12}| << \left| E_1^{(0)-E_2^{(0)}} \right|$ we can approximate the square root with a power series, obtaining the following result

$$E_1 = E_1^{(0)} + \frac{\lambda^2 |V_{12}|^2}{E_1^{(0)} - E_2^{(0)}}$$

$$E_2 = E_2^{(0)} + \frac{\lambda^2 |V_{12}|^2}{E_2^{(0)} - E_1^{(0)}} \tag{6.7}$$

Which are the perturbation-corrected eigenvalues.

Now it's simpler to grasp the «formal» development of the theory. We take our two secular equations (6.2) and (6.3), and rename the difference of eigenvalues found in the approximation (6.7) as follows $\Delta_n = E_n - E_n^{(0)}$.

Our new approximate Schrödinger equation can then be written as follows

$$\left(E_n^{(0)} - \hat{\mathcal{H}}_0 \right) |n\rangle = \left(\lambda\hat{V} - \Delta_n \right) |n\rangle \tag{6.8}$$

And now here a little precaution: $E_n^{(0)} - \hat{\mathcal{H}}_0$ and $\lambda\hat{V} - \Delta_n$ «are operators», and must be treated as such.

Back to our perturbation theory, we see right away that inverting the operator on the left isn't the way to go. It may act both on $|n_0\rangle, |n\rangle$, and therefore it's inverse is ill-defined, but the right hand side comes to our rescue, and we can impose the following condition as an Ansatz

$$\langle n_0| \left(\lambda\hat{V} - \Delta_n \right) |n\rangle = 0 \tag{6.9}$$

Now, we want to define properly the inverse of the operator on the left hand side, and we start by using a *complementary projection operator* $\hat{\phi}_n$ defined as follows

$$\hat{\phi}_n = 1 - |n_0\rangle \langle n_0| = \sum_{k \neq n} |k_0\rangle \langle k_0| \tag{6.10}$$

Now, in order to well-define the inverse operator we simply apply a projection beforehand

$$\frac{1}{E_n^{(0)} - \hat{\mathcal{H}}_0} \hat{\phi}_n = \sum_{k \neq n} \frac{1}{E_n^{(0)} - E_k^{(0)}} |k_0\rangle \langle k_0| \tag{6.11}$$

From the Ansatz (6.9) we have, evidently

$$\left(\lambda\hat{V} - \Delta_n \right) |n\rangle = \hat{\phi}_n \left(\lambda\hat{V} - \Delta_n \right) |n\rangle$$

So everything looks set up and fine, and it's tempting to find the perturbed eigenstates simply by inverting the first operator, but it simply doesn't work. Why? First of all, for $\lambda \to 0$ we «must» have $|n\rangle \to |n_0\rangle$, and Δ_0, secondly because we need to add the solution to the homogeneous equation, hence, finally, we get the following result, naming this solution $c_n |n\rangle$

$$|n\rangle = c_n(\lambda) |n_0\rangle + \frac{1}{E_n^{(0)} - \hat{\mathcal{H}}_0} \hat{\phi}_n \left(\lambda \hat{V} - \Delta_n \right) |n\rangle \tag{6.12}$$

Where we have $c_n(\lambda) \to 1$ for $\lambda \to 0$, and $c_n(\lambda) = \langle n_0|n\rangle$.
Simplifying the successive equations, we put $c_n(\lambda) = \langle n_0|n\rangle = 1$ as our normalization condition, effectively removing a common multiplicative factor that appears. Then, easing the notation, we get

$$|n\rangle = |n_0\rangle + \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} \left(\lambda \hat{V} - \Delta_n \right) |n\rangle \tag{6.13}$$

We also note that, from (6.9) that

$$\Delta_n = \lambda \langle n_0| \hat{V} |n\rangle \tag{6.14}$$

Now everything is set. What we are searching depends only on the equations (6.13) and (6.14), and using the "smallness" of λ we approximate everything using power series, hence

$$|n\rangle = \sum_{k=0}^{\infty} \lambda^k |n_k\rangle$$
$$\Delta_n = \sum_{k=0}^{\infty} \lambda^k \Delta_n^{(k)} \tag{6.15}$$

So, in order to evaluate the energy shift up to an order $\mathcal{O}(\lambda^N)$ it's sufficient to equate the coefficients of the powers of λ, putting simply the following condition $\Delta_n^{(N)} = \langle n_0| \hat{V} |n_{N-1}\rangle$. It's evident how we need to know $|n_k\rangle$ only up to $\mathcal{O}(\lambda^{N-1})$. Adding all this in (6.15), we get

$$|n_0\rangle + \lambda |n_1\rangle + \cdots = |n_0\rangle + \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} \left(\lambda \hat{V} - \lambda \Delta_n^{(1)} - \cdots \right) \left(|n_0\rangle + \lambda |n_1\rangle + \cdots \right)$$

Therefore, for $\mathcal{O}(\lambda)$ we will get the following (remembering that $\hat{\phi}_n \Delta_n^{(1)} |n_0\rangle = 0$)

$$|n_1\rangle = \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} \hat{V} |n_0\rangle \tag{6.16}$$

For $\mathcal{O}(\lambda^2)$ it gets trickier. Firstly we use the definition of $\Delta_n^{(2)}$, where

$$\Delta_n^{(2)} = \langle n_0| \hat{V} \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} \hat{V} |n_0\rangle$$

Plugging it into the power series approximation up to order 2, we get therefore

$$|n_2\rangle = \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} \hat{V} \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} |n_0\rangle -$$
$$- \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} \langle n_0| \hat{V} |n_0\rangle \frac{\hat{\phi}_n}{E_n^{(0)} - \hat{\mathcal{H}}_0} \hat{V} |n_0\rangle$$

(6.17)

Defining $\hat{\Phi} = \hat{\phi}/(E_n^{(0)} - \hat{\mathcal{H}}_0)$, we get the previous equations compacted

$$|n_1\rangle = \hat{\Phi}\hat{V} |n_0\rangle$$
$$|n_2\rangle = \hat{\Phi}\hat{V}\hat{\Phi}\hat{V} |n_0\rangle - \hat{\Phi} \left\langle \hat{V} \right\rangle_0 \hat{\Phi}\hat{V} |n_0\rangle$$

(6.18)

It's evident that there is a trend in how next-order perturbations can be found, in this not-so-simple pattern.
Written explicitly, it's evident how this works

$$|n\rangle = |n_0\rangle + \lambda \sum_{k \neq n} \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}} |k_0\rangle + \lambda^2 \sum_{k \neq n} \sum_{l \neq n} \frac{V_{kl}V_{ln}}{(E_n^{(0)} - E_k^{(0)})(E_n^{(0)} - E_l^{(0)})} |k_0\rangle -$$
$$- \lambda^2 \sum_{k \neq n} \frac{V_{nn}V_{kn}}{(E_n^{(0)} - E_k^{(0)})^2} |k_0\rangle + \cdots$$

§§ 6.1.2 Rayleigh-Schrödinger Perturbation Theory, Degenerate Case

What we have defined so far, fails when the eigenstates we perturb are degenerate, since we supposed that there was only one well-defined eigenvalue $E_n^{(0)}$ for each eigenket.
Let's now suppose that we have a system, for which there is a g-fold degeneracy, hence there are g unperturbed eigenkets $|m_0\rangle$ for one single $E_D^{(0)}$ eigenvalue. Let's define the degenerate eigenspace as $D := \left\{ |m_0\rangle \in \mathbb{H} \,\middle|\, \hat{\mathcal{H}} |m_0\rangle = E_D^{(0)} |m_0\rangle, \ m = 1, \cdots, g \right\}$. In general, the perturbation breaks the degeneracy, forming a new set of eigenkets $|l\rangle$ that do not coincide with the unperturbed set $|l_0\rangle$, although we can use the following projection

$$|l_0\rangle = \sum_{|m\rangle \in D} \langle m_0|l_0\rangle |m_0\rangle$$

Let's rewrite the Schrödinger equation for the new states $|l\rangle$, and define the projections $\hat{\pi}_0 = |m_0\rangle \langle m_0|$ and its coprojection $\hat{\pi}_1 = \hat{\mathbb{1}} - \hat{\pi}_0$. The Schrödinger equation then becomes

$$\left(E - \hat{\mathcal{H}}_0 - \lambda\hat{V} \right) |l\rangle = \left(E - E_D^{(0)} - \lambda\hat{V} \right) (\hat{\pi}_0 |l\rangle + \hat{\pi}_1 |l\rangle) = 0 \qquad (6.19)$$

We separate the equation (6.19) multiplying on the left firstly by $\hat{\pi}_0$ and then by $\hat{\pi}_1$

$$\left(E - E_D^{(0)} - \lambda\hat{\pi}_0\hat{V} \right) \hat{\pi}_0 |l\rangle - \lambda\hat{\pi}_0\hat{V}\hat{\pi}_1 |l\rangle = 0$$
$$\left(E + \hat{\mathcal{H}}_0 - \lambda\hat{\pi}_1\hat{V} \right) \hat{\pi}_1 |l\rangle - \lambda\hat{\pi}_1\hat{V}\hat{\pi}_0 |l\rangle = 0$$

(6.20)

From this equation we can then solve for $\hat{\pi}_1 \left| l \right\rangle$ and $\hat{\pi}_0 \left| l \right\rangle$

$$\hat{\pi}_1 = \hat{\pi}_1 \frac{\lambda}{E - \hat{\mathcal{H}}_0 - \lambda \hat{\pi}_1 \hat{V} \hat{\pi}_1} \hat{\pi}_1 \hat{V} \hat{\pi}_0 \left| l \right\rangle$$

$$\left(E - E_D^{(0)} - \lambda \hat{\pi}_0 \hat{V} \hat{\pi}_0 - \lambda^2 \hat{\pi}_0 \hat{V} \hat{\pi}_1 \frac{1}{E - \hat{\mathcal{H}}_0 \lambda \hat{V}} \hat{\pi}_1 \hat{V} \hat{\pi}_0 \right) \hat{\pi}_0 \left| l \right\rangle = 0 \tag{6.21}$$

The general approximation to $\mathcal{O}(\lambda^n)$ will be given from the following general expression

$$\hat{\pi}_1 \left| l_1 \right\rangle = \sum_{\left| k \right\rangle \notin D} \frac{V_{kl}}{E_D^{(0)} - E_k^{(0)}} \left| k_0 \right\rangle$$

And, henceforth, in order to solve for $\mathcal{O}(\lambda)$, we get the following equation

$$\left(E - E_D^{(0)} - \lambda \hat{\pi}_0 \hat{V} \hat{\pi}_0 \right) \hat{\pi}_0 \left| l_0 \right\rangle = 0 \tag{6.22}$$

The energy shifts $\Delta^{(1)}$ will then be the diagonal elements of the perturbation $\left\langle l_0 \right| \hat{V} \left| l_0 \right\rangle$
We can immediately ask why a λ^2 appears in (6.21). This is given simply by the substitution we made in order to get the equation, but we already know that the energy shift at the first order is $E_i^{(1)} = E_D^{(0)} + \lambda v_i$, where v_i are the eigenvalues of the operator $\hat{\pi}_0 \hat{V} \hat{\pi}_0$. We assume that the degeneracy is completely resolved after the application of the perturbation, hence we get $E_i^{(1)} = \lambda(v_i - v_j) \neq 0$. Since there isn't anymore degeneration in this system, we solve using nondegenerate Rayleigh-Schrödinger perturbation theory, obtaining the corrections

$$\hat{\pi}_0 \left| l_i^1 \right\rangle = \lambda \sum_{j \neq i} \frac{\hat{\pi}_0}{v_j - v_i} \left| l_j^0 \right\rangle \left\langle l_j^0 \right| \hat{V} \hat{\pi}_1 \frac{1}{E_D^{(0)} - \hat{\mathcal{H}}_0} \hat{\pi}_1 \hat{V} \left| l_i^0 \right\rangle \tag{6.23}$$

Since $\hat{\pi}_0 \left| l_j^0 \right\rangle$ are eigenvectors of $\hat{V}$ we get that the energy shift at the second order is simply

$$\Delta_l^{(2)} = \sum_{k \notin D} \frac{\left| V_{kl} \right|^2}{E_D^{(0)} - E_k^{(0)}} \tag{6.24}$$

§ 6.2 Variational Methods

Approximating through variational methods is done when searching for approximate ground states energies E_0 when exact values aren't available.
We start to "guess" the ground state by defining a trial ket $\left| \emptyset \right\rangle$. We then define the following

$$\left\langle \hat{\mathcal{H}} \right\rangle = \frac{\left\langle \emptyset \right| \hat{\mathcal{H}} \left| \emptyset \right\rangle}{\left\langle \emptyset | \emptyset \right\rangle} \tag{6.25}$$

Theorem 6.1. *There exists an upper bound to E_0, hence*

$$\left\langle \hat{\mathcal{H}} \right\rangle \geq E_0$$

Proof. We can expand $|\emptyset\rangle$ as follows

$$|\emptyset\rangle = \sum_{k=0}^{\infty} |k\rangle \langle k|\emptyset\rangle$$

Where, $\hat{\mathcal{H}} |k\rangle = E_k |k\rangle$, hence it's an exact eigenket.
We can write $E_k = E_k - E_0 + E_0$ and evaluating $\left\langle \hat{\mathcal{H}} \right\rangle$ we have

$$\left\langle \hat{\mathcal{H}} \right\rangle = \frac{\sum_{k=0}^{\infty} |\langle k|\emptyset\rangle|^2 E_k}{\sum_{k=0}^{\infty} \left|\langle k|\emptyset\rangle^2\right|} = \frac{\sum_{k=0}^{\infty} |\langle k|\emptyset\rangle|^2 (E_k - E_0)}{\sum_{k=0}^{\infty} \left|\langle k|\emptyset\rangle^2\right|} + E_0 \geq E_0$$

Obviously, the equality is given iff $|\emptyset\rangle$ is the exact ground eigenket $\qquad\square$

This method is really powerful, since for even a poor trial ket we have $\langle k|\emptyset\rangle \sim \mathcal{O}(\epsilon)$ and $\left\langle \hat{\mathcal{H}} \right\rangle - E_0 \sim \mathcal{O}(\epsilon^2)$.
Another way to say this is saying that if we variate $|\emptyset\rangle$, the Hamiltonian will be stationary with respect to $\delta |\emptyset\rangle$.
This method doesn't say what shape does the ket $|\emptyset\rangle$ has, hence we must guess them, using the system as a guide.
Practically, it's much more useful to define a parameter vector λ_i which will appear in the considered eigenket, and then find the minimum of $\left\langle \hat{\mathcal{H}} \right\rangle$, imposing the following equation

$$\frac{\partial \left\langle \hat{\mathcal{H}} \right\rangle}{\partial \lambda_i} = 0 \tag{6.26}$$

§ 6.3 Time Dependent Perturbation Theory

§§ 6.3.1 Dirac Interaction Picture

Let's begin considering a time dependent Hamiltonian that can be split in two parts

$$\hat{\mathcal{H}}(t) = \hat{\mathcal{H}}_0 + \hat{V}(t) \tag{6.27}$$

Where one piece is time independent and the other is time dependent. We suppose that $\hat{\mathcal{H}}_0$ is exactly solvable.
Let's suppose that at $t = 0$ the state ket is given by the following relation

$$|\alpha\rangle = \sum_n c_n(0) |n\rangle \tag{6.28}$$

Where $|n\rangle$ is the eigenvalue of $\hat{\mathcal{H}}_0$.
Our objective is to find some $c_n(t)$ such that

$$|\alpha(t)\rangle = \sum_n c_n(t) e^{-\frac{iE_n t}{\hbar}} |n\rangle \tag{6.29}$$

Now, in order to simplify our problem, we define the *Dirac picture*, or *Interaction picture*, where, having considered our Hamiltonian, we have

$$|\alpha(t)\rangle_I = e^{\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} |\alpha(t)\rangle_S \tag{6.30}$$

The observables in this picture will be defined as follows, in particular, for our perturbation $\hat{V}$, we have

$$\hat{V}_I = e^{\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \hat{V} e^{-\frac{i\hat{\mathcal{H}}_0 t}{\hbar}} \tag{6.31}$$

Remembering the relation between Schrödinger and Heisenberg picture, we have that

$$|\alpha\rangle_H = e^{\frac{i\hat{\mathcal{H}} t}{\hbar}} |\alpha(t_0)\rangle_S$$
$$|\alpha(t_0)\rangle_S = e^{-\frac{i\hat{\mathcal{H}}(t-t_0)}{\hbar}} |\alpha\rangle \tag{6.32}$$
$$\hat{A}_H = e^{\frac{i\hat{\mathcal{H}} t}{\hbar}} \hat{A} e^{-\frac{i\hat{\mathcal{H}} t}{\hbar}}$$

Hence, we have that the Dirac interaction picture will satisfy the following time dependent Schrödinger equation (remembering that $i\hbar\partial_t |\alpha(t)\rangle_I = i\hbar\partial_t(\exp\left(i\hat{\mathcal{H}}_0 t/\hbar\right) |\alpha(t)\rangle_S)$

$$i\hbar\frac{\partial}{\partial t} |\alpha(t)\rangle_I = \hat{V}_I |\alpha(t)\rangle_I \tag{6.33}$$

Where the perturbation takes the place of the Hamiltonian in the time dependent equation. It is also demonstrable that, for an observable $\hat{A}$, its interaction picture will satisfy the following differential equation

$$\frac{\mathrm{d}\hat{A}_I}{\mathrm{d}t} = \frac{1}{i\hbar}\left[\hat{A}_I, \hat{\mathcal{H}}_0\right] \tag{6.34}$$

It's obvious that this Dirac Interaction picture is the halfway between a Schrödinger and a Heisenberg representation picture.
Going back to (6.29), we hae that in interaction picture it will simply become this

$$|\alpha(t)\rangle_I = \sum_n c_n(t) |n\rangle \tag{6.35}$$

We can now write a differential equation for $c_n(t)$.

$$i\hbar\frac{\partial}{\partial t} \langle n|\alpha(t)\rangle_I = \sum_m \langle n| \hat{V}_I |m\rangle \langle m|\alpha(t)\rangle_I \tag{6.36}$$

By definition, we have $c_n(t) = \langle n | \alpha(t) \rangle_I$, henceforth

$$i\hbar \frac{\mathrm{d}c_n}{\mathrm{d}t} = \sum_m V_{nm} e^{i\omega_{nm}t} c_m(t) \tag{6.37}$$

Where we have expanded the interaction-picture time dependence, and we have by definition of frequency

$$\omega_{nm} = \frac{E_n - E_m}{\hbar} = -\omega_{mn} \tag{6.38}$$

§§ 6.3.2 Dyson Series

Usually, exact solutions for $c_n(t)$ are not available, hence we must find a suitable approximation for our solution. One good way to start is to suppose that $c_n(t)$ can be expressed via the sum of different functions as follows

$$c_n(t) = \sum_{i=0}^{\infty} c_n^{(i)}$$

Where $c_n^{(i)}$ indicates the $i-$th transition amplitude.

This problem can be attacked, using that $c_n^{(0)} = \delta_{in}$ and then using it to define a differential equation for $c_n^{(1)}(t)$ and so on. Using operator theory, this problem can be solved even in a better way.

The time evolution operator in the Dirac picture is defined as follows

$$|\alpha(t)\rangle = \hat{\mathcal{U}}_I(t) |\alpha\rangle$$

For which, we know already that it's solution to the following ODE

$$\begin{cases} i\hbar \dfrac{\mathrm{d}\hat{\mathcal{U}}_I}{\mathrm{d}t} = \hat{V}_I \hat{\mathcal{U}}_I \\ \hat{\mathcal{U}}(t_0) = \hat{\mathbb{1}} \end{cases}$$

This differential equation is equivalent to the following integral equation

$$\hat{\mathcal{U}}_I(t) = \hat{\mathbb{1}} - \frac{i}{\hbar} \int_{t_0}^{t} \hat{V}_I(t_1) \hat{\mathcal{U}}_I(t_1) \, \mathrm{d}t_1$$

Iterating, we get

$$\hat{\mathcal{U}}_I(t) = \sum_{n=0}^{\infty} \prod_{k=1}^{\infty} \left(\frac{-i\hat{\mathbb{1}}}{\hbar} \right)^n \int_{t_0}^{t_{k-1}} V_I(t_k) \, \mathrm{d}t_k \tag{6.39}$$

Which is equivalent to writing the following expression

$$\hat{\mathcal{U}}_I(t) \approx \hat{\mathbb{1}} - \frac{i}{\hbar} \int_{t_0}^t V_I(t_1) \, \mathrm{d}t_1 \approx$$

$$\hat{\mathbb{1}} - \left(\frac{-i}{\hbar}\right)^2 \int_{t_0}^t V_I(t_1)V_I(t_2) \, \mathrm{d}t_1\mathrm{d}t_2 + \cdots$$

$$\cdots + \left(\frac{-i}{\hbar}\right)^n \int_{t_0}^t \int_{t_0}^{t_1} \cdots \int_{t_0}^{t_{n-1}} V_I(t_1) \cdots V_I(t_n) \, \mathrm{d}t_1 \cdots \, \mathrm{d}t_n$$

Through this approximation, it's virtually possible to compute $\hat{\mathcal{U}}_I(t)$ to any order, and therefore $c_n(t)$. This kind of computation is fundamental in fields like atomic physics, as we will see later.

7 Identical Particles

While in classical physics, two identical particles can be distinguished, in quantum mechanics they're truly indistinguishable. Let's suppose that we have two particles, for which you have a configuration (Hilbert) space $\mathbb{H}_1$ for the first particle, and $\mathbb{H}_2$ for the second particle. The general state of the system will then be described by a ket in $\mathbb{H}_1 \otimes \mathbb{H}_2$. Hence, labeling particle 1 as α and the second as β, we then will write that a state $|s\rangle$ can be written as

$$|s\rangle = |\alpha\rangle \otimes |\beta\rangle = |\alpha\rangle \, |\beta\rangle \tag{7.1}$$

It's obvious that, since the particles in study are impossible to distinguish, that the state $|s\rangle$ can also be written also as $|\beta\rangle \otimes |\alpha\rangle$, hence, for the principle of quantum superposition, we must have that the most general state will be the following, for a two particle system

$$|s\rangle = c_1 \, |\alpha\rangle \otimes |\beta\rangle + c_2 \, |\beta\rangle \otimes |\alpha\rangle \tag{7.2}$$

This definition brings us what's known as *exchange degeneracy*. This degeneracy brings us a huge problem, since in this case, the eigenvalue of the complete basis doesn't completely define the state ket.

Before diving into the nature of exchange degeneracy, we define a new operator, called *exchange operator*, or just $\hat{P}_{ij}$. It will act as follows:

Let $|a_i\rangle \in \mathbb{H}_1$ and $|a_j\rangle \in \mathbb{H}_2$, and consider the new state $|a_i\rangle \otimes |a_j\rangle \in \mathbb{H}_1 \otimes \mathbb{H}_2$. We will have

$$\hat{P}_{ij} \, |a_i\rangle \otimes |a_j\rangle = \lambda \, |a_j\rangle \otimes |a_i\rangle$$
$$\hat{P}_{ij} = \hat{P}_{ji}$$
$$\hat{P}_{ij}^2 = \hat{\mathbb{1}} \rightarrow \lambda = \pm 1$$

In general, if we have an observable $\hat{a}$, such that

$$\hat{a}_i \, |a_i\rangle = a \, |a_i\rangle$$
$$\hat{a}_j \, |a_j\rangle = b \, |a_j\rangle$$

We get, after applying an exchange transformation

$$\hat{P}_{ij} \hat{a}_i \hat{P}_{ij}^{-1} \, |a_i\rangle \, |a_j\rangle = a \, |a_i\rangle \, |a_j\rangle$$
$$\hat{P}_{ij} \hat{a}_i \hat{P}_{ij}^{-1} \, |a_j\rangle \, |a_i\rangle = a \, |a_j\rangle \, |a_i\rangle$$

This is valid only if $\hat{P}_{ij}\hat{a}_i\hat{P}_{ij}^{-1} = \hat{a}_j$, hence, this exchange operator, applied on a system observable, changes its label, hence basically in which space of the two of the tensor space $\mathbb{H}_1 \otimes \mathbb{H}_2$ the operator $\hat{a}_i$ will act.

Let's now consider a general two-particle Hamiltonian. It will be the following

$$\hat{\mathcal{H}} = \frac{\hat{p}_1^2}{2m} + \frac{\hat{p}_2^2}{2m} + V\left(\left|x_1^i - x_2^i\right|\right) + V_e(x_1^i) + V_e(x_2^i) \tag{7.3}$$

This Hamiltonian is obviously invariant to exchange of particles, hence $\left[\hat{\mathcal{H}}, \hat{P}_{12}\right] = 0$ and $\hat{P}_{12}$ is a constant of motion

If we call the Hamiltonian's eigenket $|a_1\rangle |a_2\rangle$, we can select two main common basis eigenkets as follows

$$|s\rangle = \frac{1}{\sqrt{2}}\left(|a_1\rangle |a_2\rangle + |a_2\rangle |a_1\rangle\right)$$

$$|a\rangle = \frac{1}{\sqrt{2}}\left(|a_1\rangle |a_2\rangle - |a_2\rangle |a_1\rangle\right)$$

Where they are tied through two operators, the *symmetrization* operator and the *antisymmetrization* operator, defined as follows

$$\hat{T}_+ = \frac{1}{2}\left(\hat{\mathbb{1}} + \hat{P}_{12}\right)$$
$$\hat{T}_- = \frac{1}{2}\left(\hat{\mathbb{1}} - \hat{P}_{12}\right) \tag{7.4}$$

Hence, applied to a ket $|a_1\rangle \otimes |a_2\rangle$, we have

$$\hat{T}_\pm\left(c_1 |a_1\rangle \otimes |a_2\rangle + c_2 |a_2\rangle \otimes |a_1\rangle\right) = \frac{c_1 \pm c_2}{2}\left(|a_1\rangle \otimes |a_2\rangle \pm |a_2\rangle \otimes |a_1\rangle\right)$$

This finally gives the final symmetry of the system.

§ 7.1 Symmetrization Postulate

We will now delve shortly into quantum statistical mechanics. Here we have two statistics, *Fermi-Dirac statistics* and *Bose-Einstein statistics*. Particles that satisfy Fermi-Dirac statistics are said to be *fermions* and those who satisfy Bose-Einstein statistics are said to be *bosons*. Under exchange of two particles, we have that, if we indicate with $|b\rangle$ bosons and with $|f\rangle$ fermions, that for a system of N identical particles

$$\hat{P}_{ij}\bigotimes_{i=1}^{N}|b\rangle_i = \hat{P}_{ij}|B\rangle_i = |B\rangle_j = \bigotimes_{j=1}^{N}|b\rangle_j$$
$$\hat{P}_{ij}\bigotimes_{i=1}^{N}|f\rangle_i = \hat{P}_{ij}|F\rangle_i = -|F\rangle_j = -\bigotimes_{j=1}^{N}|f\rangle_j \tag{7.5}$$

This change of sign is dependent from the spin-wavefunction, determining that antisymmetric particle wavefunctions have half-integer spin, and symmetric particle wavefunctions have integer spin.

Empirically for fermions (half-integer spin particles), it's known that they must obey the *Pauli exclusion principle*, which states that two identical fermions cannot share the same quantum state.

For only two fermions, if we want to write the ground state wavefunction, we know that due to its antisymmetry, it must be the following

$$|GS\rangle_f = \frac{1}{\sqrt{2}}\left(|f_1\rangle|f_2\rangle - |f_2\rangle|f_1\rangle\right) \tag{7.6}$$

This is the only possible configuration. For bosons, instead, we have three possible configurations

$$|GS\rangle_b = |b_1\rangle|b_1\rangle, \quad |b_2\rangle|b_2\rangle, \quad \frac{1}{\sqrt{2}}\left(|b_1\rangle|b_2\rangle + |b_2\rangle|b_1\rangle\right) \tag{7.7}$$

§§ 7.1.1 Two Electron System

The most simple system composed by two fermions is the two-electron system. Since it's fermionic, we already know that the eigenvalue of the exchange operator must be -1.

Let's say that our base kets are specified by $|i, m_{s_i}\rangle$ where $i = 1, 2$ indicates the electron and m_{s_i} indicates the particle's spin magnetic quantum number. The most general state will then be given by a linear combination of these basis kets as follow

$$|\psi\rangle = \sum_{m_{s_1}}\sum_{m_{s_2}}|s_1, s_2, m_{s_1}, m_{s_2}\rangle\langle s_1, s_2, m_{s_1}, m_{s_2}|\psi\rangle$$

Or, in terms of wavefunctions

$$\psi_{jm}(x_1^i, x_2^i) = \sum_{m_{s_2}}\sum_{m_{s_1}}C(m_{s_1}, m_{s_2})\psi_{m_{s_1}m_{s_2}}(x_1^i, x_2^i)$$

Where with $C(m_{s_1}, m_{s_2})$ we indicated the Clebsch-Gordan coefficients for the sum of two spin 1/2 systems.

Analogously, if $\left[\hat{\mathcal{H}}, \hat{S}_{tot}^2\right] = 0$, we have that the eigenvalues (and hence eigenfunctions) of the system will be given by the tensor product $|E\rangle \otimes |sm\rangle$. Since the wavefunction associated with $|sm\rangle$ is a spinor, we will have that our wavefunction will be given by

$$\psi_{jm}(x_1^i, x_2^i) = \phi(x_1^i, x_2^i)\chi_\pm$$

With $\chi_\pm$ as our basis spinor.

Due to the properties requested by the fermion statistics, we must have that, if $\chi_\pm = |\pm\rangle$,

it «must» be one of these four

$$|\pm\rangle = \begin{cases} |+\rangle\,|+\rangle \\ \dfrac{1}{\sqrt{2}}\left(|+\rangle\,|-\rangle + |-\rangle\,|+\rangle\right) \\ |-\rangle\,|-\rangle \\ \dfrac{1}{\sqrt{2}}\left(|+\rangle\,|-\rangle - |-\rangle\,|+\rangle\right) \end{cases} \tag{7.8}$$

Applying our exchange operator we have that the first three are symmetric, which are commonly called triplet states, and the last one is antisymmetric with respect to exchange of particles, and it's called a singlet state.

Another particular relation of the particle exchange operator is obvious if we see how it acts on our kets.

We have that $\langle s_1 s_2 m_{s_1} m_{s_2}|\,\hat{P}_{12}\,|\alpha\rangle = \langle s_2 s_1 m_{s_1} m_{s_2}|\alpha\rangle$, and we also know that $\langle s_1 s_2 m_{s_1} m_{s_2}|\alpha\rangle = -\langle s_2 s_1 m_{s_2} m_{s_1}|\alpha\rangle$ from the Fermi-Dirac statistic, followed by electrons.

Hence, a full exchange operator $\hat{P}_{12}$ will be given instead by the tensor product of the spatial particle exchange operator and the spin exchange operator, as

$$\hat{P}_{12} = \hat{P}_{12}^p \otimes \hat{P}_{12}^s$$

A thorough application of this theory for $s = 1/2$ systems will be treated in a further section where it will be studied together with atomic physics.

§ 7.2 Multiparticle States

As we've already seen previously, multiparticle states can be defined as a multiple tensor product of single particle states. As we've already seen, the particle exchange operator is idempotent, i.e. $\hat{P}_{ij}^2 = \hat{\mathbb{1}}$, thus the possible eigenvalues are ± 1. It must be noted tho, that in general

$$\left[\hat{P}_{ij}, \hat{P}_{kl}\right] = \hat{P}_{ij}\hat{P}_{kl} - \hat{P}_{kl}\hat{P}_{ij} \neq 0 \tag{7.9}$$

Let's now consider a 3 particle state. We have that there are $3!$ possible combinations of the single particle states $|p_1\rangle\,|p_2\rangle\,|p_3\rangle$. If we insist on our symmetrization postulate, we have that we can either have a single completely antisymmetric state or a single fully symmetric state. This states must hence be a linear combination of 6 equally probable states, formed by the tensor product of the different particles. This state is an eigenstate for $\hat{P}_{12}, \hat{P}_{23}, \hat{P}_{13}$. Defining a new exchange operator $\hat{P}_{123} = \hat{P}_{12} \otimes \hat{P}_{13}$, we have that a completely symmetrical state can be written, (remember that if two indices are equal then there can't be a completely antisymmetric state)

Part II

Statistical Mechanics

8 Thermodynamic Systems

§ 8.1 Temperature

§§ 8.1.1 Macroscopic and Microscopic Systems

The study of any branch of natural science *must* begin with the definition of *system*

Definition 8.1.1 (System). A *system* is a finite region of space containing matter inside a closed surface, known as the *wall*.
Everything outside of the system, even other systems, that are in interaction with the system are known as the *surroundings* of the system.
A system is said to be *closed* if there's no matter flux between itself and its surroundings.
The set of the system, surroundings and all the rest is known as *universe*, and it's usually indicated with Ω.

All systems can be studied with two points of view

1. A microscopic point of view (molecular or less)

2. A macroscopic point of view (human scale or more)

Taking as our example system the cylinder chamber of a car, we can define the following *macroscopic coordinates*, i.e. descriptors of the system in a macroscopic point of view

1. Mass of gas contained

2. Volume of the chamber

3. Pressure of the gas

4. Temperature of the gas

These coordinates are *macroscopic* also because

- Do not involve assumptions about the structure of matter, fields or radiation

- Are low in number

- Are fundamental

- Can be, generally, directly measured

If we instead consider a system from the microscopic point of view, we can define the system as

1. N particles with energy states E_i

2. Particle interactions with fields and through collisions

And more.

Microscopic systems that can be considered isolated or embedded in other systems are known as *ensemble systems*. In microscopic systems, the equilibrium state is defined as the state with the highest probability, i.e. the state which will have a higher occupation number or *population*.

In general, microscopic coordinates

1. Consider the structure of matter, fields and radiation

2. Are many in number

3. Are described by mathematical models and usually not directly measurable

4. Must be calculated using the previous models

In the study of thermodynamics, in the next chapter or two, we will use the macroscopic description. The major difference with the other branches of science lays in the fact that in thermodynamics, a macroscopic quantity is always present and defined, known as *temperature*.

Generally, in thermodynamics the quantities chosen are known as *thermodynamic coordinates*, which are macroscopic coordinates that determine the internal state of the system. Systems for which thermodynamic coordinates can be defined are known as *thermodynamic systems*

§§ 8.1.2 Zeroth Law of Thermodynamics

Consider a thermodynamic system A, for which we can define two independent coordinates, (X, Y), the first being a generalized force and the other being a generalized displacement. We define:

Definition 8.1.2 (Equilibrium State). A state for which the coordinates (X, Y) are constant as long as the external conditions don't change, is known as an *equilibrium state*

Equilibrium states *depend on proximity of other systems and nature of the boundary*, if we put A in contact with another system B with coordinates (X', Y') we then can define two types of walls:

Definition 8.1.3 (Adiabatic Walls). If the walls between the system A and B are *adiabatic*, then the equilibrium states are *independent* and possible for each value of (X, Y) and (X', Y')

Definition 8.1.4 (Diathermal Walls). If the walls between the system A and B are *diathermal*, the equilibrium states of the two systems aren't independent anymore, and thus are defined only for a set of coordinates (X, Y, X', Y')

We subsequently define

Definition 8.1.5 (Thermal Equilibrium). Thermal equilibrium is defined as the state achieved by two or more systems, characterized by a restricted amount of values of system coordinates, after being put in contact through a diathermal wall.

From the definition of thermal equilibrium, an important law follows

Law 1 (0th Law of Thermodynamics). Suppose that two thermodynamic systems A and B are separated by an adiabatic wall and simultaneously in contact with a third system C through a diathermal wall. It follows that if:

- A is in thermal equilibrium with C

- B is in thermal equilibrium with C

Then, A must be in thermal equilibrium with B

Suppose that now that the system A and B are at equilibrium with each other at some coordinates (X_A, Y_A) and (X_B, Y_B). If we remove the system A, the system B will undergo a change of state to coordinates (X_2, Y_2) which must be in thermal equilibrium with the state (X_B, Y_B). It must follow then that there's a quantity, known as *temperature*, which remains constant during this transformation, thus:

Definition 8.1.6 (Temperature). We define the *temperature* as the property in common between states in thermal equilibrium. A change of state with constant temperature (i.e. in *thermal equilibrium*) is known as an *isothermal transformation* or an *isothermal process*. Temperature *must* be a scalar quantity, and it's usually indicated with T. For each possible value of T there exists a defined family of isothermal processes.

§§ 8.1.3 Temperature Measurements

In order to define a *temperature scale*, we choose a thermodynamic system for which are known its properties, known as the *thermometer*, and define a set of empirical rules for assigning a value of T for each isotherm.
Suppose that the system is well described by the generalized force X and the generalized displacement Y. Then, for defining a temperature scale we will

- Choose a convenient path in the (X, Y) plane, like $Y = const$. Then, since one of the two quantities is fixed, we must have that if the system undergoes an isothermal process, we must have that
$$\theta = \theta(X)$$
Where θ is our temperature scale

- We suppose, arbitrarily, that $\theta(X) \propto X$, thus

$$\theta_A(X) = aX$$

The scale that we defined previously has, in particular, that

$$\lim_{X \to 0} \theta_A(X) = 0$$

I.e. it's an *absolute* temperature scale. Examples of absolute scales are the *Kelvin* and *Rankine* temperature scales.

Experimentally, the standard gas for thermometric evaluations is molecular hydrogen H_2. For most thermodynamic scales, being $Y = const$ completely arbitrary, it's convenient to define $Y = Y_1 = const$ as the *triple point* of water, it being the point in which liquid water, ice and vapor exist in the same place and time. This point is measured to be at a temperature of

$$T_{TP} = 0.01°C = 273.16K$$

From the absolute arbitrary temperature scale we have defined before, defined X_{TP} the coordinate at which we have the triple point, we have

$$\theta_A(X_{TP}) = aX_{TP} = 273.16K \Longrightarrow a = \frac{273.16K}{X_{TP}}$$

Thus, in general, an absolute thermometric scale can be defined as follows

$$\theta_A(X) = 273.16 \frac{X}{X_{TP}} K \tag{8.1}$$

Thus, it's possible to define this scale in terms of pressures (generalized forces) or volumes (generalized displacements), using the triple point coordinates as reference point.

§§§ 8.1.3.1 Temperature Scales

The most common temperature scales are two in the metric system of units and two in the imperial system. If we use the definition (8.1) for the Kelvin scale, we have

Definition 8.1.7 (Celsius Temperature Scale). Defined by Anders Celsius (SWE), used in the metric system of units

$$T(°C) = T(K) - 273.15$$

Definition 8.1.8 (Rankine Temperature Scale). Absolute temperature scale defined by William Rankine (UK), used in the imperial system of units

$$T(R) = \frac{9}{5}T(K)$$

Definition 8.1.9 (Fahrenheit Temperature Scale)**.** Temperature scale define by Daniel G. Fahrenheit (GER), used in the imperial system of units

$$T\left(°F\right) = \frac{9}{5}T\left(R\right) - 459.67$$

With these definitions we can find conversions between these scales, and for commonly used scales like Fahrenheit and Celsius we get the following conversion formula

$$T\left(°F\right) = \frac{9}{5}R\left(°C\right) + 32 \tag{8.2}$$

Note that the Celsius temperature scale has the same dimensions of intervals of the Kelvin scale, thus, a temperature difference in Celsius degrees is the same in Kelvins. Thanks to future concepts, the Kelvin scale will be defined as the absolute scale of temperature, since it's tied to energetic properties of the system itself.

§ 8.2 Thermodynamic Equilibrium

§§ 8.2.1 Definition of Equilibrium

Given a thermodynamic system A, *any* change of coordinates defines a *change of state* for the system.
A non-influenced system is known as an *isolated system*, but these kinds of systems aren't important in the study of classical thermodynamics, since these systems can't be studied macroscopically.
There are various kinds of equilibrium, namely

- Mechanical equilibrium: equilibrium of forces in the system

- Chemical equilibrium: the system is in mechanical equilibrium and it doesn't undergo in spontaneous chemical reactions

- Thermal equilibrium: the system is in mechanical and chemical equilibrium, and there is no change of coordinates when the system is separated from its surroundings via a diathermal wall

- Thermodynamic equilibrium: the system is in mechanical, chemical, thermal equilibrium contemporaneously

Note that for having thermodynamic equilibrium, all the previously stated equilibriums must be satisfied, and the system doesn't undergo in changes of state.
In fact, if a system is not in mechanical equilibrium then the system is in a *non equilibrium state* and therefore thermodynamic coordinates cannot be defined, and so goes for the other equilibriums. Specifically, when the system is in *thermodynamic* equilibrium, it *does not* change state, note instead how instead *thermal* equilibrium is defined via changes of state, since in order for a system to be defined in thermal equilibrium undergoes an *isothermal* change of state, but still a change of state.

§§ 8.2.2 Thermodynamic Relationships at Equilibrium

Given any mole of gas, it's experimentally verifiable that if:

- Fixed volumes and temperature imply that the pressure can't be chosen

- Fixed pressure and temperature imply that the volume can't be chosen

- Fixed pressure and volume imply that the temperature can't be chosen

Clearly, if we use these three coordinates to describe a thermodynamic system, only one of the two can be chosen.
These relations were found empirically by Gay-Lussac and Boyle.
The first experiment, done by Gay-Lussac deals with the relationship between V and T. It has been found that, heating a solid with linear length l_0 at rest, that

$$l(T) = l_0 + \sum_{i=1}^{\infty} a_i l_0 \Delta T^i \tag{8.3}$$

With a_i being numerical coefficients depending on the material composition of the system. For small temperature variations ΔT, it can be said that

$$l(T) \simeq l_0 + a l_0 \Delta T \tag{8.4}$$

Here a, is known as the *linear thermal dilatation coefficient*. In general $a > 0$, but in nature materials with $a < 0$ are found, one of these is water, and all oxides.
Keeping the first order approximation, and using $V_0 = l_0^3$, we can say

$$V(T) \simeq l_0^3 \left(1 + 3a\Delta T\right) = V_0 \left(1 + \beta \Delta T\right) \tag{8.5}$$

The coefficient $\beta = 3a$ is the *volumetric thermal dilatation coefficient*. This experiment, when repeated with fluids, thanks to their incompressibility, gives

Theorem 8.1 (First Law of Gay-Lussac). *Given a fluid with volumetric compressibility β, contained in a volume V_0, when undergoing a change of temperature ΔT expands or contracts following this equation*

$$V(T) = V_0 \left(1 + \beta \Delta T\right) \tag{8.6}$$

If we define the absolute zero of the kelvin scale as the temperature such that $V(T_0) = 0$, we get

$$V(T) = V_0 \frac{T}{T_0} \tag{8.7}$$

If the volume is kept constant and instead the variation of pressure is measured, we have

Theorem 8.2 (Second Law of Gay-Lussac). *Given a fluid with compressibility β, if it undergoes a change of temperature at constant volume, the pressure will follow this equation*

$$p\left(T\right) = p_0\left(1 + \beta\Delta T\right) \tag{8.8}$$

If we use the kelvin scale as defined before, we get

$$p\left(T\right) = p_0\frac{T}{T_0} \tag{8.9}$$

Both these laws were derived empirically through experimentation, and give the behavior of volume and pressure with respect to changes of temperature. One might ask what happens when temperature is constant, and that's what has been found by Boyle

Theorem 8.3 (Boyle's Law). *Given a fluid undergoing pressure and volume changes in thermal equilibrium, then*

$$pV = p_0V_0 \tag{8.10}$$

Note that if a gas verifies all the previous laws, we have

$$pV = \frac{p_0V_0}{T_0}T \tag{8.11}$$

This equation gives the empirical dependencies of the three thermodynamic coordinates, (p, V, T) and is for this known as the *equation of state*.
If we also consider that the volume of the gas is $V \propto n$, where n are the *moles* of gas, one has

$$pV = nRT \tag{8.12}$$

Where $R = 8.31\text{J/Kmol}$ is a conversion constant known as the *gas constant*.
All previous experiments considered that we specifically had what's called a *hydrostatic system*, i.e. a system which has

- Uniform pressure

- Constant mass

- Surface, gravitational and electromagnetic effects can be considered negligible

§§ 8.2.3 Thermodynamic Diagrams

The previous relations can be used to determine what are known as *thermodynamic diagrams*, which permit the graphical description of changes of state in a thermodynamic system.
As we saw, of the three coordinates (p, V, T) only two are independent. Thanks to this, we can define 2D Cartesian planes which have as coordinates either pV, pT, VT. We could also define a surface, known as the pVT surface, which describes *all* the possible configurations of the system, "extruding" the pV, pT and VT

§§§ 8.2.3.1 pV Diagrams

Consider now an empirical case of water in thermal equilibrium at $94°C$ in a vessel of $2m^3$. If the vessel is sealed and there's *no air*, it's seen that water is in equilibrium with its vapor. It's notable how, if we map all the possible pV transformations of water, the diagram will be divided in three zones by what are known as *critical isotherms*.
These isotherms all coincide in a special point, known as the *critical point*. An image of such diagram follows

§§§ 8.2.3.2 pT Diagrams

Pressure-temperature diagrams, are instead useful for defining phase transition isotherms of the system. Considering water again, three main curves can be defined between the states of matter. These curves are known as

- Fusion curve

- Sublimation curve

- Vaporization curve

In the first one, along the curve, the set of states described are states of solid-liquid equilibrium, in the second curve solid-vapor equilibrium and liquid-vapor equilibrium.
At the intersection of these three lines we find the *triple point*. It's important to remember that this point is represented only in this diagram as a point, while in others is instead a curve.

§§ 8.2.4 Differential Changes of State

In order to properly define a mathematical framework for describing the thermodynamic of substances, we have to rewrite differential calculus in a way which is physically significant. Suppose that in a pV diagram, a substance undergoes a really small transition to another equilibrium state. If the volume changes by a differential $\mathrm{d}V$ and the pressure by a differential $\mathrm{d}p$, we *must* have

- For any volume V, we have $\mathrm{d}V << V$, but $\mathrm{d}V$ is big enough to be considered macroscopic

- For any pressure p, and consequent molecular perturbations to pressure δp_{mol}, we must have
$$\delta p_{mol} << \mathrm{d}p << p$$

With these considerations, both volume and pressure can be considered as mathematically continuous functions between the two equilibrium states.

Remembering that of (p, V, T) only two of the three are independent, we can define the differentials of these quantities, since

$$\begin{cases} p = p(V, T) \\ V = V(p, T) \\ T = T(p, V) \end{cases}$$

Thus

$$\begin{cases} \mathrm{d}p = \left(\dfrac{\partial p}{\partial V}\right)_T \mathrm{d}V + \left(\dfrac{\partial p}{\partial T}\right)_V \mathrm{d}T \\ \mathrm{d}V = \left(\dfrac{\partial V}{\partial p}\right)_T \mathrm{d}p + \left(\dfrac{\partial V}{\partial T}\right)_p \mathrm{d}T \\ \mathrm{d}T = \left(\dfrac{\partial T}{\partial p}\right)_V \mathrm{d}p + \left(\dfrac{\partial T}{\partial V}\right)_p \mathrm{d}V \end{cases} \tag{8.13}$$

Where, as a subscript of the parentheses, we indicated which of the coordinates is kept constant.

For volume specifically, we can define two things in particular

1. Volumetric expansivity β as

$$\beta = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p \tag{8.14}$$

2. Isothermal compressibility κ

$$\kappa = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T \tag{8.15}$$

Therefore, the differential changes of volume can be described as follows

$$\mathrm{d}\log(V) = \beta\mathrm{d}T - \kappa\mathrm{d}p \tag{8.16}$$

§ 8.3 Work

§§ 8.3.1 Hydrostatic Systems

<++>

§§ 8.3.2 Quasi-static Processes

9 Heat

<++>

§ 9.1 Adiabatic Processes

§§ 9.1.1 Internal Energy Function

§ 9.2 First Law of Thermodynamics

§§ 9.2.1 Concept of Heat

§ 9.3 Calorimetry

§§ 9.3.1 Heat Capacity

§§ 9.3.2 Calories

§ 9.4 Heat Flow

§§ 9.4.1 Heat Conduction

§ 9.5 Ideal Gases

§§ 9.5.1 Equation of State of an Ideal Gas

§§ 9.5.2 Internal Energy of an Ideal Gas

§§ 9.5.3 Heat Capacity of an Ideal Gas

§§ 9.5.4 Quasi-static Adiabatic Processes

10 Thermal Engines and Entropy

<++>

§ 10.1 Heat Conversion

§ 10.2 Engines

§§ 10.2.1 Stirling Engine

§§ 10.2.2 Internal Combustion Engines

§§§ 10.2.2.1 Otto Engine

§§§ 10.2.2.2 Diesel Engine

§ 10.3 Second Law of Thermodynamics

§§ 10.3.1 Kelvin-Plank Statement

§§§ 10.3.1.1 Refrigerators

§§ 10.3.2 Clausius Statement

§§§ 10.3.2.1 Equivalence of Kelvin-Plank and Clausius Statements

§ 10.4 Reversible Processes

§§ 10.4.1 Carnot Cycles

§§ 10.4.2 Carnot Theorem

§§ 10.4.3 Absolute Temperature

§ 10.5 Entropy

§§ 10.5.1 Integrability of Heat

§§ 10.5.2 Entropy Function

§§§ 10.5.2.1 TS Diagrams

§§ 10.5.3 Irreversible Processes

§§ 10.5.4 Principle of Increase of Entropy

11 Phase Transitions

<++>

§ 11.1 Thermodynamic Potentials

§§ 11.1.1 Maxwell Relations

§ 11.2 $T\mathrm{d}S$ Equations

§§ 11.2.1 Application of $T\mathrm{d}S$ Equations

§§ 11.2.2 Internal Energy Equations

§§ 11.2.3 Heat Capacity Equations

§ 11.3 Phase Transitions

§§ 11.3.1 First Order Transitions

§§ 11.3.2 Second Order Transitions

12 Introduction to Thermostatistics

Statistical Mechanics deals with many-body physical problems, and its main roots derive from thermodynamics.

There are three main concepts in mechanical statistics: *microstates*, *macrostates* and *statistical ensembles*

The microstate is defined by the coordinates and momenta of the system (in classical mechanics) and by the wavefunction of the system (in quantum mechanics). The collection of all microstate is referred as statistical ensemble.

§ 12.1 Probability Theory

Before diving deeply into statistical mechanics we need to grasp a basis of probability theory.

Definition 12.1.1 (Random Variable). A random variable is a quantity X which takes a value x depending upon the elements $e \in E$ called events.

In each observation X is uncertain, and only the probability of the occurrence of a given result x_i.

Definition 12.1.2 (Probability Density Function). Let X be a random variable, we define the probability density function $w(x)$ as the probability that X assumes a value x in a certain interval of values given by the following integral

$$\int_a^b w(x) \, \mathrm{d}x$$

By definition $w(x)$ is normalized.

$$\int_{-\infty}^{\infty} w(x) \, \mathrm{d}x = 1$$

Definition 12.1.3 (Mean Value). Let $w(x)$ be the probability density of a random variable X, the expectation or mean value is defined as the integral

$$\mathbb{E}(X) = \langle X \rangle = \int_{\mathbb{R}} x w(x) \, \mathrm{d}x$$

Defining a *random function* as the function of a random variable $F(X)$, one can also define

$$\langle F(X) \rangle = \int_{\mathbb{R}} F(x) w(x) \, \mathrm{d}x$$

Definition 12.1.4 (Moment). We define the $n-$th moment of a random value as follows

$$\mu_n = \langle X^n \rangle$$

Definition 12.1.5 (Characteristic Function). The characteristic function of a random variable is defined as the Fourier transform of the probability density function

$$\chi(k) = \hat{\mathcal{F}}[w](k) = \frac{1}{2\pi} \int_{\mathbb{R}} w(x) e^{-ikx} \, \mathrm{d}x = \left\langle e^{-ikX} \right\rangle$$

If all moments are well defined, we can write the characteristic function as a power series of moments

$$\chi(k) = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \langle X^n \rangle$$

For a multidimensional random function $F(\underline{X})$ which can take values f corresponding to a density function $w_F(f)$. $w_F(f)$ is defined as follows

$$w_F(f) = \langle \delta \left(F(\underline{X}) - f \right) \rangle$$

One important theorem that should be defined is the *central limit theorem*:

Theorem 12.1 (Central Limit). *Let $X_1, \ldots, X_n$ be a set of independent random variables with probability distributions $w(x_i)$, suppose that there exist the n-th moment of every variable, and define a second random variable Y as follows*

$$Y = \sum_{i=1}^{n} X_i$$

Then the probability density function of Y will be a Gaussian distribution, when $n \to \infty$

Proof. Let's define a third random variable Z as follows

$$Z = \sum_{i=1}^{n} \frac{1}{\sqrt{n}} (X_i - \langle X \rangle) = \frac{1}{\sqrt{n}} (Y - n \langle X \rangle)$$

Where by definition we have $\langle X_1 \rangle = \langle X_2 \rangle = \cdots = \langle X_n \rangle$.
The probability density function of Z is then defined as follows

$$w_Z(z) = \int \prod_{i=1}^{n} w(x_i) \delta \left(z - \frac{1}{\sqrt{n}} \left(\sum_{k=1}^{n} x_k \right) + \sqrt{n} \langle x \rangle \right) \, \mathrm{d}x_i =$$

$$= \frac{1}{2\pi} \int e^{ikz} \, \mathrm{d}k \int \prod_{i=1}^{n} w(x_i) e^{\frac{-ik \sum_{k=1}^{n} x_i}{\sqrt{n}} + ik\sqrt{n}\langle X \rangle} \, \mathrm{d}x_i =$$

$$= \frac{1}{2\pi} \int e^{ikz + ik\sqrt{n}\langle X \rangle} \left(\chi \left(\frac{k}{\sqrt{n}} \right) \right)^N$$

Where $\chi(q)$ is the characteristic function of $w(x_i)$. We can then reformulate the probability density by defining in an alternative way the characteristic function, as follows

$$\chi(q) = \exp \left(\sum_{n=1}^{\infty} \frac{(-iq)^n}{n!} C_n \right)$$

Where C_n are called the *cumulants*. They're related to the momenta of the variable, and it can be seen by comparing the Taylor series expansion of $\chi(q)$ with the previous equation. Reinserting it all into the definition of $w_Z(z)$, we get

$$w_Z(z) = \frac{1}{2\pi} \int e^{ikz - \frac{1}{2}k^2 \mu_2^2 + \ldots + k^3 \sqrt{n} + \cdots} \, \mathrm{d}k$$

Neglecting the terms that vanish for $\mathcal{O}\left(\sqrt{N} \right)$, we obtain the following

$$w_Z(z) = \frac{1}{\sqrt{2\pi \mu_2^2}} e^{-\frac{z^2}{2\mu_2^2}}$$

And since $w_Y(y) \, \mathrm{d}y = w_Z \, \mathrm{d}z$ we get, substituting the definition of z

$$w_Y(y) = \frac{1}{\sqrt{2\pi n \mu_2^2}} e^{-\frac{(y - n\langle X \rangle)^2}{2n\mu_2^2}}$$

Which is our searched Gaussian probability density for large n, which demonstrates the theorem. □

§ 12.2 Quantum Statistics

For dealing with quantum statistics we need to define a deeper mathematical apparatus. We already know that in quantum mechanics, the expected value of an observable $\hat{A}$ is defined as follows

$$\langle A \rangle = \langle \psi | A | \psi \rangle$$

And the trace of this operator as follows

$$\text{Tr}\hat{A} = \sum_n \langle n| \hat{A} |n\rangle \tag{12.1}$$

From the previous two definitions, we can then define a new operator, called the *density operator*

$$\hat{\rho} = |\psi\rangle \langle\psi| \tag{12.2}$$

Which can be used to redefine the expectation value as follows

$$\langle A \rangle = \text{Tr}\hat{\rho}\hat{A} \tag{12.3}$$

If the system in question is all in one single state $|\psi\rangle$, we have what's called a *pure ensemble*, whereas if the system is in a mixture of states, each with a probability p_i, we are in what's called a *mixed ensemble* or a *statistical mixture*.
In the case of a statistical mixture of states $|\psi_i\rangle$, we then define our density operator as follows

$$\hat{\rho} = \sum_i p_i |\psi_i\rangle \langle\psi_i| \tag{12.4}$$

Another important feature we can define for the density matrix is its behavior in time evolution. For this we start by writing both the time dependent Schrödinger equation and its adjoint

$$\begin{cases} i\hbar\dfrac{\partial}{\partial t} |\psi\rangle = \hat{\mathcal{H}} |\psi\rangle \\ -i\hbar\dfrac{\partial}{\partial t} \langle\psi| = \langle\psi| \hat{\mathcal{H}} \end{cases} \tag{12.5}$$

From the definition of $\hat{\rho}$ for statistical mixtures, inserting it into the Schrödinger equation we obtain

$$i\hbar\frac{\partial\hat{\rho}}{\partial t} = i\hbar\sum_i p_i \left(\frac{\partial}{\partial t} |\psi_i\rangle \langle\psi_i| + |\psi_i\rangle \frac{\partial}{\partial t} \langle\psi_i| \right) =$$
$$= \sum_i p_i \left(\hat{\mathcal{H}} |\psi_i\rangle \langle\psi_i| - |\psi_i\rangle \langle\psi_i| \hat{\mathcal{H}} \right) \tag{12.6a}$$

From this, substituting again into it the definition of $\hat{\rho}$, we obtain the *Von Neumann equation*

$$i\hbar\frac{\partial\hat{\rho}}{\partial t} = \left[\hat{\mathcal{H}}, \hat{\rho} \right] \tag{12.6b}$$

In Heisenberg representation, this equation appears easily remembering the definition of $\hat{\rho}$. Without delving into the calculations (which can be made directly using the definition itself of a time-evolved state), we get

$$\hat{\rho}(t) = \hat{\mathcal{U}}(t)\hat{\rho}\hat{\mathcal{U}}^\dagger(t) \tag{12.6c}$$

13 Microcanonical Ensemble

§ 13.1 The Classical Definition

A microcanonical ensemble is an ensemble which is *isolated* from the universe.
Such system can be described by a fixed number of particles in a fixed volume V with energies lying in an interval $[E, E + \delta]$, with an Hamiltonian $\mathcal{H}(q, p)$. What we wish to find is the classical density matrix of this system (i.e. the distribution function for such system). We already know that the region of phase space (the space composed by all the coordinates and the conjugated momenta) occupied by such system must be an hypervolume limited by two hypersurfaces with the following equations

$$\mathcal{H}(q, p) = E$$
$$\mathcal{H}(q, p) = E + \delta$$

(13.1)

This hypervolume is called *energy shell*. Since this state must tend to an equilibrium, we know already that every single microstate must be equiprobable and if they weren't, the density matrix would depend from other factors and would end up not commuting with the Hamiltonian, and bringing the system to never actually reach equilibrium.
Such ensemble is called the *microcanonical ensemble*, and the associated *microcanonical distribution function* has the following form

$$\rho_{MC}(q, p) = \begin{cases} \dfrac{1}{\Omega(E)\delta} & E \le \mathcal{H}(q, p) \le E + \delta \\ 0 & \mathcal{H}(q, p) \notin [E, E + \delta] \end{cases}$$

(13.2)

Where $\Omega(E)\delta$ is the volume occupied in the phase space, which we will now identify as Γ^{2n}. In the limit $\delta \to 0$ we have

$$\rho_{MC}(q, p) = \frac{1}{\Omega(E)} \delta \left(E - \mathcal{H}(q, p) \right)$$

(13.3)

From this, we can immediately determine $\Omega(E)$ by calculating the normalization of $\rho_{MC}(q, p)$.

$$\int \rho_{MC}(q, p) \, \mathrm{d}\mu = \frac{1}{h^{3N} N!} \int \rho_{MC}(q, p) \, \mathrm{d}q \mathrm{d}p = 1$$

(13.4)

Where we used the Lebesgue measure $\mathrm{d}\mu = (h^{3N} N!)^{-1} \mathrm{d}q \mathrm{d}p$, where h is a constant that comes out from the need of having a discrete phase space (see thermodynamics) and the factor $N!$ is a factor discovered by Gibbs, which results from the indistinguishability of the particles. Its absence would let the existence of an entropy of mixing gases, this is referred in literature as *Gibbs' paradox*.

From what we have then defined before, we now can write that the function $\Omega(E)$ can be evaluated with the following integral

$$\Omega(E) = \int \delta\left(E - \mathcal{H}(q,p)\right) \, \mathrm{d}\mu \tag{13.5}$$

This function is also referred as the *phase surface* of the system.
We can also define the volume of the region $\overline{\Omega}(E)$ as follows

$$\overline{\Omega}(E) = \int \Theta\left(E - \mathcal{H}(q,p)\right) \, \mathrm{d}\mu \tag{13.6}$$

From functional analysis, we already know that the following relation then holds

$$\Omega(E) = \frac{\mathrm{d}\overline{\Omega}}{\mathrm{d}E} \tag{13.7}$$

Comparing both $\overline{\Omega}(E)$ and $\Omega(E)\delta$ we can see that

$$\log\left(\Omega(E)\delta\right) = \log\left(\Omega(E)\right) + \mathcal{O}\left(\log\left(\frac{E}{n\delta}\right)\right)$$

Therefore, for large values of n we can write

$$\Omega(E)\delta = \overline{\Omega}(E) \tag{13.8}$$

§ 13.2 The Quantum Definition

Quantum mechanically, the definition of a microcanonical ensemble is an isolated system with an Hamiltonian $\hat{\mathcal{H}}$ and associated energy eigenvalues E_n with eigenstate $|n\rangle$. The microcanonical density operator then is defined as follows

$$\hat{\rho}_{MC} = \sum_n p\left(E_n\right) |n\rangle \langle n| \tag{13.9}$$

Where the probability of each energy level E_n analogously to classical mechanics, is defined as follows

$$p(E_n) = \begin{cases} \dfrac{1}{\Omega(E)\delta} & E \leq E_n \leq E + \delta \\ 0 & E_n \notin [E, E+\delta] \end{cases} \tag{13.10}$$

From the normalization of $\hat{\rho}_{MC}$ we get

$$\hat{\rho} = \frac{1}{\Omega(E)}\hat{\delta}\left(\hat{\mathcal{H}} - E\right)$$
$$\Omega(E) = \text{Tr}\hat{\delta}\left(\hat{\mathcal{H}} - E\right)$$
(13.11)

Or introducing a new notation

$$\Omega(E) = \frac{1}{\delta}\sideset{}{'}\sum_{\{|n\rangle\}} 1$$
(13.12)

Where the $\sum'$ symbol indicates that the sum is restricted only to the energy eigenvalues contained in the energy shell

§ 13.3 Entropy, Temperature and Pressure

Let ρ be an arbitrary density matrix for a system. The *entropy* of this system is defined as follows

$$S = -k_B \text{Tr}\left(\rho \log(\rho)\right) = -k_B \left\langle \log(\rho) \right\rangle$$
(13.13)

From this definition, all properties of classical entropy, already well known from macroscopic thermodynamics follow.

This definition is easily transferable to the quantum definition (where the entropy is usually called the *Von Neumann entropy*), through the quantization $\rho(q, p) \to \hat{\rho}$. Another property of entropy can be given immediately. Let $\hat{\rho}$ and $\hat{\rho}_1$ be two density operators, then

$$\text{Tr}\left(\hat{\rho}\left(\log(\hat{\rho}_1) - \log(\hat{\rho})\right)\right) \leq 0$$
(13.14)

Diagonalizing both operators, we can write

$$\text{Tr}\left(\hat{\rho}\left(\log(\hat{\rho}_1) - \log(\hat{\rho})\right)\right) = \sum_n P_n \langle n|\left(\log(\hat{\rho}_1)\right) - \log(P_n)|n\rangle = \sum_n P_n \langle n|\log\left(\frac{\hat{\rho}_1}{P_n}\right)|n\rangle =$$

$$= \sum_n \sum_\nu P_n \langle n|\nu\rangle \langle \nu|\log\left(\frac{P_{1\nu}}{P_n}\right)|\nu\rangle \langle \nu|n\rangle \leq$$

$$\leq \sum_n \sum_\nu P_n \langle n|\nu\rangle \langle \nu|\left(\frac{P_{1\nu}}{P_n} - 1\right)|\nu\rangle \langle \nu|n\rangle =$$

$$= \sum_n P_n \langle n|\left(\frac{\hat{\rho}_1}{P_n} - 1\right)|n\rangle = \text{Tr}(\hat{\rho}_1) - \text{Tr}(\hat{\rho}) \leq 0$$

§§ 13.3.1 Entropy of a Microcanonical Ensemble

Using the definition of the Von Neumann entropy, we define

$$S_{MC} = -k_B \text{Tr}\left(\hat{\rho}_{MC}\log(\hat{\rho}_{MC})\right) = -k_B \text{Tr}\left(\hat{\rho}_{MC}\log\left(\frac{1}{\Omega(E)\delta}\right)\right)$$

From the normalization condition of $\hat{\rho}_{MC}$, we obtain

$$S_{MC} = k_B \log\left(\Omega(E)\delta\right) \tag{13.15}$$

I.e. entropy is proportional to the accessible volume on the phase space (classically) or it's the logarithm of the number of accessible states (quantum mechanical)

§§ 13.3.2 Temperature and Pressure

Let's now define an isolated system composed of two interacting subsystems. The density operator for such system will be

$$\hat{\rho}_{MC} = \Omega_{1,2}^{-1}(E)\hat{\delta}\left(\hat{\mathcal{H}}_1 + \hat{\mathcal{H}}_2 - E\right) \tag{13.16}$$

Indicating the probability that the first system lies at some energy E_1 with $\omega(E_1)$, we have

$$\omega(E_1) = \left\langle \hat{\delta}\left(\hat{\mathcal{H}}_1 - E_1\right)\right\rangle = \frac{\Omega_2(E - E_1)\Omega_1(E_1)}{\Omega_{1,2}(E)} \tag{13.17}$$

Finding the extremal of such function, and putting it back to the definition of the microcanonical entropy, we get

$$\frac{\partial S_2}{\partial E_2} = \frac{\partial S_1}{\partial E_1} \tag{13.18}$$

If we now define the temperature as follows we have

$$\frac{1}{T} = \frac{\partial S}{\partial E} \tag{13.19}$$

We have

$$T_1 = T_2$$

Which is the most probable configuration.
Instead for defining pressure, we let the Hamiltonian depend from an external parameter a. We start by writing the volume in the phase space

$$\overline{\Omega}(E, a) = \int \Theta\left(E - \mathcal{H}(a)\right) \, d\Gamma \tag{13.20}$$

Taking the total differential, we have

$$d\overline{\Omega}(E, a) = \int \delta\left(E - \mathcal{H}(a)\right)\left(dE - \frac{\partial \mathcal{H}}{\partial a}\, da\right) = \Omega(E, a)\left(dE - \left\langle\frac{\partial \mathcal{H}}{\partial a}\right\rangle da\right) \tag{13.21}$$

In an alternative way we could write

$$d \log \overline{\Omega} = \frac{\Omega}{\overline{\Omega}}\left(dE - \left\langle\frac{\partial H}{\partial a}\right\rangle da\right)$$

Inserting the entropy into the equation $S(E, a) = k_B \log\big(\overline{\Omega}(E, a)\big)$ we have

$$\mathrm{d}S = \frac{1}{T}\left(\mathrm{d}E - \left\langle \frac{\partial \mathcal{H}}{\partial a} \right\rangle \mathrm{d}a\right) \tag{13.22a}$$

And, thus

$$\begin{cases} \dfrac{\partial S}{\partial E} = \dfrac{1}{T} \\[2mm] \dfrac{\partial S}{\partial a} = -\dfrac{1}{T}\left\langle \dfrac{\partial \mathcal{H}}{\partial a} \right\rangle \end{cases} \tag{13.22b}$$

By taking $a = V$ we then can define pressure.

Let's imagine an ideal gas inside an adiabatic piston with volume $V = LA$, where L is the length of the piston and A it's surface area.

The potential given by this system will be the sum of the single potentials of every N particles

$$V_{wall} = \sum_{i=1}^{N} v(x_i - L)$$

Where v is a function that is 0 for $x_i < L$ and infinite for $x_i > L$. By taking the gradient of this potential we then obtain the force applied by each particle on the walls

$$F = \sum_i F_i = -\sum_i \frac{\partial v}{\partial x_i} = \frac{\partial}{\partial L} \sum_i v(x_i - L) = \frac{\partial \mathcal{H}}{\partial L} \tag{13.23}$$

Multiplying by A^{-1}, we obtain

$$P = -\frac{\langle F \rangle}{A} = \frac{\partial \mathcal{H}}{\partial V} \tag{13.24}$$

And, therefore, we have

$$\mathrm{d}S = \frac{1}{T}\left(\mathrm{d}E + P\,\mathrm{d}V\right) \tag{13.25a}$$

And

$$\begin{cases} \dfrac{\partial S}{\partial E} = \dfrac{1}{T} \\[2mm] \dfrac{\partial S}{\partial V} = \dfrac{P}{T} \end{cases} \tag{13.25b}$$

Solving (13.25a) for $\mathrm{d}E$, we get

$$\mathrm{d}E = T\,\mathrm{d}S - P\,\mathrm{d}V \tag{13.25c}$$

Which is the *first law of thermodynamics*

§ 13.4 Examples

§§ 13.4.1 The Classical Ideal Gas

A simple example of a classical microcanonical ensemble is given by an ideal gas confined inside some volume that isolate it from the universe. Supposing that we have N molecules, we will write our Hamiltonian as follows

$$\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m} + V \tag{13.26}$$

The surface area of the energy shell will be

$$\Omega(E) = \frac{1}{h^{3N} N!} \int_V \prod_{i=1}^{N} \mathrm{d}^3 x_i \int \delta\left(E - \sum_{i=1}^{N} \frac{p_i^2}{2m}\right) \prod_{j=1}^{n} \mathrm{d}^3 p_j \tag{13.27}$$

Introducing the area of the d-dimensional unit sphere as follows

$$\int_{S^d} \mathrm{d}\sigma_d = \frac{2\sqrt{\pi^d}}{\Gamma\left(\frac{d}{2}\right)}$$

And then, evaluating the integral of the function $\overline{\Omega}(E)$, we have

$$\overline{\Omega}(E) = \frac{V^N}{h^{3N} N!} \int_{S^{3N}} \mathrm{d}\sigma_{3N} \int_0^{\sqrt{2mE}} p^{3N-1} \, \mathrm{d}p \tag{13.28}$$

Integrating directly and substituting the volume of the sphere V (considering the properties of the Euler Gamma function), we have

$$\overline{\Omega}(E) = \frac{V^N}{h^{3N} N!} \frac{(2\pi m E)^{\frac{3N}{2}}}{\left(\frac{3N}{2}\right)!} \tag{13.29}$$

Using the Stirling identity for the factorial for large values of N ($N! \approx N^N e^{-N}\sqrt{2\pi N}$), we have

$$\overline{\Omega}(E) \approx \left(\frac{V}{N}\right)^N \left(\frac{4\pi m E}{3h^2 N}\right)^{\frac{3N}{2}} e^{\frac{5N}{2}} \tag{13.30}$$

And, therefore, deriving

$$\Omega(E) = \frac{V^N}{h^{3N} N!} \frac{(2\pi m E)^{\frac{3N}{2}-1}}{\left(\frac{3N}{2}-1\right)!} \tag{13.31}$$

Which, for large values of N, becomes

$$\Omega(E) \approx \frac{3N}{2E} \left(\frac{V}{N}\right)^N \left(\frac{4\pi m E}{3h^2 N}\right)^{\frac{3N}{2}} e^{\frac{5N}{2}} \tag{13.32}$$

Starting from equation (13.31), and from the definition of entropy (13.15) we obtain the *Sakur-Tetrode* equation

$$S(E,V) = k_B N \log \left[\frac{V}{N} \left(\frac{4\pi m E}{3Nh^2} \right)^{\frac{3}{2}} e^{\frac{5}{2}} \right] \tag{13.33}$$

From the relations of entropy we can then define

$$\frac{1}{T} = \frac{\partial S}{\partial E} = \frac{3}{2}\frac{Nk_B}{E}$$
$$P = T\frac{\partial S}{\partial V} \qquad = k_B \frac{TN}{V} \tag{13.34}$$

And, from which we can write

$$E = \frac{3}{2}k_B T$$
$$PV = Nk_B T \tag{13.35}$$

Which both are the equation of state of an ideal gas.

§§ 13.4.2 Ideal Quantum Gas of Harmonic Oscillators

Let's consider a system of N non interacting quantum harmonic oscillators. The Hamiltonian of this system will be

$$\hat{\mathcal{H}} = \hbar\omega \sum_{i=1}^{N} \left(\hat{\eta}^\dagger \hat{\eta} + \frac{1}{2}\hat{\mathbb{1}} \right) \tag{13.36}$$

Thus, by definition, we have that

$$\Omega(E) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_N=0}^{\infty} \delta\left(E - \hbar\omega \sum_i \left(n_i + \frac{1}{2} \right) \right)$$
$$= \sum_{n_1=0}^{\infty} \cdots \sum_{n_N=0}^{\infty} \frac{1}{2\pi} \int e^{ik(E - \hbar\omega \sum_i (n_i + \frac{1}{2}))} \, \mathrm{d}k \tag{13.37}$$
$$= \frac{1}{2\pi} \int e^{ikE} \prod_{i=1}^{N} \frac{e^{-ik\frac{\hbar\omega}{2}}}{1 - e^{-ik\hbar\omega}} \, \mathrm{d}k$$

And finally, after evaluating the product

$$\Omega(E) = \frac{1}{2\pi} \int e^{N\left(ik\left(\frac{E}{N}\right) - \log\left(2i \sin\left(\frac{k\hbar\omega}{2}\right) \right) \right)} \, \mathrm{d}k \tag{13.38}$$

For evaluating this integral, for large N, one can use the saddle point method. We have that, where we wrote $\epsilon = E/N$ for simplifying the notation, that

$$f(k) = ik\epsilon - \log\left(2i \sin\left(\frac{k\hbar\omega}{2} \right) \right)$$

The function in study has a maximum in k_0

$$k_0 = \frac{1}{i\hbar\omega} \log\left(\frac{\epsilon + \frac{1}{2}\hbar\omega}{\epsilon - \frac{1}{2}\hbar\omega}\right)$$

And

$$f'(k_0) = i\epsilon - \frac{\hbar\omega}{2} \cot\left(\frac{k_0\hbar\omega}{2}\right)$$

$$f''(k_0) = \frac{\left(\frac{\hbar\omega}{2}\right)^2}{\sin^2\left(\frac{k_0\hbar\omega}{2}\right)}$$

We have then

$$\Omega(E) = \frac{1}{2\pi} e^{Nf(k_0)} \int e^{\frac{N}{2}f''(k_0)(k-k_0)^2} \, dk \tag{13.39}$$

The evaluation of this integral then yields

$$\Omega(E) = \exp\left[N\left(\frac{\epsilon + \frac{1}{2}\hbar\omega}{\hbar\omega} \log\left(\frac{\epsilon + \frac{1}{2}\hbar\omega}{\hbar\omega}\right) - \frac{\epsilon - \frac{1}{2}\hbar\omega}{\hbar\omega} \log\left(\frac{\epsilon - \frac{1}{2}\hbar\omega}{\hbar\omega}\right)\right)\right] \tag{13.40}$$

As we did before for the classical gas, we can write the entropy of the gas of quantum harmonic oscillators as follows

$$S(E) = k_B N \left(\frac{\frac{E}{N} + \frac{1}{2}\hbar\omega}{\hbar\omega} \log\left(\frac{\frac{E}{N} + \frac{1}{2}\hbar\omega}{\hbar\omega}\right) - \frac{\frac{E}{N} - \frac{1}{2}\hbar\omega}{\hbar\omega} \log\left(\frac{\frac{E}{N} - \frac{1}{2}\hbar\omega}{\hbar\omega}\right)\right) \tag{13.41}$$

And, therefore

$$\frac{1}{T} = \frac{k_B}{\hbar\omega} \log\left(\frac{\frac{E}{N} + \frac{1}{2}\hbar\omega}{\frac{E}{N} - \frac{1}{2}\hbar\omega}\right)$$

$$E = \frac{N\hbar\omega}{e^{\frac{\hbar\omega}{kT}} - 1} + \frac{N}{2}\hbar\omega \tag{13.42}$$

§§ 13.4.3 Paramagnetic Spin $\frac{1}{2}$ System

Another example we can use for a quantum microcanonical ensemble, is given by a system of N particles that can combine one of two possible states. This system is a good representation of a paramagnet in a magnetic field $\underline{H}$. The Hamiltonian of such system is

$$\hat{\mathcal{H}} = \mu_B H \sum_{i=1}^{N} \sigma_i \qquad \sigma_i = \pm 1 \tag{13.43}$$

The number of states of energy E is therefore

$$\Omega(E) = \sum_{\{\sigma_i = \pm 1\}} \delta\left(E - \mu_B H \sum_{i=1}^{N} \sigma_i\right) = \frac{1}{2\pi} \int \sum_{\{\sigma_i = \pm 1\}} e^{ik\left(E - \mu_B H \sum_i \sigma_i\right)} \, dk \tag{13.44}$$

Calculating the sum and using all properties of the exponential function, we get

$$\frac{1}{2\pi} \int e^{ikE} \left(2 \cos\left(-k\mu_B H\right)\right) \, \mathrm{d}k = \frac{2^N}{2\pi} \int e^{f(k)} \tag{13.45}$$

Where

$$f(k) = ikE + N \log \cos(-\mu_B H)$$

We find a maximum for this function at the point

$$-\mu_B H k_0 = \arctan\left(-\frac{iE}{N\mu_B H}\right) = \frac{i}{2} \log\left(\frac{1 - \frac{E}{N}\mu_B H}{1 + \frac{E}{N}\mu_B H}\right) \tag{13.46}$$

And, using the abbreviation $\epsilon = -\mu_B H E/N$ we have

$$\Omega(E) = 2^N e^{-\frac{N\epsilon}{2}\left(\log\left(\frac{1+\epsilon}{1-\epsilon}\right) + N\log\left(\frac{1}{\sqrt{1-\epsilon^2}}\right)\right)} \int \frac{e^{-\frac{1}{2}(-f''(k_0))(k-k_0)^2}}{2\pi} \, \mathrm{d}k \tag{13.47}$$

Integrating using the saddle point method, we have

$$\Omega(E) = \exp\left\{\left[-\frac{N}{2}\left((1+\epsilon)\log\left(\frac{1+\epsilon}{2}\right) + (1-\epsilon)\log\left(\frac{1-\epsilon}{2}\right) + \mathcal{O}(1, \log(N))\right)\right]\right\} \tag{13.48}$$

Writing $\epsilon = -\frac{E}{N\mu_B H}$ we have that the entropy of a quantum spin $\frac{1}{2}$ paramagnet is

$$S(E) = -\frac{Nk_B}{2}\left[(1+\epsilon)\log\left(\frac{1+\epsilon}{2} + (1-\epsilon)\log\left(\frac{1-\epsilon}{2}\right)\right)\right] \tag{13.49}$$

From which, we get

$$T = -\frac{2\mu_B H}{k_B \log\left(\frac{1-\epsilon}{1+\epsilon}\right)} \tag{13.50}$$

Plotting this equation, we see how we can have negative absolute temperatures for $\epsilon > 0$.
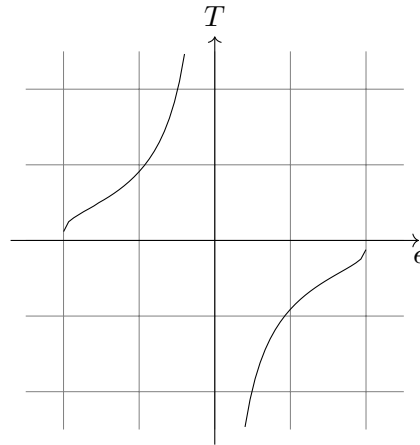
Figure 13.1: Plot of the Temperature of a quantum paramagnet, (13.50)

14 Canonical Ensemble

A canonical system can be seen as two systems, where the first one is much smaller and embedded into the second. Both system interact between each other by transferring energy. In order to work better we begin by constructing the density operator for such system. The Hamiltonian for this kind of system will then be the following

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_1 + \hat{\mathcal{H}}_2 \tag{14.1}$$

We now want to find the density operator for the first system, with Hamiltonian $\hat{\mathcal{H}}_1$. Let P_1 be the probability that the subsystem 1 is in a state $|n\rangle$ with energy eigenvalue $E_n^{(1)}$, then using the microcanonical distribution for the total system we find

$$P_1 = \sum_{\{|n\rangle_2\}}{}' \frac{1}{\Omega_{1,2}(E)\delta} = \frac{\Omega_2(E - E_n^{(1)})}{\Omega_{1,2}(E)} \tag{14.2}$$

In case that the system 2 is much bigger than the system 1 (which is our case), we can expand the logarithm of $\Omega_2(E - E_n^{(1)})$, therefore

$$P_1 \approx \frac{\Omega_2(E - \tilde{E}_1)}{\Omega_{1,2}(E)} e^{\frac{\tilde{E}_1 - E_n^{(1)}}{k_B T}} = \frac{1}{Z} e^{-\frac{E_n^{(1)}}{k_B T}} \tag{14.3}$$

The factor Z is called the *partition function* of the system, and it's calculated as

$$Z = \frac{\Omega_{1,2}(E)}{\Omega_2(\tilde{E}_2)} e^{-\frac{\tilde{E}_1}{k_B T}}$$

Or, directly as

$$Z = \sum_n e^{-\frac{E_n^{(1)}}{k_B T}} = \text{Tr}_1 e^{-\frac{\hat{\mathcal{H}}_1}{k_B T}} \tag{14.4}$$

The canonical density operator is then given by the following calculation

$$\hat{\rho}_C = \sum_n P_1 |n\rangle \langle n| = \frac{1}{Z} \sum_n e^{-\frac{E_n^{(1)}}{k_B T}} |n\rangle \langle n| = \frac{1}{Z} e^{-\frac{\hat{\mathcal{H}}_1}{k_B T}} \tag{14.5}$$

A second route which is both valid in the classical and quantum world is given as follows. We can write the following equality

$$\hat{\rho}_C = \mathrm{Tr}_2 \hat{\rho}_{MC} = \mathrm{Tr}_2 \left(\frac{\hat{\delta}\left(\hat{\mathcal{H}}_1 + \hat{\mathcal{H}}_2 - E\right)}{\Omega_{1,2}(E)} \right) = \frac{\Omega_2(E - \hat{\mathcal{H}}_1)}{\Omega_{1,2}(E)} \tag{14.6}$$

Approximating, we get

$$\hat{\rho}_C \approx \frac{\Omega_2(E - \tilde{E}_1)}{\Omega_{1,2}(E)} e^{\frac{\tilde{E}_1 - \hat{\mathcal{H}}_1}{k_B T}} \tag{14.7}$$

With this definition of density operator, the expected value for an observable in the subsystem 1 is then given as follows

$$\left\langle \hat{A} \right\rangle = \mathrm{Tr}_1 \mathrm{Tr}_2 \hat{\rho}_{MC} \hat{A} = \mathrm{Tr}_1 \hat{\rho}_C \hat{A} \tag{14.8}$$

If we transform the quantum traces to integrals in the classical case, we can then define the *partition function* of the system

$$Z = \int e^{-\frac{\mathcal{H}_1(q_1, p_1)}{k_B T}} \, \mathrm{d}\Gamma_1 \tag{14.9}$$

With the $\sum \leftrightarrow \int$ substitution the expected value of an observable is then given, obviously, by

$$\langle A(q_1, p_1) \rangle = \int \rho_C(q_1, p_1) A(q_1, p_1) \, \mathrm{d}\Gamma_1 \tag{14.10}$$

§ 14.1 Entropy of the Canonical Ensemble

From the definition of the microcanonical Von Neumann entropy we can define the entropy of the canonical ensemble as follows

$$S_C = -k_B \langle \log \hat{\rho}_C \rangle = \frac{1}{T} \left\langle \hat{\mathcal{H}} \right\rangle + k_B \log(Z) \tag{14.11}$$

Supposing that $\hat{\rho}$ corresponds to a different distribution with the same average energy, we then can define the following inequality

$$S[\hat{\rho}] = -k_B \mathrm{Tr}(\hat{\rho} \log \hat{\rho}) \leq k_B \mathrm{Tr}(\hat{\rho} \log \hat{\rho}_C) = \frac{1}{T} \left\langle \hat{\mathcal{H}} \right\rangle + k_B \log(Z) = S_C \tag{14.12}$$

I.e. the canonical ensemble has the greatest entropy of all ensembles $\hat{\rho}$ with the same average energy $\langle \hat{\rho} \rangle$.

§ 14.2 The Virial Theorem and the Equipartition Theorem

Theorem 14.1 (Classical Virial Theorem). *For a system with Hamiltonian $\mathcal{H} = T + V$, the following relation holds true*

$$\langle q^i i \partial_j V \rangle = k_B T \delta^i_j \tag{14.13}$$

Proof. Let's consider a classical system with coordinates $x_i = (q^i i, p_i)$, and we calculate the average value of the quantity $x^i \partial_j \mathcal{H}$. We have

$$\langle x^i \partial_j \mathcal{H} \rangle = \frac{1}{Z} \int x_i \partial_j \mathcal{H} e^{-\frac{\mathcal{H}}{k_B T}} \, d\Gamma$$

Applying an integration by part we obtain

$$\langle x^i \partial_j \mathcal{H} \rangle = -\frac{k_B T}{Z} \int x^i \partial_j e^{-\frac{\mathcal{H}}{k_B T}} \, d\Gamma = k_B T \delta^i_j \tag{14.14}$$

Applying this with the spatial coordinates q_i, we obtain what was searched with the theorem. $\square$

§ 14.3 Thermodynamic Quantities in the Canonical Ensemble

§§ 14.3.1 Equivalence of the Canonical and Microcanonical Macroscopic Ensembles

We start this section by stating the equivalence of the canonical and microcanonical ensembles for macroscopic systems. We have in this case that if E_1 is the most probable energy, we have

$$\langle E \rangle = E_1 \tag{14.15}$$

We start by rewriting the canonical partition functions in terms of the width of the most probable energy eigenvalue E_1 which we defined before. We have

$$Z = \frac{\Omega_{1,2}(E)}{\Omega_1(E_1)\Omega_2(E - E_1)} \Omega_1(E_1) e^{-\frac{E_1}{k_B T}} = \frac{\Omega_1(E_1)}{\omega(E_1)} e^{-\frac{E_1}{k_B T}} \tag{14.16}$$

We have tho that

$$\omega(E_1) \sim \frac{1}{\sqrt{N_1}} e^{-\frac{3(\tilde{E}_1 - E_1)^2}{4N_1 E_1}}$$

With N_1 as a normalization factor. From this we can write the partition function as

$$Z = \Omega(E_1) e^{-\frac{E_1}{k_B T}} \sqrt{N_1} \tag{14.17}$$

Evaluating now the canonical entropy, we have

$$S_C = \frac{1}{T} \left(\langle E \rangle - E_1 + k_B T \log(\Omega_1(E_1)) \right) = S_{MC}(E_1) \tag{14.18}$$

So, basically the entropy of a canonical ensemble is equal to that of the microcanonical ensemble with energy $E_1 = \langle E \rangle$. In both ensembles, then, one obtains the same identical results.

§§ 14.3.2 Thermodynamic Quantities

In these calculations, we'll leave all the indexes 1, referring to the small system inside the heat bath.

By definition, we write $\beta = (k_B T)^{-1}$, and our canonical density operator becomes

$$\hat{\rho}_C = \frac{e^{\beta \hat{\mathcal{H}}}}{Z} \tag{14.19}$$

Where

$$Z = \mathrm{Tr} e^{-\beta \hat{\mathcal{H}}} \tag{14.20}$$

We also define the *free energy* F as follows

$$F = -k_B T \log Z \tag{14.21}$$

From the definition of entropy we have

$$S_C = \frac{1}{T} \left(\langle E \rangle + k_B \log Z \right) \tag{14.22}$$

Where

$$\langle E \rangle = \left\langle \hat{\mathcal{H}} \right\rangle = -\frac{\partial}{\partial \beta} \log(Z) = k_B T^2 \frac{\partial}{\partial T} \log(Z) \tag{14.23}$$

And last but not least, pressure

$$P = -\left\langle \frac{\partial \hat{\mathcal{H}}}{\partial V} \right\rangle = k_B T \frac{\partial}{\partial V} \log(Z) \tag{14.24}$$

From this last definition of energy, it follows

$$F = \langle E \rangle - T S_C \tag{14.25}$$

Taking its total differential and evaluating, we get

$$\mathrm{d}F = -\frac{1}{T} \left(\langle E \rangle + k_B T \log(Z) \right) \, \mathrm{d}T + \left\langle \frac{\partial \hat{\mathcal{H}}}{\partial V} \right\rangle \, \mathrm{d}V$$

$$\mathrm{d}F = -S_C \, \mathrm{d}T - P \, \mathrm{d}V \tag{14.26}$$

§§ 14.3.3 Heat

In order to define the statistical meaning of *heat transfer* we need to begin with the average value of energy.

$$\langle E \rangle = \left\langle \hat{\mathcal{H}} \right\rangle = \mathrm{Tr}(\hat{\rho} \hat{\mathcal{H}}) \tag{14.27}$$

In general, we have $\langle E \rangle = \sum_i p_i E_i$ with E_i as our energy eigenstate and p_i its associated variable.

Since

$$\mathrm{d} \langle E \rangle = \sum_i E_i \, \mathrm{d} p_i + \sum_i p_i \, \mathrm{d} E_i$$

And

$$\mathrm{d} \langle E \rangle = T \, \mathrm{d} S + \left\langle \frac{\partial \hat{\mathcal{H}}}{\partial V} \right\rangle \mathrm{d} V$$

We obtain, from the definition $\mathrm{d} Q = T \, \mathrm{d} S$

$$\mathrm{d} Q = \sum_i E_i \, \mathrm{d} p_i \tag{14.28}$$

This defines the heat variation as a redistribution of the occupation probabilities of the i-th state $|i\rangle$

15 Grand Canonical Ensemble

After having considered the canonical ensemble, i.e. where one system embedded in another can only exchange energy with it (like in a heat bath), we're going to study the grand canonical ensemble, where an overall system is isolated as in the canonical ensemble, but both system can exchange energy and particles.

The probability distribution of the state variables E_1, N_1, V_1 is the following, in analogy with what we wrote for the canonical ensemble

$$\omega(E_1, N_1, V_1) = \frac{\Omega_1(E_1, N_1, V_1)\Omega_2(E - E_1, N - N_1, V - V_1)}{\Omega(E, N, V)} \tag{15.1}$$

From the equilibrium conditions $(T_1 = T_2,\ V_1 = V_2, N_1 = N_2)$ obtained from the equivalence of the logarithmic derivatives obtained from the normalization conditions of Ω, and from the differential of entropy in the variables N, V, T, we get, after defining the *chemical potential* μ

$$\mathrm{d}S = \frac{1}{T}\,\mathrm{d}E + \frac{P}{T}\,\mathrm{d}V - \frac{\mu}{T}\,\mathrm{d}N$$

$$\frac{\mu}{T} = k_B \frac{\partial}{\partial N}\log\Omega$$

We have finally

$$\mu = -k_B T \frac{\partial}{\partial N}\log\Omega = -T\frac{\partial S}{\partial N} \tag{15.2}$$

We now need to define the density operator $\hat{\rho}_G$ for the subsystem 1. The probability that in the embedded system there are N_1 particles in the state $|n\rangle$ is

$$p(N_1, E_{1n}(N_1), V_1) = \frac{\Omega_2(E - E_{1n}, N - N_1, V_2)}{\Omega(E, N, V)} = \frac{1}{Z_G}e^{-\frac{E_{1n} - \mu N_1}{k_B T}} \tag{15.3}$$

Where Z_G is the grand canonical partition function, also known as *Gibbs distribution*. We thus have

$$\hat{\rho}_G = \frac{1}{Z_G}e^{-\frac{\hat{\mathcal{H}}_1 - \mu N_1 \hat{\mathbb{1}}}{k_B T}} \tag{15.4}$$

And

$$Z_G = \mathrm{Tr}\left(e^{-\frac{\hat{\mathcal{H}}_1 - \mu N_1 \hat{\mathbb{1}}}{k_B T}}\right) = \sum_{N_1} Z(N_1)e^{\frac{\mu N_1}{k_B T}} \tag{15.5}$$

The classical limit can be achieved simply applying this abuse of notation

$$\text{Tr} \rightarrow \sum_{N_1} \int d\Gamma_{N_1}$$

From the definition of the density matrix, the entropy of the grand canonical ensemble will be

$$S_G = -k_B \langle \log \hat{\rho}_g \rangle = \frac{1}{T}(E - \mu N) + k_B \log \hat{\rho}_G \tag{15.6}$$

§ 15.1 Thermodynamic Quantities

Analogously to the free energy of the canonical ensemble, we can define the *grand potential* as follows

$$\Phi = -k_B T \log(Z_G) \tag{15.7}$$

Where, using (15.6), we can write alternatively

$$\Phi(V, T, \mu) = E - TS_G - \mu N \tag{15.8}$$

And its differential relations follow

$$\frac{\partial \Phi}{\partial T} = \frac{1}{T}(\Phi - E + \mu N) = -S_G$$
$$\frac{\partial \Phi}{\partial V} = \left\langle \frac{\partial \hat{\mathcal{H}}}{\partial V} \right\rangle = -P \tag{15.9}$$
$$\frac{\partial \Phi}{\partial \mu} = -N$$

Now, in order to avoid confusion between the ensembles and their properties, we can compose a table.

Ensemble	Microcanonical	Canonical	Grand Canonical
Situation	Isolated	Energy Exchange	Energy and Particle Exchange
Independent Variables	E, V, N	T, V, N	T, V, μ
Density Operator	$\Omega^{-1}\hat{\delta}\left(\hat{\mathcal{H}} - E\right)$	$Z^{-1}e^{-\hat{\mathcal{H}}/k_B T}$	$Z_G^{-1}e^{-(\hat{\mathcal{H}} - \mu N \hat{\mathbb{1}})/k_B T}$
Normalization	$\Omega = \text{Tr}\left(\hat{\delta}(\hat{\mathcal{H}} - E)\right)$	$Z = \text{Tr}(e^{-\hat{\mathcal{H}}/k_B T})$	$Z_G = \text{Tr}\left(e^{-(\hat{\mathcal{H}} - \mu N \hat{\mathbb{1}})/k_B T}\right)$
Thermodynamic Potential	S	F	Φ

Table 15.1: Table of the various ensembles, with the definition of their density operators, their normalization and the main thermodynamic potential for each ensemble

§ 15.2 Examples

As our main example we choose again the classical ideal gas.
We begin by calculating the partition function for N particles (keep in mind that $\beta = (k_B T)^{-1}$

$$Z_N = \frac{1}{N! h^{3N}} \int_V \int e^{-\beta \sum_{i=1}^{N} \frac{p_i^2}{2m}} \, \mathrm{d}^{3N} p \, \mathrm{d}^{3N} q \tag{15.10}$$

Noting that the integral on the coordinates is simply V^N, we can write, also noting the isotropy of momenta

$$Z_N = \frac{V^N}{N! h^{3N}} \left(\int_0^\infty e^{-\beta \frac{p^2}{2m}} \, \mathrm{d}p \right)^{3N} = \frac{V^N}{N!} \left(\frac{2m\pi}{\beta h^2} \right)^{\frac{3N}{2}} \tag{15.11}$$

Using the definition of thermal wavelength of $\lambda = h(2\pi m k_B T)^{-1/2}$, we get finally

$$Z_N = \frac{1}{N!} \left(\frac{V}{\lambda^3} \right)^{\frac{3N}{2}} \tag{15.12}$$

Writing the definition of the grand partition function, we get

$$Z_G = \sum_{N=0}^{\infty} Z_N e^{\beta \mu N} = \sum_{N=0}^{\infty} \frac{e^{\beta \mu N}}{N!} \left(\frac{V}{\lambda^3} \right)^N = e^{z \frac{V}{\lambda^3}} \tag{15.13}$$

Where we introduced z as what is usually called as *fugacity* as

$$z = e^{\beta \mu} \tag{15.14}$$

From this, we can write the grand potential and get all the thermodynamic relations of the classical gas

$$\Phi(T, V, \mu) = -k_B T \log(Z_G) = -k_B T \frac{zV}{\lambda^3} \tag{15.15}$$

Therefore, omitting all the algebraical passages and calculating μ through the derivatives of the grand potential and the definition of fugacity, we have

$$\begin{aligned}
-\frac{\partial \Phi}{\partial \mu} &= N = \frac{zV}{\lambda^3} \\
-V \frac{\partial \Phi}{\partial V} &= PV = N k_B T \\
-\frac{\partial \Phi}{\partial T} &= S = k_B N \left(\frac{5}{2} + \log \left(\frac{V}{N \lambda^3} \right) \right) \\
\mu &= -k_B T \log \left(\frac{V}{N \lambda^3} \right)
\end{aligned} \tag{15.16}$$

16 Ideal Quantum Gases

In this chapter we will derive with quantum statistical mechanics, all the thermodynamic properties of non interacting quantum gases of fermions and bosons. We begin by defining the grand potential for N non interacting and non relativistic particles confined inside a box with volume $V = L^3$. The Hamiltonian of the system is

$$\hat{\mathcal{H}} = \sum_{i=1}^{N} \frac{\hat{p}_i^2}{2m} \tag{16.1}$$

Applying periodic boundary conditions to the associated differential equation we obtain the following solution

$$\langle x_i | p_i \rangle = \phi_{p_i}(x_i) = \frac{1}{\sqrt{V}} e^{\frac{i p_i x^i}{\hbar}} \tag{16.2}$$

Where

$$p_i = \frac{2\pi\hbar}{L} \nu_i \tag{16.3}$$

Where we have $\nu_i \in \mathbb{Z}$. The single particle energy ϵ_p will be, obviously

$$\epsilon_p = \frac{p^2}{2m} \tag{16.4}$$

Now, if we consider particle spin, we find ourselves in a particular situation. As we have seen in the chapter for identical particles, we have that a multi-particle factorisable eigenket of the Hamiltonian (16.1) can then be written as follows

$$|p_1, \cdots, p_N\rangle = \mathcal{N} \sum_P (\pm 1)^P \hat{P} \bigotimes_{i=1}^{N} |p_i\rangle \tag{16.5}$$

Where $\hat{P}$ is the permutation operator, with eigenvalue 1^P and normalization $\mathcal{N} = (\prod_i N! n_{p_i}!)^{-1/2}$ for bosons, and eigenvalue $(-1)^P$ and normalization $\mathcal{N} = (N!)^{-1/2}$ for fermions. For an N-particle system, we can define the particle number as follows

$$N = \sum_p n_p \tag{16.6}$$

And the energy eigenvalue as follows

$$E(\{n_p\}) = \sum_p n_p \epsilon_p \qquad (16.7)$$

Therefore, the grand canonical partition funtion will be the following

$$Z_G = \sum_{N=0}^{\infty} \sum_{\{n_p\}} e^{-\beta(E(\{n_p\})-\mu N)} = \sum_{\{n_p\}} e^{-\beta \sum_p n_p(\epsilon_p-\mu)} = \\ = \prod_p \sum_{n_p} e^{-\beta n_p(\epsilon_p-\mu)} \qquad (16.8)$$

Therefore, summing and considering the difference between bosons and fermions

$$Z_G = \prod_p \sum_{n_p} e^{-\beta n_p(\epsilon_p-\mu)} = \begin{cases} \prod_p \dfrac{1}{1-e^{-\beta(\epsilon_p-\mu)}} & m_s \in \mathbb{Z} \\[2ex] \prod_p (1+e^{-\beta(\epsilon_p-\mu)}) & m_s \in \mathbb{F} := \left\{ m_s \in \mathbb{Q} \,\middle|\, m_s = \dfrac{n}{2},\ n \in \mathbb{Z} \right\} \end{cases} \qquad (16.9)$$

From this we van calculate directly the grand potential

$$\Phi = -\frac{1}{\beta} \log(Z_G) = \pm \frac{1}{\beta} \sum_p \log \left(1 \mp e^{-\beta(\epsilon_p-\mu)} \right) \qquad (16.10)$$

With the upper sign referring to bosons and vice versa for fermions.
The average particle number will then be

$$N = -\frac{\partial \Phi}{\partial \mu} = \sum_p \frac{1}{e^{\beta(\epsilon_p-\mu)} \mp 1} = \sum_p n(\epsilon_p) \qquad (16.11)$$

The last function $n(\epsilon_p)$ is called the Bose-Einstein distribution (for bosons) or the Fermi-Dirac distributions (for fermions). From this, we can find that it's actually the *average occupation number* of a state $|\overline{\alpha}\rangle$. In order to obtain this result we need to calculate the expectation value of $n_{\overline{\alpha}}$.

$$\langle n_{\overline{\alpha}} \rangle = \mathrm{Tr}(\hat{\rho}_G n_{\overline{\alpha}}) = \frac{\sum_{\{n_p\}} n_{\overline{\alpha}} e^{-\beta \sum_p n_p(\epsilon_p-\mu)}}{\sum_{\{n_p\}} e^{-\beta \sum_p n_p(\epsilon_p-\mu)}} = -\frac{\partial}{\partial x} \log \left(\sum_n e^{-nx} \right) = n(\epsilon_{\overline{\alpha}}) \quad (16.12)$$

From the grand potential we then get the energy of the quantum gas

$$E = \left(\frac{\partial(\Phi\beta)}{\partial\beta} \right)_{\mu\beta} = \sum_p \epsilon_p n(\epsilon_p) \qquad (16.13)$$

Considering that free particles can be considered as being confined to a space $\Delta = 2\pi\hbar V^{-1} \to \infty$, we can choose to approximate the sum over the discrete p to an integral for large volumes, using the following substitution

$$\sum_{\underline{p}}[\cdot] \to \frac{gV}{(2\pi\hbar)^3} \int [\cdot]\, \mathrm{d}^3p \tag{16.14}$$

Where g is the degeneracy factor
Using this approximation for calculating the number of particles $N = \sum_p n_p(\epsilon_p)$, we get

$$N = \frac{gV}{(2\pi\hbar)^3} \int n(\epsilon_p)\, \mathrm{d}^3p = \frac{gV}{2\pi^2\hbar^3} \int_0^\infty n(\epsilon_p)p^2\, \mathrm{d}p$$
$$= \frac{gV}{2\pi^2\hbar^3} \int_0^\infty \frac{p^2}{e^{\beta(\epsilon-\mu)} \mp 1}\, \mathrm{d}p = \frac{gVm^{\frac{3}{2}}}{\pi^2\hbar^3\sqrt{2}} \int_0^\infty \frac{\sqrt{\epsilon}}{e^{\beta(\epsilon-\mu)}}\, \mathrm{d}\epsilon \tag{16.15}$$

Where we substituted in the energy eigenvalue density. Defining the specific volume $v = V/N$ and substituting $x = \beta\epsilon$, we get

$$\frac{1}{v} = \frac{2g}{\lambda^3\sqrt{\pi}} \int_{\mathbb{R}_+} \frac{\sqrt{x}}{e^x z^{-1} \mp 1} = \frac{g}{\lambda^3} \begin{cases} g_{3/2}(z) & s \in \mathbb{Z} \\ f_{3/2}(z) & s = \frac{n}{2},\ n \in \mathbb{Z} \end{cases} \tag{16.16}$$

Where g_s, f_s are the generalized $\zeta-$functions, which are defined and analyzed in the mathematical appendix.
From this, taking the grand partition function, we have that

$$\Phi = \pm\frac{gV}{\beta(2\pi\hbar)^3} \int \log\left(1 \mp e^{-\beta(\epsilon-\mu)}\right) \mathrm{d}^3p$$
$$= \pm\frac{gVm^{\frac{3}{2}}}{\beta\pi^2\hbar^3\sqrt{2}} \int_0^\infty \log\left(1 \mp e^{\beta(\epsilon-\mu)}\right)\sqrt{\epsilon}\, \mathrm{d}\epsilon \tag{16.17}$$

Integrating by parts and remembering that $PV = -\Phi$ we get

$$-\Phi = PV = \frac{gm^{\frac{3}{2}}V\sqrt{2}}{3\pi^2\hbar^3} \int_0^\infty \frac{\epsilon^{\frac{3}{2}}}{e^{\beta(\epsilon-\mu)} \mp 1}\, \mathrm{d}\epsilon = \frac{gV}{\beta\lambda^3} \begin{cases} g_{\frac{5}{2}}(z) \\ f_{\frac{5}{2}}(z) \end{cases} \tag{16.18}$$

We also can obtain the energy E of the system as follows

$$E = \frac{gVm^{\frac{3}{2}}}{\pi^2\hbar^3\sqrt{2}} \int_{\prime}^\infty \frac{\epsilon^{\frac{3}{2}}}{e^{\beta(\epsilon-\mu)} \mp 1} \tag{16.19}$$

A quick comparison with the equation (16.18), gives the same relation that we got for the classical ideal gas

$$PV = \frac{2}{3}E \tag{16.20}$$

From the homogeneity of Φ in T, μ we can derive from the previous equations other relations, as follows

$$P = -\frac{\Phi}{V} = -T^{\frac{5}{2}}\phi\left(\frac{\mu}{T}\right)$$

$$N = VT^{\frac{3}{2}}n\left(\frac{\mu}{T}\right)$$

$$S = -\frac{\partial\Phi}{\partial T} = VT^{\frac{3}{2}}s\left(\frac{\mu}{T}\right) \qquad (16.21)$$

$$\frac{S}{N} = \frac{s}{n}$$

For an adiabatic expansion, i.e. setting $S = \alpha$, $\mu/T = \beta$, $VT^{\frac{3}{2}} = \gamma$, $PT^{-\frac{5}{2}} = \delta$, with $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, we get the adiabatic equation for an ideal quantum gas

$$PV^{\frac{5}{3}} = \eta \in \mathbb{R} \qquad (16.22)$$

Note how this differs from the classical version, since $c_p/c_v \neq 5/3$

§ 16.1 Degenerate Fermi Gas

Let's consider now the ground state of N fermions. It will correspond to a fermion gas at temperature $T = 0$K. In this situation, every single particle state will be occupied g fold, thus all momenta inside a sphere of radius p_F (the maximum momentum possible, the *Fermi momentum*) will be occupied.
The number of particles therefore will be

$$N = g\sum_{\{|p\rangle\}}{}' 1 = \frac{gV}{(2\pi\hbar)^3}\int \Theta(p_F - p)\,\mathrm{d}^3p = \frac{gVp_F^3}{6\pi^2\hbar^3} \qquad (16.23)$$

Therefore, using the particle density $n = N/V$ we get our Fermi momentum

$$p_F = \hbar\sqrt[3]{\frac{6\pi^2 n}{g}} \qquad (16.24)$$

From this, we get the *Fermi Energy*

$$\epsilon_{p_F} = \frac{p_F^2}{2m} = \frac{\hbar^2}{2m}\left(\frac{6\pi^2 n}{g}\right)^{\frac{2}{3}} \qquad (16.25)$$

The ground state energy, from these relations, will therefore be

$$E = \frac{gV}{(2\pi\hbar)^3}\int \frac{p^2}{2m}\Theta(p_F - p)\,\mathrm{d}^3p = \frac{gVp_F^5}{20m\pi^2\hbar^3} = \frac{3}{5}N\epsilon_F \qquad (16.26)$$

Using what we found in the previous section, we find that the pressure of such gas will be the following

$$P = \frac{2}{5}\epsilon_F n = \frac{\hbar^2 n^{\frac{5}{2}}}{5m}\left(\frac{6\pi^2}{g}\right)^{\frac{2}{3}} \qquad (16.27)$$

§§ 16.1.1 Complete Degeneracy Limit

Having calculated the thermodynamic properties of a quantum gas of fermions in the case of complete degeneracy (i.e., $T = 0$), we can start calculating the same properties in the *limit* of complete degeneracy, i.e. for $T \to 0$. It's easy to demonstrate that here $\mu \to \epsilon \to \epsilon_F$ and therefore

$$\Phi = -N\epsilon_F^{-\frac{3}{2}} \int_0^\infty n(\epsilon)\epsilon^{\frac{3}{2}} \, d\epsilon$$
$$N = \frac{3}{2}N\epsilon_F^{-\frac{3}{2}} \int_0^\infty n(\epsilon)\epsilon^{\frac{1}{2}} \, d\epsilon$$

(16.28)

From this, knowing already the solution of these integrals, as discussed in appendix (A.3), we can solve these integrals approximately in the limit $\beta\mu \to \infty$, and deduce some approximated conclusions for what happens thermodynamically in a Fermi gas for really low temperatures. We begin writing our integrals (called *Sommerfield integrals*) as a sum of two integrals as follows

$$I = \int_0^\mu f(\epsilon) \, d\epsilon + \int_0^\infty f(\epsilon) \left(n(\epsilon) - \Theta(\mu - \epsilon) \right) \, d\epsilon$$

(16.29)

Using a $x-$substitution with $x = \beta(\epsilon - \mu)$, extending the integral's domain over the whole real line, and Taylor approximating the function $f(\epsilon)$ around μ, we get

$$I = \int_0^\mu f(\epsilon) \, d\epsilon + \int_{\mathbb{R}} \left(\frac{1}{e^x + 1} - \Theta(-x) \right) \sum_{k=0}^\infty \frac{\beta^{-(k+1)}}{k!} \left. \frac{\partial^k f}{\partial x^k} \right|_{x=\mu} x^k \, dx$$
$$= \int_0^\mu f(\epsilon) \, d\epsilon + 2 \sum_{k=0}^\infty \frac{\beta^{-(k+1)}}{k!} \frac{\partial^k f}{\partial \mu^k} \int_0^\infty \frac{x^k}{e^x + 1} \, dx$$

(16.30)

Applying this approximation till $\mathcal{O}(T^4)$ we can write for the integrals (16.28)

$$\mu = \epsilon_F \left(1 - \frac{\pi^2}{12\beta^2} + \mathcal{O}(T^4) \right)$$
$$\Phi = -\frac{2}{5}N\epsilon_F \left(1 + \frac{5\pi^2}{12\beta^2\epsilon_F^2} + \mathcal{O}(T^4) \right)$$

(16.31)

Using $P = -\Phi/V$ we obtain immediately the energy of such gas

$$E = \frac{3}{2}PV = \frac{3}{5}N\epsilon_F \left(1 + \frac{5\pi^2}{12\epsilon_F^2\beta^2} + \mathcal{O}(T^4) \right)$$

(16.32)

And introducing the Fermi temperature as $T_F = \epsilon_F/k_B$ we get the heat capacity of this gas as

$$C_V = Nk_B \frac{\pi^2 T}{2T_F}$$

(16.33)

§ 16.2 Bose-Einstein Condensation

After having studied the Fermi gas, we begin studying a Boson gas at low temperatures, which has a particular behavior called *Bose-Einstein condensation*. This gas has $s = 0$ and $g = 1$. Due to this, in the ground state all the non-interacting bosons occupy the lowest single particle state.
In the previous sections, we found that for the particle density we have

$$\frac{\lambda}{v} = g_{\frac{3}{2}}(z) \tag{16.34}$$

This function has a maximum for a value of fugacity $z = 1$, and it's equal to $g_{\frac{3}{2}}(1) = \zeta(\frac{3}{2}) = 2.612$. Thanks to this we can define a characteristic temperature T_c, which has the following value

$$\beta_c^{-1} = \frac{2\pi\hbar^2}{m(v\zeta(\frac{3}{2}))^{\frac{2}{3}}} \tag{16.35}$$

In this case, we have that the limit $\sum_p \to \int d^3p$ isn't anymore a good approximation for $z \to 1$, since the term $p = 0$ diverges for $z = 1$. Treating it separately, we get for the particle number

$$N = \frac{1}{z^{-1} - 1} + \sum_{p \neq 0} n(\epsilon_p) \to \frac{1}{z^{-1} - 1} + \frac{V}{(2\pi\hbar)^3} \int n(\epsilon_p)\, d^3p \tag{16.36}$$

Therefore, for bosons we get, in terms of generalized $\zeta-$functions and characteristic temperature

$$N = \frac{1}{z^{-1} - 1} + \frac{Nv}{\lambda^3} g_{\frac{3}{2}}(z)$$
$$N = \frac{1}{z^{-1} - 1} + N\left(\frac{T}{T_c}\right)^{\frac{3}{2}} \frac{g_{\frac{3}{2}}(z)}{g_{\frac{3}{2}}(1)} \tag{16.37}$$

This can be seen as a sum of the number of particles in the ground state N_0 and the number of particles in excited states N_e, where

$$N_0 = \frac{1}{z^{-1} - 1}$$
$$N_e = N\left(\frac{T}{T_c}\right)^{\frac{3}{2}} \frac{g_{\frac{3}{2}}(z)}{g_{\frac{3}{2}}(1)} \tag{16.38}$$

We have that for $T > T_c$, N yields a value of $z < 1$, hence N_0 is finite and can be neglected with respect to N. For $T < T_c$ we have $z = 1 - \mathcal{O}(N^{-1})$, and when $z \to 1$, setting $z = 1$ in N_e, we obtain

$$N_0 = N\left(1 - \left(\frac{T}{T_c}\right)^{\frac{3}{2}}\right) \tag{16.39}$$

And defining the *condensate fraction* ν as follows

$$\nu = \lim_{N \to \infty} \frac{N_0}{N} \tag{16.40}$$

We get, in summary, what's called the *Bose-Einstein Condensation*, for which, at $T < T_c$ the ground state at $p = 0$ is macroscopically occupied.

$$\nu = \begin{cases} 0 & T > T_c \\ 1 - \left(\dfrac{T}{T_c}\right)^{\frac{3}{2}} & T < T_c \end{cases} \tag{16.41}$$

Evaluating the other thermodynamic quantities, we get the pressure of a Bose gas as

$$P = \begin{cases} \dfrac{1}{\beta\lambda^3} g_{\frac{5}{2}}(z) & T > T_c \\ \dfrac{1}{\beta\lambda}\zeta\left(\dfrac{5}{2}\right) = \dfrac{1}{\beta\lambda^3} 1.342 & T < T_c \end{cases} \tag{16.42}$$

And, therefore, entropy has the following expression

$$S = \frac{\partial PV}{\partial T} = \begin{cases} Nk_B \left(\dfrac{5v}{2\lambda^3} g_{\frac{5}{2}}(z) - \log(z)\right) & T > T_c \\ Nk_B \dfrac{5}{2} \left(\dfrac{T}{T_c}\right)^{\frac{3}{2}} \dfrac{g_{\frac{5}{2}}(1)}{g_{\frac{3}{2}}(1)} \end{cases} \tag{16.43}$$

Part III

Atomic Physics

17 One Electron Atoms

§ 17.1 Interaction of Particles with EM Fields

Before diving into a full computation of an electromagnetic Hamiltonian, we need to properly "quantize" it. Having $\hat{\underline{A}}$ our vector potential operator and $\hat{\phi}$ our scalar potential operator, and knowing that $\underline{P} = \underline{p} - q\underline{A}$, where $\underline{P}$ is the minimal coupling momentum of an EM field, we have that the classical Hamiltonian for an electromagnetic field is

$$\mathcal{H} = \frac{1}{2m}\left(p_i - qA_i\right)^2 + q\phi \tag{17.1}$$

Quantizing, and noting that $\left[\hat{p}_i, \hat{A}_i\right] \neq 0$ generally, we have, setting the Coulombe gauge ($\partial_i A^i = 0, \ \phi = 0$) that the Hamiltonian in study for our quantum system is the following

$$\hat{\mathcal{H}} = \frac{1}{2m}\left(-\hbar^2\nabla^2 + q^2\hat{A}^2 + 2iq\hbar\hat{\underline{A}}\cdot\nabla\right) \tag{17.2}$$

The interaction of the field with the particle can be analyzed perturbatively, and considering first order terms in $\hat{\underline{A}}$ we have

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \frac{iq\hbar}{m}\hat{\underline{A}}\cdot\nabla \tag{17.3}$$

Where we put $\hat{\mathcal{H}}_0$ as the unperturbed Hamiltonian.
Calling the perturbative piece of the Hamiltonian $\hat{W}(t)$ we have that the perturbation on a transition between a state $|a\rangle$ and $|b\rangle$ will be given by the following integral

$$W_{ba}(t) = \frac{q}{m}\langle b|\,\hat{\underline{A}}\cdot\hat{\underline{p}}\,|a\rangle \tag{17.4}$$

Where the vector potential is the solution of the following integral

$$\hat{A}_i(x_i, t) = \frac{1}{2}\int_0^\infty A_0(\omega)\epsilon_i\left(e^{i(k_i x^i - \omega t + \delta_\omega)} + e^{-i(k_i x^i - \omega t + \delta_\omega)}\right)\,\mathrm{d}\omega \tag{17.5}$$

Take now a time evolved eigenstate of the unperturbed Hamiltonian, and rewrite as a linear combination of those times an unknown time-dependent function $c_k(t)$. We can write, for a generic eigenstate $|k\rangle$, that the new perturbed state will be $|\psi\rangle$

$$|\psi\rangle = \sum_k c_k(t)e^{-\frac{iE_k t}{\hbar}}|k\rangle \tag{17.6}$$

133

Inserting everything in the time-dependent Schrödinger equation for the perturbed Hamiltonian, and simplifying terms, we have

$$i\hbar \sum_k \dot{c}_k(t) e^{-\frac{iE_k t}{\hbar}} |k\rangle = \sum_k c_k \hat{W}(t) e^{-\frac{iE_k t}{\hbar}} |k\rangle \tag{17.7}$$

Checking now the transition $|k\rangle \to |b\rangle$, and remembering the orthonormality between states, we have, putting $\omega_{bk} = (E_b - E_k)/\hbar$

$$\dot{c}_b(t) = \frac{1}{i\hbar} \sum_k c_k(t) \langle b| \hat{W} |k\rangle e^{-i\omega_{bk}t} \tag{17.8}$$

Approximating $c_k(t)$ to the first order in λ and making the assumption that the system will be in a state $|a\rangle$ for $t \leq 0$, therefore implying $c_k(t) = \delta_{ka}$ for $k \neq a, b$, we have, rearranging terms of the same order

$$\dot{c}_b(t) = \frac{1}{i\hbar} \langle b| \hat{W} |a\rangle e^{-i\omega_{ba}t} \tag{17.9}$$

Factoring out constants and time dependent parameters, we have that the coefficient for our perturbation is

$$c_b(t) = \frac{q}{2im\hbar} \int_0^\infty \int_0^t A_0(\omega) \langle b| e^{ik_i x_i} \epsilon \cdot \underline{\hat{p}} |a\rangle e^{-i(\omega_{ba}-\omega)t'} e^{i\delta_\omega} \, \mathrm{dt d\omega} + c.c. \tag{17.10}$$

If $t >> 2\pi/\omega$ we can approximate the radiation as a plane wave solution. Parting the integrals we have that one will be nonzero for $\omega_{ba} \simeq \omega$ and the second for $\omega_{ba} \simeq -\omega$. The first nonzero solution describes an absorption between the state with $|a\rangle$ and $|b\rangle$, with $E_b > E_a$, and the second instead describes an emission between the state $|a\rangle$ and $|b\rangle$. Now, indicating $\langle b| e^{ik_i x^i} \hat{p} |a\rangle$ as M_{ba} and noting that $\langle e^{i\delta_\omega} \rangle = \delta(\omega)$ we have that the transition probability will be

$$|c_b(t)|^2 = \frac{q^2}{4m^2\hbar^2} \int_0^\infty |A_0(\omega)|^2 \left| \langle b| e^{ik_i x_i} \underline{\epsilon} \cdot \underline{\hat{p}} |a\rangle \right|^2 F(t, \omega_{ba} - \omega) \mathrm{d\omega} \tag{17.11}$$

Where the function F is defined as the square modulus of the integral of the exponential, which corresponds to a sinc^2 function For $t >> 1$ we have that, for the properties of the function F, Fermi's Golden Rule, and the connection between A^2 and the intensity I, we have

$$P_{a\to b}(t) = |c_b(t)|^2 = \frac{\pi q^2 t}{2\hbar^2 \epsilon_0 \omega^2 m^2 c} I(\omega) |M_{ba}|^2 \delta(\omega_{ba} - \omega) \tag{17.12}$$

Note that this probability is linearly dependent on time, therefore, the time-weighted probability of transition between the states $|a\rangle$ and $|b\rangle$ for a system in an electromagnetic field is

$$W_{ba} = \frac{\pi q^2}{\epsilon_0 \omega_{ba}^2 m^2 c\hbar^2} I(\omega_{ba}) |M_{ba}|^2 \tag{17.13}$$

§ 17.2 Dipole Approximation

In order to evaluate dipole transitions we approximate our plane wave solution to the first order, getting $e^{ik_i x^i} \approx 1$, which gives us that $M_{ba} = \langle b | \, \hat{\epsilon} \cdot \hat{\underline{p}} \, | a \rangle$. In this approximation, we can write

$$\hat{\underline{p}} = \frac{m}{i\hbar} \left[\hat{\underline{r}}, \hat{\mathcal{H}}_0 \right]$$

Substituting in M_{ba} and using $\hat{\mathcal{H}}_0$'s hermiticity, we have, defining an *electric dipole operator* $\hat{\underline{D}} = -e\hat{\underline{r}}$

$$M_{ba}^{(1)} = \frac{im\omega_{ba}}{e} \langle b | \, \hat{\epsilon} \cdot \hat{\underline{D}} \, | a \rangle \tag{17.14}$$

And

$$W_{ba}^{(1)} = \frac{\pi I(\omega_{ba})}{\hbar^2 c \epsilon_0} \left| \langle b | \, \hat{\epsilon} \cdot \hat{\underline{D}} \, | a \rangle \right|^2 \tag{17.15}$$

Evaluating the square norm on the right, we have that in case of unpolarized light, the vector ϵ lays randomly on a sphere, and therefore we have that the matrix elements of M_{ba} will depend solely on $|b\rangle \, \hat{\underline{r}} \, |a\rangle$, as follows

$$|\langle b | \, \hat{\epsilon} \cdot \underline{r} \, | a \rangle|^2 = \frac{1}{3} \sum_k |\langle b | \, \hat{r}_k \, | a \rangle|^2 = \frac{1}{3} |\underline{r}_{ba}|^2 \tag{17.16}$$

Where we used that $\langle \cos^2(\theta) \rangle = 1/3$.
Finally, we have that in the dipole approximation

$$W_{ba} = \frac{\pi q^2 I(\omega_{ba})}{3\hbar^2 \epsilon_0 c} |\langle b | \, \hat{\underline{r}} \, | a \rangle|^2 \tag{17.17}$$

§§ 17.2.1 Dipole Selection Rules for Atomic Transitions

Consider now a Hydrogenic Hamiltonian (without spin) as our unperturbed system. In this case we choose $q = -e$.
As we have seen before, the electric dipole transition probability between two states with $E_b > E_a$ depends only on the matrix elements of $\hat{\epsilon} \cdot \underline{r}$. Rewriting everything in terms of spherical components, we have

$$\hat{r}_{\pm 1} = \frac{\hat{x} \pm i\hat{y}}{\sqrt{2}}, \quad \hat{r}_0 = \hat{z}$$

$$\hat{\epsilon_{\pm 1}} = \frac{\hat{\epsilon}_{\mathbf{x}} \pm i\hat{\epsilon}_{\mathbf{y}}}{\sqrt{2}}, \quad \hat{\epsilon_0} = \hat{\epsilon}_{\mathbf{z}} \tag{17.18}$$

Considering that Spherical Harmonics are eigenfunctions of our unperturbed Hamiltonian and can also be used to describe the components of the vectors, we have $r_{\pm 1} = \sqrt{\frac{4\pi}{3}} Y_{1,\pm 1} \hat{r}$

and $\hat{r}_0 = \sqrt{\frac{4\pi}{3}} Y_{1,0} \hat{r}$, hence we get that $\hat{\epsilon} \cdot \underline{r}_{ba}$ is

$$\sum_q \langle n'l'm' | \overline{\hat{\epsilon}_{\mathbf{q}}} \hat{r}_q | nlm \rangle = \sum_q \overline{\hat{\epsilon}_{\mathbf{q}}} \sqrt{\frac{4\pi}{3}} \iint \overline{R}_{n'l'} R_{nl} r^3 \overline{Y}_{l'}^{m'} Y_1^q Y_l^m \, \mathrm{d}\Omega \, \mathrm{d}r \tag{17.19}$$

The last integral is nonzero if and only if $m' = m+q$, hence we must have that $m - m' = \pm 1$ and $q = \pm 1$. On the other hand, due to the parity of spherical harmonics, we need that $l + l' + 1$ must be even, hence summing the coefficients using Clebsch-Gordan rules, we need that $l = l' \pm 1$, and hence $l - l' = \pm 1$. In bra-ket notation this reduces to the following calculation

$$\int \overline{Y}_{l'}^{m'} Y_1^q Y_l^m \, \mathrm{d}\Omega \rightarrow \sqrt{\frac{3(2l+1)}{4\pi(2l'+1)}} \langle l100 | l'0 \rangle \langle l1mq | l'm' \rangle \tag{17.20}$$

Which brings back our previously found selection rules, plus one more constraint on Δm. Recapping everything, we get that for a spinless system, the selection rules for a dipole transition are

Quantum Number	Permitted Transitions
l	± 1
m	$0, \pm 1$

§§ 17.2.2 Spontaneous Emission

A situation that can't be evaluated using semiclassical methods is that of spontaneous emission, since an atom can decay spontaneously, even without having some radiation stimulating the process.

We start by evaluating an atom-photon system, for which we have three probabilities of interaction between the two: absorption (B_a), emission (B_e) and spontaneous emission (A_{se}). The Bs indicate the probability in unit time that the photon induces a transition between the two states b and a with $E_b > E_a$, and A indicates the probability for unit time that the state b decays spontaneously to the state a. We already know that the Bs are tied to our previous perturbation matrix as follows

$$W_{ba} = B_{ba} \rho(\omega_{ba})$$

Where ρ is a density tied to the photon number, which has the following formal expression

$$\rho(\omega) = \frac{\hbar \omega N(\omega)}{V}$$

With V the volume considered and $N(\omega)$ the number of photons in that given frequency. Given this, the number of atoms that decay from a to b is the following

$$N_{ba} = N_a B_a \rho(\omega_{ba})$$

And, vice-versa

$$N_{ab} = N_b B_e \rho(\omega_{ba}) + N_b A_{se}$$

If the system is at equilibrium we must have $N_{ab} = N_{ba}$, hence

$$\frac{N_b}{N_a} = \frac{B_a \rho(\omega_{ba})}{B_e \rho(\omega_{ba}) + A_{se}} = e^{-\beta \hbar \omega_{ba}} \tag{17.21}$$

Where $\beta = (k_B T)^{-1}$. Knowing that ρ must follow a Black-Body law, we have, solving for ρ

$$\rho(\omega_{ba}) = \frac{A_{se}}{B_a \left(e^{\beta \hbar \omega_{ba}} - \frac{B_e}{B_a} \right)} \tag{17.22}$$

Hence

$$\frac{B_e}{B_a} = 1$$
$$\frac{A_{se}}{B_a} = \frac{\hbar \omega_{ba}^3}{\pi^2 c^3} \tag{17.23}$$

We then must have that $W_{ba}^a = W_{ba}^{se} = B_a \rho(\omega_{ba})$, and coupling it to equation (17.13), we must have that

$$W_{ba}^{se} = \frac{\pi \omega_{ba}^3 e^2}{\hbar c^3 \epsilon_0} \left| \hat{\mathbf{epsilon}} \cdot \hat{r}_{ba} \right|^2 \tag{17.24}$$

§§ 17.2.3 Thermodynamic Equilibrium

We already saw how the density of photons at a given frequency is given by the following expression, without proof.

$$\rho(\omega) = \frac{\hbar \omega^3}{\pi^2 c^3} \frac{1}{e^{\beta \hbar \omega} - 1}$$

Proof. We need to prove this relation. We start supposing that we have a system of levels with $\Delta E = \hbar \omega$ for every level in the set. The number of photons that populate this system at a temperature T is

$$N(\omega, T) = \frac{\sum_{n=0}^{\infty} n e^{-\beta n \hbar \omega} - \frac{\beta \hbar \omega}{2}}{\sum_{n=0}^{\infty} e^{-\beta n \hbar \omega - \frac{\beta \hbar \omega}{2}}} = \frac{\sum_{n=0}^{\infty} n x^n}{\sum_{n=0}^{\infty} x^n} = -x \frac{d \log(1-x)}{dx} = \frac{1}{e^{\beta \hbar \omega} - 1} \tag{17.25}$$

Where we put, for convenience, $x = e^{-\beta \hbar \omega}$

We now need to evaluate how many states we have with frequency ω. We begin considering a box with sides of length L with periodic boundary value conditions. The number of possible modes is

$$N_m(k) = 2 \frac{\frac{4}{3} \pi k^3}{\left(\frac{2\pi}{L} \right)^3}$$

Differentiating and passing to frequency ($\omega = ck$), we get

$$\mathrm{d}N_m(\omega) = \frac{V\omega^2}{\pi^2 c^3}\,\mathrm{d}\omega \tag{17.26}$$

And, henceforth, the energy density is

$$\rho(\omega)\,\mathrm{d}\omega = \frac{\hbar\omega}{V}N(\omega, T)\,\mathrm{d}N_m(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3}\frac{1}{e^{\beta\hbar\omega} - 1}\,\mathrm{d}\omega \tag{17.27}$$

$\square$

§ 17.3 Relativistic Corrections and Fine Structure

Taking the full-blown relativistically invariant Dirac Hamiltonian and considering it in the limit $v/c << 1$ and assuming infinite nuclear mass, we get the following expression

$$\hat{\mathcal{H}} = mc^2 + \frac{\hat{p}^2}{2m} - \frac{\hat{p}^4}{8m^3c^2} + \frac{1}{2m^2c^2}\frac{1}{R}\frac{\mathrm{d}V}{\mathrm{d}R}\underline{\hat{L}} \cdot \underline{\hat{S}} + \frac{\hbar^2}{8m^2c^2}\nabla^2 V(R) + V(R) \tag{17.28}$$

This Hamiltonian can be divided in three parts: The non-relativistic Hamiltonian, a kinetic relativistic correction, spin-orbit interaction and the Darwin term.

§§ 17.3.1 Relativistic Correction for the Kinetic Energy

The relativistic correction for kinetic energy is given by the expansion of $E = c\sqrt{p^2 + m^2c^2}$. Expanding, we get

$$E \simeq mc^2 + \frac{\hat{p}^2}{2m} - \frac{\hat{p}^4}{8m^3c^2} \tag{17.29}$$

We have our non perturbed Hamiltonian $\hat{\mathcal{H}}_0$ with our hydrogenoid potential summed to the correction, as follows

$$\hat{\mathcal{H}} = \frac{\hat{p}^2}{2m} - \frac{Ze^2}{4\pi\epsilon_0 r} + W_p$$
$$W_p = -\frac{\hat{p}^4}{8m^3c^2} \tag{17.30}$$

We already know that $\hat{p}^2 = 2m\left(\hat{\mathcal{H}}_0 - V\right)$, hence $\hat{p}^4 = 4m^2\left(\hat{\mathcal{H}}_0 - V\right)^2$, and

$$W_p = -\frac{1}{2mc^2}\left(\hat{\mathcal{H}}_0 - V\right)^2$$

Surprisingly, this perturbation is diagonal in the basis of the Hydrogen atom, hence we find ourselves in need of calculating only some expectation values, as follows

$$\langle W_p \rangle = -\frac{1}{2mc^2}\left\langle \hat{\mathcal{H}}_0^2 + V^2 - \left[\hat{\mathcal{H}}_0, V\right]\right\rangle \tag{17.31}$$

Which gives

$$\langle W_p \rangle = -\frac{1}{2mc^2}\left(E_n^2 + \frac{Z^2 e^4}{(4\pi\epsilon_0)^2}\left\langle\frac{1}{r^2}\right\rangle + 2E_n\frac{Ze^2}{4\pi\epsilon_0}\left\langle\frac{1}{r}\right\rangle\right) \tag{17.32}$$

The expectation values are easy to calculate and their explicit calculation is given in appendix E. For our specific case, we get that

$$\left\langle\frac{1}{r}\right\rangle = \frac{Z}{a_0 n^2}$$

$$\left\langle\frac{1}{r^2}\right\rangle = \frac{Z^2}{a_0^2 n^3 \left(l+\frac{1}{2}\right)} \tag{17.33}$$

$$\left\langle\frac{1}{r^3}\right\rangle = \frac{Z^3}{a_0^3 n^3 l \left(l+\frac{1}{2}\right)(l+1)} \quad l > 0$$

Putting it all into our matrix elements of the perturbation, and remembering that $E_n = -E_1/n^2$, with $E_1 = -Z^2/n^2\alpha^2 mc^2$, with $\alpha = e^2/\hbar c$ is the fine structure constant, we get

$$\langle W_p \rangle = -\frac{1}{2mc^2}\left(E_n^2 + \frac{Z^4 e^4}{a_0^2 n^3(4\pi\epsilon_0)^2(l+1/2)} + 2E_n\frac{Z^2 e^2}{4\pi\epsilon_0 a_0 n^2}\right) \tag{17.34}$$

Substituting for E_n, we get, finally

$$\langle W_p \rangle = -E_n\frac{Z^2\alpha^2}{2n^2}\left(\frac{3}{4} - \frac{n}{l+1/2}\right) \tag{17.35}$$

For a Hydrogen atom we have $Z = 1$, and the correction is of order $\alpha^2 E_n$. Counting that $\alpha = 137^{-1}$ the perturbation is small enough to be treated as such.

§§ 17.3.2 Darwin Term

The Darwin term $\left(\frac{\hbar^2}{8m^2 c^2}\nabla^2 V(r)\right)$, for Hydrogenoid atoms, becomes

$$\hat{W}_D = \frac{\hbar^2 Z e^2}{8\epsilon_0 m^2 c^2}\delta(r_i) \tag{17.36}$$

Where we used that $\nabla^2(1/r) = -4\pi\delta(r)$
The matrix elements of $\hat{W}_D$ will be nonzero only for $l = 0$, and will take the following form

$$\langle W_D \rangle = \frac{\hbar^2 Z e^2}{8\epsilon_0 m^2 c^2}|\phi_{nlm}(0)|^2 \tag{17.37}$$

In general $|\phi_{n00}(0)|^2 = (4\pi)^{-1}|R_{n0}|^2$, where

$$R_{nl}(\rho) = \frac{1}{(2l+1)!}\sqrt{\left(\frac{2Z}{na_0}\right)^3 \frac{(n+l)!}{2n(n-l-1)!}}e^{\frac{\rho}{2}}r^l F(l+1-n, 2l+2, \rho) \tag{17.38}$$

With $\rho = 2Zr/na_0$. We have that $|R_{n0}(0)|^2 = \frac{4Z^3}{n^3 a_0^3}$, hence

$$|\phi_{n00}(0)|^2 = \frac{Z^3}{\pi a_0^3 n^3} \tag{17.39}$$

For which, finally we get

$$\langle W_D \rangle = \frac{\hbar^2 Z^4 e^2}{8\pi\epsilon_0 a_0^3 m^2 c^2 n^3} = \frac{\hbar^2}{2m^2 c^2} \frac{Z^2}{n^3 a_0^2} \alpha^2 m c^2 = -E_n \frac{\hbar^2 Z^2}{n a_0 m^2 c^2} = -E_n \frac{Z^2 \alpha^2}{n} \tag{17.40}$$

§§ 17.3.3 Spin Orbit Coupling

The last term to evaluate, is the Spin-Orbit coupling of angular momentums, which gives the following perturbation

$$\hat{W}_{SO} = \frac{1}{2m^2 c^2} \frac{1}{r} \frac{dV}{dr} \hat{\underline{L}} \cdot \hat{\underline{S}} \tag{17.41}$$

Explicitly

$$\hat{W}_{SO} = \frac{Ze^2}{8\pi\epsilon_0 m^2 c^2 r^3} \hat{\underline{L}} \cdot \hat{\underline{S}} \tag{17.42}$$

Its matrix elements will be, then

$$\langle W_{SO} \rangle = \frac{Ze^2}{8\pi\epsilon_0 m^2 c^2} \left\langle \frac{1}{r^3} \right\rangle \left\langle \hat{\underline{L}} \cdot \hat{\underline{S}} \right\rangle \tag{17.43}$$

Adding angular momentums and passing to the basis $|jlsm_j\rangle$, we get that

$$\hat{\underline{L}} \cdot \hat{\underline{S}} = \frac{1}{2} \left(\hat{J}^2 - \hat{L}^2 - \hat{S}^2 \right)$$

And remembering that

$$\left\langle \frac{1}{r^3} \right\rangle = \frac{Z^3}{a_0^3 n^3 l(l + 1/2)(l + 1)}, \quad l > 0$$

We get

$$\langle W_{SO} \rangle = \frac{\hbar^2 Z^4 e^2}{16\pi\epsilon_0 a_0^3 n^3 l(l + 1/2)(l + 1) m^2 c^2} (j(j + 1) - l(l + 1) - s(s + 1)) \tag{17.44}$$

In which, substituting in the following values

$$E_n = -\frac{Z^2 \alpha^2 m c^2}{2n^2}$$

$$a_0 = \frac{\hbar}{mc\alpha}$$

$$\frac{e^2}{4\pi\epsilon_0} = \alpha\hbar c$$

We have

$$\langle W_{SO} \rangle = -E_n \frac{Z^2 \alpha^2}{2n^2} \frac{j(j+1) - l(l+1) - s(s+1)}{l(l+1/2)(l+1)} \tag{17.45}$$

Which, dividing it in two cases, if we have $j = l + 1/2$ or $j = l - 1/2$, we get

$$\langle W_{SO} \rangle = -E_n \frac{Z^2 \alpha^2}{2nl(l+1/2)(l+1)} \cdot \begin{cases} l & j = l + 1/2 \\ -(l+1) & j = l - 1/2 \end{cases} \tag{17.46}$$

The total, and final, perturbation given by the relativistic approximation will not depend on l, and it's given by the following formula

$$E_{nj} = E_n \left(1 + \frac{Z^2 \alpha^2}{n^2} \left(\frac{n}{j+1/2} - \frac{3}{4} \right) \right) \tag{17.47}$$

§§ 17.3.4 Fine Structure Splitting

After summing all these perturbations to our initial energy, we have that although the non-relativistic energy levels were $2n^2$ times degenerate, we have that in the Dirac theory (i.e. relativistic quantum mechanics), we have that the $n-$th level splits in n different levels, each one with its own value of j. This splitting is commonly called *fine structure splitting*, and these n levels are called *fine structure multiplets*. The dimensionless *fine structure constant* $\alpha \simeq 1/137$ controls the scale of this splitting. It's important to note how in Dirac theory, two states with the same quantum numbers n, j but with $l = j \pm 1/2$, have the same energy, where the solution still has $(-1)^l$ parity. Thus for each j we have two series of $2j + 1$ solutions with opposite parity, except for $j = n - 1/2$, for which there is only a series of solutions with parity $(-1)^{n-1}$. The splitting we talked about in this paragraph is indicated as follows with spectroscopic notation.
e.g. let's say that we have $n = 3$, hence $l = 0, 1, 2$ and $j = 1/2$, $3/2$, $5/2$. Through the perturbed energy (17.47) we will have that the $n = 3$ level will split in 5 levels, as follows

$$3s_{1/2}, 3p_{1/2}, 3p_{3/2}, 3d_{3/2}, 3d_{5/2}$$

This splitting is accompanied by a further splitting, called *hyperfine splitting*, and its contribution is called *Lamb shift*

§§ 17.3.5 Fine Structure Dipole Transitions

We already seen how the only permitted transitions in the dipole approximation are those that have $\Delta l = \pm 1$ and $\Delta s = 0$, which implies that $\Delta j = 0, \pm 1$. Since $\hat{J}$ eigenstates are linear combinations of eigenstates of $\hat{S}, \hat{L}$, we have that $\Delta j = 0$ transitions are permitted. e.g., let's see how it works for states with $l = 1$ and $s = 1/2$. There will be 6 states, where 4 will have $j = 3/2$ and 2 will have $j = 1/2$. The six states in the $|jm_j\rangle$ basis can be written as a linear combination of states $|lsm_lm_s\rangle$. Starting with $m_j = 3/2$ we have

$$\left| \frac{3}{2}, \frac{3}{2} \right\rangle = \left| 1, \frac{1}{2}, 1, \frac{1}{2} \right\rangle$$

Operating with $\hat{J}_-$ we have

$$\left|\frac{3}{2}, \frac{1}{2}\right\rangle = A\left|1, \frac{1}{2}, 0, \frac{1}{2}\right\rangle + B\left|1, \frac{1}{2}, 1, -\frac{1}{2}\right\rangle$$

Analoguously, starting from $m_j = -3/2$ and working our way up the eigenstate ladder with $\hat{J}_+$, we get

$$\left|\frac{3}{2}, -\frac{1}{2}\right\rangle = C\left|1, \frac{1}{2}, 0, -\frac{1}{2}\right\rangle + D\left|1, \frac{1}{2}, -1, \frac{1}{2}\right\rangle$$

And finally, the last two states

$$\left|\frac{1}{2}, -\frac{1}{2}\right\rangle = -\sqrt{\frac{2}{3}}\left|1, \frac{1}{2}, -1, \frac{1}{2}\right\rangle + \sqrt{\frac{1}{\sqrt{3}}}\left|1, \frac{1}{2}, 0, -\frac{1}{2}\right\rangle$$

$$\left|\frac{1}{2}, \frac{1}{2}\right\rangle = E\left|1, \frac{1}{2}, 1, -\frac{1}{2}\right\rangle + F\left|1, \frac{1}{2}, 0, \frac{1}{2}\right\rangle$$

In this situation there are 7 possible transitions, but there will be only 5 visible lines, since $\Delta j = 2$ is not permitted by the selection rules.

This can be seen as follows: before the absorption, the total angular momentum is j_i, of the electron, summed with the photon spin $s_\gamma = 1$, hence the total (initial) angular momentum will be k_i, where $|j_i - 1| < k_i < j_i + 1$. After the transition, we must have $k_f = j_f$, and since angular momentum must be conserved, we have that $k_i = k_f$, hence $|j_i - 1| < j_f < j_i + 1$, this happens analogously with emission transitions. Finally, the selection rules for dipole transitions in fine structure systems are $\Delta l = \pm 1$, $\Delta s = 0$, $\Delta j = 0, \pm 1$ where $j = 0 \to j = 0$ is not permitted

§ 17.4 Zeeman Effect

Getting back to our semiclassical EM Hamiltonian, we have that the time dependent perturbation can be written as

$$-\frac{i\hbar e}{m}\underline{\hat{A}} \cdot \nabla = \frac{e}{2m}\underline{B} \cdot \hat{\underline{L}}$$

And the quadratic perturbation as follows

$$\frac{e^2}{2m}\hat{A}^2 = \frac{e^2}{8m}\left(B^2\hat{r}^2 - (\underline{B} \cdot \hat{\underline{r}})^2\right)$$

Since in laboratories is rare to exceed 10 T of magnetic field intensity, hence since $4a_0^2 B/4\hbar \approx 10^{-6}B$ we suppose that the quadratic term is again negligible.

We define a *Magnetic dipole moment operator* as follows.

$$\hat{\underline{M}} = -\frac{e}{2m}\hat{\underline{L}} = -\frac{\mu_B}{\hbar}\hat{\underline{L}} \tag{17.48}$$

With $\mu_B = e\hbar/2m$ is *Bohr's magneton*, which has the value of around $9.27408 \cdot 10^{24}$ J/T or Am^2. The interaction Hamiltonian with this field then is

$$\hat{\mathcal{H}}_m = -\underline{\hat{M}} \cdot \underline{B} \tag{17.49}$$

To this we have to add the intrinsic magnetic moment given by the electron, which has the following shape

$$\underline{\hat{M}}_s = -\frac{g\mu_B}{\hbar}\underline{\hat{S}} \tag{17.50}$$

Where $g = 2$ is the *gyromagnetic ratio* of the electron And the following Hamiltonian

$$\hat{\mathcal{H}}_s = -\underline{\hat{M}}_s \cdot \underline{B} \tag{17.51}$$

Our final Hamiltonian that accounts for spin-orbit interactions and atom-magnetic field interaction is simply given via the Pauli equation, at which has been "attached" the spin-orbit coupling element. The Pauli equation will then be

$$\langle r|\,\hat{\mathcal{H}}\,|\psi\rangle = \left(-\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r} + \xi(r)\underline{\hat{L}} \cdot \underline{\hat{S}} + \frac{\mu_B}{\hbar}\left(\underline{\hat{L}} + 2\underline{\hat{S}}\right) \cdot \underline{B}\right)\psi(r) = E\psi(r) \tag{17.52}$$

Where, our $\psi(r)$ not only accounts for spin, but for spin-orbit coupling also.

§§ 17.4.1 Strong Fields

Strong fields are characterized for having a magnetic field intensity of $B > Z^4$ Tesla. In this situation the spin-orbit coupling term is negligible, and our equation, after applying a rotation and having $\underline{B}||\hat{\mathbf{z}}$, we have

$$\left(-\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r}\right)\psi(r_i) = \left(E - \frac{\mu_B B_z}{\hbar}\left(\hat{L}_z + 2\hat{S}_z\right)\right)\psi(r_i) \tag{17.53}$$

The perturbation is diagonal, and the shift will be

$$E_{nm_l m_s} = E_n + \mu_B B_z(m_l + 2m_s) \tag{17.54}$$

Where in this case $m_s = \pm 1/2$. Since there is no spin orbit coupling in this case, we have that levels with $m_l = 1, m_s = -1/2$ and $m_l = -1, m_s = 1/2$ coincide.
We already know from the selection rules that we must have $\Delta m_s = 0$ and $\Delta m_l = 0, \pm 1$, thus splitting the transition $n \to n'$ into three components. The two components with $\Delta m_l = \pm 1$ are called π lines, and the remaining one, with $\Delta m_l = 0$ is called the σ line. These π transitions have the following frequencies

$$\omega_{n'n}^{\pm} = \omega_{n'n} \pm \omega_L \tag{17.55}$$

Where $\omega_L = \mu_B B_z/\hbar$ is the *Larmor frequency*. This effect is known as *normal Zeeman effect*, and the π and σ lines of this effect are said to be *Lorentz triplets*. They are also observed in atoms for which $S = 0$, hence where spin-orbit coupling is absent.

§§ 17.4.2 Paschen-Back Effect

At field strengths for which spin orbit coupling is appreciable but still small with respect to the field intensity B, we have that first order perturbation theory can be applied in order to calculate the energy shift.
We have then

$$\Delta E = \int_0^\infty r^2 R_{nl}^2 \xi(r) \, \mathrm{d}r \left\langle l \frac{1}{2} m_l m_s \left| \hat{\underline{L}} \cdot \hat{\underline{S}} \right| l \frac{1}{2} m_l m_s \right\rangle = \lambda_{nl} m_l m_s \quad l \neq 0 \qquad (17.56)$$

We have that<++>

$$\lambda_{nl} = -E_n \frac{Z^2 \alpha^2}{n} \frac{1}{l(l+1/2)(l+1)} \qquad (17.57)$$

Which removes the degeneracy in l. The energy difference between two different levels, when $m_s = m_s'$ ($\Delta m_s = 0$) is then the following

$$\Delta E = E_{n'} - E_n + E_z + E_{pb} = E_{n'} - E_n + \mu_B B_z (m_l' - m_l) + (\lambda_{n'l'} m_l' - \lambda_{nl} m_l) m_s \quad (17.58)$$

§§ 17.4.3 Anomalous Zeeman Effect

For weak magnetic fields, we have what's usually called the *anomalous Zeeman effect*. In this case, the spin-orbit coupling is the dominant term, and our unperturbed Hamiltonian takes the following shape

$$\hat{\mathcal{H}}_0 = -\frac{\hbar^2}{2m} \nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r} + \xi(r) \hat{\underline{L}} \cdot \hat{\underline{S}} \qquad (17.59)$$

This Hamiltonian has exact wavefunctions, which are given via what's known as *Tensor Spherical Harmonics* (see appendix C). Our unperturbed wavefunction will then be

$$\psi_{nlsjm_j}(r,\theta,\phi) = R_{nl}(r) \mathcal{Y}_{ls}^{jm_j}(\theta,\phi) = C_{jm_j}^{lsm_l m_s} R_{nl}(r) Y_l^{m_l}(\theta,\phi) \chi_{s,m_s} \qquad (17.60)$$

Or, in Dirac notation

$$|nl\rangle \otimes |ls\rangle \otimes |jm_j\rangle = \langle lsm_l m_s | \left(|jm_j\rangle \otimes |nl\rangle \otimes |lm_l\rangle \otimes |sm_s\rangle \right) \qquad (17.61)$$

Taking the magnetic field parallel to the z axis, we have that the perturbation needed to be evaluated is

$$\hat{\mathcal{H}}' = \frac{\mu_B}{\hbar} \left(\hat{J}_z + \hat{S}_z \right) B_z \qquad (17.62)$$

And, the perturbation on the energy levels is

$$\Delta E = \mu_B m_j B_z + \frac{\mu_B}{\hbar} B_z \sum_{m_s} \int \overline{\mathcal{Y}}_{l,1/2}^{jm} \hat{S}_z \mathcal{Y}_{l,1/2}^{jm} \, \mathrm{d}\Omega \qquad (17.63)$$

From the properties of tensor spherical harmonics, we have that the integral is of easy computation, and gives the following result

$$\int \overline{\mathcal{Y}}_{l,1/2}^{l\pm1/2,m_j} \hat{S}_z \mathcal{Y}_{l,1/2}^{l\pm1/2,m_j} \, \mathrm{d}\Omega = \pm\frac{\hbar m_j}{2l+1} \tag{17.64}$$

Which gives our searched energy shift

$$\Delta E = \mu_B m_j B_z + \frac{\mu_B B_z}{2l+1} \sum_{m_s} m_j \tag{17.65}$$

Finally, getting back to our Hydrogenoid atoms, we now can write the total perturbation of the energy levels for an interaction with a constant magnetic field as follows

$$E_{njm_j} = E_n + \Delta E_{nj} + \Delta E_{m_j} \tag{17.66}$$

Where E_n is the unperturbed energy, ΔE_{nj} is the fine structure correction and ΔE_{m_j} is the weak field correction

§ 17.5 Stark Effect

The splitting of energy levels given by static electric fields is called *Stark effect*. We assume that the electric field is perpendicular to our z axis, and that the field strength is much larger than Spin-Orbit coupling.
The perturbation acting on our Hydrogenic Hamiltonian is

$$\hat{\mathcal{H}}' = eEz \tag{17.67}$$

§§ 17.5.1 Linear Stark Effect

In order to evaluate Stark shifts with perturbations at first order, we start by calculating the perturbation given to the fundamental state.
Its first order correction is

$$E_{100}^{(1)} = eE \langle 100 | \hat{z} | 100 \rangle \tag{17.68}$$

We already see that $E_{100}^{(1)} = 0$, since the integral is null, due to it being the product of an even function ($\langle x_i | nlm \rangle$) with an odd one ($z$). We therefore check for excited states, starting at $n = 2$, for which we have a fourfold degeneration given by l, with energy $E_{2lm} = -mc^2\alpha^2/8$.
In this case we have that $\langle nlm | \hat{z} | n'l'm' \rangle$ doesn't vanish, if and only if $\Delta m = 0$ and $\Delta l = \pm 1$, hence, the permitted transitions are only those between $2s$ and $2p$ states, for which our perturbation (which is real, hence Hermitian) gives the following result

$$E_{2s\to2p}^{(1)} = eE \int \overline{\psi}_{210}(x_i) z \psi_{200}(x_i) \, \mathrm{d}^3 x_i \tag{17.69}$$

This last integral gives us that $E^{(1)} = \pm 3eEa_0/Z$, for which we have two different energy levels, described by the following states

$$|1\rangle = \frac{1}{\sqrt{2}} \left(|200\rangle + |210\rangle \right)$$

$$|2\rangle = \frac{1}{\sqrt{2}} \left(|200\rangle - |210\rangle \right)$$

(17.70)

§ 17.6 Hyperfine Structure and Isotope Shifts

Hyperfine structure of energy levels appears when nucleii aren't considered anymore point charges with infinite mass. The name, *hyperfine* has been given to this structure, since the deviation of the energy levels due to this additional perturbation is much smaller than that of the fine structure shifts.

In general we can have two kinds of hyperfine effects: *isotope shifts*, that slightly deviate degenerate energy levels without splitting them, and proper hyperfine effects that break the degeneration on the levels and splits them.

§§ 17.6.1 Magnetic Dipole Hyperfine Structure

As for the electrons, we can define a *nuclear* spin, indicated with $\hat{I}$, which obey the same spin algebra of ordinary spin. As with electrons, the eigenvalues for nuclear spin can be half-integer, if the sum of the spins of the nucleons is fermionic, or integer, if the sum of the spins of the nucleons is bosonic. We denote the eigenvalues of $\hat{I}^2$ as $\hbar^2 I(I+1)$ and of $\hat{I}_z$ as $\hbar M_I$.

We can define, as for electrons, a *Nuclear Magnetic Dipole Moment*, $\hat{M}_I$ as follows

$$\hat{M}_I = \frac{g_I \mu_N}{\hbar} \hat{I}$$

(17.71)

Where g_I is the *Landé factor*, and μ_N is the *Nuclear magneton*, defined as

$$\mu_N = \frac{e\hbar}{2m_p} = \frac{m_e}{m_p} \mu_B$$

(17.72)

We now proceed to write our perturbed Hamiltonian as follows

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{\mathcal{H}}_{ND}$$

(17.73)

Since we have already solved for the fine structure Hamiltonian, it is included into the unperturbed Hamiltonian together with the Coulomb interaction.

At the zeroth order we have that the wavefunctions of $\hat{\mathcal{H}}$ are separable in electronic and nuclear variables, and are eigenfunctions of $\hat{J}^2, \hat{J}_z, \hat{I}^2, \hat{I}_z$. These wavefunction are $(2j+1)(2I+1)$ degenerate in m_j, M_I.

Examining the perturbation $\hat{\mathcal{H}}_{ND}$, we have that it will couple with both $\underline{\hat{L}}$ and $\underline{\hat{S}}$ of the electrons, hence splitting the Hamiltonian in a sum. Since this point dipole is located at the origin, we have that

$$\hat{\mathcal{H}}_{NL} = -\frac{i\hbar e}{m}\underline{\hat{A}} \cdot \nabla = \frac{\mu_0}{2\pi\hbar^2 r^3}g_I\mu_B\mu_N\underline{\hat{L}} \cdot \underline{\hat{I}} \tag{17.74}$$

For the spin-spin coupling Hamiltonian we have that, in terms of magnetic fields we can write the perturbation term as

$$\hat{\mathcal{H}}_{SS} = -\underline{\hat{M}}_s \cdot \underline{B} \tag{17.75}$$

Where

$$\underline{B} = -\frac{\mu_0}{4\pi}\left(\underline{\hat{M}}_N\nabla^2\frac{1}{r} - \nabla(\underline{\hat{M}}_N \cdot \nabla)\frac{1}{r}\right) \tag{17.76}$$

Hence

$$\hat{\mathcal{H}}_{SS} = -\frac{2\mu_0}{4\hbar^2\pi}g_I\mu_B\mu_N\left(\underline{\hat{S}} \cdot \underline{\hat{I}}\nabla^2\frac{1}{r} - (\underline{\hat{S}} \cdot \nabla)(\underline{\hat{I}} \cdot \nabla)\frac{1}{r}\right) \tag{17.77}$$

At $r = 0$ this perturbation will act only on s-states, and for $r \neq 0$ we get

$$\hat{\mathcal{H}}_{SS} = -\frac{\mu_0}{4\pi r^3}\left(\underline{\hat{M}}_s \cdot \underline{\hat{M}}_N - \frac{3}{r^2}\left((\underline{\hat{M}}_s \cdot \hat{r})(\underline{\hat{M}}_N \cdot \hat{r})\right)\right)$$

Summing the two terms, we have that $\hat{\mathcal{H}}_{ND} = \hat{\mathcal{H}}_{NL} + \hat{\mathcal{H}}_{SS}$, and we finally have

$$\hat{\mathcal{H}}_{ND} = \frac{\mu_0}{2\hbar^2\pi}g_I\mu_B\mu_N\frac{1}{r^3}\left(\underline{\hat{L}} \cdot \underline{\hat{I}} - \underline{\hat{S}} \cdot \underline{\hat{I}} + 3\frac{(\underline{\hat{S}} \cdot \hat{r})(\underline{\hat{I}} \cdot \hat{r})}{r^2}\right) r \neq 0 \tag{17.78}$$

Simplifying everything, we have, at $r = 0$

$$\hat{\mathcal{H}}_{ND} = -\frac{2\mu_0}{3\pi}\underline{\hat{M}}_s \cdot \underline{\hat{M}}_N\delta(\hat{r}) \ r = 0 \tag{17.79}$$

This last expression is called *Fermi contact interaction*. Another way to write (17.78) is to define the new operator

$$\underline{\hat{G}} = \underline{\hat{L}} - \underline{\hat{S}} + 3\frac{(\underline{\hat{S}} \cdot \hat{r})\hat{r}}{r^2} \tag{17.80}$$

And (17.78) becomes

$$\hat{\mathcal{H}}_{ND} = \frac{\mu_0}{2\hbar^2\pi}g_I\mu_B\mu_N\frac{1}{r^3}\underline{\hat{G}} \cdot \underline{\hat{I}} \tag{17.81}$$

Before proceeding in the calculation of the energy shifts, we define a *total angular momentum* for the whole system as $\underline{\hat{F}} = \underline{\hat{I}} + \underline{\hat{J}}$, and the energy shift will be given in the total angular momentum basis as follows

$$\Delta E = \frac{\mu_0}{2\hbar^2\pi}g_I\mu_B\mu_N\langle lsjIFM_F|\frac{1}{r^3}\underline{\hat{G}} \cdot \underline{\hat{I}}|lsjIFM_F\rangle \tag{17.82}$$

We can write $\hat{\underline{G}} \cdot \hat{\underline{I}}$ as $(\hat{\underline{G}} \cdot \hat{\underline{J}})(\hat{\underline{I}} \cdot \hat{\underline{J}})/\hbar^2 j(j+1)$, and knowing that $\hat{\underline{I}} \cdot \hat{\underline{J}} = 1/2 \left(\hat{F}^2 - \hat{I}^2 - \hat{J}^2 \right)$ we have the total shift as

$$\Delta E = \frac{C}{2} \left(F(F+1) - I(I+1) - j(j+1) \right) \tag{17.83}$$

Where

$$C = \frac{\mu_0}{2\hbar^2 j(j+1)\pi} g_I \mu_B \mu_N \left\langle \frac{1}{r^3} \hat{\underline{G}} \cdot \hat{\underline{J}} \right\rangle \tag{17.84}$$

Noting that $\hat{\underline{G}} \cdot \hat{\underline{J}} = \hat{L}^2$, we finally get our solution, having put a_μ as $a_0 m/\mu$, with μ the reduced mass of the electron and the nucleus

$$\Delta E_{ND} = \frac{\mu_0}{4\pi} g_I \mu_B \mu_N \frac{l(l+1)}{j(j+1)} \frac{Z^3}{a_\mu^3 n^3 l(l+1/2)(l+1)} \left(F(F+1) - I(I+1) - j(j+1) \right) l \neq 0 \tag{17.85}$$

For $l = 0$ we instead get

$$\Delta E_{ND} = \frac{2\mu_0}{3\pi} g_I \mu_B \mu_N \frac{Z^3}{a_\mu^3 n^3} \left(F(F+1) - I(I+1) - s(s+1) \right) \tag{17.86}$$

In atomic units, the final result will then be

$$\Delta E_{hyp} = \frac{m_e}{m_p} \left(\frac{\mu}{m_e} \right)^3 \frac{2g_I Z^3 \alpha^2}{n^3 (j+1)(2l+1)} \cdot \begin{cases} I + \dfrac{1}{2} & j \leq I \\ \dfrac{I(j+1/2)}{j} & j \geq I \end{cases} \tag{17.87}$$

We see that the total angular momentum quantum number behaves as $\hat{\underline{J}}$, hence $\Delta F = 0, \pm 1$ are the only permitted transitions, with $0 \rightarrow 0$ being excluded

18 Two Electron Atoms

We start by directly writing the Schrödinger equation for a two electron atom in atomic units, where

$$\hbar = 1$$

$$k_e = \frac{1}{4\pi\epsilon_0} = 1$$

$$e = 1$$

We immediately have

$$\hat{\mathcal{H}} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{Z}{r_1} - \frac{Z}{r_2} + \frac{1}{r_{12}} \tag{18.1}$$

In this case we have that the wavefunction is simmetric to spatial exchange of the two electrons (also called para-wavefunction, similarly spatially antisymmetric wavefunction are called ortho-wavefunctions).
We also must impose the Pauli exclusion principle, by taking into account the electrons' spin. We end up having our wavefunction as

$$\Psi(q_1, q_2) = \psi(x_i^1, x_i^2)\chi_{1/2, m_s}(1, 2) \tag{18.2}$$

As we know already, the basis spinor wavefunction for a system of two electrons can be either antisymmetric (singlet) or symmetric (triplet), and hence can take the following shapes

$$|00\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle\,|\downarrow\rangle - |\downarrow\rangle\,|\uparrow\rangle\right)$$

$$|11\rangle = |\uparrow\rangle\,|\uparrow\rangle$$

$$|10\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle\,|\downarrow\rangle + |\downarrow\rangle\,|\uparrow\rangle\right)$$

$$|1\downarrow 1\rangle = |\downarrow\rangle\,|\downarrow\rangle$$

Due to Pauli's exclusion principle, we have that the final wavefunction must be completely antisymmetric, hence, if we have a para wavefunction ψ_+, the final solution will take the shape

$$\Psi(q_1, q_2) = \psi_+(x_i^1, x_i^2)\chi_{00} \tag{18.3}$$

Or analoguously, if we have an ortho wavefunction

$$\Psi(q_1, q_2) = \psi_-(x_i^1, x_i^2) \cdot \begin{Bmatrix} \chi_{11} \\ \chi_{10} \\ \chi_{1-1} \end{Bmatrix} \tag{18.4}$$

Written this, one may immediately ask what's then the scheme of energy levels? As an example we take an Helium atom. We have that $Z = 2 << 40$, hence we have a set of levels of an almost independent levels, one of ortho-triplet states and one of para-singlet states. The lowest energy levels, are divided then between singlets ($S = 0$) and triplets ($S = 1$), where S is the total spin.
Let L be the sum of the eigenvalues of the square angular momentum of both electrons and M_L the sum of the eigenvalues of the projection. We can introduce the *spectroscopic terms* as a new notation for atomic energy levels. They're written as follows

$$^{2S+1}L_{M_J} \tag{18.5}$$

The term L, analogously to the term l for particles, takes the "values" $S, P, D, F, G, H, \cdots$, as l, which takes values $s, p, d, f, g, h, \cdots$. In addition, on the top left of the term there is indicated the *multiplicity of the state*, which indicates whether it's a singlet or a triplet state, in terms of total spin. On the bottom right there are indicated the possible M_J values of the considered system.

§ 18.1 Independent Particle Model

In order to get a first approach to two-electron atoms, we need to develop an approximate theory, in which the $e^- - e^-$ interaction is taken as a perturbation on the system

$$\hat{\mathcal{H}} = -\frac{1}{2}\nabla_1^2 - \frac{Z}{r_1} - \frac{1}{2}\nabla_2^2 - \frac{Z}{r_2} + \hat{\mathcal{H}}'$$

$$\hat{\mathcal{H}}' = \frac{1}{r_{12}} \tag{18.6}$$

$$\hat{\mathcal{H}} = \hat{h}_1 + \hat{h}_2 + \hat{\mathcal{H}}' = \hat{\mathcal{H}}_0 + \hat{\mathcal{H}}'$$

In this case we have that the eigenvalues of the single electron Hamiltonian ($\hat{h}_i$) are known to be the following

$$\hat{h}_i \psi_{nlm_i}(x_j^i) = -\frac{Z^2}{2n_i^2}\psi_{nlm_i}(x_j^i) \quad i = 1, 2; \; j = 1, 2, 3$$

Hence, in general, we then have the following solution for the unperturbed Hamiltonian

$$\hat{\mathcal{H}}_0 \psi^0(x_i^1, x_i^2) = -\frac{Z^2}{2}\left(\frac{1}{n_1^2} + \frac{1}{n_2^2}\right)\psi_{n_1 l_1 m_1}(x_i^1)\psi_{n_2 l_2 m_2}(x_i^2) \tag{18.7}$$

This solution adds up a new exchange degeneracy of the state. We already know tho that the final wavefunction must be the symmetrization or antisymmetrization of the two single electron wavefunctions. In Dirac notation, using $|1, 2\rangle$ for the full wavefunction and $|i\rangle$ for the single electron wavefunction, we must then have

$$|1, 2\rangle_{\pm} = \frac{1}{\sqrt{2}} \left(|1\rangle |2\rangle \pm |2\rangle |1\rangle \right) \tag{18.8}$$

Therefore, we can take the antisymmetric wavefunctions (ψ^0_-) as an approximation for ortho states and (ψ^0_+) for para states.

This consideration lets the ortho wavefunction vanish for the ground state of the system, in accord with the Pauli exclusion principle, hence there can exist a single para-singlet state as ground state, for which, the associated wavefunction is the following

$$\langle 1, 2 | 1s, 1s \rangle = \psi^0_{GS} = \psi_{1s}(x^1_i)\psi_{1s}(x^2_i) = \frac{Z^3}{\pi} e^{-Z(r_1+r_2)} \tag{18.9}$$

This result, gives then for Helium $(Z = 2)$ $E^0_{GS} = -Z^2 = -4$ a.u. $(-108.8$ eV), which corresponds to a ionization potential of $I_p = Z^2/2 = 2$ a.u $(54.4$ eV). The experimental values are not in accord with these results, which are $E^{exp}_{GS} = -2.90$ a.u. $(-79.0$ eV) and a ionization potential $I^{exp}_p = 0.90$ a.u. $(24.6$ eV).

This final result is quite far from the experimental result, since we neglected the $e^- - e^-$ interaction.

§§ 18.1.1 Central Field Approximation

We now take again our equation (18.6) and we take a new unperturbed Hamiltonian, which is the sum of the two single particle hamiltonians

$$\hat{h}'_i = -\frac{1}{2}\nabla^2_i + V(r_i) \tag{18.10}$$

And we modify the perturbation, getting this new formula

$$\hat{\mathcal{H}}' = \frac{1}{r_{12}} - \frac{Z}{r_1} - V(r_1) - \frac{Z}{r_2} - V(r_2) \tag{18.11}$$

In this new representation of the problem, we need to choose a central potential $V(r_i)$, for which we have that $\hat{\mathcal{H}}'$ is small enough. We choose his potential to be the following

$$V(r) = -\frac{Z - S}{r} = -\frac{Z_e}{r}$$

Where S is some constant, called *screening constant*. What we indicated with Z_e is commonly considered as an *effective charge*.

Since the new potential is a Coulomb potential, we already know the ground state wavefunction and the energy levels, since we might simply insert the substitution $Z \to Z_e$ in what we have already found.

Writing u_{nlm} as the single particle equation, we will have that the single-particle energy levels will depend directly on l, breaking the degeneracy on that quantum number. Our complete wavefunction for discrete excited states will then be the following

$$\psi'_{\pm}(x_i^1, x_i^2) = \frac{1}{\sqrt{2}} \left(u_{100}(x_i^1) u_{nlm}(x_i^2) \pm u_{nlm}(x_i^1) u_{100}(x_i^2) \right) \tag{18.12}$$

The total energy of the atom will then be, simply

$$E_{nl}^0 = E_{1s} + E_{nl} \tag{18.13}$$

Where E_{nl} depends on the chosen potentia $V(r_i)$, and is $E_{nl} = -Z_{nl}^2/n^2$.
In the special case of a completely screened electron we will have $S = 1$, and the energy levels will then be described, in the non-interacting electron approximation as follows

$$E_n^0 = -\frac{Z^2}{2} - \frac{(Z-1)^2}{2n^2} \tag{18.14}$$

§ 18.2 Ground State of Two Electron Atoms

§§ 18.2.1 Perturbation Theory

We have already seen that the ground state eigenket of two electron atoms, in its most general expression can be written as follows

$$|q_1, q_2\rangle = \frac{1}{\sqrt{2}} |1, 2\rangle \left(|\uparrow\rangle_1 |\downarrow\rangle_2 - |\downarrow\rangle_1 |\uparrow\rangle_2 \right) \tag{18.15}$$

Using perturbation theory, we can already get a nice guess of the ground state correction, via the calculus of the matrix elements $\langle 1s, 1s| \hat{\mathcal{H}}' |1s, 1s\rangle$, where $\hat{\mathcal{H}}' = r_{12}^{-1}$. Writing explicitly the integral, we get

$$E_0^1 = \int_0^\infty |\psi_{1s}(r_1)|^2 \frac{1}{r_{12}} |\psi_{1s}(r_2)|^2 \, dr_1 dr_2 \tag{18.16}$$

Using the following conversion, and the connection between Legendre polynomials and spherical harmonics, we have that, firstly

$$\frac{1}{r_{12}} = \sum_{l=0}^\infty \frac{\min^l(r_1, r_2)}{\max^{l+1}(r_1, r_2)} P_l(\cos\theta)$$

Hence

$$\frac{1}{r_{12}} = \sum_{l=0}^\infty \sum_{m=-l}^l \frac{4\pi}{2l+1} \frac{\min^l(r_1, r_2)}{\max^{r+1}(r_1, r_2)} \overline{Y}_l^m(\theta_1, \phi_1) Y_l^m(\theta_2, \phi_2)$$

Plugging it into the integral, and using the properties of spherical harmonics, we get

$$E_0^1 = \frac{Z^6}{\pi^2} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{16\pi^2}{2l+1} \int_0^{\infty} r_1^2 \, \mathrm{d}r_1 \int_0^{\infty} e^{-2Z(r_1+r_2)} \frac{r_m^l}{r_M^{l+1}} \delta_{l0} \delta_{m0} \, \mathrm{d}r_2$$

Contracting the indices with the sums and the Kronecker deltas, we have our final integral

$$E_0^1 = 16Z^6 \int_0^{\infty} e^{-2Zr_1} r_1^2 \, \mathrm{d}r_1 \left(\frac{1}{r_1} \int_0^{r_1} e^{-2Zr_2} r_2^2 \, \mathrm{d}r_2 + \int_{r_1}^{\infty} e^{-2Zr_2} r_2 \, \mathrm{d}r_2 \right) \quad (18.17)$$

The integrals give the value $5/128Z^5$, hence, we get our correction

$$E_0^1 = \frac{5}{8}Z \quad (18.18)$$

The final approximate energy is then

$$E_0 \approx E_0^0 + E_0^1 = -Z^2 + \frac{5}{8}Z \quad (18.19)$$

§§ 18.2.2 Variational Methods

Proceding instead using variational calculus, we set our $|1s\rangle$ wavefunction to be dependent from a parameter Z_e, which will be needed to minimize the following functional

$$E[\phi] = \frac{\langle \phi | \hat{\mathcal{H}} | \phi \rangle}{\langle \phi | \phi \rangle} \quad (18.20)$$

Where we set $|\phi\rangle = |1s\rangle_1 \otimes |1s\rangle_2$.
Calculating and introducing here the quantum virial theorem, we get that

$$\langle \phi | \frac{\hat{p}_1^2}{2m} | \phi \rangle = \langle \phi | \frac{\hat{p}_2^2}{2m} | \phi \rangle = \langle 1s | \frac{\hat{p}^2}{2m} | 1s \rangle = \frac{1}{2}Z_e^2 \quad (18.21)$$

We then also have

$$-\langle \phi | \frac{Z}{r_1} + \frac{Z}{r_2} | \phi \rangle = -2ZZ_e \quad (18.22)$$

And, due to our previous calculations

$$\langle \phi | \frac{1}{r_{12}} | \phi \rangle = \frac{5}{8}Z_e \quad (18.23)$$

The final result is our energy as a function of the parameter Z_e, which is

$$E(Z_e) = Z_e^2 - 2ZZ_e + \frac{5}{8}Z_e \quad (18.24)$$

We search for an extremal of this function deriving with respect to the parameter, and we get

$$\frac{\partial E}{\partial Z_e} = 2Z_e - 2Z + \frac{5}{8}$$
$$Z_e = Z - \frac{5}{16}$$
(18.25)

Hence, since our energy was $E_0 = -Z_e^2$, we get, in atomic units, that for Helioid atoms, the energy of the ground state will be approximated as follows

$$E_0^{GS} = -\left(Z - \frac{5}{16}\right)^2 = -Z^2 + \frac{5}{8}Z - \frac{25}{256}$$
(18.26)

Plugging in the values for Helium, we find that $E_0 = 2.84766$ au, a value that differs only by the 0.056 au from the experimental value.

Another way of finding a more precise ground state energy has been found by Hylleraas, where another set of coordinates is chosen, and the trial wavefunction has multiple variational parameters. These are the following coordinates

$$s = r_1 + r_2$$
$$t = r_1 - r_2$$
$$u = r_{12} = \|\underline{r}_1 - \underline{r}_2\|$$

The trial wavefunction is the following

$$\phi(s, t, u) = e^{-ks} \sum_{l=0}^{N} \sum_{m=0}^{N} \sum_{n=0}^{N} c_{l,2m,n} s^l t^{2m} u^m$$

Where $c_{l,2m,n}$ are linear variational parameters, and k is a nonlinear variational parameter that behaves like the effective charge Z_e.

Since the ground state must be a para state, this wavefunction must be an even function of t. Calculating for $N = 5$, we have 6 variational parameters, we have that the energy calculated is -2.90324 au, which is surprisingly close to the experimental value, differing only by $4.8 \cdot 10^{-4}$ au (0.013 eV).

§ 18.3 Excited States of Two Electron Atoms

We will now treat the excited states of Helioid atoms, with the same methods used for evaluating the ground states.

§§ 18.3.1 Perturbation Theory

Firstly, we apply perturbation theory to our system, formed by the sum $\hat{\mathcal{H}}_0 + \hat{\mathcal{H}}'$, with $\hat{\mathcal{H}}' = r_{12}^{-1}$.

We already know that these states must either be parastates or orthostates, hence we will

need our wavefunction to be either exchange-symmetric or exchange-antisymmetric. The energy perturbation will then depend on the sign of the wavefuction. We also know that $\left[\hat{\mathcal{H}}', \hat{L}_z\right] = 0$, hence we can write the following result

$$E_{\pm}^1 = \langle \psi_{\pm}^0 | \, \hat{\mathcal{H}}' \, | \psi_{\pm}^0 \rangle = J \pm K \tag{18.27}$$

Where

$$J = \int_{\mathbb{R}^3 \times \mathbb{R}^3} |\psi_{100}(\underline{r}_1)|^2 \frac{1}{r_{12}} |\psi_{nlm}(\underline{r}_2)|^2 \; \mathrm{d}^3 r_1 \mathrm{d}^3 r_2$$

$$K = \int_{\mathbb{R}^3 \times \mathbb{R}^3} \overline{\psi_{100}}(\underline{r}_1) \overline{\psi_{nlm}}(\underline{r}_2) \frac{1}{r_{12}} \psi_{100}(\underline{r}_2) \psi_{nlm}(\underline{r}_1) \; \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \tag{18.28}$$

Where $n \geq 2$. The integral J is called *Coulomb or direct integral* and K is called *Exchange integral*.

Expanding r_{12}^{-1} in spherical harmonics, we get

$$J_{nl} = \int_0^\infty R_{nl}(r_2) r_2^2 \; \mathrm{d}r_2 \int_0^\infty R_{10}^2(r_1) \frac{r_1^2}{\max(r_1, r_2)} \; \mathrm{d}r_1$$

$$K_{nl} = \frac{1}{2l+1} \int_0^\infty R_{10}(r_2) R_{nl}(r_2) r_2^2 \; \mathrm{d}r_2 \int_0^\infty R_{10}(r_1) R_{nl}(r_1) \frac{\min^l(r_1, r_2)}{\max^{l+1}(r_1, r_2)} r_1^2 \; \mathrm{d}r_1 \tag{18.29}$$

Hence, $E_{\pm}^1 = E_{nl\pm}^1$, so the energy after this correcion directly depends on n, l.
At first order, we can then write

$$E_{nl\pm} \approx E_{1n}^0 + E_{nl\pm}^1 = -\frac{Z^2}{2}\left(1 + \frac{1}{n^2}\right) + J_{nl} \pm K_{nl} \tag{18.30}$$

In order to see how this perturbation acts, we can see that J_{nl} must be always positive. Seeing that for $l = n-1$, $R_{n,n-1}$ has no nodes, hence $K_{n,n-1} > 0$. It can also be seen that, in general $K_{nl} > 0$, so, from (18.30) we see that an orthostate has an energy lower to the corresponding parastate.
This can be seen introducing spin into our calculus, as follows.

$$E_{nl\pm}^1 = J_{nl} - \frac{1}{2}\left(1 + \sigma_i^1 \sigma_2^i\right) K_{nl} \tag{18.31}$$

Where σ_i^j are the pauli matrices of the two electrons.

§§ 18.3.2 Variational Methods

Variational methods can be applied supposing that the higher-order wavefunction are orthogonal to the ground-state trial function and then calculate the variational integrals. This method is best applied on singular states.
Starting with the $2^3 S$ state of the atom, we have that, using variational parameters Z_i, Z_o as the inner and outer effective charge, we can write the state as follows

$$\left|2^3 S\right\rangle = N \left(|1s\rangle |2s\rangle - |2s\rangle |1s\rangle\right) \tag{18.32}$$

Where

$$u_{1s}(r) = e^{-Z_i r}$$
$$v_{2s}(r) = \left(1 - \frac{Z_o r}{2}\right) e^{-\frac{Z_o r}{2}}$$

(18.33)

Substituting into the variational integral and finding the minimum, we get $Z_i = 2.01$ au and $Z_o = 1.53$ au, which yeld $E_{2^3 S} = -2.167$ au. Analogously, one can use Hylleraas wavefunction in order to get a more precise result.

For $2^1 P$ and $2^3 P$ states, we use the following state kets

$$\left|2^1 P\right\rangle = N_+ \left(|1s\rangle_1 |2pm\rangle_2 + |2pm\rangle_1 |2s\rangle_2\right)$$
$$\left|2^3 P\right\rangle = N_- \left(|1s\rangle_1 |2pm\rangle_2 - |2pm\rangle_1 |2s\rangle_2\right)$$

(18.34)

With wavefunctions

$$u_{1s}(r) = e^{-Z_i r}$$
$$v_{2pm}(r) = re^{\frac{Z_o r}{2}} Y_m^1(\hat{\mathbf{r}})$$

(18.35)

The variation gives, in atomic units

$$Z_i^{2^1 P} = 2.00$$
$$Z_o^{2^1 P} = 0.97$$
$$E_{2^1 P} = -2.123$$
$$Z_i^{2^3 P} = 1.99$$
$$Z_o^{2^3 P} = 1.09$$
$$E_{2^3 P} = -2.131$$

The two theoretical values are around 3 au from the experimental measurements.

19 Many Electron Atoms

§ 19.1 Central Field Approximation

In order to treat properly many-electron atoms we need to introduce an approximation, known as *central field approximation or Hartree's method*, in which the Coulomb potential is substituted with a new spherically symmetrical *effective* potential.

Firstly, we consider an atom with N electrons and a nucleus with charge $q_N = Ze$ and infinite nuclear mass. In our Hamiltonian we (at first) neglect all small effects (Spin-Orbit coupling, relativistic effects, etc...). Having considered this, putting $r_{ij} = \left\| \underline{r}_i - \underline{r}_j \right\|$ our many-electron Hamiltonian is

$$\hat{\mathcal{H}} = \sum_{i=1}^{N} \left(-\frac{\hbar^2}{2m}\nabla_i^2 - \frac{Ze^2}{4\pi\epsilon_0 r_i} \right) + \sum_{i<j=1}^{N} \frac{e^2}{4\pi\epsilon_0 r_{ij}} \tag{19.1}$$

In this case, the term r_{ij}^{-1} is too big to be treated as a perturbation, hence, we must use what's called a central field approximation. In this approximation we map our potential $V(r)$ to a new potential defined as follows

$$V_{cf}(r) = -\frac{Z}{r} + S(r) \tag{19.2}$$

Where $S(r)$ is a spherically symmetric *screening* function.
Our Hamiltonian then becomes

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_e + \hat{\mathcal{H}}_1 = \sum_{i=1}^{n} \hat{h}_i + \sum_{i<j=1}^{N} \frac{1}{r_{ij}} - \sum_{i=1}^{N} S(r_i) \tag{19.3}$$

This way, we can divide our calculus in two. Solving for the single electron Hamiltonian $\hat{h}_i$, we have

$$\sum_{i=1}^{N} \hat{h}_i \left| \psi_e \right\rangle = E_e \left| \psi_e \right\rangle$$

$$\left| \psi_e \right\rangle = \bigotimes_{i=1}^{N} \left| nlm_l \right\rangle_i \tag{19.4}$$

$$\hat{h}_i \left| nlm_l \right\rangle_i = E_{nl} \left| nlm_l \right\rangle_i$$

Where $|nlm_i\rangle$ is our single-electron orbital. It's important to remember that these orbitals are «not» the usual Hydrogenic orbitals, since the chosen potential is quite different. Now there is only one simple thing missing, Spin. Our complete electronic wavefunction will then be the direct product between a spinor and the electronic wavefunction, as follows

$$u_{nlm_lm_s}(q_i) = u_{nlm_l}(r_i)\chi_{\frac{1}{2},m_s} \leftrightarrow |nlm_l\rangle_i \otimes \left|\frac{1}{2},\uparrow\downarrow\right\rangle \tag{19.5}$$

Now, we're ready to determine $|\psi_e\rangle$. In order to make sure that Pauli exclusion principle's requirements are solved, we know already that the total wavefunction must be zero if an electron is in the same state of another one. This is perfectly described through a determinant, called Slater determinant, which is defined as follows (let μ, ν, δ be generic states)

$$\psi_e(q_1, \cdots, q_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} u_\mu(q_1) & u_\nu(q_1) & \cdots & u_\delta(q_1) \\ u_\mu(q_2) & u_\nu(q_2) & \cdots & u_\delta(q_2) \\ \vdots & \vdots & \vdots & \vdots \\ u_\mu(q_N) & u_\nu(q_N) & \cdots & u_\delta(q_N) \end{vmatrix} \tag{19.6}$$

It's worth noting that $(N!)^{-1/2}$ is a normalizing factor for our final wavefunction. A notable example comes for He ground state wavefunction. We will have then

$$|1,2\rangle = \frac{1}{\sqrt{2}} \begin{Vmatrix} |100\rangle \otimes |\uparrow\rangle_1 & |100\rangle \otimes |\downarrow\rangle_1 \\ |100\rangle \otimes |\uparrow\rangle_2 & |100\rangle \otimes |\downarrow\rangle_2 \end{Vmatrix}$$

$$|1,2\rangle = \frac{1}{\sqrt{2}} |100\rangle_1 \otimes |100\rangle_2 \otimes \left(|\uparrow\rangle_1 \otimes |\downarrow\rangle_2 - |\downarrow\rangle_1 \otimes |\uparrow\rangle_2 \right) \tag{19.7}$$

The parity of the wavefunction ψ_e will then be $(-1)^{\sum_i l_i}$ as it should be.

Now, it's time to define new vector operators for angular momentum $\underline{\hat{L}}, \underline{\hat{S}}$ in the many-electron case, and defining both as the sum of the single electron operators we can easily demonstrate that

$$\left[\hat{\mathcal{H}}_e, \underline{\hat{S}}\right] = \left[\hat{\mathcal{H}}_e, \underline{\hat{L}}\right] = 0 \tag{19.8}$$

Therefore, we can define a new common basis between the Hamiltonian and these operator as $|LSM_lM_s\rangle$, where

$$\begin{aligned} \hat{S}^2 |LSM_lM_s\rangle &= \hbar^2 S(S+1) |LSM_lM_s\rangle \\ \hat{L}^2 |LSM_lM_s\rangle &= \hbar^2 L(L+1) |LSM_lM_s\rangle \\ \hat{S}_z |LSM_lM_s\rangle &= \hbar M_s |LSM_lM_s\rangle \\ \hat{L}_z |LSM_lM_s\rangle &= \hbar M_l |LSM_lM_s\rangle \\ \hat{\mathcal{H}}_e |LSM_lM_s\rangle &= E_{nl} |LSM_lM_s\rangle \end{aligned} \tag{19.9}$$

It's of great importance to know that in order to find the wavefunctions $\langle r_i|LSM_lM_s\rangle$, we need to find linear combinations of Slater determinants. In the case of Helium, it was fine to use a Slater determinant, since $L = S = 0$, since we were searching for $^1S_0$ wavefunctions for an electronic configuration of $1s^2$.

§§ 19.1.1 Shells and Subshells

As we have seen before, the total energy of atoms is directly determined by the electron configuration. We have already seen that the single electron energy grows with $n + l$, and since the spherical harmonics and spinors are already known from previous calculations, the problem is only to find a new radial function, where the potential is central but not coulombian.

Due to this dependency with n and l we can characterize electronic *shells* and *subshells*. A shell is composed of atoms with the same value of n and subshells are composed by atoms with the same value of n and l.

As in orbital spectroscopic nomenclature, we assign a letter to each shell, as follows

$$
\begin{aligned}
n = 0, \ l = 0 &\longrightarrow K \\
n = 1, \ l = 0 &\longrightarrow L_I \\
n = 1, \ l = 1 &\longrightarrow L_{II} \\
&\vdots
\end{aligned}
\tag{19.10}
$$

The maximum number of electrons in a subshell is $2(2l + 1)$, and if the number of electrons exactly matches it, the subshell in question is called *closed* or *filled*.

Instead, there can be maximum $2n^2$ electrons in a shell, and the "closed" or "filled" names of complete subshells transfers directly to shells.

In general, the degeneracy of a configuration (g), can be determined from the degeneracy of the shells. Let $\delta_i = 2(2l_i + 1)$ be the degeneracy of the subshell, and ν_i the number of electron occupying the same subshell with energy $E_{n_i l_i}$. Henceforth, there will be d_i ways of distributing electrons in this i-th subshell, this number is

$$
d_i = \frac{\delta_i!}{\nu_i! \, (\delta_i - \nu_i)!}
\tag{19.11}
$$

And, therefore, for an electronic configuration we will have that

$$
g = \prod_{i=1}^{n} d_i
$$

Where the index goes through all the n subshells. It's worth noting that for a closed subshell $d = 1$, easing the calculus. Taking Carbon as an example, we will have an electronic configuration of $[He]2s^2 2p^2$, hence

$$
\begin{aligned}
K &\rightarrow \nu = 2, \ \delta = 2, \ d = 1 \\
L_I &\rightarrow \nu = 2, \ \delta = 2, \ d = 1 \\
L_{II} &\rightarrow \nu = 2, \ \delta = 6, \ d = 15
\end{aligned}
\tag{19.12}
$$

Hence $g = 15$.

§§ 19.1.2 Aufbau Rule and the Periodic Table

Having defined properly the electronic structure of many electron atoms, we're able to discuss the Aufbau (building up) of atoms.

The "building up" of atoms is given by the Z electrons, that fill the shells in accordance with the Pauli exclusion principle. The *ground state configuration* of an atom is then given by distributing the electrons into n subshells. The first $n-1$ (or n if the shell is complete) subshells are filled completely, and the last subshell, if incomplete, houses the so called *valence electrons*.

The screening given by the complete shells is what makes sure that the ionization potential doesn't grow with Z, but has peaks for atoms with complete shells, that always have a ground state $^1S_0$. These atoms form the set of *noble gases*, which are chemically inert. Seeing the ionization potential tables is also evident that the ionization potential of noble gases lowers with Z, since the nucleus is bigger, and the last electrons feel less attraction from the nucleus.

The stableness given by having a complete valence shell, is the reason that the *alkalis* (Li, Na, K, Rb, Cs) and the *halogens* (F, Cl, Br, I) are extremely reactive chemically, since the first ones have one weakly bound electron more, and the last ones have a "hole", which is only a missing electron which is needed to complete the shell. The recurrence of this property is what brought chemists to build the *periodic table*, which is a table of all the known elements, ordered in base to their value of Z.

§ 19.2 Hartree-Fock Metod and Self-Consistent Fields

The basic starting point of Hartree-Fock theory is the independent particle model. The complete Hartree-Fock method accounts for the Pauli exclusion principle too, whereas the Hartree method alone doesn't.

The first thing assumed for the Hartree-Fock wavefunction is that the final N-electron wavefunction is a Slater determinant, or an antisymmetric product of electron spin-orbitals. This Slater determinant is given through the variational calculus of every single electron orbital.

Seeing this in a broader way, we can see our final wavefunction being an infinite linear combination of Slater determinants, which lets this metod to be well suited for calculus of even more complex systems like molecular orbitals and solid state physics. In order to keep things simple, we will treat only with the discussion of the ground state of a multi-electron atom, where the considered Hamiltonian is not relativistic.

Now, supposing we have a total Hamiltonian $\hat{\mathcal{H}} = \hat{\mathcal{H}}_1 + \hat{\mathcal{H}}_2$, where

$$\hat{\mathcal{H}}_1 = \sum_{i=1}^{N} \hat{h}_i = \sum_{i=1}^{N} \left(-\frac{1}{2}\nabla_i - \frac{Z}{r_i} \right)$$

$$\hat{\mathcal{H}}_2 = \sum_{i<j=1}^{N} \frac{1}{r_{ij}}$$

(19.13)

We have that the first Hamiltonian describes the sum of N electronic Hamiltonians which include nucleus-electron attraction, and the second describes $N(N-1)/2$ terms which describe the two body interaction of electrons..

According to the variational method, if we suppose that the trial wavefunction is ϕ, we have that

$$E_0 \leq E[\phi] = \langle \phi | \hat{\mathcal{H}} | \phi \rangle$$

Where we suppose $\langle \phi | \phi \rangle$, and $|\phi\rangle$ is determined through a Slater determinant of orthonormal single electron wavefunctions.

Another way to interpret this wavefunction is defining the *antisymmetrization operator* $\hat{A}$. We have

$$\hat{A} = \frac{1}{N!} \sum_P (-1)^P \hat{P}$$

$$|\phi\rangle = \sqrt{N!} \hat{A} |\phi_H\rangle \qquad (19.14)$$

$$|\phi_H\rangle = \bigotimes_{i=1}^N |u_i\rangle$$

It's obvious that since $\hat{A}$ is a linear combination of exchange operators $\hat{P}$, that it's also a projection operator.

The wavefunction $|\phi_H\rangle$ is what's called a *Hartree wavefunction*, which is simply the direct product of every single electron wavefunctions.

Since both Hamiltonians are invariant under permutation of electronic coordinates, we have that $\left[\hat{\mathcal{H}}_i, \hat{A} \right] = 0$, and therefore the calculation of the expectation values of the first Hamiltonian reduces to the following calculus, thanks to our definition of $|\phi\rangle$

$$\langle \phi | \hat{\mathcal{H}}_1 | \phi \rangle = N! \langle \phi_H | \hat{\mathcal{H}}_1 \hat{A} | \phi_H \rangle \qquad (19.15)$$

Thanks to the definition of $\hat{A}$ and $\hat{\mathcal{H}}_1$ we can reduce the calculus to the following expectation value

$$\langle \phi | \hat{\mathcal{H}}_1 | \phi \rangle = \sum_{i=1}^N \sum_P (-1)^P \langle \phi_H | \hat{h}_i \hat{P} | \phi_H \rangle = \sum_\lambda \langle u_\lambda | \hat{h}_i | u_\lambda \rangle \qquad (19.16)$$

Where the index λ runs on all possible quantum states of the single electron wavefunction Defining $I_\lambda = \langle u_\lambda | \hat{h}_i \langle u_\lambda | u_\lambda \rangle$ we have that $\langle \phi | \hat{\mathcal{H}}_1 | \phi \rangle = \sum_\lambda I_\lambda$. Analoguosly with $\hat{\mathcal{H}}_2$ we get

$$\langle \phi | \hat{\mathcal{H}}_2 | \phi \rangle = N! \langle \phi_H | \hat{\mathcal{H}}_2 \hat{A} | \phi_H \rangle \qquad (19.17)$$

Expliciting both operators, we get

$$\langle \phi | \hat{\mathcal{H}}_2 | \phi \rangle = \sum_{i<j=1}^N \sum_P (-1)^P \langle \phi_H | \frac{1}{r_{ij}} \hat{P} | \phi_H \rangle = \sum_{i<j=1}^N \langle \phi_H | \frac{1}{r_{ij}} \left(1 - \hat{P}_{ij} \right) | \phi_H \rangle \qquad (19.18)$$

Since the exchange operator in this case exchanges spin and spatial coordinates of electrons i, j, we can also write the previous equation as follows

$$\langle \phi | \hat{\mathcal{H}}_2 | \phi \rangle = \frac{1}{2} \sum_\lambda \sum_\mu \left(\langle u_\lambda u_\mu | \frac{1}{r_{ij}} | u_\lambda u_\mu \rangle - \langle u_\lambda u_\mu | \frac{1}{r_{ij}} | u_\mu u_\lambda \rangle \right) \qquad (19.19)$$

We can now define two new terms, the *direct term* $J_{\lambda\mu}$ and the *exchange term* $K_{\lambda\mu}$, where they're defined as follows

$$J_{\lambda\mu} = \langle u_\lambda u_\mu | \frac{1}{r_{ij}} | u_\lambda u_\mu \rangle$$

$$K_{\lambda\mu} = \langle u_\lambda u_\mu | \frac{1}{r_{ij}} | u_\mu u_\lambda \rangle$$

(19.20)

It's also obvious that these two matrices are real and symmetric on both indices.
This final calculus brings the following result

$$\langle \phi | \hat{\mathcal{H}}_2 | \phi \rangle = \frac{1}{2} \sum_\lambda \sum_\mu (J_{\lambda\mu} - |\lambda\mu\rangle)$$

(19.21)

The total energy functional is then the following

$$E[\phi] = \sum_\lambda I_\lambda + \frac{1}{2} \sum_\lambda \sum_\mu (J_{\lambda\mu} - K_{\lambda\mu})$$

(19.22)

Introducing the N^2 Lagrange multipliers $c_{\lambda\mu}$, we have that the variational equation that we need to solve is the following

$$\delta E - \sum_\lambda \sum_\mu c_{\lambda\mu} \delta \langle u_\lambda | u_\mu \rangle$$

(19.23)

Diagonalizing the matrix $c_{\lambda\mu}$ with an unitary transformation, we get

$$\delta E - \sum_\lambda \sum_\mu E_\lambda \delta_{\lambda\mu} \delta \langle u_\lambda | u_\mu \rangle = \delta E - \sum_\lambda E_\lambda \delta \langle u_\lambda | u_\lambda \rangle$$

(19.24)

Projecting the previous equation in $L^2(\mathbb{R}^3 \otimes \mathbb{H}_s)$, and variating the spin-orbital electronic wavefunction we have a set of integro-differential equations, known as the *Hartree-Fock equations*

$$\left(-\frac{1}{2} \nabla_i^2 - \frac{Z}{r_i} \right) u_\lambda(q_i) + \left(\sum_\mu \int \overline{u}_\mu(q_j) \frac{1}{r_{ij}} u_\mu(q_j) \, dq_j \right) u_\lambda(q_i) -$$

$$- \sum_\mu \left(\int \overline{u}_\mu(q_j) \frac{1}{r_{ij}} u_\lambda(q_j) \, dq_j \right) u_\mu(q_i) - E_\lambda u_\lambda(q_i) = 0$$

(19.25)

We can build a more compact version of the Hartree-Fock equation defining two new operators as follows

$$\hat{V}_\mu^d = \int \overline{u}_\mu(q_j) \frac{1}{r_{ij}} u_\mu(q_j) \, dq_j = \langle \mu | \frac{1}{r_{ij}} | \mu \rangle$$

$$\hat{V}_\mu^{ex} f(q_i) = \left(\int \overline{u}_\mu(q_j) \frac{1}{r_{ij}} f(q_j) \, dq_j \right) u_\mu(q_i) = \left(|\mu\rangle \langle \mu | \frac{1}{r_{ij}} \right) |f\rangle$$

(19.26)

Called respectively the *direct operator* and *exchange operator*. Otherwise we can define the *direct potential* and the *exchange potential* as follows

$$\hat{\mathcal{V}}^d = \sum_\mu \hat{V}_\mu^d = \sum_\mu \langle \mu | \frac{1}{r_{ij}} | \mu \rangle$$

$$\hat{\mathcal{V}}^{ex} = \sum_\mu \hat{V}_\mu^{ex} = \sum_\mu | \mu \rangle \langle \mu | \frac{1}{r_{ij}}$$

(19.27)

With this new addition, the Hartree-Fock equation becomes the following

$$-\frac{1}{2}\nabla_i^2 u_\lambda(q_i) - \frac{Z}{r_i}u_\lambda(q_i) + \hat{\mathcal{V}}^d u_\lambda(q_i) - \hat{\mathcal{V}}^{ex} u_\lambda(q_i) = E_\lambda u_\lambda(q_i)$$

$$\frac{1}{2}\hat{p}_i^2 | \lambda \rangle - \frac{Z}{r_i} | \lambda \rangle + \hat{\mathcal{V}}^d | \lambda \rangle - \hat{\mathcal{V}}^{ex} | \lambda \rangle = E_\lambda | \lambda \rangle$$

(19.28)

Or, defining the Hartree-Fock potential as follows

$$\hat{\mathcal{V}} = -\frac{Z}{r_i} + \hat{\mathcal{V}}^d - \hat{\mathcal{V}}^{ex}$$

(19.29)

We can rewrite the previous equation as follows

$$-\frac{1}{2}\nabla_i^2 u_\lambda(q_i) + \hat{\mathcal{V}} u_\lambda(q_i) = E_\lambda u_\lambda(q_i)$$

$$\frac{1}{2}\hat{p}^2 | \lambda \rangle + \hat{\mathcal{V}} | \lambda \rangle = E_\lambda | \lambda \rangle$$

(19.30)

§§ 19.2.1 Beryllium Ground State

A particular example of what can be calculated with the Hartree-Fock approximation is the ground state ^1S of Beryllium. The Hartree-Fock potential operator, in this case is

$$\hat{\mathcal{V}} = -\frac{4}{r_i} + \hat{V}_{1s\uparrow}^d + \hat{V}_{1s\downarrow}^d + \hat{V}_{2s\uparrow}^d + \hat{V}_{2s\downarrow}^d - \left(\hat{V}_{1s\uparrow}^{ex} + \hat{V}_{1s\downarrow}^{ex} + \hat{V}_{2s\uparrow}^{ex} + \hat{V}_{2s\downarrow}^{ex} \right)$$

(19.31)

The Hartree-Fock equations, since the spatial part of the two s electrons is be identical, separate into two coupled integro-differential equations, and we get then that $E_{1s} = E_{1s\uparrow} = E_{1s\downarrow}$ and $E_{2s} = E_{2s\uparrow} = E_{2s\downarrow}$. The general solution for these equation is given through a basis change to the basis of *Slater orbitals*, which have the following form

$$s_{nlm}(\underline{r}) = \frac{(2\alpha)^{n+\frac{1}{2}}}{\sqrt{(2n)!}} r^{n-1} e^{-\alpha r} Y_l^m(\theta, \phi)$$

(19.32)

The equations are then solved numerically.
Due to electronic dispositions it's obvious that the total Be wavefunction will be symmetric (S state), the total spin will be 0, and therefore we have a $^1S_0$ state

§ 19.3 Spin-Orbit Interactions and Fine Structure of Many-Electron Atoms

As treated before with single-electron atoms, we define an electronic Hamiltonian ($\hat{\mathcal{H}}_e$) and a perturbative Hamiltonian ($\hat{\mathcal{H}}_2$), where the first will be the sum of all the single electron Hamiltonians, and the second will be our spin-orbit perturbative Hamiltonian.
The total Hamiltonian will then be

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_e + \hat{\mathcal{H}}_1 + \hat{\mathcal{H}}_2 \tag{19.33}$$

Where $\hat{\mathcal{H}}_1$ is the electrostatic correction, and

$$\hat{\mathcal{H}}_2 = \sum_i \xi(r_i)\underline{\hat{L}}_i \cdot \underline{\hat{S}}_i = \sum_i \frac{1}{2m^2c^2}\frac{1}{r_i}\frac{\partial V}{\partial r_i}\underline{\hat{L}}_i \cdot \underline{\hat{S}}_i$$

$$\hat{\mathcal{H}}_e = \sum_i \hat{h}_i = \sum_i \left(-\frac{1}{2}\nabla_i^2 + \hat{\mathcal{V}}\right) \tag{19.34}$$

Since this atom described by $\hat{\mathcal{H}}$ is isolated, the *total parity* and $\underline{\hat{J}}$ are conserved.
We shall utilize the usual perturbation theory on the energy levels obtained from the Hartree-Fock approximation, henceforth taking the Hartree-Fock energy as our "true" energy.
As we saw before, since the Hamiltonian commutes with $\underline{\hat{J}}$ (and $\underline{\hat{L}}, \underline{\hat{S}}$), hence we can say that M_L, M_S are good quantum numbers. Hence, every level will be $(2L+1)(2S+1)$ times degenerate. As usual, every level will be indicated with the usual spectroscopic notation.

§§ 19.3.1 Determination of Possible Terms in Spin-Orbit Coupling

Using the usual angular momenta addition rules, it's possible to determine straight away that, for filled subshells $L = S = 0$. Hence, an atom with its last subshell filled must have a state $^1S_0$. In the case of ions, we toss out all filled subshells, and consider only the *optically active* electrons in order to determine the possible states of the atom. We have three main cases

1. Non-equivalent electrons (in different subshells)
   In this case, there can't be couples of optically active electrons that have the same set of quantum numbers, hence Pauli's exclusion principle is automatically satisfied.
   We find the value of L and S by summing all optically active electrons' single values.
   It can be illustrated with two simple examples

   (a) *Configuration $np\, n'p$*
       In this configuration we have $l_1 = l_2 = 1$ and $s_1 = s_2 = 1/2$, hence $L = 0, 1, 2$ and $S = 0, 1$. We therefore can have the following terms for the configuration

       $$^1S,\ ^1P,\ ^1D,\ ^3S,\ ^3P,\ ^3D$$

    (b) *Configuration $np\,n'd$*
        Here instead we have $l_1 = 1, l_2 = 2$ and $s_1 = s_2 = 1/2$, thus $L = 1, 2, 3$ and $S = 0, 1$. The possible term values are

$$^1P,\ ^1D,\ ^1F,\ ^3P,\ ^3D,\ ^3F$$

    For more than 2 optically active electrons this calculation is repeated up until all the electrons' angular momenta are summed.

2. Equivalent electrons (in the same subshell)
   This case is slightly more complicated, since Pauli's exclusion principle isn't immediately satisfied.
   The most simple case that might be encountered is the case ns^2, which forces us with a $^1S_0$ state. A slightly more complicated case is given by np^2 configurations, where the degeneracy is $g = 15$. Due to the exclusion principle, we must immediately exclude all possible states where m_l or the m_s values of two different electrons are the same. Evaluating all the 15 states, we end up with these possible quantum number couples

$$(M_L = \pm 2, M_S = 0)$$
$$(M_L = \pm 1, M_S = \pm 1),\ (M_L = \pm 1, M_S = 0),\ (M_L = 0, M_S = 0),\ (M_L = 0, M_S = \pm 1)$$

   We see immediately that we can only have $L = 2, 1, 0$, hence the terms will be S, P, D. From the configuration, and the absence of a $(2, 1)$ set of (M_L, M_S), we can immediately say that all the possible 15 states must have one of these three terms

$$^1S,\ ^1D,\ ^3P$$

3. Equivalent and non-equivalent electrons
   If an electronic configuration contains a group of equivalent electrons together with a group of non-equivalent electrons one must firstly determine the possible equivalent electron states, and then sum these states with the non-equivalent electron states. Then all the possible states can be determined.

19.3.1.0.1 Hund's Rules   A set of two empirical rules determined by Hund is fundamental in the research of the ground state configuration. According to these rules we have that

1. The term with the largest value of S has the lowest energy

2. For a given value of S, the term with the maximum possible value of L has the lowest energy

§§ 19.3.2 Fine Structure Terms and Landé's Interval Rule

Having obtained the energy level of the previous Hamiltonian, we proceed now to the second step of the calculation, and indeed add a new perturbation $\hat{\mathcal{H}}_2$ to the previous Hamiltonian. The new total Hamiltonian $\hat{\mathcal{H}}$ will not commute with $\hat{\underline{L}}, \hat{\underline{S}}$ singularly, but it will commute with $\hat{\underline{J}} = \hat{\underline{L}} \otimes \hat{\mathbb{1}}_S + \hat{\mathbb{1}}_L \otimes \hat{\underline{S}}$. Since the energy of an isolated atom cannot depend on the direction of the total angular momentum, it will have degeneracy $g = (2L + 1)(2S + 1)$ associated with ^{2S+1}L term, and an additional *fine structure splitting* characterized by the possible values of J. The new term will then be $^{2S+1}L_J$, which is $2J + 1$-fold degenerate with respect to the eigenvalues of $\hat{J}_z$. The degeneracy in M_J can be removed choosing a preferred direction in space, such as applying an external magnetic field in the Zeeman effect.

In this case, the possible values of M_J are $|L - S| \leq M_J \leq L + S$.

Let's take again the configuration $np\, n'p$. In this case, we have already seen that the possible values of the quantum numbers are the following

$$L = 0, 1, 2$$
$$S = 0, 1 \tag{19.35}$$
$$J = 0, 1, 2, 3$$

Without calculating the spin-orbit coupling splitting, we can have the following terms

$$^1S,\ ^1P,\ ^1D,\ ^3S,\ ^3P,\ ^3D$$

Applying the spin-orbit coupling we now have a additional splitting of the previous states:

$$^1S_0,\ ^1P_1,\ ^1D_2,\, ^3S_1,\ ^3P_0,\ ^3P_1,\ ^3P_2,\ ^3D_1,\ ^3D_2,\ ^3D_3$$

Another example comes from Carbon, which has a configuration $[\text{He}]2s^2 2p^2$. The two optically active electrons give rise to the following fine structure terms

$$^1S_0,\ ^1D_2,\ ^3P_0,\ ^3P_1,\ ^3P_2$$

Using Hund's Rules, we see immediately that the ground state of Carbon is $^3P_0$.

Rewriting our wavefunctions as $|LSM_LM_S\rangle$ and our spin-orbit Hamiltonian as $A\hat{\underline{L}} \cdot \hat{\underline{S}}$, we have that it's non-diagonal in this base, but we can change basis into the new basis $|LSJM_J\rangle$, where we have

$$\langle LSJM_J| \hat{\mathcal{H}}_2 |LSJM_J\rangle = \frac{1}{2}A \langle LSJM_J| \hat{J}^2 - \hat{L}^2 - \hat{S}^2 |LSJM_J\rangle$$
$$= \frac{1}{2}A \left(J(J+1) - L(L+1) - S(S+1) \right) \tag{19.36}$$

Hence, it's diagonal. From this we see that the unperturbed level splits into $2S + 1$ or $2L + 1$ if respectively we have $S \leq L$ or $S > L$. It also can be seen that $E_J - E_{J-1} = AJ$. This last result is known as the *Landé interval rule*, which holds only in L-S coupling regimes,

i.e. when $\left|\hat{\mathcal{H}}_2\right| << \left|\hat{\mathcal{H}}_1\right|$. Evaluating A, we get that if $A > 0$ the lowest energy is given by the term with the lowest value of J, and when $A < 0$ the lowest value of energy is given by the term with the highest level of J. These kinds of multiplet splitting are respectively called *regular multiplet splitting* and *inverted multiplet splitting*. Empirically, it has been established that regular multiplets occur in subshells that are less than half filled, while inverted multiplets appear in more than half filled subshells. In half filled subshells there is no multiplet splitting.

20 Electromagnetic Interactions

§ 20.1 Selection Rules

Having discussed in detail many-electron atoms, we can go forward and discuss the interaction between single atoms and external electromagnetic field.
We will consider only single photon interaction. The photon field Hamiltonian can be written as follows

$$\hat{\mathcal{H}}_\gamma = -\sum_{i=1}^{N} \frac{i\hbar e}{m} \hat{A}^i(r_i, t)\partial_i = \frac{e}{m}\hat{A}^i(r_i, t)\hat{p}_i \tag{20.1}$$

We will have $N = Z$ for neutral atoms and $N \neq Z$ for ions.
Jumping all the calculations (they're already given for single electron atoms), we get that the transition probability depends on a matrix element M_{ba}, where

$$M_{ba} = -\frac{m\omega_{ba}}{\hbar} \epsilon_j \cdot r_{ba}^j \tag{20.2}$$

The generic state $|a\rangle$ is an eigenstate of the total angular momentum and parity, hence, indicating with p the parity eigenstate, we can write M_{ba} in a different way, as follows

$$M_{ba} = -\frac{Nm\omega_{ba}}{\hbar} \epsilon^j \left\langle p'J'M'_J \right| r_j \left| pJM_J \right\rangle \tag{20.3}$$

The N comes from the fact that the electrons are indistinguishable. Inserting into M_{ba} the dipole moment operator $\underline{\hat{D}} = -\sum_j e\underline{r}_j$ we have

$$M_{ba} = \frac{m\omega_{ba}}{\hbar e} \epsilon^j \left\langle p'J'M'_J \right| \hat{D}_j \left| pJM_J \right\rangle \tag{20.4}$$

From which, the probability of spontaneous emission of a photon with polarization ϵ_i can be calculated, yielding the following result

$$W_{ab}^s \, d\Omega = \frac{\omega_{ba}}{8\pi^2\epsilon_0\hbar c^3} \left| \epsilon^j \left\langle p'J'M'_J \right| \hat{D}_j \left| pJM_J \right\rangle \right|^2 d\Omega \tag{20.5}$$

We can now define the spherical vector components of the polarization versor ϵ_j and of the electric dipole operator. Using q as an index for the three possible components $(0, \pm 1)$,

we have for the electric dipole operator

$$\hat{D}_1 = -\frac{1}{\sqrt{2}} \left(D_x + i\hat{D}_y \right) = \left\| \hat{D}_i \right\| \sqrt{\frac{4\pi}{3}} Y_1^1(\alpha, \beta)$$

$$\hat{D}_0 = \hat{D}_z = \left\| \hat{D}_i \right\| \sqrt{\frac{4\pi}{3}} Y_1^0(\alpha, \beta) \tag{20.6}$$

$$\hat{D}_{-1} = \frac{1}{\sqrt{2}} \left(\hat{D}_x - i\hat{D}_y \right) = \left\| \hat{D}_i \right\| \sqrt{\frac{4\pi}{3}} Y_1^{-1}(\alpha, \beta)$$

And for ϵ_i

$$\epsilon_1 = -\frac{1}{\sqrt{2}} \left(\epsilon_x + i\epsilon_y \right)$$

$$\epsilon_0 = \epsilon_z \tag{20.7}$$

$$\epsilon_{-1} = \frac{1}{\sqrt{2}} \left(\epsilon_x - i\epsilon_y \right)$$

Using the Wigner-Eckart theorem we know that the matrix elements of a vector operator with respect to the eigenstate of the total angular momentum (squared and through the z axis) depend only on M_J, M_J' and q through the Clebsch-Gordan coefficients $\langle JM_Jq|JM_J \rangle$, henceforth

$$\langle p'J'M_J' | \hat{D}_q | pJM_J \rangle = \frac{1}{\sqrt{2J'+1}} \langle JM_Jq|J'M_J' \rangle \langle p'J' \| \hat{D}_i \| pJ \rangle \tag{20.8}$$

The Clebsch-Gordan coefficient vanishes, unless

$$M_J + q = M_J'$$
$$|J - 1| \leq J' \leq J + 1 \tag{20.9}$$
$$J + J' \geq 1$$

Thus, obtaining the selection rule for electric dipole transition

$$\boxed{\Delta M_J = 0, \pm 1}$$
$$\boxed{\Delta J = 0, \pm 1} \tag{20.10}$$

Be cautious, transitions $J = 0 \leftrightarrow J' = 0$ are not permitted. In addition, due to *Laporte's rule*, the state from which the transition happens, «must have the opposite parity of the initial state»

§§ 20.1.1 Spin-Orbit Coupling

In the case of spin-orbit coupling (hence weak spin-orbit interaction) we can approximate the system in a way such that $\hat{L}, \hat{S}$ are conserved. We then obtain

$$\langle J'L'S'M_J' | \hat{D}_i | JLSM_J \rangle = \delta_{SS'} \langle J'L'S'M_J' | \hat{D}_i | JLSM_J \rangle \tag{20.11}$$

Which give the following selection rules

$$\boxed{\Delta L = 0, \pm 1} \text{ (0} \leftrightarrow \text{0 not allowed)}$$
$$\boxed{\Delta S = 0} \tag{20.12}$$

§ 20.2 Alkali Atoms

As we have seen for single-electron atoms, Alkali metals can be treated as such, with a single valence electron in a ns^1 orbital, shielded from the nucleus by a "core", which is composed by a closed subshell system.

Analyzing the valence electron, it's easy to see that it «always» has $l = 0$ at its lowest level, and it's subject to a potential which is Coulombian for large r. Due to this potential, there is no degeneracy in l for a given n, this degeneracy is visible only for highly excited levels, for which the atom's wavefunction is more and more Hydrogenic. The valence electron is weakly bound to the atom, and it usually needs around 5 eV to transition into the continuum.

The ground state of Alkalis, since the core is formed by a system of closed subshells and has only a valence electron, will be $(ns^1)^2 S_{1/2}$ (The core has term $^1S_0$ and you add a valence electron to it, where $l = 0, s = 1/2$). Excited states will be of the following form $(n_1 s^1)^2 S_{1/2}, \ (n_2 p^1)^2 P_{1/2,3/2}, \ (n_3 d^1)^2 D_{3/2,5/2}$ and so on.

The spectra of Alkali atoms can be determined through an approximation of the energies of single electron atoms. In atomic units, we can then write

$$E_{nl} = -\frac{1}{2} \frac{1}{(n - \mu_{nl})^2} \tag{20.13}$$

Where μ_{nl} is called *quantum defect*. This defect can be approximated to a funxtion of l, as $\mu_{nl} \simeq \delta_l$, and with this we can write a "special" quantum number $n^\star = n - \delta_l$.

This is especially useful when considering transitions. In fact, we can write that the absorption lines, in this approximation, will fall in these frequencies, at least for $p \rightarrow s$ transitions

$$\nu_{nl} = R \left(\left(\frac{1}{n_s^\star} \right)^2 - \left(\frac{1}{n_p^\star} \right)^2 \right) \tag{20.14}$$

In the case of emission lines, we can write, with $\tilde{Z} = Z - N + 1$, that the energy of the sequence of emission lines will have the following energies (in cm^{-1})

$$E_{nl} = -\frac{1}{2} \frac{\tilde{Z}^2}{(n^\star)^2} \tag{20.15}$$

From this, it's pretty easy to generalize this to the case of fine structure. The shift induced from this is given by

$$\Delta E = \frac{1}{2}\kappa_{nl}\left(j(j+1) - l(l+1) - \frac{3}{2}\right)$$

$$\kappa_{nl} = \frac{\hbar^2}{2m^2c^2}\left\langle\frac{1}{r}\frac{\partial V}{\partial r}\right\rangle$$

$$(20.16)$$

The value of κ_{nl} can be determined with $V(r)$, using Hartree-Fock's approximation.

§ 20.3 Alkaline Earths

Using the same reasoning of the Alkali atoms, and using the recursion properties of the periodic table, we can study Alkaline Earths as a special kind of two electron atoms.
In the case of small spin dependent interaction, we can take the total spin S as a good quantum number. As with Helium, all levels can be divided into singlet levels $S = 0$ and triplet levels $S = 1$. Since the electric dipole operator can't change spin, we have that the selection rule $\Delta S = 0$ must hold. For atoms with small nuclear charge Ze, spin-orbit and spin-spin coupling behave like small perturbations on the triplet states (where $\hat{S} \neq 0$) but aren't big enough in order to mix these states, hence L, S remain good quantum numbers, since both are conserved to a very good approximation.
In general, for triplet and singlet states, we can have these two following terms

$$\begin{cases} {}^{2S+1}L_{L,L\pm1} & L \neq 0 \\ {}^{1}S_1 \end{cases}$$

$$(20.17)$$

§ 20.4 Multiplet Structure

Thanks to Landé's interval rule, we can immediately determine that, in general

$$\Delta E_{J,J+1} = AJ$$

And in spin-orbit coupling regimes, it holds as

$$\Delta E_{J,J+1} = AS(2L + 1) = AL(2S + 1)$$

Given these two equations, it's now possible to determine level intensities for multiplet transitions. This is especially useful in L-S regimes, since, the number of atoms in each level is proportional to the statistical weight of the level, we get that, if we indicate with I_i the $i-$th transition intensity, we have

$$\sum_{i=1}^{n} I_i \propto (2J + 1)$$

$$(20.18)$$

A complete application of this sum rule can be done with $^3S_1 \to\ ^3P_{012}$ (note that they're permitted transitions). We get therefore the following system

$$\begin{cases} I_1 + I_2 + I_3 = 3I_S \\ I_1 = I_P \\ I_2 = 3I_P \\ I_3 = 5I_P \end{cases} \tag{20.19}$$

With I_P the proportionality coefficient for $P \to S$ transitions and I_S analogously for $S \to P$ transitions.

This simple system is basically already solved, and it immediately gives an intensity ratio of $5 : 3 : 1$ between the three transitions.

§ 20.5 Magnetic Field Interaction, Zeeman Effect

Let's now consider the perturbation applied on the energy levels of multielectron atom when applying a magnetic field. Indicating our perturbing Hamiltonian with $\hat{\mathcal{H}}_B$ we have

$$\begin{aligned} \hat{\mathcal{H}} &= \hat{\mathcal{H}}_e + \hat{\mathcal{H}}_{LS} + \hat{\mathcal{H}}_B \\ \hat{\mathcal{H}}_{LS} &= \frac{\partial V}{\partial r} \underline{\hat{L}} \cdot \underline{\hat{S}} \\ \hat{\mathcal{H}}_B &= -\underline{\hat{\mu}} \cdot \underline{\hat{B}} = \mu_B g \underline{\hat{J}} \cdot \underline{\hat{B}} = -\mu_B \underline{\hat{B}} \left(\underline{\hat{L}} + 2\underline{\hat{S}} \right) \\ \mu_B &= \frac{e\hbar}{2m_e} \end{aligned} \tag{20.20}$$

Where $\hat{\mathcal{H}}_{LS}$ is our spin-orbit coupling Hamiltonian. We might immediately define two cases.

1. $\left| \hat{\mathcal{H}}_B \right| << \left| \hat{\mathcal{H}}_{LS} \right|$, i.e. the magnetic field is much weaker than the LS coupling, also known as *Anomalous Zeeman Effect*

2. $\left| \hat{\mathcal{H}}_B \right| >> \left| \hat{\mathcal{H}}_{LS} \right|$, i.e. the magnetic field is much stronger than the LS coupling, also known as *Paschen-Back Effect*

§§ 20.5.1 Paschen-Back Effect, Strong Field

In this case we have that the magnetic field is much stronger than the LS-coupling, henceforth we can say easily that in this case the coupling is broken, i.e. the $\hat{\mathcal{H}}_{LS}$ is considered as perturbation, and therefore we're left with this Hamiltonian

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_e + \hat{\mathcal{H}}_B = \hat{\mathcal{H}}_e + \frac{\mu_B}{\hbar} B_z(\hat{L}_z + 2\hat{S}_z) + A\hat{L}_z\hat{S}_z \tag{20.21}$$

Where we chose for easier calculation our z axis as the magnetic field direction. Since, in this case, both Hamiltonians have the same eigenvectors ($|nlsm_lm_s\rangle$, we have that

$$\langle lsm_lm_s|\hat{\mathcal{H}}|lsm_lm_s\rangle = -E_n + \frac{\mu_B}{\hbar}B_z\langle lsm_lm_s|(\hat{L}_z + 2\hat{S}_z)|lsm_lm_s\rangle +$$
$$+ \langle lsm_lm_s|A\hat{L}_z\hat{S}_z|lsm_lm_s\rangle = \tag{20.22}$$
$$= -E_n + \mu_B B_z (M_L + 2M_S) + \hbar A M_L M_S$$

Where the Paschen-Back shift is $\Delta_{PB} = \mu_B B_z(M_L + 2M_S) + \hbar A M_L M_S$.

§§ 20.5.2 Anomalous Zeeman Effect, Weak Field

In this case the magnetic field is too weak in order to break the spin orbit coupling, and our Hamiltonian hasn't got $|nlsm_lm_s\rangle$ as eigenvectors, but rather $|nlsjm_j\rangle$. Hence we have that energy perturbation $\hat{\mathcal{H}}_B$ is

$$\Delta_{AZ} = \langle nlsjm_j|\frac{\mu_B}{\hbar}\underline{\hat{J}}\cdot\underline{\hat{B}}|nlsjm_j\rangle$$

Now, writing $\hat{J}_z - \hat{S}_z = \hat{L}_z$ we have that our calculation becomes the following

$$\Delta_{AZ} = \mu_B M_J B_z + \frac{\mu_B}{\hbar}B_z\langle nlsjm_j|S_z|nlsjm_j\rangle$$

Since the eigenkets are not eigenkets of $\hat{S}_z$, we use the Wigner-Eckart theorem (see appendix) in order to rewrite $\hat{S}_z$ in a easier way to manipulate it

$$\hbar^2 J(J+1)\langle nlsjm|\hat{S}_z|nlsjm\rangle = \hbar M_J\langle nlsjm|\underline{\hat{S}}\cdot\underline{\hat{J}}|nlsjm\rangle \tag{20.23}$$

Using the definition of $\underline{\hat{J}}$ we get

$$\underline{\hat{J}}\cdot\underline{\hat{S}} = \frac{1}{2}\left(\hat{J}^2 + \hat{S}^2 - \hat{L}^2\right)$$

These last operators are diagonal in our basis, hence we get, finally

$$\langle nlsjm|S_z|nlsmj\rangle = \hbar M_j\left(\frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)}\right)$$

And, therefore

$$\Delta_{AZ} = \mu_B M_J B_z\left(1 + \frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)}\right) = \mu_B g_j M_J B_z \tag{20.24}$$

Where g_j is the so called *Landé g factor*. As a recapitulation, we get that the Zeeman shift is the following for strong fields (Paschen-Back) and weak fields (Anomalous Zeeman)

$$\Delta E_{PB} = -E_n + \mu_B B_z(M_L + 2M_S) + \hbar A M_L M_S$$
$$\Delta E_{AZ} = -E_{nj} + \mu_B g_j M_J B_z \tag{20.25}$$

Part IV

Quantum Chemistry

21 Molecular Structure

§ 21.1 Introductory Remarks

In order to begin to appreciate physically how a molecule can be described is by considering a few conditions:

Obviously the nuclei of the molecules have comparatively a much bigger mass than all the electrons combined, therefore, the center of mass of the nuclei can be thought as being "fixed" in space, at a certain average distance from both atoms, called the *equilibrium spacing* which depends on the considered molecule.

Let's now consider valence electrons. They're distributed all around the molecule, and their charge distribution gives the force needed to keep the molecule bound. Now, if a is an average distance between the two nuclei, we have a rough estimate of the energy levels of the valence electrons of the molecule. Using the uncertainty principle, we get that the magnitude of the electronic energies is approximately the following

$$E_e \approx \frac{\hbar^2}{ma^2}$$

Since $a \approx 1$ Å, we have that E_e is on the order of some eV.

Let's now consider motions. A molecule, in general, can rototraslate in space, which can be reduced to the calculation of vibrations around the equilibrium distance of the molecule and the rotation of itself, after separating the center of mass.

The nuclear vibrational energies can be calculated by supposing that both nuclei are tied by an elastic force, for which the oscillation frequency is $\omega_N = (k/M)^{1/2}$ (where M is the nuclear mass, and k is a parameter). Comparing it to the electroinc oscillation, we get that the energy levels associated to vibration will have the following spacing (in relation to electronic levels)

$$\frac{E_v}{E_e} = \frac{\hbar\omega_N}{\hbar\omega_e} = \sqrt{\frac{m}{M}}$$

Hence

$$E_v \simeq \sqrt{\frac{m}{M}} E_e$$

Inserting the approximate values for the mass ratio m/M we have that E_v is around 10^2 times smaller than E_e.

Let's now consider the rotation of a diatomic molecule. In this simplified case, we have that the moment of inertia is $Ma^2/2$, and the energy associated will then be

$$E_r \simeq \frac{\hbar^2}{Ma^2} \simeq \frac{m}{M}E_e$$

Which gives has an even smaller value than E_v. Therefore we can say that $E_r < E_v < E_e$, and we can imagine E_r, E_v as a first and second order splitting of the initial energy levels E_e

§ 21.2 Born-Oppenheimer Approximation

We now consider a diatomic molecule with nuclei A and B with mass M_a, M_b with together N electrons. The internuclear coordinate will be indicated with $\underline{R} = \underline{R}_B - \underline{R}_A$ and the electronic positions with respect to the center of mass will be indicated by $\underline{r}_i$. The Schrödinger equation for this system is straightforward, and neglecting spin we have

$$\left(\hat{T}_N + \hat{T}_e + \hat{V}\right)|\psi\rangle = E|\psi\rangle \tag{21.1}$$

Where $\hat{T}_N$ is the kinetic energy of the nuclei, $\hat{T}_e$ of the electrons, and $\hat{V}$ is the potential operator. Rewriting everything explicitly, where μ is taken as the reduced mass of the two nuclei, we have

$$-\frac{\hbar^2}{2\mu}\nabla_R^2\psi(R_j, r_k) - \frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2\psi(R_j, r_k) - \left(\sum_{i=1}^{N}\frac{Z_Ae^2}{4\pi\epsilon_0|r_i - R_A|} - \right.$$

$$\left. -\sum_{i=1}^{N}\frac{Z_Be^2}{4\pi\epsilon_0|r_i - R_B|} + \sum_{i>j=1}^{N}\frac{e^2}{4\pi\epsilon_0|r_i - r_j|} + \frac{Z_AZ_Be^2}{4\pi\epsilon_0 R}\right)\psi(R_j, r_k) = E\psi(R_j, r_k) \tag{21.2}$$

This equation is clearly unsolvable. In this case tho, we can suppose that the total wavefunction $|\psi\rangle$ is the product of two wavefunctions, an electronic wavefunction, and a nuclear wavefunction, hence, respectively $|\psi\rangle = |\Phi_q\rangle \otimes |F_q\rangle$.
Here, we have that the electronic wavefunction will depend parametrically by the internuclear distance, and the wave equation searched is the following

$$\hat{T}_e|\Phi_q\rangle + \hat{V}|\Phi_q\rangle = E_q(R)|\Phi_q\rangle$$

$$-\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2\Phi_q(r_k; R) = E_q(R)\Phi_q(r_k; R) \tag{21.3}$$

These wavefunctions form a complete basis set for each R, and therefore, we can expand the total molecular wavefunction as follows

$$\psi(R_i, r_k) = \sum_{q}F_q(R)\Phi_q(r_k; R) \tag{21.4}$$

Reinserting it back on the initial equation, we have that, projecting into the $|\Phi\rangle$ basis, we have

$$\sum_q \int \overline{\Phi_s} \left(\hat{T}_N + \hat{T}_e + \hat{V} - E \right) F_q(R)\Phi_q \, \mathrm{d}^N r_i = 0 \tag{21.5}$$

Usin the fact that Φ_q is an eigenfunction for the electronic kinetic energy operator and the potential operator, we have, also considering the orthonormality condition $\langle \Phi_s | \Phi_q \rangle = \delta_{sq}$ we have

$$\sum_q \int \overline{\Phi_s} \hat{T}_N \Phi_q F_q(R) \, \mathrm{d}^N r_i + (E_s(R) - E) F_s(R) = 0 \tag{21.6}$$

Where the operator $\hat{T}_N$ acts as follows

$$\langle x_i | \hat{T}_N | \Phi_q \rangle | F_q \rangle = -\frac{\hbar^2}{2\mu} \left(F_q \nabla_R^2 \Phi_q + 2\nabla_R F_q \cdot \nabla_R \Phi_q + \Phi_q \nabla_R^2 \Phi_q \right) \tag{21.7}$$

Here comes the important piece, as we said before, the motion of the nuclei around the equilibrium value doesn't affect particularly electrons, hence all the $\nabla_R^2 \Phi_q$ parts of the differential equation can be neglected, and we're left with the *Nuclear wave equation*

$$-\frac{\hbar^2}{2\mu} \nabla_R^2 F_q(R) + (E_q(R) - E)F_q(R) = 0 \tag{21.8}$$

Thus, we have finally that the differential equation (21.2) will be solved with two different equations with the following conditions

$$-\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla_i^2 \Phi_q + V(r_k, R)\Phi_q = E_q(R)\Phi_q$$

$$-\frac{\hbar^2}{2\mu} \nabla_R^2 F_q + E_q(R)F_q = EF_q$$

Where

$$\nabla_R \Phi_q = 0$$
$$\psi(R_i, r_k) = F_q(R_i)\Phi_q(r_k; R_i)$$

It's of particular interest the fact that the energies found solving the electronic equation $E_q(R)$, while using the Born-Oppenheimer approximation, behave as a potential energy for the nuclear Schrödinger equation.

§ 21.3 Rovibronic States

Let's now analyze properly the nuclear wavefunction. Analyzing the symmetries of the system, it's obvious that this wavefunction will be the product of a radial and an angular part.
Already knowing how ∇_R^2 can be written, we have that it must be an eigenfunction of

the total angular momentum $\hat{J}^2, \hat{J}_z$, and plugging it into (21.8) we get our desired radial equation

$$-\frac{\hbar^2}{2\mu}\frac{\partial^2 \mathcal{F}_{s\nu J}}{\partial R^2} + \frac{\hbar^2}{2\mu R^2}J(J+1)\mathcal{F}_{s\nu J} + (E_s(R) - E_{s\nu J})\mathcal{F}_{s\nu J} = 0 \qquad (21.9)$$

Here ν works as a "principal" quantum number for the quantized levels of the oscillation of the two nuclei, and together with the quantum number J forms the set of *Rovibronic States*. Another problem ensues from this equation, the "potential" $E_s(R)$ hasn't been determined, and it can't properly be determined. Hence we approximate it in a power series around the equilibrium position R_0 up to the second order, and since $E'_s(R_0) = 0$, we're left with the following expression

$$E_s(R_0) = E_s(R_0) + \frac{1}{2}\frac{\partial E_s}{\partial R_0} + \mathcal{O}\big((R - R_0)^2\big) \qquad (21.10)$$

Rewriting the second derivative as k, and putting $k/\mu = \omega_0$ we see already how it behaves like a harmonic potential. We immediately write from this that the total energy will be $E_{s\nu J} = E_s(R_0) + E_\nu + E_r$, with the last one being the rotational energy

$$E_r = \frac{\hbar^2}{2\mu R_0^2}J(J+1) = BJ(J+1)$$

Where B is called the rotational constant.
Due to the harmonicity of the new $E_s(R)$ potential, we will have obviously the following result for E_ν

$$E_\nu = \hbar\omega_0\left(\nu + \frac{1}{2}\right)$$

We can then rewrite our Schrödinger equation as follows, with $\mathcal{F}_{s\nu J}Y_J^M(\theta, \phi) \to |\nu JM\rangle$

$$\hat{\mathcal{H}}_N |\nu JM\rangle = \left(E_s(R_0) + \hbar\omega_0\left(\nu + \frac{1}{2}\right) + \frac{\hbar^2}{2\mu R_0^2}J(J+1)\right)|\nu JM\rangle \qquad (21.11)$$

A better approximation is given by the empirical Morse potential $V_M(R)$, where

$$V_M(R) = D_e\left(e^{-2\alpha(R-R_0)} - 2e^{\alpha(R-R_0)}\right)$$

With D_e, α constants.
We plug it into our equation as a correction of $E_s(R)$, as $E_s(R) = E_s(\infty) + V(R)$, which then yields D_e as a minimum for $R \to \infty$, that gives its name as the *dissociation constant* for the molecule.
Approximating again $E_s(R)$ to the second order and equating the coefficients of $E_s(R)$ and $V_M(R)$, we get the following equality

$$D_e\alpha^2 = \frac{1}{2}k \qquad (21.12)$$

Now, taking $V_M(R)$ into account, we get the new vibrational energy, which has now an anharmonic correction

$$E_\nu = \hbar\omega_0 \left[\left(\nu + \frac{1}{2} \right) - \beta \left(\nu + \frac{1}{2} \right)^2 \right] \tag{21.13}$$

β is our anharmonicity constant, and from the previous considerations, we get that

$$\beta\omega_0 = \frac{\hbar\omega_0}{4D_e}$$

From the new energy, we now know that at the ground vibrational state $\nu = 0$, is not 0, which gives us the true dissociation energy D_0 as

$$D_0 = E_s(\infty) - E_s(R_0) - \frac{\hbar\omega_0}{2}$$

Or, more explicitly

$$D_0 = D_e - \frac{\hbar\omega_0}{2} \tag{21.14}$$

§§ 21.3.1 Centrifugal Distortion

Now we have determined that the oscillations of the two nuclei are quantized, hence there exists a series of values R_k, which modifies the energies of the system. We now rewrite the nuclear Hamiltonian as follows

$$\hat{\mathcal{H}}_N = \frac{\hat{p}_R^2}{2\mu} + \hat{V}_{eff} - \tilde{E}_{s\nu J} \tag{21.15}$$

Where

$$\hat{V}_{eff} = \hat{V}_M + \frac{\hat{J}^2}{2\mu R^2}$$

$$\tilde{E}_{s\nu J} = E_{s\nu J} - E_s(\infty)$$

Starting from this, we begin to evaluate the energies at $R = R_1$. We start by approximating the effective potential at the new minimum R_1 up until the 4th order

$$V_{eff}(R) \simeq V_0 + \frac{1}{2}\tilde{k}(R - R_1)^2 + c_1(R - R_1)^3 + c_2(R - R_1)^4$$

We search iteratively the exact value of R_1 by deriving $V_{eff}(R)$, and we finally get

$$R_1 \simeq R_0 + \frac{\hbar^2}{2\mu\alpha^2 R_0^3 D_e} J(J+1) \tag{21.16}$$

Where for a simpler calculus we put $c_1 = c_2 = 0$, $D_e \alpha^2 = k/2$ and $\tilde{k} = k$.

Using what we found, and applying the third and fourth order of the approximation as a second order perturbation, we finally get our effective energy $\tilde{E}_{s\nu J}$

$$\tilde{E}_{s\nu J} = -D_e + \hbar\omega_0 \left[\left(\nu + \frac{1}{2} \right) - \beta \left(\nu + \frac{1}{2} \right)^2 \right] + \frac{\hbar^2 J(J+1)}{2\mu R_0^2} -$$
$$- \frac{3\hbar^3 \omega_0 J(J+1)}{4\mu\alpha R_0^3 D_e} \left(1 - \frac{1}{\alpha R_0} \right) \left(\nu + \frac{1}{2} \right) - \frac{\hbar^4 J^2(J+1)^2}{4\mu^2\alpha^2 R_0^6 D_e}$$

(21.17)

This effective energy includes the anharmonic correction to the quantum oscillator, a rotation-vibration coupling correction and a correction to the energy of the rigid rotor. This lets us evaluate directly the centrifugal distortions, since for large ν the average distance $R > R_0$. The same goes for large values of J.

22 Molecular Spectra

§ 22.1 Diatomic Molecules

Let's now consider again the vibration of diatomic molecules, where now we also consider cases where $\Lambda \neq 0$. We then define two "new" operators, $\underline{\hat{\Lambda}}$, the angular momentum of the molecule, $\underline{\hat{N}}$, the angular momentum of the two nuclei and the associated total angular momentum operator $\underline{\hat{J}} = \underline{\hat{\Lambda}} + \underline{\hat{N}}$. We shall consider cases where the spin-orbit coupling is negligible and no relativistic corrections are considered.

We get back to the definition of $\underline{\hat{N}}$. From the definition of angular momentum, we already know that

$$\underline{\hat{N}} = \underline{\hat{R}} \wedge \underline{\hat{P}} \Longrightarrow \underline{\hat{N}} \cdot \underline{\hat{R}} = 0$$

Where $\underline{\hat{R}}$ is the internuclear distance operator.

We define the projection on the internuclear axis of the total angular momentum as follows, and we already know that it must have the same value of the projection of $\underline{\hat{\Lambda}}$ on the internuclear axis, since $\underline{\hat{N}} \cdot \underline{\hat{R}} = 0$ so

$$\hat{J}_R = \frac{\underline{\hat{J}} \cdot \underline{\hat{R}}}{\hat{R}} = \hat{\Lambda}_R$$

Considering now all the symmetries of the system, we can with ease say that $\left[\hat{\mathcal{H}}, \hat{J}_R\right] = \left[\hat{\mathcal{H}}, \hat{J}^2\right] = 0$, and we can define the eigenstate of the molecule as $|sJM\Lambda\rangle$, where s is a set of quantum numbers which do not intervene in the rotovibration of the molecule

We begin by analyzing the nuclear Schrödinger equation for the diatomic molecule, and we

get

$$\hat{\mathcal{H}}\mathcal{F}_{s\nu J} = \left(-\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial R^2} + \frac{\left\langle \hat{N}^2 \right\rangle}{2\mu R_0^2}\right)\mathcal{F}_{s\nu J} + \left(E_s(R_0) - E\right)\mathcal{F}_{s\nu J} = 0$$

$$\left\langle \hat{N}^2 \right\rangle = \left\langle \left(\hat{J}^2 - \hat{\Lambda}^2\right)\right\rangle = \hbar^2 J(J+1) + \left\langle \hat{\Lambda}^2 \right\rangle - 2\left\langle \underline{\hat{J}}\cdot\underline{\hat{\Lambda}} \right\rangle$$

$$\hat{\mathcal{H}}\mathcal{F}_{s\nu J} = \left(-\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial R^2} + \frac{\hbar^2}{2\mu R_0^2}J(J+1) + \left(E_s'(R_0) - E\right)\right)\mathcal{F}_{s\nu J} = 0 \tag{22.1}$$

$$E_s'(R_0) = E_s(R_0) + \frac{1}{2\mu R_0}\left(\left\langle \hat{\Lambda}^2 \right\rangle - 2\left\langle \underline{\hat{J}}\cdot\underline{\hat{\Lambda}} \right\rangle\right) =$$

$$= E_s(R_0) + \frac{1}{2\mu R_0}\left(\left\langle \hat{\Lambda}^2 \right\rangle - 2\hbar^2\Lambda^2\right)$$

The choice of redefining $E_s(R_0)$ is immediate, by noting how the terms defining $E_s'(R_0)$ all depend on the electronic terms of the molecule.

Another way of considering the rotation of a molecule, is given by starting from the basic approximation that a molecule could be seen as an almost completely rigid body, and therefore, in the system of the center of mass of the molecule, we can write that the kinetic energy operator will be the following in the coordinates where the inertia tensor is diagonal:

$$\hat{T} = \frac{1}{2I_a}\hat{J}_a^2 + \frac{1}{2I_b}\hat{J}_b^2 + \frac{1}{2I_c}\hat{J}_c^2 \tag{22.2}$$

Since in a diatomic molecule one of the eigenaxis will be the internuclear axis, let's say it's the c axis, we get that $I_a = I_b$, and $\hat{J}_c = \hat{\Lambda}_R$

$$\hat{T} = \frac{1}{2I_a}(\hat{J}_a^2 + \hat{J}_b^2) + \left(\frac{1}{2I_c} + \frac{1}{2I_b}\right)\hat{\Lambda}_R^2$$

This equation is the same equation of a symmetric top. For a diatomic molecule we have that $I_a = \mu R_0^2$ and that I_c depends directly on the electronic terms of the molecule
From this we have

$$E_r = \frac{\hbar^2}{2\mu R_0^2}J(J+1) + \left(\frac{1}{2I_c} + \frac{1}{2I_b}\right)\hbar^2\Lambda^2 \tag{22.3}$$

This point of view makes the generalization to polyatomic molecules much more easy to derive.

§ 22.2 Rovibrational Spectra of Diatomic Molecules

Let's begin considering interaction with the radiation field and the possible transition for the rovibronic states. We define as usual the dipole operator as the sum of the dipole moments of the nuclei and of the electrons

$$\hat{D}_j = e\left(\sum_i Z_i\hat{R}_{ij} - \sum_i \hat{r}_{ij}\right) \tag{22.4}$$

As we did before, we approximate the system by considering no coupling between angular momentums and by neglecting rotovibrational motions.

The molecular wavefunction in study will then be a product of the electronic wavefunction, a vibrational wavefunction and a rotational wavefunction, with the following quantum numbers

$$|\Psi\rangle = \frac{1}{R}|s\rangle_{e^-}|\nu\rangle|JM\Lambda\rangle \tag{22.5}$$

Where, $|\nu\rangle$ is the standard eigenfunction of the quantum harmonic oscillator.

Disregarding spin in our calculations, we define the *permanent electric dipole moment* δ of the molecule as the diagonal elements of the $\hat{D}_i$ operator.

$$\underline{\hat{D}}_{\alpha\alpha} = \langle\alpha|\underline{\hat{D}}|\alpha\rangle = \underline{\delta}$$

These matrix elements always vanish if $|\alpha\rangle$ is a nondegenerate state, but also it doesn't fade if there is an excess of charge in one of the two nuclei, which is always true when we consider heteronuclear molecules.

Since homonuclear molecules are symmetric, all elements of the dipole operator will be zero, and therefore transition can happen only if there is an electronic transition, since $J \geq |\Lambda|$. So, for the already known transition rules for $\Lambda = 0$

$$\Delta J = \pm 1$$
$$\Delta M = 0, \pm 1$$

We have, for $\Lambda \neq 0$

$$\Delta J = 0, \pm 1 \qquad \Delta J = 0 \text{ if only if } \Lambda \neq 0$$
$$\Delta M = 0, \pm 1$$
$$\Delta \Lambda = 0$$

The spectrum found will lay on the far IR or the microwave region of frequencies for diatomic molecules, which will have a definite energy

$$\hbar\Delta\omega_{\Delta J} = E_r(J+1) - E_r(J) = 2B(J+1) \tag{22.6}$$

As we said before, vibrational transitions can happen if and only if the vibrational matrix elements of the dipole operator are nonzero, i.e. if the following integral is nonzero

$$\underline{\hat{D}}_{\nu\nu'} = \langle\nu|\underline{\hat{D}}|\nu'\rangle \tag{22.7}$$

Expanding the integral in power series, and using the definition of Hermite polynomials, we have that the integral depends only on $(R - R_0)$.

$$\underline{\hat{D}}(R) \simeq \underline{\hat{D}}(R_0) + \frac{\partial\underline{\hat{D}}}{\partial R}\bigg|_{R=R_0}(R - R_0) + \cdots$$

$$I(\nu, \nu') = \int \overline{\psi_\nu}(R - R_0)\psi_\nu \, dR \tag{22.8}$$

$$\psi_\nu(x) = N_\nu H_\nu(\alpha x)e^{-\frac{\alpha^2 x^2}{2}}$$

Using the following recursion relation of the Hermite polynomials

$$2\alpha x H_\nu(\alpha x) = 2\nu H_{\nu-1}(\alpha x) + H_{\nu+1}(\alpha x)$$

We have that the integral will be nonzero if and only if $\Delta\nu = \pm 1$

The selection rules then are the following

$$
\begin{aligned}
\Delta J &= 0, \pm 1 \\
\Delta M &= 0, \pm 1 \\
\Delta \Lambda &= 0 \\
\Delta \nu &= \pm 1
\end{aligned}
$$

(22.9)

We can delve deeper into the $\Delta J = \pm 1$ transitions. We can define two different branches, i.e. R branches, if $\Delta J = -1$, and P branches when $\Delta J = 1$. The two sets will have the following separations between levels

$$
\begin{aligned}
\hbar\omega_R &= E_r(\nu + 1, J + 1) - E_r(\nu, J) = 2B(J + 1) + \hbar\omega_0 \\
\hbar\omega_P &= E_r(\nu - 1, J - 1) - E_r(\nu, J) = \hbar\omega_0 - 2BJ
\end{aligned}
$$

(22.10)

Both branches form what's usually called as *vibrational-rotational branch*, which is formed by lines spaced by $2B/h$.

There exists a third branch of the spectrum given by the anharmonicity of the oscillation of the molecule, the Q branch. If $B_\nu = B_{\nu+1}$, we have that

$$\hbar\omega_Q = E_r(\nu + 1, J) - E(\nu, J) = \hbar\omega_0$$

(22.11)

§§ 22.2.1 Raman Scattering

We get back to our consideration of homonuclear diatomic molecules. We saw that for $\Lambda = 0$ no transitions are possible due to the symmetry of the molecule, this is not properly true, since there can be a particular kind of transitions given by an inelastic scattering, called Raman scattering.

Raman scattering works as a second order process, where a photon with energy $\hbar\omega$ is absorbed from an atom or a molecule, which is excited from a state a to a state n, and then emits a second photon with energy $\hbar\omega'$ while decaying to the final state b. In this case, if $a = b$, we get again Rayleigh scattering. In other cases, using the conservation of energy we have

$$\hbar\omega' = \hbar\omega + \Delta E_{ab}$$

(22.12)

In general, this kind of scattering permits the existence of another selection rule, valid only for Raman scattering

$$\Delta J = 0, \pm 2$$

(22.13)

Let's consider now two particular cases. If the state a is the ground state of the molecule, the state b must ave higher energy, so $\omega' < \omega$, and the observed line is called a *Stokes line*, whereas in the opposite case, we have that $\omega' > \omega$ and this line is called the *Anti-Stokes line*

§ 22.3 Electronic Spectra

The electronic spectra of diatomic molecules is given by combined electronic-rovibronic transitions. The lines associated with this transitions lay on the visible or in the UV part of the light spectrum. We can immediately say that the frequency separation of these lines will be

$$\omega = \frac{(E_{s'} + E_{\nu'} + E_{r'}) - (E_s + E_\nu + E_r)}{\hbar} \tag{22.14}$$

This is obviously given by three different components, i.e. the *vibrational* or *band structure*, the *rotational structure* and the *fine structure* of the total band.

Keeping only the first two terms and ignoring rotational variations we can see that in general, the transitions will have the following energy separation

$$\hbar\omega = \hbar\omega_{s's} + \hbar\omega_0'\left(\nu' + \frac{1}{2}\right) - \hbar\omega_0\left(\nu + \frac{1}{2}\right) \tag{22.15}$$

Or, introducing again the anharmonicity of the oscillator, we get the *Deslandres formula*

$$\hbar\omega_D = \hbar\omega_{s's} + \hbar\omega_0'\left(\nu' + \frac{1}{2}\right) - \hbar\beta'\omega_0'\left(\nu' + \frac{1}{2}\right)^2 - \hbar\omega_0\left(\nu + \frac{1}{2}\right) + \hbar\beta\omega_0\left(\nu + \frac{1}{2}\right)^2 \tag{22.16}$$

The series of lines obtained from these transitions is called a ν *progression*.

Adding back the E_r term, we are obliged to use the selection rules for J.

If we now consider spin, in absence of coupling we get $\Delta S = 0$ and for transitions between Σ states we have $\Sigma^+ \leftrightarrow \Sigma^+$ and $\Sigma^- \leftrightarrow \Sigma^-$. Due to symmetry reasons we also have the selection rule $g \leftrightarrow u$.

The three branches, having considered the new selection rules and considering the centrifugal distortion

$$\begin{aligned}
\hbar\omega_P &= \hbar\omega_D + B'J(J-1) - BJ(J+1) \\
\hbar\omega_Q &= \hbar\omega_D + B'J(J+1) - BJ(J+1) \\
\hbar\omega_R &= \hbar\omega_D + B'(J+1)(J+2) - BJ(J+1)
\end{aligned} \tag{22.17}$$

It's evident that these formulas, after substituting the Deslandres equation for $\hbar\omega_D$, are not linear nor quadratic in J, but rather *parabolic* in J, which give a *Fortrat parabola*. Since $B' \neq B$ we also see that the lines aren't equally spaced, which is closely tied to the centrifugal distortion that we have already treated

§§ 22.3.1 The Franck-Condon Principle

The Franck-Condon principle is based on the idea that the atoms of the molecule do not move during the electronic transition but after. This is represented on the energy graph of the molecule as a vertical line between two electronic curves, where the centrifugal distortion gets considered.

This principle can be seen in action considering the total wavefunction $\Psi_\alpha = R^{-1}\Phi_s\psi_\nu\phi_{JM\Lambda}$ and evaluating the dipole moment operator's matrix elements

$$\underline{\hat{D}}_{\alpha'\alpha} = e\int R^{-2}\,\mathrm{d}^3R\int \overline{\Phi_{s'}\psi'_{\nu'}\phi'_{J'M'\Lambda'}}\left(\sum_{i=1}^{2}Z_iR_{ij} - \sum_i r_{ij}\right)\Phi_s\psi_\nu\phi_{JM\Lambda}\,\mathrm{d}^3r \quad (22.18)$$

From the definition of the dipole moment operator we take only the electronic part of the dipole operator

$$\underline{\hat{D}}_{e^-}(R) = -e\langle s|\sum_i \hat{r}_{ij}|s\rangle \rightarrow -e\int \overline{\Phi}_s\sum_i \hat{r}_{ij}\Phi_s\,\mathrm{d}^3r$$

The Franck-Condon principle then amounts to saying that this operator is independent of R, so that the transition amplitudes will be always proportional to the *Franck-Condon factor*

$$f_{\nu'\nu} = \int \overline{\psi'_{\nu'}}\psi_\nu\,\mathrm{d}R \quad (22.19)$$

This simply represents the overlap integral between two vibrational wavefunctions in different electronic states.

§§§ 22.3.1.1 Fluorescence and Phosphorescence

One event that can be explained using the Franck-Condon principle is fluorescence and phosphorescence, which is an effect given by some molecules, where the radiation absorbed in the near-UV gets re-emitted at a longer wavelenght, whereas in phosphorescence, it's involved the decay from an excited state to a second state with different multiplicity.
Fluorescence can be graphed as a transition from an electronic level to a second level, respecting the Franck-Condon principle. Let's take as an example a molecule in the ground state ^1X, which get excited to a second state ^1A, which slowly decays level by level up until getting back down to the ^1X state.
In this case, the effect of fluorescence can be seen as the slow decay from the excited state ^1A to the ground state ^1X, where all the decays emit photons in the range of visible and near-UV light.
Phosphorescence, instead, can be seen as a process $^1X \rightarrow^1 A \rightarrow^3 A \rightarrow^1 X$, where the molecule "slips" from the state ^1A to the state ^3A due to the nonzero multiplicity of this state.

§ 22.4 Inversion Spectrum of Ammonia ($\mathrm{NH_3}$)

The ammonia molecule is a pyramidal molecule composed by a nitrogen atom at the summit and three hydrogens at the basis
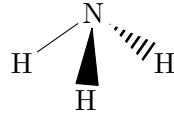
Figure 22.1: Structure of the Ammonia molecule (NH_4)

This geometry is twice degenerate, since we can invert the molecule using the nitrogen atom as an inversion point.

The vibrational modes of this molecule can be seen as an umbrella opening and closing with respect to the hydrogen plane of the molecule. The nitrogen atom can overpass this plane, hence the vibrational potential must have two wells for the two possible stable oscillation points.

Considering this well and doing a couple calculations, we can see that the transition between vibrational states $\nu = 0 \to \nu = 1$ is classically forbidden, but possible in the framework of quantum tunneling.

The tunneling of the wavefunction permits the removal of the degeneracy, hence forming a doublet.

In order to define properly this situation we consider the whole molecule as a two level system, where the upper level is the "up" configuration, where the nitrogen atom rests above the hydrogen plane, and a second "down" configuration where the nitrogen atom is below the hydrogen plane.

The wavefunctions representing these states are eigenvalues of the parity operator, and therefore we can write

$$|1\rangle = \frac{1}{\sqrt{2}} \left(|u\rangle + |d\rangle\right)$$
$$|2\rangle = \frac{1}{\sqrt{2}} \left(|u\rangle - |d\rangle\right)$$

(22.20)

We also have that the nitrogen atom could *probably* be above or below the hydrogen plane, so we can also write a third wavefunction. We will directly write the time-dependent wavefunction of the system, which is the following

$$\Psi(z,t) = \frac{1}{\sqrt{2}} \left(\psi_1(z)e^{-i\frac{E_1 t}{\hbar}} + \psi_2(z)e^{-i\frac{E_2 t}{\hbar}}\right) = \frac{1}{\sqrt{2}} \left(\psi_1(z) + \psi_2(z)e^{-2\pi i\omega t}\right)e^{-i\frac{E_1 t}{\hbar}}$$

(22.21)

Where we used $\Delta E = \hbar\omega$.

Explicating the two wavefunctions and setting $t = 1/4\pi\omega$ we get

$$\Psi(z,t) = \psi_d(z)e^{-i\frac{E_1 t}{\hbar}}$$

(22.22)

So that we have that the probability density of the wavefunction is simply $|\psi_d(z)|^2$

23 Molecular Electronic Structure

It's now time to analyze the electronic structure of diatomic molecules. We begin with the usual multielectron Hamiltonian without fine and hyperfine corrections. We choose for easier calculus the z-axis as the internuclear axis, and in this case we can immediately say that

$$\left[\hat{\mathcal{H}}, \hat{L}_z\right] = 0$$

It's also obvious that due to the nature of the system, $\hat{L}_x, \hat{L}_y, \hat{L}^2$ «do not» commute with the Hamiltonian of this system (this can be proved through direct calculus). Due to the previous statement, we know that the electronic eigenfunctions of the molecular Hamiltonian Φ_s are simultaneous eigenfunctions of $\hat{\mathcal{H}}$ and $\hat{L}_z$.

Since $\hat{L}_z \to i\hbar\partial_\phi$, we can solve directly for the angular part through separation of variables. We suppose that $\Phi_s(r_i; R) = f_r(R; r_i)\eta(\phi)$, solving then for η we have

$$\eta(\phi) = e^{iM_L\phi} \tag{23.1}$$

In molecular physics tho is common to use a new quantum number, $\Lambda = |M_L|$, therefore, we can then write for Φ_s (with appropriate normalization)

$$\Phi_s(R; r_i) = \frac{1}{\sqrt{2\pi}} f_s(r) e^{\pm i\Lambda\phi} \tag{23.2}$$

In the usual spectroscopic notation, we now have m_l for single particles, M_L for atoms and Λ for molecules, and it follows a similar pattern. It's also used $\lambda = |m_l|$ for single molecular electrons

Quantum Number	0	1	2	3...
m_l	s	p	d	f...
M_L	S	P	D	F...
Λ	Σ	Π	Δ	Φ...
λ	σ	π	δ	ϕ...

Table 23.1: Table of quantum numbers in spectroscopic notation

Let's get back to our molecular system. Through a quick sketch of a diatomic molecule it's obvious that, if the z-axis is taken as the intermolecular axis, the system is invariant for

reflections along the $x - y$ plane. Let's call this reflection operator $\hat{A}_{xy}$. Since it's a reflection we know that $\hat{A}_{xy}^2 = \hat{\mathbb{1}}$ and its eigenvalues are ± 1, and from the previous consideration that

$$\left[\hat{\mathcal{H}}, \hat{A}_{xy}\right]$$

Writing out explicitly the angular momentum operator $\hat{L}_z = \hat{\Lambda}_z$ (using $\hat{\Lambda}$ makes it easier to distinguish this operator from the usual angular momentum operators) in Cartesian form ($\hat{\Lambda}_z \to i\hbar y \partial_x - i\hbar x \partial_y$), and knowing that $\hat{A}_{xy}$ maps $y \to -y$, we also have that these two operators anticommute.

$$\left\{\hat{\Lambda}_z, \hat{A}_{xy}\right\} = 0$$

We therefore have that, for states with $\Lambda \neq 0$

$$\hat{\Lambda}_z \hat{A}_{xy} \left|\Lambda\right\rangle = -\hat{A}_{xy}\hat{\Lambda}_z \left|\Lambda\right\rangle = \pm \hbar \Lambda \left|\Lambda\right\rangle \tag{23.3}$$

Depending on the symmetry of the wavefunction, this brings up what is usually called Λ-doubling, i.e. a two-fold degeneracy on levels given by the symmetries of the system. A particular case happens when $\Lambda = 0$, in this case we can construct simultaneous eigenfunctions of $\hat{\mathcal{H}}, \hat{\Lambda}_z, \hat{A}_{xy}$, and the degeneracy is broken into two non-degenerate states, Σ^+, Σ^-, where for the first the wavefunction is unchanged on reflections along the internuclear axis, and for the second state that the wavefunction changes sign ($\hat{A}_{xy}\left|a\right\rangle = \pm\left|a\right\rangle$).
Therefore, for inversions of the kind $\underline{r}_i \to -\underline{r}_i$, we can define states which are invariant and non-invariant to this transformation, these will be indicated by either a subscript g or u, which come from the German *gerade* and *ungerade*, which mean respectively "even" and "odd".
For homonuclear diatomic molecules, the behavior of the Σ wavefunctions through the inversions $y_i \to -y_i$ and $\underline{r}_i \to -\underline{r}_i$, give the additional definition of 4 non degenerate Σ states, respectively $\Sigma_g^+, \Sigma_g^-, \Sigma_u^+, \Sigma_u^-$.
These states can be determined precisely by considering the inversion $\underline{R} \to -\underline{R}$, made by the composition of the inversion $y_i \to -y_i$ and $\underline{r}_i \to -\underline{r}_i$. The result of these inversion will be an unchanged sign for Σ_g^+, Σ_u^- wavefunctions and a changed sign for Σ_g^-, Σ_u^+ wavefunctions.

Considering now the spin of the molecular electrons, we're left with the usual operator $\hat{\underline{S}}$, with the usual eigenvalues. With spin, we can write the terms as follows

$$^{2S+1}\Lambda_{g/u}$$

The ground states are usually indicated as $X^{2S+1}\Lambda_{g/u}$, and for diatomic molecules, they are usually $X^1\Sigma^+$ and $X^1\Sigma_g^+$ for homonuclear diatomic molecules, with some exceptions[1]

---

[1]See O_2 and NO, which have as ground states $X^3\Sigma_g$ and $X^2\Pi$, respectively

§ 23.1 Approximation Methods and the H_2 Molecule

§§ 23.1.1 Linear Combination Of Atomic Orbitals (LCAO-MO)

A great example of a molecular system that can be analyzed, is the *Dihydrogen cation*, i.e. the molecule H_2^+.

Considering all particles present in the system, we can write our Hamiltonian as follows (using atomic units $\hbar = 1, k_e = 1, e = 1, m_e = 1$)

$$\hat{\mathcal{H}}_{H_2^+} = -\frac{1}{2}\nabla_r^2 - \frac{1}{r_a} - \frac{1}{r_b} + \frac{1}{R} \tag{23.4}$$

Where r, r_a, r_b are not independent, given in terms of the distance between the two nuclei R, as

$$\underline{r_a} = \underline{r} + \frac{R}{2}$$

$$\underline{r_b} = \underline{r} - \frac{R}{2}$$

The Schrödinger equation can be exactly solved, but it's useful to firstly develop the Linear Combination of Atomic Orbitals (LCAO) approximation technique.

As we have already seen, at great distances, we must have that the system is a simple hydrogen atom, hence

$$|\Phi\rangle = |1s\rangle$$

Where the $1s$ orbital takes into consideration whether it's bound to the nucleus a or b.

From this, we can construct two molecular wavefunctions with either an even (gerade) symmetry or an uneven (ungerade) symmetry.

We have

$$|g\rangle = \frac{1}{\sqrt{2}}\left(|1s\rangle_a + |1s\rangle_b\right)$$

$$|u\rangle = \frac{1}{\sqrt{2}}\left(|1s\rangle_a - |1s\rangle_b\right) \tag{23.5}$$

We then can find our molecular energy levels by plugging it into our variational equation

$$E_{g,u}(R) = \frac{\langle g, u | \hat{\mathcal{H}} | g, u \rangle}{\langle g, u | g, u \rangle}$$

Let's calculate firstly the normalization of the state $|g, u\rangle$. We have

$$\langle g, u | g, u \rangle = \frac{1}{2}\left(\langle 1s|_a \pm \langle 1s|_b\right)\left(|1s\rangle_a \pm |1s\rangle_b\right)$$

$$= \frac{1}{2}\left(\langle 1s|1s\rangle_a + \langle 1s|1s\rangle_b \pm 2\langle 1s|1s\rangle_{ab}\right) = \tag{23.6}$$

$$= 1 \pm \langle 1s|1s\rangle_{ab} = 1 \pm I(R)$$

Where $I(R) = \langle 1s|1s\rangle_{ab}$ is an overlap integral, which can be calculated, considering that

$$\langle r|1s\rangle = \psi_{1s}(r) = \frac{1}{\pi}e^{-r}$$

A piece-by-piece calculation of this and some further integrals will be given in an appendix, but for now, we have that the result is

$$I(R) = \left(1 + R + \frac{1}{3}R^2\right)e^{-R}$$

We then evaluate the numerator, keeping in mind that these two Schrödinger equations also hold

$$\left(\frac{1}{2}\hat{p}_r^2 - \frac{1}{r_a}\right)|1s\rangle_a = E_{1s}|1s\rangle_a$$

$$\left(\frac{1}{2}\hat{p}_r^2 - \frac{1}{r_b}\right)|1s\rangle_b = E_{1s}|1s\rangle_b$$

So, we then have, expanding $|g, u\rangle$ in its composing kets

$$\langle g, u|\hat{\mathcal{H}}|g, u\rangle = \frac{1}{2}\left((\langle 1s|_a \pm \langle 1s|_b)\hat{\mathcal{H}}(|1s\rangle_a \pm |1s\rangle_b)\right)$$

$$= \langle 1s|\hat{\mathcal{H}}|1s\rangle_a \pm \langle 1s|\hat{\mathcal{H}}|1s\rangle_{ab} \tag{23.7}$$

Writing the Hamiltonian and distributing once again, we get

$$\langle g, u|\hat{\mathcal{H}}|g, u\rangle = \left(E_{1s} + \frac{1}{R}\right) - \langle 1s|\frac{1}{r_b}|1s\rangle_a \pm \left(E_{1s} + \frac{1}{R}\right)\langle 1s|1s\rangle_{ab} \pm \langle 1s|\frac{1}{r_b}|1s\rangle_{ab}$$

$$= E_{1s}(1 \pm I(R)) \pm \frac{1}{R}I(R) + \frac{1}{R} - \langle 1s|\frac{1}{r_b}|1s\rangle_a \pm \langle 1s|\frac{1}{r_b}|1s\rangle_{ab} \tag{23.8}$$

Solving the two last integrals and putting everything together, we finally get

$$E_{g,u}(R) = E_{1s} + \frac{(1 + R)e^{-2R} \pm \left(1 - \frac{2}{3}R^2\right)e^{-R}}{R \pm \left(1 + R + \frac{1}{3}R^2\right)Re^{-R}} \tag{23.9}$$

From this, we can then define the two following levels

$$\hat{\mathcal{H}}_{H_2^+}|g\rangle = E_g(R)|g\rangle$$

$$\hat{\mathcal{H}}_{H_2^+}|u\rangle = E_u(R)|u\rangle$$

The first energy is the one for which the wavefunction is symmetrical ($\pm \to +$). For this wavefunction, if plotted or calculated, it's easy to see that there exists a well, for which the molecule can bond, and therefore, this molecular orbital is said to be a *bonding orbital*, indicated by the term σ_g, in the other case, for the uneven wavefunction, there is no energetic well, hence it's easy to see how this orbital doesn't contribute to the bonding of the molecule, hence it's called an *antibonding orbital*, and it's indicated by the term $\sigma_u^\star$ The second approach, after the molecule H_2^+, is the actual hydrogen molecule H_2^+, which has two electrons.

We shall build its electronic wavefunctions using what we found for the dihydrogen cation. We already know that the eigenspinors must either be a singlet state $S = \sum_i s_i = 0$

or a triplet state. We can immediately write these spinors, starting from the singlet state (remembering that $a \otimes b \neq b \otimes a$, for two vectors $a, b \in \mathbb{H}$, here due to the indistinguishability of the two electrons it only matters to know that the first ket is referring to the first electron and the second ket, obviously, to the second and last electron)

$$
\begin{aligned}
|00\rangle &= \frac{1}{\sqrt{2}} \left[|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle \right] \\
|11\rangle &= |\uparrow\rangle \otimes |\uparrow\rangle \\
|10\rangle &= \frac{1}{\sqrt{2}} \left[|\uparrow\rangle \otimes |\downarrow\rangle + |\downarrow\rangle \otimes |\uparrow\rangle \right] \\
|1-1\rangle &= |\downarrow\rangle \otimes |\downarrow\rangle
\end{aligned}
\tag{23.10}
$$

Using the two eigenfunctions $|u\rangle, |g\rangle$ we can then form four combinations of the two

$$
\begin{aligned}
|A\rangle &= [|g\rangle \otimes |g\rangle] \otimes |00\rangle \\
|B\rangle &= [|u\rangle \otimes |u\rangle] \otimes |00\rangle \\
|C\rangle &= \frac{1}{\sqrt{2}} \left[|g\rangle \otimes |u\rangle + |u\rangle \otimes |g\rangle \right] \otimes |00\rangle \\
|D\rangle &= \frac{1}{\sqrt{2}} \left[|g\rangle \otimes |u\rangle - |u\rangle \otimes |g\rangle \right] \otimes |1, M_S\rangle
\end{aligned}
\tag{23.11}
$$

It's immediate to see that, since $|g\rangle, |u\rangle$, as calculated previously for H_2^+, are either σ_g or $\sigma_u^\star$ states, that $|A\rangle, |B\rangle$ represent $^1\Sigma_g^+$, while $|C\rangle$ represents a $^1\Sigma_u^+$ state, and $|D\rangle$ represents the remaining $^3\Sigma_u^+$ states. The exact Hamiltonian for the Hydrogen molecule is, in atomic units

$$
\hat{\mathcal{H}}_{H_2} = -\frac{1}{2} \left(\nabla_1^2 + \nabla_2^2 \right) - \sum_{i=1}^{2} \frac{1}{r_{ai}} - \sum_{i=1}^{2} \frac{1}{r_{bi}} + \frac{1}{r_{12}} + \frac{1}{R}
\tag{23.12}
$$

Regrouping the two single electron Hamiltonians as $\hat{h}_i$, and remembering that

$$
\hat{h}_i |g, u\rangle = \left(E_{g,u}(R) - \frac{1}{R} \right) |g, u\rangle
$$

We get by using the eigenket $|A\rangle$, imposing without loss of generality that $\langle g, u | g, u \rangle = 1$, that the associated energy will be, after plugging everything into the Rayleigh-Ritz variational expression, that

$$
\begin{aligned}
E_A &= \langle 00 | \langle g | \otimes \langle g | \hat{\mathcal{H}} | g \rangle \otimes | g \rangle |00\rangle \\
&= \langle 00 | \langle g | \otimes \langle g | \sum_{i=1}^{2} \hat{h}_i + \frac{1}{r_{12}} | g \rangle \otimes | g \rangle |00\rangle \\
&= 2 E_g(R) - \frac{1}{R} + \langle 00 | \langle g | \otimes \langle g | \frac{1}{r_{12}} | g \rangle \otimes | g \rangle |00\rangle
\end{aligned}
\tag{23.13}
$$

The last expression, corresponds to the following integral

$$
\iint \overline{\Phi}_g(r_1) \overline{\Phi}_g(r_2) \frac{1}{r_{12}} \Phi_g(r_1) \Phi_g(r_2) \, \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 = \iint \frac{|\Phi_g(r_1)|^2 |\Phi_g(r_2)|^2}{r_{12}} \, \mathrm{d}^3 r_1 \mathrm{d}^3 r_2
$$

This integral can be computed using the approximate LCAO form, or the "exact" solution of the Schrödinger equation for the H_2^+ molecule.

Using the LCAO form of $|g\rangle$, we can then write in extensive form

$$
\begin{aligned}
|A\rangle = |g\rangle_1 \otimes |g\rangle_2 |00\rangle = \\
= \frac{1}{2} \left(|1s\rangle_a \otimes |1s\rangle_b + |1s\rangle_b \otimes |1s\rangle_a + |1s\rangle_a \otimes |1s\rangle_a + |1s\rangle_b \otimes |1s\rangle_b \right) |00\rangle
\end{aligned}
\tag{23.14}
$$

Or, using directly the Schrödinger representation of eigenstates

$$
\begin{aligned}
\langle r_{ij}|A\rangle = \Phi_A = \frac{1}{2} \left(\psi_{1s}(r_{a1})\psi_{1s}(r_{b2}) + \psi_{1s}(r_{b1})\psi_{1s}(r_{a2}) + \right. \\
\left. + \psi_{1s}(r_{a1})\psi_{1s}(r_{a2}) + \psi_{1s}(r_{b1})\psi_{1s}(r_{b2}) \right) \chi_{00}(r_1, r_2)
\end{aligned}
\tag{23.15}
$$

It's easy to see how this eigenstate can be expressed by a superposition of two states, as $|A\rangle = |A\rangle_{cov} + |A\rangle_{ion}$, which respresent two types of molecular bonding already known in chemistry: *covalent and ionic bonding.* Using the previous expression, we can write these two kinds of bonding as follows

$$
\begin{aligned}
|A\rangle_{cov} = \frac{1}{2} \left(|1s\rangle_a \otimes |1s\rangle_b + |1s\rangle_b \otimes |1s\rangle_a \right) |00\rangle \\
|A\rangle_{ion} = \frac{1}{2} \left(|1s\rangle_a \otimes |1s\rangle_a + |1s\rangle_b \otimes |1s\rangle_b \right) |00\rangle
\end{aligned}
\tag{23.16}
$$

It's immediate to see how the covalent eigenstate ($|A\rangle_{cov}$) represents a situation where the two electrons (remember that we're working in a Hilbert space of the kind $\mathbb{H}_1 \otimes \mathbb{H}_2$) are bound to both nuclei, whereas the ionic eigenstate represents a situation where both electrons are bound to one nucleus. Using the LCAO approximation, in the limit $R \to \infty$, the covalent bond yields two hydrogen atoms ($H_2 \to H + H$), whereas the second yields one proton and a negative hydrogen ion ($H_2 \to H^- + p^+$).

A better approximation can be given using the Rayleigh-Ritz variational principle, with a trial eigenket $|\lambda\rangle$, formed by an equal mixture of the two states $|A\rangle$ and $|B\rangle$, which both have symmetry $^1\Sigma_g^+$.

Noting that the state $|B\rangle$ is the following

$$
|B\rangle = \frac{1}{2} \left(|1s\rangle_A \otimes |1s\rangle_A - |1s\rangle_A \otimes |1s\rangle_B - |1s\rangle_B \otimes |1s\rangle_A + |1s\rangle_B \otimes |1s\rangle_B \right) |00\rangle
$$

The eigenstate $|\lambda\rangle$ will then have the following definition

$$
\begin{aligned}
|\lambda\rangle &= |A\rangle + \lambda |B\rangle \\
|\lambda\rangle &= \frac{1}{2} \left[(1 - \lambda) \left(|1s\rangle_A \otimes |1s\rangle_B + |1s\rangle_B \otimes |1s\rangle_A \right) \right. \\
&\quad \left. + (1 + \lambda) \left(|1s\rangle_A \otimes |1s\rangle_A + |1s\rangle_B \otimes |1s\rangle_B \right) \right] |00\rangle \\
|\lambda\rangle &= (1 - \lambda) |A\rangle_{cov} + (1 + \lambda) |A\rangle_{ion}
\end{aligned}
\tag{23.17}
$$

Using the Rayleigh Ritz variational method, it's possible to calculate the energy of the system in relation to the parameter λ, as follows

$$E(\lambda) = \frac{\langle \lambda | \hat{\mathcal{H}} | \lambda \rangle}{\langle \lambda | \lambda \rangle} \tag{23.18}$$

Finding the extremum, (i.e. where $\delta E(\lambda) = 0$, i.e. $\partial_\lambda E = 0$ and λ_0 is an extremum) it is possible to find values closer to the experimental values than the previous approximations.

§§ 23.1.2 Heiter-London Valence Bond Method (VB-MO)

Considering now the *Valence Bond method*, we approximate the wavefunction for H_2 basing ourselves on the separated atom wavefunctions.
We begin taking the triplet sigma wavefunction $|D\rangle$, for which, the covalent part is the following

$$|D\rangle_{cov} = \frac{1}{2} \left(|1s\rangle_A \otimes |1s\rangle_B - |1s\rangle_B \otimes |1s\rangle_A \right) |1M_S\rangle \tag{23.19}$$

This wavefunction has the symmetry $^3\Sigma_u^+$. Now, substituting this and $|A\rangle_{cov}$ for the $^1\Sigma_g^+$ term in the variational equation, we get that the searched energy for gerade and ungerade states is the following

$$E_{g,u}(R) = 2E_{1s} + \frac{J \pm K}{1 \pm I^2} + \frac{1}{R} \tag{23.20}$$

Where (using Schrödinger's representation), we have that

$$I = \int \psi_{1s}(r_{A1})\psi_{1s}(r_{B1}) \, \mathrm{d}^3 r_1$$

$$J = \iint |\psi_{1s}(r_{A1})|^2 \left(\frac{1}{r_{12}} - \frac{1}{r_{A2}} - \frac{1}{r_{B1}} \right) |\psi_{1s}(r_{B2})|^2 \, \mathrm{d}^3 r_1 \mathrm{d}^3 r_2$$

$$K = \iint \overline{\psi}_{1s}(r_{A1})\overline{\psi}_{1s}(r_{B2}) \left(\frac{1}{r_{12}} - \frac{1}{r_{A2}} - \frac{1}{r_{B1}} \right) \psi_{1s}(r_{A2})\psi_{1s}(r_{B1})$$

In order to get these two results we begin by inserting everything into the Rayleigh-Ritz equation.
For the $^3\Sigma_u^+$ we have

$$E_u(R) = \frac{\langle D | \hat{\mathcal{H}} | D \rangle_c}{\langle D | D \rangle_c}$$

We begin calculating the normalization factor $\langle D | D \rangle$

$$\langle D | D \rangle_c = \frac{1}{4} \left(\langle 1s |_A \otimes \langle 1s |_B - \langle 1s |_B \otimes \langle 1s |_A \right) \otimes \left(|1s\rangle_A \otimes |1s\rangle_B - |1s\rangle_B \otimes |1s\rangle_A \right) =$$

$$= \frac{1}{4} \left(2 \langle 1s | 1s \rangle_A \langle 1s | 1s \rangle_B - 2 \langle 1s | 1s \rangle_{AB} \langle 1s | 1s \rangle_{BA} \right) =$$

$$= \frac{1}{2} \left(1 - \langle 1s | 1s \rangle_{AB1}^2 \right)$$

Where, $\langle 1s|1s \rangle_{AB1} = \langle 1s|1s \rangle_{AB2} = I$

Secondly, we find the expectation value of the Hamiltonian $\left\langle \hat{\mathcal{H}} \right\rangle_{D_c}$. We begin by noting that the Hamiltonian can be separated into the sum of three Hamiltonians

$$\hat{\mathcal{H}} = \hat{h}_a + \hat{h}_b + \hat{h}_{12} = 2\hat{h}_a - \hat{h}_{12}$$

$$\hat{h}_a = \frac{\hat{p}^2}{2} - \frac{1}{r_{a1}}$$

$$\hat{h}_b = \frac{\hat{p}^2}{2} - \frac{1}{r_{b2}}$$

$$\hat{h}_{12} = \frac{1}{r_{12}} + \frac{1}{R} - \frac{1}{r_{a2}} - \frac{1}{r_{b1}}$$

In order to ease the notation in this calculation we will indicate $|1s\rangle_i \otimes$ as $|1\rangle_i$ and $\otimes |1s\rangle_i$ as $|2\rangle_i$

In this notation, we have $|D\rangle_c = 1/2(|1\rangle_A |2\rangle_B - |1\rangle_B |2\rangle_A)$, and therefore, remembering that $\hat{\mathcal{H}}^\dagger = \hat{\mathcal{H}}$ always holds, we have

$$\langle D| \hat{\mathcal{H}} |D\rangle_c = 2 \langle D| \hat{h}_a |D\rangle_c - \langle D| \hat{h}_{12} |D\rangle_c =$$

$$= 2 \left(h^a_{1212} + h^a_{2112} \right) + h^{12}_{1212} - h^{12}_{2112} =$$

$$= \left(4E_{1s} + \frac{2}{R} \right) (1 - \langle 1s|1s \rangle^2_{AB1}) + \hat{h}^{12}_{1212} - \hat{h}^{12}_{2112}$$

Writing explicitly the integrals in the last equation for the matrix elements of $\hat{h}_{12}$, we get the J integral for the first and the K integral for the second, therefore obtaining the results

$$\langle D| \hat{\mathcal{H}} |D\rangle_c = \left(4E_{1s} + \frac{2}{R} \right) (1 - I^2) + \frac{1}{2} (J - K)$$

$$\langle D|D\rangle_c = \frac{1}{2} (1 - I^2)$$

(23.21)

Putting it up altogether, we get for the $^3\Sigma_u^+$

$$E_u(R) = 2E_{1s} + \frac{1}{R} + \frac{J - K}{1 - I^2} \tag{23.22}$$

For the $^1\Sigma_g^+$ state, everything remains equal, except for the sign $-$, that becomes a $+$ due to the definition of $|A\rangle_c$, and therefore we have

$$E_g(R) = 2E_{1s} + \frac{1}{R} + \frac{J + K}{1 + I^2} \tag{23.23}$$

A question arises now, after all these calculations. One might ask what do I, J, K actually mean physically. The first is the overlap integral between the nucleus A and the nucleus B of the two electrons (it is squared since the two electrons are indistinguishable), J is the Coulomb integral which represents the interactions between the charge densities of the two electrons $(-e|\psi_{1s}(r_{A1})|^2$ and $-e|\psi_{1s}(r_{B2})|^2)$ and K is the exchange integral, each of these already known through the Hartree-Fock approximation for many-electron atoms.

§ 23.2 Homonuclear Diatomic Molecules

In general, for homonuclear diatomic molecules, we can write the electronic wavefunction for the molecular orbitals in the LCAO approximation as

$$|g, u\rangle = N_{g,u} \left(|a\rangle_{Ai} \pm |b\rangle_{Bi} \right)$$

Where $N_{g,u}$ is an appropriate normalization factor, and $|a\rangle, |b\rangle$ are two (non necessarily equal) atomic orbitals, with respect to the first or the second atom.
Studying more the molecular orbitals we have that

1. MOs with a given value of λ must connect with atomic orbitals with the same value of $|m_l|$.

2. The parity of the wavefunction (g, u) must be preserved

From spatial MOs it's possible to build spinorbitals for ech electrons, using Slater determinants.
As an example, we can write the state $|A\rangle$ for H_2 as follows

$$|A\rangle = \frac{1}{\sqrt{2}} \det_2 \begin{vmatrix} |g\rangle |\uparrow\rangle_1 & |g\rangle |\downarrow\rangle_1 \\ |g\rangle |\uparrow\rangle_2 & |g\rangle |\downarrow\rangle_2 \end{vmatrix} \tag{23.24}$$

In this case, both electrons of the H atoms can be in the bonding σ_g orbital, but in higher electron number systems, due to Pauli's principle, is not possible anymore.
For He_2, He_2^+ we have in the first case that the molecular configuration is $\sigma_g^2 \sigma_u^2$, which is unstable and hence bond-breaking. For the second case instead we're left with one less electron in the antibonding σ_u orbital, which corresponds to a weakly bond molecule.
Next in line of homonuclear diatomic molecules, there is Li_2, each atom with electronic configuration $[He]2s$. The molecule will be formed through the bonding of the $2s$ valence electron, forming a $(\sigma_g 2s)^2$ molecular orbital, which is a stable bond.
Going forward in the periodic table, we have Be_2, which must have a $\sigma_g^2 \sigma_u^2$ unstable configuration. For B_2 we have a $(\sigma_g 2p)^2$ state, given by the uncomplete LII atomic subshell, which corresponds to the $2p$ atomic orbital (the complete shell would be the $2s^2 2p$ valence shell for Boron). Going forwards, we reach the C_2 molecule. The C atom has a configuration $[He]2s^2 2p^2$. The 4 p electrons form two bonding orbitals, $\sigma_g^2 \pi_u^2$, which are both bonding, hence giving a stable molecule.
Finally we consider the O_2 molecule. Each atom has a configuration $[He]2s^2 2p^4$, which gives a molecular configuration $\sigma_g^2 \pi_u^4 \pi_g^2$. Following the Aufbau rules we know that the unfilled antibonding π_u orbital, must have one electron in the $\pi_x^\star$ and one in the $\pi_y^\star$ (remember that the star is usually used in literature to indicate an antibonding orbital, in our case the π_g orbital)

§§ 23.2.1 Valency and Chemical Bonding

It's interesting to see now what valence electrons and chemical bonds really mean in physics. Usually, in order to form a bonding molecular orbital we need to form a singlet spin state. Let's now imagine what happens when we bring an hydrogen atom close to an helium atom. Both the electrons in He are in a singlet spin state, and the Hydrogen atom can't pair with Helium, since it can't exchange with neither of the two He electrons, since such bond would have two electrons in the same state, violating Pauli's principle. We quickly study this three electron system. The total wavefunction will be the following Slater determinant

$$|\Phi\rangle = N \det_3 \begin{Vmatrix} |1s\rangle_{He} |\uparrow\rangle_1 & |1s\rangle_{He} |\downarrow\rangle & |1s\rangle_H |\uparrow\rangle_1 \\ |1s\rangle_{He} |\uparrow\rangle_2 & |1s\rangle_{He} |\downarrow\rangle_2 & |1s\rangle_H |\uparrow\rangle_2 \\ |1s\rangle_{He} |\uparrow\rangle_3 & |1s\rangle_{He} |\downarrow\rangle_3 & |1s\rangle_H |\uparrow\rangle_3 \end{Vmatrix} \tag{23.25}$$

With N our normalization constant. Substituting this wavefunction into the Rayleigh-Ritz variational expression, we have that $E(R) = J - K$, where J is the following direct integral

$$J = N^2 \langle 1s|_{He} \langle 1s|_{He} \langle 1s|_H \hat{\mathcal{H}} |1s\rangle_{He} |1s\rangle_{He} |1s\rangle_H$$

$$J = N^2 \int \overline{\psi_{1s}(1)\psi_{1s}(2)\phi_{1s}(3)} \hat{\mathcal{H}} \psi_{1s}(1)\psi_{1s}(2)\phi_{1s}(3) \ \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \mathrm{d}^3 r_3$$

$$\phi_{1s}(i) = \langle r_i|1s\rangle_H$$

$$\psi_{1s}(i) = \langle r_i|1s\rangle_{He}$$

$$\tag{23.26}$$

And K is the following exchange integral

$$K = N^2 \int \overline{\psi_{1s}(1)\psi_{1s}(2)\phi_{1s}(3)} \hat{\mathcal{H}} \psi_{1s}(3)\psi_{1s}(2)\phi_{1s}(1) \ \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \mathrm{d}^3 r_3 \tag{23.27}$$

The only exchange happening from this integral is between the electron 1 and the electron 3, which have the same spin, thus introducing a repulsion, which makes the existence of a stable HHe molecule impossible. The two electrons in the He atom are said to be *paired*, and only *unpaired* electrons contribute to chemical bonding. Due to this, since for atoms with closed valence subshells we have only paired valence electrons, they are said to be *chemically inert*. A chemical bond forms principally singlet states using unpaired electrons, forming states with $S = 0$, with the exception of the O_2 molecule, in which the two electrons in the antibonding π_g orbitals are in a relative triplet state, giving O_2's ground state a relative triplet state.

§ 23.3 Heteronuclear Diatomic Molecules

These methods of forming molecular orbitals can also be applied to heteronuclear molecules, formed by two different atoms. Since in this case there is no reflection symmetry along the internuclear axis, orbitals can't be classified as gerade or ungerade.

In general we will write our molecular orbital as a weighted combination of atomic orbitals of the atom A and the atom B

$$|\Phi\rangle = \lambda |u\rangle_A + \mu |v\rangle_B \tag{23.28}$$

A rule of thumb in order to write these molecular orbital comes from the chemical properties of the two elements, in fact, the more electronegative atom's atomic orbitals will have a greater weight, represented in a MO diagram as having a lower energy than the other atom's orbitals.

A quick way to grasp these concepts is to dive directly into the examples, in our case these will be LiH, HCl and NaCl molecules.

§§ 23.3.1 Lithium Hydride LiH

We begin by writing out the valence orbitals of these two compounds. We have that Lithium has a configuration $[\mathrm{He}]2s$, and Hydrogen has a configuration $1s$. The lowest lying molecular orbital will be the $1\sigma^2$ orbital formed by the complete K shell of Lithium, going higher in energy we find the 2σ orbital, that should be formed by the $1s$ orbital of Hydrogen and the $2s$ orbital of Lithium. Using the variational LCAO method, where as a trial function we take the superposition of the $2s$ and $2p$ orbitals of Lithium, in what is usually called *hybridization*, we can then find a lower energy eigenvalue. The new hybrid orbital is called an sp orbital, and will have a wavefunction of the following kind, if we write the atomic hydrogenoid orbitals as $|nlm\rangle$

$$|sp\rangle = c_1\,|200\rangle + c_2\,|210\rangle \tag{23.29}$$

This hybridization causes an excess negative charge, which gives a permanent electric dipole moment.

In other hydrides like BH, NH and HF, this negative charge excess still exists, but with an opposite sign.

§§ 23.3.2 Hydrogen Chloride HCl

The configuration of Chlorine is $[\mathrm{He}]3s^2 3p^5$, with the K and L shells completely filled and not participating in the chemical bonding.

Since energies of the $3s$ orbital of Chlorine aren't similar to the energies of Hydrogen's $1s$ orbital, these two do not mix, but the bond will be formed between the $1s$ orbital of Hydrogen and the $3p$ orbital of Chlorine. Only the p_z orbital can form a σ orbital, and therefore the searched bond will be formed by a superposition of the $3p_z$ of Chlorine with the $1s$ of Hydrogen.

The weight on the $3p_z$ in this case will be greater than the weight on the $1s$, therefore representing at $R \to \infty$ the combination $\mathrm{H}^+ + \mathrm{Cl}^-$, which represents a *ionic* bond

§§ 23.3.3 Sodium Chloride NaCl

Really good examples of molecules characterized by ionic bonds are compounds of an alkali atom with a halogen. In this case, the alkalis have a single valence electron outside a closed shell ($X^3 S_{1/2}$), whereas halogens miss only an electron in order to close the last shell.

This configuration makes sure that both atoms end up with a closed shell through bonding, which happens through a ionic bonding. Considering Sodium Chloride, we have that the

lowest energy configuration is given by the combination of the ions $Na^+ + Cl^-$. At small distances, this system works like inert gases, and is strongly repulsive.
Empirically, one can then write

$$E_s(R) = E_s(\infty) - \frac{1}{R} + Ae^{-cR} \tag{23.30}$$

§ 23.4 Triatomic Molecules

§§ 23.4.1 Hybridization

The construction of molecular orbitals for polyatomic molecules works in a similar way as it does for diatomic molecules.
We begin taking our molecular orbital $|\psi\rangle$ and expanding it into a basis of Gauss-wavefunctions or atomic orbitals $|\phi\rangle$

$$|\psi\rangle = \sum_{i=1}^{n} c_i |\phi\rangle$$

The coefficients c_i are determined through the Rayleigh-Ritz variational calculus. Due to the importance of geometry and symmetry groups (see Appendix G), we need this wavefunction as a basis of an irreducible translation of the molecular point groups.
We begin by searching the minimal potential energy for every wavefunction $|\psi\rangle_k$, with electronic state $|k\rangle$ for every single nucleus

$$\nabla_{R_i} V_k = 0$$

As noted before, in case we have different atoms participating in the binding, if the energy differences given by two different orbitals aren't too different, we get what's usually called hybrid orbital. Since in polyatomic molecules we mostly have heterogeneous systems, we might get different kinds of orbital hybridizations given by the intermixing of orbitals.
As we have already seen, with s and p orbitals, we can get what's called a sp orbital, if an s orbital intermixes with a p_z orbital.
It's not actually the only kind of hybridization that can happen between s and p orbitals, in fact we have a couple more possibilities.
One of these possibility is the sp^2 hybrid orbital, given by the triple mixing of an s orbital and two p_x, p_y orbitals, following the following pattern

$$|sp^2\rangle = \begin{cases} \frac{1}{\sqrt{3}} \left(|s\rangle + \sqrt{2} |p_x\rangle \right) \\ \frac{1}{\sqrt{3}} |s\rangle - \frac{1}{\sqrt{6}} |p_x\rangle + \frac{1}{\sqrt{2}} |p_y\rangle \\ \frac{1}{\sqrt{3}} |s\rangle - \frac{1}{\sqrt{6}} |p_x\rangle - \frac{1}{\sqrt{2}} |p\rangle_y \end{cases} \tag{23.31}$$

Adding up to this mix the orbital p_z, we get the sp^3 hybrid orbital, formed as follows

$$|sp^3\rangle = \begin{cases} \dfrac{1}{2}\left(|s\rangle + \sqrt{3}\,|p_z\rangle\right) \\[2mm] \dfrac{1}{2}\,|s\rangle + \sqrt{\dfrac{2}{3}}\,|p_x\rangle - \dfrac{1}{2\sqrt{3}}\,|p_z\rangle \\[2mm] \dfrac{1}{2}\,|s\rangle - \dfrac{1}{\sqrt{6}}\,|p_x\rangle + \dfrac{1}{\sqrt{2}}\,|p_y\rangle - \dfrac{1}{2\sqrt{3}}\,|p_z\rangle \\[2mm] \dfrac{1}{2}\,|s\rangle - \dfrac{1}{\sqrt{6}}\,|p_x\rangle - \dfrac{1}{\sqrt{2}}\,|p_y\rangle - \dfrac{1}{2\sqrt{3}}\,|p_z\rangle \end{cases} \tag{23.32}$$

There are more kinds of hybridization, and each one of these gives a different kind of geometry to the molecule, mainly written up on the following table

Hybrid Orbital	Geometry
sp, dp	Linear
p^2, sd	Bent
sp^2, s^2d	Trigonal Planar ($\alpha = 120°$)
p^3	Trigonal Pyramidal
sp^3	Tetrahedral
sp^3d	Bipyramidal
sp^3d^2	Octahedral

Table 23.2: Different kinds of hybridization and subsequent molecular geometry

§§ 23.4.2 Beryllium Dihydride ($\mathrm{BeH_2}$)

Our first real example will be Beryllium Dihydride. This molecule is linear, and considering that the valence shell of Beryllium is $[\mathrm{He}]2s^2$ we have that our complete molecular orbital can be written as an intermixing of s orbitals of the 3 atoms.
In formulae

$$|\psi\rangle_{BeH_2} = c_1\,|1s\rangle_H + c_2\,|2s\rangle_{Be} + c_3\,|1s\rangle_H \tag{23.33}$$

This molecular orbital has σ symmetry (it's not hard to see this, it's a superposition of s orbitals). Following the hybridization method of orbitals we can also find 3 more orbitals

$$|\psi\rangle_i = \begin{cases} |\psi\rangle_1 = |1s\rangle + \lambda\,|2s\rangle + |1s\rangle \\ |\psi\rangle_2 = |1s\rangle + \mu\,|2p_z\rangle - |1s\rangle \\ |\psi\rangle_3 = |1s\rangle - \nu\,|2s\rangle + |1s\rangle \\ |\psi\rangle_4 = -\,|1s\rangle + \omega\,|2p_z\rangle + |1s\rangle \end{cases} \tag{23.34}$$

All of these configurations represent the state $X^1\Sigma_g$ of Beryllium Dihydride.
It's easy also to verify that this molecule (like all linear molecules) can have either a $C_{\infty h}$ symmetry or a $D_{\infty h}$ symmetry
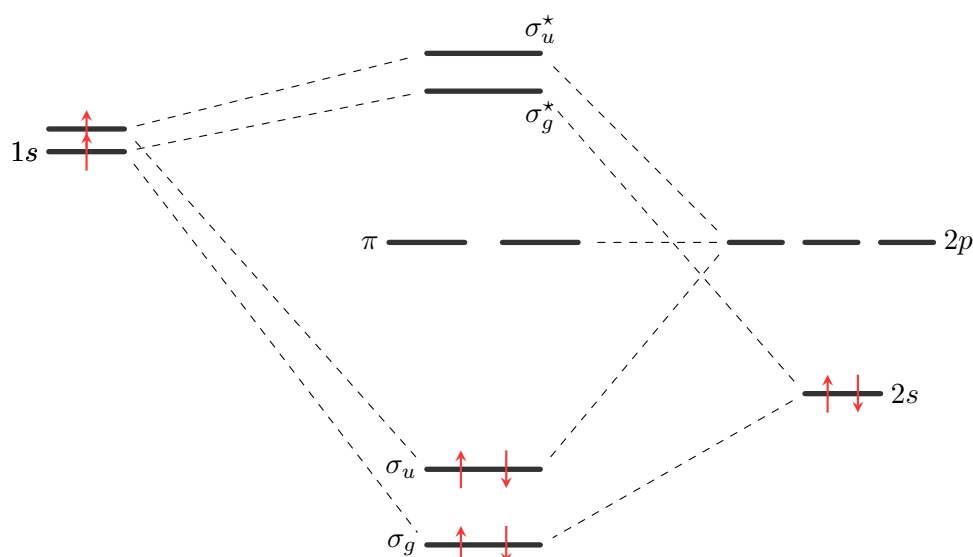
Figure 23.1: The Beryllium Hydride molecular orbital diagram

$$\text{H} \longrightarrow \text{Be} \longrightarrow \text{H}$$

Figure 23.2: The Beryllium Hydride molecule

§§ 23.4.3 Water (H_2O)

For water the situation is completely different. The two $1s$ orbitals of the hydrogen mix with the $2p$ orbitals of Oxygen, forming a sp^2 hybridization. In this case, we know already that this molecule has a bent symmetry, and the Hamiltonian will be invariant to all transformations of the group C_{2v}, due to the planar hybridization with the p_x, p_y orbitals of Oxygen. The two hybrid orbitals will then be the following 2

$$
\begin{aligned}
\left|sp^2\right\rangle_1 &= \left|1s\right\rangle_H + \lambda \left|2p_x\right\rangle_O \\
\left|sp^2\right\rangle_2 &= \left|1s\right\rangle_H + \lambda \left|2p_y\right\rangle_O
\end{aligned}
\tag{23.35}
$$

Due to the symmetry of the molecule it's possible to determine that the angle of separation of the two Hydrogen atoms is $\alpha = 105°$. Note that since this molecule is bent, usual molecular states do not apply to molecular orbitals, therefore, another notation based on group theory is used, and we get that the ground state of the water molecule is X^1A_1 and the electronic configuration is

$$2a_1^2 1b_2^2 3a_1^2 1b_1^2$$
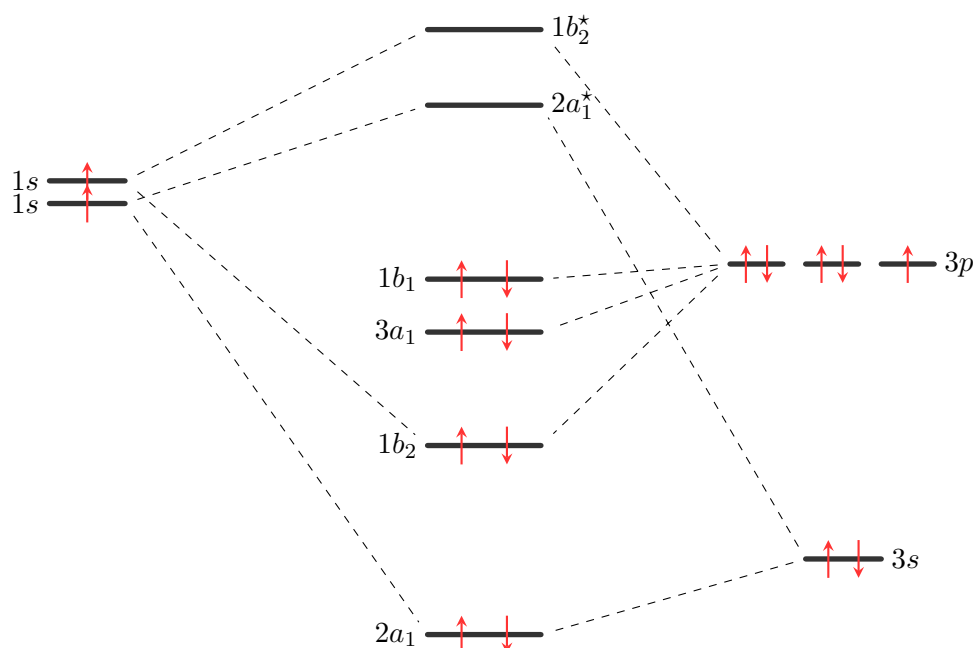
Figure 23.3: The molecular orbital diagram for water (H_2O)

Where the first two orbitals, a_1, b_2 are bonding orbitals.
The new notation for electronic structure of the molecule is given directly by group theory. This notation takes into account the bending of the molecule, which changes the energies of the usual σ, π, $\cdots$ molecular orbitals. This can be seen by taking a C_{2v} transformation and applying it to these molecular orbitals. Through this change of coordinates we see that the p orbitals transform as follows

$$
\begin{aligned}
\left|p'_x\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|p_x\right\rangle + \left|p_y\right\rangle\right) \\
\left|p'_y\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|p_x\right\rangle - \left|p_y\right\rangle\right)
\end{aligned}
\tag{23.36}
$$

Looking closely at this definition, we see that the state $|2s\rangle$ and $|2p_x\rangle$ transform into themselves, thus they have a a_1 symmetry, whereas we see that $\left|2p'_y\right\rangle$ changes sign upon this C_2 rotation, and therefore has a b_2 symmetry, and last, the state $|2p_z\rangle$ changes sign upon reflection σ_v, and therefore has b_1 symmetry.
Putting this all together we get the previous molecular configuration, which through a simple direct calculus (evaluating sign changes and symmetries) gives back the X^1A_1 spectroscopic term for water.
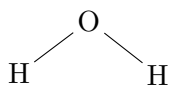
Figure 23.4: The Water molecule

§§ 23.4.4 Carbon Dioxide (CO_2) and Lone Pairs

For this molecule, we have 16 valence electrons, and we can form a bond using the $2s, 2p_x, 2p_y, 2p_z$ molecular orbitals of both atomic species.

Since the molecule is linear with an inversion point it is a $D_{\infty h}$ molecule, and he only possible σ_g atoms are given by the three $2s$ orbitals of Carbon and the two $2p_z$ orbitals of the two Oxygen atoms. We choose the internuclear axis as our z axis, and we build our orbitals as usual through the projection of $\hat{\Lambda}$ and $\hat{\underline{\lambda}}$. The hybrid atomic orbitals will then be a superposition of $2s$, $2p_z$ orbitals ($\Lambda = 0$, Σ symmetry) and of $2p_x$, $2p_y$ orbitals ($\Lambda = 1$, Π symmetry) which give a π bond.

Therefore, the electronic configuration of CO_2 will be, using the usual Aufbau rules for molecular orbitals

$$(1\sigma_g)^2(1\sigma_u^\star)^2(2\sigma_g)^2(2\sigma_u^\star)^2(1\pi_u)^4(1\pi_g^\star)^4 \tag{23.37}$$

The antibonding molecular orbitals do not contribute to the bonding, and therefore lead to a destabilization of the molecule. These are usually called (in chemistry) as *lone pairs*, and are indicated by couples of electrons which do not participate in the bonding.

The *bond order* of the molecule is given by this simple calculus

$$B_o = \frac{1}{2}\left(N_B - N_A\right) \tag{23.38}$$

Where N_B is the number of bonding electrons and N_A is the number of antibonding electrons.

From this we can immediately see that if the electrons involved in the lone pair had participated into the bonding, placing themselves into an antibonding orbital, would have brought the bond order $B_o = 0$, which means that there are *no bonds* in the molecule. This is clearly not what happens in nature in a standard atmosphere, because we are quite certain that this molecule exists!

Talking about the bond order again, it is also used in order to determine the number of bonds of a molecule, i.e. if $B_o = 3$ it means that this molecule has a triple bond, and so on.

$$O =\!\!=\!\!= C =\!\!=\!\!= O$$
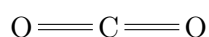
Figure 23.5: The Carbon Dioxide molecule

§ 23.5 Rayleigh-Ritz Variational Method

It's useful to check a different approach on the LCAO method, by introducing a variational form of this approximation.

We start by defining our unnormalized Hamiltonian eigenstate as follows, indicating with the subscript A the first atom, and with the subscript B the second atom.

$$|\Phi\rangle = c_1 |\psi\rangle_A + c_2 |\psi\rangle_B \tag{23.39}$$

Normalizing, and considering, without loss of generality that $\langle \psi | \psi \rangle_i = 1 \; i = A, B$, we get the new normalized eigenstate

$$|\Psi\rangle = \frac{1}{\sqrt{c_1^2 + c_2^2 + 2c_1 c_2 \langle \psi | \psi \rangle_{AB}}} \left(c_1 |\psi\rangle_A + c_2 |\psi\rangle_B \right) \tag{23.40}$$

Inserting this eigenstate in the Rayleigh-Ritz variational equation, we have

$$E = \langle \Psi | \hat{\mathcal{H}} | \Psi \rangle$$

$$E = \frac{1}{c_1^2 + c_2^2 + 2c_1 c_2 \langle \psi | \psi \rangle_{AB}} \left(\langle \psi |_A c_1 + \langle \psi |_B c_2 \right) \hat{\mathcal{H}} \left(c_1 |\psi\rangle_A + c_2 |\psi\rangle_B \right)$$

We continue calculating by bringing to the left the normalization factor

$$E \left(c_1^2 + c_2^2 + 2c_1 c_2 \langle \psi | \psi \rangle_{AB} \right) = c_1^2 \langle \psi | \hat{\mathcal{H}} | \psi \rangle_{AA} + c_2^2 \langle \psi | \hat{\mathcal{H}} | \psi \rangle_{BB} + 2c_1 c_2 \langle \psi | \hat{\mathcal{H}} | \psi \rangle_{AB}$$

$$c_1^2 E + c_2 E + 2c_1 c_2 E \langle \psi | \psi \rangle_{AB} - c_1^2 \langle \psi | \hat{\mathcal{H}} | \psi \rangle_{AA} - c_2^2 \langle \psi | \hat{\mathcal{H}} | \psi \rangle_{BB} - 2c_1 c_2 \langle \psi | \hat{\mathcal{H}} | \psi \rangle_{AB} = 0$$

Now, we impose the condition $\partial_{c_i} \langle \Psi | \hat{\mathcal{H}} | \Psi \rangle = 0$ in order to find the extremal value for the coefficients, and semplifying the constants and regrouping, we get the following system of equation

$$c_1 \left(\langle \psi | \hat{\mathcal{H}} | \psi \rangle_{AA} - E \right) + c_2 \left(\langle \psi | \hat{\mathcal{H}} | \psi \rangle_{AB} - E \langle \psi | \psi \rangle_{AB} \right) = 0$$
$$c_1 \left(\langle \psi | \hat{\mathcal{H}} | \psi \rangle_{AB} - E \langle \psi | \psi \rangle_{AB} \right) + c_2 \left(\langle \psi | \hat{\mathcal{H}} | \psi \rangle_{BB} - E \right) = 0 \tag{23.41}$$

Writing the matrix elements of the Hamiltonian as $\mathcal{H}_{ij}$, with $i, j = A, B$ and the overlap integral as S_{AB}, we have that the previous system, in the basis of the two coefficients is represented by the following matrix

$$A_{ij} = \begin{pmatrix} \mathcal{H}_{AA} - E & \mathcal{H}_{AB} - ES_{AB} \\ \mathcal{H}_{AB} - ES_{AB} & \mathcal{H}_{BB} - E \end{pmatrix}$$

It's easy to note how this matrix is given by the following matrix equation

$$\langle \psi | \hat{\mathcal{H}} | \psi \rangle_{ij} - E \langle \psi | \psi \rangle_{ij} = \mathcal{H}_{ij} - ES_{ij}$$
$$\mathcal{H}_{ij} = \begin{pmatrix} \langle \psi | \hat{\mathcal{H}} | \psi \rangle_{AA} & \langle \psi | \hat{\mathcal{H}} | \psi \rangle_{AB} \\ \langle \psi | \hat{\mathcal{H}} | \psi \rangle_{AB} & \langle \psi | \hat{\mathcal{H}} | \psi \rangle_{BB} \end{pmatrix}$$
$$S_{ij} = \begin{pmatrix} 1 & \langle \psi | \psi \rangle_{AB} \\ \langle \psi | \psi \rangle_{AB} & 1 \end{pmatrix}$$

This matrix gives the searched solution for $\det_2(A_{ij}) = 0$, hence we get the following second order linear equation on E

$$\left(\mathcal{H}_{AA} - E \right) \left(\mathcal{H}_{BB} - E \right) - \left(\mathcal{H}_{AB} - ES_{AB} \right)^2 = 0$$

Noting that $|\psi\rangle_A = |\psi\rangle_B$ we have that $\mathcal{H}_{AA} = \mathcal{H}_{BB}$, and the solution to the equation are

$$E_+(R) = \frac{\mathcal{H}_{AA} + \mathcal{H}_{AB}}{1 + S_{AB}}$$

$$E_-(R) = \frac{\mathcal{H}_{AA} - \mathcal{H}_{AB}}{1 - S_{AB}} \tag{23.42}$$

$$|c_1|^2 = |c_2|^2 \longrightarrow c_1 = \pm c_2$$

Where, expressing the matrix elements of the Hamiltonian in Schrödinger notation, we get

$$\mathcal{H}_{AA} = \int \overline{\psi}(r_A)\hat{\mathcal{H}}\psi(r_A)\,\mathrm{d}^3 r_A = J$$

$$\mathcal{H}_{AB} = \iint \overline{\psi}(r_A)\hat{\mathcal{H}}\psi(r_B)\,\mathrm{d}^3 r_A \mathrm{d}^3 r_B = K \tag{23.43}$$

$$S_{AB} = \iint \overline{\psi}(r_A)\psi(r_B)\,\mathrm{d}^3 r_A \mathrm{d}^3 r_B = I$$

Which are the already known Coulomb integral (J), Exchange integral (K) and the Overlap integral (I). The energies therefore represent the bonding-antibonding state couple

§§ 23.5.1 Hückel Theory

Further approximations can be made using this variational method. Commonly one chooses a minimal LCAO basis (overlap of $|1s\rangle$ states) in order to complete the calculations, but this is not always the best choice. An example are organic molecules and molecules with weak overlaps of orbitals (like in pi bonds). Therefore, as an example, suppose taking a planar hydrocarbon. The π bond of carbon can be thought, as for the Rayleigh-Ritz theory described before, as a linear sum of $|p_z\rangle$ states.

$$|\pi\rangle = c_1\left|p_z^1\right\rangle + c_2\left|p_z^2\right\rangle \tag{23.44}$$

The energies will therefore be the solution of the following secular equation

$$\begin{vmatrix} \mathcal{H}_{11} - ES_{11} & \mathcal{H}_{12} - ES_{12} \\ \mathcal{H}_{21} - ES_{21} & \mathcal{H}_{22} - ES_{22} \end{vmatrix} = 0 \tag{23.45}$$

In order to solve this equation, we impose the two main approximation of Hückel theory

1. We assume an orthonormal basis, therefore $S_{ij} = \left\langle p_z^i \middle| p_z^j \right\rangle = \delta_{ij}$

2. We assume that the p_z orbitals interact only with their next closest orbital, therefore $\mathcal{H}_{ij} = \alpha$ if $i = j$, or $\mathcal{H}_{ij} = \beta$ if the orbitals are adjacent. Otherwise it's considered to be zero

The secular equation therefore becomes

$$\begin{vmatrix} \alpha - E & \beta \\ \beta & \alpha - E \end{vmatrix} = (\alpha - E)^2 - \beta^2 = 0 \tag{23.46}$$

Solving for E we get $E_\pm = \alpha \pm \beta$, which gives E_+ as our bonding π_u orbital and E_- as the antibonding $\pi_g^\star$ orbital

§ 23.6 π Electron Systems and Unlocalized Orbitals

From now on we will consider a different class of molecules, where the electronic wavefunction isn't precisely localized in a certain position, permitting the existence of *unlocalized orbitals*. These particular molecules include chains of Carbon atoms, where double and single bonds alternate and aren't localized. The electrical polarizability of these molecules is much larger than the usual molecules with localized bonds. This fenomenon is given by overlapping p orbitals, i.e. π bonds.

§§ 23.6.1 Butadiene (C_4H_6)

Transbutadiene is an isomer of Butadiene (molecule with the same composition of Butadiene) which is planar. It's formed by a single σ bond between two CH radicals, which are themselves doubly bonded with two CH_2 radicals. In this case there exist 4 π bonds which are linear combinations of the p orbitals of the Carbon atoms, with the following wavefunction

$$|\pi\rangle = \sum_{n=1}^{4} c_n \, |2p\rangle_n \tag{23.47}$$

These 4 c_n coefficients can be determined using the Rayleigh-Ritz variational equation, which corresponds to the eigenvalue problem

$$|\mathcal{H}_{ij} - E S_{ij}| = 0$$

In order to solve easily these equations we impose some approximations

1. $\mathcal{H}_{ii} = \alpha$, with α a parameter

2. $\mathcal{H}_{ij} = \beta < 0$, $i \neq j$ only for adjacent atoms

3. $S_{ij} = \delta_{ij}$

These assumptions are the basis of the *Hückel method*. After some calculations, for Transbutadiene, we obtain the following wavefunction for the π bonds

$$\begin{aligned}
|\pi_1\rangle &= 0.37\,|1\rangle + 0.60\,|2\rangle + 0.60\,|3\rangle + 0.37\,|4\rangle \\
|\pi_2\rangle &= 0.60\,|1\rangle + 0.37\,|2\rangle - 0.37\,|3\rangle - 0.60\,|4\rangle \\
|\pi_3\rangle &= 0.60\,|1\rangle - 0.37\,|2\rangle - 0.37\,|3\rangle + 0.60\,|4\rangle \\
|\pi_4\rangle &= 0.37\,|1\rangle - 0.60\,|2\rangle + 0.60\,|3\rangle - 0.37\,|4\rangle
\end{aligned} \tag{23.48}$$

The orbital π_1 is completely unlocalized over the whole Transbutadiene molecule.
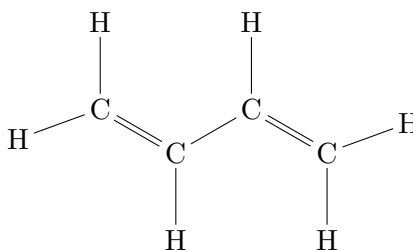
Figure 23.6: The Transbutadiene Molecule

§§ 23.6.2 Methane and Ethylene(CH_4, C_2H_4)

As a continuation to our examples, we take the Methane molecule, CH_4. Carbon bonds readily in an excited state $1s^2 2s 2p^3$ which is very close in energy, forming, for Methane, an sp^3 hybridization. We can construct the usual 4 combinations for the wavefunction of the molecular orbitals

$$|1\rangle = |2s\rangle + |2p_x\rangle + |2p_y\rangle + |2p_z\rangle$$
$$|2\rangle = |2s\rangle + |2p_x\rangle - |2p_y\rangle - |2p_z\rangle$$
$$|3\rangle = |2s\rangle - |2p_x\rangle + |2p_y\rangle - |2p_z\rangle$$
$$|4\rangle = |2s\rangle - |2p_x\rangle - |2p_y\rangle + |2p_z\rangle$$

(23.49)

Since all $|2p_i\rangle$ eigenstates are proportional to the standard euclidean basis, we can immediately determine that the direction of the maximum are respectively

$$(1, 1, 1)$$
$$(1, -1, -1)$$
$$(-1, 1, -1)$$
$$(-1, -1, 1)$$

Calculating the cosine between these directions we obtain a bond angle of $109.6°$, giving CH_4 a tetrahedral structure.

In case of a planar molecule like Ethylene, we can form sp^2 hybrids as follows

$$|1\rangle = |2s\rangle + \sqrt{2}\,|2p_x\rangle$$
$$|2\rangle = |2s\rangle + \sqrt{\frac{3}{2}}\,|p_y\rangle - \sqrt{\frac{1}{2}}\,|p_x\rangle$$
$$|3\rangle = |2s\rangle - \sqrt{\frac{3}{2}}\,|p_y\rangle - \sqrt{\frac{1}{2}}\,|2p_x\rangle$$

(23.50)
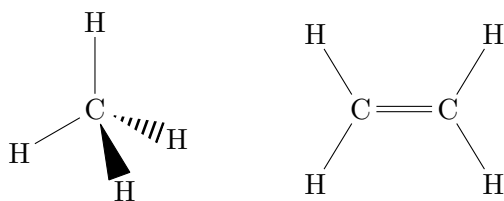
Figure 23.7: The Methane and Ethylene molecules

§ 23.7 Cyclic Molecules

§§ 23.7.1 Ideal Homonuclear Trimer

We can begin our treatment of polyatomic molecules by studying a toy problem, as the homonuclear trimer, i.e. a triangular molecule composed from 3 atoms of the same element. We begin evaluating the system by immediately finding that there is a C_3 axis of symmetry, passing through the center of the triangle. Having considered this, we already know that the electronic Hamiltonian and the $\hat{C}_3$ operator, which applicates a rotation of $\frac{2}{\pi}3$ radians, commutes $\left[\hat{\mathcal{H}}_e, \hat{C}_3\right] = 0$, thus there exists a common diagonalizing base.

We search this base by taking the minimal LCAO base orbitals. In this base, the electronic Hamiltonian $\hat{\mathcal{H}}_e$ has the energy of the single orbitals in the diagonal, and these base kets diagonalize the $\hat{C}_3$ operator.

We have

$$\hat{C}_3 \left|\alpha\right\rangle = c_m \left|\alpha\right\rangle = c_1 \left|1s\right\rangle_1 + c_2 \left|1s\right\rangle_2 + c_3 \left|1s\right\rangle_3$$

Since these eigenvalues must have norm $\|c_m\| = 1$ and must also express a rotation, have the shape of a complex exponential

$$c_i = e^{\frac{2im\pi}{3}}$$

Therefore, we can write

$$\hat{C}_3 \left|\alpha_m\right\rangle = e^{\frac{2im\pi}{3}} \left|\alpha_m\right\rangle = \left|1s\right\rangle_1 + e^{\frac{2im\pi}{3}} \left|1s\right\rangle_2 + e^{-\frac{2im\pi}{3}} \left|1s\right\rangle_3, \quad m = 0, \pm 1$$

Where we ignored general phases.

This base is made from a linear combination of bases of the electronic Hamiltonian $\hat{\mathcal{H}}_e$, hence we can easily write that $\mathrm{diag}\, \mathcal{H}_{ij}^{(e)} = (\epsilon, \epsilon, \epsilon)$, with ϵ the single atomic orbital energy, and the diagonal elements are all the same $\mathcal{H}_{ij} = -t \; \forall i \neq j$ at the same time we can also say easily that the Schrödinger equation holds

$$\hat{\mathcal{H}}_e \left|\alpha_m\right\rangle = E_m \left|\alpha_m\right\rangle$$

And therefore, applying the Hamiltonian matrix to the matrix representation of this vector (with normalization), we obtain

$$
\begin{aligned}
\mathcal{H}_{ij}\alpha_j^{(m)} &= \frac{1}{\sqrt{3}} \begin{pmatrix} \epsilon & -t & -t \\ -t & \epsilon & -t \\ -t & -t & \epsilon \end{pmatrix} \begin{pmatrix} 1 \\ e^{\frac{2im\pi}{3}} \\ e^{-\frac{2im\pi}{3}} \end{pmatrix} = \\
&= \frac{1}{\sqrt{3}} \begin{pmatrix} \epsilon - t\left(e^{\frac{2im\pi}{3}} + e^{-\frac{2im\pi}{3}}\right) \\ \epsilon - t\left(e^{\frac{2im\pi}{3}} + e^{-\frac{4im\pi}{3}}\right) \\ \epsilon - t\left(e^{\frac{4im\pi}{3}} + e^{-\frac{2im\pi}{3}}\right) \end{pmatrix} = \\
&= \frac{1}{\sqrt{3}} \left[\epsilon - t\left(e^{\frac{2im\pi}{3}} + e^{\frac{2im\pi}{3}}\right)\right] \begin{pmatrix} 1 \\ e^{\frac{2im\pi}{3}} \\ e^{-\frac{2im\pi}{3}} \end{pmatrix} = \\
&= \frac{1}{\sqrt{3}} \left[\epsilon - 2t\cos\left(\frac{2m\pi}{3}\right)\right]\alpha_j^{(m)} = E_m\alpha_j^{(m)}
\end{aligned}
\tag{23.51}
$$

Which are the energies of the three $|\alpha_m\rangle$ molecular orbitals, where two are degenerate and shifted upwards from the single orbital by $+t$ and one is shifted downwards of $-2t$.
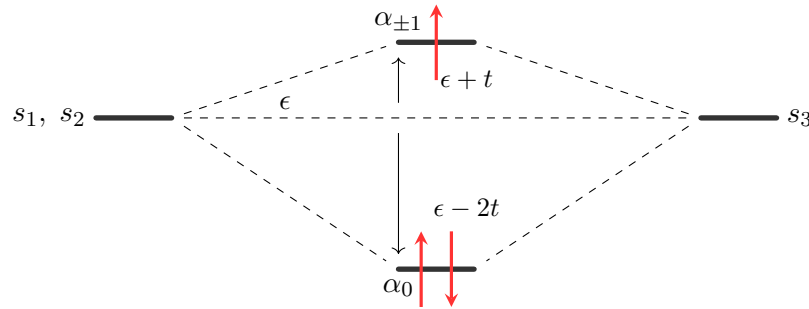
Figure 23.8: LCAO diagram for the homonuclear trimer

These energy levels though, are «not» optimized. In fact it's energetically convenient to distort the molecule from a triangular symmetry to a isosceles symmetry, breaking the degeneration for the orbitals $|\alpha_{\pm 1}\rangle$. This effect is known as *Jahn-Teller effect*.

§§ 23.7.2 Ideal Polyatomic Ring and the Tight Binding Approximation

After treating the ideal homonuclear trimer it's possible to generalize the calculations to a generic ring formed by N atoms. Suppose having a ring of such N atoms with one valence electron, this ring will be a regular polyhedron with the nuclei at its vertexes. We must obviously have $\left[\hat{\mathcal{H}}, \hat{C}_N\right] = 0$ which means that the eigenstates of the Hamiltonians must also be eigenstates of the rotation operator $\hat{C}_N$ with eigenvalue $e^{\pm 2mi\pi/N}$ with $m = 0, \pm 1, \cdots, \pm(N/2 - 1), N/2$ if $N \mod 2n = 0$ with $n \in \mathbb{N}$.

Considering a minimal LCAO basis $|1s_n\rangle = |n\rangle$, all the eigenvectors of $\hat{\mathcal{H}}$ will be therefore linear combinations of 1s orbitals of the following kind

$$|\alpha_m\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{\frac{2imn\pi}{N}} |n\rangle \tag{23.52}$$

Note how the state $|\alpha_0\rangle, |\alpha_{N/2}\rangle$ correspond respectively to the normalized sum of the orbitals (generalization of a bonding orbital) and to the alternated sum of orbitals (generalization of an antibonding orbital).

Considering the $|n\rangle$ basis as a complete orthonormal system we have that the energies of the system will be the matrix elements of the Hamiltonian, as follows

$$\epsilon_m = \langle \alpha_m | \hat{\mathcal{H}} | \alpha_m \rangle = \frac{1}{N} \sum_{n'=1}^{N} \sum_{n=1}^{N} e^{\frac{2im(n'-n)}{N}} \langle n' | \hat{\mathcal{H}} | n \rangle \tag{23.53}$$

Calling the diagonal elements of the Hamiltonian as ϵ and the off diagonal as $\langle n \pm k | \hat{\mathcal{H}} | n \rangle = -t^{(k)}$ we can safely suppose that $t^{(k)} = 0$ for $k > 1$ since the atomic eigenfunction must decay exponentially with the distance. This approximation is known as the *tight binding approximation*. Calculating the energies as we did for the trimer we get, analogously, the following expression

$$\epsilon_m = \epsilon - \frac{1}{N} \sum_{k=1}^{N-1} \sum_{n=1}^{N} t^{(k)} \left(e^{\frac{2^k im\pi}{N}} + e^{\frac{2^k im\pi}{N}} \right) = \epsilon - 2 \sum_{k=1}^{N-1} t^{(k)} \cos\left(\frac{2^k m\pi}{N} \right)$$

Including the tight binding approximation, the levels become

$$\epsilon_m = \epsilon - 2t \cos\left(\frac{2m\pi}{N} \right) \tag{23.54}$$

Note how this is a reformulation of the previously discussed Hückel theory.

All the energy levels will lay in the interval $[\epsilon - 2t, \epsilon + 2t)$ ($\epsilon + 2t$ included if the number of atoms is even).

We therefore will have, for the minimum and maximum value of the energies

$$\epsilon_m \to \begin{cases} \epsilon_0 = \epsilon - 2t \\ \epsilon_{\frac{N}{2}} = \epsilon + 2t & N \mod 2n = 0 \\ \epsilon_{\pm \frac{N-1}{2}} = \epsilon - 2t \cos\left(\frac{2\pi(N-1)}{N} \right) & N \mod (2n+1) = 0 \end{cases} \quad n \in \mathbb{N} \tag{23.55}$$

The minimum will be ϵ_0 and the remaining 2 will be the maximum levels of energy, for even and uneven N.

Generally, from this property of energy we have that the energy levels will move from the median value ϵ by $\pm 2t$, defining the band of permitted energies. Note that since we have to fill these orbitals with electrons (2 per time due to the spin degeneracy and the Pauli exclusion principle), therefore if the single atoms have only one valence electron, only half of these N orbitals will be occupied.

§§ 23.7.3 Benzene (C_6H_6)

Benzene is a planar molecule, formed by a ring of carbon atoms, for which each one of those is bonded with a hydrogen atom. This symmetry structure indicates an sp^2 hybridization of the carbon atom's orbitals.

This molecule has a C_{6v} symmetry which lets us hypothesize immediately an sp^2 hybridization of the orbitals of the carbon atoms, or more precisely of the orbitals of the Methyl group CH. Again using the symmetry of the molecule, we can write our eigenfunctions as follows, noting that there exist 3 combinations possible for a sp^2 hybridization

$$|\psi_j\rangle = \frac{1}{\sqrt{6}} \sum_{n=1}^{6} e^{\frac{i\pi n}{3}} |sp^2\rangle_j \quad j = 1, 2, 3 \tag{23.56}$$

We need simply to add to these 3 hybrid orbitals another 2 orbitals participating in the bonding, which are given by the CH group.

Solving everything, we find that the sp^2 hybrids are well localized in the molecule, and instead the remaining $2p_z$ orbitals, which participate in the bonding, are loosely bound in a unlocalized π bond.
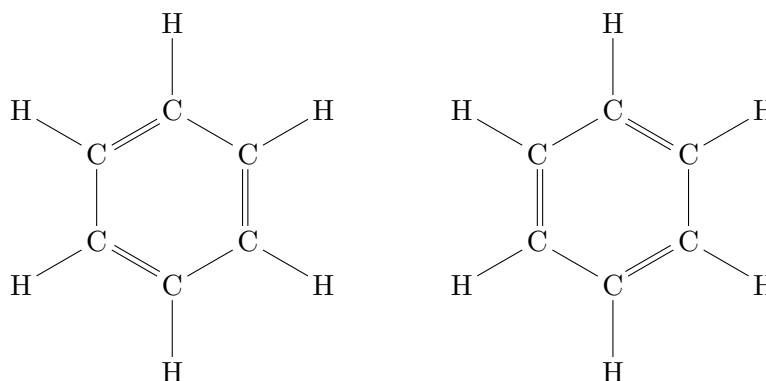
Figure 23.9: The Benzene molecule and its resonance of the double π bonds

Another way to treat this organic molecule, is to use its rotational symmetry for $C - C$ bonds and tight binding for tying everything up. Since there are 6 C atoms we have that there will be 6 α_m orbitals, combination of the basis orbitals. Mathematically

$$|\alpha_m\rangle = \frac{1}{\sqrt{6}} \sum_{n=1}^{6} e^{\frac{2im\pi}{6}} |n\rangle \tag{23.57}$$

Adding up the hydrogens and supposing that only the $2s2p$ orbitals of carbon will partake in the bonding, considering also the three previously mentioned sp^2 hybrid orbitals, we will have a grand total of 30 atomic orbitals. Since carbon is tetravalent and hydrogen is univalent, we will have $4 * 6 + 6 = 30$ electrons. Due to the doubly degenerate nature of molecular orbitals, only the first 15 will be completely occupied, leaving 15 free orbitals. It's not easy to solve easily a 30×30 Hamiltonian for such problem, but it's possible to

greatly reduce the dimensionality of the problem. We have that the interacting orbitals in this system must be 4 for the carbons and only one for hydrogen.

We identify with α, β the two orbitals of carbon coplanar with the ring, the first in the clockwise direction and vice versa, with p the perpendicular p_z orbital, with γ the orbital between carbon and hydrogen, and with s the remaining hydrogen orbital.

Using therefore these $\alpha, \beta, \gamma, p, s$ orbitals end up, using the previous formula, with the following general expression for their linear combination

$$|C\alpha_m\rangle = \frac{1}{\sqrt{6}} \sum_{n=1}^{6} e^{\frac{2im\pi}{6}} |\alpha_n\rangle$$

$$|C\beta_m\rangle = \frac{1}{\sqrt{6}} \sum_{n=1}^{6} e^{\frac{2im\pi}{6}} |\beta_n\rangle$$

$$|C\gamma_m\rangle = \frac{1}{\sqrt{6}} \sum_{n=1}^{6} e^{\frac{2im\pi}{6}} |\gamma_n\rangle \qquad (23.58)$$

$$|Cp_m\rangle = \frac{1}{\sqrt{6}} \sum_{n=1}^{6} e^{\frac{2im\pi}{6}} |p_n\rangle$$

$$|Hs_m\rangle = \frac{1}{\sqrt{6}} \sum_{n=1}^{6} e^{\frac{2im\pi}{6}} |s_n\rangle$$

Approximating with the tight binding model we have that 2 of each of these orbitals are orthogonal for different m and have 0 Hamiltonian matrix element, imposing these two constraints

$$\langle I|J\rangle = \delta_{mm'} S_{IJ}(m)$$
$$\langle I|\hat{\mathcal{H}}|J\rangle = \delta_{mm'} \mathcal{H}_{IJ}(m) \qquad (23.59)$$

This reduces the complete 30×30 problem in 6 5×5 problems.

To this first reduction, we apply the tight binding approximation therefore rendering the S matrix nonzero only on the diagonal and on the first elements above and below the diagonal. We apply the same reasoning for the Hamiltonian, giving finally

$$S_{IJ} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & S_{\alpha\beta} & 0 & 0 \\ 0 & S_{\beta\alpha} & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & S_{\gamma s} \\ 0 & 0 & 0 & S_{s\gamma} & 1 \end{pmatrix}$$

$$\mathcal{H}_{IJ} = \begin{pmatrix} \epsilon_m & 0 & 0 & 0 & 0 \\ 0 & \epsilon_{\alpha\beta\gamma} & -t_{CC\sigma} & 0 & 0 \\ 0 & -t_{CC\sigma} & \epsilon_{\alpha\beta\gamma} & 0 & 0 \\ 0 & 0 & 0 & \epsilon_{\alpha\beta\gamma} & -t_{CH\sigma} \\ 0 & 0 & 0 & -t_{CH\sigma} & \epsilon_s \end{pmatrix} \qquad (23.60)$$

Where $I, J = p, \alpha, \beta, \gamma, s$ and

$$\epsilon_m = \epsilon_p - 2t_{pp\pi} \cos\left(\frac{2m\pi}{6}\right)$$

$$\epsilon_{\alpha\beta\gamma} = \langle C\alpha| \hat{\mathcal{H}} |C\alpha\rangle = \langle C\beta| \hat{\mathcal{H}} |C\beta\rangle = \langle C\gamma| \hat{\mathcal{H}} |C\gamma\rangle \qquad (23.61)$$

$$\epsilon_s = \langle Hs| \hat{\mathcal{H}} |Hs\rangle - t_{CC\sigma} = \langle C\alpha| \hat{\mathcal{H}} |C\beta\rangle$$

$$-t_{CH\sigma} = \langle C\gamma| \hat{\mathcal{H}} |Hs\rangle$$

Note how the energies for the sigma bonds are localized to either the $C - C$ or $C - H$ bonds and are localized to one single orbital, while the energy for the p_z orbitals of each carbon are dependent on m and are delocalized over the whole ring.
Ordering the orbital energies and filling up the orbitals using the usual rules we have that the two hybrid sp^2 orbitals σ_{CH}, σ_{CC} are completely filled, while the delocalized π_{CC} orbital is filled only in the lower levels with $m = 0, \pm1$<++>

24 Introduction to Solid State Physics

§ 24.1 Infinite Linear Chain

The first idea of solid we can imagine is the direct generalization of an ideal cyclic molecule with N atoms, where N is a pretty big number but not infinite, in general $N \sim 10^7 - 10^8$. This regular polygon with so many vertices can be approximated without problems to a chain long $2\pi R \simeq Na$ where a is the distance between two atoms. Taking a single atom of the chain as our coordinate origin we see easily that a finite rotation of $2\pi N^{-1}$, for $N \to \infty$ is almost indistinguishable from a discrete translation of a along an infinite chain.

From this symmetry we can immediately say, that if we define a translation operator $\hat{T}_a$ that moves this chain, if we write the Hamiltonian of the system as $\hat{\mathcal{H}}$, we have that $\left[\hat{\mathcal{H}}, \hat{T}_a \right] = 0$. Since we're generalizing the ideal polyatomic cyclic molecule for N univalent atoms, we take our usual minimal LCAO basis. We will have $|\alpha_m\rangle$ orbitals, where $m = 0, \cdots, N$, and our translation operator on our basis kets will act as follows

$$\hat{T}_a |\alpha_m\rangle = e^{\frac{2i\pi m}{N}} |\alpha_m\rangle \tag{24.1}$$

The induced phase, which depends on m, can now be redefined with a new quantum number, due to the properties of this chain.

We must have, for the electronic wavefunction

$$\psi_{e^-}(x) = \psi_{e^-}(x + L), \quad L = Na \tag{24.2}$$

Supposing the electron completely free we have that the wavefunction must be of the form $\psi \propto e^{ikx}$. Imposing the previous condition in order to constrain the electron wavefunction to the chain we must have

$$Ae^{ikx} = Ae^{ik(x+Na)} \Longrightarrow k = \frac{2\pi m}{Na}, \ m = 0, \cdots, N \tag{24.3}$$

This is the quantization relation for the chain, which gives us a new quantum number k_m, that due to its definition is called the «wavenumber» of the chain, where

$$k_m = \begin{cases} \dfrac{2\pi m}{Na} = 0, \cdots, \pm \left(\dfrac{N-1}{N} \right) \dfrac{\pi}{a} & N = 2k+1, \ k \in \mathbb{N} \\[4mm] \dfrac{2\pi m}{Na} = 0, \cdots, \pm \dfrac{\pi}{a} & N = 2k \end{cases} \tag{24.4}$$

It's easy to see from its definition that $k_m \in [-\pi/a, \pi/a]$, which is called the «1st Brillouin zone» . It's also obvious that for $N \to \infty$, $k_m \to k$, where k is a continuous variable defined in the same Brillouin interval.

With this definition, we will have that applying $\hat{T}_a$ on our basis states, we have

$$\hat{T}_a \ket{\alpha_k} = e^{ik_m a} \ket{\alpha_k} \tag{24.5}$$

Where we substituted $m \to k_m$, our new quantum (wave)number. As we said before the Hamiltonian commutes with the translation operator, therefore, since $\ket{\alpha_k}$ are eigenstates of $\hat{T}_a$, they must also be eigenstates of the Hamiltonian, such that

$$\begin{aligned}
\braket{k|k'} &= \delta_{kk'} \\
\bra{k} \hat{\mathcal{H}} \ket{k'} &= \epsilon_k \delta_{kk'}
\end{aligned} \tag{24.6}$$

Where $\ket{\alpha_k} = \ket{k}$.

Since basically everything is equal to the case of the N-atomic cyclic molecule, supposing that $\ket{\alpha_m} = \ket{1s}_m$ (i.e. $\ket{\alpha}$ is a minimal LCAO base for every single atom) we can say, imposing the tight binding approximation, that

$$\ket{k} = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{ikna} \ket{1s}_n \tag{24.7}$$

$$\epsilon_k = \epsilon_0 - 2t \cos(ka)$$

All of this is generalizable to 2 and 3 dimensions, with the definition of a Bravais lattice

Definition 24.1.1 (Bravais Lattice). A Bravais lattice is defined as a set of discrete points which can be described with a lattice vector $\underline{R} = n_1 \underline{a}_1 + n_2 \underline{a}_2 + n_3 \underline{a}_3$, where $n_1, n_2, n_3 \in \mathbb{Z}$ and $\underline{a}_i$ are linearly independent basis vectors. Therefore

$$\begin{aligned}
R &= na &&\text{in 1D} \\
\underline{R} &= n_1 \underline{a}_1 + n_2 \underline{a}_2 &&\text{in 2D} \\
\underline{R} &= n_1 \underline{a}_1 + n_2 \underline{a}_2 + n_3 \underline{a}_3 &&\text{in 3D}
\end{aligned} \tag{24.8}$$

And with the usage of the Bloch theorem, which states

Theorem 24.1 (Bloch Theorem). *Given a Hamiltonian with a lattice-periodic potential* $V(\underline{r}) = V(\underline{r}+\underline{R})$, *where* $\underline{R}$ *is a lattice vector, the eigenfunctions of the system will be of the following shape*

$$\psi_{\underline{k}}(\underline{r}) = e^{i\underline{k}\underline{r}} u_{\underline{k}}(\underline{r}) \tag{24.9}$$

Where $u_{\underline{k}}$ *is a lattice-periodic function. The wavefunction* $\psi_{\underline{k}}$ *is known as a Bloch wave.*

All this summed up, for a 3D lattice, we will have

$$\ket{\underline{k}} = \frac{1}{\sqrt{N_1 N_2 N_3}} \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} \sum_{n_3=1}^{N_3} e^{ia(k_x n_1 + k_y n_2 + k_z n_3)} \ket{n_1 n_2 n_3} \tag{24.10}$$

$$\epsilon_{\underline{k}} = \epsilon_0 - 2t \cos(k_x a) - 2t \cos(k_y a) - 2t \cos(k_z a)$$

§ 24.2 Fermi Level

As seen in Statistical Mechanics, the electrons, being fermions with $s = 1/2$ must obey the Fermi-Dirac statistic, where the density of states is some function $g(\epsilon)$. The number of states will be given by the integral of this probability density function $g(\epsilon)$, remembering that if $\mathrm{d}N$ is the number of states between $\epsilon, \epsilon + \mathrm{d}\epsilon$ and g is the degeneracy of the levels, we have

$$\int_{-\infty}^{\infty} g(\epsilon)\mathrm{d}\epsilon = \frac{N}{g}$$

Since for electrons we have a degeneracy of 2, since $m_s = \pm 1/2$ we can say that in general, the number of states possible is $2g(\epsilon)\mathrm{d}\epsilon$. It's obvious that the Fermi-Dirac distribution must be a Dirac delta for discrete spectra (like the QHO), where

$$g(\epsilon) = \sum_i \delta(\epsilon - \epsilon_i)$$

Where i is a quantum number.

In general, for solids we will search the density of states in a certain energy interval, given by $g(\epsilon)/V$ where V is either the length, surface or volume (in 1D, 2D or 3D) of the lattice. Due to the high number of atoms in the crystals we treat it's common to employ the following approximation, where we will move everything from the discrete sums to a continuous integral, where in 3 dimensions, we have

$$\frac{1}{V} \sum_i \rightarrow \frac{1}{(2\pi)^3} \int_B \mathrm{d}^3 k$$

Where we also approximated $\underline{k_m} \rightarrow \underline{k}$, considering it as a continuous quantum number. Note that B is Brillouin's first zone.

Using the definition of $\epsilon_{\underline{k}}$ given in (24.10) the integral is solvable in 1D, and therefore, we get

$$g(\epsilon) = \begin{cases} \dfrac{1}{\pi a} \dfrac{1}{\sqrt{4t^2 - (\epsilon - \epsilon_0)}} & |\epsilon - \epsilon_0| < 2t \\ 0 & |\epsilon - \epsilon_0| > 2t \end{cases} \tag{24.11}$$

The Fermi level of a crystal, is defined as the level with energy ϵ_F such that the highest possible state is occupied by fermions at $T = 0$. For electrons we have, supposing a univalent crystal with only one basis orbital (a one level system) for N atoms, we have that

$$N_{e^-}(\epsilon_F) = N = 2 \int_{-\infty}^{\epsilon_F} g(\epsilon)\mathrm{d}\epsilon \tag{24.12}$$

Note that the integral must be equal to the number of atoms, since that's the number of electrons in the system, since each atom is univalent.

By sheer logic, since each cell of the crystal can be occupied by two electrons, we have that $N/2$ cells are occupied and $N/2$ are unoccupied, therefore at $T = 0\ K$, noting that the

energy band is symmetric with respect to ϵ_0 ($|\cos(x)| \leq 1$), we can immediately say that having the first $N/2$ cells occupied, it implies that the Fermi level must lay at the center of the band, giving $\epsilon_F = \epsilon_0$.

If this is generalized to two or three dimensional lattices, we can define a Fermi surface, i.e. a surface Σ_F defined as follows

$$\Sigma_F := \left\{ \underline{k} \in \mathbb{R}^3 \,|\, \epsilon_{\underline{k}} = \epsilon_F \right\}$$

I.e. it's the set of all wavevectors $\underline{k}$ such that $\epsilon_{\underline{k}} = \epsilon_0$.

For a tight-binding Hamiltonian in 3D, where all the atoms are in an univalent lattice, we can immediately say that since $\epsilon_0 = \epsilon_F$, the Fermi surface will be the solution of the following equation

$$\cos(k_x a) + \cos(k_y a) + \cos(k_z a) = 0 \tag{24.13}$$

§ 24.3 Free Electron Approach to Metals

§§ 24.3.1 Drude Theory of Conduction

The idea of conduction theorized by Drude is a completely classical idea.

Suppose having a metal with Z valence electrons and N atoms with charge $+Z_a e$. These electrons move almost freely on the surface of the metal, where their only interaction is through collisions with the ions.

Supposing these collisions as "conduction collisions" we can write Newton's equation of motion for this system

$$m\underline{a} = e\underline{E} \Longrightarrow \underline{a} = -\frac{e\underline{E}}{m} \tag{24.14}$$

By integrating on t and including $\langle \underline{v_0} \rangle = 0$ by hypothesis, we have that on average, for each electron

$$\langle \underline{v_{e^-}} \rangle = \left\langle \frac{e\underline{E}}{m} t \right\rangle = \frac{e\underline{E}}{m} \tau \tag{24.15}$$

Where τ is the "relaxation time" of the metal, i.e. the average time between the collisions of the electrons.

By definition of current we can see that it will be proportional to the average velocity of the electrons times the number density of electrons in the metal, therefore, writing Ohm's law

$$\underline{J} = -ne \langle \underline{v_{e^-}} \rangle = \sigma \underline{E} \Longrightarrow \sigma = \frac{ne^2}{m} \tau \tag{24.16}$$

Where σ is the conductivity of the metal.

Since the system is completely classical, we have that the electrons will follow a Maxwell-Boltzmann distribution of velocities, and since $\langle E \rangle = 3k_B T/2$ we also have that the electron speed is tied to the temperature of the metal with the following formula

$$\langle \underline{v_{e^-}} \rangle (T) = \sqrt{\frac{3k_B T}{m}} \tag{24.17}$$

The mean free path of the electrons on the metal will obviously be given by the formula $\lambda = |\langle \underline{v}_{e^-} \rangle| \tau$.

§§ 24.3.2 Sommerfeld Theory of Metals and Conduction

The Sommerfeld theory is basically a quantum approach to Drude's classical theory.
The first idea behind this is to evaluate the free electrons of the solid inside a box of volume $V = L^3$. The Schrödinger equation for such system will be

$$\frac{\hbar^2}{2m} \nabla^2 \psi = \epsilon_{\underline{k}} \psi \tag{24.18}$$
$$\psi(\underline{r} + \underline{L}) = \psi(\underline{r})$$

This equation has a solution a plane wave with normalization $V^{-1/2}$, and a quantization of the wavenumber $\underline{k}_n = 2\pi L^{-1}(n_x, n_y, n_z)$. Note that since we didn't consider a lattice structure the electrons are completely free and $\underline{k}$ isn't confined to Brillouin's first zone.
From this we can say that since the particles considered must obey Fermi-Dirac's statistic, we have that a $T = 0$ the density of states will be confined inside a minimal volume V_F in k space it, where

$$V_F = \frac{4}{3}\pi k_F^3$$

The Fermi momentum is implicitly defines Fermi's momentum k_F, where

$$\epsilon_F = \frac{\hbar^2 k_F}{2m} \Longrightarrow k_F = \sqrt{\frac{2m}{\hbar^2}\epsilon_F}$$

The density of states will be $\rho(k) = 2V/8\pi^3$ for electrons ($g = 2$) and the total number of states will therefore be

$$N_s = \rho(k)V_F = \frac{V}{3\pi^2}k_F^3$$

Which, if we write $n = N_s/V$ as the number density of states gives

$$k_F = \sqrt[3]{3n\pi^2} \tag{24.19}$$

Rearranging everything using (24.19) we end up having the following expressions for the Fermi energy in terms of the density of electrons

$$\epsilon_F(n) = \frac{\hbar^2}{2m}(3n\pi^2)^{\frac{2}{3}}$$
$$n(\epsilon_F) = \frac{1}{3\pi^2}\left(\frac{2m\epsilon_F}{\hbar^2}\right)^{\frac{3}{2}} \tag{24.20}$$
$$n(k_F) = \frac{k_F^3}{3\pi^2}$$

Another intrinsic characteristic of the metal is the Fermi temperature T_F of such, which is simply $T_F = k_B^{-1}\epsilon_F$.

Using statistical mechanics we can approximate the energy distribution for electrons, and expanding around ϵ_F we get

$$g(\epsilon) = \frac{4}{3}\frac{n}{\epsilon_F}\sqrt{\frac{\epsilon}{\epsilon_F}} \qquad \epsilon \geq 0$$

Inserting this into the Fermi-Dirac distribution we have that

$$\mu \approx \epsilon_F - \frac{\pi^2}{6}\frac{g'(\epsilon_F)}{g(\epsilon_F)} = \epsilon_F\left[1 - \frac{1}{3}\left(\frac{\pi}{2}\frac{T}{T_F}\right)\right]$$

Indicating that since $T/T_F << 1$ at room temperature, Sommerfeld's free electron approach is a good approximation for what happens in a metal, since $\epsilon \approx \epsilon_F$
Suppose now that we want to calculate another property of the metal itself, the specific heat at constant volume.
Using statistical mechanics again we have that the energy per single electron is given by

$$\frac{E}{N} = \frac{u}{n} = \frac{u_0}{n} + \frac{\pi^2}{6}\frac{2g(\epsilon_F)}{n}(k_BT)^2$$

Where $g(\epsilon_F)$ is our well known Fermi-Dirac distribution evaluated at the Fermi energy. Using the well known formula for calculating the specific heat we get

$$c_V = \frac{\partial u}{\partial T} = \frac{\pi^2}{2}\left(\frac{k_BT}{\epsilon_F}\right)k_BT = \frac{\pi^2}{2}\left(\frac{T}{T_F}\right)nk_B$$

§ 24.4 blah blah

hi this is a test <++>

Part V

Nuclear and Particle Physics

25 Scattering and Cross Sections

§ 25.1 A Small Intro to Some Wacky Units

Since in particle physics usually we deal with particular calculations, it's preferable to avoid using the SI system of units, and instead pass to what I like to call, the system of «God given units» , where the most common fundamental constants are taken as unitary, i.e. $\hbar = c = 1$.

With this choice is also common to write energy in terms of eV, i.e., using $c = 1$ and $E = \gamma mc^2$, we have that

$$[E] = [m] = eV$$

Using the dispersion relation we also get

$$E = p^2 + m^2 \Longrightarrow [p] = eV$$

Therefore mass, energy and momentum are all expressed in eV. Using that $1\,J = (1.602)^{-1} \times 10^{19}\,eV$ we get the conversion value

$$1\,kg = 5.6 \times 10^{35}\,eV \tag{25.1}$$

In these units we have that the mass of the electron m_e and the mass of the proton m_p are

$$m_e = 9.109 \times 10^{-31}\,kg = 0.511\,MeV$$
$$m_p = 1.673 \times 10^{-27}\,kg = 938.3\,MeV \tag{25.2}$$

The second consequence of taking $\hbar = c = 1$ is that time can also be expressed in terms of eV. In fact since $[\hbar] = Js$ in the SI system, and $\hbar = 1$ in the GGS[1], we have

$$\hbar = 1.055 \times 10^{-34}\,Js = 6.583 \times 10^{-22}\,MeVs$$

Therefore

$$1\,s = \frac{1}{\hbar}\,MeV^{-1} = 1.519 \times 10^{21}\,MeV^{-1} \tag{25.3}$$

---

[1]God Given System

225

Combining both $\hbar c$ we have $[\hbar c] = MeV m$, therefore we can think of expressing distances with this unit. Multiplying the constants we get

$$\hbar c = 197.35 \times 10^{-15} \text{ MeVm} = 197.35 \text{ MeVfm}$$

This implies that

$$1 \text{ fm} = 5.608 \text{ GeV}^{-1} \tag{25.4}$$

In these units is also quick to see that

$$\alpha = \frac{e^2}{4\pi\epsilon_0 \hbar c} \frac{1}{137} \tag{25.5}$$

§ 25.2 Cross Section

Consider a beam of particles colliding with a target which is long d. The N_p particles of the beam will react with the N_t particles of the target N_R times in some time T.
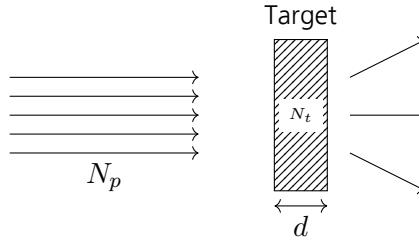
Figure 25.1: Stylization of a beam of particles colliding with a target thick d

Considering the previous data, we can immediately say that the probabilty of interaction between a beam particle and a target particle in a time T, if the particle density of the target is n_t, is

$$\sigma = \frac{1}{n_t d} \frac{N_R/T}{N_p/T} \approx \frac{1}{n_t d} \frac{\dot{N}_R}{\dot{N}_p} \tag{25.6}$$

This probability is commonly known as the «cross-section» for the reaction.

Consider now a beam with cross sectional surface S
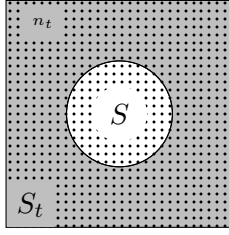
Since we have $S < S_t$ as in figure (25.2) we have that the target is going to get hit in a fraction of his surface. Considering that the target is long d, we have that the amount of particles that might get hit by the beam will be

$$N_t = n_t S d \implies n_t d = \frac{N_t}{S} \quad (25.7)$$

Figure 25.2: Cross section of the beam with surface S hitting a target with surface S_t and particle density n_t

Inserting it back into the formula of the cross section we have

$$\sigma = \frac{\dot{N}_R}{N_t} \frac{1}{S\dot{N}_p} \quad (25.8)$$

Since $S\dot{N}_p$ gives the amount of particles passing through a cross-sectional surface of the beam per unit time (i.e., a flux), we can define the flux of particles ϕ_p as follows

$$\phi_p = S\frac{\mathrm{d}N_p}{\mathrm{d}t} \quad (25.9)$$

I.e.

$$\sigma = \frac{1}{N_t\phi_p} \frac{\mathrm{d}N_R}{\mathrm{d}t} \quad (25.10)$$

Noting also that if the particle beam is long L, if its particle density is n_p we have, if N_p is the total number of beam particles moving with average velocity v_p

$$N_p = n_p v_p S \mathrm{d}t$$

And therefore, since N_p is fixed

$$\phi_p = \frac{1}{S}\frac{\mathrm{d}N_p}{\mathrm{d}t} = n_p v_p$$

Which gives

$$\sigma = \frac{1}{n_p v_p N_t} \frac{\mathrm{d}N_R}{\mathrm{d}t}$$

And solving for $\dot{N}_R$ we have

$$\frac{\mathrm{d}N_R}{\mathrm{d}t} = \sigma n_t d \frac{\mathrm{d}N_p}{\mathrm{d}t} \quad (25.11)$$

Which is the differential equation that gives the number of reactions per unit time. Note how $[\sigma] = L^2$, therefore it has units of m^2 in the S.I., in particle physics these units are quite

uncomfortable since the cross sections evaluated are infinitesimally small, and another unit is used, the «barn» . As follows there are the conversions

$$1 \text{ b} = 10^{-28} \text{ m}^2 \tag{25.12}$$

For common reactions we have

$$
\begin{aligned}
\text{p} + \text{p} &\rightarrow \text{X} & \sigma_{pp} &\approx 10 \text{ mb} = 10^{-30} \text{ m}^2 \\
\nu_{\text{e}} + \text{p} &\rightarrow \text{X} & \sigma_{\nu p} &\approx 10 \text{ fb} = 10^{-14} \text{ b} \\
\chi + \text{p} &\rightarrow \text{X} & \sigma_{\chi p} &\approx 10^{-2} \text{ fb} = 10^{-17} \text{ b} = 10^{-45} \text{ m}^2
\end{aligned}
\tag{25.13}
$$

§§ 25.2.1 Crossed Beam Interaction and Luminosity

Consider two beams colliding frontally like they'd do in the LHC, a 27km particle accelerator ring.

Figure 25.3: Schematization of the two beams colliding at some point inside a particle collider

The target we're now considering is a second beam with cross sectional surface $S_2 = S_1 = S$ and N_2 particles. Rewriting the formulas using $N_p = N_1$ and $N_t = N_2$ we get

$$\frac{dN_R}{dt} = \sigma n_2 d \frac{dN_1}{dt} = \frac{\sigma}{S} N_2 \frac{dN_1}{dt}$$

Using $\phi_1 = S^{-1} \dot{N}_1$ we simply get

$$\frac{dN_R}{dt} = \sigma \phi_1 N_2 \tag{25.14}$$

We can define a new quantity, f_{int}, i.e. the «interaction frequency» of the beam, we have that the flux can be redefined as follows

$$\phi_1 = \frac{N_1 f_{int}}{S} \tag{25.15}$$

And this gives

$$\frac{dN_R}{dt} = \sigma \frac{N_1 N_2 f_{int}}{S} = \sigma \mathcal{L} \tag{25.16}$$

The new quantity $\mathcal{L}$ is called the «instantaneous luminosity» of the collider and it has units of $L^{-2}T^{-1}$ and is commonly expressed therefore in Hz b^{-1} or Hz cm^{-2}.
Integrating the previous formula we get the «integrated luminosity» L, measured in b^{-1}

$$L = \int_{t_0}^{t_1} \mathcal{L} dt$$

§§ 25.2.2 Higgs Bosons, Exclusive and Inclusive Cross Sections

Going back to what we said for the luminosity of a detector, for the well known Large Hadron Collider we have

$$\mathcal{L}_{LHC} \approx 10^{34} \, \frac{\text{Hz}}{\text{cm}^2} = 100 \, \frac{\text{Hz}}{\text{b}}$$

More specifically, during the years and the various runs it varied. The highest value was reached with run 3 of the LHC in 2018, where

$$L_{LHC} = 163 \text{ fb}^{-1}$$

This value can be used to estimate the number of events happened for a certain reaction. Take as an example the creation of Higgs boson

$$p + p \rightarrow H + X$$

The number of Higgs produced will be given using the integrated luminosity of LHC in run 3 and the cross section for the creation of $H + X$ particles

$$N_H = L\sigma_{HX}$$

Here $\sigma_{HX} \approx 10^{-10}$ b and $\sqrt{s} = 13$ TeV.

$$N_H = 163 \cdot 10^{-10} \text{bfb}^{-1} = 163 \cdot 10^5$$

Therefore, in run 3, between 2015 and 2018 around 16 million Higgs bosons were produced in the LHC.

It's important tho to note that there are various ways a Higgs boson can decay, which will be what is going to be measured at the detector. These decays will be

$$H \rightarrow bb$$
$$H \rightarrow \gamma\gamma$$
$$H \rightarrow ZZ$$
$$H \rightarrow WW$$

The probability for having one of these decays is known as the «branching factor» BF. Suppose that now we want to see how many reactions of the kind $H \rightarrow \gamma\gamma$ happened in run 3, then we will have to weigh the total number of Higgs bosons with its branching factor

$$N_{\gamma\gamma} = N_H BF_{H \rightarrow \gamma\gamma} \tag{25.17}$$

Suppose instead that we want to analyze a more complex decay of the Higgs boson

$$H \rightarrow ZZ$$
$$ZZ \rightarrow \begin{cases} e^+ + e^- \\ \mu^+ + \mu^- \\ \tau^+ + \tau^- \end{cases} \tag{25.18}$$

The branching factor for the $H \to ZZ$ is $BF_{H \to ZZ} = 3\%$, which means that

$$N_{ZZ} = 163 \cdot 10^5 \cdot 3 \cdot 10^{-2} = 489 \cdot 10^3$$

This is not enough to check how many Higgs have been produced, since only the $Z + Z \to e^+ + e^-$ reaction is measured, and the total chain reaction we gotta consider becomes

$$
\begin{aligned}
p + p &\longrightarrow H + X \\
&\longrightarrow Z^\pm + Z^\pm \\
&\longrightarrow e^+ + e^- \\
&\longrightarrow e^+ + e^-
\end{aligned}
$$
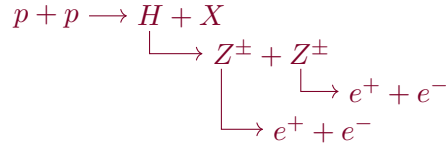
Therefore, the real number of measured Higgs bosons $N_H^\star$ will be

$$N_H^\star = \sigma_H L BF_{ZZ} BF_{Zee}^2 = 163 \cdot 10^5 \cdot 9 \cdot 10^{-4} \approx 432$$

I.e., if the LHC's CMS detector was a perfect detector we would have measured *at most* 432 Higgs bosons in the 4 full years of the third run.

A major problem now comes into play. What if I don't know the cross section for the reaction σ_{HX}?. In general we have that the number of particles is given by the number of counts of the detector, the integrated luminosity is measured directly, and the branching factors can be either measured or determined theoretically. Therefore, simply inverting and adding a factor ε determining the efficiency of the detector, which accounts for imperfections in its surface, we have

$$\sigma_{HX} = \frac{N_H}{L BF_{HZZ} BF_{Zee}^2} \frac{1}{\varepsilon}$$

This is not enough for a proper determination of values, since in the LHC there are various reactions that get measured. We have that for each reaction with cross section σ_i, the total cross section of the $p + p \to X$ reaction is

$$\sigma_{ppX} = \sum_{i=1}^{n} \sigma_i \tag{25.19}$$

The cross sections on the right are known as «exclusive cross sections» and the one on the left is the «total cross section» for the measured event.

Note that in searching for the Higgs boson, we actually want to measure the photons from pair annihilation of the two electrons of the previous reaction, since their energies will peak around $m_H c^2$, giving the complete reaction

$$
\begin{aligned}
p + p &\longrightarrow H + X \\
&\longrightarrow Z^\pm + Z^\pm \\
&\longrightarrow e^+ + e^- \longrightarrow \gamma + \gamma \\
&\longrightarrow e^+ + e^- \longrightarrow \gamma + \gamma
\end{aligned}
$$
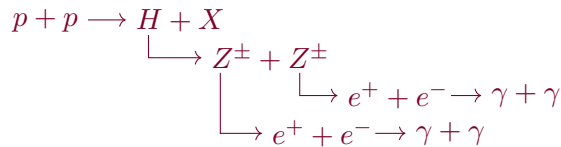
Generally we're seeing the decay of a state $|pp\rangle$ into a state $|e^+e^-e^+e^-\rangle$, passing through a $Z^\pm$ state. Using SR we have

$$
\begin{aligned}
P^\mu_{Z_1} &= p^\mu_{e^+_1} + p^\mu_{e^-_1} \\
P^\mu_{Z_2} &= p^\mu_{e^+_2} + p^\mu_{e^-_2}
\end{aligned}
\tag{25.20}
$$

The 4 moment of the 4 electrons is measured, and using the well known formula for the invariant mass we have that

$$
m_{ZZ} = \sqrt{P^\mu_{(1)} P^{(2)}_\mu}
$$

Which, also gives the energies of the emitted photons in the center of mass of the decay. Measuring various combinations of such at LHC in 2012, finally a peak in counts of reactions at around $E^\star_{\gamma\gamma} \approx 125$ GeV was measured, which corresponds to the measured mass of this famous boson.

§§ 25.2.3 Mean Free Path

Starting again for the usual and now well known relation for giving the number of reactions per unit time for some beam of particles hitting a target long l, we can imagine to generalize everything to differential lengths $l \to \mathrm{d}x$, giving us

$$
\frac{\mathrm{d}N_R}{\mathrm{d}t} = \sigma n_t \frac{\mathrm{d}N_p}{\mathrm{d}t} \mathrm{d}x
$$

The interaction probability will be

$$
P_{int} = \frac{\dot{N}_R}{\dot{N}_p} = \sigma n_t \mathrm{d}x = \sigma \frac{N_t}{S}
$$

Define now an «absorption coefficient» μ. This coefficient indicates the amount of particles of the beam that get "absorbed" by the target per unit length. Using some reasoning with the previous formulae we must have that it must depend linearly to the linear density of particles of the target n_t, and therefore

$$
\mu = \sigma n_t
\tag{25.21}
$$

We have now that the pdf for the interaction energy is

$$
P_{int} = \mu \mathrm{d}x
\tag{25.22}
$$

This, must also be proportional to the derivative of the flux, since the flux must lower by some quantity while passing through the target. Finally we have

$$
\frac{\mathrm{d}\phi_p}{\mathrm{d}x} = -P_{int}\phi_p = -\mu\phi_p \mathrm{d}x
\tag{25.23}
$$

Solving the ODE we get

$$
\phi_p(x) = \phi_0 e^{-\mu x}
\tag{25.24}
$$

Noting that μ has units of L^{-1} we can imagine to define a length λ as the inverse of this attenuation factor, known as the «mean free path» of the particle.

The main utility of this length is that it can be used to determine the interaction cross section experimentally. In fact, imagine sending a flux of particles through N targets with increasing lengths d_i as in the next figure
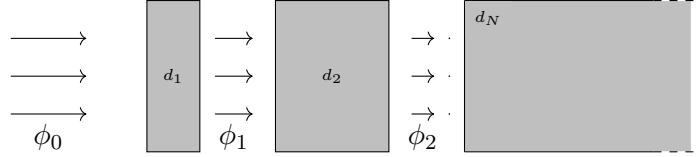
Figure 25.4: Example sketch of the experiment

Since $\phi \propto e^{-\mu x}$ we could use an exponential fit to find μ. Using $\mu = \sigma n_t$ it's possible to estimate σ if the properties of the target element are known

§ 25.3 Differential Cross Section

Consider now a realistic approach for the collision of a beam of particles with a target. In this realistic approach the detectors will occupy part of the space around the target, and therefore there will be some preferred angles in order to detect properly the particle.

This final angle depends on the flight angle of the scattered particle.
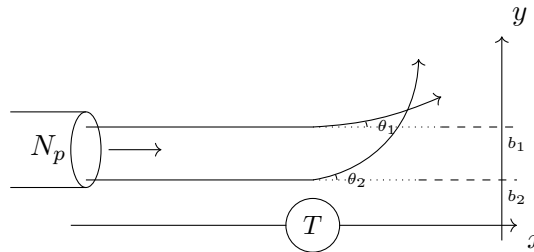
Figure 25.5: Sketch of the scattering process in study

In the previous figure it's possible to see the sketch of this kind of scattering. The two parameters b_1, b_2 are known as the «impact parameters» of the two particles, while the two angles θ_1, θ_2 are the «flight angles» of the two particles.

In general, it's not hard to believe that this flight angle will depend on the impact parameter b which is the vertical distance between the target and the unperturbed particle path's y height.

From this supposition we can immediately say

$$\begin{cases} g(b) = \theta \\ f(\theta) = b \end{cases} \Longrightarrow db \propto d\theta$$

Imagining a toroidal detector around the target we can transform this reasoning into 3D, where $\mathrm{d}\theta \to \mathrm{d}\Omega = \sin\theta\mathrm{d}\theta\mathrm{d}\varphi$. In this case, noting that $b\mathrm{d}b = \sin\theta\mathrm{d}\theta$ we get a new «differential cross-section» $\mathrm{d}\sigma$

$$\mathrm{d}\sigma = b\mathrm{d}b\mathrm{d}\theta \tag{25.25}$$

Manipulating this a bit we get

$$\mathrm{d}\sigma = \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\mathrm{d}\Omega = \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}|\sin\theta\mathrm{d}\theta\mathrm{d}\varphi| = b\mathrm{d}b\mathrm{d}\varphi$$

Or, rearranging everything

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{b}{|\sin\theta|}\left|\frac{\mathrm{d}b}{\mathrm{d}\theta}\right| \tag{25.26}$$

Where the absolute value comes from $\sigma > 0$.

Example 25.3.1 (An Easy Example of Differential Cross-Section). As an example it's really exemplar finding the differential cross section for a particle hitting a rigid sphere with radius R.
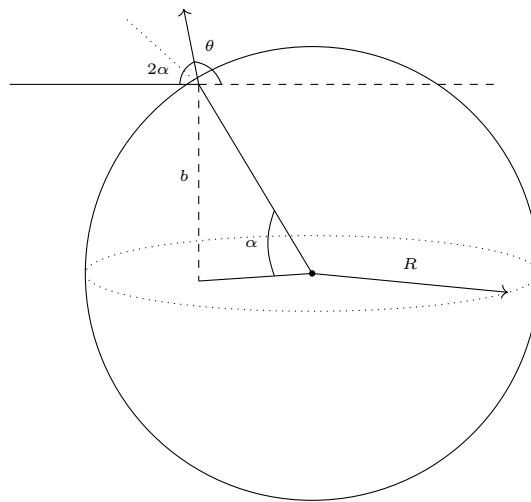
Figure 25.6: Quick sketch of the particle beam colliding with the cited sphere

This rigid sphere can be imagined as a potential wall in 3D spherical coordinates, where, simply

$$\mathcal{U}(R) = \begin{cases} 0 & r > R \\ \infty & r < R \end{cases}$$

It's obvious from the picture also that:

$$\begin{cases} b = R\sin\alpha \\ 2\alpha + \theta = \pi \end{cases}$$

Solving the second equations for $\alpha(\theta)$ we get that

$$\alpha = \frac{\pi - \theta}{2}$$

And therefore

$$\sin(\alpha) = \cos\left(\frac{\theta}{2}\right) \Longrightarrow b = R\cos\left(\frac{\theta}{2}\right)$$

Deriving b with respect to θ and getting its absolute value we get

$$\left|\frac{\mathrm{d}b}{\mathrm{d}\theta}\right| = \frac{R}{2}\left|\sin\left(\frac{\theta}{2}\right)\right|$$

And, simply substituting everything into (25.26) we get

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{R^2}{2|\sin\theta|}\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) = \frac{R^2}{4}$$

Having now a simple differential equation for σ we have finally

$$\sigma = \frac{R^2}{4}\int_{4\pi}\mathrm{d}\Omega = \pi R^2$$

Which means that, if the beam has cross sectional surface S, the probability of interaction of the beam with the sphere is

$$P_{int} = \frac{\pi R^2}{S}$$

This is in complete accord with the impulsive idea that the interaction probability in this case will be given from the exposed surface of the sphere divided by the surface of the beam, giving without problems a function of reactions in terms of cross-sectional surface of the beam

$$N_R(S) = \frac{\sigma}{S} = \frac{\pi R^2}{S}$$

Having said all of this one might rightfully ask how would someone measure the differential cross section.
Start with hitting the target in question with a beam of particles with known flux, then the number of counts per unit time will be the usual well known formula

$$\frac{\mathrm{d}N_R}{\mathrm{d}t} = \sigma n_t d\frac{\mathrm{d}N_p}{\mathrm{d}t}$$

Substituting inside this the differential quantities we have

$$\frac{\mathrm{d}N_R}{\mathrm{d}t} = n_t d\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\frac{\mathrm{d}N_p}{\mathrm{d}t}\mathrm{d}\Omega$$

Expressing $n_t d$ in terms of differential surface we can say immediately that

$$\frac{n_t}{S}\frac{\mathrm{d}N_p}{\mathrm{d}t}\mathrm{d}S = \phi_p(x)N_t$$

Which gives

$$\frac{\mathrm{d}N_R}{\mathrm{d}t} = \phi_p(x)N_t\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\mathrm{d}\Omega$$

Since every term on the right is known and the number of counts per unit time is measured by the detector we can solve for the differential cross section, getting

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{\phi_p(x)N_t}\frac{\mathrm{d}N_R}{\mathrm{d}t}\frac{1}{\mathrm{d}\Omega} \tag{25.27}$$

26 Nuclear Physics

§ 26.1 First Discoveries

At the end of the '800s we managed to discover radioactivity, and it was divided as follows

- X rays, discovered by Röntgen in 1895

- Natural radioactivity observed in phosphorescent materials as ^{238}U salts, of which

  - α rays

  - β rays

  - γ rays

The firsts, X-rays, were known for passing easily through matter and leave traces in photographic tables, they're now known as high energy photons
The last, natural radiation in the shape of α, β and γ radiation are now today known as emission of particular particles by atomic nuclei

1. γ rays, today known as photons for which $E > E_X$

2. β rays, today known as electrons and their antimatter counterpart, the positron

3. α rays, now known as Helium nuclei, which are emitted only from heavy nuclei

§§ 26.1.1 Thompson and the Discovery of the Electron

One of the first experiments was conducted by Thompson, Milliken et. al which studied the nature of β particles using a CRT as in picture

$$V_C << V_A$$
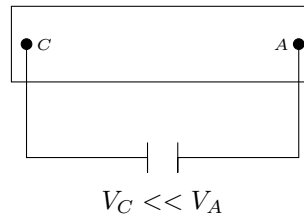
Figure 26.1: An example of a CRT's functioning, the cathode C and the anode A are enclosed in a glass tube

The Thompson-Milliken experiment goes as follows: Filling the CRT tube with different gases it's possible to measure events in function of the kind of gas, its pressure, the potential difference $V_A - V_C$ and obviously in function of the material used to produce the cathode. The observations that the scientists reported were

- A green luminescence close to the anodes

- Electric shocks present also with small ΔV

- The electrical screening between the anode and cathode grows with ΔV

- All these effects are independent from the presence of a magnetic field B

They ended up with an hypothesis: CRT rays must be charged particles.
Thompson went forward proving this setting a new experiment, set up as in the following picture
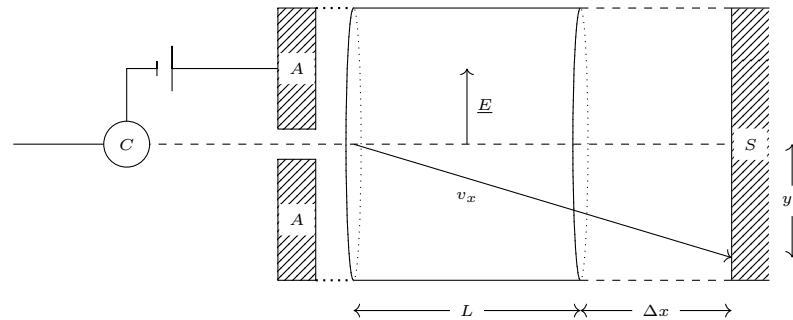
Figure 26.2: Scheme of the Thompson experiment. The electrons emitted from the cathode C travel through a cylindrical capacitor long L at a constant velocity v_x. Interacting with the electric field the path of the radiation variates and this variation is measured in a screen S at the end

In this experiment, we have $V_A >> V_C$ and therefore the electrons are accelerated from C to A, from which continue traveling with constant velocity v_x. The y component of the

velocity gets perturbed by an electrostatic force, which gives us, inside the cylinder

$$\begin{cases} a_x = 0 \\ a_y = \dfrac{q}{m} E \end{cases} \tag{26.1}$$

Which implies

$$\begin{cases} x(t) = v_x t \\ y(t) = \dfrac{q}{2m} E t^2 \end{cases}$$

Since the capacitor is long L we have that when exiting the capacitor at a time $t_L = L/v_x$ the particle will be at the following y position with v_y velocity

$$\begin{cases} y(t_L) = \dfrac{q}{2m} \dfrac{E L^2}{v_x^2} \\ v_y(t_L) = \dfrac{q}{m} \dfrac{E L}{v_x} \end{cases}$$

The final measured y position on the screen, y_s will be given by an easy computation

$$y_s = y(t_L) + v_y(t_L) \frac{\Delta x}{v_x}$$

Where Δx is the horizontal distance between the end of the capacitor and the screen. We get therefore

$$y_s = \frac{qE}{2m} \frac{L^2}{v_x^2} + \frac{qE}{m} \frac{L\Delta x}{v_x^2} = \frac{qE}{m} \frac{L}{v_x^2} \left(\frac{L}{2} + \Delta x \right) \tag{26.2}$$

The problem with this equation is that v_x is unknown and therefore it's not possible to find q/m.

Accounting for an orthogonal magnetic field B and applying the Lorentz force it's possible to find that

$$v_x = \frac{E}{B}$$

And therefore

$$\frac{q}{m} = \frac{y_s}{L \left(\frac{L}{2} + \Delta x \right)} \frac{E}{B} \tag{26.3}$$

This is a property of the projectile since it's invariant with the gas and the components of the anodes, and Nobel in 1906 calculated

$$\frac{q}{m} = 1.76 \cdot 10^{11} \ \text{C/kg}$$

A subsequent experiment by Milliken et al. in 1923 used a falling drop of oil in order to calculate the charge of this new particle. The setup of the experiment was as follows
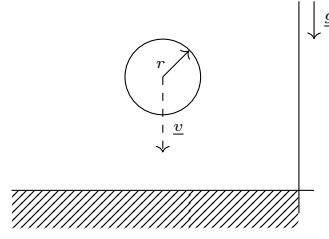
Figure 26.3: Experimental setup of the falling droplet of oil

This experiment used a falling drop of oil. The mass of the drop is quickly calculated assuming a spherical drop

$$m_p = \frac{4\pi}{3} r^3 \rho_{oil}$$

Using hydrodynamics the drag force's modulus is quickly calculated accounting for the viscosity η as follows

$$F_d = -6\pi \eta r v_0$$

Where v_0 is experimentally measured as the terminal velocity of the droplet. The classical equilibrium is reached when the weight of the droplet is balanced by the hydrodynamic drag

$$\frac{4\pi}{3} \pi r^3 \rho_{oil} g = 6\pi r v_0$$

Solving for r we get a measurable formula for the radius of the droplet, i.e.

$$r = \sqrt[3]{\frac{\eta v_0}{2\rho_{oil} g}} \tag{26.4}$$

Applying now an electric field opposite to the motion of the droplet we get a new terminal velocity v_1 which can be experimentally measured, gives a relation between the known measured variables and the charge, as follows

$$qE = 2\pi r \left(\frac{2r^2}{3} \rho_{oil} g - 3\eta v_1 \right)$$

Substituting for r and with some algebra we get

$$qE = 18\eta \pi \sqrt{\frac{\eta v_0}{2\rho_{oil} g}} (v_0 - v_1)$$

Since E is known and the variation in terminal velocities of the droplet is directly measured we get that the charge of this new particle is

$$q = \frac{18\eta \pi}{E} \sqrt{\frac{\eta v_0}{2\rho_{oil} g}} \Delta v = 1.59 \cdot 10^{-19} \text{ C} \tag{26.5}$$

The final value reached from Milliken is so precise that its relative error is in the 1% from the modern known value.

Mixing the result of these two experiment it's possible to evaluate the mass of these beta particles, which gives

$$m_\beta = 0.911 \cdot 10^{-30} \text{ Kg} \approx 511 \text{ KeV} \tag{26.6}$$

This beta particle is now a well known fundamental particle that we treated already in depth in the previous chapters, which is the electron.

§§ 26.1.2 Discovery of the Nucleus

After the discovery of the electron and its properties with β radiation, Rutherford et al. proceeded with a new experiment trying to discern the physics behind α particles. It was known that for some atom, with A big enough, we can have a radioactive decay through the emission of an α particle, in the following reaction

$$^A_Z\text{N} \rightarrow ^{A'}_{Z'}\text{M} + \alpha \tag{26.7}$$

Note that today it's known that $\alpha = ^4_2\text{He}$, $m_\alpha = 3.7$ GeV and $A' = A - 4$, $Z' = Z - 2$. These α particles emitted naturally by these heavy nuclei have a center of mass energy of around 5 MeV. Using special relativity we know that they aren't relativistic, in fact we have

$$E_{cm} = E_\alpha - m_\alpha = (\gamma - 1)m_\alpha \Longrightarrow \gamma \approx 2.35 \cdot 10^{-3}$$

Rutherford's experiment was initially in order to determine which of the nuclear models was true. At the time there were two main contending ideas, one being Thompson's idea, where the electrons $q < 0$ were situated inside this nucleus with $q_n > 0$ and the total atom was neutral (the so called Pancake nucleus) whereas Rutherford supposed of an atom with the positive charges all in the center and electrons orbited these positive charges, rendering in general the atom neutral.
This experiment is extremely similar to the setup (25.5), where the target is our nucleus with $q > 0$.
The experimental setup that Rutherford et al. used consisted of a radioactive α emitter bar targeting these particles to a foil of gold $_{79}\text{Au}$ which is surrounded by a $150°$ spherical detector, which would measure the differences in the deflection angles.
The source chosen was Radium Bromide, RaBr_2, an alpha emitter radioactive compound.
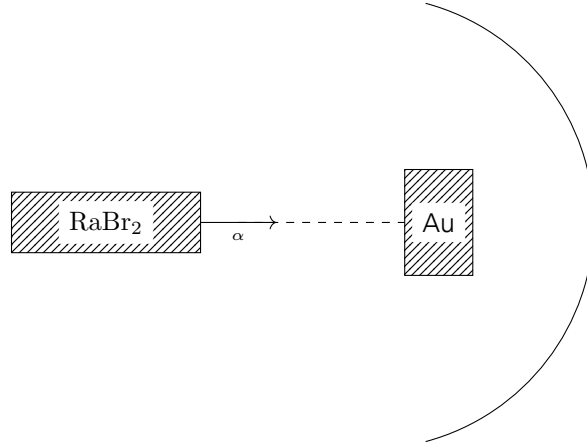
Figure 26.4: The setup was as follows, the radioactive source was placed at some distance from a gold plate at the center of a semi-spherical detector, which would then be used to determine the deviation angles of the deflected α particles

The main scope of the experiment was to measure the numbers of reaction with respect to the deviation angle.

This can be done by firstly finding the differential cross section. The experimental setup makes it much easier since the spherical symmetry of the system easies the calculation.

Since the angular momentum of the particles and energy is conserved we can begin by writing the Lagrangian of the system

$$\mathcal{L}_\alpha = \frac{m_\alpha}{2}\left(\dot{r}^2 + r^2\dot{\varphi}^2\right) - \mathcal{U}(r) \tag{26.8}$$

Thanks to the spherical symmetry we have that $\dot{\varphi}$ is cyclical and its momentum will be conserved, i.e.

$$p_\varphi = mr^2\dot{\varphi} = L$$

And the energy of the system is

$$E = p\dot{q} - \mathcal{L}_\alpha = \frac{1}{2}m_\alpha\dot{r}^2 + \frac{1}{2}m_\alpha r^2\dot{\varphi}^2 + \mathcal{U}(r)$$

Substituting the cyclical coordinate with the conserved quantity L we have

$$E = \frac{1}{2}m\dot{r}^2 + \frac{L^2}{2mr^2} + \mathcal{U}(r) = T(\dot{r}) + \mathcal{U}_{eff}(r) \tag{26.9}$$

Where $\mathcal{U}_{eff}(r)$ is an effective potential that includes the centrifugal potential.

Solving for $\dot{r}$ we get thanks to the conservation of energy the following differential equation

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \sqrt{\frac{2}{m_\alpha}}\sqrt{E - \frac{L^2}{2mr^2} - \mathcal{U}(r)}$$

In order to evaluate in terms of the deflection angle, using p_φ we can substitute $\mathrm{d}t$ with $\mathrm{d}\varphi$, using the straightforward substitution

$$\mathrm{d}\varphi = \frac{L}{m_\alpha r^2}\mathrm{d}t$$

Which gives

$$\mathrm{d}\varphi = \frac{\frac{L}{m_\alpha r^2}}{\sqrt{2m_\alpha}\sqrt{(E - \mathcal{U}(r)) - \frac{L^2}{r^2}}}\mathrm{d}r$$

It's now useful to pass everything to the impact parameters for the α particle.
The conservation of energy permits us to write without problems that

$$\begin{cases} p_i = m_\alpha v_\infty \\ E_\infty = E_0 = \frac{1}{2}m_\alpha v_\infty^2 \\ L = \|r_\infty \wedge p_\infty\| = bm_\alpha v_\infty \end{cases} \tag{26.10}$$

Where r_∞ is the distance of the particle from the nucleus of gold, v_∞ is the velocity "at infinity", i.e. the velocity of the non-interacting particle, and b is the already known impact parameter.
Substituting into the differential of φ we have that

$$2m_\alpha \left(\frac{1}{2}m_\alpha v_\infty^2 - \mathcal{U}(r) - \frac{b^2 m^2 v_\infty^2}{r^2}\right) = m_\alpha^2 v_\infty^2 \left(1 - \frac{\mathcal{U}(r)}{E_\infty} - \frac{b^2}{r^2}\right)$$

Substituting it back and integrating we have

$$\varphi(r) = \int_{r_{min}}^{\infty} \frac{b}{r\sqrt{1 - \frac{\mathcal{U}(r)}{E_\infty} - \frac{b^2}{r^2}}}\mathrm{d}r$$

The potential is the usual Coulomb potential, which in natural units is

$$\mathcal{U}(r) = \frac{\alpha Z_p Z_n}{r} = \frac{A}{r}$$

Substituting it back into the integral we get a solvable integral

$$\varphi(r) = \int_{\infty}^{r_{min}} \frac{b}{r\sqrt{1 - \frac{A}{E_\infty r} - \frac{b^2}{r^2}}}\mathrm{d}r = \arccos\left[\frac{\frac{A}{2E_\infty b}}{\sqrt{1 + \left(\frac{A^2}{2E_\infty b}\right)}}\right] \tag{26.11}$$

Writing $B = A/2E_\infty b$ lets us invert the function in terms of the impact parameter, giving us

$$b(\varphi) = \frac{A}{2E_\infty}\tan(\varphi) \tag{26.12}$$

It's quick to see that $\varphi = \pi/2 - \theta/2$, therefore

$$b(\theta) = \frac{A}{2E_\infty} \cot\left(\frac{\theta}{2}\right) \tag{26.13}$$

Deriving with respect to θ and inserting it into the equation for the differential cross section we get

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2 Z_p^2 Z_n^2}{16 E_\infty^2} \csc^4\left(\frac{\theta}{2}\right) \tag{26.14}$$

Rutherford confirmed this cross section, and went forward estimating the value of r_{min}. Using $E_\infty = E_\alpha = 5$ MeV and the conservation of energy we have

$$\mathcal{U}(r_{min}) = \frac{\alpha Z_p Z_n}{r_{min}} = \frac{1}{2} m_\alpha v_\infty^2 = 5 \text{ MeV}$$

I.e.

$$r_{min} = \frac{\alpha Z_p Z_n}{5} \text{ MeV}^{-1} = 0.23 \text{ MeV}^{-1}$$

Converting into more usual units we have that $r_{min} = 46$ fm, which from the experiment it was confirmed that $r_{min} < 30$ fm.

§§ 26.1.3 Discovery of the Proton and of the Neutron

Again from Rutherford et al. in 1918, still using α radiation, the proton was discovered. Consider the following nuclear reaction

$$\alpha + {}^{14}_{7}\text{N} \to {}^{17}_{8}\text{O} + X$$

This reaction included an artificial nuclear transmutation and the emission of an unknown particle X on which a spectrography was executed and the measured q/m was compatible with H^+. This particle was called the proton, with symbol p. The reaction was now completed

$$\alpha + {}^{14}_{7}\text{N} \to {}^{17}_{8}\text{O} + p$$

Continuing on this path, Chadwick et al. studied another reaction, for which an unknown neutral particle was discovered. The reaction is

$$\alpha + {}^{9}_{4}\text{Be} \to {}^{12}_{6}\text{C} + X$$

The creation of this particle was observed also with ${}^{9}_{4}\text{Li}$ and ${}^{9}_{4}\text{B}$ This particle is heavily piercing, and therefore two main hypotheses were considered

1. The particle is a photon

2. It's a new heavy neutral particle

The idea for evaluating which is true was to consider the scattering with a proton from which to measure the impulse of p and therefore evaluate the mass of this unknown particle. The measured impulse of the proton after the scattering event had velocities $\beta \approx 0.1$ and therefore can be considered weakly relativistic.

Curie et al. went forward proposing the idea that the X particle was a photon with an energy of around 50 MeV, which tho contradicted previous experiments for which the energy of such photon would have been of the order of one MeV.

Chadwick et al. finally concluded that this was a new neutral massive particle, for which $m_n \approx m_p \pm 10\%$. This particle was called the neutron, and is between the early fundamental blocks of nuclear physics together with p, e^- and α particles.

§§ 26.1.4 Modern Considerations and Experiments

In our modern understanding of nuclear physics we know 7 main properties of atomic nuclei

- Mass A

- Charge Z, number of protons

- Number of neutrons $N = A - Z$

- Nuclear spin

- Magnetic moment μ

- Electric moments and quadrupole moments $'m$

- Isospin

All the chemical properties are tied to the charge of the nucleus Z and therefore compose the periodic table.

With these properties one can use the two main ones, A, Z in order to build a couple and from there identify a nucleus. In general, a nuclear object with mass and charge (A, Z) will be called a «nuclide» .

An «isotope» is a nuclide of some determined element (A', Z) for which although the mass is different, the charge is the same. Vice-versa one can define «isobars» , or a nuclide of some determined element (A, Z') for which charge is different but the mass is the same.

The properties of a nucleus can be found out with different experiments.

Starting for mass one can repeat the Thompson experiment using a mass spectrometer, while for charge one can use an X ray spectroscopy of the internal electron shells of the atom. Other experiments include tests for finding the radius of a nuclide, repeating the Chadwick-Rutherford experiment with either highly energetic electrons in order to have a smaller resolution, or using μ-mesic studies of the atom with muons instead of electrons, which all give finer details on the nucleus.

The final conclusion from all the various experiments are that

- The matter distribution is proportional to the charge distribution, i.e. $A \propto Z$

- The nucleus is in good approximation a sphere, for which $R \approx R_0 \sqrt[3]{A}$ with $R_0 = 1.1$ fm

- The volume enclosed in a nucleus is proportional to the mass of such, i.e. $V_n \propto A$

§ 26.2 Nuclear Binding Energy

§§ 26.2.1 Stability and Radioactivity

Suppose having some nuclide (A, Z) with unknown mass. In general, we can say without doubt that the total mass of the nuclide will be smaller than the sum of the masses of the components.

$$M(A, Z) < Zm_p + (A - Z)m_n$$

This is immediately obvious in a relativistic context, in fact we're not yet accounting for the binding energy of the nucleons, which is for sure negative, and together with that, we're not accounting for the electronic binding energy.
With some simple calculations and noting that obviously the binding energy of the electrons can be neglected, we have that the binding energy B of such nucleon will be

$$B(A, Z) = Zm_p + (A - Z)m_n - M(A, Z) \tag{26.15}$$

The experimental determination of this value can be made via a spectrometer for some stable nucleus, and using nuclear reactions for unstable radioactive nuclei.
The shape of the binding energy function for nuclei, must include a stable region of nuclides for $A \approx 60$. This corresponds to the fact that Fe is the most stable nuclide, which is well known from astrophysical processes in stars.
In general we have that for

- $A \geq 30$, $\frac{B}{A} \approx 8$ MeV

- $A = 60$, $\frac{B}{A} \approx 8.5$ MeV

- $A > 60$, $\frac{B}{A} \to 7.5$ MeV

But,what is precisely B/A? Quantum mechanically nuclides are bound states of neutrons and protons, that behave exactly as a quantum bound state would.
The main caveat of this is that since the nuclei are quite energetic, all excited states emit high energy photons. This is the proper origin of gamma radiation.
Take some nuclide in some excited state $|E^\star\rangle$. This state is unstable and will decay towards a $|E\rangle$ ground state.
Suppose that the level corresponding to $|E^\star\rangle$ is as in the following diagram
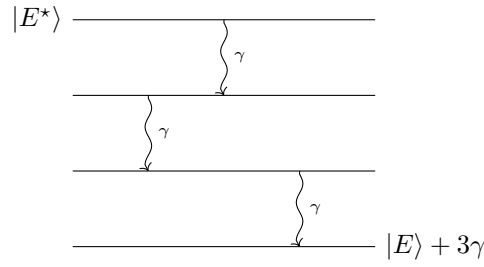
Figure 26.5: Level diagram for a gamma emission from a nuclide

The half-life of an excited nuclear state can be $10^{-17}\text{s} \lesssim \tau^\star_{1/2} \lesssim 10^2\text{y}$. In general the gamma decay has an exponential nature. Given some nuclide ^A_ZX we have

$$N\left(^A_Z\text{X}^\star\right) = N_0 e^{-\frac{t}{\tau}}$$

The half life will then be half e-folding time of the previous equation.
It's important to note that not all ground states of nuclides are stable, in fact there exist various radionuclides, or naturally radioactive elements, like ^3_1H, also known as Tritium, $^{238}_{92}\text{U}$, etc.
Due to the composition of nuclei, made with Z positive charges and $A - Z$ neutral particles, it's already obvious that electromagnetism doesn't explain their existence, therefore we must account for a new force that we now call the «strong force» .
Going back to stable and unstable nuclei that, if Z_s is the stable charge value and N is the number of neutrons we already can determine two possible radioactive decays.

1. Nuclei with $N > Z_s$. In these nuclei a neutron decays into a proton via β^- decay, with the following reaction

$$^A_Z\text{X}_N \rightarrow ^A_{Z+1}\text{Y}_{N-1} + e^- + \bar{\nu}_e \tag{26.16}$$

2. Nuclei with $N < Z_s$. In these nuclei a proton decays into a neutro via β^+ decay.

$$^A_Z\text{X}_N \rightarrow ^A_{Z-1}\text{Y}_{N+1} + e^+ + \nu_e \tag{26.17}$$

3. Another possibility is the inverse-β process, also known as electron capture, which is not a decay.

$$^A_Z\text{X}_N + e^- \rightarrow ^A_{Z-1}\text{Y}_{N-1} + \nu_e \tag{26.18}$$

Another radioactive process, together with γ and β radiation, which a nucleus can use to reach stability, is α decay. This type of decay happens usually for $A \gtrsim 180$ and is typical

for $A > 200$ nuclides. This process corresponds to the emission of a ^4_2He nucleus, as in the following reaction

$$^A_Z\text{X}_N \rightarrow \,^{A-4}_{Z-2}\text{Y}_{N-2} + \,^4_2\text{He}_2 \tag{26.19}$$

An example would be the decay of $^{238}_{92}\text{U}$ as in

$$^{238}_{92}\text{U} \rightarrow \,^{234}_{90}\text{Th} + \,^4_2\text{He}$$
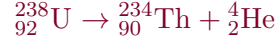
In general we have that $E_\alpha \approx 4.2$ Mev.

§§ 26.2.2 Binding Energy and Nuclear Structure

For calculating the binding energy of a nucleus we need to account for various correlation factors

- Strong interactions

- Electromagnetic interactions

- Quantum mechanical effects

The last ones are directly tied to the Heisenberg indetermination principle and Pauli's exclusion principle, due to the nucleons being $s = 1/2$ particles.
The indetermination principle makes sure that these nucleons can't be still in one point and therefore are freely moving inside a spherical potential barrier that corresponds to the nuclear boundary.
Supposing $T > 0$ we have that $E_k >> k_B T \neq 0$. This energy is deeply tied to the Fermi energy of the ensemble, and we can say that the potential barrier is

$$\mathcal{U} = \epsilon_F + \frac{B}{A}$$

For studying the effect on nucleons we directly delve into a quantum mechanical study of a spherical well.
Solving the Schrödinger equation we get, in natural units

$$\psi(x) = A\sin(px)$$

Imposing the boundary conditions at the center and border of the well we get in 3 dimensions that the permitted values of momentum for a nucleon are

$$\begin{aligned}
p_x &= \frac{\pi n_x}{L} \\
p_y &= \frac{\pi n_y}{L} \\
p_z &= \frac{\pi n_z}{L}
\end{aligned} \tag{26.20}$$

Inserting into the energy of a free particle (i.e., inside the barrier), we have

$$E_n = \frac{\pi^2}{2mL^2}n^2 \tag{26.21}$$

In order to find the Fermi energy of the system we must know how many states for a given value of momentum or energy.

Taking a spherical slice of phase space, we have that for $p \in [p + \mathrm{d}p]$ there will be $\mathrm{d}n$ states, which accounting for the spherical symmetry of the system give

$$\mathrm{d}n = \frac{1}{8} \frac{4\pi p^2 \mathrm{d}p}{\pi^3 / L^3} = \frac{V}{(2\pi)^3} 4\pi p^2 \mathrm{d}p \tag{26.22}$$

Dividing both sides by $\mathrm{d}p$ we get that

$$\frac{\mathrm{d}n}{\mathrm{d}p} \propto p^2$$

Which implies that the derivative of the particle number in the infinitesimal shell with respect to the energy is actually an energy density

$$\frac{\mathrm{d}n}{\mathrm{d}E} = \rho(E) \propto \sqrt{E}$$

Integrating this energy density we must have, since the particles are A fermions with $s = 1/2$ and $g_s = 2$

$$n = A = g \int_0^{\epsilon_F} \mathrm{d}n$$

Dividing for electrons and protons we have

$$n_p = Z = \int_0^{\epsilon_F^p} \mathrm{d}n_p$$

$$n_n = A - Z = \int_0^{\epsilon_F^n} \mathrm{d}n_n$$

For which, integrating, we have

$$Z = \frac{2V}{(2\pi)^2} \int_0^{p_F^p} 4\pi p^2 \mathrm{d}p = \frac{2V}{(2\pi)^3} \frac{4\pi p_F^3}{3} \tag{26.23}$$

Using $R = R_0 A^{1/3}$ we have

$$Z = \frac{4}{9} \pi p_F^3 A R_0^3$$

Which implies

$$p_{F,p} = \frac{1}{R_0} \sqrt[3]{\frac{9Z}{4\pi A}} \tag{26.24}$$

And for neutrons, substituting with $n_p = A - Z$

$$p_{F,n} = \frac{1}{R_0} \sqrt[3]{\frac{9(A - Z)}{4\pi A}} \tag{26.25}$$

Inserting into $E = p^2/2m$ we get that the Fermi energy of the nucleus will be

$$\epsilon_F = \frac{p_F^2}{2m} \approx 30 \text{ MeV} \tag{26.26}$$

The average kinetic energy will then be the integral of the energy density with respect to the number of particles divided by the number of particles A

$$\langle K \rangle = \frac{2}{A} \left(\int_0^{p_{Fp}} \frac{p^2}{2m_p} \mathrm{d}n_p + \int_0^{p_{Fn}} \frac{p^2}{2m_n} \mathrm{d}n_n \right) \tag{26.27}$$

Using

$$\mathrm{d}n = \frac{2V}{(2\pi)^3} 4\pi p^2 \mathrm{d}p$$

We get the following integral

$$\langle K \rangle = \frac{4\pi V}{A(2\pi)^3} \left(\frac{1}{m_p} \int_0^{p_{F,p}} p^4 \mathrm{d}p + \frac{1}{m_n} \int_0^{p_{F,n}} p^4 \mathrm{d}p \right) \tag{26.28}$$

The integral is of direct solution, giving us

$$\langle K \rangle = \frac{V}{2\pi^2 A} \left(\frac{p_{F,p}^5}{5m_p} + \frac{p_{F,n}^5}{5m_n} \right) = \frac{4R_0^3}{3\pi} \left(\frac{p_{F,p}^5}{10m_p} + \frac{p_{F,n}^5}{10m_n} \right) \tag{26.29}$$

Where note that we used $V = (4/3)\pi A R_0^3$
Substituting what we found for the Fermi momentums of neutrons and protons and using $m_n \simeq m_p = m \approx 1$ GeV we get

$$\langle K \rangle = \frac{2}{15m\pi R_0^2} \left(\frac{9}{4\pi} \right)^{\frac{5}{3}} \left[\left(\frac{Z}{A} \right)^{\frac{5}{3}} + \left(\frac{A-Z}{A} \right)^{\frac{5}{3}} \right] \tag{26.30}$$

The last equation can be expanded with power series into the following approximate result

$$\langle K \rangle \approx k \left(A + \frac{5(A-2Z)^2}{9A} \right) \tag{26.31}$$

§§ 26.2.3 Nuclear Drop Model

In order to setup all the possible correlations to the binding energy of a nucleus it's possible to rewrite the binding energy formula including everything.
All these possible contributors include

1. A volume term

2. An electromagnetic interaction term

3. A surface term

4. Semiclassical electromagnetic corrections

5. Kinetic corrections

This model of the nucleus used is known as the «liquid drop model» of the nucleus and gives us a semi-empirical formula for finding the binding energy of the nucleus.
Starting from the volume term, we have that $V \propto A$, therefore our first term will be

$$B_V = a_V A$$

For the electromagnetic interaction term, considering that $\mathcal{U}_{EM} \simeq 2^{-1}A(A-1)$ we have

$$B_C = a_C A^2$$

The third term, the surface term, considers a loss of energy on the surface of the nucleus, and noting that $S \propto A^{2/3}$ we have

$$B_S = a_S A^{\frac{2}{3}}$$

The fourth term is slightly more complex. Consider a charged classical sphere with charge Z and radius $R = R_0 A^{1/3}$. The (constant) charge density inside the volume will be

$$\rho = \frac{Ze}{V}$$

This implies that the electromagnetic energy will be

$$E_{EM} = \int \rho V(r)\mathrm{d}^3 r \propto R^5$$

Calculating properly the integral we have

$$E_{EM} = \frac{Z^2 e^2}{15} \frac{9}{(4\pi)^2 R_0 A^{\frac{1}{3}}}$$

I.e. this energetic corrections gives us $E_{EM} \propto Z^2 A^{-1/3}$, giving us our electromagnetic correction term

$$B_{EM} = -a_{EM} Z^2 A^{-\frac{1}{3}}$$

The final term comes directly from the formula (26.31), which immediately gives the following correction term

$$B_F = -a_F \frac{(A-2Z)^2}{A}$$

Adding all contributions we get the «Bethe-Weiszacker formula» , a semi-empirical formula for evaluating the nuclear binding energies of nucleons (A, Z)

$$B(A, Z) = a_V A - a_S A^{\frac{2}{3}} - a_{EM} Z^2 A^{-\frac{1}{3}} - a_F \frac{(A-2Z)^2}{A} \tag{26.32}$$

The constants are found through fitting from experimental results, and have the following approximate values.

$$a_V \approx 16 \text{ MeV}$$
$$a_S \approx 18 \text{ MeV}$$
$$a_{EM} \approx 0.7 \text{ MeV}$$
$$a_F =\approx 93 \text{ MeV}$$

The formula tho, is systematically different from the experimental results, noting that a full term accounting for spin is missing. This term will be either positive or negative depending on the values of A and Z. In general

$$\delta = \begin{cases} +\delta & A, Z, A - Z \text{ even} \\ 0 & A \text{ uneven} \\ -\delta & A \text{ even}, Z, A - Z \text{ uneven} \end{cases}$$

The principal characteristics for this formula is that for isobars it follows a parabolic path

$$B(A, Z) \to B(Z, Z^2)$$

Using the mass formula

$$M(A, Z) = Zm_p + (A - Z)m_n - B(Z, Z^2)$$

We can now calculate the maximum of the mass function, in order to find that the minimum Z with A constant for the mass is

$$Z_{min} \simeq \frac{A}{2} \frac{1}{1 + 0.0076A^{2/3}} \tag{26.33}$$

Connecting this to our previous relationship between Z and β decays we have that

- $Z < Z_{min}$ implies a $\beta^- -$active radionuclide

$$^A_Z\text{X} \to ^A_{Z+1}\text{Y} + e^- + \overline{\nu}_e$$

- $Z > Z_{min}$ implies a $\beta^+ -$active radionuclide

$$^A_Z\text{X} \to ^A_{Z-1}\text{Y} + e^+ + \nu_e$$

§ 26.3 Alpha Decay

The alpha decay of nucleus is a radioactive process that happens for $A > 200$, for which $E_\alpha \approx 5$ MeV.
The reaction is a two body decay

$$ {}^A_Z X \to {}^{A-4}_{Z-2} Y + \alpha $$
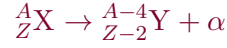
The decay half-life of an α-active nucleus is a strong function of the kinetic energy of the particle $K_\alpha \approx 5$ MeV. We have empirically that this half-life can be evaluated via the «Geiger-Nuttall law»

$$ \log \tau_{1/2} = a - b \log \sqrt{K_\alpha} \tag{26.34} $$

Since K_α can go from 4 MeV to extremely large values $\tau_{1/2}$ can also change by various orders of magnitude.
Let's evaluate the kinematics of this decay in the center of mass of ${}^A_Z X$. For the conservation of 4-momentum we get

$$ (M_X, 0) = \left(M_Y + m_\alpha + K_\alpha + K_Y, \underline{p}_\alpha + \underline{p}_Y \right) $$

From this we immediately get $p_Y = p_\alpha = 0$ and that

$$ Q = M_X - M_Y - m_\alpha = K_\alpha + K_Y = \frac{p^2}{2m_\alpha} \left(1 + \frac{m_\alpha}{M_Y} \right) $$

Noting that $M_Y >> m_\alpha$ since we're considering nuclei with $A \simeq 200$ we can approximate and then write an expression for K_α

$$ K_\alpha = \frac{Q}{1 + \frac{m_\alpha}{M_Y}} \simeq Q \left(1 - \frac{4}{A} \right) \tag{26.35} $$

Where we used $M \simeq A$
It's obvious that for this reaction to happen we need $Q > 0$. Since Q is an energy, also the binding energy has an important role in this decay. We therefore have

$$ Q = B_Y(A - 4, Z - 2) + B_\alpha(4, 2) - B_X(A, Z) \tag{26.36} $$

The constraint $Q > 0$ imposes that

$$ B_X(A, Z) < B_Y(A - 4, Z - 2) + B_\alpha(4, 2) \approx B_Y(A - 4, Z - 2) + 28 \text{ MeV} $$

From what we found we can also say that

$$ \frac{\partial B}{\partial A} < 0 \qquad \text{for } A > 60 $$

Ignoring $B_\alpha(4, 2)$ we have that an alpha decay can happen already for $A > 60$, which is not experimentally supported since it's seen only for nuclides with $A > 200$.

Using Geiger-Nuttall we have an empirical table which connects Q with the half-life of the radionuclide $\tau_{1/2}$

$$\begin{cases} A \simeq 140 & Q \simeq 0 \Longrightarrow \tau_{1/2} \to \infty \\ A \simeq 200 & Q \simeq 4 \text{ MeV} \\ A \simeq 240 & Q \simeq 8 \text{ MeV} \end{cases}$$

I.e. for $200 < A < 240$ alpha decay is possible, but with a long $\tau_{1/2}$. Spontaneous alpha decays are experimentally seen from ^{209}Bi, which is the radionuclide with the longest known half-life.

§§ 26.3.1 Quantum Tunneling and α Decay

A quantum mechanical model for alpha decay can be constructed starting from the fact that α is a strongly bound state with $B_\alpha \simeq 28$ MeV and the decaying nucleus is a heavy nuclide with $A \simeq 200$.

The decay can be thought as having the α particle bound in a potential well created by $^{A-4}_{Z-2}Y$, where

$$\mathcal{U}_Y(r) = \begin{cases} -\mathcal{U}_0 & r < R_0 A^{1/3} \\ \dfrac{2\alpha(Z-2)e^2}{r} & r > R_0 A^{1/3} \end{cases} \tag{26.37}$$

This can be graphed with an energy/distance graph as follows
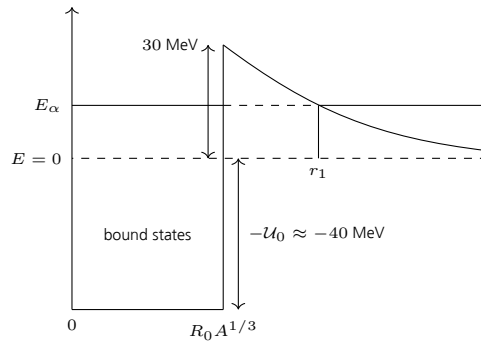
Figure 26.6: Energy level diagram of the potential well inside $^A_Z Y$ and the coulombian barrier at $r = R_0 A^{1/3}$

Approximating the coulombian barrier as a step potential from R to r_1 with height $\mathcal{U}_0$ we can solve the Schrödinger equation inside and outside the well, getting two free-particle solutions and a decaying exponential solution inside.

In the first and third zone (outside the barrier) the solution will obviously be the free particle one, and using $\hbar = c = 1$ we can immediately write the Schrödinger equation for both

$$\frac{1}{2m_\alpha} \frac{\mathrm{d}^2 \psi_{1,3}}{\mathrm{d}r^2} + E_\alpha \psi_{1,3}(r) = 0$$

Which gives the following solutions

$$\psi_1(r) = Ae^{ipr} + Be^{-ipr}$$
$$\psi_3(r) = Ee^{ipr}$$

Note that we don't have a wave traveling backwards outside the barrier, since we're interested only in the transmitted particle.

In the second region, instead we get the following Schrödinger equation

$$-\frac{1}{2m_\alpha}\frac{\mathrm{d}^2\psi_2}{\mathrm{d}r^2} + (\mathcal{U}_0 - E_\alpha)\,\psi_2(r) = 0$$

With exponential solution

$$\psi_2(r) = Ce^{ikr} + De^{-ikr}$$

Note that since in the first and third regions we have a free particle we have

$$E_\alpha = \frac{p^2}{2m_\alpha} \Longrightarrow p = \sqrt{2m_\alpha E_\alpha}$$

In the second region instead

$$E_\alpha = \mathcal{U}_0 - \frac{k^2}{2m_\alpha} \Longrightarrow k = \sqrt{2m_\alpha(\mathcal{U}_0 - E_\alpha)}$$

The boundary conditions needed in order for ψ_i to be a wavefunction are that $\psi \in L^2$ inside and outside the well. This means that the wavefunction and its first derivative must be continuous on the walls of the potential. More specifically, using $R = R_0 A^{1/3}$

$$\begin{cases} \psi_1(R) = \psi_2(R) \\ \psi_1'(R) = \psi_2'(R) \\ \psi_2(r_1) = \psi_3(r_1) \\ \psi_2'(r_1) = \psi_3'(r_1) \end{cases} \tag{26.38}$$

Shifting the potential barrier by R and writing $R - r_1 = L$ as the length of the barrier, we get by substituting the wavefunctions inside the system

$$\begin{cases} A + B = C + D \\ ip\,(A - B) = k\,(C - D) \\ Ce^{kL} + De^{-kL} = Ee^{ipL} \\ k\left(Ce^{kL} - De^{-kL}\right) = ipEe^{ipL} \end{cases} \tag{26.39}$$

The tunneling probability T will be then the norm squared of the amplitude of the outgoing particle E divided by the norm squared of the amplitude of the incoming particle A.

The solution of this system is a long and tedious algebra task, which we will now undertake.

No unfortunately the answer isn't obvious so here's the calculations.

Firstly we begin by rearranging the equations. Dividing the second row by ip and summing it to the first we get a new way of writing the continuity condition at the first wall

$$A + B = C + D$$

$$2A = \left(1 + \frac{k}{ip}\right)C + \left(1 - \frac{k}{ip}\right)D$$

We continue by taking the fourth row and dividing it by k, and then summing and subtracting it from the third row, getting a new (nice) condition for C and D

$$2Ce^{kL} = Ee^{ipL}\left(1 + \frac{ip}{k}\right)$$

$$2De^{-kL} = Ee^{-ipL}\left(1 - \frac{ip}{k}\right)$$

Dividing out the constants on these we get the new third and fourth rows of the system, leaving us with this intermediate solution

$$
\begin{cases}
A + B = C + D \\[2mm]
2A = \left(1 + \dfrac{k}{ip}\right)C + \left(1 - \dfrac{k}{ip}\right)D \\[3mm]
C = \dfrac{E}{2}e^{ipL-kL}\left(1 + \dfrac{ip}{k}\right) \\[3mm]
D = \dfrac{E}{2}e^{ipL+kL}\left(1 - \dfrac{ip}{k}\right)
\end{cases}
\tag{26.40}
$$

We continue by plugging in the second row the third and fourth of the new system, getting this algebraic monster

$$2A = \left(1 + \frac{k}{ip}\right)\left(1 + \frac{ip}{k}\right)\frac{E}{2}e^{ipL-kL} + \left(1 - \frac{k}{ip}\right)\left(1 - \frac{ip}{k}\right)\frac{E}{2}e^{ipL+kL}$$

Rearranging after some algebra, we have

$$2A = \frac{E}{2}e^{ipL}\left[\left(2 + \frac{k}{ip} + \frac{ip}{k}\right)e^{-kL} + \left(2 - \frac{k}{ip} - \frac{ip}{k}\right)e^{kL}\right]$$

Fixing the exponentials inside the square brackets we have

$$2A = \frac{E}{2}e^{ipL}\left[2\left(e^{kL} + e^{-kL}\right) + \left(\frac{k}{ip} + \frac{ip}{k}\right)e^{-kL} - \left(\frac{k}{ip} + \frac{ip}{k}\right)e^{kL}\right]$$

Working on the sums inside the parentheses we have

$$\frac{k}{ip} + \frac{ip}{k} = \frac{k^2 - p^2}{ipk} = -\frac{i\left(k^2 - p^2\right)}{pk}$$

Which, reinserted back into the equation, gives

$$2A = \frac{E}{2}e^{ipL}\left[2\left(e^{kL} + e^{-kL}\right) + \frac{i(k^2 - p^2)}{kp}\left(e^{kL} - e^{-kL}\right)\right]$$

Which in terms of hyperbolic functions is

$$2A = \frac{E}{2}e^{ipL}\left[4\cosh(kL) + 2i\frac{k^2 - p^2}{kp}\sinh(kL)\right]$$

Dividing by $2E$ both sides we have

$$\frac{A}{E} = e^{ipL}\left[\cosh(kL) + i\frac{k^2 - p^2}{2kp}\sinh(kL)\right]$$

Taking the square modulus of this we get the inverse of the tunneling probability T, which remembering how complex numbers behave is

$$T^{-1} = \left|\frac{A}{E}\right|^2 = \cosh^2(kL) + \frac{(k^2 - p^2)^2}{4k^2p^2}\sinh^2(kL)$$

Using $\cosh^2(x) = \sinh^2(x) + 1$ we can rewrite it as follows

$$T^{-1} = 1 + \left(1 + \frac{(k^2 - p^2)^2}{4k^2p^2}\right)\sinh^2(kL) \tag{26.41}$$

Continuing the calculations on the term inside the parentheses we have

$$1 + \frac{(k^2 - p^2)}{4k^2p^2} = \frac{k^4 + p^4 - 2k^2p^2 + 4k^2p^2}{4k^2p^2} = \frac{\left(k^2 + p^2\right)^2}{4k^2p^2}$$

Substituting $k = \sqrt{2m_\alpha(\mathcal{U}_0 - E_\alpha)}$ and $p = \sqrt{2m_\alpha E_\alpha}$ we have

$$\frac{(k^2 + p^2)^2}{4k^2p^2} = \frac{4m_\alpha^2\mathcal{U}_0^2}{16m_\alpha^2 E_\alpha(\mathcal{U}_0 - E_\alpha)} = \frac{\mathcal{U}_0^2}{4E_\alpha(\mathcal{U}_0 - E_\alpha)}$$

Which gives

$$T^{-1} = 1 + \frac{\mathcal{U}_0^2}{4E_\alpha(\mathcal{U}_0 - E_\alpha)}\sinh^2(kL)$$

Which gives our final tunneling probability for an alpha particle jumping a coulombian potential barrier

$$T = \frac{1}{1 + \frac{\mathcal{U}_0^2}{4E_\alpha(\mathcal{U}_0 - E_\alpha)}\sinh^2\left(L\sqrt{2m_\alpha(\mathcal{U}_0 - E_\alpha)}\right)} \tag{26.42}$$

Gamow, which was the first to propose quantum tunneling as an answer to alpha decay, continued the calculation approximating the tunnel probability for small values of kL (note that $k \simeq 430$ MeV and $L \approx 40$ fm), and using

$$\sinh^2(kL) \approx \frac{1}{4}e^{2kL}$$

We get that

$$T \approx 4\frac{\mathcal{U}_0^2 - (2E_\alpha - \mathcal{U}_0)^2}{\mathcal{U}_0}e^{-2kL}$$

Or, expressing the multiplicative constant in terms of momenta as

$$\begin{cases} \mathcal{U}_0 = \dfrac{1}{2m_\alpha}(p^2 + k^2) \\ E_\alpha = \dfrac{p^2}{2m_\alpha} \end{cases}$$

We get

$$T \approx 16\frac{k^2 p^2}{(k^2 + p^2)^2}e^{-2kL} = Ae^{-2G}$$

The constant $G = kL$ is the so called «Gamow factor». Integrating for all possible r_1 we have that the total tunnel probability will be the product of all probabilities, we have

$$T = B\prod_{i=1}^{n} T_i$$

And therefore

$$G = \int_{r_0}^{r_1} \sqrt{2m_\alpha(\mathcal{U}_0 - E_\alpha)}\,\mathrm{d}r \tag{26.43}$$

§§ 26.3.2 Radioactivity and Units

Given a radioactive material we have two main useful informations from which to base the units we need

1. Activity, i.e. the number of decays per second

2. Effects on biological tissue from α, β, γ radiation

The first can be seen in two ways: one being the actual number of decays per second, which is completely unrelated to energy, and the second the energy produced by the decay products.
For the first the most common used units are two, one being the Becquerel (Bq), which corresponds to 1 decay per second, and the second being the Curie (Ci), with the following definition

$$1\text{ Bq} = 1\text{ dec/s}$$
$$1\text{ Ci} = 37\text{ GBq} = 3.7 \cdot 10^{10}\text{ Bq}$$

The effects on biological tissues is then evaluated by the actual ionizing power of the radioactive products, in units of Coulomb per kg of ionized air. A derived unit used commonly is the Röntgen, which is defined as follows

$$1\text{ R} = 2.58 \cdot 10^{-4}\text{ C/kg}$$

This units gives a measure on the exposure to ionizing radiation.

More useful on determining possible biological effects is the absorbed dose and the equivalent dose, the first being the amount of energy per kg actually absorbed by the body, independent of the type of radiation. The two most common units are the Gray (Gy) and the Radiation Absorbed Dose (rad), which are defined as follows

$$1\ \mathrm{Gy} = 1\ \mathrm{J/kg}$$
$$1\ \mathrm{rad} = 0.01\ \mathrm{Gy}$$

Including a radiation weighting factor (W_R) to these two units we can now distinguish between the various radiation kinds and the potential damage caused by ionizing radiation, the two derived units are the Sievert (Sv) and the Röntgen Equivalent Man (rem), which are by definition

$$1\ \mathrm{Sv} = W_R \cdot 1\ \mathrm{Gy}$$
$$1\ \mathrm{rem} = W_R \cdot 1\ \mathrm{rad}$$

All these units can be then multiplied by a second factor W_T which weights the radiation dose per each different tissue.

As an example, note that $W_R = 20$ for alpha particles, therefore 1 Gy of exposure to alpha radiation corresponds to a weighted equivalent dose of 20 Sv In order to have a deeper understanding of the dosage units it's useful to check the lethal dose of radiation.

The lethal dose is defined as a radiation dose (expressed in Sieverts) expected to cause death in 50% of an exposed population within 30 days if received within 30 days, and it's denoted as $LD_{50/30}$. Such dose is in the following range of

$$LD_{50/30} = 400 \to 450\ \mathrm{rem} = 4 \to 5\ \mathrm{Sv}$$

Note that a yearly dose of natural background radiation corresponds to around 2.4 mSv, which is roughly $0.27\ \mu\mathrm{Sv/h}$

§ 26.4 Nuclear Reactions

Let a, b be two nuclides and c, d the resulting nuclides formed from a reaction. In general a nuclear reaction takes the following shape

$$a + b \to c + d + Q \tag{26.44}$$

Q is known as the «Q-value» of a reaction and corresponds to the energy absorbed from the environment or emitted into the environment after the reaction takes place, and for the previous general reaction evaluates to

$$Q = m_a + m_b - m_c - m_d$$

More generally, if there are R reagents with masses m_i and N resulting particles with masses f_i, we have that

$$Q = \sum_{i=1}^{R} m_i - \sum_{i=1}^{N} f_i \tag{26.45}$$

All reactions can be divided into two subcategories depending on the sign of Q.

1. Exothermic reactions, with $Q > 0$

2. Endothermic reactions, with $Q < 0$

As we have demonstrated before, a decay process is necessarily exothermic, since $Q > 0$ corresponds to the threshold energy for such reaction to happen.

Considering an exothermic reaction, henceforth $Q > 0$, we have that the masses of the reagent particles partially convert to kinetic energy of the resulting particles, whereas if the reaction is endothermic ($Q < 0$), the kinetic energy of the reacting particles converts to mass for the resulting particles.

Consider now a really famous endothermic reaction

$$\alpha + {}^{14}_{7}\text{N} \rightarrow {}^{17}_{8}\text{O} + p$$

This reaction has $Q = -1.19$ MeV and needs $K_\alpha \gtrsim 5$ MeV in order to take place, which is the energy needed for the α particle to overcome the Coulomb barrier. This is the main reason for the need of accelerating particles.

In order to accelerate charged particles it's possible to use electric fields, but the main problem comes with neutral particles like neutrons.

A base reaction we can take as an example is the following

$$p + {}^{3}_{7}\text{Li} \rightarrow \alpha + \alpha$$
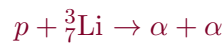
In this reaction $Q = 17.5$ MeV, which implies that the reaction is exothermic and therefore the kinetic energy of the proton is transferred into the kinetic energy of the alpha particles. This corresponds to an indirect acceleration of α.

For neutrons one such reaction is the one used by Chadwick in 1932

$$\alpha + {}^{9}_{4}\text{Be} \rightarrow {}^{12}_{6}\text{C} + n$$
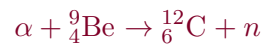
Here $Q = 5.71$ MeV and both the Beryllium and Carbon isotopes are produced at rest, which implies an indirect acceleration of the neutron. This discovery led the beginning of the studies of nuclear reactions with neutrons, which don't interact with the Coulomb field. The main results obtained in the early days of neutron physics were that the reaction

$$n + {}^{A}_{Z}\text{X}$$

Had three main results: elastic scattering, neutron capture or an induced fission of the nuclide X

§§ 26.4.1 Uranium Fission and Nuclear Power Plants

The most known result of neutron physics was the discovery of nuclear fission, and in particular the fission of Uranium.

The first observation of Uranium fission happened in 1938, where Hahn and Strassmann observed the following reaction

$$n + {}^{238}_{92}\text{U} \rightarrow {}^{137}_{56}\text{Ba} + {}^{84}_{36}\text{Kr} + 17n$$

Since heavier nuclei have more neutrons than needed for stability, the end result of this reaction is a net release of neutrons.

The next year Meitner and Frisch continued the study on the more general decay

$$^A_Z\text{N} \rightarrow {}^{A-A_2}_{Z-Z_2}\text{X} + {}^{A_2}_{Z_2}\text{Y}$$
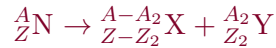
The main question remained on why this reaction isn't spontaneous, since for $A > 60$, $\partial_A \overline{B} < 0$ which should favor a spontaneous fission.

Considering the case of Uranium fission we have that $Q \approx 210$ MeV, but still spontaneous fission reactions are observed mostly for $A \geq 300$.

A proposed explanation for the fission of uranium was nuclear deformation, where the mostly spherical nucleon gets deformed into an ellipsoid containing the two child products. The final volume of the deformed nucleus is

$$V_d = \frac{4\pi}{3}ab^2$$

Where $a = R(1 + \epsilon)$ is the radial deformation. This gives $ab^2 = R^3 \Longrightarrow b = R(1 + \epsilon)^{-1}$.

Reevaluating the Bethe-Weiszäcker formula for the binding energy of the nucleon with this we get that the variation of the binding energy is

$$\Delta B = -a_S A^{2/3}\left(\frac{2}{5}\epsilon^2\right) - a_C Z^2 A^{1/3}\left(-\frac{1}{5}\epsilon^2\right)$$

In order to then have a fission of the nucleus it's necessary to insert in the system an energy $E_f?\Delta B$ in order to let the system overcome this energy barrier.

Imposing $\Delta B \approx 0$ we get that

$$\frac{Z^2}{A} \approx 50$$

Suggesting that the fission can be induced with really little energy, or it can also happen through quantum tunneling across the barrier.

Considering again the $Q-$value for a fission reaction, we generally have

1. $A \approx 300$, $Q - E > 0$ and the fission is spontaneous

2. $A \approx 240$, $E - Q \approx 6$ MeV, the fission can happen with little energy or through quantum tunneling

3. $A \approx 100$, $E - Q \approx 60$ MeV, the fission reaction doesn't occur.

For Uranium in particular we have $A = 238$ (99.3% in nature, $A = 235$ 0.7% in nature) and therefore the fission reaction can happen both through quantum tunneling and via low

energy neutron bombardment.

Considering a neutron bombardment of $^{235}_{92}$U we have the following self sustaining chain reaction

$$n + {}^{235}_{92}\text{U} \rightarrow {}^{236}_{92}\text{U}^\star \rightarrow {}^{139}_{56}\text{Ba} + {}^{94}_{32}\text{Kr} + 3n$$

The neutron absorption reaction has $Q = 6.5$ MeV and $E = 6.2$ MeV, with cross section $\sigma_{235\text{U},n} \approx 580$ b, for which $K_n = 0.025$ eV are needed. This makes ^{235}U an easily fissile material. Considering the more abundant ^{238}U we get $Q_{238} = 4.8$ MeV, $E = 6.6$ MeV, and therefore $K_n = 1.8$ MeV is needed to make the reaction happen.

In general, for a complete fission reaction of an uranium nucleus we have $Q = 200$ MeV. Considering $m_p \approx m_n \approx 1$ GeV we have that $M_\text{U} \approx 240$ GeV, and in this reaction we have an efficiency of energy released

$$\eta = \frac{Q}{M_\text{U}} = \frac{200 \text{ MeV}}{240 \text{ GeV}} \approx 10^{-3}$$

This corresponds that the possible energetic output from the fission of 1 g of uranium is

$$Q = 200 \text{ MeV} \cdot \left(\frac{1 \text{ g}}{A} N_A\right) \approx 5 \cdot 10^{23} \text{ MeV} \approx 10^{11} \text{ J}$$

Note that this corresponds to 3 times the energy output from the combustion of 1 ton of carbon.

This high energy potential output gives the possibility of energy production using nuclear reactors with fissile fuels.

In modern nuclear reactors the chain reaction is stabilized using either heavy or light water, which acts as a neutron moderator (slows the fast neutron from the fission in order to thermalize them and let them sustain a chain reaction).

This is useful since the neutron-proton and the neutron-neutron cross section is quite high, implying that the loss of energy is major.

§ 26.5 Nuclear Fusion

§§ 26.5.1 Quantum Tunneling and Fusion

For lighter nuclei ($A = a < 60$) the opposite of the fission reaction is possible, i.e. a «nuclear fusion» . The reaction in question has the following shape

$$^a_{Z_1}\text{X} + {}^b_{Z_2}\text{Y} \rightarrow {}^{a+b}_{Z_1+Z_2}\text{N} + Q \tag{26.46}$$

The main idea for having a successful fusion reaction is that a coulomb barrier with maximum as $R = R_X + R_Y$ must be overcame, i.e.

$$\mathcal{U}_{max}(R) = \frac{\alpha Z_1 Z_2}{R_X + R_Y}$$

Considering a $p + p$ fusion reaction we have that $\mathcal{U}_{max} \propto \frac{\alpha}{2R_0} = 550$ keV, which is our minimal kinetic energy needed by one proton to jump the coulomb barrier of the other

proton and complete the reaction.

Take now a second reaction with $A > 1$ nuclei.

$$^{12}_{6}\text{C} + ^{12}_{6}\text{C} \rightarrow ^{24}_{12}\text{Mg}$$

Here $Q = 13.9 \text{ MeV}$. The barrier evaluates to 9.3 MeV, which gets us to reason if it's really a convenient fusion reaction when the net gain of energy is 4.3 MeV. The efficiency of this kind of reaction (for this one is $\eta \propto 10^{-4}$) is proportional to Q and inversely proportional to the mass of the sum of the fusing nuclei, which lets us hope that Q decreases slowly when mass decreases.

Fortunately this is mostly the case and we can talk about nucleosynthesis in stars.

§§ 26.5.2 Stellar Fusion and Nucleosynthesis

In stars of the main sequence the most common fusion reaction observed is the $p - p$ chain reaction, where proton are fused creating Helium and energy which keeps the star from collapsing.

The chain reaction is summarized as follows

$$p + p \rightarrow ^{2}_{1}\text{H} + e^{+} + \nu_e + 0.42 \text{ MeV}$$
$$p + ^{2}_{1}\text{H} \rightarrow ^{3}_{2}\text{He} + \gamma + 5.49 \text{ MeV} \tag{26.47}$$
$$^{3}_{2}\text{He} + ^{3}_{2}\text{He} \rightarrow ^{4}_{2}\text{He} + p + p + 12.86 \text{ MeV}$$
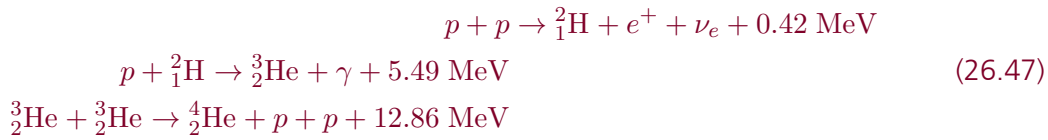
The last reaction continues the cycle and produces a stable Helium nucleus together with around 13 MeV.

The total reaction is

$$4p \rightarrow ^{4}_{2}\text{He} + 2e^{+} + 2\nu_e + 2\gamma + 24.7 \text{ MeV}$$
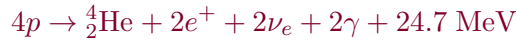
Considering he total Q-value of this reaction is $Q \approx 24 \text{ MeV}$ since we need to consider the extremely weak interaction of neutrinos, which basically amounts to a loss energy of $2m_\nu$. It's important to note that stellar cores are rich in atoms, therefore the total $Q \approx 26 \text{ MeV}$ since it's necessary to consider electron-positron interactions.

All of this considered, we have

$$\eta_{pp} = \frac{Q}{4m_p} \approx 6.5 \cdot 10^{-3}$$

Making this an efficient fusion reaction, which explains why it's so common in stars.

Since this reaction creates alpha particles, when the proton fuel for the reaction ends and the density of helium is enough, the pp chain reaction ends, causing a contraction in the star and an increase of temperature which permits a new chain fusion, the α chain fusion process.

This process bases itself on the following single reaction

$$\alpha + \alpha \rightarrow ^{8}_{4}\text{Be} \rightarrow \alpha + \alpha$$

The first reaction has $Q \approx -100 \text{ keV}$, and the nuclide of ^8Be is highly unstable with half life $\tau_{1/2} \approx 8 \cdot 10^{-17}s$. This nuclide is fundamental for the reaction, since we have the following chain emerging from the $\alpha-$Beryllium fusion

$$\begin{aligned} \alpha + {}^8_4\text{Be} &\to {}^{12}_6\text{C} + \gamma \\ \alpha + {}^{12}_6\text{C} &\to {}^{16}_8\text{O} + \gamma \\ \alpha + {}^{16}_8\text{O} &\to {}^{20}_{10}\text{Ne} + \gamma \\ \alpha + {}^{20}_{10}\text{Ne} &\to {}^{24}_{12}\text{Mg} + \gamma \end{aligned} \tag{26.48}$$

The next tightly bound nucleus is $^{12}_6\text{C}$ and it's the next nucleus to be fused inside the stellar core.

This process continues up until $A = 60$ with ^{60}Fe. The iron-iron fusion is endothermic and therefore needs energy to proceed, causing a stellar core-collapse.

Heavier nuclides are created then with supernovas and neutron capture

§ 26.6 Beta Decay

Beta decay is a kind of decay that at first was thought to be a two-body decay with the emission of an electron in the following reaction

$$^A_Z\text{X} \to {}^A_{Z+1}\text{Y} + e^-$$

Plotting the experimental results for decays per energy, we clearly see that it's not compatible with a 2-body decay.

A quick relativistic calculation gives in the center of mass of X

$$\sqrt{s} = M_x = E_Y + E_e \qquad p_Y = p_e = p$$

Substituting $E = \sqrt{m^2 + p^2}$ we have

$$M_x = \sqrt{M_Y^2 + p^2} + \sqrt{m_e^2 + p^2}$$

And solving for p^2

$$p^2 = \frac{\left(M_X^2 + M_Y^2 - m_e^2\right)^2}{4M_X^2} - M_Y^2$$

Using the approximation $M_X \approx M_Y \approx Am_p >> m_e$ we have $p^2 \approx 0$ and therefore $E_Y \approx M_Y + K_Y$ and

$$E_X = M_Y + K_Y + E_e$$

And therefore $E_e \approx M_X - M_Y$ which implies that energy isn't conserved in this process. Another way of evaluating that this reaction is not possible is by checking the angular momentum of the particles in this reaction. We have

$$\hbar L_X = \hbar L_Y + \frac{\hbar}{2}$$

Here angular momentum is obviously not conserved.

An idea for solving this problem was adding a new child particle, firstly it was though that it was a photon, but since $S_\gamma = \hbar$ this is not the case.

Pauli in 1930 supposed the existence of a new fermion ($S = \hbar/2$) with a small mass and a neutral charge in order to balance the reaction.

Fermi proceeded to the theorization of the interaction of this new particle, called the neutrino ν for its properties.

In Fermi's theory, beta decay then became a three body process with the following reaction

$$\ _{Z}^{A}\text{X} \to\ _{Z+1}^{A}\text{Y} + e^- + \overline{\nu} \tag{26.49}$$

This kind of interaction, due to its weak nature was called the «weak interaction» , where firstly the following kind of reactions are studied using Fermi's theory and also explain the process of electron capture.

§§ 26.6.1 Unstable States and Decay Rates

Consider a group of particles N_0 situated in an unstable state. As we already know, after a time t an exponential number of particles will have decayed, giving

$$N(t) = N_0 e^{-\frac{t}{\tau}} = N_0 e^{-\Gamma t}$$

The value $\Gamma = 1/\tau$ is known as the decay rate of the unstable state.

Suppose that this state is an exponential solution of the Schrödinger equation with imaginary energy $E_0 + i\Gamma$ where E_0 is the unperturbed state.

The evolved state will then be

$$\psi(\underline{r}, t) = \psi_0(\underline{r}) e^{-iE_n t} e^{-\frac{\Gamma}{2} t}$$

Where ψ_0 is the solution to the TISE unperturbed Hamiltonian.

We're interested to see the behavior of the wavefunction in the energy space, and therefore we perform an inverse Fourier transform of the following quantity

$$e^{-i\left(E_0 + i\frac{\Gamma}{2}\right)t} = \int_{\mathbb{R}} \varphi(E) e^{-iEt} \mathrm{d}E$$

Where $\varphi(E)$ is the searched wavefunction in the energy space. Antitrasforming we have

$$\varphi(E) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i(E-E_0)t + \frac{\Gamma}{2} t} \mathrm{d}t \tag{26.50}$$

Considering that for $t < 0$ the unstable state doesn't exist and modifying accordingly the integration bounds we get

$$\varphi(E) = \frac{1}{2\pi} \left(\frac{1}{\frac{\Gamma}{2} - i(E - E_0)} \right)$$

The probability density function for this energy wavefunction will be a Breit-Wigner Lorentzian distribution

$$|\varphi(E)|^2 = \frac{1}{4\pi^2} \frac{1}{\frac{\Gamma^2}{4} + (E - E_0)^2}$$

Which will draw a bell curve with width Γ. This property gives the decay rate Γ its second name, the «decay width» .

Note that for $\Gamma/m \ll 1$ the particle is clearly stable, since for the indetermination principle $\Gamma\tau = \hbar$.

Suppose now having an unstable particle in its center of mass, which for $t > 0$ decays into two child particles
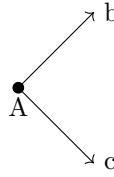
Figure 26.7: Example drawing of the $A \to b + c$ two body decay

For this decay we have

$$\sqrt{s} = M_A = \sqrt{(p^\mu p_\mu)_a + (p^\mu p_\mu)_c}$$

This $\sqrt{s}$ is well determined since for $\Gamma \to 0$ $|\varphi|^2 = \delta(E - m)$.

We continue by writing the unstable state as a weak perturbation $\hat{\mathcal{H}}_I$ on a Hamiltonian $\hat{\mathcal{H}}_0$ for which the TISE holds for the unperturbed state $|n\rangle$

$$\hat{\mathcal{H}}_0 |n\rangle = E_n |n\rangle$$

We suppose that $|n\rangle$ is an orthonormal eigenvector basis for this Hamiltonian, and we begin evaluating the new perturbed state using undetermined coefficients which depend on time, after evolving the unperturbed state

$$\hat{\mathcal{H}} |\psi\rangle = \left(\hat{\mathcal{H}}_0 + \hat{\mathcal{H}}_I\right) |\psi\rangle = \hat{E} |\psi\rangle$$

$$|\psi\rangle = \sum_n a_n(t)\hat{\mathcal{U}} |n\rangle$$

We proceed by inserting the new state inside a TDSE, using $\hat{E} = i\partial_t$, getting

$$\hat{E} |\psi\rangle = i \sum_n \dot{a}_n \hat{\mathcal{U}} |n\rangle + \sum_n a_n(t) E_n \hat{\mathcal{U}} |n\rangle = \hat{\mathcal{H}} |\psi\rangle$$

Using the first result on the unperturbed Hamiltonian and $\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{\mathcal{H}}_I$ we have immediately confronting the two results that

$$i \sum_n \dot{a}_n(t)\hat{\mathcal{U}} |n\rangle = \hat{\mathcal{H}}_I |\psi\rangle$$

Using the orthonormality of $|n\rangle$ ($\langle k|n\rangle = \delta_{nk}$), we have by multiplying both sides by $\langle k|$ that

$$i\sum_n \dot{a}_n(t)\hat{\mathcal{U}}\langle k|n\rangle = \langle k|\hat{\mathcal{H}}_I|\psi\rangle = \sum_n a_n(t)\langle k|\hat{\mathcal{H}}_I\hat{\mathcal{U}}|n\rangle$$

Rearranging things and evaluating the sums, and defining $\langle n|\hat{\mathcal{H}}_I|n\rangle = \hat{V}_{nk}$ we have

$$i\dot{a}_n(t)e^{-iE_k t} = \sum_n a_n(t)\hat{V}_{nk}e^{-iE_n t}$$

We define for convenience $\hat{\mathcal{M}} = -i\hat{V}$, which gives us the final differential equation for the variation

$$\frac{da_k}{dt} = \hat{\mathcal{M}}_{kn}e^{-i(E_n - E_k)t} \tag{26.51}$$

Where

$$\hat{\mathcal{M}}_{kn} = \langle k|\hat{\mathcal{M}}|n\rangle = -i\langle k|\hat{V}|n\rangle = -i\langle k|\hat{\mathcal{H}}_I|n\rangle \tag{26.52}$$

Considering that the perturbation is small we can say without problems that $a_n(0) = \delta_{mk}$ for some initial state m, and considering a weak time dependence, we can also say that $a_m(t) \approx 1$
With these hypotheses $\hat{\mathcal{M}}_{mk}$ can be considered weakly dependent on time (adiabatic) and the integration will give

$$a_k(t) = \hat{\mathcal{M}}_{mk}\int_0^T e^{i(E_k - E_m)t}dt$$

Taking k as our final state and m our initial state the transition probability per unit time will be

$$P_{i\to f} = \lim_{T\to\infty}\frac{|a_f(t)|^2}{T} \tag{26.53}$$

Which is

$$\lim_{T\to\infty}\frac{\left|\hat{\mathcal{M}}_{mk}\right|^2}{T}\int_0^T e^{i(E_f - E_i)t}dt\int_0^T e^{-i(E_f - E_i)\tau}d\tau$$

Executing a change of variables $T \to T - T/2$ the integrals on the right become easily identifiable as the Fourier transforms of a complex exponential, which gives $\hat{f}(x) = 2\pi\delta(x)$, and therefore

$$P_{i\to f} = \lim_{T\to\infty}\frac{\left|\hat{\mathcal{M}}_{mk}\right|^2}{T}2\pi T\delta(E_f - E_i)$$

Which finally gives «Fermi's golden rule» for time dependent perturbations

$$P_{i\to f} = 2\pi\left|\hat{\mathcal{M}}_{mk}\right|^2\delta(E_f - E_i) \tag{26.54}$$

Since there are various possibilities for this two body decay, it's necessary to then consider a statistical approach for the energies, and we need to evaluate the density of states in the

phase space, for the volume $E_f - E_i$. In this case, considering then all possible decays, we have finally the transition rate between the initial and the final state

$$\Gamma_{fi} = \int P_{i \to f} \mathrm{d}n = \int P_{i \to f} \frac{\mathrm{d}n}{\mathrm{d}E_f} \mathrm{d}E_f = 2\pi \left| \hat{\mathcal{M}}_{fi} \right|^2 \int \delta(E_f - E_i) \frac{\mathrm{d}n}{\mathrm{d}E_f} \mathrm{d}E_f$$

Integrating the delta on the right we have finally

$$\Gamma = 2\pi \left| \hat{\mathcal{M}}_{fi} \right|^2 \rho(E_i) \tag{26.55}$$

Where $\hat{\mathcal{M}}_{fi}$ depends on the interaction Hamiltonian, whereas $\rho(E_i)$ is the density of states in the phase space for the two body decay.

§§ 26.6.2 Beta Decay Rate

Using what we found before for unstable states we can immediately imagine to apply it to Fermi's theory of β decay. In this case we have a reaction of the type

$$n \to p + e^- + \overline{\nu}$$

Here, we have $|i\rangle = |n\rangle$ and $|f\rangle = |pe^-\overline{\nu}\rangle$. The decay rate as for Fermi's golden rule is

$$\Gamma = 2\pi \left| \hat{\mathcal{M}}_{fi} \right|^2 \rho(E)$$

Where, in this case

$$\hat{\mathcal{M}}_{fi} = -i \int \overline{\psi}_p \overline{\psi}_e \overline{\psi}_\nu G_F \psi_n \mathrm{d}^3 r \tag{26.56}$$

Where G_F is our interaction parameter.
Suppose to normalize the wavefunctions on a nuclear volume V, $\psi \propto V^{-1/2}$. With this normalization we also have that $[G_F] = E^{-2}$.
Considering the total decay the wavefunctions for $e^-, \overline{\nu}$ can be considered without loss of generality as plane waves due to their non-interacting nature after the decay, therefore

$$\psi_e = \frac{1}{\sqrt{V}} e^{-i\underline{p}\underline{r}}$$
$$\psi_\nu = \frac{1}{\sqrt{V}} e^{-i\underline{q}\underline{r}} \tag{26.57}$$

Where we chose $\underline{p}_e = \underline{p}$ and $\underline{p}_\nu = \underline{q}$.
Approximating the interaction factor G_F as a constant as for Fermi's theory we have finally

$$\hat{\mathcal{M}}_{fi} = -i\frac{G_F}{V} \int \overline{\psi}_p \psi_n e^{i(\underline{p}+\underline{q})\underline{r}} \mathrm{d}^3 r$$

Another approximation can be made by checking that $Q = m_n - m_p - m_e - m_\nu \leq 1$ and therefore $p \simeq q \leq 1$ MeV, and $r \approx 1$ fm, which gives $(p+q)r \approx 5/1000$.

Considering this the exponential in this integral can be expanded with a power series and approximated to the first order, giving

$$\hat{\mathcal{M}}_{fi} = -i\frac{G_F}{V}\int \overline{\psi}_p \psi_n \mathrm{d}^3 r = -\frac{iG_F N}{V} \tag{26.58}$$

Where N is the so called «nuclear term» $N = \langle p|n \rangle$.
Reinserting it into the decay rate equation we have

$$\Gamma = \frac{2\pi G_F^2 |N|^2}{V^2}\rho(E)$$

We only need to find $\rho(E)$ in order to complete the calculations. Considering that for the conservation of 3-momentum $\underline{p}_e + \underline{p}_\nu + \underline{p}_X = 0$, and for the conservation of energy

$$\rho(E) = \left(\frac{\mathrm{d}n}{\mathrm{d}E_f}\right)_{E_i} = \int \delta(E_f - E_i)\mathrm{d}n$$

$$\mathrm{d}n = \frac{V^2}{(2\pi)^6}\mathrm{d}^3 p \mathrm{d}^3 q \tag{26.59}$$

In the center of mass of X we have then

$$E_i = \sqrt{s} = M_X, \quad E_f = E_Y + E_e + E_\nu = M_Y + K_Y + E_e + E_\nu$$

Considering that the masses of X and Y are much greater than the mass of the electron and neutrino we have that $K_Y \approx 0$ and we have $M_X \approx M_Y + E_e + E_\nu$. The total available energy for the reaction will then be

$$E_T = M_X - M_Y = E_e + E_\nu$$

And, therefore

$$Q = M_X - M_Y - m_e - m_\nu \approx E_T$$

In the limit case where X $= p$ and Y $= n$ we have $E_T \approx 1$ MeV, and

$$\delta(E_f - E_i) = \delta(E_T - E_e - E_\nu)$$

Approximating $m_\nu = 0$ we have $E_\nu = q$ and therefore $q^2 \mathrm{d}q = E_\nu^2 \mathrm{d}E_\nu$. For the electron we have instead

$$p = \sqrt{E_e^2 - m_e^2} \qquad p^2 \mathrm{d}p = pE_e \mathrm{d}E_e$$

Substituting in the delta integral we have

$$(4\pi)^2 \iint \delta(E_f - E_i)p^2 q^2 \mathrm{d}p\mathrm{d}q = (4\pi)^2 \iint \delta(E_T - E_e - E_\nu)E_\nu^2 E_e p\,\mathrm{d}E_\nu \mathrm{d}E_e = (4\pi)^2 \int_0^{E_T} pE_e(E_T - E_e)^2 \mathrm{d}E_e$$

Substituting p we get finally for the decay rate

$$\Gamma_\beta = \frac{G_F^2 |N|^2}{2\pi^3}\int_0^{E_T} E_e\sqrt{E_e^2 - m_e^2}(E_T - E_e)^2 \mathrm{d}E_e$$

For the limit case of a hyperrelativistic electron we can say with ease that $m_e << E_e$ and therefroe we can approximate the integral as follows

$$\int_0^{E_T} E_e^2 (E_T - E_e)^2 \mathrm{d}E_e = \frac{E_T^5}{3} + \frac{E_T^5}{5} - \frac{E_T^5}{2}$$

Summing up and inserting into the decay rate we get «Sargent's rule» for beta decay with highly energetic electrons

$$\Gamma = \frac{G_F^2 |N|^2}{60\pi} E_T^5 \tag{26.60}$$

Since here $E_T \approx M_X - M_Y$ we have that the phase space must grow with $M_X - M_Y$

§§ 26.6.3 Experimental Estimate of m_ν

So far the hypotheses we managed to stack up from beta decay are

1. $m_\nu \neq 0$

2. $\Gamma \propto p E_e (E_T - E_e)^2$, $E_T = M_X - M_Y - m_\nu - m_e = E_e + E_\nu$

It's clear that reducing E_T we have the maximum possible E_e. We define here the Kurie function $K(E)$ as

$$K(E) = \sqrt{\frac{1}{pE_e} \frac{\mathrm{d}\Gamma}{\mathrm{d}E}} \propto (E_T - E_e)^2$$

Evaluating this function for a minimal E_T it's possible to resolve $E_T - m_\nu$ and evaluate the neutrino mass.
Note that $K(E) \propto (E_T - E_e)$ and therefore we expect a linear decay of this function up until E_e^{max}. Experimentally it has been observed that at $\min(E_T)$ instead the function evaluates to $E_T - m_\nu$ effectively giving an estimate for the mass of the neutrino

§§ 26.6.4 Cross Section for Beta Decays

The reaction considered up until now

$$n \rightarrow p + e^- + \overline{\nu}$$

Can also be reversed, indicating that another weak reaction is possible

$$\overline{\nu} + p \rightarrow n + e^+ \tag{26.61}$$

This scattering reaction is experimentally observed, but we must

- generate neutrinos

- choose a p-rich target

- observe n, e^+ and count the number of events

- measure the scattering cross section σ

The first question we get is: Can Fermi's 4-fermion theory evaluate theoretically σ? Starting again from cross sections we have, by definition that the number of observed reactions per unit time is

$$\frac{\mathrm{d}N_R}{\mathrm{d}t} = \sigma n_T \frac{\mathrm{d}N_p}{\mathrm{d}t}\mathrm{d}x$$

Where N_p is the number of incoming projectiles, and n_T is the particle density of the target. Evaluating everything we have

$$\frac{\mathrm{d}N_R}{\mathrm{d}t} = \frac{\sigma}{S} \underbrace{n_T S \mathrm{d}x}_{N_T}$$

Also using $S^{-1}\mathrm{d}N_p/\mathrm{d}t = \phi_p$, i.e. the flux of incoming projectiles, we have

$$\frac{1}{N_T}\frac{\mathrm{d}N_R}{\mathrm{d}t} = \sigma\phi_p = \sigma v_p n_p = \sigma\frac{N_p}{V}$$

Where n_p, N_p are respectively the number density and total number of particles of the projectile beam, and v_p is the relative velocity of the projectiles with respect to the target. Fixing the previous equation we have

$$\frac{1}{N_T N_p}\frac{\mathrm{d}N_R}{\mathrm{d}t} = \sigma\frac{v_p}{V}$$

Using $v_p = p_p/E_p$ and noting that on the left we have the number of reaction per unit time per single projectile on a single target, we have

$$\frac{1}{N_T N_p}\frac{\mathrm{d}N_R}{\mathrm{d}t} = \sigma\frac{p_p}{V E_p} = \Gamma(a + b \to c + d) = 2\pi\left|\hat{\mathcal{M}}_{fi}\right|^2\rho(E)$$

Therefore we have

$$\sigma(a + b \to c + d) = 2\pi\left|\hat{\mathcal{M}}_{fi}\right|^2\frac{V}{v_p}\rho(E) \tag{26.62}$$

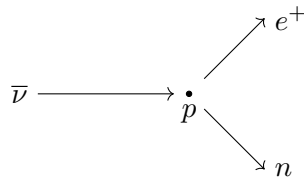For our reaction we have the following scattering reaction

Figure 26.8: Scattering process in study

Moving to the center of mass of the initial $\bar{\nu} + p$ state we have

$$\beta\gamma = \frac{p}{\sqrt{s}} = \frac{E_\nu}{\sqrt{m_p^2 + 2E_p m_\nu}}$$

Where we approximated $m_\nu = 0$ therefore $p_\nu = E_\nu$.

The velocity of the ν with respect to the p targets is therefore

$$v_i = \frac{p_\nu^\star}{E_\nu^\star} + \frac{p_p^\star}{E_p^\star}$$

Inserting it all into the cross-section equation we get

$$\sigma = \frac{2\pi G_F^2 |N|^2}{V} \frac{E_\nu^\star E_p^\star}{p_\nu^\star E_p^\star + p_p^\star E_\nu^\star} \rho(E_f)$$

In order to calculate the density of states in the final decayed state we need to consider the conservation of 3-momentum, which implies $p_n^\star = p_e^\star = p^\star$, and accounting that

$$\rho(E_f) = \frac{\mathrm{d}n}{\mathrm{d}E_f} = \frac{V}{2\pi^2} p^2 \frac{\mathrm{d}p}{\mathrm{d}E_f}$$

Using $E_f = \sqrt{p^2 + m_e^2} + \sqrt{p^2 + m_n^2}$ gives

$$\frac{\mathrm{d}E_f}{\mathrm{d}p^\star} = \frac{p(E_n^\star + E_e^\star)}{E_n^\star E_e^\star} = v_f$$

Where v_f is the velocity of the positrons with respect to the neutrons. Inverting and inserting it into the cross-section equation, using $E_f = E_n^\star + E_e^\star$ we have

$$\sigma = \frac{G_F^2 |N|^2 (p^\star)^2}{\pi v_i v_f} = \frac{G_F^2 |N|^2}{\pi} \frac{E_p^\star E_\nu^\star E_p^\star E_e^\star}{E_f \left(p_\nu^\star E_p^\star + p_p^\star E_p^\star\right)} p^\star \tag{26.63}$$

Where $p^\star = p_e^\star = p_\nu^\star$ is the momentum of the final two decay products, the positron and the neutron.

Inserting the experimental values of $G_F = 1.17 \cdot 10^{-5} \text{ GeV}^{-2}$, $v_i \approx v_f \approx 1$, we get

$$\sigma(p^\star)^2 \approx 10^{-37} \text{ cm}^2/\text{GeV}^2$$

Considering that the $Q-$value for this reaction is $Q \approx 1.8 \text{ MeV}$ we have that this process is possible only if $E_\nu > 1.8 \text{ MeV}$ which gives, for the cross-section at threshold level

$$\sigma(p^\star)^2 \approx 10^{-43} \text{ cm}^2/\text{GeV}^2$$

In order to evaluate what it means for experimental tests of this decay we need to account for the mean free path of the reaction. Supposing a light water target, for which $A_{\text{H}_2\text{O}} = 18 \text{ g/mol}$ we have $n_T = A^{-1} N_A$, for which we get

$$\frac{1}{\lambda} = \sigma n_T = E \cdot 10^{-43} \cdot 3 \cdot 10^{22} \text{ MeV}^{-1}\text{cm}^{-1}$$

Using $E \approx 1 \text{ MeV}$ we get an estimate of $\lambda \approx 10^{19} m$ which is an absurdly large mean free path for such reaction.

What is needed therefore for building experiments on this reaction is to note that accelerated neutrinos will have a higher E_ν, and therefore will contribute to lower this. A second fix is the usage of high density targets together with heavy neutrinos emitters.

One way of accomplish this is considering Uranium fission. In 1956 Reines, Cowen et al. were the first to build a nuclear reactor for the production of energy, with a thermal power output of $1000 \text{ MW} = 10^9 \text{ J/s} = 6 \cdot 10^{27} \text{ eV/s}$.

Considering $Q_{fis} = 200 \text{ MeV}$ and a production of around $N_\nu = 6$ neutrinos per reaction we have

$$N_R = \frac{P_{reactor}}{Q} = 3 \cdot 10^{19} \text{ Hz}$$

Which implies an average production of 10^{20} Hz neutrinos, with $E_\nu \approx 3 \text{ MeV}$.

Not too far from the reactor core we have that the flux of neutrinos per solid angle of detector is

$$\phi_\nu = \frac{\mathrm{d}N_\nu}{\mathrm{d}t\mathrm{d}S} = 10^{13} \text{ cm}^{-2}\text{s}^{-1}$$

Using a $CaCl_2 + H_2O$ target the reaction we expect to observe is an extra production of $\gamma + \gamma$ corresponding to the following reaction between the decay products and the electrons of the target

$$\bar{\nu} + p \rightarrow n + e^+$$
$$e^+ + e^- \rightarrow \gamma + \gamma$$
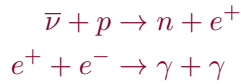
There is a second $\gamma + \gamma$ peak that needs to be accounted, since neutrons thermalize in the collisions with the fluid, it's possible to also have a neutron capture reaction inside the target, as follows

$$n + {}^A_Z X \rightarrow {}^{A+1}_Z Y^\star \rightarrow {}^{A+1}_Z Y + \gamma + \gamma$$

Where $E_\gamma \approx 6 \text{ MeV}$. Turning on the reactor it's possible to see a sharp increase of $\gamma + \gamma$ reactions, corresponding to the neutrino scattering, confirming the existence of the neutrino.

27 Electromagnetic Interactions

§ 27.1 Particle Detection

We have described nuclear physics with three main interactions: Electromagnetic, Strong (Nuclear) and Weak (Fermi's Theory).

All experimental results consisted in measuring the impulse of decay products, the energy of photons produced and the detection of neutrons. All these measures obviously need the existence of a particle detector, i.e. a device for a quantitative measure of some information or trace of particles.

The usual trace detected is the energy released by the particles, although ideally a particle shouldn't lose energy for being detected.

Suppose that a particle has and energy E and loses a ΔE. There are three possible scenarios

1. The particle loses a small $\Delta E/E$ fraction, this loss can be ignored and the particle can be detected using a charged particle tracker

2. The particle loses a considerable fraction of energy, in this case it's possible to study the process and find the correct value of E

3. $\Delta E/E = 1$ and the particle loses *all* energy. In this case we use «calorimetry»

Consider now the detector. The detector itself, being made of matter, interacts with a projectile particle. Depending on the interaction type there are three important considerations to make

1. EM interactions have an infinite radius of interaction

2. Strong interactions need a $b \approx 1 \text{ fm}$

3. Weak interactions have an infinitesimally small radius of interaction and therefore they're *almost* negligible

From these three considerations we can say without doubt $\sigma_{EM} >> \sigma_S >> \sigma_W$. Note that for $q = 0$ and $q \neq 0$ require different detection strategies

§§ 27.1.1 Charged Particles and Ionization

Charged particles with $q \neq 0$ itσpossible to define a clear distinction

- Hadrons, which interact strongly

- Leptons, which interact weakly

Leptons include quarks q, muons μ, tauons τ and electrons e. The detection strategy for these particles is using elastic collisions as in Rutherford's experiment.
The two kinds of collisions we can consider are collision with nuclei and nuclear collisions with electrons.
In the first collision the nucleus behaves like a wall for the electron, therefore $\Delta E_e << 1$, where

$$\Delta E = \frac{\Delta p^2}{2m_N}$$

The second collision we have the nucleus as a projectile, for which $m_p = M >> m_e$ and the electron ricochets with a really big energy variation

$$\Delta E_e >> 1$$

Here the interaction with the electron is dominant.
The two main results of this collision are the following

1. The electron collects energy and the atom or molecule excites, when de-exciting it emits a photon

2. If $M >> m_e$ and therefore $\Delta E_p >> 1$ and the atom can be ionized and it's then possible to measure the charge q, for which $q \propto \Delta E_p \propto E_p$

The main loss of energy for charged particles is ionization.
The loss of energy for ionization can be described as a function of density ρ, momentum $\underline{p}$, Energy E, charge q, electron momenta $\underline{p}_e$ and average ionization charge $\langle I \rangle$, all depending on the traveled space Δx made by the charged particle inside the target.
The ionization loss per unit length was studied in depth by Bethe Bloch et al. using QED in 1930, but a non relativistic limit can be studied using Bohr's atom.
Considering that collisions are a stochastic process depending on Δx, for the central limit theorem we can expect that the loss of energy follows a Gaussian distribution.
For small targets this distribution becomes the Landau distribution and it follows closely a Poisson statistics.

§§ 27.1.2 Bohr Formula

Considering a non relativistic case for an atom with charge Ze we have that energy transitions are the usual

$$\Delta E = \hbar(\omega_2 - \omega_1)$$

Where the energy levels are

$$E_n = -\frac{\alpha Z e}{2n^2}$$

Using a semiclassical approximation we can say that the classical radius of the electron is, from $m_e c^2 = \mathcal{U}_{EM}$ $r_e = \alpha/m_e$.

Putting ourselves in the reference frame of the projectile we have that the transferred momentum to the electron is

$$\Delta \underline{p} = \int \underline{F} \mathrm{d}t = \int \frac{\underline{F}}{v} \mathrm{d}x$$

Dividing the force into the parallel and orthogonal components we have that the parallel component is zero due to symmetry evaluations, and therefore

$$\Delta p_\perp = \frac{1}{v} \int F_\perp \mathrm{d}x = \frac{e}{v} \int E_\perp \mathrm{d}x \tag{27.1}$$

Considering a cylinder long L with radius b where b is the impact parameter will then be the surface integral of the perpendicular component of the electric field, which will then be divided into the sum of the top and bottom circles of the cylinders and the boundary of such The flux of the field is then

$$\phi(\underline{E}_\perp) = \int_{\Sigma_1} \underline{E}_\perp \mathrm{d}\underline{S} + \int_{\Sigma_2} \underline{E}_\perp \mathrm{d}\underline{S} = 2\pi b \int E_\perp \mathrm{d}x$$

Solvin the integral we have that

$$\Delta p_\perp = \frac{2Ze^2}{4\pi\epsilon_0 bv} = Z \frac{e^2}{4\pi\epsilon_0 b^2} \frac{2b}{v}$$

The parameter $2b/v$ is known as the «collision time» .

In the non-relativistic approximation we have that the energy variation T is

$$T = \frac{\Delta p * 2}{2m_e} = \left(\frac{e^2}{4\pi\epsilon_0}\right)^2 \frac{4}{b^2 v^2} \frac{Z^2}{2m_e}$$

Using $r_e = \frac{e^2}{4\pi\epsilon_0} \frac{1}{m_e c^2}$ we get

$$T = 2m_e c^2 \frac{Z^2 r_e^2}{b^2 \beta^2} \tag{27.2}$$

This equation gives the relation between the impact parameter b and the transferred energy T.

For a fact we know that this transferred energy is equal to $-\Delta E_p$ which is the lost energy by the projectile, in our case the nucleus.

Inverting for the impact parameter we get

$$b^2 = 2m_e c^2 v_e^2 \frac{Z^2}{\beta^2 T} \tag{27.3}$$

Since $\mathrm{d}\sigma = 2\pi b \mathrm{d}b$ we have

$$\mathrm{d}\sigma = 2\pi m_e c^2 r_e^2 \frac{Z^2}{\beta^2 T^2} \mathrm{d}T \tag{27.4}$$

Which is the cross section for the process in study.

It's obvious that $\sigma'(T) \propto T^{-2}$ and therefore processes with small T are more probable. Note that $T \propto b^{-2}$.

Since T is the lost energy in a single collision, the total lost energy will be $\Delta E_t = \Delta E N$ where N is the total number of collisions, and

$$N = 2\pi b n_e \mathrm{d}b \mathrm{d}x$$

And therefore

$$\frac{\mathrm{d}^2 E}{\mathrm{d}b\mathrm{d}x} = 2\pi n_3 b \Delta E = 4\pi m_e c^2 r_e^2 n_e \frac{Z^2}{\beta^2 b} \tag{27.5}$$

In the end, the searched energy loss is

$$\frac{\mathrm{d}E}{\mathrm{d}x} = A \log\left(\frac{b_{max}}{b_{min}}\right)$$

Therefore, we have

$$\frac{\mathrm{d}E}{\mathrm{d}x} \propto \frac{Z^2}{\beta^2} \log\left(\frac{b_{min}}{b_{max}}\right)$$

Bohr continued this calculation evaluating this in the quantum mechanical case of the atom considering that $b/v \propto t_{ion}$, i.e. the b/v is proportional to the ionization time of the electron. Considering therefore the electron as if it was free we have

$$\frac{b}{v} \approx \gamma t_{det} = \frac{\gamma}{\langle \omega \rangle} = \gamma \frac{\hbar}{I}$$

Where t_{det} is the detection time, with I being the average ionization energy.

Therefore

$$b_{max} = \gamma \frac{\hbar v}{I} = \frac{\hbar \gamma \beta c}{I}$$

For the indetermination principle we can say that $\Delta x \approx \hbar/p_e$ and therefore

$$b_{min} \approx \Delta x = \frac{\hbar}{p_e} = \frac{\hbar}{\gamma \beta m c}$$

Substituting it into (27.5) we have

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = 4\pi m_e c^2 r_e^2 n_e \frac{Z^2}{\beta^2} \log \frac{\gamma^2 \beta^2 m_e c^2}{I}$$

Using $n_e = \rho A N_A Z/A$ we get

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = 4\pi m_e c^2 r_e^2 N_A \rho \frac{Z^3}{\beta^2 A} \log \frac{\gamma^2 \beta^2 m_e c^2}{I} \tag{27.6}$$

Note that the minus sign is there since the energy gained comes from the loss of energy of the projectile, due to conservation of energy.

Approximating the constants to 0.9 MeV and dividing $-\frac{\mathrm{d}E}{\mathrm{d}x}$ by the density of the material we get «Bohr's formula»

$$-\frac{1}{\rho}\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{0.3}{2}\ \mathrm{MeV}\frac{Z^2}{\beta^2}\log\frac{\beta^2\gamma^2 m_e c^2}{I} \tag{27.7}$$

This formula has major problems, since it doesn't account for quantum mechanics and relativity and isn't universal for all materials.

Bethe and Bloch, then reevaluated the calculation using Quantum Electrodynamics, finding that

$$-\frac{1}{\rho}\frac{\mathrm{d}E_{QFD}}{\mathrm{d}x} = (0.3\ \mathrm{MeV})\frac{Z^3}{A\beta^2}\left(\log\frac{2m_e\beta^2\gamma^2 c^2\omega_{max}}{I^2} - \beta^2 - \frac{\delta}{2}\right) \tag{27.8}$$

Where $\omega_{max} \propto \max E_e^\star$, i.e. $\omega_{max} = 2m_e c^2\beta^2\gamma^2$.

We immediately notice $-\beta^2$ as a relativistic correction and $\frac{\delta}{2}$ as a correction for density and polarization effects of the object, all together corrected by an electric screening factor.

This formula, permits us with a measurement of $\frac{\mathrm{d}E}{\mathrm{d}x}$ and p permits to identify the projectile particle, with a formula that holds up for $\beta\gamma$ from $0.1 \to 1000$

§§ 27.1.3 Residual Path

Roche proceeded to use the Bethe-Bloch equation for evaluating the so called «residual path» , i.e. the path for which the projectile loses all its energy. This is given by

$$R(E) = \int_E^0 \mathrm{d}x = \int_E^0 \frac{\mathrm{d}E}{-\frac{\mathrm{d}E_{QED}}{\mathrm{d}x}} = \int_0^E \frac{\mathrm{d}E}{\frac{\mathrm{d}E_{QED}}{\mathrm{d}x}} \tag{27.9}$$

From Bethe-Bloch's equation it's quick to deduce that $\beta\gamma$ small include heavy losses of energy.

It's of note that before the actual stopping of the projectile particle, the Bethe-Bloch formula finds its maximum in what's known the «Bragg peak» .

This spike in energy is quite useful in tumor treating in what's known as Hadrontherapy, where hadrons, usually p or ^{12}C, are shot in a localized region and by energy dissipation releases a sharp amount of energy where shot.

In general, considering an elastic collision with nuclei, $\frac{\mathrm{d}E}{\mathrm{d}x} \approx 0$ due to collisions being elastic, but this leads to a big $\Delta\underline{p}$, which implies an angular deviation as in Rutherford's experiment. The angular deviations are random, and it makes this a stochastic process.

Consider a thick target, for the theorem of the central limit we can say that the distribution of angular deviation is proportional to a Gaussian distribution

$$f(\Delta\theta) \propto G(\langle\theta\rangle = 0, \sigma = \sqrt{\langle\theta^2\rangle})$$

The $n-$th moment of the distribution will be

$$\langle\theta^n\rangle = \frac{\int \theta^n\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\mathrm{d}\Omega}{\int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\mathrm{d}\Omega}$$

Where σ is the Rutherford cross section, which is proportional to $\csc^2(\theta/2)$. Using $d\Omega \approx 2\pi\theta d\theta$ we can say without many doubts that

$$\frac{d\sigma}{d\Omega} \propto \frac{d\sigma}{d\theta}$$

Calculating the second momenta we get the expected value for a multiple coulombian scattering, which is

$$\langle \theta^2 \rangle = 21 \text{ MeV} \frac{Z}{\beta cp}\sqrt{\frac{x}{\chi_0}}$$

Where x is the depth traveled in the medium (in cm), and χ_0 is the «radiation length», which is

$$\frac{1}{\chi_0} = 4r_e^2\alpha\rho\frac{N_A Z^2}{A} \log\left(183 Z^{-1/3}\right) \tag{27.10}$$

§ 27.2 Cherenkov Effect

The Cherenkov effect is an effect similar to a sonic boom, for which a massive particle passes through a medium at a velocity higher than the speed of light in that medium $c_n = c/n$. This effect is due to polarization effects.
Consider a projectile moving at $v_x > c_n$ in some medium. The Cherenkov radiation is accompanied by the emission of photons perpendicular to the light cone at some angle θ_c. Consider now the process after a time t. The radius of the boom circle created by this "superluminal" particle is $L = v_y t = ct/n$, supposing the particle has traveled some space $\Delta x = \beta c\Delta t$ we have

$$L = \Delta x \cos\theta_c \Longrightarrow \cos\theta_c = \frac{1}{\beta n} \tag{27.11}$$

The angle θ_c can be experimentally measured, and a measure of θ_c and p.
Note that the constraint $\cos\theta_c < 1$ we get that this process exists if and only if $\beta \geq 1/n$
Tying this angle to the momentum of the particle we begin by noting that

$$\beta\gamma = \frac{\beta}{\sqrt{1-\beta^2}} = \frac{p}{m}$$

We have

$$\frac{1}{\beta} = \sqrt{1 + \frac{m^2}{p^2}}$$

Substituting what we found before for $\cos\theta_c$ we have

$$\cos\theta_c = \frac{1}{\beta n} = \frac{1}{n}\sqrt{1 + \frac{m^2}{p^2}} > 1 \tag{27.12}$$

Which implies that the second constraint on Cherenkov radiation is that $p \geq p_{th}$. In the limit case of $o_c = p + m$, $\beta_{th} = n^{-1}$ we get

$$\frac{1}{n} = \sqrt{1 + \frac{m^2}{p_{th}^2}}$$

Which, since $n > 1$ gives

$$p_{th} = \frac{m}{\sqrt{n^2 - 1}} \tag{27.13}$$

For $p >> p_{th}$ it's clear from formulas and experimental measurements that it's not possible to distinguish between different particles from their Cherenkov angle.

In order to find the number of Cherenkov photons emitted in the process we can use classical electrodynamics, getting

$$\frac{\mathrm{d}^2 N}{\mathrm{d}x\mathrm{d}E} = \frac{\alpha Z^2}{\hbar c}\left(1 - \frac{1}{\beta^2 n^2(E)}\right) = \frac{\alpha Z^2}{\hbar c}\sin^2\theta_c \tag{27.14}$$

Where $n(E) = n(\lambda)$ is the refraction index of the medium.

Using $E = 2\pi\hbar c/\lambda$ we have changing variables

$$\frac{\mathrm{d}^2 N}{\mathrm{d}x\mathrm{d}\lambda} = \frac{2\pi\hbar c}{\lambda^2}\frac{\mathrm{d}^2 N}{\mathrm{d}x\mathrm{d}E} = \frac{2\pi\alpha Z^2}{\lambda^2}\sin^2\left(\theta_c(\lambda)\right) \tag{27.15}$$

In the visible spectrum we have $E \approx 2\text{ eV } (1.8 \rightarrow 3.1\text{ eV})$

§ 27.3 Loss of Energy for Electrons and Positrons

Since we're dealing with the loss of energy of charged particles we have to use the Bethe-Bloch loss of energy equation.

We have that tat the minimum of the functon $\beta\gamma \approx 3$ and therefore $p \approx 1.5\text{ MeV}$ for electrons.

This impilies that the relativistic climb of the electrons is really slow.

Considering heavy projectiles like nuclei. The electrons will get strongly accelerated and will emit radiation, as for the Larmor formula we have that the radiated power is

$$P_L = \frac{2}{3}\frac{e^2}{m^2 c^3}|\dot{\underline{v}}|^2 \tag{27.16}$$

Extending it to a Lorentz invariant formulation we have

$$P_L = -\frac{2}{3}\frac{e^2}{m^2 c^3}\frac{\mathrm{d}p_\mu}{\mathrm{d}\tau}\frac{\mathrm{d}p^\mu}{\mathrm{d}\tau}$$

Using

$$\frac{\mathrm{d}p_\mu}{\mathrm{d}\tau}\frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} = \frac{1}{c^2}\left|\frac{\mathrm{d}E}{\mathrm{d}\tau}\right|^2 - \left|\frac{\mathrm{d}\underline{p}}{\mathrm{d}\tau}\right|^2$$

And substituting $E = \gamma m c^2$, $\underline{p} = \gamma m c \underline{\beta}$ we have

$$P_L = \frac{2e^2}{3c}\gamma^6\left(\left(\frac{\mathrm{d}\underline{\beta}}{\mathrm{d}t}\right)^2 - \left(\underline{\beta} \times \frac{\mathrm{d}\underline{\beta}}{\mathrm{d}\tau}\right)^2\right) \tag{27.17}$$

Which, β, γ are the relativistic parameters of a charged particle.
There are two limit cases, one of linear acceleration, and one of perpendicular acceleration of the particles.
For linear acceleration we have

$$\underline{\dot{\beta}} \| \underline{\beta}$$
$$\underline{\dot{\beta}} \times \underline{\beta} = 0$$
$$P_L = \frac{2e^2}{3c}\gamma^6\dot{\beta}^2$$

And for perpendicular acceleration

$$\underline{\dot{\beta}} \perp \underline{\beta}$$
$$P_L = \frac{2e^2}{c}\gamma^6\left(\dot{\beta}^2 - \beta^2\dot{\beta}^2\right)$$

Writing the addition on the parentheses as $\beta\dot{\beta}^2/\gamma$ we get that the Larmor radiation power of the electron on a perpendicular acceleration motion (circular motion, ndr.) is proportional to the energy of the particle divided by its mass

$$P_L \propto \gamma^4 = \frac{E^4}{m^4}$$

This is important for the determination of the particle emitting this Larmor radiation.
Suppose you have two charged projectiles with $m_1 \neq m_2$ and $E_1 = E_2$, since $P_L \propto m^{-4}$ we have a huge increase in the Larmor power for the lighter (charged) particle.
As an example take E fixed and evaluate the radiated power of protons and electrons, we have

$$\frac{P_L^e}{P_L^p} = \frac{m_e^4}{m_p^4} \approx 1.6 \cdot 10^{13}$$

I.e. electrons emit radiation 13 orders of magnitude more powerful than the one emitted by protons

§ 27.4 Bremsstrahlung

Bremsstrahlung, or «braking radiation» translated from German, is an effect dominant in electrons, and approximately 0 for $m > m_e$ as long as E isn't big enough.
From Bethe-Bloch's formula we have that the lost energy is

$$-\frac{\mathrm{d}E}{\mathrm{d}x} \propto \frac{E}{\chi_0}$$

Where χ_0 is defined in (27.10) and it's a property of the medium. Solving the approximate ODE we have

$$E(x) = E_0 e^{-\frac{x}{\chi_0}} \tag{27.18}$$

At $x = \chi_0$ we have E_0/e which implies that energy decreases by about 30%. In a standard length χ_0 there therefore around 63% of loss of energy. This loss of energy is known as «Bremmstrahlung» .

Confronting it with the ionization energy loss for electrons and positrons we have

$$-\frac{\mathrm{d}E_{tot}}{\mathrm{d}x} = -\left(\frac{\mathrm{d}E_{ion}}{\mathrm{d}x} + \frac{\mathrm{d}E_{Brem}}{\mathrm{d}x} \right)$$

Inserting the Bethe-Bloch formula we have

$$R_{B/i} = \frac{\frac{\mathrm{d}E_{Brem}}{\mathrm{d}x}}{\frac{\mathrm{d}E_{ion}}{\mathrm{d}x}} \approx \frac{K_e Z}{1200 m_e}$$

Where $K_e = E_e - m_e \approx E_e$. It's defined as «critical energy of the medium» the value of energy such that $R_{B/i} = 1$. For electrons we have

$$E_c = \frac{600}{Z} \text{ MeV}$$

This is useful for evaluating the ionization minimum for materials, as

$$I_{min} = \frac{E_c}{\chi_0} \approx (3.5 \text{ MeV}) \frac{Z}{A}$$

§ 27.5 Photons in Matter

There are three major processes for photons that depend directly on the energy of the photon γ

1. At low energy, the photoelectric effect dominates, with reaction

$$\gamma + {}^A_Z \text{X} \rightarrow {}^A_{Z-1} \text{X} + e^- \tag{27.19}$$

2. At a higher energy, the Compton scattering process dominates, with reaction

$$\gamma + e^- \rightarrow \gamma + e^- \tag{27.20}$$

3. At high energies, the pair production effect dominates, with reaction

$$\gamma + {}^A_Z \text{X} \rightarrow e^+ + e^- + {}^A_Z \text{X} \tag{27.21}$$

The total cross-section for photon interaction will obviously be a function depending on E_γ and Z of the interacting nucleus, since at different energies (and also at different Z) different processes can happen

§§ 27.5.1 Photoelectric Effect

Considering the photoelectric effect we have a reaction of the kind

$$\gamma + {}^A_Z X \rightarrow {}^A_{Z-1} X + e^-$$
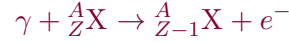
In this effect we have

$$\sqrt{s} = \sqrt{m_e^2 + m_X^2 + 2\left(E_e E_X - \underline{p}_e \underline{p}_X\right)}$$

The second part $m_X^2 + 2\left(E_e E_X - \underline{p}_e \underline{p}_X\right)$ is the impulse absorbed by the nucleus. Since $m_X \gg m_e$ we have $\Delta E_X = 0$ and all the impulse of γ gets transferred to the electron. The quantum effect to take note in this case is the ionization energy of the electron, which gets absorbed by e^- in order to escape the atomic bound state, therefore the gained energy is

$$E_e = E_\gamma - I$$

Inverting the relation in terms of E_γ and noting that $E_e = K_e$ we have

$$E_\gamma = K_e + I$$

This effect is dominant for $E_\gamma < m_e \approx 100\text{ keV}$, and the cross section is

$$\sigma_{\gamma X} \propto \begin{cases} \dfrac{Z^5}{E_\gamma^3} & E_\gamma < m_e \\[2mm] \dfrac{Z^5}{E_\gamma} & E_\gamma > m_e \end{cases} \tag{27.22}$$

§§ 27.5.2 Compton Effect

For photon energies $E_\gamma \gg I$ we have that $\sigma_{\gamma X}$ gets particularly small and the Compton effect becomes the dominant EM process.
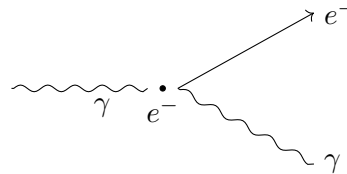

Figure 27.1: Sketch of the Compton scattering, the photon and the electron get scattered by an angle θ

Using the transformations of angle equation in special relativity we have that the new energy of the photon after the scattering is

$$E'_\gamma = \frac{E_\gamma}{1 + \frac{E_\gamma}{m_e}(1 - \cos\theta)} \tag{27.23}$$

The maximum energy transfer will happen at $\theta_{max} = \pi$, for which

$$E_\gamma(\theta = \pi) = \frac{E_\gamma}{1 + \frac{2E_\gamma}{m_e}} \tag{27.24}$$

And when $E_\gamma \approx m_e$, where solving we have

$$E'_\gamma \approx \frac{E_\gamma}{3} = \frac{m_e}{3} \tag{27.25}$$

In this last case, for $\theta = \pi$, $E_\gamma \approx m_e$, the photon loses around 66% of its initial energy, transferring it all to the electron.
This effect is dominant for $E_\gamma > m_e$ and its cross section is

$$\sigma_{\gamma e^-} \propto \frac{Z}{E_\gamma} \tag{27.26}$$

§§ 27.5.3 Pair Production

Considering the pair production process we might think to write it as a process where a photon decays into an electron and a positron

$$\gamma \to e^- + e^+$$

Evaluating $\sqrt{s}$ for the LHS and RHS of the process we see clearly that this is impossible, since

$$\sqrt{s} = m_\gamma = 0 \neq m_{e^-}^2 + m_{e^+} + 2\left(E_{e^-}E_{e^+} - \underline{p}_{e^-}\underline{p}_{e^+}\right)$$

Adding a nucleus $^A_Z X$ on both sides we instead get

$$\sqrt{s} = m_X^2 + 2E_\gamma m_X$$

Where $2E_\gamma m_X$ is a recoil term of the nucleus given by the conservation of momentum.
It's obvious that there is a threshold energy for this process, for which

$$E_\gamma \geq m_{e^-} + m_{e^+} \approx 1.022\ \text{MeV} \tag{27.27}$$

Taken all these process at once we have that for $10\ \text{eV} \leq E_\gamma \leq 100\ \text{keV}$ the photoelectric effect is dominant, for $100\ \text{keV} < E_\gamma \leq 10\ \text{MeV}$ the Compton effect is dominant with a maximum for $E_\gamma = m_e$. And lastly for $E_\gamma > 10\ \text{MeV}$ the pair production effect is the most predominant effect, starting from $E_\gamma = 1.022\ \text{MeV}$

§§ 27.5.4 Attenuation Length for Photons

Consider now a beam of photons going through some medium. We have that the intensity of the beam will depend on the distance traveled inside the medium as for the equation

$$I(x) = I_0 e^{-\frac{x}{\lambda}} = I_0 e^{-\mu x} \tag{27.28}$$

Where μ is the well known attenuation coefficient. Considering the relation between the cross section and the attenuation coefficient we have that

$$\mu = \mu_{e^+ e^-} + \mu_{\gamma e^-} + \mu_{\gamma X} = \sigma n = \left(\sum_i \sigma_i \right) n \tag{27.29}$$

Where σ is the total cross section of the photon interaction. We have that for $E > 100 \text{ MeV}$ σ is mostly constant, and therefore

$$\frac{1}{x_\gamma} = \sigma n \approx \frac{7}{9} \frac{1}{\chi_0} \tag{27.30}$$

Where $1/x_\gamma$ is the attenuation length of the photon, which is deeply tied to the radiation length of the medium. Note that n is the refraction coefficient.
Since we're in the range of $E_\gamma > 10 \text{ MeV}$ the dominant process is pair production, and from this we can calculate the cross section for pair production as

$$\sigma_{e^+ e^-} = \frac{7A}{9 \rho \chi_0 N_A} \propto Z^2 \log \left(183 Z^{-1/3} \right) \tag{27.31}$$

Note that positrons have the exact same parameters of electrons, excluding the charge which is opposite.

§§ 27.5.5 Electromagnetic Showers in Mediums

Consider a beam of high energy e^+, e^- in some medium, these particles will produce high energy photons through Bremsstrahlung radiation. These high energy photons will then produce pairs of e^+, e^- creating a shower.
Note that the positron-electron pair will lose 30% of their initial energy due to Bremsstrahlung and the photons lose around 60% of their initial energy due to the previous listed scattering processes.
This process is a stochastic shower which continues up until $E_i > E_c$. This process is determined by χ_0, which gives the average distance traveled before the doubling of particles, which cause E to half, and it will continue up until there is no more Bremsstrahlung effect (E_i reaches the level of Compton scattering).
This process is studied in function of the depth $t = x/\chi_0$ and $-E_0^{-1} \frac{dE}{dx}$. The profile recovered for this stochastic process is

$$F(t) \propto \frac{E_0}{E_c} t^a e^{-bt} \tag{27.32}$$

And from this we have that the maximum depth the shower will reach is

$$t_{max} = \log\left(\frac{E_0}{E_c}\right) + c \tag{27.33}$$

Where a, b, c are parameters given by energetic corrections and fitting of data.
The cone created by the shower is given by multiple coulomb scattering, which gives no energy loss (elastic scattering) but only an angular reaction. The radius of the cylinder containing 90% of the particles interacting in the shower is

$$R_M = \frac{21}{E_c}\chi_0 \tag{27.34}$$

Which is known as «Moliere radius».

§§ 27.5.6 Hadronic Interactions

Consider now the case of hadrons, particles that interact both electromagnetically and strongly, like $p, \pi^{\pm}, K^{\pm}, n$ (protons, pions, kaons, neutrons). These particles can interact strongly with the nuclei in the medium. In general

1. For low energies ($2 \text{ GeV} \leq E \leq 5 \text{ GeV}$) the scattering is elastic and there is no energy loss

2. For intermediate energies ($5 \text{ GeV} \leq E \leq 100 \text{ GeV}$) there is an EM interaction with the medium which transfers around 100 MeV of energy

3. For high energies ($E > 100 \text{ GeV}$) the hadrons interact strongly with the nucleus, and similarly to the EM case, a shower happens

For the intermediate energy levels we have that $\sigma_n \approx \pi R_N^2 \approx \pi \text{ fm}^2 \approx 30 \text{ mb}$ and the cross section is inelastic, since the collision is inelastic. σ grows with energy, $\sigma \approx \sigma_0 A^{2/3}$.
For the high energy case, considering a beam of hadrons with intensity $I(x) = I_0 \exp(-x/\lambda)$ we can say that

$$\frac{1}{A} = \sigma_h n = \sigma\rho\frac{N_A}{A}$$

Where λ is the nuclear interaction length, also known as the usual mean free path between the inelastic collisions. In general $\sigma_n \approx 1 \text{ b} < \sigma_{EM}$ therefore $\lambda > \chi_0$

28 Particle Detectors

Consider a general reaction

$$p + p \rightarrow H + q_1 + q_2$$

Where H is the Higgs boson which decays as usual into two Z^0 bosons which decay in electron-positron or muon-antimuon couples, and q_1, q_2 are hadron swarms.

The ideal objective is to measure the 4-momentum of all particles. For this various kinds of detectors are used

§ 28.1 Trackers

Tracker detectors for charged particles which measure p through ionization. They rebuild the trace of the particle and use the radius of curvature of the path and use it for finding p.

In this case $\frac{dE}{dx}$ must be measurable.

In order to have this reduction of E from ionization we have from Bethe-Bloch that

$$-\frac{dE}{dx} \propto C\rho \frac{Z}{A} f(\beta\gamma)$$

Inside the detector we have that the reduction of energy is

$$-dE \propto \rho dx$$

And therefore we need ρ small (i.e. a gas) and a thin dx

§§ 28.1.1 Cloud Chambers

Cloud chambers are a simple kind of tracker detector composed by a container filled with saturated vapor.

The charged particles passing through the vapor ionize it, creating bubbles around the ions. After that the track can be photographed and then the needed values can be measured.

This type of particle detector was used to discover the positron. In that cloud chamber a $6 \ \mathrm{mm}$ thick slab of Pb was used as a target to slow the particles and a magnetic field was applied.
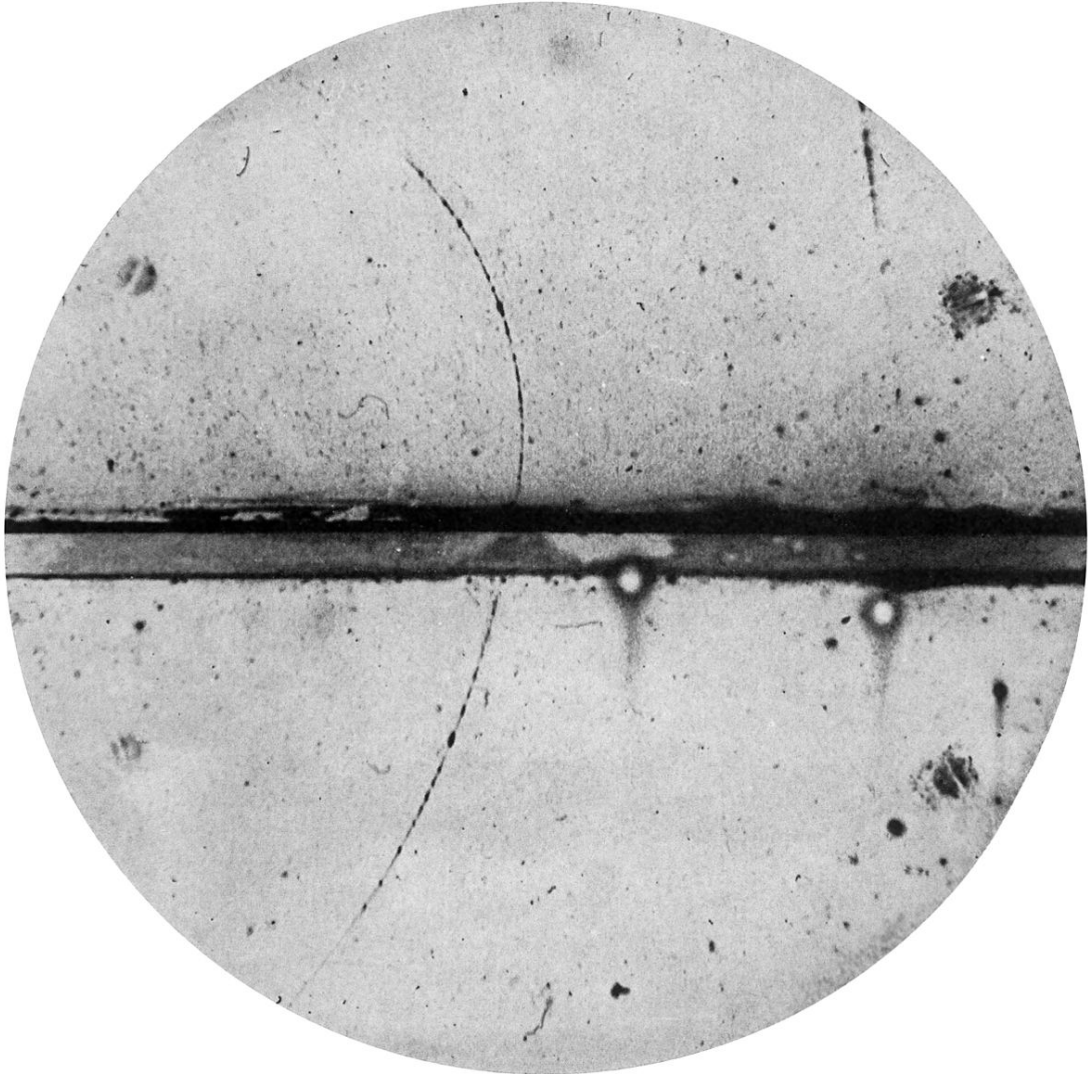
Figure 28.1: The picture taken to the cloud chamber in 1932 by Anderson et al. that shown the production of a positron

From the photo it's possible to identify a particle going through the lead from the bottom, with $p_i = 63$ MeV and with $p_f = 23$ MeV. Using $\chi_{0,\mathrm{Pb}} = 5.6$ mm we have that

$$E(\chi_0) \approx \frac{E_0}{3}$$

Supposing that the particle is a proton we would have

$$\beta\gamma = \frac{p}{m_p} = \frac{63}{1000} \approx 0.06$$

Since it's clear from the experiment that the particle passed more than 6 cm, we have that it must be an unknown particle, the positron in this case. Doing the same calculation with m_e we get $\beta\gamma \approx 120$.

A slight modification of this is the bubble chamber, which uses high pressure liquids to create the saturated vapor. This detector has a much greater spatial resolution with $\delta x \approx 100~\mu\text{m}$

§§ 28.1.2 Nuclear Emulsions

Another kind of tracker detector used is the nuclear emulsion, where gelatinous slabs of silver bromide AgBr are in a suspension on the slab.

The ionization of charged particles releases electrons that make silver shine engraving the track permanently on the slab.

This type of detector has a resolution of $\delta x \approx 1~\mu\text{m}$, but it's slow to analyze, and therefore it's useful for processes with low frequencies.

§ 28.2 Ionization Detectors

Ionization detectors function by catching the ionization charge of the particles passing by. The measured signal depends on $\frac{\mathrm{d}E}{\mathrm{d}x}$ and on the potential difference ΔV created by two plates that enclose a noble gas.

Chapak in 1962 proposed the construction of such detector with cathode planes and parallel anode wires. The spectral resolution of such detector is $\delta x \approx 300~\mu\text{m}$

§§ 28.2.1 Drift Chambers

Drift chambers are ionization detectors made by a multiwire proportional chamber which measures the time needed for a signal to arrive from one side to another

§§ 28.2.2 Silicon Detectors

Silicon detectors are another kind of ionization detectors. They are really quick and have a high resolution. They use semiconductors.

Considering that ρ_{Si} is big, the detectors are quite thin, with a resolution of $\delta x \approx 10~\mu\text{m}$. Note that their resolution is deeply tied to the magnetic field B and thickness of the detector L, in fact we have

$$\frac{\delta P}{P} = \frac{P}{0.3BL^2}\delta x \tag{28.1}$$

I.e., the resolution gets smaller when P grows. Note that

$$P \approx 0.3BR$$

Where P is expressed in GeV, B in T and R in m

§ 28.3 Calorimeters and Energy Measures

The passage of photons, electron positron pairs and hadrons creates EM or hadronic swarms inside mediums. Their ionization can excite molecules or atoms inside this medium.
One kind of calorimeter is the scintillator, which records the de-excitation photons of the swarms (in the visible, close UV spectra). There are two major kinds of scintillators

1. Organic or plastic scintillators made mainly from anthracene, with a response time of 10^{-8} s

2. Inorganic or crystalline scintillators, made mainly from $\mathrm{NaI}, \mathrm{CsI}, \mathrm{PbWO_4}$ with a response time of 10^{-6} s

The main problem with scintillators is that the medium reabsorbs part of these de-excitation photons, giving fewer photons to measure, they activate mainly in the close UV region.
The main solution for solving the first problem is using doping materials in order to increase the scintillation photons.
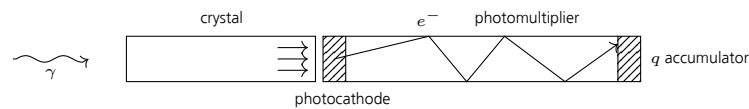

Figure 28.2: Tiny scheme describing how a scintillator functions

As in the picture, the photocatode is a piece sensible to photons and photoelectric effects. It creates a landslide effects on diodes that amplify the photo-electron (photoelectric process' child particle).
For around 10 nodes it's possible to have a gain of 10^4 till 10^7 depending on the value of ΔV. The charge measured q is proportional to the number of photo-electrons and therefore proportional to the energy of the photons.
A detector made only of scintillators and photomultipliers is known as a «homogeneous calorimeter» .
The calorimeters use long crystals, such that $N\chi_0 \approx 20\chi_0$ in order to contain more or less all the energy of the EM swarm. Therefore, for a crystal with $\chi_0 = 3$ cm in a scintillator it'd be long at least 50 cm.
Hadronic swarms, on the counterpart are regulated by the interaction length $\lambda >> \chi_0$, therefore using homogeneous calorimeters would be prohibitive since they'd need to be really long.
For hadrons usually «sample calorimeters» which use a first scintillator block and multiple absorbent blocks, permitting an absorption of the hadrons, since $\lambda^{-1} = \rho\sigma$. Inside the absorber the swarm develops faster and we have

$$\delta q \propto \sqrt{N} \propto \sqrt{N}$$

In calorimeters we have

$$\frac{\delta E}{E} \approx \frac{a}{\sqrt{E}} \tag{28.2}$$

Where a is a measured constant depending on the calorimeter used. This constant is known as the characteristic constant of the calorimeter, and it's known that $a_{hom} < a_{sample}$. Note also that a higher energy input grants a better resolution, as clear from (28.2)

§ 28.4 Particle Accelerators

A great example, and the first, of particle accelerators, is the well known cathodic tube, which accelerates electrons using a variation of tension. It's a typical example of linear accelerator.
Other famous linear accelerators are PEP-II and BaBAR which managed to reach a $\sqrt{s} = 90 \text{ GeV}$

§§ 28.4.1 Cyclotrons

Cyclotron accelerators are a type of accelerators which use a magnetic field in order to accelerate circularly charged particles. It was first suggested by Lawrence in 1959 and accelerate ions emitted at the center of a circular object composed by two semicircular "Dee".
The frequency of a cyclotron is readily calculable using classical electromagnetism, and it's equal to

$$\nu_c = \frac{qB}{2\pi m} \tag{28.3}$$

This comes since inside the cyclotron there is an uniformly accelerated circular motion caused by ΔV. Using Newton's second law we have

$$F = ma = m\frac{v^2}{r} = qvB \Longrightarrow \frac{v}{r} = \frac{qB}{m}$$

The period of a revolution is fixed by B, and equals

$$\Delta T = \frac{1}{\nu_c} = \frac{2\pi m}{qB}$$

Note that this formula is not relativistic. In the relativistic formulation, using $\Delta t = \gamma \delta \tau$ we have

$$\nu_c = \frac{qB}{2\pi \gamma m} \qquad \gamma = \frac{E}{m} \tag{28.4}$$

It's quick to see that the cyclotron frequency depends on the velocity of the particle (in the relativistic case), and therefore ultrarelativistc e^- are not suitable for a cyclotron.
Ions, being much heavier and slower, are a better particle candidate for use inside cyclotrons. We have

$$T_{max} = \frac{1}{2}mv_{max}^2 = \frac{1}{2}m\omega^2 r^2$$

And therefore, setting $r = R$ with R being the radius of the cyclotron, we get that the maximum kinetic energy reached by the ions inside the cyclotron is

$$T_{max} = \frac{(qBR)^2}{2m} \tag{28.5}$$

For relativistic particles, since if v grows $1/\gamma$ decreases, we might suppose to decrease the cyclotron's frequency in order to keep valid the previous relation, keeping B constant
This kind of variable-frequency cyclotron is known as a synchro-cyclotron, which use variable E fields to reach this result.
Another solution is to variate B while keeping ν_c constant in order to compensate for γ^{-1}, these accelerators are known as synchrotrons.
Synchrocyclotrons are usually used for accelerating particles from 10 to 900 MeV, which is a relatively small acceleration

§§ 28.4.2 Synchrotrons

Synchrotrons are the relativistic counterpart of cyclotrons. They work by fixing the radius R and variating B for compensating for γ^{-1} in the synchrotron frequency formula

$$\nu_s = \frac{qB}{2\pi\gamma m} \tag{28.6}$$

This kind of accelerator doesn't need to create poles like the cyclotrons and its "Dees" and a uniform B field is not needed in all the accelerator, therefore there are multiple dipoles along the synchrotrons.
The principal limit of synchrotron is Larmor radiation, also known as synchrotron radiation, with power

$$P_L = \frac{e^2}{6\pi\epsilon_0 c^3}\gamma^4 aP2 \tag{28.7}$$

Using $\gamma = E/m$ and $p = v^2/R$ we get

$$P_L = \frac{e^2 E^4}{6\pi\epsilon_0 R^2 m^4} \tag{28.8}$$

Which, for a period $\Delta T = 2\pi R/c$ gives that the energy lost in synchrotron radiation is

$$\Delta E_{lost} \propto \frac{E^4}{m^4 R^2}$$

In order to balance this energy loss it's needed to make bigger synchrotrons. The biggest (so far) is the Large Hadron Collider in Geneva, a synchrotron with $R = 4.3 \text{ km}$. For electrons in the LHC ($m_e \approx 500 \text{ keV}$) we have that the lost energy is proportional to

$$\Delta E_{e,LHC} \approx 88.5\frac{E^4}{4300} \tag{28.9}$$

Which imposes that for high energy electrons all energy is used to compensate for radiative losses inside the synchrotron.

For LEP@CERN, an experiment lasted from 1988 till the early 2000, there was $\sqrt{s} = 90 \text{ GeV}$, for a single beam energy of 45 GeV. It used the same tunnel of LHC and its radiative losses per lap were

$$\Delta E_{LEP} = 84 \text{ MeV/lap}$$

Note that

$$\nu_{LEP} = \frac{c}{2\pi R} = \frac{c}{27 \text{ km}} \approx 10^6 \text{ Hz}$$

Note that for what we have seen it's impossible to accelerate electrons to TeV ranges without increasing the radius of the synchrotron. From 2000 onwards, using protons, a $\sqrt{s} = 13 \text{ TeV}$ has been reached, which corresponds to 6.5 MeV per beam.

A new project is planned, the FCC, a supercollider with a circumference of 100 km built in Geneva. With these radius it's possible to reach $\sqrt{s} = 50 \text{ TeV}$.

29 Basics of Quantum Field Theory

§ 29.1 Perturbations and Feynman Diagrams

The first question that comes to mind when talking about particle accelerators is why.
The idea is quite simple and comes from $E^2 = m^2 + p^2$, therefore we can use higher E for converting it into mass.
This lets us discover heavier particles, create more particles via inelastic collisions and to obviously discover new particles.
In general we treat a reaction of the kind

$$a + b \rightarrow c + d(+f + g + \cdot)$$

We need to evaluate the initial and final states for such reactions, and Fermi's golden rule comes in handy for this. For FGR, given a transition probability P_{fi} between an initial state and a final state is

$$\lim_{T \to 0} \frac{P_{fi}}{T} = 2\pi \left| \hat{\mathcal{M}}_{fi} \right|^2 \delta \left(E_f - E_i \right) \tag{29.1}$$

Note that different disposition of particles in the final state changes only the kinematics of the final state, since the 4-momentum is always conserved.
What this actually mean is that $\hat{\mathcal{M}}_{fi}$ is independent from the kinematics of the process.
Remember that the decay rate of a reaction is

$$\Gamma_{fi} = \int 2\pi \left| \hat{\mathcal{M}}_{fi} \right|^2 \delta(E_f - E_i) \mathrm{d}n = 2\pi \left| \hat{\mathcal{M}}_{fi} \right|^2 \rho(E) \tag{29.2}$$

Where $\rho(E)$ is the density of states in the phase space and is equal to

$$\rho(E) = \frac{\mathrm{d}n}{\mathrm{d}E_f} \tag{29.3}$$

It's important to note that

$$\hat{\mathcal{M}}_{fi} = -i \langle f | \hat{\mathcal{H}}_I | f \rangle \tag{29.4}$$

Where $\hat{\mathcal{H}}_I$ is the perturbation Hamiltonian for a system $\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{\mathcal{H}}_I$.
Using perturbation theory for such system we have

$$\hat{\mathcal{H}}_0 \psi_n = E_n \psi_n i \frac{\partial \psi}{\partial t} = \left(\hat{\mathcal{H}}_0 + \hat{\mathcal{H}}_I \right) \psi$$

$$\psi = \sum_n a_n(t) \psi_n e^{-iE_n t} \tag{29.5}$$

For a transition $|i\rangle \to |f\rangle$ we impose the following approximation conditions on the functions $a_n(t)$

$$\begin{cases} a_i(t) = 1 & k = i \\ a_k(0) = 0 & k \neq i \end{cases} \tag{29.6}$$

Taking the second equation in (29.5) and multiplying on the left by $\langle k|$ we get the following Schrödinger equation for a $k-$th state

$$\dot{a}_k(t) = -i \int \overline{\psi}_k \hat{\mathcal{H}}_I \psi_i e^{i(E_k - E_i)t} \mathrm{d}^3 r \tag{29.7}$$

Integrating and rewriting on the RHS the definition of $\hat{\mathcal{M}}_{ki}$, and imposing $|k\rangle = |f\rangle$ we have that

$$P_{fi} = |a_f(t)|^2 = \left| \int_0^t \hat{\mathcal{M}}_{fi} e^{i(E_f - E_i)t} \mathrm{d}t \right|^2 \tag{29.8}$$

Going over to the second order terms of the perturbation and reinserting everything inside the Schrödinger equation, where we write the $a_n(t)$ we found before, we have

$$\dot{a}_k(t) = -i\hat{V}_{ki} e^{i(E_k - E_i)t} + (-i)^3 \sum_{n \neq i} \frac{\hat{V}_{kn} \hat{V}_{ni}}{E_n - E_i} e^{i(E_k - E_i)t} = -i\hat{V}_{ki}^{(2)} e^{i(E_k - E_i)t} \tag{29.9}$$

Where $\hat{V}_{ki}^{(2)}$ is the second order perturbation and is equal to

$$\hat{V}_{ki}^{(2)} = \hat{V}_{ki} + (-i)^3 \sum_{n \neq i} \frac{\hat{V}_{kn} \hat{V}_{ni}}{E_i - E_n} \tag{29.10}$$

$$\hat{V}_{ki} = \langle k| \hat{\mathcal{H}}_I |i\rangle$$

The second order transition probability is then

$$\Gamma_{fi}^{(2)} = 2\pi \left| \hat{\mathcal{M}}_{fi}^{(2)} \right|^2 \rho(E) \tag{29.11}$$

Using FGR we have for the conservation of energy $\delta(E_f - E_i) \Longrightarrow E_f = E_i$ but only to the first order.
Considering a scattering like Rutherford scattering we have considering second order corrections

$$\hat{\mathcal{M}}_{fi} \sim \sum_{n \neq i} \frac{\hat{V}_{fn} \hat{V}_{ni}}{E_i - E_n}$$

The system in this approximation will jump from the first state $|i\rangle$ to accessible states $|n\rangle$ to the final state $|f\rangle$ via the transition matrices $\hat{V}_{ni}$, $\hat{V}_{fn}$.

By definition we have $E_n \neq E_i \neq E_f$, therefore energy is not conserved in this situation. This is fixed by imposing that states with $E_n \gg E_i$ and $E_n \gg E_i$ are improbable.

A good example of showing graphically these perturbations is using «Feynman diagrams», graphs where each vertex corresponds to a degree of perturbation.

At the first order of perturbation we can write as an example the following transition

$$\left|e^+ e^-\right\rangle \rightarrow |\gamma\rangle$$

Which corresponds to a pair annihilation reaction $e^+ + e^- \rightarrow \gamma$. The first order perturbation will be graphed as
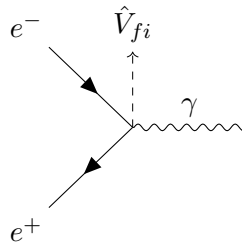

Figure 29.1: Feynman representation of the first order perturbation $\hat{V}_{fi}$.

This graph is a ≈ 1 dimensional graph. The only dimension accounted here is time, which flows from right to left. Matter is drawn as arrows flowing with time and antimatter (see e^+) is drawn with arrows that flow against time. The vertices represent the actual interaction matrix $\hat{V}_{fi}$, in this case only for a 1st order perturbation.

With a quick check of this process we see that

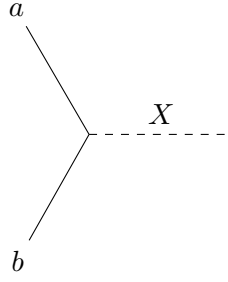1. The incoming particles are a positron and an electron

2. The only outgoing particle is a photon

And therefore

$$\sqrt{s_i} = 2m_e^2 \neq 0 = \sqrt{s_f}$$

Therefore the conservation of energy given by $\delta(E_f - E_i)$ is not valid.

More generally, with a reaction $a + b \rightarrow X$, the diagram would be drawn as

Figure 29.2: Feynman diagram for the process $a + b \rightarrow X$

Note that this process can only happen if $s_i = m_a^2 + m_b^2 = m_X^2 = s_f$, and therefore doesn't happen for $s_i > s_f$.

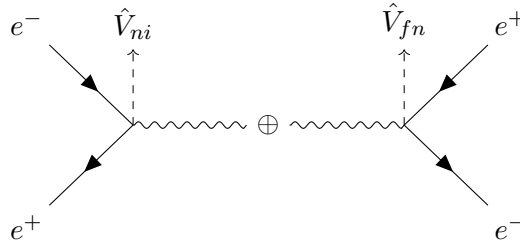For second order processes the diagram for the interaction described in (29.1) becomes

Figure 29.3: Second order diagram considered as the sum of the two vertexes corresponding to the transitions $|i\rangle \rightarrow |n\rangle$ and $|n\rangle \rightarrow |f\rangle$. The photon connecting the two diagram is known as «virtual» due to its non-physical and non measurable energy E_n

This diagram is exactly drawn as the sum of two single vertex diagrams corresponding to a transition $|i\rangle \rightarrow |n\rangle \rightarrow |f\rangle$. The photon inside can have E_n that aren't possible otherwise, such as $E_n \neq p^2 + m^2$, the so called «off shell» energies.

What happened in this scattering process, a $e^+e^- \rightarrow e^+e^-$ elastic scattering, is that a virtual photon mediates the process. Basically the electron and positron annihilate creating a virtual photon which re-decays into an electron and a positron, which get measured as outgoing particles.

Note that for Heisenberg this is possible. In fact $\Delta E \Delta t \approx \hbar$ imply that $\Delta t \leq \hbar / \Delta E$, and therefore for $t < \Delta t$ the ΔE violations are possible, as long as charge and quantum numbers are conserved.

§§ 29.1.1 Electrodynamic Processes

For describing electromagnetic processes we can build three basic vertexes which can be used to build up higher order diagrams
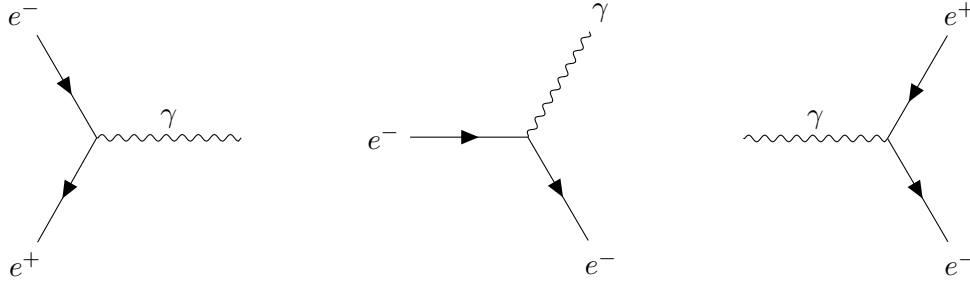
All these three basic vertexes share one main thing: charge is conserved.

Consider now Rutherford scattering, this process can be described via the following diagram
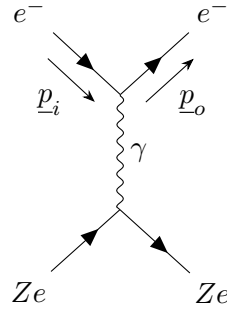
Figure 29.4: Feynman diagram for Rutherford scattering

From this diagram we can immediately see that $Q_i = e(Z - 1) = Q_f$, since we used two fundamental vertexes. In order to get something more from this diagram, especially how to grasp the perturbations from the vertexes we need to do some calculations on the initial state $|i\rangle$ and the final state $|f\rangle$. We begin by Born-approximating the wavefunction of the incoming electron as a planar wave, therefore

$$
\psi_i = \frac{1}{\sqrt{V}} e^{i\underline{p}_i \underline{r}}
$$
$$
\psi_0 = \frac{1}{\sqrt{V}} e^{i\underline{p}_o \underline{r}}
$$

(29.12)

Therefore, having an electromagnetic perturbation given by the potential of the nucleus, we have that our perturbation is

$$
\hat{V}_{fi} = -\langle f| \frac{Z\alpha}{r} |i\rangle = -\frac{Z\alpha}{V} \int \frac{1}{r} e^{i(\underline{p}_i - \underline{p}_o)\underline{r}} \mathrm{d}^3 r
$$

Writing $\underline{q} = \underline{p}_i - \underline{p}_f$ and $\underline{q}\underline{r} = qr\cos\theta$ and transforming the integral into spherical coordinates we have

$$
\hat{V}_{fi} = -\frac{Z\alpha}{V} \int_0^\infty r\mathrm{d}r \int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi e^{iqr\cos\theta} \sin\theta \mathrm{d}\theta
$$

Taking the third integral and writing $\mathrm{d}\cos\theta = -\sin\theta$ we have

$$\int_0^\pi \sin\theta e^{iqr\cos\theta}\mathrm{d}\theta = \frac{1}{iqr}\int_{-1}^1 e^{iqr\cos\theta}\mathrm{d}\cos\theta = \frac{1}{iqr}\left(e^{iqr} - e^{-iqr}\right)$$

Reinserting it into the integral and integrating with respect to φ we have

$$\hat{V}_{fi} = -\frac{2\pi Z\alpha}{iqV}\int_0^\infty \left(e^{iqr} - e^{-iqr}\right)\mathrm{d}r$$

The last integral can be calculated using something similar to Feynman's integration trick, by multiplying the function by a dummy function $e^{-\varepsilon r}$ and taking the limit for $\varepsilon \to 0$ for getting back the last result. Substituting and integrating we have

$$\hat{V}_{fi} = -\frac{2Z\pi\alpha}{iqV}\lim_{\varepsilon\to 0}\left(-\frac{1}{iq-\varepsilon} - \frac{1}{iq+\varepsilon}\right) = -\frac{2Z\pi\alpha}{iqV}\frac{2iq}{q^2}$$

Simplifying, we get

$$\hat{V}_{fi} = -\frac{-4Z\pi\alpha}{Vq^2} \tag{29.13}$$

Using $\hat{\mathcal{M}}_{fi} = -i\hat{V}_{fi}$ we get the transition matrix, and the transition probability as

$$\hat{\mathcal{M}}_{fi} = \frac{4iZ\pi\alpha}{Vq^2}$$
$$\left|\hat{\mathcal{M}}_{fi}\right|^2 = \frac{16Z^2\pi^2\alpha^2}{V^2q^4} \tag{29.14}$$

Which implies $\sigma \propto Z^2\alpha^2 q^{-4}$.

Going back to the diagram and remembering that each vertex represents a perturbation we have, for the vertex of the nucleus a charge of Ze and a contribute of $\sqrt{\alpha}$, for the other one we have another contribution $\sqrt{\alpha}$ with charge e, connected by a virtual photon, which contributes for the moment with a so called «propagator» . The photon propagator is proportional to q^{-2}, and therefore, simply by looking at the two vertexes in natural units $(e = 1)$, we get

$$\hat{\mathcal{M}}_{fi} = \sqrt{\alpha}\frac{1}{q^2}Z\sqrt{\alpha} = \frac{Z\alpha}{q^2} \tag{29.15}$$

Which is what we found up to a factor of $-i4\pi$.

Consider now as a second example of these rules the process of Compton scattering. The diagram for this process will be
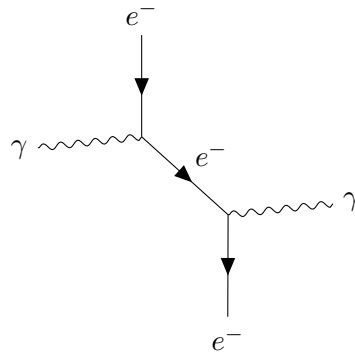
Figure 29.5: Second order diagram for Compton scattering

Using Feynman rules on vertexes we have that $\hat{\mathcal{M}}_{fi} \approx \sqrt{\alpha}\sqrt{\alpha} = \alpha$, and therefore we can immediately suppose that $\sigma \sim \left|\hat{\mathcal{M}}_{fi}\right|^2 \approx \alpha^2$.

One quick thought about conservation laws makes clear that if the virtual particle in the diagram *must* be an electron, because the number of leptons must be conserved in all vertexes. Note that if the virtual particle was a baryon like a proton, it also wouldn't be right since the number of baryons isn't conserved.

Also consider a quick thing, if the outgoing particles were switched in the diagram, the result would be the same, although we must consider this diagram's contribution in the final calculations.

Take now the Bremsstrahlung radiation, this process corresponds to the following third order diagram For this diagram we have $\hat{\mathcal{M}}_{fi} \propto Z\sqrt{\alpha}\sqrt{\alpha}\sqrt{\alpha}\frac{1}{q^2}$, therefore $\sigma \propto Z^2\alpha^3/q^4$.
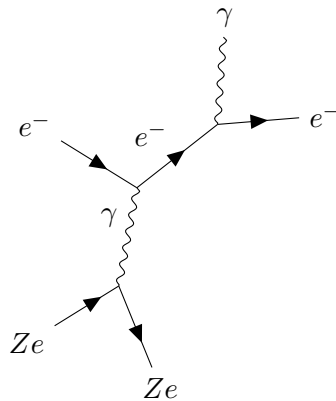
Figure 29.6: Bremsstrahlung effect third order Feynman diagram

Confronting the obtained cross section to the one for Rutherford scattering we get $\sigma_R \propto \alpha^2$, and $\sigma_{Brem} \propto \alpha^3$, with $\alpha = 1/137$.

Going down this path of describing electrodynamic processes with Feynman diagrams we impact ourselves in a new Feynman rule for electromagnetic interactions: vertexes can't

have multiple photons reaching them, therefore this diagram is impossible
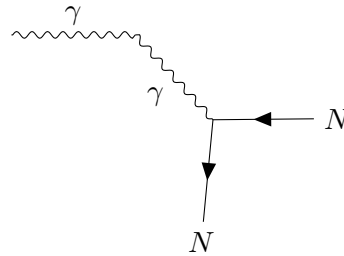
Figure 29.7: An impossible diagram

And now one might think, how can I build a pair production diagram? The answer is: add a virtual particle between the photonic vertexes. The searched diagram then is
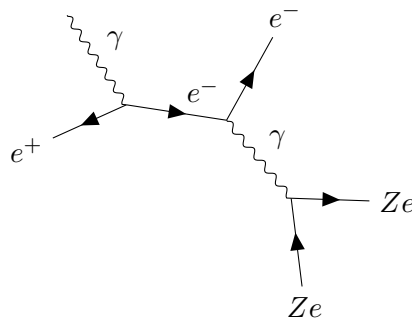
Figure 29.8: Maybe fix this? idk not sure probably pair production

By just looking at the diagram we have $\hat{\mathcal{M}}_{fi} \propto Z\sqrt{\alpha}\sqrt{\alpha}\sqrt{\alpha}q^{-2} = Z\alpha^{3/2}/q^2$. Another interesting process is Bhabha scattering, $e^+e^- \to e^+e^-$. This diagram is pretty simple to draw

e^- e^-

γ

e^+ e^+
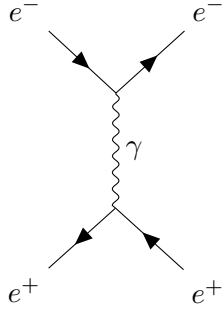
Figure 29.9: Bhabha scattering diagram

e^- e^+

γ

e^+ e^-

Figure 29.10: A symmetric version of the same bhabha scattering diagram

By checking the nodes of this diagram we get immediately that $\sigma \propto \alpha^2/q^4$

§ 29.2 Klein-Gordon Equation and the Yukawa Potential

Yukawa in 1935, using the idea of virtual mediator particles went on trying to explain the nuclear force between nucleons. Experimetally it had been seen that it was a short range force, and that there is a symmetry between neutrons and protons.

From classical EM we know that the electrostatic potential generated by a pointlike charge at the origin solves the inhomogeneous Poisson equation

$$\nabla^2 V = -e\delta^3(r) \tag{29.16}$$

The solution is an integral retarded potential

$$V(r) = \int_V \frac{\rho(\underline{r}')}{|\underline{r} - \underline{r}'|} \mathrm{d}^3 r' \tag{29.17}$$

For time dependent potentials, defining $\Box = \partial^\mu \partial_\mu = \partial_t^2 - \nabla^2$ the Maxwell equations are

$$\Box\,(\underline{E}, \underline{B}) = 0 \tag{29.18}$$

Where, in the potential formulation become, writing a 4-potential $A_\mu = (\phi, \underline{A})$

$$\frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = \rho$$
$$\frac{\partial^2 \underline{A}}{\partial t^2} - \nabla^2 \underline{A} = \underline{J} \tag{29.19}$$

Taking only the first of the two equations, we might think to quantize this equation imposing $i\partial_t \to \hat{E}$ and $-i\nabla \to \hat{\underline{p}}$, getting the following equation

$$\Box\phi = \left(\hat{E}^2 - \hat{p}^2\right)\phi = 0 \tag{29.20}$$

It's immediately clear that we must have $E = p$, and therefore this equation works only for massless particles.

Since we're dealing with massive particles we might immediately think to substitute $E^2 = m^2 + p^2$, getting what is known as the «Klein-Gordon equation», which satisfies a massive mediation of the potential

$$\left(\hat{E}^2 - \hat{p}^2 + m^2\right)\phi = \left(\Box + m^2\right)\phi = 0 \tag{29.21}$$

Fitting it into the stationary case of the nucleon with a nuclear charge $g \neq 0$ at $\underline{r} = 0$ we get the following equation

$$\left(\nabla^2 - m^2\right)\phi(\underline{r}) = -g\delta^3(\underline{r}) \tag{29.22}$$

Which has for solution the Yukawa potential for the strong interaction

$$\phi(r) = -\frac{g}{4\pi r}e^{-mr} \tag{29.23}$$

This potential corresponds to a shielded couloomb interaction.

In natural units we have that mr is adimensional, and using $\Delta E \Delta t \approx \hbar$ with $'DE \approx mc^2, \Delta t \approx R/c$ with $R \approx 1.5$ fm we get that

$$mc^2 \approx \frac{\hbar c}{R} \approx 150 \text{ MeV}$$

The mediating particle for this process must have a mass of $m \approx 150$ MeV, so Yukawa theorized the existence of a mesotron particle which mediates the nuclear force.

From experiments we now know that this mesotron, or better known as pion, isn't an elementary particle, and therefore Yukawa's model is effective in explaining these interactions but it's not a fundamental one.

In general a charged pion π^2 is a meson (quark-antiquark bound state), with state $|u\bar{d}\rangle$.

We might think to find this pion propagator using Yukawa's potential as a perturbation to Born states, which gives

$$\langle f| - \frac{g}{4\pi r}e^{-mr}|i\rangle = -i\frac{g^2}{4\pi}\frac{1}{q^2 + m^2} \tag{29.24}$$

This propagator is similar to the photon propagator, and using $\alpha_{EM} = e^2/4\pi$ we might think to construct a "fine structure constant" for strong interaction, which is $\alpha_S = g^2/4\pi$.

In general for a massive potential we have a propagator of the following kind

$$\langle f|\hat{\mathcal{H}}_I|i\rangle = -i\alpha\frac{1}{q^2 + m^2} \tag{29.25}$$

Supposing a force with $m^2 >> q^2$ we can see immediately that the momentum exchanged between the interacting particles is negligible with respect to the mass of the mediating particle, which lets us approximate the propagator to

$$\langle f|\hat{\mathcal{H}}_I|i\rangle \approx -i\alpha\frac{1}{m^2} \tag{29.26}$$

§§ 29.2.1 Weak Interactions

We might think to apply this approximation to Fermi's interactions, with $\hat{\mathcal{H}}_I = G_F$. The weak interaction that Fermi studied is a finite range interaction. Let's suppose that this interaction is mediated by a massive particle with charge g_w and mass m_w, the potential for such interaction is analoguous to Yukawa's potential

$$V_w = -\frac{g_w}{4\pi r}e^{-m_w r} \tag{29.27}$$

Writing $\hat{\mathcal{H}}_I = g_w V_w = G_F$ we get from experimental values

$$G_F = \frac{g_w^2}{4\pi m_w^2} = 1.16 \cdot 10^{-5} \text{ MeV}^{-2} \tag{29.28}$$

Which suggests a mass of the weak mediator, in terms of measurable quantities, of

$$m_w = \left(\frac{g_w}{e}\right)^2 \frac{\alpha_{EM}}{G_F} \approx \left(\frac{g_w}{e}\right)^2 10^2 \text{ MeV} \tag{29.29}$$

In order to discover this new W particle a new particle accelerator was built at LEP, which was the combination of a proton-synchrotron and an antiproton-synchrotron, which was used to verify the following theoretical reaction

$$p + \bar{p} \rightarrow W^+ + X \tag{29.30}$$

Considering a cross beam interaction with two targets we have

$$E_p = E_{\bar{p}} = 270 \text{ GeV}, \qquad \sqrt{s} = 2E_p = 540 \text{ GeV}$$

The confirmation of this reaction awarded a Nobel prize in 1984 to Rubbiz and Van der Meer.

So far we have listed three fundamental forces: Electromagnetism, strong force and the weak force, all mediated by particles, respectively $\gamma, g, W^\pm/Z^0$. The second mediator is known as the «gluon» and it's a massless particle, which is not predicted by Yukawa's theory, although it's effective to explain nuclear phenomena.

Only one fundamental force is missing a mediator particle, which is gravity. The idea of a «graviton» particle which for now there haven't been any experimental verifications.

§ 29.3 Symmetries

As we know from Nöther's theorem, each constant of motion corresponds to a simmetry of the system.

Considering reactions $a + b \rightarrow c + d$, we immediately know that a conserved quantity in the reaction is a symmetry of the transition Hamiltonian $\hat{\mathcal{H}}_I$.

There are 4 kinds of symmetry we might consider

1. Continuous symmetries

   - Temporal traslations, which correspond to E conservation
   - Spatial translations, which correspond to p conservation
   - Spatial rotations, which correspond to $\underline{L}$ conservation

2. Gauge symmetries, which correspond to q conservation

3. Fundamental symmetries, like the lepton number conservation, baryon number conservation, etc.

4. Non-spatial rotations, which correspond to the conservation of a new quantity known as Isospin $\hat{\underline{I}}$

§§ 29.3.1 Discrete Transformations

Going to the quantum world with our symmetries we can list immediately three discrete symmetries

- Parity reflection, $\hat{P} : \hat{r} \to -\hat{r}$

- Charge conjugation, $\hat{C} : q \to -q$

- Time inversion, $\hat{T} : t \to -t$

These laws are multiplicative, and as an example the matter-antimatter symmetry corresponds to a $\hat{C}\hat{P}$ transformation.

§§ 29.3.2 Leptonic Number

Consider a weak interaction like the β decay, with reaction

$$n \to p + e^- + X$$

Considering Reines-Cowan's experiment which supposes a reaction $X + p \to n + e^+$, a possible reaction with $Q > 0$, we have a problem, and the reaction was never observed. It was theorized the existence of a leptonic number associated with the electron L_e, for which $\Delta L_e = 0$ in reactions.
Leptonic matter accounts for $L = 1$, while leptonic antimatter for $L = -1$. In leptonic matter also neutrinos are accounted, and we can make a list of couples lepton/neutrino.

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix} L_e = 1 \quad \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix} L_\mu = 1 \quad \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix} L_\tau = 1 \tag{29.31}$$

And antilepton/antineutrino

$$\begin{pmatrix} e^+ \\ \overline{\nu}_e \end{pmatrix} L_e = -1 \quad \begin{pmatrix} \mu^+ \\ \overline{\nu}_\mu \end{pmatrix} L_\mu = -1 \quad \begin{pmatrix} \tau^+ \\ \overline{\nu}_\tau \end{pmatrix} L_\tau = -1 \tag{29.32}$$

For all interactions, a violation of the conservation of the leptonic number was never observed, and therefore it's possible to assume

$$\Delta L_i = 0 \qquad i = e, \mu, \tau \tag{29.33}$$

§§ 29.3.3 Baryonic Number and Isospin

Segré with its experiment discovered the antiproton, was searching for the proof of the existence of the following reaction

$$p + p \to p + p + p + \bar{p} \tag{29.34}$$

This reaction corresponds to an inelastic scattering of two protons, but there are also other reactions possible

$$p + p \to p + p + \pi^+ + \pi^-$$
$$p + p \to p + p + \pi^0$$
$$p + p \to p + p$$

From his eperiment it was discovered that $m_p = m_{\bar{p}}$, but also a new question arises. Why is this reaction not observed?

$$p + p \to p + \bar{p} + \pi^+ + \pi^+$$

This non-possibility of the previous reaction brings us to the conservation of the baryonic number B.

A baryon (antibaryon) is a composite particle composed by three quarks (antiquarks)

$$B = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \qquad \overline{B} = \begin{pmatrix} \bar{q}_1 \\ \bar{q}_2 \\ \bar{q}_3 \end{pmatrix} \tag{29.35}$$

For each baryon a baryonic number N is associated, which evaluates to ± 1 if the particle considered is either a baryon or an antibaryon. Particles like the pions $\pi^\pm, \pi^0$ are known as «mesons» and are composed of a quark/antiquark bound state

$$M = \begin{pmatrix} q_1 \\ \bar{q}_2 \end{pmatrix} \tag{29.36}$$

Counting $1/3$ for each quark and $-1/3$ for each antiquark we have that mesons contribute with $N_B = 0$ and each baryon with $N_B = \pm 1$ as we said before.

The question remains, why doesn't that reaction happen, since the baryonic number is conserved? In order to explain this we have to dwelve deeper into the physics of this.

Since $m_n - m_p \approx 1 \, \text{MeV}$, nuclear interactions *do not* distinguish neutrons from protons. Due to the non-degeneration theorem we know that there exists a new degree of freedom of the system with an associated quantum number in order to account to this strong symmetry

between n, p.

This new degree of freedom is the «Isospin» , which its algebra corresponds exactly to the spin algebra, a rotation group.

We have for protons and neutrons

$$|p\rangle = |I, I_3\rangle = \left|\frac{1}{2}, \frac{1}{2}\right\rangle$$

$$|n\rangle = |I, I_3\rangle = \left|\frac{1}{2}, -\frac{1}{2}\right\rangle$$

(29.37)

Accounting for this, nuclear interactions are invariant under isospin transformations.

Note that for pions

$$\left|\pi^+\right\rangle = |11\rangle$$

$$\left|\pi^0\right\rangle = |10\rangle$$

$$\left|\pi^-\right\rangle = |1-1\rangle$$

(29.38)

§ 29.4 Isospin

Isospin was a property first introduced by Heisenberg in order to explain strong nuclear interactions between protons and neutrons.

This property behaves algebrically as angular momentum and it can be used for classifying known Hadrons, estimate strong cross section and to theorize states which are not yet observed.

In strong interactions isospin is conserved, indicating that it's a symmetry of the strong interaction Hamiltonian.

Let's begin considering a bound neutron-proton state, deuterium ^2_1H. As we know from before we have

$$|p\rangle = |I, I_z\rangle = \left|\frac{1}{2}, \frac{1}{2}\right\rangle$$

$$|n\rangle = |I, I_z\rangle = \left|\frac{1}{2}, -\frac{1}{2}\right\rangle$$

Using the angular momentum summation rules we have that, for a state $|np\rangle$ we have $I = 0, 1$ which implies the existence of three states plus one with total $I_z = -1, 0, 1$, these triplet states correspond to the couple $|pp\rangle, |np\rangle, |nn\rangle$ plus $|np\rangle$ and are

$$|pp\rangle = |p\rangle |p\rangle = |1, 1\rangle$$

$$|np\rangle = \frac{1}{\sqrt{2}} (|n\rangle |p\rangle + |n\rangle |p\rangle) = |0, 1\rangle$$

$$|nn\rangle = |n\rangle |n\rangle = |1, -1\rangle$$

(29.39)

The previous states are known, as for spin, triplet symmetric isospin states. The additional remaining state is the singlet antisymmetric state

$$|np\rangle = \frac{1}{\sqrt{2}} (|n\rangle |p\rangle - |n\rangle |p\rangle) = |0, 0\rangle$$

(29.40)

Experimentally triplet states are not observed, therefore we can safely assume that for a deuterium nucleus, (deuteron, d) we have

$$|d\rangle = |np\rangle = |00\rangle \tag{29.41}$$

§§ 29.4.1 Pion-Nucleon Scattering

Consider the scattering between pions and nucleons

$$\pi^{\pm,0} + (p,n) \to \pi^{\pm,0} + (p,n) \tag{29.42}$$

Considering the isospin for the system and noting that $I_\pi = 1$, $I_{p,n} = 1/2$ we have that for the initial state

$$|i\rangle = |1,a\rangle + \left|\frac{1}{2},b\right\rangle = \alpha\left|\frac{3}{2},a+b\right\rangle + \beta\left|\frac{1}{2},a+b\right\rangle \tag{29.43}$$

The scattering process can be seen as the addition of an interaction Hamiltonian $\hat{\mathcal{H}}_S$ which is invariant for isospin, and considering a final state

$$|f\rangle = \gamma\left|\frac{3}{2}\right\rangle + \delta\left|\frac{1}{2}\right\rangle \tag{29.44}$$

We have that, using the selection rule $\Delta I = 0$

$$\langle f|\hat{\mathcal{H}}_S|i\rangle = a\left\langle\frac{3}{2}\right|\hat{\mathcal{H}}_S\left|\frac{3}{2}\right\rangle + b\left\langle\frac{1}{2}\right|\hat{\mathcal{H}}_S\left|\frac{1}{2}\right\rangle = a\hat{\mathcal{M}}_3 + b\hat{\mathcal{M}}_1 \tag{29.45}$$

The final cross section will then be

$$\sigma \propto |a|^2\left|\hat{\mathcal{M}}_3\right|^2 + |b|^2\left|\hat{\mathcal{M}}_1\right|^2 + ab\hat{\mathcal{M}}_1\hat{\mathcal{M}}_3 \tag{29.46}$$

Using Clebsch-Gordan coefficients we can already see how these transition matrices will decompose.
Noting that pions are isospin 1 and nucleons are isospin $1/2$ system we have a $\mathbf{3}\otimes\mathbf{2}$ system. From the table B we have that such state will decouple in a symmetric quadruplet and an antisymmetric doublet (remembering that $I_z^\pi = 1, 0 - 1$ and $I_z^{(n,p)} = 1/2, -1/2$) giving for $\pi^+ + p \to \pi^+ + p$ and $\pi^- + n \to \pi^- + n$

$$\sigma \propto \left|\hat{\mathcal{M}}_3\right|^2 \tag{29.47}$$

Due to symmetry one immediately expects that $N(\pi^- + p) = N(\pi^+ + n)$ where

$$\sigma_{\pi^- p} \propto \frac{1}{3}\left|\hat{\mathcal{M}}_3\right|^2 + \frac{2}{3}\left|\hat{\mathcal{M}}_1\right|^2 + \frac{2}{9}\hat{\mathcal{M}}_1\hat{M}_3 \tag{29.48}$$

Expanding for the other possible reaction and remembering that $\hat{\mathcal{H}}_S$ is orthonormal in the total isospin basis the remaining cross sections (and therefore also the number of reactions) are easily calculable

§§ 29.4.2 Isospin and Charge

Consider a nucleus (A, Z), using $I_z^p = 1/2$, $I_z^n = -1/2$ we have that the total $z-$projection of isospin will be

$$I_z = \frac{Z}{2} - \frac{A - Z}{2} \tag{29.49}$$

Inverting for Z we have a direct connection between isospin and charge

$$Z = I_z + \frac{A}{2} \tag{29.50}$$

Considering that A is the total number of neutrons and protons in the nucleus, we have from Gell-Mann-Nishijima that for a generic system with baryonic number B the total charge will be given by the following equation

$$Q = I_3 + \frac{B}{2} \tag{29.51}$$

This clearly also works for mesons, in fact taking π^+ as an example, for which $Q = 1$

$$Q_{\pi^+} = 1 + \frac{0}{2} = 1$$

The later discovery of strange particle broke this formula, which was fixed by defining a *hypercharge*, which is defined as

$$Y = B + S \tag{29.52}$$

With B being the baryonic number and S the strangeness number. The introduction of the quark model with charm, strange, top and bottom quarks then defined the hypercharge as the sum of the quantum numbers of all the previous quarks, giving for the Nishijima equation

$$Q = I_z + \frac{B + s + c + b + t}{2} = I_z + \frac{Y}{2} \tag{29.53}$$

Note that that strange and bottom quarks reduce s, b by 1 like antitop and anticharm quarks reduce charm and top quantum numbers. Remember that this is purely a convention

Part VI

Appendices

A   Mathematical Methods

§ A.1 Properties of the ϵ_{ijk} Tensor

We have defined in the chapter on 3D system, the Levi-Civita tensor, a completely antisymmetrical unit tensor that can be used to define cross products using tensor notation, and ease the calculus of multiple cross products. Some properties of this tensor, that can be particularly useful, are the following ones

$$\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl} \tag{A.1a}$$

$$\epsilon_{ijk}\epsilon_{ijl} = 2\delta_{kl} \tag{A.1b}$$

$$\epsilon_{ijk}\epsilon_{ijk} = 6 \tag{A.1c}$$

In general, we can write the following identity

$$\epsilon_{ijk}\epsilon_{lmn} = \det\begin{pmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{pmatrix} = $$

$$= \delta_{il}\left(\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}\right) - \delta_{im}\left(\delta_{jl}\delta_{kn} - \delta_{jn}\delta_{kl}\right) + \delta_{in}\left(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}\right) \tag{A.1d}$$

For a matrix a_{ij}, we can write its determinant in two ways using the Levi-Civita tensor

$$\det(a_{ij}) = \epsilon_{i_1\cdots i_n}a_{1i_1}\cdots a_{ni_n} \tag{A.1e}$$

$$\det(a_{ij}) = \frac{1}{n!}\epsilon_{i_1\cdots i_n}\epsilon_{j_1\cdots j_n}a_{i_1j_1}\cdots a_{i_nj_n} \tag{A.1f}$$

Coupling these rules to the properties of vector products, we get the following identity for operators

$$\hat{\underline{a}} \times (\hat{\underline{b}} \times \hat{\underline{c}}) = \hat{\underline{b}}(\hat{\underline{a}} \cdot \hat{\underline{c}}) - (\hat{\underline{a}} \cdot \hat{\underline{b}})\hat{\underline{c}} + \left[\hat{\underline{a}}_j, \hat{\underline{b}}\right]\hat{\underline{c}}_j \tag{A.2a}$$

Proof. Let's write the cross product in tensor notation, using the Levi-Civita tensor

$$\epsilon_{kji}\hat{a}^j\epsilon_{ilm}\hat{b}^l\hat{c}^m = \epsilon_{kji}\epsilon_{ilm}\hat{a}^j\hat{b}^l\hat{c}^m = -\epsilon_{ijk}\epsilon_{ilm}\hat{a}^j\hat{b}^l\hat{c}^m$$

$$= -\left(\hat{\delta}_{jl}\hat{\delta}_{km} - \hat{\delta}_{jm}\hat{\delta}_{kl}\right)\hat{a}^j\hat{b}^l\hat{c}^m = \hat{a}_j\hat{b}_k\hat{c}_j - \hat{a}_j\hat{b}_j\hat{c}_k = $$

$$= \left[\hat{a}_j, \hat{b}_k\right]\hat{c}_j + \hat{b}_k\hat{a}_j\hat{c}_j - \hat{a}_j\hat{b}_j\hat{c}_k$$

315

Where we used the usual index contraction rules with the kronecker symbol, where $\delta_{ij}a^j = a_i$. Usually it's made in terms of the metric tensor $g_{ij} = \langle e_i|e_j \rangle$ as basis vectors, but in non relativistic cases we simply have $g_{ij} = \delta_{ij}$ $\square$

Another useful for vector products of operators that is useful is the following

$$\underline{\hat{a}} \times \underline{\hat{b}} = -\underline{\hat{b}} \times \underline{\hat{a}} + \epsilon_{ijk}\left[\hat{a}^j, \hat{b}^k\right] \tag{A.2b}$$

Proof. This proof is straightforward using the antisimmetry of the tensor ϵ_{ijk}, and yields the following result

$$\underline{\hat{a}} \times \underline{\hat{b}} = \epsilon_{ijk}\hat{a}^j\hat{b}^k = \epsilon_{ijk}\left(\left[\hat{a}^j, \hat{b}^k\right] + \hat{b}^k\hat{a}^j\right) = -\epsilon_{ikj}\hat{b}^k\hat{a}^j + \epsilon_{ijk}\left(\left[\hat{a}^j, \hat{b}^k\right]\right)$$

$\square$

Although a proof for the next relation won't be given, we'll list it due to its usefulness in quantum mechanical calculations

$$\left(\epsilon_{ijk}\hat{a}^j\hat{b}^k\right)^2 = \hat{a}^2\hat{b}^2 - \left(\hat{a}^i\hat{b}_i\right) - \hat{a}_j\left[\hat{a}^j, \hat{b}^k\right]\hat{b}_k + \hat{a}_j\left[\hat{a}^k, \hat{b}_k\right]\hat{b}^j - \hat{a}_j\left[\hat{a}^k, \hat{b}^j\right]\hat{b}_k - \hat{a}_j\hat{a}_k\left[\hat{b}^k, \hat{b}^j\right]$$
$$\tag{A.2c}$$

In the special case of $\left[\hat{a}_i, \hat{b}_j\right] = \gamma\delta_{ij}$ with $\gamma \in \mathbb{C}$ and $\left[\hat{b}^i, \hat{b}^j\right] = 0$, it yields the following special case

$$\left(\epsilon_{ijk}\hat{a}^j\hat{b}^k\right)^2 = \hat{a}^2\hat{b}^2 - \left(\hat{a}^i\hat{b}_i\right)^2 + \gamma\hat{a}^i\hat{b}_i \tag{A.2d}$$

Now, considering the curl as a vector product between the nabla operator and a vector field ($\nabla \times \underline{F}$), we can also write the following identity

$$\nabla \times \underline{F} = \epsilon_{ijk}\frac{\partial F^k}{\partial x^j} \tag{A.3}$$

In general coordinates, indicating the Jacobian as $J := \det(\partial_j x_i)$, we can write the following identity

$$\nabla \times \underline{F} = J\epsilon_{ijk}\frac{\partial F^k}{\partial x^j} \tag{A.4}$$

Which generalizes the vector product.

§ A.2 Special Functions

§§ A.2.1 Spherical Harmonics

Spherical Harmonics appear in the chapter of quantum angular momentum, where they're defined as the eigenfunctions of the orbital angular momentum.

Simplifying our calculus, we give the definition, and actual calculation of spherical harmonics as the actual search for the eigenfunctions of orbital angular momentum.

So, using directly spherical coordinates in 3 dimensions, we already have that the $\hat{L}_z$ is written simply as follows

$$\hat{L}_z \to -i\hbar\frac{\partial}{\partial\phi} \tag{A.5}$$

Since $\left[\hat{L}_z, \hat{L}^2\right] = 0$, we also write out $\hat{L}^2$, as it is needed to solve our equation

$$\hat{L}^2 \to -\hbar^2\left(\csc^2\theta\frac{\partial^2}{\partial\phi^2} + \csc\theta\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\right) \tag{A.6}$$

By definition, we get that the eigenfunctions must solve the following equations «simultaneously»

$$\hat{L}_z f(\theta, \phi) = -i\hbar\frac{\partial f}{\partial\phi} = \hbar m f(\theta, \phi)$$

$$\hat{L}^2 f(\theta, \phi) = -\hbar^2\left(\csc^2\theta\frac{\partial^2}{\partial\phi^2} + \csc\theta\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\right) f(\theta, \phi) = \hbar^2 l(l+1) f(\theta, \phi) \tag{A.7}$$

Due to symmetry, f must also be cyclic in $\phi \mod 2\pi$, hence $f(\theta, \phi) = f(\theta, \phi + 2\pi)$. Due to the shape of the system of differential equation, we suppose that f is really the product of two function of the single variables, hence we imply that

$$f(\theta, \phi) = g(\theta)h(\phi)$$

The first equation is of immediate solution in this way, and we have that

$$-i\hbar g(\theta)\frac{\partial h}{\partial\phi} = \hbar m g(\theta)h(\phi) \tag{A.8a}$$

Hence, finally, solving the equation we get

$$h(\phi) = g(\theta)ke^{im\phi} \tag{A.8b}$$

k is simply a multiplicative constant, and without loss of generality can be set to be $k = 1$
In order for $h(\phi) = h(\phi + 2\pi)$ to be true, then $m \in \mathbb{N}$.
The second equation isn't that simple to solve, and after applying the separation of variables and, utilizing the property of $h(\phi)$, for which $h(\phi) \neq 0 \,\forall\phi$, we get

$$\csc\theta\frac{\partial}{\partial\theta}\sin\theta\frac{\partial g}{\partial\theta} - m^2\csc^2\theta g(\theta) = -\lambda g(\theta) \tag{A.9}$$

There is a way in order to "simplify" this differential equation. We define the following nontrivial sostitution

$$\begin{cases} \xi = \cos\theta \\ F(\xi) = g(\theta) \end{cases}$$

The equation (A.9) then becomes a general form of the Legendre differential equation

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\left((1-\xi^2)\frac{\mathrm{d}F}{\mathrm{d}\xi}\right) - \frac{m^2}{1-\xi^2}F(\xi) + \lambda F(\xi) = 0 \tag{A.10}$$

In order to solve (A.10) firstly we impose $m = 0$, and then we suppose that $F(\xi)$ is analytical, and hence it holds that

$$F(\xi) = \sum_{k=0}^{\infty} a_k \xi^k$$

Using a method analogous to the analytical solution of the quantum harmonic oscillator, we derive, substitute our indexes and plug the result inside our differential equation. We get then the following equation

$$\sum_{k=0}^{\infty}\left(a_k(\lambda - k(k+1)) + (k+1)(k+2)a_{k+2}\right)\xi^k = 0 \tag{A.11}$$

This is true «if and only if» the following recurrence relation is true

$$-\frac{\lambda - k(k+1)}{(k+1)(k+2)}a_k = a_{k+2} \tag{A.12}$$

In our quantum case, it's evident that $\lambda = l(l+1)$. Looking at the recurrence relation, we see that it's actually the recurrence relation of the Legendre polynomials, that are defined by the Rodrigues' formula

$$P_l(\xi) = \frac{1}{2^l l!}\frac{\mathrm{d}^l}{\mathrm{d}\xi^l}\left(\xi^2 - 1\right) \tag{A.13}$$

For completeness, we add that using this formula, we get that the k-th element of the succession a_k is, for even $l + k$, the following (with k fixed)

$$a_k = \frac{1}{2^l l!}(-1)^{\frac{l-k}{2}}\binom{l}{\frac{1}{2}(l+k)}\prod_{\alpha=1}^{l}(k+\alpha)$$

It's evident that $P_l(\pm 1) = (\pm 1)^l$. For $m \neq 0$ we have that the equation is solvable only for $m \leq l$, and it's solved by the *Legendre functions of the first kind*, defined as follows

$$P_l^m(\xi) = \frac{(1-\xi^2)^{\frac{m}{2}}}{2^l l!}\frac{\mathrm{d}^{l+m}}{\mathrm{d}\xi^{l+m}}\left(\xi^2 - 1\right)^l \tag{A.14}$$

Having finally solved the second equation, we remember that the eigenfunction f is the product of g and h, and it's exactly the spherical harmonics $Y_l^m(\theta, \phi)$, which their normalized counterpart is defined as follows

$$Y_l^m(\theta, \phi) = \sqrt{\frac{2(l+1)(l-m)!}{4\pi(l+m)!}}(-1)^m e^{im\phi}P_l^m(\cos\theta) \tag{A.15}$$

Since it's possible to define $m \leq 0$ but $m \geq -l$, we have this property which is impossible for the associated Legendre Functions, and, we get that

$$Y_l^{-m}(\theta, \phi) = (-1)^m \overline{Y}_l^m(\theta, \phi) \tag{A.16}$$

This condition, finally our searched bounds for m, which can take only values in a set $I \subset \mathbb{Z}$, where $I = [-l, l]$

§§ A.2.2 Confluent Hypergeometric Function and Laguerre Polynomials

The *Confluent Hypergeometric Function* $F(\alpha, \gamma, z)$ is an analytic complex function, defined as follows

$$F(\alpha, \gamma, z) = \sum_{n=0}^{\infty} \frac{(\alpha)_n}{(\gamma)_n} \frac{z^n}{n!} \quad \forall z \in \mathbb{C} \tag{A.17}$$

Where $\alpha \in \mathbb{C}$, $\gamma \in \mathbb{Z}$. The misterious application on α and γ is called the *Pochhammer symbol* and it's defined as follows

$$(a)_n = \begin{cases} 1 & n = 0 \\ \prod_{k=0}^{n-1}(a+k) & n > 0 \end{cases} \tag{A.18}$$

The need for this function comes from the theory of differential equations. Let's define a linear differential operator as follows

$$\hat{L}_F = z\frac{\mathrm{d}^2}{\mathrm{d}z^2} + (\gamma - z)\frac{\mathrm{d}}{\mathrm{d}z} - \alpha \tag{A.19}$$

The only function that solves the equation $\hat{L}_F f(z) = 0$ is a superposition of confluent hypergeometric functions, where

$$f(z) = c_1 F(\alpha, \gamma, z) + c_2 z^{1-\gamma} F(-\alpha + \gamma + 1, \alpha - \gamma, z) \tag{A.20}$$

A curious property is given if $\alpha = -n$ where $n \in \mathbb{N}$. Then, the confluent hypergeometric function is a polynomial, defined as follows

$$F(-n, \gamma, z) = \frac{z^{1-\gamma}e^z}{(\gamma)_n}\frac{\mathrm{d}^n}{\mathrm{d}z^n}\left(e^{-z}z^{\gamma+n-1}\right) \tag{A.21}$$

If we also have $\gamma = m$ and $m \in \mathbb{N}$ then the polynomial is defined differently, as follows

$$F(-n, m, z) = \frac{(-1)^{m+1}e^z}{(m)_n}\frac{\mathrm{d}^{m+n-1}}{\mathrm{d}z^{m+n-1}}\left(e^{-z}z^n\right) \tag{A.22}$$

In the added special case where $0 \leq m \leq n$, we get a particular case, which is fundamental in the physics of hydrogenoid atoms, for which the confluent hypergeometric function

is directly proportional to *Laguerre polynomials*, which describe the radial part of the wavefunction. Hence, we have

$$F(-n, m, z) \propto L_n^m(z)$$

Where, $L_n^m(z)$ are the generalized Laguerre polynomials, and are defined from the confluent hypergeometric function as follows

$$L_n^m(z) = \frac{(-1)^m(n!)^2}{m!(n-m)!}F(-(n-m), m+1, z) \tag{A.23}$$

Substituting this in (A.22), we get the Rodrigues formula for generalized Laguerre polynomials

$$L_n^m(z) = \frac{(-1)^m n!}{(n-m)!}e^z z^{-m}\frac{\mathrm{d}^{n-m}}{\mathrm{d}z^{n-m}}\left(e^{-z}z^n\right) \tag{A.24}$$

For $m = 0$ we get the Laguerre polynomials, which, from the previous formula are then defined as follows

$$L_n(z) = e^z\frac{\mathrm{d}^n}{\mathrm{d}z^n}\left(e^{-z}z^n\right) \tag{A.25}$$

Hence, from the confluent hypergeometric function, we can define both as follows

$$L_n^m(z) = (-1)^m n!\binom{n}{m}F(-n+m, m+1, z)$$

$$L_n(z) = n!F(-n, 1, z) \tag{A.26}$$

§ A.3 Generalized Riemann ζ Functions and Bernoulli Numbers

In quantum statistical mechanics we end up analyzing integrals of the generalized ζ functions, defined as follows

$$\left.\begin{array}{c}g_s(z)\\f_s(z)\end{array}\right\} = \frac{1}{\Gamma(s)}\int_{\mathbb{R}_+}\frac{x^{s-1}}{e^x z^{-1}\mp 1}\,\mathrm{d}x \tag{A.27}$$

It's possible to evaluate the integrals of some values for the two functions through direct evaluation and usage of the properties of the Riemann ζ function, as follows for $g_s(1)$

$$g_s(1) = \frac{1}{\Gamma(s)}\int_{\mathbb{R}_+}\frac{x^{s-1}}{e^x + 1}\,\mathrm{d}x = \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k^s} = \sum_{k=1}^{\infty}\frac{1}{k^s} - 2\sum_{n=1}^{\infty}\frac{1}{(2n)^s}$$

$$g_s(1) = (1 - 2^{1-s})\zeta(s) \tag{A.28}$$

And equivalently for $f_s(1)$

$$f_s(1) = \frac{1}{\Gamma(s)}\int_{\mathbb{R}_+}\frac{x^{s-1}}{e^x - 1}\,\mathrm{d}x = \sum_{k=1}^{\infty}\frac{1}{k^s} = \zeta(s) \tag{A.29}$$

From the residue theorem tho, we can write the integral form of the Riemann ζ

$$\zeta(s) = \frac{1}{4i} \int_\gamma \frac{\cot(\pi z)}{z^s} \, \mathrm{d}z \tag{A.30}$$

Before continuing, we introduce the following formula

$$\frac{1}{2} \cot\left(\frac{z}{2}\right) = 1 - \sum_{n=1}^\infty \frac{B_{2n} z^{2n}}{(2n)!} \tag{A.31}$$

Where B_n are the Bernoulli numbers. Therefore, for even numbers $s = 2k \in \mathbb{Z}$ we can write

$$\zeta(2k) = \frac{(2\pi)^k}{2(2k)!} B_k \tag{A.32}$$

Using then the definitions (A.28) and (A.29), we can directly calculate the following integrals, that often pop out in quantum statistical mechanics calculations

$$g_{2k}(1) = (2^{2k-1} - 1)\zeta(2k) = \pi^{2k}(2^{2k-1} - 1)\frac{B_k}{(2k)!}$$

$$\Gamma(2k)g_{2k}(1) = (2k-1)!\zeta(2k) = \pi^{2k}(2^{2k-1} - 1)\frac{B_{2k}}{2k} \tag{A.33}$$

$$\Gamma(2k)f_{2k}(1) = (2k-1)!\zeta(2k) = \frac{(2\pi)^{2k}}{4k} B_k$$

§ A.4 Saddle Point Method

The saddle point method of approximation of integrals comes from the problem of analyzing the asymptotic behavior of an integral depending from a parameter λ for $\lambda \to \infty$, where the integral is defined as $I(\lambda) : \mathbb{R} \to \mathbb{C}$, with

$$I(\lambda) = \int_\gamma e^{\lambda f(z)} g(z) \, \mathrm{d}z \tag{A.34}$$

Where $f, g : D \to \mathbb{C}$, $f, g \in H(D)$ (they are both holomorphic function), and γ is a piecewise smooth curve such that $\{\gamma\} \subset D$.
Without loss of generality we can take $\lambda \in \mathbb{R}$, $\lambda > 0$.
Before going directly for the complete calculus in $\mathbb{C}$, we begin by supposing that the integral is in a real interval $[a, b]$ and both functions f, g are smooth in this interval.
Supposing that the function f has a maximum in a point $t_0 \in [a, b]$, we approximate around this point with a Taylor expansion at the second order and substitute it back into the integral

$$f(t) = f(t_0) + \frac{1}{2} f''(t_0)(t - t_0)^2 + \mathcal{O}\left((t - t_0)^3\right)$$

And therefore

$$I(\lambda) = \int_a^b e^{\lambda\left(f(t_0)-\frac{1}{2}|f''(t_0)|(t-t_0)^2+\mathcal{O}((t-t_0)^3)\right)}\left(g(t_0) + \mathcal{O}((t-t_0))\right)\,\mathrm{d}t \tag{A.35}$$

Using the substitution $u = \sqrt{\lambda|f''(t_0)|}(t-t_0)$ we obtain

$$I(\lambda) = \frac{e^{\lambda f(t_0)}g(t_0)}{\sqrt{\lambda|f''(t_0)|}}\int_{-(t_0-a)\sqrt{\lambda|f''(t_0)|}}^{(b-t_0)\sqrt{\lambda|f''(t_0)|}} e^{-\frac{1}{2}u^2+\mathcal{O}\left((u(\lambda|f''(t_0)|)^{-\frac{1}{2}})^3\right)}\left(1 + \mathcal{O}\left((u(\lambda|f''(t_0)|)^{-\frac{1}{2}})\right)\right)\,\mathrm{d}t \tag{A.36}$$

Assuming that $g(t_0) \neq 0$ and approximating the integral with $\sqrt{2\pi}$ we obtain that asymptotically we obtain

$$I(\lambda) \simeq \sqrt{\frac{2\pi}{\lambda|f''(t_0)|}}e^{\lambda f(t_0)}g(t_0) \tag{A.37}$$

Without approximating the integral of the exponential, we can impose the variable substitution $u = f(t_0) - f(t)$ and evaluating in the intervals $[a, t_0) \cup (t_0, b]$ where the function $f(t)$ is strictly monotonous, we have that

$$I(\lambda) = e^{\lambda f(t_0)}\left(\int_0^{f(t_0)-f(b)} G_+(u)e^{-\lambda u}\,\mathrm{d}(u) - \int_0^{f(t_0)-f(a)} G_-(u)e^{-\lambda u}\,\mathrm{d}u\right) \tag{A.38}$$

Where, the two functions $G_\pm$ are defined as follows

$$G_\pm(u) = \left.\frac{g(t)}{f'(t)}\right|_{t=f_\pm^{-1}(f(t_0)-u)}$$

Expanding the function g/f' we get

$$\frac{g(t)}{f'(t)} = \frac{g(t_0)}{f''(t_0)}(t-t_0)^{-1} + \left(\frac{g'(t_0)}{f''(t_0)} - \frac{g(t_0)f'''(t_0)}{2(f''(t_0))^2}\right) + \mathcal{O}(t-t_0) \tag{A.39}$$

Therefore, inserting the previous u-substitution, we get that $(t-t_0) = \pm\sqrt{2u/|f''(t_0)|}$ and therefore we get

$$G_\pm(u) = \mp\frac{g(t_0)}{\sqrt{2|f''(t_0)|}}u^{-\frac{1}{2}} + \left(\frac{g'(t_0)}{f''(t_0)} - \frac{g(t_0)f'''(t_0)}{2\left(f'(t_0)\right)^2}\right) + \mathcal{O}(\sqrt{u}) \tag{A.40}$$

For $\lambda \to \infty$ we then obtain

$$\int_0^{f(t_0)-a_i} G_\pm(u)e^{-\lambda u}\,\mathrm{d}u = \pm\frac{g(t_0)}{\sqrt{2|f''(t_0)|}}\sqrt{\frac{\pi}{\lambda}} + \frac{1}{\lambda}\left(\frac{g'(t_0)}{f''(t_0)} - \frac{g(t_0)f'''(t_0)}{2\left(f''(t_0)\right)^2}\right) + \mathcal{O}\left(\lambda^{-\frac{3}{2}}\right) \tag{A.41}$$

In conclusion, we obtain

$$I(\lambda) = \sqrt{\frac{2\pi}{\lambda|f''(t_0)|}}e^{\lambda f(t_0)}g(t_0) + \mathcal{O}\left(\lambda^{-\frac{3}{2}}e^{\lambda f(t_0)}\right) \tag{A.42}$$

Now, going back to the complex case, we have parameterizing the curve γ, that

$$I(\lambda) = \int_a^b e^{\lambda f(\gamma(t))} g(\gamma(t)) \gamma'(t) \, \mathrm{d}t \tag{A.43}$$

In order to make sure that the saddle point approximation is applicable, we chose a specific parametrization of this curve, which must be inside an open ball from a simple critical point (saddle point) $z_0 \in D$ of the function $f(z)$. We can therefore choose two curves, one of "steepest climb" (γ_+) and one of "steepest descent" (γ_-), orthogonal to each other and defined as follows

$$\begin{cases} \gamma_+(t) = z_0 + (t - t_0)e^{-\frac{i}{2}\mathrm{Arg}f''(z_0)} & t_0 - \epsilon < t < t_0 + \epsilon \\ \gamma_-(t) = z_0 + (t - t_0)e^{i\left(\frac{\pi}{2} - \frac{1}{2}\mathrm{Arg}f''(z_0)\right)} & t_0 - \epsilon < t < t_0 + \epsilon \end{cases} \tag{A.44}$$

In these curves, we have that

$$\mathfrak{Re}\, f(\gamma_\pm(t)) = \mathfrak{Re}\, f(z_0) \pm \frac{1}{2}\left|f''(z_0)\right|(t - t_0)^2 + \mathcal{O}\left((t - t_0)^3\right)$$

$$\mathfrak{Im}\, f(\gamma_\pm(t)) = \mathfrak{Im}\, f(z_0) + \mathcal{O}\left((t - t_0)^3\right)$$

Using an homotopy transformation we can map $\gamma \to \eta$, where η is a new curve that coincides with γ_- for $t \in [t_0 - \epsilon/2, t_0 + \epsilon/2]$, coincides with γ outside the open ball $B_\epsilon(z_0)$ and joins back with γ on the frontier of the ball.
We then get

$$I(\lambda) = \int_\eta e^{\lambda f(z)} g(z) \, \mathrm{d}z =$$

$$= \int_a^{t_0 - \frac{\epsilon}{2}} e^{\lambda f(\eta(t))} g(\eta(t)) \eta'(t) \, \mathrm{d}t + \int_{t_0 - \frac{\epsilon}{2}}^{t_0 + \frac{\epsilon}{2}} e^{\lambda f(\gamma_-(t))} g(\gamma_-(t)) \gamma_-'(t) \, \mathrm{d}t \tag{A.45}$$

$$+ \int_{t_0 + \frac{\epsilon}{2}}^b e^{\lambda f(\eta(t))} g(\eta(t)) \eta'(t) \, \mathrm{d}t$$

Applicating now the saddle point approximation, we get that for $\lambda \to \infty$

$$I(\lambda) = \sqrt{\frac{2\pi}{\lambda|f''(z_0)|}} e^{\lambda f(z_0)} g(z_0) e^{i\left(\frac{\pi}{2} - \frac{1}{2}\mathrm{Arg}f''(z_0)\right)} + \mathcal{O}\left(\lambda^{-\frac{3}{2}} e^{\lambda f(z_0)}\right) \tag{A.46}$$

If we have more than one simple critical point for $f(z)$ along γ, the solution will be a sum of terms of the kind (A.46)

B Clebsch-Gordan Table

In this appendix, we give the table of Clebsch-Gordan coefficients used in the addition of angular momenta, in order to switch from a common basis between $\hat{J}_1^2, \hat{J}_2^2, \hat{J}_{z1}, \hat{J}_{z2}$ to the new "total" common basis of $\hat{J}_1^2, \hat{J}_2^2, \hat{J}^2, \hat{J}_z$, where the total angular momentum operator is defined as follows

$$\underline{\hat{J}} = \underline{\hat{J}}_1 \otimes \hat{\mathbb{1}} + \hat{\mathbb{1}} \otimes \underline{\hat{J}}_2$$
$$\hat{J}_z = \hat{J}_{z1} \otimes \hat{\mathbb{1}} + \hat{\mathbb{1}} \otimes \hat{J}_{z2} \tag{B.1}$$
$$\hat{J}^2 = \left(\underline{\hat{J}}_1 \otimes \hat{\mathbb{1}} + \hat{\mathbb{1}} \otimes \underline{\hat{J}}_2 \right)^2$$

Its eigenvalues are the following

$$\hat{J}^2 \left| j_1, j_2, j, m \right\rangle = \hbar^2 j(j+1) \left| j_1, j_2, j, m \right\rangle$$
$$\hat{J}_z \left| j_1, j_2, j, m \right\rangle = \hbar m \left| j_1, j_2, j, m \right\rangle \tag{B.2}$$

And are subject to the following constraints

$$m = m_1 + m_2, \quad -j \leq m \leq j$$
$$0 \leq |j_1 - j_2| \leq j \leq j_1 + j_2 \tag{B.3}$$

Note: A square-root sign is to be understood over *every* coefficient, *e.g.*, for $-8/15$ read $-\sqrt{8/15}$.

Notation:

J	J	...
M	M	...
m_1	m_2	
m_1	m_2	Coefficients
.	.	
.	.	

$$Y_1^0 = \sqrt{\frac{3}{4\pi}}\cos\theta$$

$$Y_1^1 = -\sqrt{\frac{3}{8\pi}}\sin\theta\, e^{i\phi}$$

$$Y_2^0 = \sqrt{\frac{5}{4\pi}}\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)$$

$$Y_2^1 = -\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta\, e^{i\phi}$$

$$Y_2^2 = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\sin^2\theta\, e^{2i\phi}$$

$$Y_\ell^{-m} = (-1)^m Y_\ell^{m*}$$

$$d_{m,0}^\ell = \sqrt{\frac{4\pi}{2\ell+1}}\,Y_\ell^m\, e^{-im\phi}$$

$$\langle j_1 j_2 m_1 m_2 | j_1 j_2 JM\rangle = (-1)^{J-j_1-j_2}\langle j_2 j_1 m_2 m_1 | j_2 j_1 JM\rangle$$

$$d_{m',m}^j = (-1)^{m-m'} d_{m,m'}^j = d_{-m,-m'}^j$$

$$d_{0,0}^1 = \cos\theta \qquad d_{1/2,1/2}^{1/2} = \cos\frac{\theta}{2} \qquad d_{1,1}^1 = \frac{1+\cos\theta}{2}$$

$$d_{1/2,-1/2}^{1/2} = -\sin\frac{\theta}{2} \qquad d_{1,0}^1 = -\frac{\sin\theta}{\sqrt{2}}$$

$$d_{1,-1}^1 = \frac{1-\cos\theta}{2}$$

$$d_{3/2,3/2}^{3/2} = \frac{1+\cos\theta}{2}\cos\frac{\theta}{2}$$

$$d_{3/2,1/2}^{3/2} = -\sqrt{3}\,\frac{1+\cos\theta}{2}\sin\frac{\theta}{2} \qquad d_{2,2}^2 = \left(\frac{1+\cos\theta}{2}\right)^2$$

$$d_{3/2,-1/2}^{3/2} = \sqrt{3}\,\frac{1-\cos\theta}{2}\cos\frac{\theta}{2} \qquad d_{2,1}^2 = -\frac{1+\cos\theta}{2}\sin\theta$$

$$d_{3/2,-3/2}^{3/2} = -\frac{1-\cos\theta}{2}\sin\frac{\theta}{2} \qquad d_{2,0}^2 = \frac{\sqrt{6}}{4}\sin^2\theta \qquad d_{1,1}^2 = \frac{1+\cos\theta}{2}(2\cos\theta - 1)$$

$$d_{1/2,1/2}^{3/2} = \frac{3\cos\theta - 1}{2}\cos\frac{\theta}{2} \qquad d_{2,-1}^2 = -\frac{1-\cos\theta}{2}\sin\theta \qquad d_{1,0}^2 = -\sqrt{\frac{3}{2}}\sin\theta\cos\theta$$

$$d_{1/2,-1/2}^{3/2} = -\frac{3\cos\theta + 1}{2}\sin\frac{\theta}{2} \qquad d_{2,-2}^2 = \left(\frac{1-\cos\theta}{2}\right)^2 \qquad d_{1,-1}^2 = \frac{1-\cos\theta}{2}(2\cos\theta + 1) \qquad d_{0,0}^2 = \left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)$$

Clebsch-Gordan coefficient tables for:
$1/2 \times 1/2$, $1 \times 1/2$, 2×1, 1×1, $2 \times 1/2$, $3/2 \times 1/2$, $3/2 \times 1$, $3/2 \times 3/2$, $2 \times 3/2$, 2×2.

C Tensor Spherical Harmonics

In this appendix, we will treat the particular theme of spin-orbit angular momenta addition, and tensor spherical harmonics.
We define our total angular momentum as follows

$$\hat{\underline{J}} = \hat{\underline{L}} \otimes \hat{\mathbb{1}} + \hat{\mathbb{1}} \otimes \hat{\underline{S}} \tag{C.1}$$

We then define our common basis between these operators and their projections, as in the general rules of angular momenta addition, using Clebsch-Gordan coefficients

$$|lsjm\rangle = \sum_{m_l=-l}^{l} \sum_{m_s=-s}^{s} \langle lsm_l m_s | lsjm \rangle \, |lsm_l m_s\rangle$$

By definition we then have that the eigenvalues of $\hat{J}^2, \hat{J}_z$ can then take the following values

$$|l - s| \leq j \leq l + s$$
$$m = m_l + m_s$$

We then recall that, in Schrödinger representation, the eigenfunctions of $\hat{\underline{L}}$ and $\hat{\underline{S}}$ are the following

$$\begin{cases} \langle x_i | \hat{S}^2 | sm_s\rangle = \hbar^2 s(s+1) \langle x_i | sm_s\rangle & \langle x_i | \hat{S}_z | sm_s\rangle = \hbar m_s \langle x_i | sm_s\rangle \\ \langle x_i | \hat{L}^2 | lm_l\rangle = \hbar^2 l(l+1) \langle x_i | lm_l\rangle & \langle x_i | \hat{L}_z | lm_l\rangle = \hbar m_l \langle x_i | lm_l\rangle \end{cases} \tag{C.2}$$

Where $\langle x_i | sm_s\rangle = \chi_{sm_s}$ is the basis spinor and $\langle x_i | lm_l\rangle = Y_l^{m_l}(\theta, \phi)$ are the spherical harmonics.
In the new basis of common eigenvectors of $\hat{J}^2, \hat{J}_z, \hat{L}^2, \hat{S}^2$ we then define the *tensor spherical harmonics* $\langle x_i | lsjm \rangle = \mathcal{Y}_{jm}^{ls}(\theta, \phi)$ as follows

$$\begin{cases} \hat{J}^2 \mathcal{Y}_{jm}^{ls}(\theta, \phi) = \hbar^2 j(j+1) \mathcal{Y}_{jm}^{ls}(\theta, \phi) & \hat{J}_z \mathcal{Y}_{jm}^{ls}(\theta, \phi) = \hbar m \mathcal{Y}_{jm}^{ls}(\theta, \phi) \\ \hat{L}^2 \mathcal{Y}_{jm}^{ls}(\theta, \phi) = \hbar^2 l(l+1) \mathcal{Y}_{jm}^{ls}(\theta, \phi) & \hat{S}^2 \mathcal{Y}_{jm}^{ls}(\theta, \phi) = \hbar^2 s(s+1) \mathcal{Y}_{jm}^{ls}(\theta, \phi) \end{cases} \tag{C.3}$$

By definition then, we can define the tensor spherical harmonics as follows

$$|lsjm\rangle = \sum_{m_l=-l}^{l} \sum_{m_s=-s}^{s} \langle lsm_lm_s|jm\rangle \, |lm_l\rangle \otimes |sm_s\rangle$$

$$\mathcal{Y}_{jm}^{ls}(\theta,\phi) = \sum_{m_l=-l}^{l} \sum_{m_s=-s}^{s} \langle lsm_lm_s|jm\rangle \, \langle x_i|lm_l\rangle \otimes \langle x_i|sm_s\rangle$$

(C.4)

Or, indicating the Clebsch-Gordan coefficients as $C_{jm}^{lsm_lm_s}$

$$\mathcal{Y}_{jm}^{ls}(\theta,\phi) = C_{jm}^{lsm_lm_s} Y_l^{m_l}(\theta,\phi)\chi_{sm_s}$$

§ C.1 Spin $1/2$ System

The overall problem simplifies enormously for $s = \frac{1}{2}$ systems. The permitted values of j are $l + 1/2$ and $l - 1/2$ and, the Clebsch-Gordan coefficients, are then simply the following

j	$m_s = \frac{1}{2}$	$m_s = -\frac{1}{2}$
$l + \frac{1}{2}$	$\sqrt{\frac{l+m+\frac{1}{2}}{2l+1}}$	$\sqrt{\frac{l+m+\frac{1}{2}}{2l+1}}$
$l - \frac{1}{2}$	$-\sqrt{\frac{l+m+\frac{1}{2}}{2l+1}}$	$\sqrt{\frac{l+m+\frac{1}{2}}{2l+1}}$

$$\mathcal{Y}_{l\pm\frac{1}{2},m}^{l\frac{1}{2}}(\theta,\phi) = \frac{1}{\sqrt{2l+1}}\begin{pmatrix} \pm\sqrt{l\pm m+\frac{1}{2}}Y_l^{m-\frac{1}{2}}(\theta,\phi) \\ \sqrt{l\mp m+\frac{1}{2}}Y_l^{m+\frac{1}{2}}(\theta,\phi) \end{pmatrix}$$

(C.5)

Which can then be manipulated in order to get various useful informations, as spin-orbit coupling eigenvalues, which appear in atomic physics in the relativistic approximation of hydrogenoid atoms.

D Calculus of $\left\langle r^k \right\rangle_{lnm}$ Integrals

We know already that the wavefunction for a Hydrogen atom (non-normalized) is

$$\psi_{nlm}(x_i) = R_{nl}(r)Y_l^m(\theta, \phi) \tag{D.1}$$

Setting $a_0 = \frac{4\hbar^2 \pi \epsilon_0}{me^2}$ as our Bohr radius, we get that R_{nl} for $1s, 2s, 2p$ states are

$$R_{10} = 2 \left(\frac{a_0}{Z}\right)^{-\frac{3}{2}} e^{-\frac{Zr}{a_0}}$$

$$R_{20} = 2 \left(\frac{2a_0}{Z}\right)^{-\frac{3}{2}} \left(1 - \frac{Zr}{2a_0}\right) e^{-\frac{Zr}{2a_0}} \tag{D.2}$$

$$R_{21} = \sqrt{\frac{1}{3}} \frac{Zr}{a_0} \left(\frac{2a_0}{Z}\right)^{-\frac{3}{2}} e^{-\frac{Zr}{2a_0}}$$

Hence, polynomials times an exponential with r/a_0 as a variable. The integral we have to calculate will then be

$$\left\langle r^d \right\rangle_{nlm} = \int_0^\infty r^{d+2} |R_{nl}(r)|^2 \, \mathrm{d}r \tag{D.3}$$

Which are of the kind

$$I(k, p) = \int_0^\infty r^k e^{-\frac{Zpr}{a_0}} \, \mathrm{d}r \tag{D.4}$$

Where $k, p \in \mathbb{N}$. Using a coordinate transformation $z = Zpr/a_0$ we have that the integral reduces to a Euler Gamma integral for whole numbers

$$I(k, p) = \left(\frac{a_0}{Zp}\right)^{k+1} \int_0^\infty z^k e^{-z} \, \mathrm{d}z = \left(\frac{a_0}{Zp}\right)^{k+1} k! \tag{D.5}$$

With this in mind, we calculate the especially useful expectation values for r^{-i}, with $i = 1, 2, 3$ on $1s, 2s, 2p$ states of the Hydrogen atom, for which $Z = 1$. These expectation values pop

up when evaluating relativistic corrections for atoms.

$$\left\langle \frac{1}{r} \right\rangle_{1s} = \frac{1}{a_0}$$

$$\left\langle \frac{1}{r} \right\rangle_{2s} = \frac{1}{2a_0^3} \left(I(1,1) - \frac{1}{a_0} I(2,1) + \frac{1}{4a_0^2} I(3,1) \right) = \frac{1}{4a_0}$$

$$\left\langle \frac{1}{r} \right\rangle_{2p} = \frac{1}{24a_0^5} I(3,1) = \frac{1}{4a_0}$$

(D.6)

Lowering k by one we obtain the expectation values for r^{-2}

$$\left\langle \frac{1}{r^2} \right\rangle_{1s} = \frac{2}{a_0}$$

$$\left\langle \frac{1}{r^2} \right\rangle_{2s} = \frac{1}{2a_0^3} \left(I(0,1) - \frac{1}{a_0} I(1,1) + \frac{1}{4a_0} I(2,1) \right) = \frac{1}{4a_0^2}$$

$$\left\langle \frac{1}{r^2} \right\rangle_{2p} = \frac{1}{24a_0^5} I(2,1) = \frac{1}{12a_0^2}$$

(D.7)

For r^{-3} we will evaluate only the level $2p$, since this expectation value pops up with spin orbit interaction, which is null for ns states

$$\left\langle \frac{1}{r^3} \right\rangle_{2p} = \frac{1}{24a_0^5} I(1,1) = \frac{1}{24a_0^3}$$

(D.8)

E Periodic Table

Periodic Table of the Elements

Key:
- Atomic Number
- Atomic Mass
- Symbol
- Name
- Electron Configuration

Z	Mass	Symbol	Name	Configuration
1	1.008	H	Hydrogen	$1s^1$
2	4.003	He	Helium	$1s^2$
3	6.941	Li	Lithium	$[He]2s^1$
4	9.012	Be	Beryllium	$[He]2s^2$
5	10.811	B	Boron	$[He]2s^22p^1$
6	12.011	C	Carbon	$[He]2s^22p^2$
7	14.007	N	Nitrogen	$[He]2s^22p^3$
8	15.999	O	Oxygen	$[He]2s^22p^4$
9	18.998	F	Fluorine	$[He]2s^22p^5$
10	20.180	Ne	Neon	$[He]2s^22p^6$
11	22.990	Na	Sodium	$[Ne]3s^1$
12	24.305	Mg	Magnesium	$[Ne]3s^2$
13	26.982	Al	Aluminum	$[Ne]3s^23p^1$
14	28.086	Si	Silicon	$[Ne]3s^23p^2$
15	30.974	P	Phosphorus	$[Ne]3s^23p^3$
16	32.066	S	Sulfur	$[Ne]3s^23p^4$
17	35.453	Cl	Chlorine	$[Ne]3s^23p^5$
18	39.948	Ar	Argon	$[Ne]3s^23p^6$
19	39.098	K	Potassium	$[Ar]4s^1$
20	40.078	Ca	Calcium	$[Ar]4s^2$
21	44.956	Sc	Scandium	$[Ar]3d^14s^2$
22	47.88	Ti	Titanium	$[Ar]3d^24s^2$
23	50.942	V	Vanadium	$[Ar]3d^34s^2$
24	51.996	Cr	Chromium	$[Ar]3d^54s^1$
25	54.938	Mn	Manganese	$[Ar]3d^54s^2$
26	55.845	Fe	Iron	$[Ar]3d^64s^2$
27	58.933	Co	Cobalt	$[Ar]3d^74s^2$
28	58.693	Ni	Nickel	$[Ar]3d^84s^2$
29	63.546	Cu	Copper	$[Ar]3d^{10}4s^1$
30	65.38	Zn	Zinc	$[Ar]3d^{10}4s^2$
31	69.723	Ga	Gallium	$[Ar]3d^{10}4s^24p^1$
32	72.631	Ge	Germanium	$[Ar]3d^{10}4s^24p^2$
33	74.922	As	Arsenic	$[Ar]3d^{10}4s^24p^3$
34	78.971	Se	Selenium	$[Ar]3d^{10}4s^24p^4$
35	79.904	Br	Bromine	$[Ar]3d^{10}4s^24p^5$
36	84.798	Kr	Krypton	$[Ar]3d^{10}4s^24p^6$
37	84.468	Rb	Rubidium	$[Kr]5s^1$
38	87.62	Sr	Strontium	$[Kr]5s^2$
39	88.906	Y	Yttrium	$[Kr]4d^15s^2$
40	91.224	Zr	Zirconium	$[Kr]4d^25s^2$
41	92.906	Nb	Niobium	$[Kr]4d^45s^1$
42	95.95	Mo	Molybdenum	$[Kr]4d^55s^1$
43	98.907	Tc	Technetium	$[Kr]4d^55s^2$
44	101.07	Ru	Ruthenium	$[Kr]4d^75s^1$
45	102.906	Rh	Rhodium	$[Kr]4d^85s^1$
46	106.42	Pd	Palladium	$[Kr]4d^{10}$
47	107.868	Ag	Silver	$[Kr]4d^{10}5s^1$
48	112.414	Cd	Cadmium	$[Kr]4d^{10}5s^2$
49	114.818	In	Indium	$[Kr]4d^{10}5s^25p^1$
50	118.711	Sn	Tin	$[Kr]4d^{10}5s^25p^2$
51	121.760	Sb	Antimony	$[Kr]4d^{10}5s^25p^3$
52	127.6	Te	Tellurium	$[Kr]4d^{10}5s^25p^4$
53	126.904	I	Iodine	$[Kr]4d^{10}5s^25p^5$
54	131.249	Xe	Xenon	$[Kr]4d^{10}5s^25p^6$
55	132.905	Cs	Cesium	$[Xe]6s^1$
56	137.328	Ba	Barium	$[Xe]6s^2$
57	138.905	La	Lanthanum	$[Xe]5d^16s^2$
58	140.116	Ce	Cerium	$[Xe]4f^15d^16s^2$
59	140.908	Pr	Praseodymium	$[Xe]4f^36s^2$
60	144.243	Nd	Neodymium	$[Xe]4f^46s^2$
61	144.913	Pm	Promethium	$[Xe]4f^56s^2$
62	150.36	Sm	Samarium	$[Xe]4f^66s^2$
63	151.964	Eu	Europium	$[Xe]4f^76s^2$
64	157.25	Gd	Gadolinium	$[Xe]4f^75d^16s^2$
65	158.925	Tb	Terbium	$[Xe]4f^96s^2$
66	162.500	Dy	Dysprosium	$[Xe]4f^{10}6s^2$
67	164.930	Ho	Holmium	$[Xe]4f^{11}6s^2$
68	167.259	Er	Erbium	$[Xe]4f^{12}6s^2$
69	168.934	Tm	Thulium	$[Xe]4f^{13}6s^2$
70	173.055	Yb	Ytterbium	$[Xe]4f^{14}6s^2$
71	174.967	Lu	Lutetium	$[Xe]4f^{14}5d^16s^2$
72	178.49	Hf	Hafnium	$[Xe]4f^{14}5d^26s^2$
73	180.948	Ta	Tantalum	$[Xe]4f^{14}5d^36s^2$
74	183.84	W	Tungsten	$[Xe]4f^{14}5d^46s^2$
75	186.207	Re	Rhenium	$[Xe]4f^{14}5d^56s^2$
76	190.23	Os	Osmium	$[Xe]4f^{14}5d^66s^2$
77	192.217	Ir	Iridium	$[Xe]4f^{14}5d^76s^2$
78	195.085	Pt	Platinum	$[Xe]4f^{14}5d^96s^1$
79	196.967	Au	Gold	$[Xe]4f^{14}5d^{10}6s^1$
80	200.592	Hg	Mercury	$[Xe]4f^{14}5d^{10}6s^2$
81	204.383	Tl	Thallium	$[Xe]4f^{14}5d^{10}6s^26p^1$
82	207.2	Pb	Lead	$[Xe]4f^{14}5d^{10}6s^26p^2$
83	208.980	Bi	Bismuth	$[Xe]4f^{14}5d^{10}6s^26p^3$
84	208.982	Po	Polonium	$[Xe]4f^{14}5d^{10}6s^26p^4$
85	209.987	At	Astatine	$[Xe]4f^{14}5d^{10}6s^26p^5$
86	222.018	Rn	Radon	$[Xe]4f^{14}5d^{10}6s^26p^6$
87	223.020	Fr	Francium	$[Rn]7s^1$
88	226.025	Ra	Radium	$[Rn]7s^2$
89	227.028	Ac	Actinium	$[Rn]6d^17s^2$
90	232.038	Th	Thorium	$[Rn]6d^27s^2$
91	231.036	Pa	Protactinium	$[Rn]5f^26d^17s^2$
92	238.029	U	Uranium	$[Rn]5f^36d^17s^2$
93	237.048	Np	Neptunium	$[Rn]5f^46d^17s^2$
94	244.064	Pu	Plutonium	$[Rn]5f^67s^2$
95	243.061	Am	Americium	$[Rn]5f^77s^2$
96	247.070	Cm	Curium	$[Rn]5f^76d^17s^2$
97	247.070	Bk	Berkelium	$[Rn]5f^97s^2$
98	251.080	Cf	Californium	$[Rn]5f^{10}7s^2$
99	254	Es	Einsteinium	$[Rn]5f^{11}7s^2$
100	257.095	Fm	Fermium	$[Rn]5f^{12}7s^2$
101	258.1	Md	Mendelevium	$[Rn]5f^{13}7s^2$
102	259.101	No	Nobelium	$[Rn]5f^{14}7s^2$
103	262	Lr	Lawrencium	$[Rn]5f^{14}6d^17s^2$
104	261	Rf	Rutherfordium	$[Rn]5f^{14}6d^27s^2$
105	262	Db	Dubnium	$[Rn]5f^{14}6d^37s^2$
106	266	Sg	Seaborgium	$[Rn]5f^{14}6d^47s^2$
107	264	Bh	Bohrium	$[Rn]5f^{14}6d^57s^2$
108	269	Hs	Hassium	$[Rn]5f^{14}6d^67s^2$
109	268	Mt	Meitnerium	$[Rn]5f^{14}6d^77s^2$
110	269	Ds	Darmstadtium	$[Rn]5f^{14}6d^87s^2$
111	272	Rg	Roentgenium	$[Rn]5f^{14}6d^97s^2$
112	277	Cn	Copernicium	$[Rn]5f^{14}6d^{10}7s^2$
113	unknown	Uut	Ununtrium	$[Rn]5f^{14}6d^{10}7s^27p^1$
114	289	Fl	Flerovium	$[Rn]5f^{14}6d^{10}7s^27p^2$
115	289	Uup	Ununpentium	$[Rn]5f^{14}6d^{10}7s^27p^3$
116	unknown	Lv	Livermorium	$[Rn]5f^{14}6d^{10}7s^27p^4$
117	unknown	Uus	Ununseptium	$[Rn]5f^{14}6d^{10}7s^27p^5$
118	unknown	Uuo	Ununoctium	$[Rn]5f^{14}6d^{10}7s^27p^6$

Lanthanide Series: 57–71
Actinide Series: 89–103

Group labels: 1 IA 1A · 2 IIA 2A · 3 IIIB 3B · 4 IVB 4B · 5 VB 5B · 6 VIB 6B · 7 VIIB 7B · 8 VIII 8 · 9 VIII 8 · 10 VIII 8 · 11 IB 1B · 12 IIB 2B · 13 IIIA 3A · 14 IVA 4A · 15 VA 5A · 16 VIA 6A · 17 VIIA 7A · 18 VIIIA 8A

F Basic Chemistry

§ F.1 Basic Notions

§§ F.1.1 Moles and Molar Mass

Given some chemical compound $X = M_aN_b$ it's possible to find the molar mass M_X by summing the single masses of the atomic components. The masses are found in the periodic table (E) precisely on the top right of the box of the element considered, and is expressed in g/mol. Therefore, for the compound X the total molar mass is

$$M_X = am_M + bm_N \tag{F.1}$$

Where m_M, m_N are the masses of the single atoms. Note that for compounds with more than two composing elements, this calculation easily generalizes to a simple sum.
Given an amount m_X of the compound, the total amount of moles n_X contained is

$$n_X = \frac{m_X}{M_X} \tag{F.2}$$

And the total number of molecules is

$$p = n_X N_A \tag{F.3}$$

Where N_A is the Avogadro number, $N_A = 6.022 \cdot 10^{23}$ mol^{-1}

Example F.1.1. Given $m = 12.49$ g of $C_3H_8O_3$, find the amount of moles and how many particles of the compound are contained in this mass.

$$m_C = 12.011 \text{ g/mol}$$
$$m_H = 1.008 \text{ g/mol}$$
$$m_O = 15.999 \text{ g/mol}$$

Therefore

$$M = 3m_C + 8m_H + 3m_O = 92.094 \text{ g/mol}$$

Which means

$$n = \frac{m}{M} = \frac{12.49}{92.094} \text{ mol} = 0.136 \text{ mol}$$

The number of particles is then

$$p = nN_A = 8.19 \cdot 10^{22}$$

§§ F.1.2 Elemental Analysis

If the compound itself is unknown, with some data it's possible to find the chemical formula for the compound.

Suppose that the unknown compound X is composed by $n-$ secondary compounds in unknown quantities, say two without loss of generality $X = A_x B_y$. If the percentage concentration of the two elements is known, we have that for 100 g of compound there will be exactly $\%_A, \%_B$ grams of the single elements.

The two unknown coefficients x, y then can be determined through a *minimal combination* of the moles of the single atoms. We have then, that in 100 g of X there will be

$$n_A = \frac{\%_A}{m_A}$$
$$n_B = \frac{\%_B}{m_B} \tag{F.4}$$

The minimal combination will then made by taking $n_m = \min\{n_A, n_B\}$, and the coefficients will be

$$x = \left\lceil \frac{n_A}{n_m} \right\rceil$$
$$y = \left\lceil \frac{n_B}{n_m} \right\rceil \tag{F.5}$$

Note that if the value inside the rounding operator isn't close to a natural number r ($|n_i/n_m - r| > 0.25$) the coefficient can't be properly determined. The solution is found by taking a number $k \in \mathbb{N}$ and multiplying all the values inside the rounding operator. This shall be repeated until there aren't any dubious values.

Example F.1.2 (An Organic Compound). An unknown composite X is experimentally known to contain C, H, O with the following percentages:

$$\%_C = 55.81$$
$$\%_H = 7.02$$
$$\%_O = 37.17$$

Find the chemical formula for the compound $C_a H_b O_c$ We have that in 100 g of X

$$n_C = \frac{\%_C}{m_C} = \frac{55.81 \text{ g}}{12.01 \text{ g/mol}} = 4.65 \text{ mol}$$
$$n_H = \frac{\%_H}{m_H} = \frac{7.02 \text{ g}}{1.01 \text{ g/mol}} = 6.96 \text{ mol}$$
$$n_O = \frac{\%_O}{m_O} = \frac{37.17 \text{ g}}{16.00 \text{ g/mol}} = 2.32 \text{ mol}$$

And therefore $\min\{n_C, n_H, n_O\} = n_O$ and the minimal combination gives the following coefficients

$$a = \left\lceil \frac{n_C}{n_O} \right\rceil = 2$$

$$b = \left\lceil \frac{n_H}{n_O} \right\rceil = 3$$

$$c = 1$$

The complete formula is therefore C_2H_3O

G Symmetry and Point Groups

We can categorize symmetry transformations in various ways, but principally, we have 3 fundamental transformations

1. Rotations through an axis

2. Reflections on a plane

3. Parallel displacements

It's obvious that the last kind of transformation, in our field, will have sense only in infinite mediums like lattices and solids, hence, for molecules, we'll be interested principally in the first two: *rotations* and *reflections*.

Let's start by defining a rotation operator $\hat{C}_n$, where $n \in \mathbb{N}$. The whole number n is called *order of symmetry* of the considered axis, which means that the system will be again in the initial transformation after n rotations of $\alpha = 2\pi/n$ degrees.

We can immediately determine two fundamental properties of the rotation operator from this

$$\hat{C}_1 = 1$$

$$\hat{C}_n^n = \hat{\mathbb{1}}$$

The second operator we will define, is the plane reflection operator $\hat{\sigma}$, which operates through a reflection of the system on a predetermined plane, and finally we can define the inversion operator $\hat{I}$, which corresponds to a complete inversion of the coordinate system used.

§ G.1 Group Theory

In order to delve deeper into the theory of symmetry, we need to define what a group is mathematically and how does it work, since symmetries of the system arise from the invariance of the Hamiltonian to transformations pertaining to one of these groups.

So let's begin by calling the set G a group. In order to be such, it has to have the following properties

1. Identity: $\hat{\mathbb{1}} \in G$

339

2. Associativity: $(\hat{a}\hat{b})\hat{c} = \hat{a}(\hat{b}\hat{c}) \; \forall \hat{a}, \hat{b}, \hat{c} \in G$

3. Non Commutativity: $\hat{a}\hat{b} \neq \hat{b}\hat{a} \; \forall \hat{a}, \hat{b} \in G$

4. Existence of the inverse: $\forall \hat{a} \in G \; \exists! \hat{a}^{-1} \in G : \hat{a}^{-1}\hat{a} = \hat{1} = \hat{a}\hat{a}^{-1}$

5. $\forall \hat{a}, \hat{b} \in G \; (\hat{a}\hat{b})^{-1} = \hat{b}^{-1}\hat{a}^{-1}$

A group is said abelian, if and only if $\forall \hat{a}\hat{b} \in G$, $\hat{a}\hat{b} = \hat{b}\hat{a}$, hence it's *commutative*.
From an abelian group, one can define a cyclic group, which is an abelian group for which $\forall \hat{a} \in G, \exists n \in \mathbb{N} : \hat{a}^n = \hat{1}$. The integer n is called the *order* of the group G, and it's usually indicated as $\{G\}$.

Definition G.1.1 (Finite Group). A finite group is a group G for which there exists a finite number of elements.

Definition G.1.2 (Subgroup). A subset $H \subset G$ is a subgroup, if and only if has all the properties of a group.

Definition G.1.3 (Complex). Let $H \subset G$ be a subgroup of a finite group (hence, a finite subgroup). Since $H \subset G$, there exists a finite number of elements $G_i \in G$, for which $G_i \notin H$. We can then define a new element of G which is not an element of H by multiplying all elements of H with a single G_i.
The new algebraic construction is called *complex*.
In general, if G is a group, and $L, H \subset G$ are subgroups of G, for which $L = G \setminus H$, $\forall \hat{l}_i \in L, \; \hat{h} \in H$, the product $\hat{h}\hat{l}_i$ generates another n subgroups of G called *complexes*.
Since these all are finite groups by hypothesis, if $\{L\} = l, \; \{H\} = h$, then $\{G\} = \{H\}\{L\}$.

Definition G.1.4 (Direct Product). A direct product of two groups A, B, indicated as $A \otimes B$, is defined as follows

$$A \otimes B := \left\{ \hat{a} \in A, \hat{b} \in B : \; \hat{a}\hat{b} \in A \otimes B \right\}$$

We can now start talking about *molecular point groups*. We start by defining rotation groups.
C_n **group**
The group C_n is the group of symmetries around a rotation axis of the n-th order.

S_{2n} **group**
The S_{2n} group is the *rotoreflection* group of a single axis of the n-th order. Two special cases of this group are the S_2 and S_{4p+2} group, which are

$$S_2 := \{\hat{1}, \hat{I}\} = C_i$$
$$S_{4p+2} = C_{2p+1} \otimes C_i$$

C_{nh} group

The C_{nh} group is the group of simmetries around an axis of the n-th order, coupled with a symmetry through a plane perpendicular to the axis. This group can be seen as a complex created from n $\hat{C}_n$ elements and one $\hat{\sigma}_h$ element as follows. It's also obvious that it's Abelian

$$\hat{C}_{nh} = \hat{C}_n^k \otimes \hat{\sigma}_h = \hat{\sigma}_h \otimes \hat{C}_n^k$$

A special case is the group C_{1h}, which is formed by two elements

$$C_{1h} := \{\hat{\mathbb{1}}, \hat{\sigma}_h\} = C_s$$

C_{nv} group

The C_{nv} group is similar to the C_{nh} group, but the n-th order symmetry axis lies on the symmetry plane. This generates $n - 1$ more planes of symmetry separated by an angle $\alpha = \pi/n$

D_n group

The D_n group is formed from a symmetry axis of order n, which is perpendicular to an axis of order 2. I.e. it's formed from an axis of order n and $n - 1$ axes of order 2 separated from an angle $\alpha = \pi/n$. A special case is given by the group $D_2 = V$, which is formed from an two axes of order 2 perpendicular to each other

D_{nh} group

The D_{nh} group is a complex formed from the product $D_n \otimes C_s$, i.e. for every order 2 axis there lies a plane perpendicular to such axis.

D_{nd} group

Starting from the D_{nh} group, one can define the group D_{nd} by taking the planes at a separation of $\alpha = \pi/2n$ radians.

T group

The T group is the group formed from the symmetries of the tetrahedron. This group can be formed from the group V (D_2), coupled with 4 oblique C_3 axes.

T_d group

The T_d group is formed from the T group by adding a plane of symmetry to the tetrahedron, it's formally given by the product $T_d = T \otimes C_s$

T_h group

The T_h group, instead is formed from the T group through addition of a center of symmetry, i.e. a point of inversion. This is described by the product $T_h = T \otimes C_i$

O group

The O group is the group given by the symmetries of the octahedron. Adding a center of symmetry, we obtain the group $O_h = O \otimes C_i$.

H Special Relativity

§ H.1 Principle of Relativity

The principle of relativity states a quite simple but deep affirmation: «All interaction propagate at a constant speed independent from the chosen frame of reference» . This speed is usually denoted as c and it's informally known as the speed of light, which has the following value (in SI units)

$$c = 2.998 \times 10^8 \text{ m/s} \tag{H.1}$$

In the part on classical mechanics we always intended between the lines that all interactions are instantaneous and therefore we'd have $c \to \infty$ formally. This can be interpreted as taking classical mechanics as an approximation of Einstein's relativity for which $v/c << 1$, which is the case for our really slow classical particles.

Note that this constant speed of propagation precludes that time isn't universal, and it is frame dependent. In order to understand this it's useful to get two coordinate frames K and $\tilde{K}$, where one is moving with respect to the other with a constant speed V.

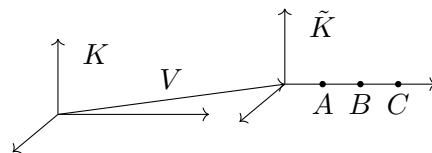Suppose now that a point A emits a signal towards two other points B and C

Figure H.1: The two frames K and $\tilde{K}$

In the frame $\tilde{K}$, where A is at rest, we see that the signal reaches both points at the same time, but the same CANNOT be true for the other system, since the relativity principle would be violated. Thinking in a different way, suppose that you're standing at the origin of the K system. If the velocity of the signal is constant in all reference frames we can for sure say that it's so where we're standing, therefore we end up seeing B moving towards the signal and C moving away from it, both with speed V. In this system we therefore must see a delay in when the two points receive such signal.

Although counterintuitive we're experimentally more than sure that this is actually a better approximation of nature than our beloved Newtonian mechanics.

§ H.2 Spacetime

Since time it's not anymore an universal thing and behaves itself as a coordinate, we can now think of our universe as a $4D$ manifold with time as a new coordinate. This is known as «Minkowsky Spacetime» or in short as «Spacetime» . This new definition follows:

Definition H.2.1 (Event). Given a spacetime with coordinates (ct, x, y, z) with c the speed of light, we define a point in spacetime as an «event» in such.
Since time only "flows" one way, we have that for every particle corresponds a wordline which connects all the events pertaining to such. Note that events are also known as «universe points»

Given the principle of relativity one might also ask rightfully how to formulate mathematically all of this, bringing out some invariants that might help with further derivations. Take again the previous system and call l the distance traveled by the signal after being emitted from A. Calling t_1 and t_2 the emission time and the arrival time respectively, we have that for obvious reasons

$$l = c(t_1 - t_2) \tag{H.2}$$

But, we can also write as follows

$$l = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \tag{H.3}$$

With (x_1, y_1, z_1) being the departure coordinates and (x_2, y_2, z_2) the arrival coordinates in K In $\tilde{K}$, analogously we have

$$\tilde{l} = c(\tilde{t}_2 - \tilde{t}_1)$$
$$\tilde{l} = \sqrt{(\tilde{x}_2 - \tilde{x}_1)^2 + (\tilde{y}_2 - \tilde{y}_1)^2 + (\tilde{z}_2 - \tilde{z}_1)^2} \tag{H.4}$$

Tying up both equations we end with the following result

$$\begin{cases} c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2 = 0 \\ c^2(\tilde{t}_2 - \tilde{t}_1)^2 - (\tilde{x}_2 - \tilde{x}_1)^2 + (\tilde{y}_2 - \tilde{y}_1)^2 + (\tilde{z}_2 - \tilde{z}_1)^2 = 0 \end{cases} \tag{H.5}$$

In "layman" words this basically means, that the following quantity

$$s_{12}^2 = c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2 \tag{H.6}$$

Called, «interval» , is a «relativistic invariant» , and therefore invariant with respect to changes of coordinate frames in the context of special relativity.
From (H.5) we have that if the two points are infinitesimally close to eachother we can define the infinitesimal interval as

$$ds^2 = c^2 dt - dx^2 - dy^2 - dz^2 \tag{H.7}$$

The invariance of such differential quantity is easy to show considering the previous case we stated where $\mathrm{d}s = \mathrm{d}\tilde{s} = 0$ we have, using basic intuition that

$$\mathrm{d}s^2 = a(V)\mathrm{d}\tilde{s}^2 \qquad \text{(H.8)}$$

Where $a(V)$ is a function of the relative velocity between the two considered frames. It cannot depend on direction due to the isotropy of space.

Consider now three inertial reference frames K, K_1, K_2, and let V_1, V_2 be the velocities of the frames K_1, K_2. We can therefore say, using (H.8) that

$$\begin{aligned} \mathrm{d}s^2 &= a(V_1)\mathrm{d}s_1^2 = a(V_2)\mathrm{d}s_2^2 \\ \mathrm{d}s_1^2 &= a(V_{12})\mathrm{d}s_2^2 \end{aligned} \qquad \text{(H.9)}$$

Where we defined the velocity between K_1, K_2 as V_{12}. Rewriting the equation we have

$$\mathrm{d}s^2 = a(V_1)a(V_{12})\mathrm{d}s_2^2 = a(V_2)\mathrm{d}s_2^2$$

Equating the coefficients of the differential $\mathrm{d}s_2$, we have

$$a(V_{12}) = \frac{a(V_2)}{a(V_1)} \qquad \text{(H.10)}$$

The previous equation then might be true if and only if $a(V_{12})$ depends only on the angle between the velocities V_1, V_2. This cannot be true due to the isotropy of spacetime, as we stated for the previous problem, and therefore $a(V)$ might only be a constant function. Taking $a(V_{12}) = 1$ for consistency between frames of reference, we have finally demonstrated that the differential spacetime interval is invariant

$$\mathrm{d}s = \mathrm{d}\tilde{s} \qquad \text{(H.11)}$$

This definition of $\mathrm{d}s$ gives rise to three kinds of intervals:

1. «Spacelike intervals» if $s_{12}^2 < 0$

2. «Timelike intervals» if $s_{12}^2 > 0$

3. «Light-like intervals» if $s_{12}^2 = 0$

These three distinctions let us answer two previously impossible questions: is it possible to find a reference frame where two events happen at the same time or at the same place in our three-dimensional perception?. The answer is surprisingly *yes*. It depends on the kind of the interval between the two points.

Let's work with the first assumption, taken two events in spacetime E_1, E_2, defined $t_{12} = t_2 - t_1$ and l_{12} as our usual 3D distance between the events, we have

$$s_{12}^2 = c^2 t_{12}^2 - l_{12}^2$$

Let's now search a system where $l'_{12} = 0$. In order to have this, using that $s_{12} = s'_{12}$ we have

$$s_{12} = c^2 t_{12} - l_{12}^2 = c^2 t'^2_{12} = s'^2_{12} > 0$$

I.e. the spacetime interval between the frame of reference at rest with respect to the two events and the new unknown frame of reference is timelike.

Analogously, if we wanted to find a new system where the two events happen at the same time, we might have set $t'_{12} = 0$, therefore getting

$$s_{12} = c^2 t_{12} - l_{12}^2 = l'^2_{12} = s'^2_{12} < 0 \tag{H.12}$$

§§ H.2.1 Spacetime Diagrams

The idea of spacetime and absoluteness of the velocity of interactions can be described well by a 2D spacetime diagram. Taken an origin for our system of coordinates (ct, x) we have that, considering v as the slope of a constant wordline, that $|v| < c$.
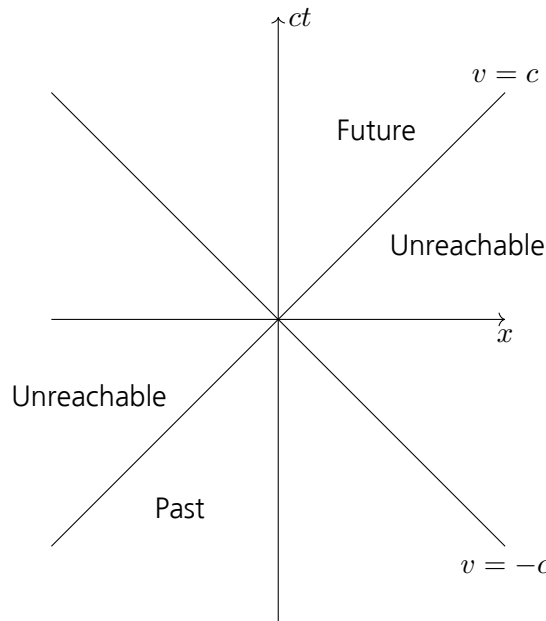

Figure H.2: Simple spacetime diagram. Note how all the events beyond the asymptote (or «horizon») $v = \pm c$ are inaccessible from 0

Thought in higher dimensions we have that all the past and future of an event are enclosed inside a cone bordered by our horizon $|v| = c$ which separates physical impossibilities from the actual physical past and future of what we're considering.

Note that if $v = \pm c$ we must have $x = \pm ct$, giving us a spacelike interval for our diagram. Considering instead past and future it's also easy to see that the past is always spacelike,

since $c^2t^2 - x^2 < 0$, and that the future is always timelike. Note also that past and future must be absolute

§ H.3 Proper Time

Since time is not a relativistic invariant, we need to search for a good substitute of it. Given a clock fixed at the origin of some inertial frame K'. After some time dt, the clock has moved (in our system) by the following quantity

$$\sqrt{dx^2 + dy^2 + dz^2}$$

By definition, in K' this clock is at rest, therefore we have

$$dx' = dy' = dz' = 0$$

Imposing the invariance of intervals we have that

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2 = c^2dt'^2 \tag{H.13}$$

Therefore, it must be true that

$$dt' = dt\sqrt{1 - \frac{dx^2 + dy^2 + dz^2}{c^2dt^2}} \tag{H.14}$$

This is the expression for the passing of time in the system where the clock is at rest, and it's called the «proper time» of the clock, usually indicated with τ. Writing the sum of differentials as dr^2 and using the definition of v^2, we have that

$$d\tau = dt\sqrt{1 - \frac{v^2}{c^2}} = \frac{ds}{c} \tag{H.15}$$

Integrating and using the fundamental theorem of calculus, we have that a given time interval will be "felt" differently by the clock, where

$$\Delta\tau = \int_{\tau_1}^{\tau_2} \sqrt{1 - \frac{v^2}{c^2}}dt < \Delta t \tag{H.16}$$

This tells us that a moving clock will tick slower than a clock at rest (note also on how this definition depends directly on the chosen frame).
This difference of measured time is known as «time dilation» .

§ H.4 Formalization of the Principle of Relativity

All of what we found before can be crammed into the most fundamental element of relativity: coordinate transformations.
Consider two reference frames K, (ct, x, y, z) and K', (ct, x, y, z). Mathematically, what

we call interval is the usual 4D distance in a seminegative definite metric, and due to its invariance we must have that all coordinate transformations between these two systems must be rototraslations (isometries). Translations can be immediately ignored since they only move the origin of the system, and therefore we choose our faithful rotations in order to find these coordinate transformation laws.

All the possible rotations are between the planes xy, xz, yz and tx, ty, tz. All rotations xy, xz, yz are our usual 3D rotations and are of no use, therefore we choose the rotations tx, ty, tz. Taking tx as the chosen one we have that the spacetime interval is

$$s^2 = c^2 t^2 - x^2$$

Therefore, all searched rotations *must preserve* this relationship. The first idea one might have is to look at the symmetry of the system and deduce immediately that such rotation must be hyperbolic in nature. We therefore define the following

$$\begin{pmatrix} x \\ ct \end{pmatrix} = \begin{pmatrix} \cosh\psi & \sinh\psi \\ \sinh\psi & \cosh\psi \end{pmatrix} \begin{pmatrix} x' \\ ct' \end{pmatrix} \tag{H.17}$$

Taking $x' = 0$ it all reduces to this single equation

$$\frac{x}{ct} = \frac{V}{c} = \tanh\psi \tag{H.18}$$

It's common to indicate such value with the pure number β, called the «Lorentz Boost» , where

$$\beta = \frac{V}{c}$$

Solving (H.18) we have that

$$\beta = \frac{\sinh\psi}{\sqrt{1 + \sinh^2\psi}} = 0 \Longrightarrow \sinh^2\psi = \frac{\beta}{\sqrt{1 - \beta^2}} \tag{H.19}$$

And

$$\cosh^2\psi = 1 + \sinh^2\psi \Longrightarrow \cosh\psi = \frac{1}{\sqrt{1 - \beta^2}} = \gamma \tag{H.20}$$

Where γ is known as the «Lorentz/Gamma Factor» .

Substituting back into (H.18) we have back our searched transformations

$$\begin{pmatrix} x \\ ct \end{pmatrix} = \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} x' \\ ct' \end{pmatrix} \tag{H.21}$$

Note that the inverse transformation is simply given imposing $\beta \to -\beta$.

The complete transformation between the two reference frames will finally be a 4D linear system as follows

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} \tag{H.22}$$

These transformations are known as «Lorentz Transformations» and are the fundamental transformations between frames of reference in special relativity. These transformations formalize the principle of relativity. For $v << c$ these transformations bring back the usual Galilean transformations corrected by a first order factor in c, as we expected

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & \beta & 0 & 0 \\ \beta & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} \tag{H.23}$$

§§ H.4.1 Length Contraction and Time Dilation

Using Lorentz transformations it's possible to mathematically formalize all relativistic effects. One of such is known as «length contraction» , where the measured length of an object depends on the chosen reference frame.
As a matter of example take a "rigid" rod in a system K, long Δx, and consider the system K' where the rod is at rest. In this system we have

$$\Delta x' = x'_1 - x'_2 = \gamma(x_2 - x_1) - \gamma\beta c(t_2 - t_1) = \gamma\Delta x - \gamma\beta c\Delta t \tag{H.24}$$

Since we're measuring the length directly, we can say without problems that $\Delta t = 0$, and we get

$$\Delta x' = \gamma\Delta x = \frac{\Delta x}{\sqrt{1 - \beta^2}} = \frac{\Delta x}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{H.25}$$

Therefore, for $\beta \neq 0$ we have $\Delta x' < \Delta x$. We call $\Delta x = l_0$ as the proper lenght of this rod. Note that a major consequence of this is that a rigid body in the classical sense of the term cannot be conceived in Special Relativity.
A second effect that we stated before and didn't formalize properly is that of time dilation. Taken a clock at rest in a system K' and two events happening at some coordinate (x', y', z') of K'. We have that the time elapsed between the two events will be $\Delta t' = t'_2 - t'_1$, and therefore, using Lorentz transformations we get, in K

$$\Delta t = \gamma\left(t'_1 + \frac{\beta}{c}\Delta x'\right) \tag{H.26}$$

Imposing that the events happen at the same place (x', y', z') we have $\Delta x' = 0$ and therefore

$$\Delta t = \gamma\Delta t' \tag{H.27}$$

Therefore, the clock in the still frame is measuring smaller time intervals, and the time measured is dilated.

§§ H.4.2 Velocity Transformations

As we have seen velocities have an upper bound which is the speed of light. It's possible to find the transformations of velocities from the transformations (H.21) and applying them to differentials.
We have

$$
\begin{pmatrix} \mathrm{d}t \\ \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}z \end{pmatrix} = \begin{pmatrix} \gamma & \frac{\beta\gamma}{c} & 0 & 0 \\ \frac{\beta\gamma}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{d}t' \\ \mathrm{d}x' \\ \mathrm{d}y' \\ \mathrm{d}z' \end{pmatrix}
\tag{H.28}
$$

Rearranging the terms we have finally

$$
\begin{cases}
v_x = \dfrac{v_x' + \beta c}{1 + \frac{\beta}{c} v_x'} \\[3mm]
v_y = \dfrac{v_y'}{\gamma \left(1 + \frac{\beta}{c} v_x'\right)} \\[3mm]
v_z = \dfrac{v_z'}{\gamma \left(1 + \frac{\beta}{c} v_x'\right)}
\end{cases}
\tag{H.29}
$$

Approximating for $v \ll c$ we get the usual velocity composition formula with an added relativistic correction

$$
\begin{cases}
v_x \approx v_x' + V \left(1 - \dfrac{v_x^{2'}}{c^2}\right) \\[3mm]
v_y \approx v_y' - v_x' v_y' \dfrac{\beta}{c} \\[3mm]
v_z \approx v_z' - v_x' v_z' \dfrac{\beta}{c}
\end{cases}
\tag{H.30}
$$

Or, in vector form

$$
v^i = v^{i'} + V^i - \frac{v^{i'}}{c^2}(V^i v_i')
\tag{H.31}
$$

Note how v and v' are tied asymetrically in the transformation. Consider now a simple planar motion in the xy plane, where $v^i = (v_x, v_y, 0)$, we can find the law of transformation of angles considering that v^i can be rewritten in polar coordinates, as follows

$$
\begin{cases}
v_x = v \cos\theta \\
v_y = v \sin\theta \\
v_z = 0
\end{cases}
$$

Applying the transformations, we have

$$
\begin{cases}
v\cos\theta = \dfrac{v'\cos\theta' + \beta c}{1 + \frac{\beta}{c}v'\cos\theta'} \\[2ex]
v\sin\theta = \dfrac{v'\sin\theta'}{\gamma\left(1 + \frac{\beta}{c}v'\cos\theta'\right)}
\end{cases}
\tag{H.32}
$$

Where we used that the motion in the new system will be still planar.
Rewritten in other terms, we have

$$
\tan\theta = \frac{\frac{v'\sin\theta'}{\gamma\left(1 + \frac{\beta}{c}v'\cos\theta'\right)}}{\frac{v'\cos\theta' + \beta c}{1 + \frac{\beta}{c}v'\cos\theta'}} = \frac{v'\sin\theta'}{\gamma\left(v'\cos\theta' + \beta c\right)}
\tag{H.33}
$$

Which explicitates the change of direction of velocity between different coordinate systems.

§ H.5 4-Vectors

[1]

As we have already suggested before, the 4-tuple $x^\mu = (ct, x, y, z)$ can be seen as a set of coordinates in spacetime, or as a radius vector. The square of vectors in spacetime can be seen as a non-euclidean scalar product as follows

$$
x^\mu x_\mu = g_{\mu\nu}x^\mu x^\nu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2
\tag{H.34}
$$

Where $g_{\mu\nu}$ is the metric tensor of spacetime

$$
g_{\mu\nu} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{pmatrix}
\tag{H.35}
$$

From what we wrote for special relativity itself, we have a new definition

Definition H.5.1 (4-Vector). A «4-vector» is a 4-tuple that transforms between coordinate frames using Lorentz transformations, as

$$
a^\mu = \Lambda^\mu_\nu a^\nu
\tag{H.36}
$$

Where λ^μ_ν is the already defined transformation matrix of the Lorentz transformations.

---

[1]From here on, all greek indexes $(\mu, \nu, \sigma, \cdots)$ are to be intended as spacetime indexes, and latin indexes $(i, j, k, \cdots)$ as usual 3D indexes if not otherwise stated

Using the metric tensor one can transforms between covariant vectors and contravariant vectors using $a_\mu = g_{\mu\nu}a^\nu$, and due to the semidefinite signature of the metric one has that $a^i = -a_i$, where a^i is the spatial part of the vector. Note also that inserting it into the formula for a scalar product ($a^\mu b_\mu$) one gets back what we had defined before.

It's also possiible to define 4-scalars, which are relativistic invariants. One of such 4-scalars is the square of a 4-vector or the scalar product between 2 4-vectors.

Another way of writing 4-vectors is with a tuple composed as follows

$$a^\mu = (a^0, a^i) \tag{H.37}$$

Where the first component is known as the «polar» component of the 4-vector, and the second is known as the «axial» component of the 4-vector. Therefore we can write

$$\begin{aligned} x^\mu &= (ct, x^i) \\ x_\mu &= (ct, -x_i) \end{aligned} \tag{H.38}$$

§§ H.5.1 4-Velocity and 4-Acceleration

It's possible to define a 4-vector analogue to the velocity of a particle. Indicating with τ the proper time we define the 4-velocity u^μ as

$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \tag{H.39}$$

Since $\mathrm{d}\tau = \frac{c}{\gamma}\mathrm{d}t$ we have

$$u^\mu = \frac{\gamma}{c}\frac{\mathrm{d}x^\mu}{\mathrm{d}t}$$

In other words

$$u^\mu = \left(\gamma, \frac{\gamma}{c}v^i\right)$$

Note that the square of u^μ is a relativistic invariant and special in nature due to its unitary value, in fact

$$u^\mu u_\mu = \gamma^2 - \gamma^2\frac{v^2}{c^2} = 1$$

The 4-acceleration w^μ is defined analogously derivating again with respect to the proper time, hence

$$w^\mu = \frac{\gamma}{c}\,\mathrm{d}u^\mu\,t = \left(\frac{\gamma}{c}\frac{\mathrm{d}\gamma}{\mathrm{d}t}, \frac{\gamma}{c^2}\frac{\mathrm{d}\gamma v^i}{\mathrm{d}t}\right) \tag{H.40}$$

Deriving with respect to time we have firstly that

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t} = \frac{v^i a_i}{\left(1 - \frac{v^2}{c^2}\right)^{\frac{3}{2}}} = \frac{\gamma^3}{c^2}v^i a_i$$

And therefore

$$w^\mu = \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = \frac{\gamma}{c}\left(\frac{\gamma^3}{c^2}v^i a_i, \frac{\gamma}{c^3}v^j a_j v^i + \frac{\gamma}{c}a^i\right) \tag{H.41}$$

It's possible to demonstrate that $w^\mu u_\mu = 0$, i.e. that 4-velocity and 4-acceleration are always mutually orthogonal. In fact

$$\frac{\mathrm{d}}{\mathrm{d}\tau}u^\mu w_\mu = \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau}u_\mu + \frac{\mathrm{d}u_\mu}{\mathrm{d}\tau}u^\mu = 2u^\mu w_\mu = 0$$

§ H.6 Exercises

Exercise H.6.1 (Uniformly Accelerated Motion). Solve the motion of an uniformly accelerated particle in the context of Special Relativity.
Consider that the 4-acceleration is constant only in the frame comoving with the particle.
S O L U T I O N
We have that in the comoving frame $\gamma = 1$ and $v = 0$, therefore

$$w^\mu = \left(0, \frac{\dot{v}^i}{c^2}\right)$$

Since a is constant we rotate the 3D system in order to get $a \parallel x$, therefore getting

$$w^\mu = \left(0, \frac{a}{c^2}, 0, 0\right)$$

Note that we can also define a 4-scalar

$$w^\mu w_\mu = -\frac{a^2}{c^2}$$

Changing to the fixed frame of reference, we have

$$w^{\mu\prime} = \frac{\gamma}{c}\left(\frac{\gamma^3}{c^2}v^i \dot{v}_i, \frac{\gamma^3}{c^2}v^j \dot{v}_j v^i + \frac{\gamma}{c}\dot{v}^i\right) = \frac{\gamma^4}{c^2}\left(\frac{v^i \dot{v}_i}{c}, \frac{v^2}{c^2}\dot{v}^i + \frac{\dot{v}^i}{\gamma^2}\right)$$

Using that

$$\left(\frac{v^2}{c^2} + \frac{1}{\gamma^2}\right)\dot{v}^i = \dot{v}^i$$

We end up with the following simplified result

$$w^{\mu\prime} = \frac{\gamma^4}{c^2}\left(\frac{1}{c}\dot{v}^i v_i\right)$$

Which gives us the following differential equation

$$w^\mu w_\mu = \frac{\gamma^8}{c^4}\left(\frac{1}{c^2}(v^i \dot{v}_i)^2\right) - \frac{\gamma^8}{c^4}\dot{v}^2 = -\frac{a^2}{c^4}$$

Simplifying the LHS we get

$$\frac{\gamma^8}{c^4}\left(\frac{v^2}{c^2}\dot{v}^2 - \dot{v}^2\right) = \frac{\gamma^8}{c^4}\left(\frac{v^2}{c^2} - 1\right) = -\frac{\gamma^6}{c^4}\dot{v}^2$$

Therefore, putting it back into the first equation, we get

$$-\gamma^6\dot{v}^2 = -a^2 \Longrightarrow \gamma^3\frac{dv}{dt} = a$$

Note that using the derivative of γ with respect to time we can rewrite the LHS as the derivative of a product, in fact

$$\frac{d(\gamma v)}{dt} = \frac{\gamma^3}{c^2}v^2\dot{v} + \gamma\dot{v} = \dot{v}\left(\frac{\gamma^3}{c^2}v^2 + \gamma\right) = \gamma^3\dot{v}\left(\frac{v^2}{c^2} + \frac{1}{\gamma^2}\right) = \gamma^3\dot{v}$$

Therefore, finally

$$\frac{d(\gamma t)}{dt} = a \Longrightarrow \gamma v(t) = at + c$$

Imposing that $v(0) = 0$ we get $c = 0$ and therefore, solving for $v(t)$, we have

$$\frac{v(t)}{\sqrt{1 - \frac{v^2}{c^2}}} = at \Longrightarrow v^2 = a^2t^2 - \frac{a^2t^2}{c^2}v^2 \Longrightarrow v^2 = a^2t^2\left(1 + \frac{a^2t^2}{c^2}\right)^{-1}$$

Therefore

$$v(t) = \frac{at}{\sqrt{1 + \frac{a^2t^2}{c^2}}}$$

Then, by direct integration we can find $x(t)$

$$x(t) = \int \frac{at}{\sqrt{1 + \frac{a^2t^2}{c^2}}}dt = \frac{c^2}{2a}\int \frac{1}{\sqrt{1 + w^2}}dw = \frac{c^2}{2a}\left(2\sqrt{1 + w} + k\right)$$

Where we used the substitution $w = \frac{a^2t^2}{c^2}$. Imposing the initial condition that $x(0) = 0$ we get $k = -1$, and therefore

$$x(t) = \frac{c^2}{a}\left(2\sqrt{1 + \frac{a^2t^2}{c^2}} - 1\right)$$

The proper time of the particle is

$$\tau = \frac{1}{c}\int_{s_0}^{s}ds = \int_{t_0}^{t}\frac{1}{\gamma}dt = \int_{0}^{t}\sqrt{1 - \frac{v^2}{c^2}}dt$$

From the definition of $v(t)$ we have that

$$\gamma = \frac{1}{1 - \frac{a^2t^2}{c^2\left(1 + \frac{a^2t^2}{c^2}\right)}}$$

Therefore our integral becomes

$$\tau = \int_0^t \sqrt{1 - \frac{a^2 t^2}{c^2 \left(1 + \frac{a^2 t^2}{c^2}\right)}} \, dt = \frac{a}{c} \int_0^{\frac{a}{c}t} \sqrt{1 - \frac{z^2}{1 + z^2}} \, dz = \frac{a}{c} \int_0^{\frac{a}{c}t} \frac{1}{\sqrt{1 + z^2}} \, dz = \frac{a}{c} \arcsin\left(\frac{at}{c}\right)$$

Where we used the substitution $\frac{at}{c} = z$

§ H.7 Relativistic Least Action Principle

In order to use a relativistic version of the least action principle, we need to impose the principle of relativity into its formulation. This is done via imposing that $\mathcal{S}$ must be a relativistic invariant.

It's already obvious that the simplest invariant differential in special relativity is ds therefore we must have that $d\mathcal{S} \propto ds$, and we can rewrite our action as follows

$$\mathcal{S} = -\alpha \int_a^b ds \tag{H.42}$$

Where $\alpha \in \mathbb{R}$ is a parameter that depends directly on the properties of the particle. The minus sign added there is needed in order to make sure that $\mathcal{S}$ as an extremal in $[s(a), s(b)]$. We now proceed to find a Lagrangian as per usual, therefore we need to transform the wordline integral into a time integral. Since time is NOT a relativistic invariant we can only choose one "time", which is the proper time of the reference frame, which is also proportional to the interval differential.

Using $ds = \frac{c}{\gamma} dt$ we have

$$\mathcal{S} = -\alpha c \int_{\tau_1}^{\tau_0} \frac{1}{\gamma} dt = \int_{\tau_0}^{\tau_1} -\alpha c \sqrt{1 - \frac{v^2}{c^2}} \, dt \tag{H.43}$$

What we have inside the integral symbol on the RHS is the Lagrangian of the system, which depends on the parameter α.

In order to determine this parameter we need to have that for $v/c << 1$ our relativistic Lagrangian must become our known classical Lagrangian. We have for $\beta \to 0$

$$\mathcal{L} = -\frac{\alpha c}{\gamma} = -\alpha c \sqrt{1 - \beta^2} \approx -\alpha c + \frac{\alpha v^2}{2c} = \frac{m v^2}{2} \tag{H.44}$$

Comparing the terms with v^2 we have that $\alpha = mc$, and therefore we have

$$\mathcal{S} = -mc \int_{\tau_0}^{\tau_1} \frac{1}{\gamma} dt \Longrightarrow \mathcal{L} = -\frac{mc}{\gamma} \tag{H.45}$$

Where γ is the already known Lorentz factor.

§§ H.7.1 Relativistic Energy and Momentum

The easiest way possible to define relativistic energy and relativistic momentum is by using our well known identities in Lagrangian mechanics. We must have therefore, for the 3-momentum

$$p_i = \frac{\partial \mathcal{L}}{\partial v^i} \tag{H.46}$$

Since we know that $\mathcal{L} = -mc/\gamma$ the calculation is straightforward. We have

$$\frac{\partial \mathcal{L}}{\partial v^i} = -mc\frac{\partial}{\partial v^i}\sqrt{1 - \frac{v^2}{c^2}} = \frac{mv^i}{\sqrt{1 - \frac{v^2}{c^2}}} = \gamma m v^i$$

As usual, using the Lagrangian for defining our energy we have that $p^i v_i - \mathcal{L} = E$, i.e.

$$E = \gamma m v^2 + \frac{mc^2}{\gamma} = \gamma\left(mv^2 + \frac{mc^2}{\gamma^2}\right) =$$
$$= \gamma\left(mv^2 + mc^2 - mv^2\right) = \gamma mc^2 = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{H.47}$$

Note that due for this definition, if $v = 0$, $E \neq 0$. We have $E_0 = mc^2$, and this is known as the «rest energy» of the given particle.

This shape for our energy also demonstrates that mass is not conserved anymore in relativity. In fact, given a body comprised of multiple particles we have that E_0^b will be, if the mass of the body is M

$$E_0^b = Mc^2 \tag{H.48}$$

But, since in the rest energy of the single particles composing the body we also have to add all the interaction energies between the body, we will therefore have

$$E_0^b = Mc^2 \neq \sum_i m_i c^2 \tag{H.49}$$

Therefore, finally $M \neq \sum_i m_i$, which makes our previous point.

After defining Lagrangian and energy, the next step we can make is to find the Hamiltonian of a relativistic particle. From the definition of energy and momentum we can write the following system

$$\begin{cases} E = \gamma mc^2 \\ p^i = \gamma m v^i \end{cases} \tag{H.50}$$

Manipulating γ and the whole system, we have

$$\gamma(p) = \sqrt{1 + \frac{p^2}{m^2 c^2}} \tag{H.51}$$

Plugging it into the definition of E we have

$$E = \gamma(p)mc^2 = mc^2\sqrt{1 + \frac{p^2}{m^2 c^2}} \tag{H.52}$$

By definition of Hamiltonian function this is our $\mathcal{H}$. Rewriting in a different way we have

$$\mathcal{H} = \gamma(p)mc^2 = mc^2\sqrt{1 + \frac{p^2}{m^2c^2}} = \sqrt{m^2c^4 + p^2c^2} = c\sqrt{m^2c^2 + p^2} \tag{H.53}$$

Again, for $\frac{o^2}{m^2c^2} = \beta(p) << 1$ we get the classical counterpart, plus the relativistic rest energy.

Note how, using the previous equations we have that for $v \to c$ $E \to \infty$ if $m \neq 0$, and defining the existence of massless particles is not obvious. Using the Hamiltonian formulation we immediately see that if $m = 0$ we have

$$\mathcal{H} = E = pc \tag{H.54}$$

It's also obvious from this that the only velocity such particle can have is $v = c$. These kinds of particles are known as ultrarelativistic particles, and photons are one example of such. Note that it's possible to approximate the energy of massive particles with their ultrarelativistic counterparts in case where the rate between the rest energy and the total energy of the particle is small enough for the needs. I.e.

$$mc^2 << E \Longrightarrow E \approx pc \tag{H.55}$$

§ H.8 Relativistic Hamilton Jacobi Equation and 4-Vector Formulation

It's possible to rewrite the relativistic least action principle using 4-vector notation. For what we wrote in the previous section in the part on 4-vectors, we can imagine to define the infinitesimal interval as a 4-scalar via the definition of the infinitesimal 4-radius vector $\mathrm{d}x^\mu = (c\mathrm{d}t, \mathrm{d}r^i)$. We have

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu \Longrightarrow \mathrm{d}s = \sqrt{g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu} \tag{H.56}$$

We add to this the boundary conditions for the action in spacetime, as $\delta x^\mu(a) = \delta x^\mu(b) = 0$, and we get, for our least action principle

$$S = -mc\int_a^b \mathrm{d}s = -mc\int_a^b \sqrt{g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu} \tag{H.57}$$

Variating the action we have, firstly

$$\delta\sqrt{g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu} = \frac{g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}\delta x^\nu}{\sqrt{g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu}} = \frac{\mathrm{d}x^\mu}{\mathrm{d}s}\delta\mathrm{d}x^\nu = u_\mu\delta\mathrm{d}x^\mu$$

Then

$$\delta S = -mc\int_a^b u_\mu\delta\mathrm{d}x^\mu$$

Using as usual integration by parts in order to move the differentials we have, implicitly using the boundary conditions

$$\delta\mathcal{S} = mc\int_a^b \frac{\mathrm{d}u_\mu}{\mathrm{d}s}\delta x^\mu \mathrm{d}s \tag{H.58}$$

Imposing lastly the least action principle we get the relativistic equation of motion for a free particle

$$\frac{\mathrm{d}u^\mu}{\mathrm{d}s} = 0 \tag{H.59}$$

Considering instead the second condition $\delta x^\mu(b) = \delta x^\mu$ nonzero we get the usual definition of action as a function of (spacetime) coordinates

$$\delta\mathcal{S} = -mcu_\mu\delta x^\mu \Longrightarrow \frac{\partial\mathcal{S}}{\partial x^\mu} = -mcu_\mu$$

By comparation with the classical definition, the derivative of the action with respect to the coordinates is defined as the generalized momentum of the system. Since in this case we're using 4-vectors and the derivative of the action with respect to the 4-position is a 4-vector itself (not really in general relativity, but in SR it's true) we have a new definition for the momentum, the «4-momentum» of a relativistic system

$$p_\mu = \partial_\mu\mathcal{S} = mcu_\mu \tag{H.60}$$

Using the definition of 4-velocity as $u_\mu = \gamma\left(1, -\frac{1}{c}v^i\right)$ we have, remembering that $E = \gamma mc^2$

$$p_\mu = \left(\frac{E}{c}, -\gamma mv^i\right) = \left(\gamma mc, -p^i\right) \tag{H.61}$$

Using instead that $\partial_0 = c^{-1}\partial_t$ we have

$$p_0 = \partial_0\mathcal{S} = \frac{1}{c}\frac{\partial\mathcal{S}}{\partial t} = \frac{E}{c} \tag{H.62}$$

Where we used the classical conclusion that $\partial_t\mathcal{S} = E$. Using the already known Lorentz transformations we have that energy and momentum, since they're tied by a 4-vector, they aren't invariants and transform as follows

$$\begin{cases} E = \gamma\left(E' + \beta cp'_x\right) \\ p_x = \gamma\left(p'_x + \frac{\beta}{c}E'\right) \\ p_y = p'_y \\ p_z = p'_z \end{cases} \tag{H.63}$$

Using that $u^\mu u_\mu = 1$ we can define a relativistic invariant for 4-momentum, as follows

$$p^\mu p_\mu = m^2c^2 u^\mu u_\mu = m^2c^2 \tag{H.64}$$

We can also find a 4-force definition by deriving with respect to proper time.

$$f^\mu = \frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} = \frac{\gamma}{c}\frac{\mathrm{d}p^\mu}{\mathrm{d}t} = \frac{\gamma}{c}\left(\dot{\gamma}mc, \frac{\mathrm{d}p^i}{\mathrm{d}t}\right) \tag{H.65}$$

Using the already known derivative of the Lorentz factor we have, finally

$$f^\mu = \frac{\gamma}{c}\left(\frac{\gamma^3}{c}mv^i a_i, f^i\right) = \left(\frac{\gamma}{c^2}f^i v_i, \frac{\gamma}{c}f^i\right) = \left(\frac{\gamma}{c^2}W, \frac{\gamma}{c}f^i\right) \tag{H.66}$$

Where W is the already well known work of the force.
The Hamilton-Jacobi equation can be defined from (H.64), and it simply becomes

$$\partial_\mu \mathcal{S} \partial^\mu \mathcal{S} = m^2 c^2 \tag{H.67}$$

Or, in explicit form

$$\frac{1}{c^2}\left(\frac{\partial \mathcal{S}}{\partial t}\right)^2 - (\nabla \mathcal{S})^2 = m^2 c^2 \tag{H.68}$$

§ H.9 Laboratory and Center of Mass Reference Frames

We have already proven that in special relativity there are no preferred reference frames, therefore we might already choose the ones that ease our calculations.
In the field of relativistic particle collisions (like particle physics) we have two main choices of reference frames

1. The laboratory reference frame

2. The center of mass reference frame

The first one is defined as the reference frame of the resting observer of the event, while the second is the reference frame of the center of mass of the system of particles interacting.[2]
For understanding properly consider the collision of two particles m_1, m_2 where the second is a target particle at rest in the lab frame. We have, before the collision

$$p_1^\mu = \left(\frac{E_1}{c}, p_1^i\right)$$
$$p_2^\mu = (m_2 c, 0) \tag{H.69}$$

And after the collision

$$P^\mu = p_1^\mu + p_2^\mu = \left(\frac{E_1}{c} + m_2 c, p_1^i\right) \tag{H.70}$$

---

[2]I will use a star indicating the center of mass r.f. the rest will be intended as being in the laboratory reference frame

By definition of center of mass, we have that this reference frame will be the one for which $P^i = 0$, i.e., going back to our two particles pre-collision

$$p_1^{\mu\star} = \left(\frac{E_1^\star}{c}, p^{i\star}\right), \qquad p_2^{\mu\star} = \left(\frac{E_2^\star}{c}, -p^{i\star}\right) \tag{H.71}$$

I.e.

$$P^{\mu\star} = \left(\frac{1}{c}\left(E_1^\star + E_2^\star\right), 0\right) \tag{H.72}$$

§ H.10 Invariant Mass

Consider a system of n particles with momentum $p_{(k)}^\mu = (E_k, p_{(k)}^i)$ and let the sum of all 4-momentums be P^{μ}[3].

Definition H.10.1 (Invariant Mass). Given the previous system of particles, the relativistic invariant of the total 4-momentum is defined as «invariant mass» $\sqrt{s}$ of the system, so

$$\sqrt{s} = \sqrt{P^\mu P_\mu} = \sqrt{\left(\sum_k E_k\right)^2 - \left(\sum_k p_k\right)^2} = \sum_k E_k^\star = E^\star = M^\star \tag{H.73}$$

Where we used that $\sum_k p_{(k)}^i = 0$ in the $\star-$system, also known before as the center of mass system

Consider now the case of a particle colliding into a target particle. We have

$$p_1^\mu = \left(E_1, p_1^i\right), \quad p_2^\mu = (m_2, 0), \quad P^\mu = \left(E_1 + m_2, p_1^i\right) \tag{H.74}$$

And

$$P^\mu P_\mu = (E_1 + m_2)^2 - p_1^2 = E_1^2 + m_2^2 + 2E_1 m_2 - p_1^2 = P^{\mu\star} P_{\mu\star} \tag{H.75}$$

Note that $E_1^2 = m_1^2$. Therefore, putting it all together, for a particle colliding into a target, the invariant mass will be

$$\sqrt{s} = \sqrt{P^\mu P_\mu} = \sqrt{m_1^2 + m_2^2 + 2E_1 m_2} = E^\star = M^\star \tag{H.76}$$

Note that if we have $m_1, m_2 << E_1$, ie $\beta \approx 1$ and the particles are ultrarelativistic, then the invariant mass formula can be approximated as follows

$$\sqrt{s} \approx \sqrt{2E_1 m_2} \tag{H.77}$$

In case that both particle have $p_{(k)}^i \neq 0$ in the lab system, we have

$$p_1^\mu = \left(E_1, p_1^i\right), \quad p_2^\mu = \left(E_2, p_2^i\right), \quad P^{\mu\star} = (E_1^\star + E_2^\star, 0) \tag{H.78}$$

---

[3]From now on we will work in God-given units, where $c = 1$

Therefore

$$P^\mu P_\mu = E_1^2 + E_2^2 + 2E_1 E_2 - \left(p_1^2 + p_2^2 + 2p_1 p_2 \cos\theta\right)$$
$$P^{\mu^\star} P_{\mu^\star} = (E_1^\star + E_2^\star)^2 = P^\mu P_\mu$$

Or in simpler terms

$$\sqrt{s} = \sqrt{m_1^2 + m_2^2 + 2\left(E_1 E_2 - p_1 p_2 \cos\theta\right)} = E_1^\star + E_2^\star \tag{H.79}$$

If $m_1, m_2 << E_1$, so in the ultrarelativistic case, we have

$$\sqrt{s} \approx \sqrt{2E_1 E_2 \left(1 - \cos\theta\right)} \tag{H.80}$$

Note that in cases like particle colliders, like the LHC, SuperKamiokande or Fermilab, we have that the collision is frontal, i.e. $\theta = \pi$ and therefore the invariant mass formula becomes extremely easier to remember, especially if the particles are of the same kind ($E_1 = E_2$)

$$\sqrt{s} \approx 2\sqrt{E_1 E_2} = 2E \tag{H.81}$$

§§ H.10.1 Transformations of the Invariant Mass

Let's go back to what we had defined before for the invariant mass. We have that $\sqrt{s} = E^\star$, therefore, considering the total 4-momentum in the $\star-$system, we can write without problems

$$P^{\mu^\star} = \left(\sqrt{s}, 0\right) \tag{H.82}$$

This, as every 4-momentum, transforms with Lorentz transformations. Consider the boost along the x-axis without loss of generality, and transform towards the lab system.

$$\begin{pmatrix} \sqrt{s} \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sum_k E_k \\ \sum_k p_k \\ 0 \\ 0 \end{pmatrix} \tag{H.83}$$

Expanding the system and keeping only the two nonzero lines we have

$$\begin{cases} \gamma \sum_k E_k - \beta\gamma \sum_k p_k = \sqrt{s} \\ \gamma \sum_k p_k - \beta\gamma \sum_k E_k = 0 \end{cases} \tag{H.84}$$

From the second row we have

$$\sum_k p_k = \beta \sum_k E_k \Longrightarrow \beta = \frac{\sum_k p_k}{\sum_k E_k} = \frac{P}{E} \tag{H.85}$$

And, therefore

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\left(\frac{P}{E}\right)^2}} = \sqrt{\frac{E^2}{E^2-P^2}} = \frac{E}{\sqrt{s}} \tag{H.86}$$

Which, wrapped up, gives a different way to interpret the Lorentz boost and Lorentz factor

$$\begin{cases} \gamma = \dfrac{E}{\sqrt{s}} \\[2mm] \beta = \dfrac{P}{E} \end{cases} \tag{H.87}$$

§ H.11 Transverse Momentum and Transformation of Angles

Consider now a particle moving along the z axis. Transforming from the lab system to the $\star-$system we have, considering spherical polar coordinates

$$\begin{pmatrix} E \\ p\sin\theta\cos\varphi \\ p\sin\theta\sin\varphi \\ p\cos\theta \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & \beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} E^\star \\ p^\star\cos\theta^\star\cos\varphi^\star \\ p^\star\sin\theta^\star\sin\varphi^\star \\ p\cos\theta^\star \end{pmatrix} \tag{H.88}$$

Where γ, β are the Lorentz factors of the lab system..

Definition H.11.1 (Transverse Momentum). We define the «transverse momentum» $p_\perp$ as the 3-momentum orthogonal to the z axis.
In general it's defined as the 2-vector $p_\perp = (p_x, p_y)$, i.e.

$$p_\perp = \begin{pmatrix} p_x \\ p_y \end{pmatrix} = \begin{pmatrix} p\sin\theta\cos\varphi \\ p\sin\theta\sin\varphi \end{pmatrix} \tag{H.89}$$

Applying the transformation it's obvious that $p_\perp = p_\perp^\star$, therefore $p_\perp$ is a relativistic invariant.

From this last relativistic invariant, taking the square we therefore must have

$$p_\perp^2 = (p_\perp^\star)^2 \Longrightarrow p^2\sin\theta = (p^\star)^2\sin^2\theta^\star \tag{H.90}$$

Note how φ disappears from the calculations, giving $\varphi = \varphi^\star$, this means that the azimuthal angle is another relativistic invariant of motion.
Applying now the transformation on p_z we have

$$p_z = p\cos\theta = \gamma\left(\beta E^\star + p^\star\cos\theta^\star\right) \tag{H.91}$$

Using that p_y, φ are relativistic invariants we can write the following system

$$p_y = p \sin\theta \sin\varphi = p^\star \sin\theta^\star \cos\varphi^\star$$
$$p_z = p \cos\theta = \gamma \left(\beta E^\star + p^\star \cos\theta^\star\right) \tag{H.92}$$

Which, solving for θ gives

$$\frac{p_y}{p_z} = \tan\theta = \frac{\sin\theta^\star}{\gamma \left(\beta E^\star + p^\star \cos\theta^\star\right)} \tag{H.93}$$

Rewriting the denominator we get

$$\tan\theta = \frac{\sin\theta^\star}{\gamma_0 \left(\beta_0 \frac{E^\star}{p^\star} - \cos\theta^\star\right)} = \frac{\sin\theta^\star}{\gamma_0 \left(\frac{\beta_0}{\beta^\star} - \cos\theta^\star\right)} = \tag{H.94}$$

Where we defined the boost of the center of mass with respect to the $\star$ energy-momentum as follows

$$\beta^\star = \frac{p^\star}{E^\star} \tag{H.95}$$

From this it's possible to define 3 major cases for the transformation of angles between the lab and the $\star$ system after a collision.
1) $\beta > \beta^\star$
If $\beta > \beta^\star$ we have

$$\frac{\beta}{\beta^\star} - \cos\theta^\star > 0$$

Which implies that $\forall \theta^\star \in [0, \pi]$, $\theta \in [0, \pi/2]$.
This means that the particle after the collision, in the lab system will be observed as moving forwards with a flight angle θ between $0, \pi/2$ with respect to the initial motion.
Since also $\theta = 0$ for $\theta^\star = 0, \pi$, we have that there must exist a maximum flight angle $\theta_{max} < \pi/2$.
Deriving the previous equation with respect to $\theta^\star$ we get

$$\frac{d\tan\theta}{d\theta^\star} = \frac{1 + \frac{\beta}{\beta^\star} \cos\theta^\star}{\left(\frac{\beta}{\beta^\star} - \cos\theta^\star\right)^2} = 0$$

Which gives

$$\cos\theta^\star_{max} = -\frac{\beta^\star}{\beta} \tag{H.96}$$

Shoving it back into the equation for the tangent, we have

$$\tan\theta_{max} = \frac{\beta^\star}{\gamma\sqrt{\beta^2 - (\beta^\star)^2}} \tag{H.97}$$

Using $\gamma E^\star + \beta\gamma p^\star \cos\theta^\star = E$ we also have

$$E(\theta_{max}) = \gamma\left(E^\star - \beta^\star p^\star\right) = \gamma\left(\frac{m^2 - (p^\star)^2}{E^\star}\right) = \gamma\frac{m^2}{E^\star} = m\frac{\gamma}{\gamma^\star} \tag{H.98}$$

Where we defined

$$\gamma^\star = \frac{1}{\sqrt{1 - (\beta^\star)^2}} = \frac{m}{E^\star} \tag{H.99}$$

2) $\beta < \beta^\star$

In this case the velocity of the particle in the lab system can stop the center of mass, giving $\theta \geq \pi/2$. Note that there is no maximum angle since the derivative of the tangent is always positive.

3) $\beta = \beta^\star$

In this case $\cos\theta^\star = -1$ and it corresponds to a single possible angle $\theta = \pi/2$. Here the particle moves opposite to the center of mass ($\theta^\star = \pi$), while in the lab it's at rest.

§ H.12 Decays and Threshold Energy

Consider now a particle colliding with a target at rest, such that after the collision n particles are produced.

Definition H.12.1 (Threshold Energy). The «threshold energy» of the reaction is defined as the minimal kinetic energy T_{th} that the projectile needs in order to produce all the n particles at rest in the $\star$ system.

In order to find this T_{th} we have that the invariant mass in the final state is

$$\sqrt{s} = \sum_{f=1}^{n}\left(T_f^\star + m_f\right) \tag{H.100}$$

Whereas in the initial state

$$s = (E_i + m_T)^2 - p^i p_i = 2m_T E_i + E_i^2 + m_p^2 \tag{H.101}$$

Where m_T is the mass of the target and m_p is the mass of the projectile.

Writing $T_i = (\gamma - 1)E_i = E_i - m_i$ We can rewrite the invariant mass before the collision as follows

$$s = 2m_T T_i + 2m_T m_i + (m_i^2 + m_T^2)$$

Which gives

$$\sqrt{s} = \sqrt{2m_T T_i + (m_i + m_T)^2} \tag{H.102}$$

Equating $\sqrt{s}$ before and after the collision (it's a relativistic invariant) we have the following equality

$$\left(\sum_{f=1}^{n}\left(T_f^\star + m_f\right)\right)^2 = 2m_T T_i + (m_i + m_T)^2 \tag{H.103}$$

Solving for T_i we get

$$T_i = \frac{\left(\sum_{f=1}^{n}(T_f^\star + m_f)\right)^2 - (m_i + m_T)^2}{2m_T} \tag{H.104}$$

Setting $T_f^\star = 0$ we find the value of T_i such that the particles are produced at rest in the center of mass, i.e. the threshold energy for the reaction.

$$T_{th} = \frac{\left(\sum_{f=1}^{n} m_f\right)^2 - (m_i + m_T)^2}{2m_T} \tag{H.105}$$

Note that if $T_i < T_{th}$ the reaction is *kinematically impossible*.

§ H.13 Elastic Scattering

In the event of elastic scattering between particles, the classical conservation of energy and momentum in special relativity translattes into the conservation of the 4-momentum, i.e.

$$p_1^\mu + p_2^\mu = p_1^{\mu'} + p_2^{\mu'} \tag{H.106}$$

Consider now the scattering between an electron e^- and the nucleus of an atom A. We have that if in the lab A is at rest, we can write

$$p_{e^-}^\mu = \left(E_{e^-}, p_{e^-}^i\right), \quad p_A^\mu = (M, 0) \tag{H.107}$$

After the collision we get

$$p_{e^-}^{\mu'} = \left(E_{e^-}', p_{e^-}^{i'}\right), \quad p_A^{\mu'} = \left(E_A, p_A^{i'}\right) \tag{H.108}$$

We must have $P^\mu = P^{\mu'}$, therefore, using $P^\mu P_\mu = P^{\mu^\star} P_{\mu^\star}$ we have

$$P^\mu P_\mu = p_{e^-}^\mu p_\mu^{e^-} + p_A^\mu p_\mu^A + 2p_A^\mu p_\mu^{e^-} = m_{e^-}^2 + M^2 + 2p_{e^-}^{\mu'} p_{\mu'}^{e^-} \tag{H.109}$$

Note that experimentally only the electron gets measured after the scattering, therefore it's convenient to write the following

$$p_A^{\mu'} = p_{e^-}^\mu + p_A^\mu - p_{e^-}^{\mu'}$$

Therefore

$$p_{e^-}^\mu p_\mu^A = p_{e^-}^{\mu'} \left(p_\mu^{e^-} + p_\mu^A - p_{\mu'}^{e^-}\right) = p_{e^-}^{\mu'} p_\mu^{e^-} + p_{e^-}^{\mu'} p_\mu^A - m_{e^-}^2$$

In the lab system we have that the previous equation becomes

$$E_{e^-} M = E_{e^-}' E_{e^-} + E_{e^-}' M - m_{e^-}^2 - p_{e^-}^{i'} p_i^{e^-}$$

If the electron is ultrarelativistic we have

$$EM = E'E + E'M - pp'\cos\theta \Longrightarrow E' = \frac{E}{1 + \frac{E}{M}(1 - \cos\theta)} \tag{H.110}$$

Where θ is the final scattering angle

§ H.14 N-body Decays

Consider a particle of mass M decaying in n particles with masses m_i. Considering the reference system where M is at rest, we can say using the conservation of energy, that

$$M = \sqrt{s} \sum_{i=1}^{n} E_i^\star = \sum_{i=1}^{n} \sqrt{(p_i^\star)^2 + m_i^2} \geq \sum_{i=1}^{n} m_i \qquad \text{(H.111)}$$

Note that this defines also a threshold energy for the decay

$$M \geq \sum_{i=1}^{n} m_i$$

Consider now, without loss of generality, a 2 particle decay $a \rightarrow b + c$. We have

$$M = \sqrt{s} = E_b^\star + E_c^\star$$
$$p_c^\star = p_b^\star \qquad \text{(H.112)}$$

Using (H.111) we have that

$$M = \sqrt{p^2 + m_a^2} + \sqrt{p^2 + m_b^2}$$

Therefore, juggling a bit with the equation, we end up with

$$M^2 + m_b^2 - m_c^2 = 2M\sqrt{(p^\star)^2 + m_b^2}$$
$$M^4 + (m_b^2 - m_c^2) + 2M^2(m_b^2 - m_c^2) = 4M^2\left((p^\star)^2 + m_b^2\right)$$

Solving for $p^\star$ we get

$$p^\star = \frac{1}{2M}\sqrt{M^4 + (m_b^2 - m_c^2)^2 - 2M^2(m_b^2 + m_c^2)} \qquad \text{(H.113)}$$

Using the dispersion relation for $E^\star$ we get

$$E_b^\star = \frac{M^2 + (m_b^2 - m_c^2)}{2M} \qquad \text{(H.114)}$$

And using $\sqrt{s} = M$, we have

$$E_b^\star = \frac{s + m_b^2 - m_c^2}{2\sqrt{s}} \qquad \text{(H.115)}$$

The calculation is completely analogous for $E_c^\star$, and we get

$$E_c^\star = \frac{s - m_b^2 - m_c^2}{2\sqrt{s}} \qquad \text{(H.116)}$$

It's obvious that this decay has only one possible decay energy, and hence it's called «monoenergetic» . This is not true for decays with $n \geq 2$.

In the $\star$ system the decay is isotropic due to the conservation of 3-momentum, but the direction of one particle with respect to the other is fixed, where

$$p_b^\star = -p_c^\star \Longrightarrow \theta_{bc}^\star = \pi$$

In this case $\star$ coincides with the lab system, therefore $\theta_{bc} = \pi$.

Consider now another case, where the particle decays in flight, with $\beta \neq 0$. In this case the angle measured in the lab system will be obviously different from the one in the $\star$ system. Using what we got before for defining $\beta, \beta^\star, \gamma$ using momentum and energy, we have our usual transformation of angles for the first particle at $\theta^\star$ and the second at $\pi - \theta^\star$

$$\tan \theta_{bc} = \frac{\sin \theta^\star}{\gamma_a \left(\beta_a \beta_{bc}^\star \pm \cos \theta^\star \right)} \tag{H.117}$$

In general, with n particles, we have three cases

1) $\beta_a > \beta_i$

In the lab system the i-th child particle is emitted forwards, with a maximum angle $\theta_{max} < \pi/2$, corresponding to the angle $\theta_i^\star = \arccos \left(-\beta/\beta_i^\star \right)$

2) $\beta_a < \beta_i$

The i-th child particle is emitted forwards in the lab system if and only if $\cos \theta_i^\star > -\beta_a/\beta_i^\star$. At that value the particle gets emitted at $\theta_i = \pi/2$, while if $\cos(\theta_i^\star) < -\beta_a/\beta_i^\star$ it gets emitted backwards.

Bibliography

[BJ90]   B. H. Bransden and C. J. Joachain. «Physics of Atoms and Molecules». Longman Scientific & Technical, John Wiley & Sons Inc., 1990.

[BS17]   G. B. Bachelet and V. D. P. Servedio. «Elementi di Fisica Atomica, Molecolare e dei Solidi». Aracne Editrice, 2017.

[Dem03] W. Demtröder. «Molekülphysik, Teoretische Grundlagen und Experimentelle Methoden». Oldenbourg Verlag München Wien, 2003. isbn: 3486249746.

[DL17]   C. Dionisi and E. Longo. «Fisica Nucleare e Subnucleare». 2017.

[Gri05] D. J. Griffiths. «Introduction to Quantum Mechanics». Pearson Education International, 2005.

[Gri87] D. J. Griffiths. «Introduction to Elementary Particles». WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, 1987. isbn: 9783-527-40601-2.

[LLP89] L. D. Landau, E. M. Lifshits, and L. P. Pitajevsky. «Course of Theoretical Physics, Vol.3, Non-relativistic Quantum Mechanics». MIR, 1989.

[Par11] R. Paramatti. «Cinematica Relativistica». 2011.

[Par18] M. Tanabashi et al. (Particle Data Group). «Particle Physics Booklet». Particle Data Group, LBNL, CERN, 2018. url: http://pdg.lbl.gov/.

[Pd16]   L. E. Picasso and E. d'Emilio. «Lectures in Quantum Mechanics». Springer International Publishing, 2016.

[Pre13] C. Presilla. «Elementi di Analisi Complessa, Funzioni di una Variabile». Springer, 2013. isbn: 9788847055001.

[PT15]   S. Patrì and M. Testa. «Fondamenti di Meccanica Quantistica». Edizioni Nuova Cultura, 2015.

[Sak94] J. J. Sakurai. «Modern Quantum Mechanics». Addison Wesley Publishing Company, 1994.

[Sch16] F. Schwabl. «Statistical Mechanics». Springer-Verlag Berlin Heidelberg, 2016.

[Sol98] E. Massó i Soler. «Curs de Relativitat Especial». Universitat Autònoma de Barcelona, Servei de Publicacions, 1998. isbn: 84-490-1284-8.

[ZD97]   Mark W. Zemansky and Richard H. Dittman. «Heat and Thermodynamics».
McGraw-Hill, 1997. isbn: 0-07-017059-2.