Theroetical Physics

Notes from master degree

Luca Morelli

Copyright © 2023 Luca Morelli

PUBLISHED BY LUCA MORELLI

LUCAMORELLI.NETLIFY.APP

Licensed under the Creative Commons Attribution-NonCommercial 3.0 Unported License (the "License"). You may not use this file except in compliance with the License. You may obtain a copy of the License at `http://creativecommons.org/licenses/by-nc/3.0`. Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

First printing, WIP

Contents

IV Quantum field theory

Introduction

These notes are the result of my own studies during my master dregree in theoretical physics at the University of Bologna. I will try to cover most of the topics I have learned (and I'm learining) tring to explain everything in the best way I can.
I want to divide these notes in *parts* by different topics.

I Fundaments of theoretical physics

1. The fundaments of physics

1.1 The principles of physics

There is no physics without measurements, otherwise it would be just mathematics.
As theoretical physicists there is no need to know the whole procedure behind experiments and measurements, we just need to know that what we are studying can be measured and therefore our assumptions must be verified by some kind of experimental result.
Measurements must, from the definition of the scientific method, be reproducible: this means that everybody, everywhere in the universe should be able to replicate the results of a certain experiment. Since the outcome of an experiment is determined by the laws of nature and everybody should agree on such outcome, it seems reasonable to think that laws of nature should be the same for everybody.
Actually this isn't so obvious and thus require some further understanding of the reality.
In order to proceed we remind the intrinsic experimental nature of physics, in fact we are going to define some *principles*: those are fundamentally different from the axioims of mathematic. The principles of physics are based always on some empirical observation and thus they are not arbitrary.

1.1.1 Space and time

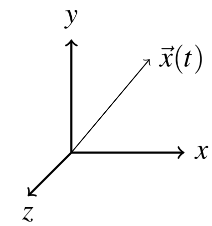
First we are going to assume the existence of space and time: this assumption is actually obvious since we experience every day the space, in which we live in, and the flow of time. It is less obvious which mathematical proprieties to give to those two concepts. We will assume from empirical evidence that:

- **space** is *3 dimensional, isotropic, homogeneous* and it obeys *euclidean geometry*;
- **time** is *1 dimensional, isotropic and homogeneous* too.

Given these proprieties space and time can be studied as an affine space (since they are isotropic and homogeneous) that we call **universe**.

A body moving in space will result in a curve in the universe. If we want to describe that motion it is convenient to use a set of coordinate in a vector space: this can be done by choosing an arbitrary point in time, from which we will start to measure time itself, and arbitrary point in space, as the null vector, and then three linearly independent vectors as basis. Let's observe that all of these choices are arbitrary only if space and time are isotropic and homogeneous, as we assumed in the beginning.
The procedure described is the mathematical equivalent of choosing a **reference frame** which allow us to describe motion as the curve:

$$\vec{x}(t), \qquad \vec{x} \in \mathbb{R}^3,\ t \in \mathbb{R}.$$



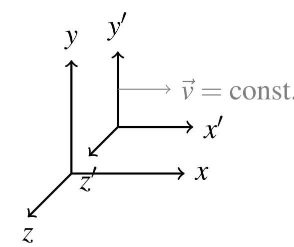
The vector $\vec{x}(t)$ in the $\mathbb{R}^3$ vector space of the reference frame.

1.1.2 Inertial reference frames and relativity

Now we have all the tools to discuss the problem, that we mentioned at the beginning of this chapter, of having the same laws of nature for everybody.

First we need to define a special type of reference frame, called **inertial**.

> **Definition 1.1.1** An inertial reference frame is such if and only if the motion of its origin, seen in another inertial reference frame, is uniform and rectilinear.



The primed reference frame moving with constant velocity $\vec{v}$ and thus being inertial.

This definition might seem not well-structured, since we need a preexisting reference frame. Using other principles of physics we will acquire new ways to determine if a reference frame is inertial.

The notion of inertial reference frame is totally useless until we postulate the **principle of relativity**, as the solution of our initial problem.

> **Principle 1.1.1 — of Relativity.** All the laws of physics are the same in all inertial reference frames.

In this way it makes totally sense to describe phenomena and pretend to determine the universal laws of nature from experiments and thus physics can exist only if these principles hold, or some kind of their generalization.

What we have assumed until now represent the most fundamental axioms of physics: every theory, from classical mechanics, to relativity, up to quantum mechanics, needs these assumptions and by adding others they can be constructed.

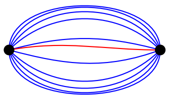
1.1.3 Transformations of inertial reference frames

1.2 Principle of least action

The most important fundamental principle of physics is, without any doubt, the **principle of least action**. This principle gives us a way to describe the time evolution of a system: firstly of mechanical systems and then, by its generalization, to fields.

In order to describe a mechanical system we define a function L called **lagrangian**, which depends on the *generalized coordinates* of the systems $q_1, q_2, \dots, q_n$ and their time derivatives $\dot{q}_1, \dot{q}_2, \dots, \dot{q}_n$[1].

[1] $\dot{q} \triangleq \frac{dq}{dt}$.



Different evolutions of a system through time; only the red one minimize the action and thus is the real one.

Principle 1.2.1 — Least action. The time evolution of a mechanical system, described by n generalized coordinates $q_1, q_2, \dots, q_n$, between t_1 and t_2, with fixed initial and final coordinates, is given by the functions that minimize the **action** functional $\mathcal{S}$.

$$\delta\mathcal{S} = \delta\int_{t_1}^{t_2} dt\, L(q_1(t), q_2(t), \dots, q_n(t), \dot{q}_1(t), \dot{q}_2(t), \dots, \dot{q}_n(t), t) = 0. \tag{1.1}$$

In this way we just need to find the right lagrangian to describe the system in the right framework (classical or relativistic) and then the approach to integrate the system will be the same.

1.2.1 Euler-Lagrange equations

Given the principle 1.2.1 we can derive a set of differential equations which solutions are the equations of motion, this can be done through the **Euler-Lagrange equations**.

Theorem 1.2.1 The functions that make stationary the action functional (defined in (1.1)) are the solutions of the differential equations given by:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \qquad \forall i \in \{1, 2, \dots, n\}. \tag{1.2}$$

Proof. We will use the shorthand for all the generalized coordinates $q = (q_1, q_2, \dots, q_n)$. In order to find the set of functions $q(t)$ which make stationary the action we will consider a second arbitrary set of functions $\delta q(t) = (\delta q_1(t), \delta q_2(t), \dots, \delta q_n(t))$ such as these are arbitrary small in the interval (t_1, t_2) and they vanish for $t = t_1$ and $t = t_2$. This last requirement is such that $q(t) + \delta q(t)$ has the same initial and finale coordinates of $q(t)$ in the interval $[t_1, t_2]$.
The variation of the action given by $q(t) \to q(t) + \delta q(t)$ is:

$$\delta\mathcal{S} = \mathcal{S}[q(t) + \delta q(t)] - \mathcal{S}[q(t)] = \int_{t_1}^{t_2} dt\, \Big[L\big(q(t) + \delta q(t), \dot{q}(t) + \delta\dot{q}(t), t\big) - L\big(q(t), \dot{q}(t), t\big)\Big].$$

Expanding $L\big(q(t) + \delta q(t), \dot{q}(t) + \delta\dot{q}(t), t\big)$ in Taylor series the integral becomes (at the first order of expansion):

$$\delta\mathcal{S} \approx \int_{t_1}^{t_2} dt\, \left[\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta\dot{q}\right] = \int_{t_1}^{t_2} dt\, \sum_{i=0}^{n}\left[\frac{\partial L}{\partial q_i}\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\delta\dot{q}_i\right]$$

From integration by parts of the terms containing $\delta\dot{q}_i = \frac{d}{dt}\delta q_i$ and considering that $\delta q_i(t_1) = \delta q_i(t_2) = 0$, we get:

$$\delta\mathcal{S} \approx \left[\frac{\partial L}{\partial \dot{q}_i}\delta q_i\right]_{t_1}^{t_2} + \int_{t_1}^{t_2} dt\, \sum_{i=0}^{n}\left[\frac{\partial L}{\partial q_i}\delta q_i - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\delta q_i\right] = \int_{t_1}^{t_2} dt\, \sum_{i=0}^{n}\left[\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right]\delta q_i = 0$$

In order to have $\delta\mathcal{S} = 0$ the integrand must vanish and, considering that δq_i is an arbitrary function, all the terms in square brackets should independently go to zero, resulting in (1.2). ■

Let's observe two consequences of this formulation of the principle of least action:

- since in the lagrangian appear only first time derivatives, the differential equations resulting from (1.2) are second order ordinary differential equations and thus every solution is unique given initial positions and velocities.
- if in the lagrangian doesn't appear one coordinate q_k then:

$$\frac{\partial L}{\partial q_k} = 0 \Rightarrow \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} = 0 \Rightarrow \frac{\partial L}{\partial \dot{q}_k} = \text{const.}$$

1.2.2 Noether's theorem

1.2.3 Euler-Lagrange equations for fields

As we have already said, the principle of least action (Principle 1.2.1) can be generalized in order to describe the time evolution of a field[2]. In this case the system will be described by a **lagrangian density** $\mathcal{L}$, a function of the field φ, its spacial and temporal first partial derivatives and space-time coordinates.

Notazione 1.1. *We will use the shorthand notation* $x^\mu = (t,x,y,z)$ *and* $\partial_\mu = (\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$*. This notation in taken from special relativity and will become clearer after the introduction of 4-vectors.*

Principle 1.2.2 — Least action. Given a lagrangian density $\mathcal{L}(\varphi, \partial_\mu\varphi, x^\mu)$, defined on a volume Ω in space-time, if φ and $\partial_\mu\varphi$ vanish on the frontier $\partial\Omega$ then the measurable field is the one that minimize the action functional $\mathcal{S}$:

$$\delta\mathcal{S} = \delta\int_\Omega d^4x\, \mathcal{L}(\varphi, \partial_\mu\varphi, x^\mu) = 0. \tag{1.3}$$

We can now derive the Euler-Lagrange equations for fields, using the same approach we used to obtain (1.2).

Theorem 1.2.2 The functions that make stationary the action functional (defined in (1.3)) are the solutions of the differential equations given by:

$$\sum_{\mu=0}^{3} \partial_\mu \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi} - \frac{\partial\mathcal{L}}{\partial\varphi} = 0. \tag{1.4}$$

Proof. Let's consider another field $\delta\varphi$, defined on Ω and arbitrarily small, such that it vanishes on the frontier $\partial\Omega$. For the field $\varphi + \delta\varphi$ all the hypothesis of the principle of least action 1.2.2 are still valid. Furthermore, all the coordinates are not changed.
In order to be the minimizing field of the action functional, the variation $\delta\mathcal{S}$, caused by $\varphi \to \varphi$, has to vanish:

$$\delta\mathcal{S} = \mathcal{S}[\varphi+\delta\varphi] - \mathcal{S}[\varphi] = \int_\Omega d^4x \left[\mathcal{L}(\varphi+\delta\varphi, \partial_\mu\varphi + \partial_\mu\delta\varphi, x^\mu) - \mathcal{L}(\varphi, \partial_\mu\varphi, x^\mu)\right] = 0.$$

[2]Fields (maps from $\mathbb{R}^n$ to $\mathbb{R}$) are usually used to describe continuous systems, such as a rope vibrating, but will be necessary in order to describe particles in quantum field theory.

Taylor expanding at the first order the integrand (with respect to $\delta\varphi$):

$$\delta\mathcal{S} \approx \int_\Omega d^4x \left[\frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \sum_{\mu=0}^{3}\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\partial_\mu\delta\varphi\right].$$

Using integration by parts and the generalized theorem of calculus we get:

$$\delta\mathcal{S} \approx \int_\Omega d^4x \left[\frac{\partial\mathcal{L}}{\partial\varphi} - \sum_{\mu=0}^{3}\partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\right]\delta\varphi + \int_{\partial\Omega} d\sigma\, \sum_{\mu=0}^{3}\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}n_\mu\delta\varphi,$$

where $d\sigma$ is the surface element of $\partial\Omega$ and n_μ are the components of the normal vector in space-time to $d\sigma$. The last integral is equal to zero, because $\delta\varphi$ vanishes on $\partial\Omega$, thus to have $\delta\mathcal{S} = 0$ for every arbitrary $\delta\varphi$ has to vanish the term in square brackets:

$$\frac{\partial\mathcal{L}}{\partial\varphi} - \sum_{\mu=0}^{3}\partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi} = 0.$$

■

1.3 The phase space

2. Special relativity

3. Mechanics

4. Quantum mechanics

II Statistical mechanics

5. Basic thermodynamics concepts

5.1 Thermodynamic systems

Classical mechanics is able to predicts the motion of bodies using the differential equation given by the principle of least action 1.1. Usually these are some kind of second order ordinary differential equations, thus, in order to get a certain solution, we should use initial conditions such as the initial position and the initial velocity. This approach results in exact solutions only if the system is made up of few particles (usually less than 3) or if we impose other assumption, such as that all the particles form a rigid body.
Studying in this way a gas or a fluid is practically impossible, such systems will have a number of particles of the order of at least $10^2$3, in these cases, it is not just mathematically impossible to get some exact solutions, but it becomes impossible to manage so many initial conditions (at least 6 for every particle). In the 19^{th} century it was developed a new branch of physics that tried to study these systems using, apparently, non-mechanical quantities: thermodynamics.

Every thermodynamic system is characterized by the so-called **thermodynamic variables**: quantities defined empirically that describe microscopically the system. There are two main types of these:

- **extensive variables** that scale with the system (energy, entropy, volume, polarization, ...);
- **intensive variables** that doesn't scale with the system (temperature, pressure, chemical potential, ...).

Every the extensive variable has an intensive conjugate (such as volume and pressure) and vice versa.
The variables of a system are connected each other by the **state equation of the system**.

5.2 The laws of thermodynamic

In general thermodynamic variables are connected by the **laws of thermodynamic**. These hold for every system and are the empirical axioms of thermodynamics.

Law 5.2.1 — Zeroth. Two systems in thermal contact have the same empirical temperature at equilibrium.

As we can see the zeroth law defines what does equilibrium means in thermodynamics, and to do so it defines **temperature** as the variable that determines if two systems are at equilibrium.

Law 5.2.2 — First. The variation of internal energy of a system is given by:

$$dE = \delta Q - \delta L + \mu dN, \tag{5.1}$$

where Q is the heat, L is the work, μ the chemical potential and N the number of particles of the system.

This law connects the main variables of the system together. We should notice that, while E and N are actual thermodynamic variables, the same is not true for Q and L, that are not some kind of intrinsic proprieties of the system, but the just describe how energy is transferred. This is reflected in the fact that in (5.1) these two are not exact differentials, but they depend on the type of transformation the system is subject to.

Law 5.2.3 — Second. The variation of internal entropy of a system is such:

$$ds \geq \frac{\delta Q}{T}, \tag{5.2}$$

where the equality holds only if the system undergoes a **reversible process**.

The second law of thermodynamic define what is a reversible process giving an exact formulation of what entropy is in that case. Further, it tells us that in the case of reversible processes the heat becomes a exact differential. This law can have others equivalent formulations that state that heat always moves from a hotter body to a colder one.

Law 5.2.4 — Third. The variation of entropy goes to 0 as the temperature goes too for any **reversible isothermal** process.

The last law actually states that it is not possible to reach the **absolute zero** temperature using a finite amount of reversible processes.
We can now use the first principle, together with the observation that energy must be extensive and the second law for reversible processes, to get the explicit form of this state function:

$$E = TS - PV + \mu N. \tag{5.3}$$

From this we also get explicit relations between intensive and extensive quantities:

$$T = \left.\frac{\partial E}{\partial S}\right|_{V,N}, \qquad P = -\left.\frac{\partial E}{\partial V}\right|_{S,N}, \qquad \mu = \left.\frac{\partial E}{\partial N}\right|_{V,S}. \tag{5.4}$$

5.3 Thermodynamic potential

Using the Legrand transform of the energy (5.3) it is easy to obtain other functions, that will be really useful later on, called **thermodynamic potentials**:

- **Helmholtz free energy**: $F(T,V,N) = E - TS$;
- **Enthalpy**: $H(S,P,N) = E + PV$;
- **Gibbs free energy**: $G(T,P,N) = E - TS + PV$;
- **Grand potential**: $\Omega(T,V,\mu) = E - TS - \mu N$.

These can be used with the second law ($ds \geq \frac{\delta Q}{T}$) to get relations that define equilibrium for the system and how it evolves to reach it. Using the Helmholtz free energy, for example:

$$\begin{aligned} dF &= dE - SdT - TdS = \delta Q - \delta L + \mu dN - SdT - TdS \\ &\leq \delta Q - \delta L + \mu dN - SdT - \delta Q = -PdV + \mu dN - SdT. \end{aligned}$$

If we take T,N,V constants we get a minimum principle (since having them constants implies that their differentials are zero):

$$dF \leq 0,$$

which implies that the system, at constant temperature, volume, and amount of matter, will evolve in such a way that F is decreasing.
Studying where F has a minimum, we can get the conditions to have equilibrium:

$$\begin{cases} \frac{\partial^2 F}{\partial T^2}\big|_V < 0, \\ \frac{\partial^2 F}{\partial V^2} > 0, \\ \frac{\partial^2 F}{\partial T \partial N} = \frac{\partial \mu}{\partial T} \leq 0. \end{cases}$$

6. Classical statistical mechanics

6.1 The microcanonical ensemble

The first statistical approach to the study of a thermodynamic system is the **microcanonical ensemble**. The systems we will study are **perfectly isolated** and thus characterized by fixed energy, volume and amount of matter.
In phase space, such systems will evolve on surfaces of constant energy $\mathcal{H}(q_i, p_i) = E$.
We will define a probability distribution on phase space, such that $\rho(q_i, p_i)dpdq$ will be the probability of finding the system in the microstate (q_i, p_i).

The microcanonical probability distribution is built on the assumption of **a priori uniform probability** on the surface of constant energy S_E. In this way the probability distribution reads:

$$\rho_{mc}(q_i, p_i) = \frac{\delta(\mathcal{H} - E)}{\omega(E)}, \tag{6.1}$$

where δ is the Dirac's delta function and $\omega(E)$ is area of the surface S_E.
Using this probability we can evaluate the average of some quantities defined on the phase space:

$$\langle f \rangle_{mc} = \int_{\mathcal{M}^N} d\Omega \, f(q_i, p_i) \frac{\delta(\mathcal{H} - E)}{\omega(E)} = \frac{1}{\omega(E)} \int_{S_E} dS_E \, f(q_i, p_i).$$

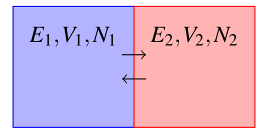
6.1.1 Microcanonical entropy

We want to define the entropy of the microcanonical ensemble: this should be extensive (as thermodynamic entropy) and should be a function of the energy. Since combining two systems gives a total volume in phase space equals to the product of the volume of the two isolated systems, we will define:

$$S_{mc} = k_b \log \Gamma(E), \tag{6.2}$$

because the logarithm turns multiplication into sums.

Let's consider two systems characterized by (E_1, V_1, N_1) and (E_2, V_2, N_2). These two systems can exchange energy and thus together the form an isolated system. We will assume that the total energy of the system is the sum of the energies of the two, because in the thermodynamic limit every energy contribution due to interaction on the surface of contact will be negligible.



The two systems that we are studying.

The phase space of the whole system is described by $(q_i^1, p_i^1, q_j^2, p_j 2)$ (the coordinates of one system and of the other one) and will have a volume element $d\Gamma = d\Gamma_1 d\Gamma_2$.
The area of the surface of constant energy (E) in the total phase space is:

$$\begin{aligned}
\omega(E) &= \int_{\mathcal{M}} d\Lambda_1 d\Lambda_2 \delta(\mathcal{H} - E) \\
&= \int dE_1 \int dS_{E_1} \int dE_2 \int dS_{E_2} \delta(\mathcal{H} - E) \\
&= \int dE_1 \int dE_2 \delta(\mathcal{H} - E) \omega(E_1) \omega(E_2) \\
&= \int_0^E dE_1 \omega(E_1) \omega(E - E_1) \leq \max_{E_1 \in [0,E]} \omega(E_1) \omega(E - E_1) E.
\end{aligned}$$

Defining E^* as the energy that maximize the above expression:

$$\omega(E) \Delta E \leq \omega(E_1^*) \omega(E - E_1^*) E \Delta E.$$

Now, for some ΔE it is true:

$$\Delta E \, \omega(E_1^*) \omega(E - E_1^*) \leq \omega(E),$$

therefore

$$(\Delta E)^2 \omega(E_1^*) \omega(E - E_1^*) \leq \Delta E \, \omega(E) \leq \omega(E_1^*) \omega(E - E_1^*) E \Delta E.$$

Considering the volume $\Gamma(E) = \Delta E \, \omega(E)$ the above inequality reads:

$$\Gamma_1(E_1^*) \Gamma_2(E - E_1^*) \leq \Gamma(E) \leq \Gamma_1(E_1^*) \Gamma_2(E - E_1^*) \frac{E}{\Delta E},$$

taking the logarithm of all the sides:

$$\log \Gamma_1(E_1^*) + \log \Gamma_2(E - E_1^*) \leq \log \Gamma(E) \leq \log \Gamma_1(E_1^*) + \log \Gamma_2(E - E_1^*) + \log \frac{E}{\Delta E}.$$

In the thermodynamic limit the term $\log \frac{E}{\Delta E} \to 0$, thus at equilibrium the entropy of each system is maximum and the total entropy[1] is given by the sum of the entropy of each subsystem:

$$s(E) = s_1(E_1^*) + s_2(E_2^*).$$

[1] Since we are at the thermodynamic limit we cannot use the entropy because it would be infinite, we thus use the entropy per unit volume.

If we consider a variation of energy of the subsystems $\delta E_1 = -\delta E_2$ that will cause a variation of the volume Γ, this last one must be zero at equilibrium:

$$\begin{aligned}
&\delta\Lambda = \left.\frac{\partial \Gamma_1}{\partial E_1}\right|_{E_1^*} \delta E_1 \Gamma_2 + \Gamma_1 \left.\frac{\partial \Gamma_2}{\partial E_2}\right|_{E_2^*} \delta E_2 = 0 \\
&\Rightarrow \left(\frac{1}{\Gamma_1} \left.\frac{\partial \Gamma_1}{\partial E_1}\right|_{E_1^*} - \frac{1}{\Gamma_2} \left.\frac{\partial \Gamma_2}{\partial E_2}\right|_{E_2^*} \right) \delta E_1 = 0 \\
&\Rightarrow \left.\frac{\partial \log \Gamma_1}{\partial E_1}\right|_{E_1^*} = \left.\frac{\partial \log \Gamma_2}{\partial E_2}\right|_{E_2^*}
\end{aligned}$$

This shows that the microcanonical entropy we have defined, at equilibrium, plays the same role as the thermodynamic entropy, since $\frac{\partial S}{\partial E}\big|_{V,N} = \frac{1}{T}$.

Lastly, we can show that we can evaluate the entropy in two other ways: first using that in the microcanonical case $\Gamma(E) = \Delta E \omega(E)$ with some arbitrary small δE, we get

$$S = k_B \log \Gamma = k_B \log \omega.$$

Given this observation we can prove that the **Boltzmann universal formula** holds:

$$\langle \log \rho_{mc} \rangle = \int_{\mathcal{M}} d\Gamma \rho_{mc} \log \rho_{mc} = \int_{S_E} dS_E \frac{\log \frac{1}{\omega(E)}}{\omega(E)} = \log \frac{1}{\omega(E)} \int_{S_E} \frac{dS_E}{\omega(E)} = -\frac{S}{k_B}.$$

■ **Example 6.1 — The ideal gas.** We will now study, as an example, the ideal gas. As ideal gas we mean N classical, indistinguishable, non-interacting particles in a finite volume V. The Hamiltonian is $\mathcal{H} = \sum_{i=1}^{n} \frac{p_i^2}{2m}$.
In order to calculate $\omega(E)$ we can use that, given $\Sigma(E)$ the volume in phase space whose points have energy lower than E, $\omega(E) = \frac{\partial \Sigma}{\partial E}$.

$$\begin{aligned}
\Sigma(E) &= \int_{0<\mathcal{H}<E} \frac{\Pi_i^N d^3 q_i d^3 p_i}{h^{3N}} = \frac{1}{h^{3N}} \int_V \Pi_i^N d^3 q_i \int_{0<\sum p_i^2<2mE} \Pi_i^N d^3 p_i \\
&= \frac{V^N}{h^{3N}} \Omega_{3N}(\sqrt{2mE}),
\end{aligned}$$

where $\Omega_{3N}(\sqrt{2mE})$ is the volume of a $3N$-dimensional sphere with radius $\sqrt{2mE}$, ad h is some constant that has the dimension of an action. Ignoring the all the multiplicative constants $\Omega_{3N}(\sqrt{2mE}) \propto (2mE)^{\frac{3}{2}N}$, thus:

$$\frac{\partial \Omega}{\partial E} = k \frac{3N}{2} (2mE)^{\frac{3N}{2}-1} 2m = \frac{3N}{2E} \Omega \;\Rightarrow\; \omega(E) = \frac{\partial \Sigma}{\partial E} = \frac{3N}{2E} \Sigma.$$

From this result we can observe that, in the thermodynamic limit $N \to \infty$:

$$\frac{\log \omega(E)}{N} = \frac{\log \Sigma(E)}{N} + \frac{\log \frac{3N}{2E}}{N} \to \frac{\log \Sigma(E)}{N}.$$

This relation, we have proved just for the ideal gas, holds for every system.
In this way we can calculate entropy using $\Sigma(E)$; to evaluate it we need the explicit form of $\Omega_{3N}(\sqrt{2mE}) = \frac{(2\pi mE)^{\frac{3N}{2}}}{N\Gamma(\frac{3N}{2})}$, thus we get:

$$S = k_B \log \Sigma(E) = k_B \left\{ N \log V \left(\frac{2\pi mE}{h^2} \right)^{\frac{3}{2}} - \log N - \log \Gamma\left(\frac{3N}{2} \right) + \log \frac{2}{3} \right\}.$$

Considering the thermodynamic limit and the Stirling's approximation it reads:

$$S = k_B \left\{ N \log V \left(\frac{2\pi mE}{h^2} \right)^{\frac{3}{2}} - \frac{3N}{2} \log \frac{3N}{2} + \frac{3N}{2} \right\} = \frac{3}{2} N k_B + N k_B \log V \left(\frac{4\pi mE}{3Nh^2} \right)^{\frac{3}{2}}.$$

We should notice that the entropy we obtained in not extensive, this is due to the fact that we are not considering the indistinguishability of the particle, and so we are double counting some states in the integration process. Dividing Σ by $N!$ we discard all the double counted states and so (using Stirling' approximation) we obtain:

$$S = \frac{5}{2} N k_B + N k_B \log \frac{V}{N} \left(\frac{4\pi mE}{3Nh^2} \right)^{\frac{3}{2}},$$

which is now extensive.
From the entropy we have obtained we can find the internal energy:

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_{V,N} = \frac{3}{2} \frac{N k_B}{E} \quad \Rightarrow \quad E = \frac{3}{2} K_B T;$$

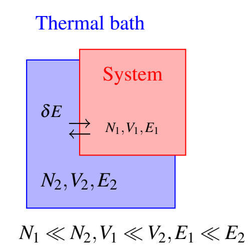
and the state equation:

$$\frac{P}{T} = \left. \frac{\partial S}{\partial V} \right|_{E,N} = \frac{N k_B}{V} \quad \Rightarrow \quad PV = N K_B T.$$

6.2 The canonical ensemble

We now want to study systems that can exchange energy, and thus they don't have a fixed one. Those systems are better studied by the **canonical ensemble**.
In order to study such systems we will introduce a thermal bath (denoted by the numer 2), much bigger than our system (denoted by the number 1). The two systems can exchange energy one to the other, but together they form an isolated system. These two are at thermal equilibrium ($T_1 = T_2 = T$). In this way we can tackle this problem using the microcanonical approach we have just developed in the last section 6.1. The total system will have energy $E = E_1 + E_2$ and thus the microcanonical probability distribution will be:

$$\rho_{mc} = \frac{\delta(\mathcal{H}_1 + \mathcal{H}_2 - E_1 - E_2)}{\omega(E_1 + E_2)}.$$



A simple sketch of the system and the thermal bath.

We now want to find the marginal distribution depending on just the variables of our system. To do so we have to integrate over the phase space of the thermal bath:

$$\int_{\mathcal{M}_2} \Pi_j^{N_2} dq_j^{(2)} dp_j^{(2)} \rho_{mc}(q_i^{(1)}, p_i^{(1)}, q_j^{(2)}, p_j^{(2)}) = \int_{\mathcal{M}_2} \Pi_j^{N_2} dq_j^{(2)} dp_j^{(2)} \frac{\delta(\mathcal{H}_1 + \mathcal{H}_2 - E_1 - E_2)}{\omega(E_1 + E_2)},$$

the Dirac's delta fixes the energy of the thermal bath in terms of the energy of our system. Recalling that $S = k_B \log \omega$, the integral above results in:

$$\Rightarrow \quad \frac{\omega_2(E_2 = E - E_1)}{\omega(E)} = \frac{1}{\omega(E)} \exp\left\{ \frac{1}{k_B} S_2(E - E_1) \right\}.$$

We now recall that the thermal bath is built as "bigger" than the system, mathematically we mean $E_1 \ll E_2$, therefore $E_1 \ll E$. This construction let us Taylor expand the exponential with respect to E_1, since $\frac{\partial S}{\partial E} = \frac{1}{T}$:

$$\frac{1}{\omega(E)} \exp\left\{ \frac{1}{k_B} \left[S_2(E) - \frac{\partial S_2}{\partial E} E_1 \right] \right\} = \frac{\exp\left\{ \frac{S_2(E)}{k_B} \right\}}{\omega(E)} \exp\left\{ -\frac{E_1}{k_B T} \right\}.$$

We have found the marginal probability distribution of our system, called canonical distribution. Since we don't want always have to use the entropy and energy of the bath, we put all the terms depending on those in a constant that we can evaluate using the normalization of the function.

$$\rho_C(q_i, p_i) = \frac{e^{-\beta \mathcal{H}}}{Z_N}, \qquad \beta = \frac{1}{k_B T}, \quad Z_N = \int_{\mathcal{M}^N} d\Omega \, e^{-\beta \mathcal{H}}. \tag{6.3}$$

We could think that Z_N, called the **partition function**, is just a constant (indeed it is a constant with respect of our distribution) but we should notice that it is dependent on the temperature of the system, due to β, the number of particles of the system and the volume of the system, due to the integration over $\mathcal{M}^N$. We will see that, for this reason, Z_N is related to a thermodynamic potential that depends on the same variables.

Let's observe that, as we saw in the microcanonical framework, this approach over counts the states of the system, assuming that all the particles are always distinguishable (as in the lattice of a crystal). To solve this over counting we have to divide the volume element in phase space by a factor $N!$, in this way $Z_N \to \frac{Z_N}{N!}$.

Lastly, let's consider a system made up of different types of particles, each one with its hamiltonian:

$$N = \sum_i N_i, \qquad \mathcal{H} = \sum_i \mathcal{H}_i.$$

Given Ω_i the volume element of the phase space of the i-esim system (scaled to account indistinguishability and remove the dimension), the total partition function reads:

$$\begin{aligned} Z_N &= \int_{\mathcal{M}^N} \Pi_i d\Omega_i \, e^{-\beta \sum_i \mathcal{H}_i} = \left(\int_{\mathcal{M}^{N_1}} d\Omega_1 e^{-\beta \mathcal{H}_1} \right) \left(\int_{\mathcal{M}^{N_2}} d\Omega_2 e^{-\beta \mathcal{H}_2} \right) \ldots \\ &= \Pi_i Z_{N_i}. \end{aligned}$$

This approach can be used to evaluate the partition function of a system of N identical particle as $Z_N = Z^N$, where Z is the partition function of one single particle.

6.2.1 Partition function and free energy

As already mentioned, the partition function depends on V,T,N. There is a thermodynamic potential, the Helmholtz's free energy, which depends on the same variables. We will show that defining an arbitray function $F(T,V,N)$, such that

$$Z_N(T,V) \triangleq e^{-\beta F(T,V,N)},$$

this function behave exactly as the Helmholtz's free energy.

First, from the above definition, the normalization condition of Eq. (6.3) reads

$$1 = \int d\Omega \frac{e^{-\beta\mathcal{H}}}{Z_N} = \int d\Omega\, e^{-\beta(\mathcal{H}-F)},$$

differentiating this expression with respect to β we get:

$$\begin{aligned}
0 &= \int d\Omega \frac{e^{-\beta\mathcal{H}}}{Z_N} = \int d\Omega \frac{\partial}{\partial\beta} e^{-\beta(\mathcal{H}-F)} = \int d\Omega \left[-(\mathcal{H}-F) + \beta\frac{\partial F}{\partial\beta}\right] e^{-\beta(\mathcal{H}-F)} \\
&= -\int d\Omega \frac{e^{-\beta\mathcal{H}}}{Z_N}\mathcal{H} + F + \beta\frac{\partial F}{\partial\beta} = -E + F + \beta\frac{\partial F}{\partial\beta} \\
&\Rightarrow\ E = F + \beta\frac{\partial F}{\partial\beta} = F - T\frac{\partial F}{\partial T}.
\end{aligned}$$

If F is the Helmholtz's free energy we should recall the relation

$$\frac{\partial F}{\partial T} = \frac{F-E}{T} = -S,$$

therefore, using the Boltzmann's universal formula as definition of entropy:

$$\begin{aligned}
S &= -k_B\langle \log\rho_C\rangle = -k_B\int d\Omega\, \rho_C\log\rho_C = -k_B\int d\Omega\, \rho_C\log\frac{e^{-\beta\mathcal{H}}}{Z_N} \\
&= -k_B\int d\Omega\, \rho_C\big(-\beta\mathcal{H} - \log Z_N\big) = \frac{E}{T} + k_B\frac{\beta}{\beta}\log Z_N \\
&= \frac{E}{T} + \left(-\frac{1}{T}\right)\left(-\frac{1}{\beta}\log Z_N\right) = \frac{E}{T} - \frac{F}{T},
\end{aligned}$$

which shows that these two definitions of the free energy and entropy lead to the same relations that holds in thermodynamic.

From these caluclation we have found the follwing relations, that can be used to easily get thermodynamic variables from the partition function:

$$\begin{aligned}
F &= -\frac{1}{\beta}\log Z_N, \\
E &= \int d\Omega\, \mathcal{H}\frac{e^{-\beta\mathcal{H}}}{Z_N} = -\frac{1}{Z_N}\int d\Omega\, \frac{\partial}{\partial\beta}e^{-\beta\mathcal{H}} = -\frac{1}{Z_N}\frac{\partial Z_N}{\partial\beta} = -\frac{\partial\log Z_N}{\partial\beta}.
\end{aligned}$$

6.2.2 Equipartition theorem

Using the canonical framework it is possible to derive a really useful result that can help to study systems without the need of determining the whole partition function. This is the **equipartition theorem** which allow to calculate the internal energy of a system just by studying its hamiltonian. We will now illustrate its generalized form.

Theorem 6.2.1 — Generalized equipartition theorem. Let $\xi_1 \in [a,b]$ be one of the canonical coordinates or momenta and ξ_j $j>1$ denotes all the others.
If holds that $\int \Pi_{j\neq 1} d\xi_j [\xi_1 e^{-\beta\mathcal{H}}]_a^b = 0$ then:

$$\left\langle \xi_1 \frac{\partial \mathcal{H}}{\partial \xi_1} \right\rangle = k_B T. \tag{6.4}$$

Proof. We will start by the normalization condition of the canonical distribution:

$$1 = \int d\Omega \frac{e^{-\beta\mathcal{H}}}{Z_N} = \frac{1}{Z_N} \int d\xi_1 (\Pi_{j\neq 1} d\xi_j) e^{-\beta\mathcal{H}}.$$

Since $\frac{d}{d\xi_1}\left[\xi_1 \exp(-\beta\mathcal{H})\right] = \exp(-\beta\mathcal{H}) - \beta\xi_1 \exp(-\beta\mathcal{H}) \frac{\partial\mathcal{H}}{\partial\xi_1}$ we can integrate by parts with respect to ξ_1:

$$\begin{aligned} 1 &= \frac{1}{Z_N} \int (\Pi_{j\neq 1} d\xi_j) \left(d\left[\xi_1 e^{-\beta\mathcal{H}}\right] + \beta \xi_1 e^{-\beta\mathcal{H}} \frac{\partial \mathcal{H}}{\partial \xi_1} d\xi_1 \right) \\ &= \frac{1}{Z_N} \int (\Pi_{j\neq 1} d\xi_j) \left[\xi_1 e^{-\beta\mathcal{H}}\right]_a^b + \frac{\beta}{Z_N} \int (\Pi_{j\neq 1} d\xi_j) \xi_1 e^{-\beta\mathcal{H}} \frac{\partial \mathcal{H}}{\partial \xi_1} d\xi_1. \end{aligned}$$

The fist integral is zero by hypothesis and the second correspond to the mean value of $\xi_1 \frac{\partial\mathcal{H}}{\partial\xi_1}$:

$$\frac{1}{\beta} = k_B T = \int d\Omega \frac{e^{-\beta\mathcal{H}}}{Z_N} \xi_1 \frac{\partial \mathcal{H}}{\partial \xi_1} = \left\langle \xi_1 \frac{\partial \mathcal{H}}{\partial \xi_1} \right\rangle$$

■

The standard equipartition theorem, can be obtained as corollary of the previous theorem.

Corollary 6.2.2 — Equipartition theorem. The internal energy of a system has a $\frac{k_B T}{2}$ contribute for each quadratic term that appears in the hamiltonian.

Proof. Let's consider a quadratic hamiltonian in ξ_1:

$$\mathcal{H} = A\xi_1^2 + G(\xi_2, \xi_3, \dots),$$

with G arbitrary function. Clearly ξ_1 satisfies the hypothesis of the generalized equipartition theorem 6.2.1, thus we can calculate:

$$k_B T = \left\langle \xi_1 \frac{\partial \mathcal{H}}{\partial \xi_1} \right\rangle = \langle \xi_1 2A\xi_1 \rangle \quad \Rightarrow \quad \langle A\xi_1^2 \rangle = \frac{k_B T}{2}.$$

■

Example 6.2 — The energy of the ideal gas. As we already discussed, an ideal gas is made up of non-interacting classical and indistinguishable particles. Therefore, its hamiltonian is the sum of N free particle hamiltonian:

$$\mathcal{H} = \sum_i^N \frac{\vec{p}_i^2}{2m}.$$

Using the equipartition theorem 6.2.2 we then get that the internal energy of the system is $E = \frac{3}{2}Nk_BT$. The same result can be obtained studying the partition function:

$$\begin{aligned}
Z_N &= \frac{Z_1^N}{N!} \\
Z_1 &= \frac{1}{h^3}\int_V d^3q \int_{\mathbb{R}^3} d^3p\, e^{-\beta\frac{\vec{p}^2}{2m}} = \frac{V}{h^3}\int_{\mathbb{R}^3} d^3p\, e^{-\beta\frac{\vec{p}^2}{2m}} \\
&= \frac{V}{h^3}\left(\int_{-\infty}^{+\infty} d^3p\, e^{-\beta\frac{p^2}{2m}}\right)^3 = \frac{V}{h^3}\left(\frac{2m\pi}{\beta}\right)^{\frac{3}{2}} = V\left(\frac{2m\pi k_BT}{h^2}\right)^{\frac{3}{2}}, \\
Z_N &= \frac{V^N}{N!}\left(\frac{2m\pi k_BT}{h^2}\right)^{\frac{3}{2}N}, \\
E &= -\frac{\partial \log Z_n}{\partial \beta} = \frac{3}{2}Nk_BT.
\end{aligned}$$

Example 6.3 — The energy of the ultrarelativistic gas. We can consider an ultra relativistic gas as a gas of relativistic particles whose kinetic energy is much bigger than the energy due to the rest mass, thus this last can be neglected:

$$\mathcal{H} = \sum_i^N \sqrt{c^2\vec{p}_i^2 + m^2c^4} \approx \sum_i^N c|\vec{p}_i|.$$

Since the hamiltonian isn't quadratic, we have to use the generalized equipartition theorem 6.2.1: notice that this hamiltonian satisfy its hypothesis because it is always positive, and thus $[|\vec{p}_i|\exp(-\beta\sum_i^N c|\vec{p}_i|)]_{-\infty}^{+\infty} = 0$. From the theorem we get:

$$\left\langle p_{i,x}\frac{\partial \mathcal{H}}{\partial p_{i,x}}\right\rangle = \left\langle p_{i,x}c\frac{p_{i,x}}{|\vec{p}_i|}\right\rangle = k_BT.$$

Summing this expression over all the components of the momentum of one particle:

$$3k_BT = \left\langle c\frac{p_{i,x}^2 + p_{i,y}^2 + p_{i,z}^2}{|\vec{p}_i|}\right\rangle = \left\langle c\frac{p_i^2}{|\vec{p}_i|}\right\rangle = \langle c|\vec{p}_i|\rangle.$$

Lastly, summing over all the particles we get:

$$\sum_i^n \langle c|\vec{p}_i|\rangle = \langle \mathcal{H}\rangle = 3Nk_BT.$$

■ **Example 6.4 — 1-D solid.** As the last example of the canonical ensemble framework we will discuss the case of a 1-D lattice of atoms, forming a sort of solid. Each atom is fixed to its own equilibrium position, and can oscillate around it. The lattice is therefore created by harmonic interactions between neighbors atoms, resulting in a total hamiltonian of the form:

$$\mathcal{H} = \sum_i \left(\frac{p_i^2}{2m} + \frac{m\omega^2}{2} q_i^2 \right).$$

We will find the total partition function by studying the single particle system:

$$Z_1 = \int_{-\infty}^{+\infty} dq\, e^{-\beta \frac{m\omega^2}{2} q^2} \int_{-\infty}^{+\infty} dp \frac{e^{-\beta \frac{p_i^2}{2m}}}{h} = \frac{\pi}{h} \sqrt{\frac{2k_B T}{m\omega^2} 2mk_B T} = \frac{2\pi k_B T}{\omega h}.$$

We must notice that every single particle is not indistinguishable from the others, because the lattice structure gives to each particle a precise position in it, thus we can distinguish two particles by their positions. The total partition function is therefore:

$$Z_N = Z_1^N = \left(\frac{2\pi k_B T}{\omega h} \right)^N.$$

We can now obtain the main thermodynamic variables:

$$\begin{aligned}
F &= -\frac{1}{\beta} \log Z_n = -Nk_B T \log \left(\frac{2\pi k_B T}{\omega h} \right) \\
E &= -\frac{\partial \log Z_N}{partial \beta} = Nk_B T, \\
S &= \frac{E - F}{T} = Nk_B T + Nk_B T \log \left(\frac{2\pi k_B T}{\omega h} \right) \\
P &= -\frac{\partial F}{\partial V}\bigg|_{T,N} = 0.
\end{aligned}$$

Let's observe that the pressure of this system is always zero: this can be interpreted as a consequence of the infinite volume of this system (since we have integrated over the whole configuration space)-

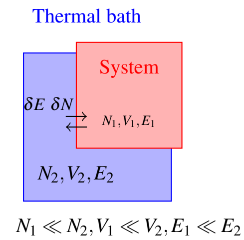
6.3 Grand canonical ensemble

In this section we will discuss the framework used to study open systems. In this way we will have to account for a non-fixed number of particles. For this reason, the distribution describing our system is defined over each phase space of n particles, going from 1 to ∞.

We will derive the grand canonical partition function by studying our system (denoted by (1)) embedded (at equilibrium) in a thermal bath (denoted by (2)) such that the whole universe (system plus bath) is a closed system, therefore described by the canonical partition function.

$$\rho_c(q_i^{(1)}, p_i^{(1)}, q_j^{(2)}, p_j^{(2)}) = \frac{e^{-\beta(\mathcal{H}_1+\mathcal{H}_2)}}{Z_N}.$$

We can fix the number of particles in each system in order to integrate over the phase space of the bath and thus get the marginal distribution of our system (with N_1 particles). Doing this integration we will assume the normalization constant of the distribution of the system to be $\frac{N!}{N_1!N_2!}$, and we will later check that our assumption was the right one.



A simple sketch of the system and the thermal bath.

$$\begin{aligned}
\rho_{gc} &= \frac{N!}{N_1!N_2!} \int_{\mathcal{M}^{N_2}} \Pi_j dq_j^{(2)}\, dp_j^{(2)} \rho_c \\
&= \frac{N!}{N_1!N_2!} \int_{\mathcal{M}^{N_2}} \Pi_j dq_j^{(2)}\, dp_j^{(2)} \frac{e^{-\beta\mathcal{H}_1} e^{-\beta\mathcal{H}_2}}{\int d\Gamma\ \exp(-\beta\mathcal{H}_1 + \mathcal{H}_2)} \frac{h^{dN}}{h^{dN}} \\
&= \frac{Z_{N_2}}{Z_N} \frac{e^{-\beta\mathcal{H}_1}}{N_1! h^{dN_1}}.
\end{aligned}$$

This is the grand canonical distribution for our system. We will now check our assumption by integrating it over the phase space of the system with N_1 particles and then summing up all the contribution of all the integrals with different numbers of particles (up to the total number of particles of the universe). We expect that the result of this operation should be equals to 1.

$$\sum_{N_1=0}^{N} \int_{\mathcal{M}^{N_1}} \rho_{gc} d\Omega_{N_1} = \sum_{N_1=0}^{N} \frac{Z_{N_2}}{Z_N} \int_{\mathcal{M}^{N_1}} \frac{e^{-\beta\mathcal{H}_1}}{N_1! h^{dN_1}} d\Omega_{N_1} = \sum_{N_1=0}^{N} \frac{Z_{N_2} Z_{N_1}}{Z_N}.$$

This expression doesn't add up to one, since each partition function is integrated over different phase spaces: this problem is solved by considering the bath infinitely large (in the thermodynamic limit). We firstly define

$$\phi(q) = \int_{\mathbb{R}^{dN_1}} \Pi_i dp_i^{(1)} \int_{\mathbb{R}^{dN_2}} \Pi_j dp_j^{(2)} e^{-\beta(\mathcal{H}_1+\mathcal{H}_2)},$$

in this way the above sum reads:

$$\begin{aligned}
\sum_{N_1=0}^{N} \int_{\mathcal{M}^{N_1}} \rho_{gc} d\Omega_{N_1} &= \sum_{N_1=0}^{N} \frac{N!}{N_1!N_2!} \frac{\int_{\mathcal{M}^{N_1}} d\Gamma_1 e^{-\beta\mathcal{H}_1} \int_{\mathcal{M}^{N}_{N_2}} d\Gamma_2 e^{-\beta\mathcal{H}_2}}{\int d\Gamma\ \exp(-\beta\mathcal{H}_1 + \mathcal{H}_2)} \\
&= \sum_{N_1=0}^{N} \frac{N!}{N_1!N_2!} \frac{\int_{V_1} \Pi_i dq_i^{(1)} \int_{V_2} \Pi_j dq_j^{(2)} \phi(q)}{\int_{V_1+V_2} \Pi_i dq_i^{(1)} \int_{V_1+V_2} \Pi_j dq_j^{(2)} \phi(q)}.
\end{aligned}$$

We define $\langle\phi\rangle_V$ as the mean value of ϕ over the volume of integration, normalized by dividing by such volume, in this way the sum reads:

$$\sum_{N_1=0}^{N} \int_{\mathcal{M}^{N_1}} \rho_{gc} d\Omega_{N_1} = \sum_{N_1=0}^{N} \frac{V_1^{N_1} V_2^{N_2}}{(V_1+V_2)^{N_1+N_2}} \frac{\langle\phi\rangle_{V_1V_2}}{\langle\phi\rangle_V} = \sum_{N_1=0}^{N} \left(\frac{V_1}{V}\right)^{N_1} \left(\frac{V_2}{V}\right)^{N_2} \frac{\langle\phi\rangle_{V_1V_2}}{\langle\phi\rangle_V}.$$

In the thermodynamic limit the volumes go to infinity, as for the number of particles, thus the tho averages tend to the same limit, we then get:

$$\lim_{V\to\infty}\sum_{N_1=0}^{N}\int_{\mathcal{M}^{N_1}}\rho_{gc}d\Omega_{N_1}\approx\lim_{N\to\infty}\sum_{N_1=0}^{N}\left(\frac{V_1}{V}\right)^{N_1}\left(\frac{V_2}{V}\right)^{N_2}=\lim_{N\to\infty}\left(\frac{V_1+V_2}{V}\right)^{N}=1$$

Therefore, in the thermodynamic limit our initial assumption holds.

Lastly we need to remove all the explicit dependencies by the variables of the thermal bath, using the definition of the Gibb's free energy, given for the canonical ensemble, we can get:

$$\frac{Z_{N_2}}{Z_N}=\exp\left\{-\beta F(T,N-N_1,V-V_1)+\beta F(T,N,V)\right\}\approx\exp\left\{\beta\frac{\partial F}{\partial N}N_1+\beta\frac{\partial F}{\partial V}V_1\right\},$$

in which we have used Taylor expansion to the first order, considering the bath larger than our system. If we define $\mu=\frac{\partial F}{\partial N}$, the chemical potential, and reminding that from classical thermodynamic $P=-\frac{\partial F}{\partial V}$, the grand canonical distribution reads:

$$\rho_{gc}=\frac{e^{-\beta\mathcal{H}_1}}{N_1!h^{dN_1}}e^{\beta\mu N_1}e^{-\beta PV_1}.$$

We can now remove the label of the system and define the **fugacity**

$$z=e^{\beta N}$$

and the **grand potential**

$$\Omega=-PV=E-TS-\mu N,$$

using the normalization condition we get:

$$\begin{aligned}1&=\sum_{N=0}^{\infty}\int_{\mathcal{M}^N}d\Omega_N e^{-\beta\mathcal{H}}e^{\beta\mu N}e^{-\beta PV}=e^{-\beta PV}\sum_{N=0}^{\infty}e^{\beta\mu N}\int_{\mathcal{M}^N}d\Omega_N e^{-\beta\mathcal{H}}\\&=e^{\Omega}\sum_{N=0}^{\infty}z^N Z_N.\end{aligned}$$

Defining the **grand canonical partition function**

$$\mathcal{Z}=\sum_{N=0}^{\infty}z^N Z_N\quad\Rightarrow\quad\Omega=-\frac{\log\mathcal{Z}}{\beta},\tag{6.5}$$

the grand canonical distribution reads:

$$\rho_{gc}=\frac{e^{-\beta\mathcal{H}}e^{\beta\mu N}}{\mathcal{Z}}.\tag{6.6}$$

Having the full distribution allows us to calculate means values of the mechanical variables of the system. These can be expressed in terms of the canonical means:

$$\begin{aligned}<f>_{gc}&=\sum_{N=0}^{\infty}\int_{\mathcal{M}^N}d\Omega_N\frac{e^{-\beta\mathcal{H}}e^{\beta\mu N}}{\mathcal{Z}}f(q_i,p_i)=\frac{1}{\mathcal{Z}}\sum_{N=0}^{\infty}Z_N e^{\beta\mu N}\int_{\mathcal{M}^N}d\Omega_N\frac{e^{-\beta\mathcal{H}}}{Z_N}f(q_i,p_i)\\&=\frac{1}{\mathcal{Z}}\sum_{N=0}^{\infty}Z_N z^N<f>_c.\end{aligned}$$

We now calculate all the main thermodynamic variables:

$$
\begin{aligned}
E =< \mathcal{H} >&= \sum_{N=0}^{\infty} \int_{\mathcal{M}^N} d\Omega_N \frac{e^{-\beta\mathcal{H}} e^{\beta\mu N}}{\mathcal{Z}} \mathcal{H} = \frac{-1}{\mathcal{Z}} \sum_{N=0}^{\infty} z^N \int_{\mathcal{M}^N} d\Omega_N \frac{\partial}{\partial\beta} e^{-\beta\mathcal{H}} \\
&= \frac{-1}{\mathcal{Z}} \frac{\partial}{\partial\beta} \sum_{N=0}^{\infty} z^N \int_{\mathcal{M}^N} d\Omega_N e^{-\beta\mathcal{H}} \bigg|_{z \text{ fixed}} = -\frac{1}{\mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial\beta} \bigg|_{z \text{ fixed}}, \\
< N > &= \sum_{N=0}^{\infty} \int_{\mathcal{M}^N} d\Omega_N \frac{e^{-\beta\mathcal{H}} e^{\beta\mu N}}{\mathcal{Z}} N = \sum_{N=0}^{\infty} e^{\beta\mu N} N \frac{Z_N}{\mathcal{Z}} \int_{\mathcal{M}^N} d\Omega_N \frac{e^{-\beta\mathcal{H}}}{Z_N} \\
&= \sum_{N=0}^{\infty} e^{\beta\mu N} N \frac{Z_N}{\mathcal{Z}} = \frac{1}{\mathcal{Z}} \sum_{N=0}^{\infty} z^N N Z_N = \frac{z}{\mathcal{Z}} \frac{\partial}{\partial z} \sum_{N=0}^{\infty} z^N Z_N \bigg|_{\beta \text{ fixed}} \\
&= \frac{z}{\mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial z} \bigg|_{\beta \text{ fixed}}.
\end{aligned}
$$

Lastly we recover entropy using the Boltzmann's universal formula:

$$
\begin{aligned}
S &= -k_B < \log \rho_{gc} >= -k_B \sum_{N=0}^{\infty} \int_{\mathcal{M}^N} d\Omega_N \frac{e^{-\beta\mathcal{H}} e^{\beta\mu N}}{\mathcal{Z}} [-\beta\mathcal{H} + \beta\mu N - \log \mathcal{Z}] \\
&= k_B \beta E - k_B \beta\mu < N > + k_B \log \mathcal{Z} \frac{\beta}{\beta} = \frac{E - \mu N - \Omega}{T} = S_{\text{Th}},
\end{aligned}
$$

which shows the constituency of all the definition of the thermodynamic variables.

6.3.1 Van deer Waals' law

We are going to derive the **Van deer Waals' law**, for real gasses, using the grand canonical framework. In order to do so we will need to approximate the grand canonical partition function using a so-called **virial expansion**.

Considering a small fugacity we will truncate the summation defining the partition function:

$$
\mathcal{Z} = 1 + zZ_1 + z^2 Z_2 + z^3 Z_3 + \cdots \approx 1 + zZ_1.
$$

Considering a gas of free particles, as we already derived, the partition function of the single particle is:

$$
Z_1 = V \left(\frac{2m\pi k_B T}{h^2} \right)^{\frac{3}{2}} = \frac{V}{\lambda_T^3} \quad \Rightarrow \quad \mathcal{Z} \approx 1 + z \frac{V}{\lambda_T^3}.
$$

Since we are considering an approximation of the first order in z, we can evaluate the logarithm of $\mathcal{Z}$ truncating at the same order its Taylor's approximation

$$
\log \mathcal{Z} \approx \log \left(1 + z \frac{V}{\lambda_T^3} \right) \approx z \frac{V}{\lambda_T^3},
$$

using the definitions given in the last section:

$$
PV = \frac{\log \mathcal{Z}}{\beta} = z \frac{V}{\lambda_T^3 \beta}, \quad N = z \frac{\partial \log \mathcal{Z}}{\partial z} \bigg|_{\beta} = z \frac{V}{\lambda_T^3} \quad \Rightarrow \quad PV = k_B T N.
$$

Therefore, the first order virial approximation results in the state equation of the ideal gas. We can notice that this also shows that a low fugacity (the limit that we are using for the approximation) corresponds to a very dilute gas.

Let's truncate such approximation at the second order, in order to do so we need to evaluate the partition function of two particles (that can interact):

$$\begin{aligned}
Z_2 &= \frac{1}{2!}\int_V\int_{\mathbb{R}^3}\frac{d^3q_1d^3p_1}{h^3}\int_V\int_{\mathbb{R}^3}\frac{d^3q_2d^3p_2}{h^3}\exp\left\{-\beta\frac{\vec{p}_1^2+\vec{p}_2^2}{2m}-\beta U(|\vec{q}_1-\vec{q}_2|)\right\}\\
&= \frac{1}{2\lambda_T^6}\int_V\int_V d^3q_1d^3q_2e^{-\beta U(|\vec{q}_1-\vec{q}_2|)}\\
&\text{Changing variables } \vec{R}=\frac{\vec{q}_1+\vec{q}_2}{2},\quad \vec{r}=\vec{q}_1-\vec{q}_2,\\
&= \frac{1}{2\lambda_T^6}\int_V\int_V d^3Rd^3r\, e^{-\beta U(|\vec{r}|)} = \frac{V}{2\lambda_T^6}\int_V d^3r\, e^{-\beta U(|\vec{r}|)} = \frac{V}{2\lambda_T^6}J_2(\beta).
\end{aligned}$$

In our approximation the grand canonical partition function reads

$$\mathcal{Z}\approx 1+z\frac{V}{\lambda_T^3}+z^2\frac{V}{2\lambda_T^6}J_2(\beta)$$

and thus

III

Relativistic Quantum Mechanics

7. First quantization of special relativity

7.1 From classical to relativistic quantum mechanics

Classical mechanics is based on the quantization of classical observable by promoting them to operators, with their own commutation relations. De Broglie suggested that every quantum object can be described in terms of waves proprieties related to the energy and the momentum of the object:

$$E = h\nu, \ \vec{p} = h\vec{k}, \qquad \Rightarrow \qquad \psi(\vec{x},t) = \exp\left(\frac{i}{\hbar}P^{\mu}x_{\mu}\right), \quad P^{\mu} = \left(\frac{E}{c},\vec{p}\right).$$

Schrödinger was the first to propose this quantization procedure in order to find the equation of motion of these objects: in a simpler way than what he did, we can observe that the differentiation operators applied to ψ act as operators with eigenvalues related to the components of P^{μ}.

$$i\hbar\frac{\partial}{\partial t}\psi = E\psi = \frac{\vec{p}^2}{2m}\psi = (-i\hbar\vec{\nabla})(-i\hbar\vec{\nabla})\frac{\psi}{2m} = -\frac{\hbar^2}{2m}\vec{\nabla}^2\psi.$$

In this way we've got the **Schrödinger equation**. The same approach can be used in order to derive the relativistic equations of motion:

$$E = \sqrt{p^2c^2 + m^2c^4} \quad \rightarrow \quad i\hbar\frac{\partial}{\partial t}\psi = \sqrt{-\hbar^2c^2\vec{\nabla}^2 + m^2c^4}\,\psi.$$

The differential equation is hard to solve since it is not so clear what should be the square root of a derivative, to solve this problem we could Taylor expand the square root, but this leads to a non-local theory.
Klein and Gordon tried a different approach quantizing $E^2 = p^2c^2 + m^2c^4$:

$$-\hbar^2\frac{\partial^2}{\partial t^2}\psi = \left(-\hbar^2c^2\vec{\nabla}^2 + m^2c^4\right)\psi, \tag{7.1}$$

this is the **Klein-Gordon equation**.
In equation (7.1) it appears a constant term given by $\frac{mc}{\hbar^2}$, this is called the **reduced Compton wavelength**.
Using natural units ($\hbar = c = 1$) and defining the box operator $\Box = \partial^\mu \partial_\mu$, equation (7.1) reads:

$$(\Box - m^2)\psi = 0. \tag{7.2}$$

We will now study the proprieties of the solutions of equation (7.2).
Using the ansatz $\psi(x^\mu) \propto e^{iA_\mu x^\mu}$ we get:

$$(-\partial^\nu \partial_\nu + m^2)e^{iA_\mu x^\mu} = (A^\nu A_\nu + m^2)e^{iA_\mu x^\mu} = 0.$$

Since the exponential is a positive function, in order to be a solution, A^μ must satisfy the energy-momentum relation of a particle with mass m, and thus it should be $A^\mu = P^\mu$:

$$P^0 = \pm\sqrt{\vec{p}^2 + m^2} \triangleq \pm E_p, \tag{7.3}$$

which is the so called **mass-shell condition**.
This relation gives rise to a strange behavior of the relativistic free particle: since P^0 is the energy of the particle, this equation shows that wave functions with **negative energies** are allowed by the theory. Defining

$$\begin{cases} \psi^+_{\vec{p}}(x^\mu) = e^{-iE_p t} e^{i\vec{p}\cdot\vec{x}} \\ \psi^-_{\vec{p}}(x^\mu) = e^{iE_p t} e^{-i\vec{p}\cdot\vec{x}} \end{cases}$$

we can get the general solution (7.2) by superposition of these:

$$\psi(x^\mu) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} \left(a(\vec{p}) e^{-iE_p t} e^{i\vec{p}\cdot\vec{x}} + b^*(\vec{p}) e^{iE_p t} e^{-i\vec{p}\cdot\vec{x}} \right). \tag{7.4}$$

We should observe that, since $\psi^+_{\vec{p}}(x^\mu) = (\psi^-_{\vec{p}}(x^\mu))^*$, the complex conjugate of a wave function given by (7.4) is:

$$\psi(x^\mu)^* = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} \left(b(\vec{p}) e^{-iE_p t} e^{i\vec{p}\cdot\vec{x}} + a^*(\vec{p}) e^{iE_p t} e^{-i\vec{p}\cdot\vec{x}} \right).$$

We will see that the negative energy components should be interpreted as an antiparticle or the same particle going backwards in time.

Lastly, we will see that the probability interpretation of quantum mechanics can fail due to this negative energy states. Multiplying the Klein-Gordon equation (7.2) by $\psi*$ and combining it with its complex conjugate we get the probability conservation law:

$$\begin{aligned} \psi^*(\Box - m^2)\psi - \psi(\Box - m^2)\psi^* &= \psi^*\Box\psi - \psi\Box\psi^* \\ &= \partial_\mu(\psi^*\partial^\mu\psi - \psi\partial^\mu\psi^*) = 2im\partial_\mu J^\mu = 0, \end{aligned}$$

$$J^\mu \triangleq \frac{1}{2im}(\psi^*\partial^\mu\psi - \psi\partial^\mu\psi^*).$$

The component J^0 is therefore the probability density, which can became negative for negative energies. The field should be interpreted as an object whose quanta are the particles we observe.
This interpretation leads directly to quantum field theory (Part IV).

7.2 The Yukawa potential

Since we are going to interpret the wave function resulting form the Klein-Gordon equation (7.2) as we interpret the electromagnetic waves that arise from Maxwell's equations, we could try to include some source terms in this theory too.
The first attempt we can try is a point source (like a point charge) stationary:

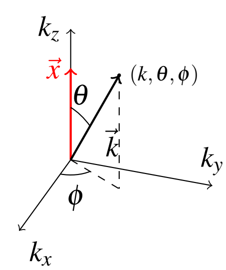
$$(\Box - m^2)\psi(x^\mu) = g\delta^3(x^\mu) \qquad \xrightarrow[case]{Stationary} \qquad \left(\vec{\nabla}^2 - m^2\right)\psi(\vec{x}) = g\delta^3(\vec{x}).$$

In order to solve this equation let's consider the Fourier transform of a generic solution and the Fourier representation of the Dirac delta function:

$$\psi(\vec{x}) = \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}}\tilde{\psi}(\vec{p}), \qquad \delta^3(\vec{x}) = \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}}.$$

Plugging these two in the differential equation, and using the linearity of the nabla operator, we get:

$$\begin{aligned}
&\left(\vec{\nabla}^2 - m^2\right)\int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}}\tilde{\psi}(\vec{k}) = g\int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}},\\
\Rightarrow \qquad &\int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}}(-k^2 - m^2)\tilde{\psi}(\vec{k}) = g\int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}};\\
\Rightarrow \qquad &\tilde{\psi}(\vec{k}) = \frac{-g}{k^2 + m^2}.
\end{aligned}$$



Spherical coordinates used for the integration over the k-space.

Using the spherical coordinate in the k-space we can now calculate the explicit form of the solution ψ:

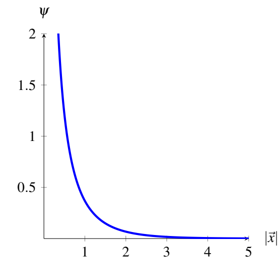
$$\begin{aligned}
\psi(\vec{x}) &= \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}} \frac{-g}{k^2 + m^2}\\
&= \frac{-g}{(2\pi)^3}\int_0^\infty k^2 dk \int_0^\pi d(-\cos\theta)\int_0^{2\pi} d\phi \frac{e^{ikx\cos\theta}}{k^+ m^2}\\
&= \frac{-g}{(2\pi)^2}\int_0^\infty k^2 dk \int_0^\pi d(-\cos\theta)\frac{e^{ikx\cos\theta}}{k^+ m^2}\\
&= \frac{-g}{(2\pi)^2}\int_0^\infty \frac{k^2}{k^2 + m^2}\frac{e^{ikx\cos\theta}}{-ikx}\bigg|_{-1}^{+1} dk = \frac{-2g}{(2\pi)^2 x}\int_0^\infty \frac{k\sin(kx)}{k^2 + m^2}\\
&= \frac{-g}{(2\pi)^2 x}\int_{-\infty}^{+\infty} \frac{k\sin(kx)}{k^2 + m^2} dk.
\end{aligned}$$

In order to solve this last integral we can use integration over the complex plane:

$$\begin{aligned}
\frac{-g}{(2\pi)^2 x}\int_{-\infty}^{+\infty} \frac{k\sin(kx)}{k^2 + m^2} dk &= \mathfrak{Im}\left\{\frac{-g}{(2\pi)^2 x}\int_{-\infty}^{+\infty} \frac{ke^{ikx}}{k^2 + m^2} dk\right\}\\
&= \mathfrak{Im}\left\{\frac{-g}{(2\pi)^2 x} 2\pi i \operatorname*{Res}_{z=im}\left(\frac{ze^{izx}}{z^2 + m^2}\right)\right\}\\
&= \mathfrak{Im}\left\{\frac{-g}{(2\pi)^2 x} 2\pi i \frac{ime^{-mx}}{2im}\right\} = -g\frac{e^{-mx}}{4\pi x}.
\end{aligned}$$

The solution we have found is the so-called **Yukawa potential**, it describes the strong interactions inside the nuclei of atoms (using $m = m_\pi$) but it is something more general. Notice that for fields of massless particles this becomes similar to the Coulomb potential.

$$\psi(\vec{x}) = -g\frac{e^{-m|\vec{x}|}}{4\pi|\vec{x}|}. \tag{7.5}$$



Plot of the Yukawa potential.

7.2.1 Green's functions for the generalized Klein-Gordon equation

We want now to generalize Yukawa's approach to any kind of sources:

$$(\Box - m^2)\psi(x^\mu) = J(x^\mu). \tag{7.6}$$

In order to give a general solution of this equation we can introduce the Green's function $G(x^\mu, y^\mu)$, such that:

$$\psi(x^\mu) = \psi_0(x^\mu) + \int d^4y\, G(x^\mu - y^\mu)J(y^\mu), \qquad (\Box - m^2)G(x^\mu) = \delta^4(x^\mu),$$

where ψ_0 is a solution of the homogeneous equation.
In this way ψ is the general solution we seek for:

$$\begin{aligned}(\Box - m^2)\psi(x^\mu) &= (\Box - m^2)\psi_0(x^\mu) + (\Box - m^2)\int d^4y\, G(x^\mu - y^\mu)J(y^\mu)\\ &= \int d^4y\, (\Box - m^2)G(x^\mu - y^\mu)J(y^\mu) = \int d^4y\, \delta^4(x^\mu - y^\mu)J(y^\mu) = J(x^\mu).\end{aligned}$$

Lastly we need to find the explicit form of G by solving the differential equation above: introducing its Fourier transform and the Dirac's delta representation:

$$\begin{aligned}&G(x^\mu) = \int \frac{d^3p}{(2\pi)^4} e^{iP_\mu x^\mu}\tilde{G}(P^\mu), \qquad \delta^4(x^\mu) = \int \frac{d^4k}{(2\pi)^4} e^{iP_\mu x^\mu},\\ \Rightarrow\quad &(\Box - m^2)\int \frac{d^3p}{(2\pi)^4} e^{iP_\mu x^\mu}\tilde{G}(P^\mu) = \int \frac{d^4k}{(2\pi)^4} e^{iP_\mu x^\mu}\\ \Rightarrow\quad &\int \frac{d^3p}{(2\pi)^4} e^{iP_\mu x^\mu}(-P^\mu P_\mu - m^2)\tilde{G}(P^\mu) = \int \frac{d^4k}{(2\pi)^4} e^{iP_\mu x^\mu}\\ \Rightarrow\quad &\tilde{G}(P^\mu) = \frac{-1}{P^\mu P_\mu + m^2}.\end{aligned}$$

We now introduce the **propagator**:

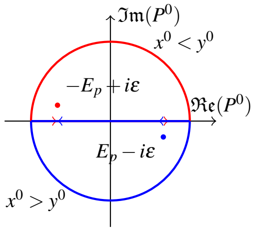
$$\Delta(x^\mu - y^\mu) = -iG(x^\mu - y^\mu) = -i\int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}} \int \frac{dP^0}{2\pi} \frac{e^{-iP^0(x^0 - y^0)}}{P^\mu P_\mu + m^2}.$$

To solve this integral we have to impose the **Stuckelberg-Feynman prescription**:

$$P^\mu P_\mu + m^2 = -(P^0 + E_p)(P^0 - E_p) \qquad \xleftarrow[\varepsilon\to 0]{for} \qquad -(P^0 + E_p + i\varepsilon)(P^0 - E_p + i\varepsilon),$$

where $E_p = \sqrt{\vec{p}^2 + m^2}$.
This prescription, as we will see, imposes that particles with positive energies propagate

forward in time and those with negative energies propagates backwards.
Using this prescription we can integrate in the complex plane with the poles translated by a bit up and down the real axis:



Path used in the integration in the complex plane.

$$\Delta(x^\mu - y^\mu) = \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}} \left[\frac{e^{-iE_p\left(x^0-y^0\right)}}{2E_p} \theta\left(x^0 - y^0\right) + \frac{e^{iE_p\left(x^0-y^0\right)}}{2E_p} \theta\left(y^0 - x^0\right) \right],$$

where $\theta(x)$ is the Heaviside's step function, that is given by the fact that when $x^0 > y^0$ Jordan's theorem wants the integration path to go in the lower part of the complex plane, thus we just have the contribution of the pole $z = E_p - i\varepsilon$, and when $x^0 < y^0$ we will have only the contribution of the other pole $z = -E_p + i\varepsilon$.

Lastly we should concern about the interpretation of this result: since y^0 is the time at which the "point source" of the Green's function is located, the term resulting from $x^0 > y^0$ should be interpreted as a particle going forward in time (from y^0 to x^0) and, as already anticipated, we can recognize that this has positive energy; the term resulting from $x^0 < y^0$ should be interpreted as a particle going backwards in time and this has clearly a negative energy.

8. Dirac's equation

8.1 Deriving the Dirac's equation

As we saw in the previous chapter, the Klein-Gordon equation (7.1) is unable to describe a relativistic theory of quantum mechanics that, as for the classical formulation from Schrödinger, can result in a probabilistic interpretation of the wave function. Furthermore, we will see that the new theory that we are going to describe, predicts precisely all of those aspects that are shown by experiments but that in classical quantum mechanics (as well for the Klein-Gordon theory) needs to be "artificially" added.

Dirac understood that absence of a definite positive probability density, in the Klein-Gordon equations, was caused by the non-linearity of the differential equation, with respect to time. Since quantizing the energy momentum relation of the form $E = \sqrt{c\vec{p}^2 + m^2c^4}$ had its own problems, as we already discussed in Section 7.1, Dirac tried a different approach: he assumed that was possible to obtain a linear relation between momentum and energy using a set of four hermitian[1] matrix:

$$E = c\vec{\alpha} \cdot \vec{p} + \beta mc^2, \qquad \alpha_1, \alpha_2, \alpha_3, \beta \in M_{4\times 4}(\mathbb{C}). \tag{8.1}$$

Given this assumption, we need to have it satisfy the energy momentum relation:

$$\begin{aligned} E^2 &= (c\alpha^i p^i + \beta mc^2)(c\alpha^j p^j + \beta mc^2) \\ &= cp^i p^j \alpha^i \alpha^j + \beta^2 m^2 c^4 + mc^3 p^i \left(\alpha^i \beta + \beta \alpha^i\right) \\ &= c^2 \vec{p}^2 + m^2 c^4. \end{aligned}$$

In order to have this last equality satisfied we need:

[1]These must be hermitian since the whole expression has to be real: in this way E is hermitian too and thus its components, since it is a multiple of the identity matrix 4×4, are real.

- $\beta^2 = 1 \quad \Rightarrow \quad \{\beta,\beta\} = 2$;
- $\alpha^i\beta + \beta\alpha^i = 0 \quad \Rightarrow \quad \{\beta,\alpha^i\} = 0,\ i = \{1,2,3\}$;
- since $p^i p^j$ is symmetric, the symmetric part[2] of $\alpha^i\alpha^j$ should result in a Kronecker's delta

$$\frac{\alpha^i\alpha^j + \alpha^j\alpha^i}{2} = \delta^{ij} \quad \Rightarrow \quad \{\alpha^i\alpha^j\} = 2\delta^{ij},\ i,j = \{1,2,3\}.$$

The notation $\{A,B\}$ is the anticommutator: $AB+BA$.
This structure defines a *Clifford algebra* and can be used to derive the explicit form of these matrices (we use a shorthand notation for 4×4 matrices):

$$\beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad \alpha^i = \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix}, \tag{8.2}$$

where σ^i are the 2×2 Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8.3}$$

We now need to quantize equation (8.1). In order to do so we promote observable to operators acting on wave functions (as in Section7.1):

$$E \rightsquigarrow i\hbar\partial_t, \qquad \vec{p} \rightsquigarrow -i\hbar\vec{\nabla}.$$

Before plugging this operator inside the equation (8.1) we have to stress the fact that those operators are multiplied to some 4×4 matrices (Eq. (8.2)) and thus we cannot expect to have a scalar wave function. It is natural to use a vector of 4 wave functions, called **spinor**. The resulting equation reads:

$$i\hbar\partial_t\psi = (-ic\hbar\alpha^i\partial_i + \beta mc^2)\psi.$$

We now want to express this equation in a covariant form: to do so we define:

$$\gamma^0 = -i\beta, \qquad \gamma^j = -i\beta\alpha^j, \qquad j = \{1,2,3\}. \tag{8.4}$$

In this way, multiplying ,at the left, both sides of the above equation by $\frac{\beta}{c\hbar}$, remembering that $\beta^2\mathbb{1}$ and plugging the matrices 8.4, we get:

$$\begin{aligned}
&\frac{i\beta}{c}\partial_t\psi = \left(-i\beta\alpha^i\partial_i + \frac{mc}{\hbar}\right)\psi \\
&\Rightarrow \left(\gamma^0\partial_0 + \gamma^i\partial_i\frac{mc}{\hbar}\right)\psi = 0 \\
&\Rightarrow \left(\gamma^\mu\partial_\mu + \frac{mc}{\hbar}\right)\psi = 0.
\end{aligned}$$

Lastly we use the natural units with the Feynman's slash convention $\not{\partial} = \gamma^\mu\partial_\mu$, it reads:

$$(\not{\partial} + m)\psi = 0. \tag{8.5}$$

[2]Any symmetric tensor contracted with an antisymmetric one always vanishes.

We remind that $\psi(x^\mu)$ is a vector of four wave functions:

$$\psi(x^\mu) = \begin{pmatrix} \psi_1(x^\mu) \\ \psi_2(x^\mu) \\ \psi_3(x^\mu) \\ \psi_4(x^\mu) \end{pmatrix}, \qquad \psi_i : \mathbb{R}^4 \to \mathbb{C}.$$

Let's now check if this equation has some kind of probability current and density. First we can multiply, to the left, the equation (8.5) by $\psi^\dagger$, then we subtract the conjugate transpose of the same equation, multiplied by ψ to the left:

$$\begin{aligned} 0 &= \psi^\dagger \left[\gamma^\mu \partial_\mu + m\right]\psi - \left[\left(\gamma^\nu \partial_\nu + m\right)\psi\right]^\dagger \psi \\ &= \psi^\dagger \left[\gamma^\mu \partial_\mu + m\right]\psi - \left[\partial_\nu \psi^\dagger \gamma^{\dagger\nu} + \psi^\dagger m\right]\psi \\ &= \psi^\dagger \gamma^\mu \partial_\mu \psi + \partial_\mu \psi^\dagger \gamma^\mu \psi = \partial_\mu \left(\psi^\dagger \gamma^\mu \psi\right). \end{aligned}$$

We have obtained a conserved quantity in a form of *continuity equation*, as it happens with the Schrödinger equation, thus we can conclude that Dirac's approach restore the probability interpretation of quantum mechanics. It is more convenient to multiply the continuity equation by $i\beta$, in this way $\left(i\beta\gamma^0 = \beta^2 = \mathbb{1}\right)$ the probability density reads $\psi^\dagger\psi$ as in classical quantum mechanics.

8.1.1 Some proprieties of the γ matrices

We will now discuss the main proprieties of the matrix formalism 8.4 introduced by Dirac. First of all we should notice that, since β and α^i form a Clifford algebra, the corresponding $\gamma^0 = -i\beta,\ \gamma^j = -i\beta\alpha^j$ form the same structure:

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}.$$

Furthermore, from the hermiticity of β and α^i and the anticommutation relation $\beta\alpha^i = -\alpha^i\beta$, it is easy to see that γ^0 is antihermitian and γ^i are hermitian:

$$\left(\gamma^0\right)^\dagger = \left(-i\beta\right)^\dagger = i\beta = -\gamma^0, \quad \left(\gamma^j\right)^\dagger = \left(-i\beta\alpha^j\right)^\dagger = i\alpha^j\beta = -i\beta\alpha^j = \gamma^j.$$

These relations can be summarized in the compact form

$$\left(\gamma^\mu\right)^\dagger = \gamma^0\gamma^\mu\gamma^0 \quad \text{or} \quad \left(\gamma^\mu\right)^\dagger = -\beta\gamma^\mu\beta.$$

The second proprieties of these matrices is that are traceless, for example, using that $\left(\gamma^2\right)^2 = \mathbb{1}$, the anticommutation and the cyclic propriety of the trace:

$$Tr\,\gamma^1 = Tr\,\gamma^1\left(\gamma^2\right)^2 = -Tr\,\gamma^2\gamma^1\gamma^2 = -Tr\,\gamma^1\left(\gamma^2\right) = -Tr\,\gamma^1 \quad \Rightarrow \quad Tr\,\gamma^1 = 0.$$

Lastly we should introduce the **chirality matrix** γ^5, defined by

$$\gamma^5 = -i\gamma^0\gamma^1\gamma^2\gamma^3, \tag{8.6}$$

which has the following proprieties:

$$\{\gamma^5, \gamma^\mu\} = 0, \quad \left(\gamma^5\right)^2 = \mathbb{1}, \quad \left(\gamma^5\right)^\dagger = \gamma^5, \quad Tr\,\gamma^5 = 0. \tag{8.7}$$

This matrix is needed to define the **chiral projectors**

$$P_L = \frac{\mathbb{1}-\gamma^5}{2}, \quad P_R = \frac{\mathbb{1}+\gamma^5}{2} \tag{8.8}$$

that allows the Dirac spinor to be divided into its right-handed and left-handed components, called **Weyl spinors**.

8.2 Plane wave solution of Dirac's equation

We will now study the problem of finding the free particle solutions of the Dirac's equation. In order to do so we will suppose that a solution will be in the form:

$$\psi(x^\mu) = \omega(P^\mu)e^{iP_\mu x^\mu},$$

where P^μ is an arbitrary 4-vector and ω is the **polarization spinor**.
Plugging this ansatz in the Eq. (8.5) we get the following condition:

$$0 = \big(\gamma^\mu \partial_\mu + m\big)\omega(P^\mu)e^{iP_\nu x^\nu} = \big(i\gamma^\mu P_\mu + m\big)\omega(P^\mu)e^{iP_\nu x^\nu}.$$

By multiplication by $\big(-i\gamma^\nu P_\nu + m\big)$ and observing that

$$P^\mu P^\nu \gamma^\mu \gamma^\nu = P^\mu P^\nu \frac{\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu}{2} = P^\mu P^\nu \eta^{\mu\nu} = P^\mu P_\mu,$$

we than get the condition needed to have our ansatz to satisfy the Dirac's equation:

$$\big(P^\mu P_\mu + m^2\big)\omega(P^\mu) = 0,$$

where we can recognize the mass-shell condition and thus that P^μ has to be the 4-momentum.
Let's now consider a rest particle, therefore its 4-momentum will be $P^\mu = (E,0,0,0)$. In this case the first condition we obtained reads:

$$0 = (i\gamma^0 P_0 + m)\omega = (-iE\gamma^0 + m)\omega = (E\beta + m)\omega.$$

In matrix form:

$$\begin{pmatrix} E & 0 & 0 & 0 \\ 0 & E & 0 & 0 \\ 0 & 0 & E & 0 \\ 0 & 0 & 0 & E \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \omega_4 \end{pmatrix} = \begin{pmatrix} m & 0 & 0 & 0 \\ 0 & m & 0 & 0 \\ 0 & 0 & -m & 0 \\ 0 & 0 & 0 & -m \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \omega_4 \end{pmatrix},$$

it is clear that even Dirac's equation allows some negative energy states ($E = -m$) which are characterized by different polarization spinor with respect to those with positive energy

$$\omega = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \to \text{Positive energy}, \qquad \omega = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \to \text{Negative energy}.$$

Solutions for a particle at rest will then read:

$$\psi(x^\mu) = \omega(P^\mu)e^{-imt},$$

we will obtain solutions describing moving particles performing boots on these solutions.

8.3 Non-relativistic limit of the Dirac's equation

In order to test the validity of this theory we can study the limit for which this should reduce to the classical one and verify if those two are consistent. Furthermore, this approach will give us easier equation that can be used as correction to the classical theory.

First of all we have to take into accounts that classical quantum mechanics make use of 2-d spinors, and for this reason we divide the Dirac's spinors in two components (we have again introduced non-natural units because we will need to study the limit $c \to \infty$):

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_3 \\ \psi_3 \\ \psi_4 \end{pmatrix} e^{-\frac{i}{\hbar}mc^2 t} = \begin{pmatrix} \varphi \\ \chi \end{pmatrix} e^{-\frac{i}{\hbar}mc^2 t}, \quad \text{such that} \quad \varphi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \ \chi = \begin{pmatrix} \psi_3 \\ \psi_4 \end{pmatrix}.$$

Using this notation we get that the Dirac's equation (8.5) reads ($\vec{p}$ will be used as a shorthand for the momentum operator):

$$\begin{aligned}
& i\hbar\partial_t \begin{pmatrix} \varphi \\ \chi \end{pmatrix} e^{-\frac{i}{\hbar}mc^2 t} = (c\vec{\alpha}\cdot\vec{p} + \beta mc^2) \begin{pmatrix} \varphi \\ \chi \end{pmatrix} e^{-\frac{i}{\hbar}mc^2 t} \\
& i\hbar\left[\begin{pmatrix} \dot{\varphi} \\ \dot{\chi} \end{pmatrix} + mc^2 \begin{pmatrix} \varphi \\ \chi \end{pmatrix} \right] e^{-\frac{i}{\hbar}mc^2 t} = \left[c \begin{pmatrix} 0 & \vec{p}\cdot\vec{\sigma} \\ \vec{p}\cdot\vec{\sigma} & 0 \end{pmatrix} \begin{pmatrix} \varphi \\ \chi \end{pmatrix} + mc^2 \begin{pmatrix} \varphi & 0 \\ 0 & -\chi \end{pmatrix} \right] e^{-\frac{i}{\hbar}mc^2 t} \\
& \Rightarrow \begin{cases} i\hbar\dot{\varphi} + mc^2\varphi = c\vec{\sigma}\cdot\vec{p}\chi + mc^2\varphi \\ i\hbar\dot{\chi} + mc^2\chi = c\vec{\sigma}\cdot\vec{p}\varphi - mc^2\chi \end{cases} .
\end{aligned}$$

The first equation we have obtained contains some terms that cancel out, while the second one is the one that should be treated in order to get the non-relativistic limit:

$$\begin{cases} i\hbar\dot{\varphi} = c\vec{\sigma}\cdot\vec{p}\chi \\ i\hbar\dot{\chi} + 2mc^2\chi = c\vec{\sigma}\cdot\vec{p}\varphi \end{cases} ,$$

this limit can be achieved by considering $c \to \infty$ and by neglecting smaller term, such as the term $i\hbar\dot{\chi}$ which will have a negligible contribution with respect to the mass term, which contains c^2

$$\begin{cases} i\hbar\dot{\varphi} = c\vec{\sigma}\cdot\vec{p}\chi \\ 2mc^2\chi = c\vec{\sigma}\cdot\vec{p}\varphi \end{cases} .$$

Operating the right algebraic substitution we finally get the non-relativistic equation of motion, know as the **free Pauli's equation**:

$$i\hbar\dot{\varphi} = \frac{(\vec{\sigma}\cdot\vec{p})^2}{2m}\varphi,$$

which, remembering that $\sigma_i\sigma_j = \mathbb{1}\delta_{ij} + i\varepsilon_{ijk}\sigma_k$, is just the classical Schrödinger's equation.

We will now see that, considering electromagnetism, we can recover, in this classical limit, a theory which predicts the spin approach of Pauli, which is normally imposed in

classical quantum mechanics.
To do so we will need to use the generalized momentum:

$$\pi_\mu = P_\mu - \frac{e}{c}A_\mu, \quad P^\mu \to \pi^m u,$$

this technique takes the name of *minimal substitution*. In this way, the Dirac's equation reads:

$$i\hbar\partial_t \begin{pmatrix}\varphi\\ \chi\end{pmatrix} e^{-\frac{i}{\hbar}mc^2 t} = (c\vec{\alpha}\cdot\vec{\pi} + \beta mc^2 + e\phi)\begin{pmatrix}\varphi\\ \chi\end{pmatrix} e^{-\frac{i}{\hbar}mc^2 t},$$

where $\vec{\pi} = -i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}$.
Following the same approach as before we now get:

$$\begin{cases} i\hbar\dot{\varphi} = c\vec{\sigma}\cdot\vec{\pi}\chi + e\phi\varphi \\ i\hbar\dot{\chi} + 2mc^2\chi = c\vec{\sigma}\cdot\vec{p}\varphi + e\phi\chi \end{cases},$$

where the terms $i\hbar\dot{\chi}$ and $e\phi\chi$ are negligible in the non-relativistic limit $c\to\infty$. In this limit this reads:

$$\begin{cases} i\hbar\dot{\varphi} = c\vec{\sigma}\cdot\vec{\pi}\chi + e\phi\varphi \\ 2mc^2\chi = c\vec{\sigma}\cdot\vec{p}\varphi \end{cases} \qquad \Rightarrow \qquad i\hbar\dot{\varphi} = \frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m} + e\phi\varphi.$$

Now we cannot recover the Schrödinger's equation we could get by just classically adding the electromagnetic potentials terms. Expanding the squared term we get:

$$\sigma^i\sigma^k\pi^i\pi^k = (\mathbb{1}\delta^{ij} + i\varepsilon^{ijk}\sigma^k)\pi^i\pi^k.$$

Since ε^{ijk} is antisymmetric, only the antisymmetric part of $\pi^i\pi^j$ will not vanish, thus:

$$\sigma^i\sigma^k\pi^i\pi^k = \pi^i\pi^i + i\varepsilon^{ijk}\sigma^k\frac{\pi^i\pi^j - \pi^j\pi^i}{2} = \pi^i\pi^i + \frac{i}{2}\varepsilon^{ijk}\sigma^k[\pi^i,\pi^j].$$

The commutator can be evaluated observing that $[\partial^i,\partial^j] = [A^i,A^j] = 0$ and that

$$[p^i,A^j]\varphi = -i\hbar\partial^i(\varphi A^j) + i\hbar A^j\partial^i\varphi = -i\hbar A^j\partial^i\varphi - i\hbar\varphi\partial^i A^j + i\hbar A^j\partial^i\varphi = -i\hbar\varphi\partial^i A^j,$$

thus giving:

$$[\pi^i,\pi^j] = -\frac{e}{c}\big([p^i,A^j] + [A^i,p^j]\big) = -\frac{e}{c}\big([p^i,A^j] - [p^j,A^i]\big) = -\frac{e}{c}\big(\partial^i A^j - \partial^j A^i\big).$$

The equation we've got in the non-relativistic limit, using the cyclic proprieties of ε_{ijk}, now reads:

$$i\hbar\dot{\varphi} = \frac{1}{2m}\left\{\vec{\pi}^2 - \frac{\hbar e}{c}\varepsilon^{ijk}\partial^i A^j\sigma^k\right\}\varphi + e\phi\varphi,$$

which recalling $\varepsilon^{ijk}\partial^i A^j = \vec{\nabla}\times\vec{A} = \vec{B}$ becomes

$$i\hbar\dot{\varphi} = \frac{1}{2m}\left\{\vec{\pi}^2 - \frac{\hbar e}{c}B^k\sigma^k\right\}\varphi + e\phi\varphi.$$

We can now compute $\vec{\pi}^2$ which, remembering that $\vec{L} = \vec{r} \times \vec{p}$ and that for constant magnetic fields $\vec{A} = \frac{1}{2}\vec{B} \times \vec{r}$, becomes:

$$\begin{aligned}
\vec{\pi}^2 &= \vec{p}^2 + \frac{e^2}{c^2}\vec{A}^2 - \frac{e}{c}(\vec{A}\cdot\vec{p} + \vec{p}\cdot\vec{A}) = \vec{p}^2 + \frac{e^2}{c^2}\vec{A}^2 - \frac{e}{2c}(\vec{B}\times\vec{r}\cdot\vec{p} + \vec{p}\cdot\vec{B}\times\vec{r}) \\
&= \vec{p}^2 + \frac{e^2}{c^2}\vec{A}^2 - \frac{e}{c}\vec{r}\times\vec{p}\cdot\vec{B} = \vec{p}^2 + \frac{e^2}{c^2}\vec{A}^2 - \frac{e}{c}\vec{L}\cdot\vec{B}.
\end{aligned}$$

The full equation reads:

$$i\hbar\dot{\varphi} = \frac{1}{2m}\left\{\vec{p}^2\frac{e^2}{c^2}\vec{A}^2 - \frac{e}{c}\vec{L}\cdot\vec{B} - \frac{\hbar e}{c}\vec{B}\cdot\vec{\sigma}\right\}\varphi + e\phi\varphi,$$

which lastly, defining $\vec{S} = \frac{\hbar}{2}\vec{\sigma}$, gives the Pauli's equation with electromagnetic interactions

$$i\hbar\dot{\varphi} = \frac{1}{2m}\left\{\vec{p}^2 + \frac{e^2}{c^2}\vec{A}^2 - \frac{e}{c}\left(\vec{L}\cdot\vec{B} - 2\vec{B}\cdot\vec{S}\right)\right\}\varphi + e\phi\varphi,$$

In this way we have show that the spin interactions that are imposed in classical quantum mechanics, arise naturally in the Dirac's formalism, even predicting the gyromagnetic factor of 2 that is measured.

8.3.1 The spin operator

We now generalize the spin operator definition, that we have just given and that is valid in the classical regime, to the Dirac's formalism:

$$\vec{S} = \frac{\hbar}{2}\begin{pmatrix}\vec{\sigma} & 0 \\ 0 & \vec{\sigma}\end{pmatrix} = \frac{\hbar}{2}\Sigma, \qquad \Sigma^i = -\frac{i}{2}\varepsilon^{ijk}\alpha^j\alpha^k.$$

Neither this operator, neither the angular momentum operator $L^i = \varepsilon^{ijk}x^j p^k$, commute with the hamiltonian of the system (which is given by the Dirac's equation (8.5)), only their combination $\vec{J} = \vec{S} + \vec{L}$ commutes with the hamiltonian. Using natural units:

$$\begin{aligned}
\mathcal{H} &= \alpha^l p_l + \beta m, \\
[\mathcal{H}, L^i] &= [\alpha^l p_l, \varepsilon^{ijk}x^j p^k] + [\beta m, \varepsilon^{ijk}x^j p^k] \\
&= \alpha^l [p_l, x^j]\varepsilon^{ijk}p^k = -i\alpha^l\delta^j_l\varepsilon^{ijk}p^k = -i\varepsilon^{ijk}\alpha^j p^k, \\
[\mathcal{H}, S^i] &= -\frac{i}{4}[\alpha^l p_l, \varepsilon^{ijk}\alpha^j\alpha^k] - \frac{i}{4}[\beta m, \varepsilon^{ijk}\alpha^j\alpha^k], \\
[\beta, \alpha^j\alpha^k] &= \alpha^j[\beta, \alpha^k] + [\beta, \alpha^j]\alpha^k \\
&= \{\beta, \alpha^j\}\alpha^k - \alpha^j\{\beta, \alpha^k\} = 0, \\
[\alpha^l, \alpha^j\alpha^k] &= \{\alpha^l, \alpha^j\}\alpha^k - \alpha^j\{\alpha^l, \alpha^k\} \\
&= 2\delta^{lj}\alpha^k - 2\delta^{lk}\alpha^j \\
[\mathcal{H}, S^i] &= -\frac{i}{4}\varepsilon^{ijk}(2\delta^{lj}\alpha^k - 2\delta^{lk}\alpha^j)p^l = \frac{i}{4}\varepsilon^{ijk}(2\delta^{lk}\alpha^j + 2\delta^{lk}\alpha^j)p^l \\
&= i\varepsilon^{ijk}\alpha^j p^k \\
&\Rightarrow \qquad [\mathcal{H}, J^i] = [\mathcal{H}, L^i + S^i] = 0.
\end{aligned}$$

This shows that degeneracy of energy states can be removed by studying the spectrum of the total angular momentum. This is crucial since, when studying atomic spectra, only the introduction of spin (that the Klein-Gordon equation doesn't consider) can reproduce the right number of emission/absorption lines.

8.4 Relativistic invariance of the Dirac's equation

In this section we will study the invariance of the Dirac's equation, this will naturally lead us to the internal transformation of the wave function under a Boost, therefore giving us the solution of a particle in motion.

First of all, we should remind all the well known transformations of the object we are already familial with:

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu}, \qquad \partial^{i}_{\mu} = \Lambda_{\mu}{}^{\nu}\partial_{\nu}.$$

We want to find a linear transformation that maps the spinor ψ to the one that satisfy the Dirac's equation in the new reference frame, we will call this transformation $S(\Lambda) \in M_{4\times 4}\mathbb{C}$:

$$\left(\gamma^{\mu}\partial'_{\mu} + m\right)\psi' = \left(\gamma^{\mu}\Lambda_{\mu}{}^{\nu}\partial_{\nu} + m\right)S(\Lambda)\psi = 0.$$

By multiplying to the left by S^{-1} we obtain

$$\left(S^{-1}(\Lambda)\gamma^{\mu}\Lambda_{\mu}{}^{\nu}S(\Lambda)\partial_{\nu} + m\right)\psi = 0,$$

which should be again some sort of Dirac equation in the old frame of reference. In order to be so $S^{-1}(\Lambda)\gamma^{\mu}\Lambda_{\mu}{}^{\nu}S(\Lambda)$ should be equal to the Dirac's gammas, giving

$$S^{-1}(\Lambda)\gamma^{\mu}S(\Lambda) = \Lambda^{\mu}{}_{\nu}\gamma^{\nu},$$

that will be used to derive the explicit form of S.

We will study what happens under an infinitesimal transformation of the Lorentz Group:

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu} \quad \Rightarrow \quad S(\Lambda) = \mathbb{1} + \frac{i}{2}\omega_{\mu\nu}\Sigma^{\mu\nu},$$

where ω are the parameters (small due to the infinitesimal nature) of the transformation (6 in total, such as the angles of rotation or the rapidity of the boost) and Σ are some appropriate matrices yet to find. We can show that, from the definition of the Lorentz group (all the transformations that preserve the Minkowski metric), both $\omega_{\mu\nu}$ and $\Sigma^{\mu\nu}$ are antisymmetric. Using the condition obtained by the Dirac's equation:

$$\begin{aligned}
&\left(\mathbb{1} + \frac{i}{2}\omega_{\alpha\beta}\Sigma^{\alpha\beta}\right)\gamma^{\mu}\left(\mathbb{1} + \frac{i}{2}\omega_{\gamma\delta}\Sigma^{\gamma\delta}\right) = (\delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu})\gamma^{\nu} \\
&\Rightarrow \gamma^{\mu} - \frac{i}{2}\omega_{\alpha\beta}[\Sigma^{\alpha\beta}, \gamma^{\mu}] + O(\omega^2) = \gamma^{\mu} + \omega^{\mu}{}_{\nu}\gamma^{\nu} \\
&\Rightarrow \omega_{\alpha\beta}[\Sigma^{\alpha\beta}, \gamma^{\mu}] = 2i\omega^{\mu}{}_{\nu}\gamma^{\nu} = 2i\eta^{\mu\alpha}\omega_{\alpha\beta}\gamma^{\beta}.
\end{aligned}$$

Considering the antisymmetry of $\omega_{\alpha\beta}$ (which implies that only the antisymmetric part of the tensor, it is contracted to, doesn't vanish) we get:

$$[\Sigma^{\alpha\beta}, \gamma^{\mu}] = i\big(\eta^{\mu\alpha}\gamma^{\beta} - \eta^{\mu\beta}\gamma^{\alpha}\big).$$

This commutation relation can be used in order to obtain the explicit form Σ:

$$\begin{aligned}
i\big(\eta^{\mu\alpha}\gamma^{\beta} - \eta^{\mu\beta}\gamma^{\alpha}\big) &= i\left(\frac{1}{2}\{\gamma^{\alpha}, \gamma^{\mu}\}\gamma^{\beta} - \gamma^{\alpha}\frac{1}{2}\{\gamma^{\beta}, \gamma^{\mu}\}\right) \\
&= -\frac{i}{4}\left(\gamma^{\alpha}\{\gamma^{\beta}, \gamma^{\mu}\} - \{\gamma^{\alpha}, \gamma^{\mu}\}\gamma^{\beta}\right) + \frac{i}{4}\left(\{\gamma^{\alpha}, \gamma^{\mu}\}\gamma^{\beta} - \gamma^{\alpha}\{\gamma^{\beta}, \gamma^{\mu}\}\right) \\
&= -\frac{i}{4}[\gamma^{\alpha}\gamma^{\beta}, \gamma^{\mu}] + \frac{i}{4}[\gamma^{\beta}\gamma^{\alpha}, \gamma^{\mu}] = -\frac{i}{4}\Big[(\gamma^{\alpha}\gamma^{\beta} - \gamma^{\beta}\gamma^{\alpha}), \gamma^{\mu}\Big] \\
&\Rightarrow \Sigma^{\alpha\beta} = -\frac{i}{4}(\gamma^{\alpha}\gamma^{\beta} - \gamma^{\beta}\gamma^{\alpha}).
\end{aligned}$$

Now that we have found the explicit form of Σ we can utilize the exponential map to get any kind of transformation of the spinorial wave function:

$$S(\Lambda) = e^{\frac{i}{2}\omega_{\mu\nu}\Sigma^{\mu\nu}}.$$

Example 8.1 — Rotations in the spinorial space. A generic rotation of the spacial coordinates, along the z-axis of some angle φ is made by the matrix:

$$\Lambda^{\mu}{}_{\nu} = \left(e^{\omega}\right)^{\mu}{}_{\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi & 0 \\ 0 & -sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Considering the Taylor expansion of this rotation, up to the first order, we can get the expression for the infinitesimal rotation and therefore of the ω matrix:

$$\Lambda \approx \delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu} \quad \Rightarrow \quad \omega^{\mu}{}_{\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\varphi.$$

We can now substitute this matrix in the exponential map in order to get the rotation in the spinorial space:

$$
\begin{aligned}
S(\varphi) &= \exp\left\{\frac{i}{2}\omega_{\mu\nu}\Sigma^{\mu\nu}\right\} = \exp\left\{\frac{1}{8}\omega_{\mu\nu}(\gamma^{\mu}\nu^{\beta} - \gamma^{\nu}\gamma^{\mu})\right\} = \exp\left\{\frac{1}{4}\omega_{\mu\nu}\gamma^{\mu}\gamma^{\nu}\right\} \\
&= \exp\left\{\frac{1}{4}(\omega_{12}\gamma^1\gamma^2 + \omega_{21}\gamma^2\gamma^1)\right\} = \exp\left\{\frac{\varphi}{4}(\gamma^1\gamma^2 - \gamma^2\gamma^1)\right\} \\
&= \exp\left\{\frac{\varphi}{4}(-\beta\alpha^1\beta\alpha^2 + \beta\alpha^2\beta\alpha^1)\right\} = \exp\left\{\frac{\varphi}{4}(\alpha^1\beta\beta\alpha^2 + \alpha^1\beta\beta\alpha^2)\right\} \\
&= \exp\left\{\frac{\varphi}{2}\alpha^1\alpha^2\right\},
\end{aligned}
$$

in which we have used that $\{\alpha^i, \alpha^j\} = 2\delta^{ij},\ \{\beta, \alpha^i\} = 0,\ \{\beta, \beta\} = 2.$
Now, a direct computation of $\alpha^1\alpha^2$ gives:

$$
\begin{pmatrix} 0 & \sigma^1 \\ \sigma^1 & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix} = i\begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix},
$$

therefore the rotation now reads:

$$
S(\varphi) = \exp\left\{i\frac{\varphi}{2}\begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma_3 \end{pmatrix}\right\} = \cos\left(\frac{\varphi}{2}\right) + i\sin\left(\frac{\varphi}{2}\right)\begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma_3 \end{pmatrix}
$$

This expression shows how spinors behave differently from the usual vectors, since any 2π rotation won't result in the identity transformation.
This result can be generalized to account for any rotation along any axis $\vec{n}$ (versor of the axis):

$$
S(\varphi, \vec{n}) = \exp\left\{i\frac{\varphi}{2}\vec{n}\cdot\vec{\sigma}\right\} = \cos\left(\frac{\varphi}{2}\right) + i\sin\left(\frac{\varphi}{2}\right)\vec{n}\cdot\vec{\sigma}.
$$

Notice that rotations are unitary transformations.

All the calculations that we have made for rotations can be done again in order to find boosts: starting with a finite boost along the x-axis, defining the rapidity ω:

$$
\Lambda^{\mu}{}_{\nu} = \begin{pmatrix} \cosh\omega & -\sinh\omega & 0 & 0 \\ -\sinh\omega & \cosh\omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \Rightarrow \quad \omega_{01} = -\omega_{10} = \omega.^{3}
$$

The boost for the Dirac's spinor reads:

$$
S(\Lambda) = \exp\left\{\frac{\omega}{2}\gamma^0\gamma^1\right\} = \exp\left\{-\frac{\omega}{2}\alpha^1\right\} = \cosh\left(\frac{\omega}{2}\right) - \sinh\left(\frac{\omega}{2}\right)\alpha^1.
$$

This transformation is clearly not unitary, but it is what we call *pseudo-unitary*, since it

---

[2]Here we have lowered the first index using the metric matrix, thus we get different signs for ω_{01} and ω_{10}, as needed by the antisymmetry of $\omega_{\mu\nu}$.

satisfies $S^\dagger = \beta S^{-1}\beta$. In fact, it is easy to prove:

$$\begin{aligned}
\gamma^{\mu\dagger} &= -\beta\gamma^\mu\beta \\
\Sigma^{\mu\nu\dagger} &= \left(-\frac{i}{4}[\gamma^\mu,\gamma^\nu]\right)^\dagger = \frac{i}{4}[\gamma^{\nu\dagger},\gamma^{\mu\dagger}] = \frac{i}{4}\beta[\gamma^\nu,\gamma^\mu]\beta \\
&= \frac{i}{4}\beta[\gamma^\mu,\gamma^\nu]\beta = \frac{i}{4}\beta\Sigma^{\mu\nu}\beta.
\end{aligned}$$

8.4.1 The moving particle

As we already have anticipated, we can now use the spinorial representation of the Lorentz boost to get the *free moving particle solution* of the Dirac's equation.

Considering a moving particle, we can find the rapidity of the boost that gets us in the comoving reference frame with the particle:

$$\begin{aligned}
P^\mu &= (E,\vec{p}) = (m\gamma,\vec{\beta}\gamma), \quad \sinh\omega = \beta\gamma, \quad \cosh\omega = \gamma \\
\tanh\frac{\omega}{2} &= \frac{e^{\frac{\omega}{2}} - e^{-\frac{\omega}{2}}}{e^{\frac{\omega}{2}} + e^{-\frac{\omega}{2}}} = \frac{e^{\frac{\omega}{2}} - e^{-\frac{\omega}{2}}}{e^{\frac{\omega}{2}} + e^{-\frac{\omega}{2}}} \times \frac{e^{\frac{\omega}{2}} + e^{-\frac{\omega}{2}}}{e^{\frac{\omega}{2}} + e^{-\frac{\omega}{2}}} = \frac{e^\omega - e^{-\omega}}{2+e^\omega+e^{-\omega}} \\
&= \frac{\sinh\omega}{1+\cosh\omega} = \frac{\beta\gamma}{1+\gamma} = \frac{|\vec{p}|}{m+E}, \\
\cosh\frac{\omega}{2} &= \frac{e^{\frac{\omega}{2}} + e^{-\frac{\omega}{2}}}{2} = \sqrt{\frac{\left(e^{\frac{\omega}{2}} + e^{-\frac{\omega}{2}}\right)^2}{4}} = \sqrt{\frac{2+e^\omega+e^{-\omega}}{4}} \\
&= \sqrt{\frac{1+\cosh\omega}{2}} = \sqrt{\frac{1+\gamma}{2}} = \sqrt{\frac{m+E}{2m}}, \\
&\Rightarrow \quad S(\Lambda) = \sqrt{\frac{m+E}{2m}}\left(\mathbb{1} - \frac{|\vec{p}|}{m+E}\alpha^1\right).
\end{aligned}$$

We can now generalize to any moving reference frame, by considering that, in the reference frame in which the particle is at rest, any observer that sees the particle moving with $\vec{\beta}$ is moving with $-\vec{\beta}$:

$$S(\Lambda) = \sqrt{\frac{m+E}{2m}}\left(\mathbb{1} + \frac{\vec{p}\cdot\vec{\alpha}}{m+E}\right). \tag{8.9}$$

Just for the ease of calculation, we can express the (8.9) in matrix form:

$$\vec{p}\cdot\vec{\alpha} = \begin{pmatrix} 0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0 \end{pmatrix} p_1 + \begin{pmatrix} 0&0&0&-i\\0&0&i&0\\0&-i&0&0\\i&0&0&0 \end{pmatrix} p_2 + \begin{pmatrix} 0&0&1&0\\0&0&0&-1\\1&0&0&0\\0&-1&0&0 \end{pmatrix} p_3 = \begin{pmatrix} 0&0&p_3&p_1-ip_2\\0&0&p_1+ip_2&-p_3\\p_3&p_1-ip_2&0&0\\p_1+ip_2&-p_3&0&0 \end{pmatrix},$$

which, defining $p_\pm = p_1 \pm ip_2$, simplifies to

$$\vec{p}\cdot\vec{\alpha} = \begin{pmatrix} 0&0&p_3&p_-\\0&0&p_+&-p_3\\p_3&p_-&0&0\\p_+&-p_3&0&0 \end{pmatrix}.$$

Let's now apply this transformation to a Dirac spinor of a particle at rest:

$$\psi_1'(x^\mu) = \sqrt{\frac{m+E}{2m}}\left(\mathbb{1}+\frac{\vec{p}\cdot\vec{\alpha}}{m+E}\right)\begin{pmatrix}1\\0\\0\\0\end{pmatrix} e^{iP^\mu x_\mu} = \sqrt{\frac{m+E}{2m}}\begin{pmatrix}1\\0\\\frac{p_3}{m+E}\\\frac{p_+}{m+E}\end{pmatrix} e^{iP^\mu x_\mu},$$

where we also had to transform the argument of the exponential using Lorentz transformations.

8.4.2 Fermionic bilinears

In this section we will tackle the problem of finding invariant quantities under the internal transformations of the Dirac's spinor. In order to so we remind that all the internal transformation can be expressed as:

$$S(\Lambda) = \exp\left\{\frac{i}{2}\omega_{\mu\nu}\Sigma^{\mu\nu}\right\}, \qquad \Sigma^{\mu\nu} = -\frac{i}{4}[\gamma^m u, \gamma^\nu],$$

furthermore we will also need to remember that every Dirac's matrix anticommute with all the others.

First, let's define the **Dirac conjugate** of a spinor as:

$$\bar{\psi} = \psi^\dagger\beta = (\beta\psi)^\dagger, \tag{8.10}$$

we will see later on that this object will play a fundamental role to get invariant quantities. Those objects will transform as:

$$\bar{\psi}' = (\psi^\dagger)'\beta = (S(\Lambda)\psi)^\dagger\beta = \psi^\dagger S(\Lambda)^\dagger\beta = \psi^\dagger\beta\beta S(\Lambda)^\dagger\beta = \bar{\psi}\beta S(\Lambda)^\dagger\beta,$$

which using the pseudo-unitary propriety of S reads:

$$\bar{\psi}' = \bar{\psi}S(\Lambda)^{-1}.$$

It is easy to see that, due to this transformation law, the product $\bar{\psi}\psi$ is invariant under the Lorentz transformations group:

$$\bar{\psi}'\psi' = \bar{\psi}S(\Lambda)^{-1}S(\Lambda)\psi = \bar{\psi}\psi,$$

we will call this and the other invariant and covariant quantities of a similar form **fermionic bilinears**.
We should now stress that the product $\psi^\dagger\psi$, that we interpret as a conserved density (usually of probability), is not Lorentz-invariant. In fact this product is actually the first component of a 4-vector:

$$J^\mu = i\bar{\psi}\gamma^m u\psi \quad \Rightarrow \quad J^0 = i\bar{\psi}\gamma^0\psi = i\psi^\dagger\beta(-i\beta) = \psi^\dagger\psi.$$

We should now prove that J^μ transforms as a 4-vector:

$$(J^\mu)' = i\bar{\psi}'\gamma^\mu\psi' = i\bar{\psi}S(\Lambda)^{-1}\gamma^\mu S(\Lambda)\psi = i\bar{\psi}\gamma^\nu\psi\Lambda^\mu{}_\nu,$$

where we have used the definition of the matrix S that we found in the previous section $S(\Lambda)^{-1}\gamma^\mu S(\Lambda) = \Lambda^\mu{}_\nu \gamma^\nu$.

We can build others fermionic bilinears using all the other matrices of the Dirac's formalism, we will use in these proofs that γ^5 anticommutes with γ^μ and thus it commutes with $S(\Lambda)$ (which is a combination of the product $\gamma^\mu\gamma^\nu$):

- $\bar{\psi}\gamma^5\psi$, called **pseudo-scalar**,
$$\bar{\psi}'\gamma^5\psi' = \bar{\psi}S^{-1}\gamma^5 S\psi = \bar{\psi}\gamma^5\psi;$$
- $\bar{\psi}\gamma^\mu\gamma^5\psi$, called **axial-vector**,
$$\bar{\psi}'\gamma^\mu\gamma^5\psi' = \bar{\psi}S^{-1}\gamma^\mu\gamma^5 S\psi = \bar{\psi}S^{-1}\gamma^\mu S\gamma^5\psi = \bar{\psi}\gamma^\nu\gamma^5\psi\Lambda^\mu{}_\nu;$$
- $\bar{\psi}\Sigma^{\mu\nu}\psi$, which transforms as a 4-tensor,
$$\begin{aligned}\bar{\psi}'\Sigma^{\mu\nu}\psi' &= \bar{\psi}S^{-1}\left(-\frac{i}{4}\right)[\gamma^\mu,\gamma^\nu]S\psi = -\frac{i}{4}\bar{\psi}[S^{-1}\gamma^\mu S, S^{-1}\gamma^\nu S]\psi \\ &= \bar{\psi}\left(-\frac{i}{4}\right)[\gamma^\alpha,\gamma^\beta]\psi\Lambda^\mu{}_\alpha\Lambda^\nu{}_\beta = \bar{\psi}\Sigma^{\alpha\beta}\psi\Lambda^\mu{}_\alpha\Lambda^\nu{}_\beta.\end{aligned}$$

Let's notice that we have proven that $S(\Lambda)$ is the transformation of the Dirac's matrices that leaves them invariant under Lorentz transformations:
$$\gamma^\mu = \Lambda^\mu{}_\nu S(\Lambda)\gamma^\nu S^{-1}(\Lambda).$$

8.5 Discrete transformation

We will now study a precise class of transformations, called **discrete**, which are part of the transformations group $O(1,3)$ but aren't part of the connect part of it $SO^+(1,3)$. The Dirac's equation (8.5) will have some symmetries under these transformations.

The first transformation is the **parity transform**, defined in the following way:
$$x^\mu = \begin{pmatrix} t \\ \vec{x} \end{pmatrix} \longmapsto \begin{pmatrix} t' \\ \vec{x}' \end{pmatrix} = \begin{pmatrix} t \\ -\vec{x} \end{pmatrix} \quad \Rightarrow \quad \mathbf{P}^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{8.11}$$

Under this transformation the differentiation operator transforms as $\partial'_\mu = \mathbf{P}_\mu{}^\nu\partial_\nu$, we can suppose that exist a parity transformation $\mathcal{P}$ of the spinor, in this way the Dirac's equation transforms:
$$(\gamma^\mu\partial'_\mu + m)\psi' = (\gamma^\mu\mathbf{P}_\mu{}^\nu\partial_\nu + m)\mathcal{P}\psi = 0.$$

Since the Dirac's equation must hold in all internal reference frame, we can use the above transformed equation to derive the explicit form of $\mathcal{P}$:
$$\mathcal{P}^{-1}(\gamma^\mu\mathbf{P}_\mu{}^\nu)\mathcal{P} = \gamma^\nu \quad \Rightarrow \quad \mathcal{P}^{-1}(\gamma^\mu)\mathcal{P} = \mathbf{P}^\mu{}_\nu\gamma^\nu = \begin{pmatrix} \gamma^0 \\ -\gamma^1 \\ -\gamma^2 \\ -\gamma^3 \end{pmatrix}.$$

This suggests that $\mathcal{P}$ must commute with γ^0 and anticommute with γ^i, in this way it is natural (and the easiest choice) to choose $\mathcal{P} = \beta$.
We can now study how transforms all the objects, that we have introduced, under parity:

$$\begin{aligned}
\bar{\psi} = \psi^\dagger \beta &\xrightarrow{\mathcal{P}} \bar{\psi}' = (\beta\psi)^\dagger \beta = \psi^\dagger = \bar{\psi}\beta \\
\bar{\psi}\psi &\xrightarrow{\mathcal{P}} \bar{\psi}'\psi' = \bar{\psi}\beta\beta\psi = \bar{\psi}\psi \\
\bar{\psi}\gamma^5\psi &\xrightarrow{\mathcal{P}} \bar{\psi}'\gamma^5\psi' = \bar{\psi}\beta\gamma^5\beta\psi = -\bar{\psi}\beta\beta\gamma^5\psi = -\bar{\psi}\gamma^5\psi \\
\bar{\psi}\gamma^\mu\psi &\xrightarrow{\mathcal{P}} \bar{\psi}'\gamma^\mu\psi' = \bar{\psi}\beta\gamma^\mu\beta\psi = \mathbf{P}^\mu{}_\nu\bar{\psi}\gamma^\mu\psi \\
\bar{\psi}\gamma^\mu\gamma^5\psi &\xrightarrow{\mathcal{P}} \bar{\psi}'\gamma^\mu\gamma^5\psi' = \bar{\psi}\beta\gamma^\mu\gamma^5\beta\psi = -\mathbf{P}^\mu{}_\nu\bar{\psi}\gamma^\mu\gamma^5\psi
\end{aligned}$$

In this way we have shown why $\bar{\psi}\gamma^5\psi$ is called pseudo-scalar and $\bar{\psi}\gamma^\mu\gamma^5\psi$ axial-vector (since they don't transform exactly as a scalar or a vector).
If we now reintroduce the chiral projectors (8.8) (notice that since S commutes with γ^5 the chiral components of every spinor don't mix under transformations of $SO^+(1,3)$[4]) we can observe that parity transformations mix the chiral components of a spinor (since β and γ^5 anticommute):

$$\mathcal{P}P_L\psi = \mathcal{P}\left(\frac{1-\gamma^5}{2}\psi\right) = \beta\frac{1-\gamma^5}{2}\psi = \frac{1+\gamma^5}{2}\beta\psi = P_R\psi.$$

Let's now study the second type of discrete transformation, the **time reversal**. As we have done for the parity transformation, the time reversal can be described by:

$$x^\mu = \begin{pmatrix} t \\ \vec{x} \end{pmatrix} \longmapsto \begin{pmatrix} t' \\ \vec{x}' \end{pmatrix} = \begin{pmatrix} -t \\ \vec{x} \end{pmatrix} \quad \Rightarrow \quad \mathbf{T}^\mu{}_\nu = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{8.12}$$

Transforming the Dirac's equation (8.5) we could still try the same approach used with the parity transformation, but in that way we would find out that it wouldn't work out. The right approach is to suppose that spinors transform as $\psi = \mathcal{T}\psi^*$.

$$(\gamma^\mu\partial'_\mu + m)\psi' = (\gamma^\mu\mathbf{T}_\mu{}^\nu\partial_\nu + m)\mathcal{T}\psi^* = 0$$

This can be turned into the defining equation of $\mathcal{T}$:

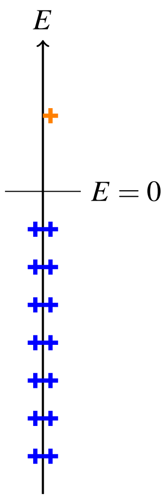
$$\mathcal{T}^{-1}(\gamma^\mu\mathbf{T}_\mu{}^\nu)\mathcal{T} = \gamma^{\nu*} \quad \Rightarrow \quad \mathcal{T}^{-1}(\gamma^\mu)\mathcal{T} = \mathbf{T}^\mu{}_\nu\gamma^{\nu*} = \begin{pmatrix} -\gamma^{0*} \\ \gamma^{1*} \\ \gamma^{2*} \\ \gamma^{3*} \end{pmatrix} = \begin{pmatrix} \gamma^0 \\ -\gamma^1 \\ \gamma^2 \\ -\gamma^3 \end{pmatrix}$$

This suggests that, up to a multiplicative factor, $\mathcal{T}$ should be equals to $\gamma^1\gamma^3$, since it must anticommute with these.

[4]Due to this the chiral components form the two irreducible representations of the Lorentz group.

8.5.1 Charge symmetry and the hole theory

In this section we will discuss the last discrete symmetry, which is **charge conjugation**. This is the symmetry of particle and antiparticle, and was suggested by the existence of negative energies states: Dirac proposed that we could interpret those states, in a certain sense, as antiparticles.



A simple representation of the Dirac's sea idea. Blue crosses are the negative energy solutions while the orange ones are the positive.

To understand this interpretation we should introduce the concept of the **Dirac's sea**, or simply the collection of all the negative energy states, which, being experimentally inaccessible, are considered all full. All the regular particles that we observe are, energetically speaking, above this sea (since they have positive energy) and thus they cannot lose their energy and jump inside the sea (because it is already full).
What could actually happen is to observe a **hole** in the sea, it is easy to see that this must be an antiparticle: in fact a particle (now with negative energy) could jump into the hole with the result that the energy and the charge of the system has to be those of the vacuum, thus zero.

$$\begin{cases} E(\text{hole}) + (-E_p) = E_{\text{Vac}} = 0 \\ Q(\text{hole}) + Q = Q_{\text{Vac}} = 0 \end{cases} \quad \Rightarrow \quad \begin{cases} E(\text{hole}) = E_p \\ Q(\text{hole}) = -Q \end{cases} .$$

In this way we can see that holes has a positive energy (so that we don't contradict the fact that we don't measure negative energy particles) and the opposite charge of the regular particle, which corresponds to our experimental observation about antiparticles.
We could think of an antiparticle-particle couple as a negative energy state getting excited and thus jumping outside the sea creating a positive energy particle and a hole.
Notice that the sea should have an infinite energy, but we have set that to be zero, predicting the renormalization of the vacuum energy of quantum field theory. However, we will see that this interpretation will be replaced by the results of quantum field theory itself.

Let's study what happens to particles and antiparticles in the Dirac's equation. Introducing electromagnetic interactions we expect the Dirac's equation to hold both for positive charges ψ and negative charges ψ_C:

$$[\gamma^\mu(\partial_\mu - ieA_\mu) + m]\psi = 0, \qquad [\gamma^\mu(\partial_\mu + ieA_\mu) + m]\psi_C = 0.$$

To obtain these couple of equations we can try to take the complex conjugate of the one for positive cherges:

$$[\gamma^{\mu *}(\partial_\mu + ieA_\mu) + m]\psi^* = 0,$$

this equation suggests that we have to find a proper way to transform $\gamma^{\mu *}$ to the normal ones. Introducing a transformation $\mathcal{A}$ such as:

$$\mathcal{A}\gamma^{\mu *}\mathcal{A}^{-1} = \gamma^\mu \quad \Rightarrow \quad [\mathcal{A}^{-1}\gamma^\mu\mathcal{A}(\partial_\mu + ieA_\mu) + m]\psi^* = 0,$$

by multiplying all the equation, to the left, by $\mathcal{A}$ we get the equation for the antiparticle, which is thus described by $\psi_C = \mathcal{A}\psi^*$.
However, is usually done is to find ψ_C using $\bar{\psi}$:

$$\psi_C = \mathcal{A}\psi^* = \mathcal{A}\beta\beta\psi^* = \mathcal{A}\beta(\psi^\dagger\beta)^t = \mathcal{A}\beta\bar{\psi}^t,$$

we thus define the charge conjugation operation $\mathcal{C} = \mathcal{A}\beta$.
The transformation law for γ^μ now reads:

$$\gamma^m u = \mathcal{C}\beta\gamma^{\mu *}(\mathcal{C}\beta)^{-1} = \mathcal{C}\beta\gamma^{\mu *}\beta\mathcal{C}^{-1} = -\mathcal{C}\gamma^{\mu t}\mathcal{C}^{-1},$$

since $\beta\gamma^\mu\beta = \gamma^{\mu\dagger}$.
This suggests that γ^0 and γ^2 should anticommute with $\mathcal{C}$, while the other two should commute

$$\mathcal{C}\gamma^\mu\mathcal{C}^{-1} = -\gamma^{\mu t} = \begin{pmatrix} -\gamma^0 \\ \gamma^1 \\ -\gamma^2 \\ \gamma^3 \end{pmatrix} \qquad \Rightarrow \qquad \mathcal{C} = \gamma^0\gamma^2.$$

Here we have used that γ^0 and γ^2 are symmetric while the other two are antisymmetric.

Let's study what happens to a Weyl's spinor under this transformation:

$$(\psi_L)_C = \mathcal{C}\bar{\psi}_L^t = \mathcal{C}(\beta P_L\psi^*) = \mathcal{C}P_R(\beta\psi^*) = P_R(\psi_L)_C,$$

so the charge conjugate of a left-handed spinor is right-handed (and vice versa).

8.6 Propagator of the Dirac's equation

We want to find the propagator of the Dirac's equation: such fucntion, as we already discussed for the Klein-Gordon equation, is a green function of the differential equation that we are studing. Such function should then satisfy:

$$(\partial\!\!\!/ + m)S(x-y) = \delta^4(x-y).$$

Again, to solve this differential equation we can use the Fourier transform:

$$S(x-y) = \int \frac{d^4p}{(2\pi)^4} e^{ip(x-y)}\tilde{S}(x-y),$$

plugging it into the above differential equation we get

$$(i\, p\!\!\!/ + m)\tilde{S}(p) = \mathbb{1}.$$

This algebraic equation leads to the solution:

$$\tilde{S}(p) = \frac{-\,p\!\!\!/ + m}{p^2+m^2},$$

which should be used with the Feynman prescription in order to get the propagator:

$$-iS(x-y) = \int \frac{d^4p}{(2\pi)^4} e^{ip(x-y)} \frac{-\,p\!\!\!/ + m}{p^2+m^2-i\varepsilon}.$$

8.7 Action of the Dirac's equation

We will end our discussion on the Dirac's equation by studying the action that, through the principle of least action, leads to such equation. Being a complex field, the action functional had to depend on both the ψ and its conjugate, however, in order to obtain a fermionic scalar Lorentz invariant, it will depend on $\bar{\psi}$:

$$\mathcal{S}[\psi,\bar{\psi}] = \int d^4x[-\bar{\psi}(\not{\partial}+m)\psi]. \tag{8.13}$$

Evaluating the variation $\delta\mathcal{S}$, under the assumption that the fields vanish on the frontier of the domain, we can check that this is actually the right action:

$$\delta\mathcal{S} = \int d^4x[-\bar{\delta}\psi(\not{\partial}+m)\psi - \bar{\psi}(\not{\partial}+m)\delta\psi]$$

Integrating by parts

$$= \int d^4x[-\bar{\delta}\psi(\not{\partial}+m)\psi + (\bar{\psi}\ \overleftarrow{\not{\partial}}+m\bar{\psi})\delta\psi - \partial_\mu(\bar{\psi}\gamma^\mu\delta\psi)] = 0,$$

here the last term vanishes due to boundary conditions while the other two must vanish giving the equations of motion (8.5).

We can now use the Nother's theorem on this action. The action (8.13) is clearly (being a bilinear) invariant under transformations of the $U(1)$ group $\psi \to \psi' = e^{i\alpha}\psi$. The variation of the field under this transformation is

$$\delta_\alpha\psi = \psi' - \psi = (1+i\alpha+O(\alpha^2))\psi - \psi \backsimeq i\alpha\psi, \qquad \delta_\alpha\bar{\psi} \backsimeq -i\alpha\bar{\psi},$$

by promoting α to a function of the coordinates, the variation of the action reads:

$$\delta_{\alpha(x)}\mathcal{S} = \int d^4x(-\partial_\mu)(\bar{\psi}\gamma^\mu\psi\alpha(x)) = -\int d^4x[(\partial_\mu\alpha)(i\bar{\psi}\gamma^\mu\psi) + \alpha\partial_\mu(i\bar{\psi}\gamma^\mu\psi)] = 0,$$

which gives us the conserved current $\partial_\mu J^\mu = \partial_\mu(i\bar{\psi}\gamma^\mu\psi) = 0$. This current can be generalized to the symmetry $U(1)\times SU(N)$ of a system of N particles.

Let's now study the action that can describe the chiral nature of fermions.
In order to do so we should study how different Weyl components mix in bilinear products (remind that $P^2_{R/L=P_{R/L}}$):

$$\begin{aligned}
\bar{\psi}\psi_{R/L} &= \bar{\psi}\frac{1\pm\gamma^5}{2}\psi = \psi^\dagger\beta\left(\frac{1\pm\gamma^5}{2}\right)^2\psi = \psi^\dagger\frac{1\mp\gamma^5}{2}\beta\frac{1\pm\gamma^5}{2}\psi \\
&= \bar{\psi}_{L/R}\psi_{R/L}, \\
\bar{\psi}\gamma^\mu\psi_{R/L} &= \bar{\psi}\gamma^\mu\frac{1\pm\gamma^5}{2}\psi = \psi^\dagger\beta\gamma^\mu\left(\frac{1\pm\gamma^5}{2}\right)^2\psi = \psi^\dagger\frac{1\pm\gamma^5}{2}\beta\gamma^\mu\frac{1\pm\gamma^5}{2}\psi \\
&= \bar{\psi}_{R/L}\gamma^\mu\psi_{R/L}.
\end{aligned}$$

Using these observations it is straight forward to see that the action (8.13) reads:

$$\mathcal{S} = \int d^4x[-\bar{\psi}_L\gamma^\mu\partial_\mu\psi_L - \bar{\psi}_R\gamma^\mu\partial_\mu\psi_R - m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)]. \tag{8.14}$$

Notice that components gets mixed only by the mass term of this equation, for this reason we cannot describe massive particles that are fully right or left-handed.

Lastly we should study a particular mass term that we could add to this action, called **Majorana mass**:

$$\mathcal{L}_{\text{Maj}} = \frac{M}{2}\psi^t \mathcal{C}^{-1}\psi + \frac{M}{2}(\psi^t \mathcal{C}^{-1}\psi)^\dagger, \tag{8.15}$$

this term is Lorentz invariant, since an infinitesimal Lorentz transformation is given by (remind that $\mathcal{C}^{-1}\gamma^\mu \mathcal{C} = -\gamma^{\mu\ t}$)

$$\begin{aligned}
\delta\psi &= \frac{i}{2}\omega_{\mu\nu}\Sigma^{\mu\nu}\psi, \\
\delta\psi^t &= \frac{i}{2}\omega_{\mu\nu}\psi^t\Sigma^{\mu\nu\ t} = -\frac{i}{2}\omega_{\mu\nu}\psi^t\mathcal{C}^{-1}\Sigma^{\mu\nu}\mathcal{C},
\end{aligned}$$

each variation cancels out the other. This term breaks the $U(1)$ symmetry of the system, since a constant phase would appear in the action. This can be used to reproduce particles that, experimentally, break this symmetry. We could thus define a **Majorana spinor**

$$\mu(x) = \psi_L(x) + (\psi_L)_C(x),$$

which can describe particle with a physical mass given by $\mu_{1,2}^2 = m^2 + M^2$.

9. Particles with spin greater than 1

9.1 Pauli-Fierz equations

In this chapter we will introduce the equations that govern the motion of all quantum particles. The two previous chapters described the two main equations that describe spin 0 particles (Klein-Gordon) and spin $\frac{1}{2}$ particles (Dirac). All the other types of particles can be described by these two equations, with some other condition: together they form the **Pauli-Fierz equations**.

Bosons, or particles with integer spin s, are described by a set of s spinors (such that $\phi_{\mu_1\ldots\mu_s}$ is totally symmetric) that obeys Klein-Gordon equation (7.1), the Pauli-Fierz equations are in this case:

$$\begin{cases} (\Box - m^2)\phi_{\mu_1\ldots\mu_s} = 0, \\ \partial^\mu \phi_{\mu\mu_2\ldots\mu_s} = 0, \\ \eta^{\mu_1\mu_2}\phi_{\mu_1\mu_2\ldots\mu_s} = 0 \end{cases} . \tag{9.1}$$

We should now notice that the solutions of these equations are too many with respect to the $2s+1$ degree of freedom that they should have.

Fermions, or particles with half-odd spin $s = \frac{1}{2} + s$, are described by s spinors (such that $\phi_{\mu_1\ldots\mu_s}$ is totally symmetric) governed by the Dirac's equation (8.5), the Pauli-Fierz equations are in this case:

$$\begin{cases} (\not{\partial} + m)\psi_{\mu_1\ldots\mu_s} = 0, \\ \partial^\mu \psi_{\mu\mu_2\ldots\mu_s} = 0, \\ \gamma^\mu \psi_{\mu\mu_2\ldots\mu_s} = 0 \end{cases} . \tag{9.2}$$

Again we have more degrees of freedom of the expected ones.

9.1.1 Plane waves for bosonic Pauli-Fierz equations

Consider a particle of spin s, we will suppose that the solution of the equations (9.1) are of the form:

$$\phi_{\mu_1\dots\mu_s}(x) = \varepsilon_{\mu_1\dots\mu_s}(p)e^{ipx}.$$

We want to further study this ansatz, therefore we will plug it inside all the 3 equations: the first equation leads to the mass-shell condition of the particle

$$\varepsilon_{\mu_1\dots\mu_s}(-p^2-m^2)(p)e^{ipx} = 0 \qquad \Rightarrow \qquad p^2+m^2 = 0.$$

We can now proceed to study the particle in its rest reference frame, there the second equation read:

$$p^{\mu}\varepsilon_{\mu\mu_2\dots\mu_s} = 0 \qquad \Rightarrow \qquad m\varepsilon_{0\mu_2\dots\mu_s} = 0,$$

due to symmetry we easily see that all time components vanish and we are left with $\varepsilon_{i_1\dots i_s} \neq 0$.
Last equation gives:

$$\varepsilon^{i}_{ii_2\cdots_s} = 0,$$

this implies that these tensors are traceless and symmetric, thus the degrees of freedom are less than what we could expect by analyzing the indices, that reduce to the expected $2s+1$.

9.2 Spin 1 particles

Spin 1 particles are described by a one index spinor A_μ that obeys just the first two equations of (9.1):

$$\begin{cases} (\Box - m^2)A_\mu = 0, \\ \partial^\mu A_\mu \end{cases},$$

these equations are also called **Proca equations**.
We can get these equations from the following lagrangian (with $F^{\mu\nu} = \partial^\mu A\nu - \partial^\nu A^\mu$):

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}m^2A^\mu A_\mu. \tag{9.3}$$

Evaluating the variation of the action given by this lagrangian we get:

$$\begin{aligned}\delta\mathcal{S} &= \int d^4x\left[-\frac{1}{2}F_{\mu\nu}\delta F^{\mu\nu} - m^2A_\mu\delta A^\mu\right] = \int d^4x\left[-\frac{1}{2}F_{\mu\nu}(\partial^\mu\delta A^\nu - \partial^\nu\delta A^\mu) - m^2A_\mu\delta A^\mu\right] \\ &= \int d^4x\left[-F_{\mu\nu}\partial^\mu\delta A^\nu - m^2A_\mu\delta A^\mu\right] = \int d^4x\left[\partial^\mu F_{\mu\nu} - m^2A_\nu\right]\delta A^\nu = 0,\end{aligned}$$

equations of motion are therefore given by

$$\partial^\mu F_{\mu\nu} = m^2 A_\nu.$$

We can differentiate this equation, and using the antisymmetry of $F^{\mu\nu}$ we get the second one of (9.1)

$$\partial^\nu \partial^\mu F_{\mu\nu} = m^2 \partial^\nu A_\nu = 0,$$

while the first one can be obtained inserting the definition of $F_{\mu\nu}$ inside the above equation of motion

$$\partial^\mu F_{\mu\nu} = \partial^\mu(\partial_\mu A_\nu - \partial_\nu A_\mu) = \Box A_\nu = m^2 A_\nu.$$

We can now study the plane wave solutions of this system: as we already saw in the previous section, the equation (9.2) will result in some conditions for the solution to hold:

$$\begin{cases} (-p^2 - m^2)\varepsilon_\mu(p)e^{ipx} = 0, \\ ip^\mu \varepsilon_\mu(p)e^{ipx} = 0 \end{cases} \qquad \Rightarrow \qquad \begin{cases} p^2 + m^2 = 0, \\ p^\mu \varepsilon_\mu(p) = 0 \end{cases},$$

we can study the last condition in comoving reference frame, in which $p^\mu = (m,0,0,0)$, that reads $m\varepsilon_0 = 0$. Therefore, we see that the polarization of the solutions has only 3 degree of freedom.

Lastly we can study the propagator, which is defined by the differential equation for the Green function

$$[(-\Box + m^2)\eta^{\mu\nu} + \partial^\mu \partial^\nu] G_{\nu\lambda}(x-y) = \delta^\mu_\lambda \delta^4(x-y),$$

which can be obtained as an alternative equation of motion from (9.3).
Using Fourier transform, we can get as solution:

$$G_{\mu\nu}(x-y) = \int \frac{d^4p}{(2\pi)^4} e^{ip(x-y)} \tilde{G}_{\mu\nu}(p), \qquad \tilde{G}_{\mu\nu}(p) = \frac{\eta_{\mu\nu} + \frac{p_\mu p_\nu}{m^2}}{p^2 + m^2}.$$

Notice that this propagator is not well-defined anymore in the massless limit: we are going to study this condition in the next section.

9.2.1 Massless case

10. Path integrals

10.1 A heuristic approach to path integrals

In this section we will introduce the main ideas of behind path integral considering one of the most fundamental experiments of quantum mechanics, the double slit experiment for electrons.
We will now, temporarily, forget the wave interpretation of quantum mechanics, and we will think to electrons as particles. Each one, classically, should follow a precise path, passing through one or the other slit, we will instead think that they can go through all possible path Γ_i. To each one we will associate one amplitude $A(\Gamma_i)$, that can be used to recover the probabilistic interpretation of quantum mechanics: in this way the probability of a particle to reach a certain point on the screen of the experimental setup is given by

$$P = \left| \sum_{\text{All path}} A(\Gamma_i) \right|^2.$$

We can test this idea (in a non-formal way) proposing that $A(\Gamma_i) = e^{\frac{i}{\hbar}\mathcal{S}[\Gamma_i]}$, we will show that this guess reproduces the interference pattern of the double slit experiment.
We will consider a path, from the first slit, that has a length of D, while the path that goes to the same point on the screen but passing through the second slit has a length $D+d$, with $D \gg d$. Considering electrons moving at a constant velocity, we can write the action evaluated on the two different path:

$$\mathcal{S}(\Gamma_1) = \frac{m}{2}\left(\frac{D}{T}\right)^2 T, \qquad \mathcal{S}(\Gamma_2) = \frac{m}{2}\left(\frac{D+d}{T}\right)^2 T = \frac{mD^2}{2T} + \frac{mDd}{T} + O(d^2),$$

where T is the time of flight of the particle. In this way we can get the amplitude associated with this point on the screen by

$$A = A(\Gamma_1) + A(\Gamma_2) = e^{i\frac{mD^2}{2T\hbar}}\left(1 + e^{i\frac{mDd}{T}}\right),$$

from this one we can clearly see that the probability has a maximum if $\frac{mDd}{T} = 2\pi n$. We can recognize that $m\frac{D}{T}$ is the momentum of the particle and then $\frac{\hbar}{p} = \lambda$, therefore

$$\frac{d}{\lambda} = 2\pi n \qquad n \in \mathbb{N},$$

this de Borglie relation shows how this path integral interpretation is an appropriate interpretation of quantum mechanics. In fact, using the approximation $d \approx l \sin\theta$, with l distance between the two slits, we obtain the formula describing the interference path that we experimentally observe.
The ideas coming from this experiment can be generalized to an infinite amount of slits on an infinite number of screens (so that there is a slit in each point of space), in this way we define the path integral

$$A = \int Dx\, e^{\frac{i}{\hbar}\mathcal{S}[x]}.$$

Notice that considering a classical system (for which $\mathcal{S} \gg \hbar$) the least action principle assures us that, considering a trajectory $x(t)$, the action

$$\frac{\mathcal{S}[x(t) + \delta x(t)]}{\hbar} \approx \frac{\mathcal{S}[x(t)]}{\hbar} + \frac{\delta\mathcal{S}[x(t)]}{\hbar}$$

is equals to $\frac{\mathcal{S}[x(t)]}{\hbar}$ nearby the classical solution, while it gets arbitrarily large far from it. In this way we get constructive interference around the classical solution while, every elsewhere destructive interference makes vanish all the other paths.

10.2 1-D path integrals

We will now show in a rigorous way that path integrals can be used to describe the evolution of a 1-D quantum mechanical system. To do so we will start by the definition of transition amplitude in regular quantum mechanics

$$A(\psi_i \to \psi_f, t) = \langle\psi_f|\, e^{-\frac{i}{\hbar}\hat{\mathcal{H}}t}\,|\psi_f\rangle\,,$$

from this one we introduce the completeness relation of the position operator twice

$$\langle\psi_f| \int dx_f \int dx_i\, |x_f\rangle\, \langle x_f|\, e^{-\frac{i}{\hbar}\hat{\mathcal{H}}t}\,|x_i\rangle\, \langle x_i|\,|\psi_f\rangle = \int dx_f \int dx_i \psi_f^*(x_f)\, \langle x_f|\, e^{-\frac{i}{\hbar}\hat{\mathcal{H}}t}\,|x_i\rangle\, \psi_f(x_i),$$

in this way we will have to calculate just the scalar product $\langle x_f|\, e^{-\frac{i}{\hbar}\hat{\mathcal{H}}t}\,|x_i\rangle$.
In order to evaluate this scalar product using the idea of checking every possible path (this

is equivalent to using the infinite slits that we introduced previously) we will divide N times the time $t \to \varepsilon = \frac{t}{N}$

$$\langle x_f| e^{-\frac{i}{\hbar}\hat{\mathcal{H}}t} |x_i\rangle = \langle x_f| \left(e^{-\frac{i}{\hbar}\hat{\mathcal{H}}\varepsilon}\right)^N |x_i\rangle ,$$

and then we inter $N-1$ identities that we are again going to expand using completeness of the position operator (we are going to relabel $x_i = x_0$ and $x_f = x_N$)

$$\langle x_f| e^{-\frac{i}{\hbar}\hat{\mathcal{H}}\varepsilon}\mathbb{1}\dots\mathbb{1}e^{-\frac{i}{\hbar}\hat{\mathcal{H}}\varepsilon} |x_i\rangle = \int \left(\prod_{k=1}^{N-1} dx_k\right)\left(\prod_{k=1}^{N} \langle x_k| e^{-\frac{i}{\hbar}\hat{\mathcal{H}}\varepsilon} |x_{k-1}\rangle\right).$$

The same procedure can be repeated N times with the momentum operator and its eigenkets

$$\begin{aligned}
&\int \left(\prod_{k=1}^{N-1} dx_k\right)\left(\prod_{k=1}^{N} \frac{dp_k}{2\pi\hbar}\right)\left(\prod_{k=1}^{N} \langle x_k|p_k\rangle \langle p_k| e^{-\frac{i}{\hbar}\hat{\mathcal{H}}\varepsilon} |x_{k-1}\rangle\right) \\
&= \int \left(\prod_{k=1}^{N-1} dx_k\right)\left(\prod_{k=1}^{N} \frac{dp_k}{2\pi\hbar}\right)\left(\prod_{k=1}^{N} e^{\frac{i}{\hbar}p_k x_k} \langle p_k| e^{-\frac{i}{\hbar}\hat{\mathcal{H}}\varepsilon} |x_{k-1}\rangle\right),
\end{aligned}$$

in this way we have to evaluate a different scalar product; we should notice that the form of this one helps us in the calculation since the operator $\hat{\mathcal{H}}$ is a function of momentum and position operators.

Until now, we have considered N slits, therefore we will now impose a limiting procedure, for which $\varepsilon \to 0$ and $N \to \infty$. In this limit we can consider the infinitesimal action of the time evolution operator as $e^{-\frac{i}{\hbar}\hat{\mathcal{H}}\varepsilon} \approxeq 1 - \frac{i}{\hbar}\hat{\mathcal{H}}\varepsilon$, thus

$$\langle p_k| \left[1 - \frac{i}{\hbar}\varepsilon\left(\frac{\hat{p}^2}{2m} + V(\hat{x})\right)\right] |x_{k-1}\rangle = \langle p_k|x_{k-1}\rangle \left[1 - \frac{i}{\hbar}\varepsilon\left(\frac{\hat{p}_k^2}{2m} + V(x_{k-1})\right)\right],$$

in this way the previous integral reads

$$\lim_{n\to\infty} \int \left(\prod_{k=1}^{N-1} dx_k\right)\left(\prod_{k=1}^{N} \frac{dp_k}{2\pi\hbar}\right) \exp\left\{\frac{i}{\hbar}\sum_{k=1}^{N}\left(p_k(x_k - x_{k-1}) - \varepsilon\mathcal{H}(p_k, x_{k-1})\right)\right\},$$

if we multiply and divide the sum by ε the sum in the exponential we obtain the Riemann definition of an integral

$$\lim_{n\to\infty} \int \left(\prod_{k=1}^{N-1} dx_k\right)\left(\prod_{k=1}^{N} \frac{dp_k}{2\pi\hbar}\right) \exp\left\{\frac{i}{\hbar}\sum_{k=1}^{N}\varepsilon\left(p_k\frac{x_k - x_{k-1}}{\varepsilon} - \mathcal{H}(p_k, x_{k-1})\right)\right\},$$

which we can recognize as the action (in the limit of $N \to \infty$ the integrand becomes the Legrand transform of the hamiltonian). We define **path integral** as this limit and we will indicate those by the following notation:

$$\int Dx(t)Dp(t)e^{\frac{i}{\hbar}\mathcal{S}[x(t),p(t)]}.$$

We want now to solve this kind of integral, this can be easily done when the hamiltonian contains only quadratic forms, since in those cases we can use the gaussian integral

$$\int_{-\infty}^{+\infty} dx\, e^{-\alpha\frac{x^2}{2}} = \sqrt{\frac{2\pi}{\alpha}},$$

that can be analytically continuated in all the complex plane.
With this tool we can integrate over the momentum space (since energy is quadratic with respect to momentum): to do so we complete the square and recognizing $\alpha = \frac{i\varepsilon}{\hbar m}$ we obtain the **path integral over configuration space**

$$\begin{aligned}
&\lim_{n\to\infty}\int\left(\prod_{k=1}^{N-1}dx_k\right)\left(\prod_{k=1}^{N}\frac{dp_k}{2\pi\hbar}\right)\exp\left\{\frac{i}{\hbar}\sum_{k=1}^{N}\varepsilon\left(p_k\frac{x_k-x_{k-1}}{\varepsilon}-\frac{p_k^2}{2m}-V(x_{k-1})\right)\right\}\\
&=\lim_{n\to\infty}\int\left(\prod_{k=1}^{N-1}dx_k\right)\left(\prod_{k=1}^{N}\frac{1}{2\pi\hbar}\sqrt{\frac{2\pi\hbar m}{i\varepsilon}}\right)\exp\left\{\frac{i}{\hbar}\sum_{k=1}^{N}\varepsilon\left(m\frac{(x_k-x_{k-1})^2}{2\varepsilon^2}-V(x_{k-1})\right)\right\}\\
&=\lim_{n\to\infty}\int\left(\prod_{k=1}^{N-1}dx_k\right)\left(\frac{m}{2\pi\hbar i\varepsilon}\right)^{\frac{N}{2}}\exp\left\{\frac{i}{\hbar}\sum_{k=1}^{N}\varepsilon\left(\frac{m}{2}\left(\frac{x_k-x_{k-1}}{\varepsilon}\right)^2-V(x_{k-1})\right)\right\}\\
&=\int Dx(t)e^{\frac{i}{\hbar}\mathcal{S}[x(t)]}.
\end{aligned}$$

Last, considering the free particle, we can again use the gaussian integral, over configuration space, to obtain

$$\int Dx(t)e^{\frac{i}{\hbar}\mathcal{S}[x(t)]}=\sqrt{\frac{m}{2\pi i\hbar t}}e^{\frac{i}{\hbar}m\frac{(x_f-x_i)^2}{2t}}.$$

One could show that this path integral satisfy the free Schrödinger equation in one dimension.

10.3 A general way to solve path integrals

We will now how to derive solutions of these path integrals in a more general way. Consider the path integral

$$A=\int Dx\, e^{\frac{i}{\hbar}\mathcal{S}[x]},$$

we introduce $x_c(t)$, the classical trajectory of motion, and small variations from it given by $\delta q(t)$ that vanish at initial and final time. The path integral can be computed with respect to q, since x_c is fixed

$$A=\int Dq\, e^{\frac{i}{\hbar}\mathcal{S}[x_c+\delta q]},$$

to do so we want to split the action in the classical part and the variation part; for a free particle this is easy since

$$\int dt\frac{m}{2}(\dot{x}_c^2+2\dot{x}_c\dot{q}+\dot{q}^2)$$

can be manipulated by integrating by parts $2\dot{x}_c\dot{q}$ which becomes proportional to the classical free Newton's equation $\ddot{x}_c=0$, thus $\mathcal{S}[x_c+\delta q]=\mathcal{S}[x_c]+\mathcal{S}[q]$, and then

$$A=e^{\frac{i}{\hbar}\mathcal{S}[x_c]}\int Dx\, e^{\frac{i}{\hbar}\mathcal{S}[q]}=ke^{\frac{i}{\hbar}\mathcal{S}[x_c]},$$

with $k=\int Dx\, e^{\frac{i}{\hbar}\mathcal{S}[q]}$ constant. This procedure reproduces the exact result that we derived previously with gaussian integrals.

IV Quantum field theory

11. Introduction to Quantum Filed Theory

11.1 Why quantum field theory?

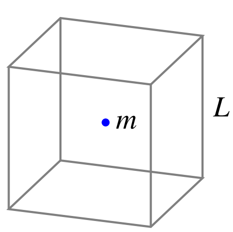
As we already discussed in Part III, in the first decades of the 19$^{\text{th}}$ century physics discovered that nature is described by two main theories:

- **special relativity**, at high energies;
- **quantum mechanics**, at short distances.

Therefore, relativistic quantum mechanics (Part III) was developed by promoting relativistic observable to operators acting in a Hilbert space. This approach, even explaining lots of different phenomena, revealed to be inconsistent and not powerful enough:

- it doesn't allow to change the number of particles of a system;
- it can violate causality;
- it predicts infinite states with negative energy.

It is easy to see that basics concepts of quantum mechanics and relativity leads to the need of a framework in which the total number of particles can change.
Let's consider a box, of side L, and a particle, of mass m, inside it.
From the Heisenberg's uncertainty principle:



A particle in a box of equal sides.

$$\Delta x \Delta p \geq \hbar \qquad \Rightarrow \qquad \Delta p \geq \frac{\hbar}{L},$$

since the uncertainty on the position of the particle is given by the fact that we just know that it is inside the box.
From relativity, we can use the energy impulse relation, that for high velocities is approximate by:

$$E^2 = p^2c^2 + m^2c^4 \approx p^2c^2 \qquad \Rightarrow \qquad \Delta E \approx c\Delta p.$$

Combining these two result we get the energy uncertainty of the particle if it's moving at relativistic speed:

$$\Delta E \geq \frac{\hbar c}{L}.$$

If we now make the box smaller we will reach a length scale for which the energies fluctuations can be bigger than the rest energy of the particle mc^2 and thus we could have more than just one particle.

Definition 11.1.1 Given some particle of mass m, we define the **Compton's wave length**:

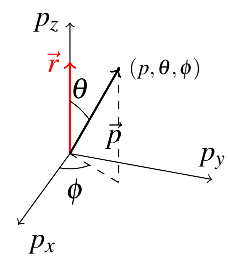
$$\lambda_C = \frac{\hbar}{mc}, \tag{11.1}$$

which is the length scale at which the notion of *"one single particle"* brakes.

Let's now show that relativistic quantum mechanics violates causality, letting have a non-zero probability of having particles with higher speed compared to c.
We will calculate the probability of a particle to move from $\vec{x}$ to $\vec{y}$ in a time t using the relativistic time evolution operator:

$$\begin{aligned} P(\vec{x} \to \vec{y}, t) &= |\langle \vec{x}| \, e^{-\frac{i}{\hbar}\sqrt{\hat{p}^2c^2+m^2c^4}t} \, |\vec{y}\rangle|^2 = \left| \int \frac{d^3p}{(2\pi\hbar)^3} \langle \vec{x}| \, e^{-\frac{i}{\hbar}\sqrt{\hat{p}^2c^2+m^2c^4}t} \, |\vec{p}\rangle \langle \vec{p}| \, |\vec{y}\rangle \right|^2 \\ &= \left| \int \frac{d^3p}{(2\pi\hbar)^3} \langle \vec{x}|\vec{p}\rangle \, e^{-\frac{i}{\hbar}\sqrt{p^2c^2+m^2c^4}t} \langle \vec{p}|\vec{y}\rangle \right|^2 = \left| \int \frac{d^3p}{(2\pi\hbar)^3} e^{-\frac{i}{\hbar}\sqrt{p^2c^2+m^2c^4}t} e^{\frac{i}{\hbar}\vec{p}(\vec{x}-\vec{y})} \right|^2. \end{aligned}$$

We will now set $\hbar = c = 1$, $\sqrt{p^2c^2+m^2c^4} = \omega_p$ and $\vec{x}-\vec{y} = \vec{r}$.
Using spherical coordinates, with $\vec{r} \parallel z-axis$:

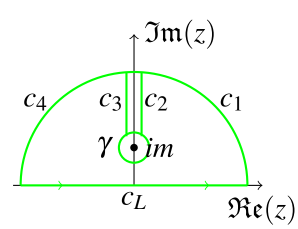


Spherical coordinates used for the integration over the momenta space.

$$\begin{aligned} &\int_0^\infty \frac{p^2 dp}{(2\pi)^3} \int_0^\pi d(-\cos\theta) \int_0^{2\pi} d\phi e^{-i\omega_p t + ipr\cos\theta} \\ &= \int_0^\infty \frac{p^2 dp}{(2\pi)^3} \frac{2\pi}{ipr} (e^{ipr} - e^{-ipr}) e^{-i\omega_p t} \\ &= \frac{-i}{(2\pi)^2 r} \int_{-\infty}^{+\infty} dp \; p e^{ipr} e^{-i\sqrt{p^2+m^2}t}. \end{aligned}$$

To solve this last integral we will use complex integration and Cauchy's theorem.
We will have to pay attention to this integration since the square root term makes the integrand polidrome, for this reason the path of integration cannot cross the imaginary axis over the point im (where we decide to put our branch cut).

$$\oint dz \; z e^{izr} e^{-i\sqrt{z^2+m^2}t} = \left(\int_{c_1} + \int_{c_2} + \int_{c_3} + \int_{c_4} + \int_{\gamma} + \int_{c_L} \right) z e^{izr} e^{-i\sqrt{z^2+m^2}t} dz = 0$$



Path used in the integration in the complex plane.

The integral that we need is given by integration over c_L in the limit of $c_L \to \mathbb{R}$, thus we need to evaluate all the other ones:

- using the Darboux's inequality it is easy to see that the integral over γ goes to 0 if the radius of the circumference (ε) goes too
$$\left|\int_\gamma z e^{izr} e^{-i\sqrt{z^2+m^2}t} dz\right| \leq 2\pi\varepsilon \sup_{z\in\gamma}\left|z e^{izr} e^{-i\sqrt{z^2+m^2}t}\right| \leq 2\pi m\varepsilon^2 \left|e^{-mr} e^{it\sqrt{2im\varepsilon e^{i\theta_{\max}}}}\right|;$$
- for the path c_4 we can express the integrand as
$$z e^{izr} e^{-i\sqrt{z^2+m^2}t} = z e^{i\Re\mathfrak{e}(z)r} e^{-i\Re\mathfrak{e}(\sqrt{z^2+m^2})t} e^{-\Im\mathfrak{m}(z)r} e^{\Im\mathfrak{m}(\sqrt{z^2+m^2})t},$$
since the c_4 is at the left of the branch cut $\Im\mathfrak{m}(\sqrt{z^2+m^2}) < 0$, thus the integrand vanishes for $|z| \to \infty$;
- for the path c_1 we cannot use the same approch as for c_4 but, since we want to show that the probability $P(\vec{x} \to \vec{y}, t)$ for $|\vec{x} - \vec{y}| = r > ct$ in non-zero, we can assume $r \gg t$ (we fixed $c = 1$), thus:
$$e^{-\Im\mathfrak{m}(z)r} e^{\Im\mathfrak{m}(\sqrt{z^2+m^2})t} \approx e^{-\Im\mathfrak{m}(z)r} \xrightarrow{|z| \to\infty} 0;$$
- the integrals over c_2 and c_3 are the only non-vanishing
$$\int_{c_L} z e^{izr} e^{-i\sqrt{z^2+m^2}t} dz = -\left(\int_{c_2} + \int_{c_3}\right) z e^{izr} e^{-i\sqrt{z^2+m^2}t} dz = -2\int_m^L y e^{-yr} \sinh(\sqrt{y^2-m^2}) dy$$

In the limit $c_L \to \mathbb{R}$ we can now use these results to get:
$$P(\vec{x} \to \vec{y}, t) \xrightarrow{|\vec{x}-\vec{y}| \gg ct} \left|\frac{2i}{(2\pi)^2 r}\int_m^{+\infty} y e^{-yr} \sinh(\sqrt{y^2-m^2}) dy\right|^2 \neq 0.$$
To make more evident that this probability is non-vanishing we can use that for $t > 0$:
$$\begin{aligned} &\sinh(\sqrt{y^2-m^2}) < e^{\sqrt{y^2-m^2}} < e^y t \\ \Rightarrow \quad &\frac{2i}{(2\pi)^2 r}\int_m^{+\infty} y e^{-yr} \sinh(\sqrt{y^2-m^2}) dy < \frac{2i}{(2\pi)^2 r}\int_m^{+\infty} y e^{-y(r-t)} dy \\ &= \frac{2i}{(2\pi)^2 r}\frac{m(r-t)+1}{(r-t)^2} e^{m(r-t)} \neq 0. \end{aligned}$$

11.2 Quantum field theory framework

The main idea behind quantum field theory is that each fundamental particle has its own associated quantum field (such as photons and the electromagnetic field). Those particles arise as quanta of excitations of their respective fields around the ground state or vacuum. This approach ensures:

- **Locality**: all interactions are local;
- **Causality**: nothing can escape the light cone;
- **Particle/antiparticle annihilation and creation**;
- **Identical particle**: all particle of the same field are identical;
- **Bosons and Fermions**: symmetry or antisymmetry of the wave function of a particle is not imposed as in classical quantum mechanics.

In order to describe our systems as fields we will promote the fields themselves to operators acting on a so-called **Fock's space**. Then we will have to impose commutation relations in order to quantize the system.

11.2.1 Units and scales

Before introducing the first models of quantum field theory, we will discuss the relations between the units that we will use and the scales of energy of the universe, to understand where we will be able to apply this theory.
The three main physical constants of nature are: the **speed of light**, the **Planck's constant** and the **Newton's constant**.

$$[c] = \frac{L}{T}, \qquad [\hbar] = M\frac{L^2}{T}, \qquad [G] = \frac{L^3}{MT}.$$

From now on we will use the so-called **natural units**: $c = \hbar = 1$.

$$\Rightarrow \qquad [c] = [\hbar] = 1, \qquad [G] = \frac{1}{M^2}.$$

In this system everything has dimensions of a power of mass unit, we define the notation of *mass-dimension* as:

$$[x] = d \quad \Rightarrow \quad [x] = M^d, \qquad \text{thus} \qquad [c] = [\hbar] = 0,\ [G] = -2,\ [\lambda_C] = -1.$$

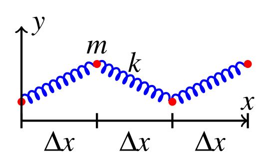
Lastly we define the **Planck mass** M_P and **Planck length** l_P:

$$M_P = G^{-\frac{1}{2}} \backsimeq 10^{19}\text{Gev}, \qquad l_P = M_P^{-1} \backsimeq 10^{-35}\text{m}.$$

The Planck mass defines the upper limit for which gravity is predominant over the other interactions. At lower energies ($M_{GUT} \backsimeq 10^{16}$GeV) we reach the point where electromagnetism, the weak and the strong interactions exists as independent interactions, exiting the **grand unified theories** energy scales. From 10^3 Gev (energy used in **LHC**) to 0.5 Mev we have the energy scale of the **standard model of particles**. At around 1 meV we find the energy of a neutrino and the energy associated to the **cosmological constant**.
Lastly we have the energy associated to the Hubble constant, or associated to the size of the universe ($10^{26}m$), which is 10^{-32} eV.
To have in mind this scale can help us to remember the energies for which we have tested the theoretical framework that we are going to build and for which it can hold.

11.3 The mechanical model of a quantum field

In this section we will build, from classical mechanics, a model of a field and then we will proceed to quantize it. In this way we will reproduce the steps needed to build a quantum field but always having in mind the classical meaning of it.
The model we are going to study is the mechanical model of a string. First of all we will consider it as N coupled harmonic oscillators, in the form of N massive identical particles, that can move only in the y direction, connected each other by springs to form a ring (in this way the last particle is connected to the first).
The lagrangian reads:



Representation of the system of N coupled oscillators.

$$L = T - U = \sum_{i=1}^{N} \frac{m\dot{y}_i^2}{2} - \sum_{i=1}^{N} \frac{k}{2}\left[y_i - y_{i-1}\right]^2,$$

for reasons that will be clearer later, we define $v^2 = \frac{k\Delta x^2}{m}$ and then we manipulate the lagrangian to get:

$$L = \frac{m}{2}\sum_{i=1}^{N}\left[\dot{y}_i^2 - v^2\left(\frac{y_i - y_{i-1}}{\Delta x}\right)^2\right].$$

Using Euler-Lagrange equations (1.2) we get N coupled second order differential equations:

$$\ddot{y}_i(t) = -\left(\frac{v}{\Delta x}\right)^2\big(2y_i(t) - y_{i+1}(t) - y_{i-1}(t)\big).$$

In order to decouple these equations, since we are assuming periodic conditions on this system, we will perform a *discrete Fourier Transform*:

$$\begin{aligned}
y_j(t) &= \frac{1}{\sqrt{N}}\sum_{s=0}^{N} e^{i\frac{2\pi}{N}sj}\tilde{y}_s(t),\\
\ddot{y}_j(t) &= \frac{1}{\sqrt{N}}\sum_{s=0}^{N} e^{i\frac{2\pi}{N}sj}\ddot{\tilde{y}}_s(t)\\
&= -\left(\frac{v}{\Delta x}\right)^2\frac{1}{\sqrt{N}}\sum_{s=0}^{N} e^{i\frac{2\pi}{N}sj}\big(2\tilde{y}_s(t) - e^{i\frac{2\pi}{N}s}\tilde{y}_s(t) - e^{-i\frac{2\pi}{N}s}\tilde{y}_s(t)\big)\\
&= -\left(\frac{v}{\Delta x}\right)^2\frac{1}{\sqrt{N}}\sum_{s=0}^{N} e^{i\frac{2\pi}{N}sj}\big(2 - e^{i\frac{2\pi}{N}s} - e^{-i\frac{2\pi}{N}s}\big)\tilde{y}_s(t)\\
&= -\left(\frac{v}{\Delta x}\right)^2\frac{2}{\sqrt{N}}\sum_{s=0}^{N} e^{i\frac{2\pi}{N}sj}\left(1 - \cos\left(\frac{2\pi s}{N}\right)\right)\tilde{y}_s(t)\\
&= -\frac{1}{\sqrt{N}}\sum_{s=0}^{N} e^{i\frac{2\pi}{N}sj}\left(\frac{2v}{\Delta x}\sin\left(\frac{\pi s}{N}\right)\right)^2\tilde{y}_s(t).
\end{aligned}$$

In this way we have obtained N decoupled harmonic oscillators:

$$\ddot{\tilde{y}}_s(t) = -\left(\frac{2v}{\Delta x}\sin\left(\frac{\pi s}{N}\right)\right)^2\tilde{y}_s(t) \quad\Rightarrow\quad \tilde{y}_s(t) = A_s e^{-i\omega_s t}, \quad \omega_s = \frac{2v}{\Delta x}\sin\left(\frac{\pi s}{N}\right).$$

Using again the Fourier transform we now can get an explicit solution:

$$y_j(t) = \frac{1}{\sqrt{N}}\sum_{s=0}^{N} e^{i\frac{2\pi}{N}sj}A_s e^{-i\omega_s t} = \frac{1}{\sqrt{N}}\sum_{s=0}^{N} A_s e^{i\frac{2\pi}{N}sj - i\omega_s t}, \quad \omega_s = \frac{2v}{\Delta x}\sin\left(\frac{\pi s}{N}\right),$$

this is the superposition of N harmonic waves each with its own waves number and velocity:

$$k_s = \frac{2\pi s}{N\Delta x}, \quad v_s = \frac{\omega_s}{k_s} = v\sin\left(\frac{\pi s}{N}\right)\frac{N}{\pi s}.$$

We should now notice that this model allows only $\frac{N}{2} - 1$ independent oscillators: the last oscillator has $\omega_s = 0$, for $s = \frac{N}{2}$ ω_s is maximum and for $s > \frac{N}{2}$ all the ω_s are the same of

those with $s < \frac{N}{2}$ (since the sine function is symmetric in the interval $[0,\pi]$).
Due to the fact that the equation of motion must be real, we can see that:

$$\begin{aligned}
y_j(t) &= \frac{1}{\sqrt{N}} \sum_{s=0}^{\frac{N}{2}-1} e^{i\frac{2\pi}{N}sj} \tilde{y}_s(t) + \frac{1}{\sqrt{N}} \sum_{s=\frac{N}{2}+1}^{N} e^{i\frac{2\pi}{N}sj} \tilde{y}_s(t) + \tilde{y}_{\frac{N}{2}}(t) + \tilde{y}_N(t) \\
&= \frac{1}{\sqrt{N}} \sum_{s=0}^{\frac{N}{2}-1} \left[e^{i\frac{2\pi}{N}sj} \tilde{y}_s(t) + e^{i\frac{2\pi}{N}j(N-s)} \tilde{y}_{N-s}(t) \right] + \tilde{y}_{\frac{N}{2}}(t) + \tilde{y}_N(t) \\
&= \frac{1}{\sqrt{N}} \sum_{s=0}^{\frac{N}{2}-1} \left[e^{i\frac{2\pi}{N}sj} \tilde{y}_s(t) + e^{-i\frac{2\pi}{N}js} \tilde{y}_{N-s}(t) \right] + \tilde{y}_{\frac{N}{2}}(t) + \tilde{y}_N(t) \\
&= y_j^* = \frac{1}{\sqrt{N}} \sum_{s=0}^{\frac{N}{2}-1} \left[e^{-\frac{2\pi}{N}sj} \tilde{y}_s^*(t) + e^{i\frac{2\pi}{N}js} \tilde{y}_{N-s}^*(t) \right] + \tilde{y}_{\frac{N}{2}}^*(t) + \tilde{y}_N^*(t) \\
&\Rightarrow \qquad \tilde{y}_s = \tilde{y}_{N-s}^*, \qquad \tilde{y}_s^* = \tilde{y}_{N-s}, \qquad \forall s \neq \frac{N}{2}, N.
\end{aligned}$$

This decoupling process can be used to diagonalize the lagrangian, which reads:

$$L = m \sum_{s=1}^{\frac{N}{2}-1} \left[|\dot{\tilde{y}}_s|^2 - \omega_s^2 |\tilde{y}_s|^2 \right].$$

We will now proceed to quantize this simple system. In order to do so we remind that the simple quantum oscillator, with an hamiltonian operator

$$\hat{\mathcal{H}} = \frac{\hat{P}^2}{2m} + \frac{\omega^2}{2} \hat{y}^2,$$

can be quantized introducing the **creation/annihilation operators**:

$$\begin{aligned}
\hat{a} &= \sqrt{\frac{\omega}{2}} \hat{y} + \frac{i}{\sqrt{2\omega}} \hat{P}, & \hat{a}^\dagger &= \sqrt{\frac{\omega}{2}} \hat{y} - \frac{i}{\sqrt{2\omega}} \hat{P}, \\
[\hat{a}, \hat{a}^\dagger] &= \hat{\mathbb{1}}, & [\hat{\mathcal{H}}, \hat{a}^\dagger] &= \omega \hat{a}^\dagger, \\
\hat{\mathcal{H}} &= \omega \left(\hat{N} + \frac{1}{2} \right), & \hat{N} &= \hat{a}^\dagger \hat{a}.
\end{aligned}$$

Every eigenstate of the hamiltonian is created using the $\hat{a}^\dagger$ operator: given the ground state $|0\rangle$ it easy to deduce that:

$$\hat{\mathcal{H}} \left(\hat{a}^\dagger \right)^n |0\rangle = (E_n + \omega) \left(\hat{a}^\dagger \right)^n |0\rangle, \quad \Rightarrow \quad \left(\hat{a}^\dagger \right)^n |0\rangle = \sqrt{n!}\, |n\rangle.$$

In order to quantize our system we will promote every $\tilde{y}_i$ to an operator:

$$\hat{\mathfrak{Re}} \tilde{y}_s = \frac{\hat{a}_s^{(R)} + \hat{a}_s^{(R)\dagger}}{\sqrt{2\omega_s}}, \qquad \hat{\mathfrak{Im}} \tilde{y}_s = \frac{\hat{a}_s^{(I)} + \hat{a}_s^{(I)\dagger}}{\sqrt{2\omega_s}}, \qquad s \in [1, N/2] \cap \mathbb{N}.$$

In this way each harmonic oscillator will have its own Hilbert space $\mathcal{H}_{\omega_s}$, the whole system will be described by a more complicated combination of these Hilbert spaces. We than define the Hilbert space $\mathcal{H}_k = \otimes_{s=1}^{k}\mathcal{H}_{\omega_s}$, a generic state is now described in a Fock Space $\mathcal{F} = \oplus_{i=1}^{N}\mathcal{H}_i$. The state of the system will be a linear combination of vectors in the form:

$$\left(\hat{a}_{i_1}^{(R)\dagger}\right)^{n_1^R}\left(\hat{a}_{i_2^R}^{(R)\dagger}\right)^{n_2}\dots\left(\hat{a}_{i_1}^{(I)\dagger}\right)^{n_1^I}\left(\hat{a}_{i_2}^{(I)\dagger}\right)^{n_2^I}\dots|0\rangle = |n_i^R, n_2^R, \dots, n_i^I, n_2^I, \dots\rangle\,,$$

where the vacuum state $|0\rangle$ should be intended as

$$|0\rangle = |0\rangle_{\omega_1}^{R}\otimes|0\rangle_{\omega_2}^{R}\otimes\dots|0\rangle_{\omega_1}^{I}\otimes|0\rangle_{\omega_2}^{I}\otimes\dots\,.$$

In this way, for example, the state of a system with 1 particle with energy ω_1 and 2 with energy ω_4 will be represented by:

$$\hat{a}_{\omega_1}^{(R)\dagger}\left(\hat{a}_{\omega_4}^{(R)\dagger}\right)^2|0\rangle = |1,0,0,2,0,\dots\rangle\,.$$

In this way the hamiltonian of the whole system becomes:

$$\hat{\mathcal{H}} = \sum_{s=1}^{\frac{N}{2}-1}\omega_s\left(\hat{N}_s + 1\right),$$

it is now possible to observe a non-trivial aspect of the theory that we are building, in fact the energy of the vacuum state, when there are many particles in the system, is not negligible due to the constant identity operator that appear in the hamiltonian.

We will now take the continuum limit, that considering $N\to\infty$ and $\Delta x\to 0$ will lead us to the generalization of the N solutions of our system to a field:

$$j\Delta x \quad \xrightarrow[\Delta x\to 0]{N\to\infty} \quad x, \qquad y_s(t) \quad \xrightarrow[\Delta x\to 0]{N\to\infty} \quad \psi(x,t).$$

The mechanical properties of the string, such as the velocity of the waves, also will be affected by this limiting procedure:

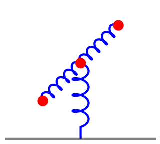
$$v_s = \frac{\omega_s}{k_s} = v\sin\left(\frac{\pi s}{N}\right)\frac{N}{\pi s} \xrightarrow[\Delta x\to 0]{N\to\infty} v, \qquad \omega_s = 2\frac{v}{\Delta x}\sin\left(\frac{\pi s}{N}\right) \xrightarrow[\Delta x\to 0]{N\to\infty} vk.$$

Quantizing this system we can obtain phonons which obey some sort of "relativistic massless energy" (considering $v = c$):

$$\omega_k = cK \quad\Rightarrow\quad \hbar\omega_k = \hbar cK \quad\Rightarrow\quad E_p = pc,$$

in this way we could interpret this filed, that arises from the string model, as a relativistic massless particle.

Modifying the initial discrete lagrangian, we could add a term to introduce mass. This term can be introduced by inserting a new elastic interaction which attracts every point mass to its own equilibrium point, independently of the others masses.



A sketch of the system with the interaction needed to consider massive particles.

$$L = \sum_{i=1}^{N}\left[\frac{m}{2}\dot{y}_i^2 - \frac{m}{2}v^2\left(\frac{y_i - y_{i-1}}{\Delta x}\right)^2 - \frac{K_\mu}{2}y_j^2\right].$$

To end this introduction we will study the continuum limit of this modified largrangian. The sum over all the point masses becomes an integral over the length of the string:

$$L = \int dx \left[\frac{m}{2} \left(\frac{\partial \psi}{\partial t} \right)^2 - \frac{m}{2} v^2 \left(\frac{\partial \psi}{\partial x} \right)^2 - \frac{K_\mu}{2} \psi^2 \right].$$

The integrand we have obtained is a **lagrangian density** and we will see that this will become the object describing the proprieties of the fields. This particular one is the lagrangian density of the **Klein-Gordon field**, which, as we supposed, describes massive relativistic spin zero particles. By the Euler-Lagrange equations this lagrangian density gives as equations of motion the Klein-Gordon equation :

$$m \partial_\mu \partial^\mu \psi + K_\mu \psi = 0.$$

Again this is a series of infinitely many coupled harmonic oscillators that can be decoupled by the Fourier Transform: each one of these will have a relativistic energy of a massive particle with a precise momentum.

12. Quantization of the Klein-Gordon field

12.1 The second quantization

In the discussion of relativistic quantum mechanics (Part III) we have derived the Klein-Gordon equation

$$(\Box + m^2)\varphi\big(x^\mu\big) = 0$$

from the first quantization of special relativity. In this section we will attempt the second quantization procedure of this equation, in the same manner we described introducing quantum field theory.

As we already mentioned in the previous section, the lagrangian density of the Klein-Gordon equation is given by:

$$\mathcal{L} = \frac{1}{2}\partial_\mu \partial^\mu \varphi - \frac{m^2}{2}\varphi^2, \tag{12.1}$$

where m is the mass of a single particle described by the field φ.
We now need to quantize the fields by promoting them to operators: these operators will be some functions of space but not of time, since the time evolution will be generated by a Schrödinger equation. We remind that in classical quantum mechanics we promote position and momentum to operators imposing the commutation relations given by the Poisson brackets

$$q_i, p_j = \frac{\partial L}{\partial \dot{q}_j} \quad \longrightarrow \quad \hat{q}_i, \hat{p}_j \quad \text{with} \quad [q_i, p_j] = i\delta_{ij}.$$

In order to promote the field φ, and its conjugate field $\pi = \frac{\partial \mathcal{L}}{\partial \partial_0 \varphi}$, we will impose similar commutation relations:

$$\varphi_i, \pi^j \quad \longrightarrow \quad \hat{\varphi}_i, \hat{\pi}^j \quad \text{with} \quad [\hat{\varphi}_i(\vec{x}), \hat{\pi}^j(\vec{y})] = i\delta^3(\vec{x}-\vec{y})\delta_i^j.$$

As we have done with the elastic string, we have to decouple all the harmonic oscillators that arise from the second derivative in time in the Klein Gordon equation. To do so we use the Fourier Transform in the differential equation itself:

$$\varphi(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}} \tilde{\varphi}(\vec{p},t) \quad \Rightarrow \quad \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}} \left(\frac{\partial^2}{\partial t^2} + \vec{p}^2 + m^2 \right) \tilde{\varphi}(\vec{p},t) = 0.$$

In this way we have recovered an infinite series of harmonic oscillators, characterized by $\omega_{\vec{p}} = \sqrt{\vec{p}^2 + m^2}$, which are the Fourier transform of the solutions $\tilde{\varphi}$. In this way every harmonic oscillator will be quantized by its own creation and annihilation operators:

$$\begin{cases} \hat{q} = \frac{1}{\sqrt{2\omega}}\left(\hat{a} + \hat{a}^\dagger\right) \\ \hat{p} = -i\sqrt{\frac{\omega}{2}}\left(\hat{a} - \hat{a}^\dagger\right) \end{cases} \qquad \Rightarrow \qquad \begin{cases} \hat{\tilde{\varphi}}^{\vec{p}} = \frac{1}{\sqrt{2\omega_{\vec{p}}}}\left(\hat{a}_{\vec{p}} + \hat{a}^\dagger_{\vec{p}}\right) \\ \hat{\tilde{\pi}}_{\vec{p}} = -i\sqrt{\frac{\omega_{\vec{p}}}{2}}\left(\hat{a}_{\vec{p}} - \hat{a}^\dagger_{\vec{p}}\right) \end{cases} .$$

These creation and annihilation operators must satisfy their own commutation relations (which are just generalizations of the classical ones):

$$[\hat{a}_{\vec{p}}, \hat{a}_{\vec{q}}] = [\hat{a}^\dagger_{\vec{p}}, \hat{a}^\dagger_{\vec{q}}] = 0 \quad \text{and} \quad [\hat{a}_{\vec{p}}, \hat{a}^\dagger_{\vec{q}}] = (2\pi)^3\delta^3(\vec{p}-\vec{q}).$$

By the Fourier transform we then get the two field operators as:

$$\hat{\varphi} = \int \frac{d^3p}{(2\pi)^3} \frac{e^{i\vec{p}\cdot\vec{x}}}{\sqrt{2\omega_{\vec{p}}}} \left(\hat{a}_{\vec{p}} + \hat{a}^\dagger_{\vec{p}}\right), \qquad \hat{\pi} = -i \int \frac{d^3p}{(2\pi)^3} \sqrt{\frac{\omega_{\vec{p}}}{2}} e^{i\vec{p}\cdot\vec{x}} \left(\hat{a}_{\vec{p}} - \hat{a}^\dagger_{\vec{p}}\right),$$

that, since the integration is over the whole momenta space and the solutions are the same for $\vec{p}$ or $-\vec{p}$ (they depend on its modulus), can be rewritten as

$$\hat{\varphi} = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left(\hat{a}_{\vec{p}} e^{i\vec{p}\cdot\vec{x}} + \hat{a}^\dagger_{\vec{p}} e^{-i\vec{p}\cdot\vec{x}} \right), \qquad \hat{\pi} = -i \int \frac{d^3p}{(2\pi)^3} \sqrt{\frac{\omega_{\vec{p}}}{2}} \left(\hat{a}_{\vec{p}} e^{i\vec{p}\cdot\vec{x}} - \hat{a}^\dagger_{\vec{p}} e^{-i\vec{p}\cdot\vec{x}} \right). \tag{12.2}$$

We can verify that these satisfy the commutation conditions with a simple calculation:

$$\begin{aligned} [\hat{\varphi}_i(\vec{x}), \hat{\pi}^j(\vec{y})] &= \frac{-i}{2} \int \frac{d^3p\, d^3q}{(2\pi)^6} \sqrt{\frac{\omega_{\vec{q}}}{\omega_{\vec{p}}}} \left[\hat{a}_{\vec{p}} e^{i\vec{p}\cdot\vec{x}} + \hat{a}^\dagger_{\vec{p}} e^{-i\vec{p}\cdot\vec{x}}, \hat{a}_{\vec{q}} e^{i\vec{q}\cdot\vec{x}} - \hat{a}^\dagger_{\vec{q}} e^{-i\vec{q}\cdot\vec{x}} \right] \\ &= \frac{i}{2} \int \frac{d^3p\, d^3q}{(2\pi)^6} \sqrt{\frac{\omega_{\vec{q}}}{\omega_{\vec{p}}}} (2\pi)^3 \delta^3(\vec{p}-\vec{q}) \left\{ e^{i(\vec{p}\cdot\vec{x}-\vec{q}\cdot\vec{y})} + e^{-i(\vec{p}\cdot\vec{x}-\vec{q}\cdot\vec{y})} \right\} \\ &= \frac{i}{2} \int \frac{d^3p}{(2\pi)^3} \left\{ e^{i(\vec{p}\cdot\vec{x}-\vec{p}\cdot\vec{y})} + e^{-i(\vec{p}\cdot\vec{x}-\vec{p}\cdot\vec{y})} \right\} = i\delta(\vec{x}-\vec{y}). \end{aligned}$$

We now want to find the hamiltonian of the system, in order to quantize it:

$$\begin{aligned}\mathcal{H} &= \int d^3x(\pi\partial_t\varphi - \mathcal{L}) = \int d^3x\left(\pi^2 - \frac{1}{2}\partial_\mu\partial^\mu\varphi - \frac{m^2}{2}\varphi^2\right)\\ &= \int d^3x\left(\pi^2 - \frac{1}{2}\pi^2 + \frac{1}{2}(\vec{\nabla}\varphi)^2 - \frac{m^2}{2}\varphi^2\right)\\ &= \int d^3x\left(\frac{1}{2}\pi^2 + \frac{1}{2}(\vec{\nabla}\varphi)^2 - \frac{m^2}{2}\varphi^2\right).\end{aligned}$$

The quantization procedure consists in the substitution of the fields with their operational counterparts, that we have previously derived:

$$\begin{aligned}\hat{\mathcal{H}} &= \int d^3x\left(\frac{1}{2}\hat{\pi}^2 + \frac{1}{2}(\vec{\nabla}\hat{\varphi})^2 - \frac{m^2}{2}\hat{\varphi}^2\right)\\ &= \frac{1}{2}\int\frac{d^3x\,d^3p\,d^3q}{(2\pi)^6}\left[-\frac{\sqrt{\omega_{\vec{p}}\omega_{\vec{q}}}}{2}\left(\hat{a}_{\vec{p}}e^{i\vec{p}\cdot\vec{x}} - \hat{a}^\dagger_{\vec{p}}e^{-i\vec{p}\cdot\vec{x}}\right)\left(\hat{a}_{\vec{q}}e^{i\vec{q}\cdot\vec{x}} - \hat{a}^\dagger_{\vec{q}}e^{-i\vec{q}\cdot\vec{x}}\right)+\right.\\ &\quad - \frac{\vec{p}\cdot\vec{q}}{2\sqrt{\omega_{\vec{p}}\omega_{\vec{q}}}}\left(\hat{a}_{\vec{p}}e^{i\vec{p}\cdot\vec{x}} - \hat{a}^\dagger_{\vec{p}}e^{-i\vec{p}\cdot\vec{x}}\right)\left(\hat{a}_{\vec{q}}e^{i\vec{q}\cdot\vec{x}} - \hat{a}^\dagger_{\vec{q}}e^{-i\vec{q}\cdot\vec{x}}\right) + \frac{m^2}{2\sqrt{\omega_{\vec{p}}\omega_{\vec{q}}}}\times\\ &\quad\left.\times\left(\hat{a}_{\vec{p}}e^{i\vec{p}\cdot\vec{x}} + \hat{a}^\dagger_{\vec{p}}e^{-i\vec{p}\cdot\vec{x}}\right)\left(\hat{a}_{\vec{q}}e^{i\vec{q}\cdot\vec{x}} + \hat{a}^\dagger_{\vec{q}}e^{-i\vec{q}\cdot\vec{x}}\right)\right].\end{aligned}$$

This integral can be solved with ease if we integrate first with respect to x, using the Fourier representation of the Dirac's delta function:

$$\begin{aligned}&\frac{1}{2}\int\frac{d^3p\,d^3q}{(2\pi)^3}\left[\delta^3(\vec{p}+\vec{q})\big(\hat{a}_{\vec{p}}\hat{a}_{\vec{q}} + \hat{a}^\dagger_{\vec{p}}\hat{a}^\dagger_{\vec{q}}\big)\left(-\frac{\sqrt{\omega_{\vec{p}}\omega_{\vec{q}}}}{2} - \frac{\vec{p}\cdot\vec{q}}{2\sqrt{\omega_{\vec{p}}\omega_{\vec{q}}}} + \frac{m^2}{2\sqrt{\omega_{\vec{p}}\omega_{\vec{q}}}}\right)+\right.\\ &\qquad\left.-\,\delta^3(\vec{p}-\vec{q})\big(\hat{a}_{\vec{p}}\hat{a}^\dagger_{\vec{q}} + \hat{a}^\dagger_{\vec{p}}\hat{a}_{\vec{q}}\big)\left(-\frac{\sqrt{\omega_{\vec{p}}\omega_{\vec{q}}}}{2} - \frac{\vec{p}\cdot\vec{q}}{2\sqrt{\omega_{\vec{p}}\omega_{\vec{q}}}} - \frac{m^2}{2\sqrt{\omega_{\vec{p}}\omega_{\vec{q}}}}\right)\right]\\ &= \frac{1}{2}\int\frac{d^3p}{(2\pi)^3}\left[\big(\hat{a}_{\vec{p}}\hat{a}_{\vec{p}} + \hat{a}^\dagger_{\vec{p}}\hat{a}^\dagger_{\vec{p}}\big)\left(-\frac{\omega_{\vec{p}}}{2} + \frac{\vec{p}^2}{2\omega_{\vec{p}}} + \frac{m^2}{2\omega_{\vec{p}}}\right)+\right.\\ &\qquad\left.-\big(\hat{a}_{\vec{p}}\hat{a}^\dagger_{\vec{p}} + \hat{a}^\dagger_{\vec{p}}\hat{a}_{\vec{p}}\big)\left(-\frac{\omega_{\vec{p}}}{2} - \frac{\vec{p}^2}{2\omega_{\vec{p}}} - \frac{m^2}{2\omega_{\vec{p}}}\right)\right].\end{aligned}$$

We should now notice that the first term of the integrand vanishes since

$$-\frac{\omega_{\vec{p}}}{2} + \frac{\vec{p}^2}{2\omega_{\vec{p}}} + \frac{m^2}{2\omega_{\vec{p}}} = \frac{1}{2\omega_{\vec{p}}}(-\omega^2_{\vec{p}} + \vec{p}^2 + m^2) = \frac{1}{2\omega_{\vec{p}}}(-\vec{p}^2 - m^2 + \vec{p}^2 + m^2) = 0,$$

and thus the integral reads:

$$\frac{1}{2}\int\frac{d^3p}{(2\pi)^3}\big(\hat{a}_{\vec{p}}\hat{a}^\dagger_{\vec{p}} + \hat{a}^\dagger_{\vec{p}}\hat{a}_{\vec{p}}\big)\left(\frac{\omega_{\vec{p}}}{2} + \frac{\vec{p}^2}{2\omega_{\vec{p}}} + \frac{m^2}{2\omega_{\vec{p}}}\right) = \frac{1}{2}\int\frac{d^3p}{(2\pi)^3}\omega_{\vec{p}}\big(\hat{a}_{\vec{p}}\hat{a}^\dagger_{\vec{p}} + \hat{a}^\dagger_{\vec{p}}\hat{a}_{\vec{p}}\big).$$

Using the commutator for the creation annihilation operator we can manipulate the last expression

$$\hat{a}_{\vec{p}}\hat{a}^\dagger_{\vec{p}} + \hat{a}^\dagger_{\vec{p}}\hat{a}_{\vec{p}} = [\hat{a}_{\vec{p}}, \hat{a}^\dagger_{\vec{p}}] + 2\hat{a}^\dagger_{\vec{p}}\hat{a}_{\vec{p}} = 2\hat{a}^\dagger_{\vec{p}}\hat{a}_{\vec{p}} + (2\pi)^3\delta^3(0),$$

therefore the hamiltonian operator reads:

$$\hat{\mathcal{H}} = \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\omega_{\vec{p}}\hat{a}^\dagger_{\vec{p}}\hat{a}_{\vec{p}} + \int d^3p\,\omega_{\vec{p}}\delta^3(0) \tag{12.3}$$

The first term that appears in the equation (12.3) is responsible for the energy contribution of each particle in the system, while the second one represent the energy of the vacuum. It is easy to see that in fact the eigenvalue of $|0\rangle$ is non-zero (as in classical quantum mechanics):

$$\hat{\mathcal{H}}|0\rangle = \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\omega_{\vec{p}}\hat{a}^\dagger_{\vec{p}}\hat{a}_{\vec{p}}|0\rangle + \int d^3p\,\omega_{\vec{p}}\delta^3(0)|0\rangle = \int d^3p\,\omega_{\vec{p}}\delta^3(0)|0\rangle .$$

This term is actually divergent for two reasons:

- $\delta^3(0)$ (**IR Divergence**) is some kind of infinity and it is due to the fact that vacuum has energy and the space we are considering is infinite, this can be shown considering Fourier representation of the delta

$$d^3(0) = \lim_{L\to\infty}\int_{L/2}^{L/2} d^3x\, e^{i\vec{p}\cdot\vec{x}}\bigg|_{\vec{p}=0} = \lim_{L\to\infty} L^3$$

  and in this way we can treat this divergence restricting the spacial domain of our system to a box;
- $\int d^3p\,\omega_{\vec{p}} \to \infty$ (**UV Divergence**), this divergence is due to incostincencies of the theory at high energies probably due to gravitational interactions, this last type of divergence is treated by introducing a *cut-off* where the theory stops to work.

Furthermore, we could define the **normal ordering** hamiltonian

$$:\hat{\mathcal{H}} := \hat{\mathcal{H}} - \langle 0|\hat{\mathcal{H}}|0\rangle ,$$

this one removes the divergent part of the hamiltonian leaving just energies meant as differences between the vacuum energy and the one of the system. This hamiltonian can be obtained by manipulation of the classical hamiltonian, before quantization.

Lastly, let's consider the state

$$|\vec{p}\rangle = \hat{a}^\dagger_{\vec{p}}|0\rangle ,$$

letting the hamiltonian act on this (in normal ordering) we get:

$$\begin{aligned}\hat{\mathcal{H}}|\vec{p}\rangle &= \frac{1}{2}\int \frac{d^3q}{(2\pi)^3}\omega_{\vec{q}}\hat{a}^\dagger_{\vec{q}}\hat{a}_{\vec{q}}\hat{a}^\dagger_{\vec{p}}|0\rangle = \frac{1}{2}\int \frac{d^3q}{(2\pi)^3}\omega_{\vec{q}}\hat{a}^\dagger_{\vec{q}}([\hat{a}_{\vec{q}},\hat{a}^\dagger_{\vec{p}}] + \hat{a}^\dagger_{\vec{p}}\hat{a}_{\vec{q}})|0\rangle \\ &= \frac{1}{2}\int \frac{d^3q}{(2\pi)^3}\omega_{\vec{q}}\hat{a}^\dagger_{\vec{q}}[\hat{a}_{\vec{q}},\hat{a}^\dagger_{\vec{p}}]|0\rangle = \frac{1}{2}\int \frac{d^3q}{(2\pi)^3}\omega_{\vec{q}}\hat{a}^\dagger_{\vec{q}}|0\rangle(2\pi)^2\delta^3(\vec{p}-\vec{q}) \\ &= \omega_{\vec{p}}\hat{a}^\dagger_{\vec{q}}|0\rangle ,\end{aligned}$$

which shows that the acting of the creation operator on the vacuum correspond to the creation of a particle. The same procedure can be used to show that the momentum operator $\hat{\vec{p}}$ acting on $|\vec{p}\rangle$ returns an eigenvalue of $\vec{p}$:

$$\hat{\vec{p}} = -\int d^3x\,\hat{\pi}\vec{\nabla}\hat{\varphi} = \int \frac{d^3p}{(2\pi)^3}\vec{p}\hat{a}^\dagger_{\vec{p}}\hat{a}_{\vec{p}}.$$

The same is true for the number operator

$$\hat{N} = \int \frac{d^3p}{(2\pi)^3} \hat{a}^\dagger_{\vec{p}} \hat{a}_{\vec{p}},$$

which returns the total number of particles in the Fock space. Whiteout interactions $\hat{N}$ commutes with the hamiltonian and therefore the number of particles is conserved. We should now notice that, being all the creation operators commuting, if they create particles with different momentum, all the states that can be described by this theory are symmetric with respect to exchange of two particles:

$$|\vec{q},\vec{p}\rangle = \hat{a}^\dagger_{\vec{q}} \hat{a}^\dagger_{\vec{p}} |0\rangle = \hat{a}^\dagger_{\vec{p}} \hat{a}^\dagger_{\vec{q}} |0\rangle = |\vec{p},\vec{q}\rangle .$$

We should now notice that this formalism is not Lorentz covariant, in fact, if we study the completeness relation, which should hold in every reference frame, it is easy to see that it is made up of to non-Lorentz invariant parts:

$$\hat{\mathbb{1}} = \int \frac{d^3p}{(2\pi)^3} |\vec{p}\rangle \langle \vec{p}| .$$

The natural choice to get an invariant internal expression is to integrate over d^4P, which transform with the absolute value of the determinant of a Lorentz transformation, which is 1, thus d^4P is actually Lorentz invariant.
We need to reduce this integral to the previous one, this can be done using a single Dirac's delta that fixes the energy $P^0 = \sqrt{\vec{p}^2 + m^2}$, paying attention not to include negative energies (using a θ Heaviside function):

$$\hat{\mathbb{1}} = \int \frac{d^4P}{(2\pi)^3} \delta((P^0)^2 - \vec{p}^2 - m^2)\theta(P^0) |\vec{p}\rangle \langle \vec{p}| .$$

The delta function can be manipulated resulting in:

$$\hat{\mathbb{1}} = \int \frac{d^3p}{(2\pi)^3 E_{\vec{p}}} |\vec{p}\rangle \langle \vec{p}| ,$$

that can have the same form of the original completeness relation by scaling every state by its energy $|p\rangle = \sqrt{E_{\vec{p}}} |\vec{p}\rangle$.

12.2 Charges and symmetries of the field

We will now study a system made up of 2 Klein-Gordon real fields, then we will see that symmetries in the lagrangian of these two fields lead naturally to the finding of some discrete conserved charge, that can be interpreted as an electric charge.

The lagrangian of a system of two Klein-Gordon field is composed by the sum of the lagrangian of both fields alone:

$$\mathcal{L} = \sum_{i=1}^{2} \left(\frac{1}{2} \partial_\mu \varphi_i \partial^\mu \varphi_i - \frac{m_i}{2} \varphi_i^2 \right).$$

This lagrangian will give raise to two different (due to the different masses) and independent fields, as solutions of two different and independent Klein-Gordon equations. Each system can be quantized, as we already have done, defining its own set of creation/annihilation operators that will result in two distinct hamiltonian, 3-momentum and number operators (one for each field):

$$\hat{\mathcal{H}}_i = \int \frac{d^3p}{(2\pi)^3}\omega_{i,\vec{p}}\hat{a}^\dagger_{i,\vec{p}}\hat{a}_{i,\vec{p}}, \quad \hat{\vec{p}}_i = \int \frac{d^3p}{(2\pi)^3}\vec{p}_i\hat{a}^\dagger_{i,\vec{p}}\hat{a}_{i,\vec{p}}, \quad \hat{N}_i = \int \frac{d^3p}{(2\pi)^3}\omega_{i,\vec{p}}\hat{a}^\dagger_{i,\vec{p}}\hat{a}_{i,\vec{p}}.$$

From these we can build their total version, accounting for particles of both the fields:

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_1 + \hat{\mathcal{H}}_2, \quad \hat{\vec{p}} = \hat{\vec{p}}_1 + \hat{\vec{p}}_2, \quad \hat{N} = \hat{N}_1 + \hat{N}_2.$$

In this construction we don't have degenerates, since two particles with the same momentum will have different energies, and so different hamiltonian eigenvalues, and their state will be in different eigenspace.
The same is not true anymore if we consider to fields with the same mass. In this case, even though the system is now degenerate, the lagrangian acquires a new symmetry. In fact, we can think to those two fields as a 2-D real vector $\vec{\varphi}$:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\vec{\varphi}^T\partial^\mu\vec{\varphi} - \frac{m}{2}\vec{\varphi}^T\vec{\varphi}, \tag{12.4}$$

it is now clear that all the rotation in the space of $\vec{\varphi}$ won't change the lagrangian. Therefore, the Lagrangian has a symmetry of the group $SO(2)$ that we can exploit in order to use the Nother's theorem.
Let's consider an infinitesimal rotation:

$$R = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \quad \Rightarrow \quad \delta R = \begin{pmatrix} 1 & \theta \\ -\theta & 1 \end{pmatrix} = \mathbb{1} + = \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix},$$

acting on the fields this gives us the variations

$$\begin{pmatrix} \varphi_1' \\ \varphi_2' \end{pmatrix} = \begin{pmatrix} 1 & \theta \\ -\theta & 1 \end{pmatrix}\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} = \begin{pmatrix} \varphi_1 + \theta\varphi_2 \\ \varphi_2 - \theta\varphi_1 \end{pmatrix} \quad \Rightarrow \quad \begin{cases} \delta\varphi_1 = \theta\varphi_2 \\ \delta\varphi_2 = -\theta\varphi_1 \end{cases}.$$

Now, Nother's theorem gives us a conserved current:

$$J^\mu = \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi_i}\delta\varphi_i = \partial^\mu\varphi_1\theta\varphi_2 - \partial^\mu\varphi_2\theta\varphi_1.$$

Neglecting θ (which is just a constant multiplicative factor), we can get a conserved charge from the first component of the current:

$$Q = \int d^3x\, J^0 = \int d^3x\, (\dot{\varphi}_1\varphi_2 - \dot{\varphi}_2\varphi_1), \qquad \frac{dQ}{dt} = 0.$$

We now have to quantize this observable, this can be done plugging in the quantized fields (12.2):

$$\hat{Q} = \int d^3x \, (\hat{\pi}_1 \hat{\varphi}_2 - \hat{\pi}_2 \hat{\varphi}_1)$$

$$\int d^3x \, \hat{\pi}_i \hat{\varphi}_j = \frac{-i}{2} \int \frac{d^3x \, d^3p \, d^3q}{(2\pi)^6} \sqrt{\frac{\omega_{\vec{q}}}{\omega_{\vec{p}}}} \left(\hat{a}_{i,\vec{q}} e^{i\vec{q}\cdot\vec{x}} - \hat{a}^\dagger_{i,\vec{q}} e^{-i\vec{q}\cdot\vec{x}} \right) \left(\hat{a}_{j,\vec{p}} e^{i\vec{p}\cdot\vec{x}} + \hat{a}^\dagger_{j,\vec{p}} e^{-i\vec{p}\cdot\vec{x}} \right)$$

$$= \frac{-i}{2} \int \frac{d^3p \, d^3q}{(2\pi)^3} \sqrt{\frac{\omega_{\vec{q}}}{\omega_{\vec{p}}}} \left[(\hat{a}_{i,\vec{q}} \hat{a}_{j,\vec{p}} - \hat{a}^\dagger_{i,\vec{q}} \hat{a}^\dagger_{j,\vec{p}}) \delta^3(\vec{p}+\vec{q}) + (\hat{a}_{i,\vec{q}} \hat{a}^\dagger_{j,\vec{p}} - \hat{a}^\dagger_{i,\vec{q}} \hat{a}_{j,\vec{p}}) \delta^3(\vec{p}-\vec{q}) \right]$$

$$= \frac{-i}{2} \int \frac{d^3p}{(2\pi)^3} \left[(\hat{a}_{i,-\vec{p}} \hat{a}_{j,\vec{p}} - \hat{a}^\dagger_{i,-\vec{p}} \hat{a}^\dagger_{j,\vec{p}}) + (\hat{a}_{i,\vec{p}} \hat{a}^\dagger_{j,\vec{p}} - \hat{a}^\dagger_{i,\vec{p}} \hat{a}_{j,\vec{p}}) \right]$$

$$\hat{Q} = \frac{-i}{2} \int \frac{d^3p}{(2\pi)^3} \Big[(\hat{a}_{1,-\vec{p}} \hat{a}_{2,\vec{p}} - \hat{a}^\dagger_{1,-\vec{p}} \hat{a}^\dagger_{2,\vec{p}}) + (\hat{a}_{1,\vec{p}} \hat{a}^\dagger_{2,\vec{p}} - \hat{a}^\dagger_{1,\vec{p}} \hat{a}_{2,\vec{p}}) +$$

$$- (\hat{a}_{2,-\vec{p}} \hat{a}_{1,\vec{p}} - \hat{a}^\dagger_{2,-\vec{p}} \hat{a}^\dagger_{1,\vec{p}}) - (\hat{a}_{2,\vec{p}} \hat{a}^\dagger_{1,\vec{p}} - \hat{a}^\dagger_{2,\vec{p}} \hat{a}_{1,\vec{p}}) \Big]$$

$$= \frac{-i}{2} \int \frac{d^3p}{(2\pi)^3} \left\{ [\hat{a}_{1,-\vec{p}}, \hat{a}_{2,\vec{p}}] + [\hat{a}^\dagger_{2,-\vec{p}} \hat{a}^\dagger_{1,\vec{p}}] + 2(\hat{a}_{1,\vec{p}} \hat{a}^\dagger_{2,\vec{p}} - \hat{a}_{2,\vec{p}} \hat{a}^\dagger_{1,\vec{p}}) \right\}$$

$$= -i \int \frac{d^3p}{(2\pi)^3} (\hat{a}_{1,\vec{p}} \hat{a}^\dagger_{2,\vec{p}} - \hat{a}_{2,\vec{p}} \hat{a}^\dagger_{1,\vec{p}})$$

To prove this expression we used the fact that operators of different fields commutes.

We should observe that the simple conservation law leaves this quantity determined up to a multiplicative factor (which can be used to set the units of this charge) and a constant. This last constant can be removed using normal ordering:

$$\hat{\hat{Q}} = (\hat{Q} + c) \quad \Rightarrow \quad \langle 0 | \hat{\hat{Q}} | 0 \rangle = c \quad \Rightarrow \quad : \hat{\hat{Q}} := \hat{\hat{Q}} - \langle 0 | \hat{\hat{Q}} | 0 \rangle = \hat{Q},$$

in this way the ambiguity that this constant could bring is removed.

Let's now find the spectrum of the charge operator we have just introduced. In order to do so we introduce the operators:

$$\hat{a}_\pm \triangleq \frac{1}{\sqrt{2}} (\hat{a}_{1,\vec{p}} \pm i \hat{a}_{2,\vec{p}}), \qquad \hat{a}^\dagger_\pm \triangleq \frac{1}{\sqrt{2}} (\hat{a}^\dagger_{1,\vec{p}} \mp i \hat{a^\dagger}_{2,\vec{p}}).$$

For these operators hold the following commutation relations:

$$[\hat{Q}, \hat{a}_{\pm,\vec{q}}] = -i \int \frac{d^3p}{(2\pi)^3} ([\hat{a}_{1,\vec{p}} \hat{a}^\dagger_{2,\vec{p}}, \hat{a}_{\pm,\vec{q}}] - [\hat{a}_{2,\vec{p}} \hat{a}^\dagger_{1,\vec{p}}, \hat{a}_{\pm,\vec{q}}])$$

$$= -i \int \frac{d^3p}{(2\pi)^3} (\hat{a}_{1,\vec{p}} [\hat{a}^\dagger_{2,\vec{p}}, \hat{a}_{\pm,\vec{q}}] + [\hat{a}_{1,\vec{p}}, \hat{a}_{\pm,\vec{q}}] \hat{a}^\dagger_{2,\vec{p}} - \hat{a}_{2,\vec{p}} [\hat{a}^\dagger_{1,\vec{p}}, \hat{a}_{\pm,\vec{q}}] - [\hat{a}_{2,\vec{p}}, \hat{a}_{\pm,\vec{q}}] \hat{a}^\dagger_{1,\vec{p}})$$

$$= -\frac{i}{\sqrt{2}} \int \frac{d^3p}{(2\pi)^3} (\mp i \hat{a}_{1,\vec{p}} (2\pi)^3 \delta^3(\vec{p}-\vec{q}) + 0 + \hat{a}_{2,\vec{p}} (2\pi)^3 \delta^3(\vec{p}-\vec{q}) - 0)$$

$$= \frac{1}{\sqrt{2}} (\mp \hat{a}_{1,\vec{q}} - i \hat{a}^\dagger_{2,\vec{q}}) = \mp \hat{a}_\pm,$$

$$
\begin{aligned}
[\hat{Q}, \hat{a}^\dagger_{\pm,\vec{q}}] &= -i \int \frac{d^3p}{(2\pi)^3} ([\hat{a}_{1,\vec{p}}\hat{a}^\dagger_{2,\vec{p}}, \hat{a}^\dagger_{\pm,\vec{q}}] - [\hat{a}_{2,\vec{p}}\hat{a}^\dagger_{1,\vec{p}}, \hat{a}^\dagger_{\pm,\vec{q}}]) \\
&= -i \int \frac{d^3p}{(2\pi)^3} (\hat{a}_{1,\vec{p}}[\hat{a}^\dagger_{2,\vec{p}}, \hat{a}^\dagger_{\pm,\vec{q}}] + [\hat{a}_{1,\vec{p}}, \hat{a}^\dagger_{\pm,\vec{q}}]\hat{a}^\dagger_{2,\vec{p}} - \hat{a}_{2,\vec{p}}[\hat{a}^\dagger_{1,\vec{p}}, \hat{a}^\dagger_{\pm,\vec{q}}] - [\hat{a}_{2,\vec{p}}, \hat{a}^\dagger_{\pm,\vec{q}}]\hat{a}^\dagger_{1,\vec{p}}) \\
&= -\frac{i}{\sqrt{2}} \int \frac{d^3p}{(2\pi)^3} (0 + \hat{a}^\dagger_{2,\vec{p}}(2\pi)^3\delta^3(\vec{p}-\vec{q}) - 0 \pm i\hat{a}^\dagger_{1,\vec{p}}(2\pi)^3\delta^3(\vec{p}-\vec{q})) \\
&= \frac{1}{\sqrt{2}}(\pm\hat{a}_{1,\vec{q}} - i\hat{a}^\dagger_{2,\vec{q}}) = \pm\hat{a}^\dagger_{\pm}.
\end{aligned}
$$

To prove these relations we have used the commutators:

$$
[\hat{a}^\dagger_{i,\vec{p}}, \hat{a}^\dagger_{j,\vec{q}}] = [\hat{a}_{i,\vec{p}}, \hat{a}_{j,\vec{q}}] = 0, \qquad [\hat{a}^\dagger_{i,\vec{p}}, \hat{a}_{i,\vec{q}}] = (2\pi)^3\delta^3(\vec{p}-\vec{q})\delta_{ij}.
$$

Considering a state $|s\rangle$ with charge q ($\hat{Q}|s\rangle = q|s\rangle$), we can use the commutation relations above to prove that $\hat{a}^\dagger_\pm$ is a ladder operator of the charge (the same can be done for $\hat{a_\pm}$):

$$
\hat{Q}\hat{a}^\dagger_\pm|s\rangle = ([\hat{Q}, \hat{a}^\dagger_\pm] + \hat{a}^\dagger_\pm\hat{Q})|s\rangle = (\pm\hat{a}^\dagger_\pm + \hat{a}^\dagger_\pm\hat{Q})|s\rangle = (q\pm 1)\hat{a}^\dagger_\pm|0\rangle.
$$

Since these ladder operators are linear combination of the ladder operators of the hamiltonian, 3-momentum and number operators, all the states generated by $\hat{a}^\dagger_\pm$ from vacuum are simultaneous eigenstates of $\hat{\mathcal{H}}, \hat{\vec{p}}, \hat{N}$ and $\hat{Q}$. This last one operator removes the degeneracy that was acquired considering two identical Klein-Gordon field, since there will be state with positive and others with negative charge (given the same momentum).
Therefore, every particle is fully described by its mass, its momentum and its charge. Opposite charge versions of the same particle will be interpreted as particle and antiparticle. Notice that this system can be reduced to a single real field only if the particles are all chargeless:

$$
0 = Q = \int d^3x(\dot{\varphi}_1\varphi_2 - \dot{\varphi}_2\varphi_1) \quad \Rightarrow \quad \varphi_1 = \varphi_2.
$$

To end the discussion of charges and fields we can introduce a new single complex filed, built using the previous two:

$$
\varphi = \frac{1}{\sqrt{2}}(\varphi_1 + i\varphi_2), \qquad \varphi^* = \frac{1}{\sqrt{2}}(\varphi_1^* + i\varphi_2^*).
$$

This field allows us to write the lagrangian density (12.4) as:

$$
\mathcal{L} = \partial_\mu\varphi^*\partial_\mu\varphi - m^2\varphi^*\varphi, \tag{12.5}
$$

this one is clearly invariant under $U(1)$ transformations, which is equivalent to the $O(2)$ symmetry of the system of two real fields we have just discussed. This symmetry leads again to a conserved current and charge:

$$
\begin{aligned}
J^\mu &= \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\delta\varphi + \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi^*}\delta\varphi^* = (\partial^\mu\varphi^*)i\theta\varphi - (\partial^\mu\varphi)i\theta\varphi^* \\
&= i\theta(\varphi\partial^\mu\varphi^* - \varphi^*\partial^\mu\varphi), \\
Q &= \int d^3x\, J^0 = \int d^3x\, (\varphi\partial_t\varphi^* - \varphi^*\partial_t\varphi) = \int d^3x\, [\varphi\pi^* - \varphi^*\pi], \\
\hat{Q} &= \int d^3x\, (\hat{\varphi}\hat{\pi}^* - \hat{\varphi}^*\hat{\pi}).
\end{aligned}
$$

Now, we can calculate these operators using the $\hat{a}_{\pm,\vec{p}}, \hat{a}^\dagger_{\pm,\vec{p}}$ operators, since they are defined as the same linear combination of creation/annihilation operators as the fields $\varphi,\ \varphi^*$:

$$\begin{cases} \hat{\varphi}(\vec{x}) = \frac{1}{\sqrt{2}}\big(\hat{\varphi}_1(\vec{x}) + i\hat{\varphi}_2(\vec{x})\big) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\left(\hat{a}_{+,\vec{p}}e^{i\vec{p}\cdot\vec{x}} + \hat{a}^\dagger_{-,\vec{p}}e^{-i\vec{p}\cdot\vec{x}}\right), \\ \hat{\varphi}^*(\vec{x}) = \frac{1}{\sqrt{2}}\big(\hat{\varphi}_1(\vec{x}) - i\hat{\varphi}_2(\vec{x})\big) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\left(\hat{a}_{-,\vec{p}}e^{i\vec{p}\cdot\vec{x}} + \hat{a}^\dagger_{+,\vec{p}}e^{-i\vec{p}\cdot\vec{x}}\right) \end{cases}.$$

Using these (and the operator $\hat{\pi}$) we can recover all the operators we have already used, that in normal ordering read:

$$\hat{Q} = \int \frac{d^3p}{(2\pi)^3}(\hat{a}^\dagger_{+,\vec{p}}\hat{a}_{+,\vec{p}} + \hat{a}^\dagger_{-,\vec{p}}\hat{a}_{-,\vec{p}}) = \hat{N}_+ - \hat{N}_-,$$
$$\hat{N}_\pm = \int \frac{d^3p}{(2\pi)^3}\hat{a}^\dagger_{\pm,\vec{p}}\hat{a}_{\pm,\vec{p}}.$$

The fields operators now represent particles/antiparticles couples, as $\hat{a}^\dagger_{+,\vec{p}}$ creates a particle and $\hat{a}^\dagger_{+,\vec{p}}$ an antiparticle. We should observe that these are always created with positive energies.

12.3 Heisenberg picture

Up until now, we have used the so-called **Schrödinger picture**, in which the time evolution is manged by a Schrödinger equation and a time evolution operator:

$$i\frac{d}{dt}\left|\vec{p}(t)\right\rangle = \hat{\mathcal{H}}\left|\vec{p}(t)\right\rangle, \qquad \left|\vec{p}(t)\right\rangle = \hat{U}(0,t)\left|\vec{p}(t)\right\rangle = e^{-iE_{\vec{p}}t}\left|\vec{p}(0)\right\rangle.$$

We will now introduce the **Heisenberg picture**, in which the time evolution is managed by every operator, in this way our field operators will be time-dependent:

$$\hat{O}_H(t) = e^{i\hat{\mathcal{H}}t}\hat{O}_S e^{-i\hat{\mathcal{H}}t}.$$

In this framework the time evolution of every operator is determined by its commutator with the hamiltonian at a time fixed:

$$\begin{aligned} \frac{d}{dt}\hat{O}(t) &= \frac{de^{i\hat{\mathcal{H}}t}}{dt}\hat{O}_S e^{-i\hat{\mathcal{H}}t} + e^{i\hat{\mathcal{H}}t}\hat{O}_S\frac{de^{-i\hat{\mathcal{H}}t}}{dt} \\ &= i\hat{\mathcal{H}}\hat{O}_H(t) - i\hat{O}_H(t)\hat{\mathcal{H}} = i[\hat{\mathcal{H}},\hat{O}_H(t)]. \end{aligned}$$

Recalling the commutators in Schrödinger picture

$$\begin{cases} [\hat{\varphi}(\vec{x}),\hat{\varphi}(\vec{y})] = [\hat{\pi}(\vec{x}),\hat{\pi}(\vec{y})] = 0 \\ [\hat{\varphi}(\vec{x}),\hat{\pi}(\vec{y})] = i\delta^3(\vec{x}-\vec{y}) \end{cases}$$

we can get the time evolution equation in the Heisenberg picture ($\varphi(t,\vec{x}) = \varphi(x)$)

$$\begin{aligned} \frac{\partial\hat{\varphi}(x)}{\partial t} &= i[\hat{\mathcal{H}},\hat{\varphi}(x)] = \frac{i}{2}\left[\int d^3y(\hat{\pi}^2 + (\vec{\nabla}\hat{\varphi})^2 + m^2\hat{\varphi}^2),\hat{\varphi}\right]_{\text{fixed t}} \\ &= \frac{i}{2}\int d^3y(\hat{\pi}[\hat{\pi},\hat{\varphi}] + [\hat{\pi},\hat{\varphi}]\hat{\pi}) = \frac{1}{2}\int d^3y(\hat{\pi}\delta^3(\vec{x}-\vec{y}) + \delta^3(\vec{x}-\vec{y})\hat{\pi}) \\ &= \hat{\pi}(x), \end{aligned}$$

$$
\begin{aligned}
\frac{\partial \hat{\pi}(x)}{\partial t} &= i[\hat{\mathcal{H}}, \hat{\pi}(x)] = \frac{i}{2}\left[\int d^3y\left(\hat{\pi}^2 + \left(\vec{\nabla}_y\hat{\varphi}\right)^2 + m^2\hat{\varphi}^2\right), \hat{\pi}\right]_{\text{fixed t}} \\
&= \frac{i}{2}\int d^3y([(\vec{\nabla}_y\hat{\varphi})^2, \hat{\pi}] + m^2[\hat{\varphi}^2, \hat{\pi}]) \\
&= \frac{i}{2}\int d^3y(\vec{\nabla}_y\hat{\varphi}\cdot[\vec{\nabla}_y\hat{\varphi}, \hat{\pi}] + [\vec{\nabla}_y\hat{\varphi}, \hat{\pi}]\cdot\vec{\nabla}_y\hat{\varphi} + m^2[\hat{\varphi}^2, \hat{\pi}]) \\
&= -\int d^3y(\vec{\nabla}_y\hat{\varphi}\cdot\vec{\nabla}_y\delta^3(\vec{x}-\vec{y}) + m^2\hat{\varphi}\delta^3(\vec{x}-\vec{y})) \\
&\qquad \text{Integrating by parts} \\
&= \nabla_x^2\hat{\varphi}(x) - m^2\hat{\varphi}(x).
\end{aligned}
$$

Combining these two we can get a differential equation for $\hat{\varphi}$:

$$\ddot{\hat{\varphi}} = (\vec{\nabla}^2 - m^2)\hat{\varphi},$$

which we can easily recognize as the Klein-Gordon equation (7.1), confirming that our second quantization procedure is consistent with the starting point of this theory.

We can now study the time evolution of the creation/annihilation operators:

$$(\hat{a}_{\vec{p}})_H = e^{i\hat{\mathcal{H}}t}\hat{a}_{\vec{p}}e^{-i\hat{\mathcal{H}}t} = \left([e^{i\hat{\mathcal{H}}t}, \hat{a}_{\vec{p}}] + \hat{a}_{\vec{p}}e^{i\hat{\mathcal{H}}t}\right)e^{-i\hat{\mathcal{H}}t},$$

in order to proceed in this calculation we should evaluate the above commutator, which, using the Taylor series definition of the exponential of an operator, reduces to the following calculations

$$
\begin{aligned}
[\hat{\mathcal{H}}, \hat{a}_{\vec{p}}] &= \int\frac{d^3q}{(2\pi)^3}\omega_{\vec{q}}[\hat{a}^\dagger_{\vec{q}}\hat{a}_{\vec{q}}, \hat{a}_{\vec{p}}] = \int\frac{d^3q}{(2\pi)^3}\omega_{\vec{q}}(\hat{a}^\dagger_{\vec{q}}[\hat{a}_{\vec{q}}, \hat{a}_{\vec{p}}] + [\hat{a}^\dagger_{\vec{q}}, \hat{a}_{\vec{p}}]\hat{a}_{\vec{q}}) \\
&= -\int\frac{d^3q}{(2\pi)^3}\omega_{\vec{q}}(2\pi)^3\delta^3(\vec{p}-\vec{q})\hat{a}_{\vec{q}} = -\omega_{\vec{p}}\hat{a}_{\vec{p}}, \\
\Rightarrow \qquad \hat{\mathcal{H}}\hat{a}_{\vec{p}} &= \hat{a}_{\vec{p}}(\hat{\mathcal{H}} - E_{\vec{p}}), \\
\Rightarrow \qquad \hat{\mathcal{H}}^n\hat{a}_{\vec{p}} &= \hat{a}_{\vec{p}}(\hat{\mathcal{H}} - E_{\vec{p}})^n, \\
\Rightarrow \quad [e^{i\hat{\mathcal{H}}t}, \hat{a}_{\vec{p}}] &= e^{i\hat{\mathcal{H}}t}\hat{a}_{\vec{p}} - \hat{a}_{\vec{p}}e^{i\hat{\mathcal{H}}t} = \hat{a}_{\vec{p}}e^{i(\hat{\mathcal{H}}-E_{\vec{p}})t} - \hat{a}_{\vec{p}}e^{i\hat{\mathcal{H}}t} = (e^{-E_{\vec{p}}t} - 1)\hat{a}_{\vec{p}}e^{i\hat{\mathcal{H}}t}, \\
\Longrightarrow \qquad (\hat{a}_{\vec{p}})_H &= e^{-iE_{\vec{p}}t}\hat{a}_{\vec{p}}, \qquad (\hat{a}^\dagger_{\vec{p}})_H = e^{-E_{\vec{p}}t}\hat{a}^\dagger_{\vec{p}}.
\end{aligned}
$$

Therefore, the field operator now reads:

$$\varphi(t,\vec{x}) = \int\frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\left(\hat{a}_{\vec{p}}e^{-i(E_{\vec{p}}t-\vec{p}\cdot\vec{x})} + \hat{a}^\dagger_{\vec{p}}e^{+i(E_{\vec{p}}t-\vec{p}\cdot\vec{x})}\right) = \int\frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\left(\hat{a}_{\vec{p}}e^{-ipx} + \hat{a}^\dagger_{\vec{p}}e^{+ipx}\right),$$

where px stands for $P^\mu x_\mu$.
We shall observe that in this way we obtained a Fourier expansion in time independent operators and in which time dependency is managed by the exponential terms.

Now that we have introduced time dependency, we should check if everything in our theory is Lorentz consistent. We have already checked the Lorentz invariance of this

framework, we still have to check if it holds causality.
First of all, let's notice that two operators should always commute if they represent observables measured in two space-time events which are reciprocally outside their light cones. In fact, in these cases the two measures cannot be dependent on each other (in the sense that there cannot be any correlation due to causality).

$$[\hat{O}_1(x), \hat{O}_2(y)] = 0 \quad \forall x, t \text{ such that } (x-y)^2 = (x^\mu - y^\mu)(x_\mu - y_\mu) < 0. \tag{12.6}$$

In order to check this requirement we define the propagator

$$\begin{aligned}\Delta(x-y) = [\hat{\varphi}(x), \hat{\varphi}(y)] &= \int \frac{d^3p}{(2\pi)^3}\frac{d^3q}{(2\pi)^3}\frac{(2\pi)^3}{2\sqrt{E_{\vec{p}}E_{\vec{q}}}}\delta^3(\vec{p}-\vec{q})(e^{-ipx+iqy} - e^{ipx-iqy}) \\ &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}(e^{-ip(x-y)} - e^{ip(x-y)}).\end{aligned} \tag{12.7}$$

We should notice that this operator is Lorentz invariant, due to the invariance of the measure of the integral and of the argument of the exponential.

Let's study the (12.7) inside the light-cone, therefore we will consider an inertial reference frame in which the events x, y happens at the same point in space, furthermore we will require $x-y$ to be a space-like 4-vector:

$$(x-y)^2 = t^2 > 0.$$

In this case the propagator (12.7) reads:

$$\Delta(x-y) = \frac{4\pi}{(2\pi)^3}\int_0^\infty dp \frac{\vec{p}^2}{2\sqrt{\vec{p}^2+m^2}}(e^{-i\sqrt{\vec{p}^2+m^2}t} - e^{i\sqrt{\vec{p}^2+m^2}t}),$$

this integral results in a combination of the Bessel's functions $J_1(mt)$ and $Y_1(mt)$, in the limit of $t \to \infty$ those tend to

$$\Delta(x-y) \sim (e^{-imt} - e^{imt}) \neq 0.$$

In this way we see that event inside the light-cone can be non-vanishing.

Considering $x-y$ time-like, in a reference frame where the two events happens at the same time, such that:

$$(x-y)^2 = -r^2 < 0.$$

Now, in this case the (12.7) reads:

$$\begin{aligned}\Delta(x-y) &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\sqrt{\vec{p}^2+m^2}}(e^{-i\vec{p}\cdot\vec{r}} - e^{i\vec{p}\cdot\vec{r}}) \\ &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\sqrt{\vec{p}^2+m^2}}e^{-i\vec{p}\cdot\vec{r}} - \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\sqrt{\vec{p}^2+m^2}}e^{i\vec{p}\cdot\vec{r}} \\ &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\sqrt{\vec{p}^2+m^2}}e^{-i\vec{p}\cdot\vec{r}} - \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\sqrt{\vec{p}^2+m^2}}e^{-i\vec{p}\cdot\vec{r}} = 0,\end{aligned}$$

in which we have changed variables in the second integral $\vec{p} \to -\vec{p}$. This shows that the propagator always vanishes outside the light-cone granting that causality holds (since observables depending on the field outside and inside the light-cone cannot be correlated).

Let's now try to compute the probability of a particle to exit the light cone, this will be evaluated as the modulus squared of

$$\begin{aligned}
D(x-y) &= \langle 0|\, \hat{\varphi}^\dagger(x)\hat{\varphi}(y)\, |0\rangle \\
&= \int \frac{d^3p\, d^3q}{(2\pi)^6} \frac{1}{2\sqrt{E_{\vec{p}}E_{\vec{q}}}} \langle 0|\, (\hat{a}_{\vec{p}} e^{-ipx} + \hat{a}^\dagger_{\vec{p}} e^{ipx})(\hat{a}_{\vec{q}} e^{-iqy} + \hat{a}^\dagger_{\vec{q}} e^{iqy})\, |0\rangle \\
&= \int \frac{d^3p\, d^3q}{(2\pi)^6} \frac{1}{2\sqrt{E_{\vec{p}}E_{\vec{q}}}} \langle 0|\, \hat{a}_{\vec{p}} \hat{a}^\dagger_{\vec{q}}\, |0\rangle\, e^{-ipx+iqy} \\
&= \int \frac{d^3p\, d^3q}{(2\pi)^6} \frac{1}{2\sqrt{E_{\vec{p}}E_{\vec{q}}}} \langle 0|\, ([\hat{a}_{\vec{p}}, \hat{a}^\dagger_{\vec{q}}] + \hat{a}^\dagger_{\vec{q}} \hat{a}_{\vec{p}})\, |0\rangle\, e^{-ipx+iqy} \\
&= \int \frac{d^3p\, d^3q}{(2\pi)^6} \frac{e^{-ipx+iqy}}{2\sqrt{E_{\vec{p}}E_{\vec{q}}}} \langle 0|\, ([\hat{a}_{\vec{p}}, \hat{a}^\dagger_{\vec{q}}])\, |0\rangle = \int \frac{d^3p\, d^3q}{(2\pi)^3} \frac{e^{-ipx+iqy}}{2\sqrt{E_{\vec{p}}E_{\vec{q}}}} \delta^3(\vec{p}-\vec{q}) \\
&= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} e^{-ip(x-y)},
\end{aligned}$$

imposing that $(x-y)^2 < 0$.
To simplify this calculation we can evaluate this integral in a reference frame where the two events happens at the same time

$$\begin{aligned}
D(x-y) &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} e^{i\vec{p}\cdot\vec{r}} = \frac{1}{(2\pi)^3} \int_0^{2\pi} d\varphi \int_{-1}^{1} d(cos\theta) e^{i|\vec{p}|rcos\theta} \int_0^\infty \frac{|\vec{p}|^2 dp}{2E_{\vec{p}}} \\
&= \frac{2\pi}{(2\pi)^3} \int_0^\infty dp \frac{e^{ipr} - e^{-ipr}}{2irE_{\vec{p}}} p = \frac{-i}{2(2\pi)^2 r} \int_{-\infty}^{+\infty} dp \frac{p}{\sqrt{p^2+m^2}} e^{ipr}.
\end{aligned}$$

To proceed we need to integrate over the complex plane, in the same way as we did in the introduction, obtaining a non-vanishing integral:

$$D(x-y) = -\frac{2i}{2(2\pi)^2 r} \int_m^{+\infty} dp \frac{-ye^{iyr}}{i\sqrt{y^2-m^2}} = \frac{1}{(2\pi)^2 r} \int_m^{+\infty} dy \frac{ye^{iyr}}{i\sqrt{y^2-m^2}} \neq 0.$$

This seems to contradict what we previously proved about causality, since there is a non-zero probability of a particle to exit the light-cone.
What really happens is that a particle can actually exit the light-cone but it cannot generate correlations, since we have proved that $\Delta(x-y)$ would vanish.
Now, we could physically interpret this strange behavior as the interference of another particle entering the light-cone, since it holds:

$$\langle 0|\, \Delta(x-y)\, |0\rangle = \langle 0|\, \hat{\varphi}^\dagger(x)\hat{\varphi}(y)\, |0\rangle - \langle 0|\, \hat{\varphi}(y)\hat{\varphi}^\dagger(x)\, |0\rangle = D(x-y) - D(y-x) = 0.$$

Considering a complex Klein-Gordon field we would get that the particle entering the light cone is actually the antiparticle of the field.

13. Quantization of the Dirac field

13.1 Dirac's representation

In this chapter we are going to quantize the Dirac's field, which is the one that obeys the Dirac's equation (8.5). We should pay attention, since in this approach we are going to use a different metric, for space-time, and a different basis of the Dirac representation.
We will define the Clifford algebra of the gamma matrices by $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$, so that the matrices are:

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1}_{2\times 2} \\ \mathbb{1}_{2\times 2} & 0 \end{pmatrix}, \qquad \gamma^\mu = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}.$$

It can be easily seen that all the operators, such as parity, that depends on β now depend on γ^0, since now it is this last one which obeys the commutators of β and α. We will show now an example.
A generic Lorentz transformation for a spinor ψ will be of the form

$$\psi \xrightarrow{Lorentz} \psi' = \left(e^{-\frac{i}{2}\omega_{\mu\nu}\Sigma^{\mu\nu}}\right)\psi,$$

with $\Sigma^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu]$, but we will use a different matrix representation of the Lorentz group, for which the transformation of ψ becomes

$$\psi \xrightarrow{Lorentz} \psi' = \left(e^{\frac{1}{2}\omega_{\mu\nu}S^{\mu\nu}}\right)\psi,$$

with $S^{\mu\nu} = i\Sigma^{\mu\nu} = -\frac{1}{4}[\gamma^\mu, \gamma^\nu]$.
The matrices $S = e^{\frac{1}{2}\omega_{\mu\nu}S^{\mu\nu}}$, has we have already discussed in the part III, these are pseudo

unitary, and with these different convention this propriety reads

$$S^\dagger = \gamma^0 S^{-1} \gamma^0.$$

Lastly, we can see that the Dirac conjugate has a different definition too

$$\bar{\psi} = \psi^\dagger \gamma^0.$$

Knowing all of these, we can start to discuss Dirac's theory.

13.2 Dirac's equation revisited

In this chapter we want to interpret Dirac's spinors formalism as a representation of the Lorentz group, and try to build the equations of motion from an action principle. Therefore, we should prove again all the results, such as the behavior of fermionic bilinears. Since those proof would be totally identical to what we have already proven during the discussion of relativistic quantum mechanics (Part III), we will just use those results.

A good candidate for the action must be Lorentz invariant and a scalar quantity. Furthermore, we want this action to depend on the field and, at least, its first derivatives. Therefore, the first two candidates we should consider are $\bar{\psi}\psi$ and $\bar{\psi}\gamma^\mu\partial_\mu\psi$; in analogy to the Klein-Gordon action, we could try to use $\bar{\psi}\psi$ in order to describe the mass of the particle. Lastly we want this action to be real, therefore each term that we are considering must be hermitian:

$$\begin{aligned}
(\bar{\psi}\psi)^\dagger &= \psi^\dagger(\psi^\dagger\gamma^0)^\dagger = \psi^\dagger\gamma^0\psi = \bar{\psi}\psi \\
(\bar{\psi}\gamma^\mu\partial_\mu\psi)^\dagger &= \partial_\mu\psi^\dagger(\gamma^\mu)^\dagger\bar{\psi}^\dagger = \partial_\mu\psi^\dagger\gamma^0\gamma^\mu\gamma^0\gamma^0\psi = \partial_\mu\bar{\psi}\gamma^\mu\psi,
\end{aligned}$$

this last term can be integrated by parts (letting vanish all border terms) inside the action resulting in:

$$-\bar{\psi}\gamma^\mu\partial_\mu\psi.$$

In order to remove the minus sign that appears we can multiply this term by i.

Given these observations, we can guess the lagrangian to be:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi, \tag{13.1}$$

notice that, since the mass dimension of the action has to be 0 and $[d^4x] = -4$ while $[\partial_m u] = 1$, this guess implies that $[\psi] = [\bar{\psi}] = \frac{3}{2}$.

It is now straightforward to use the Euler-Lagrange equations and obtain the equation of motion for this quantum system:

$$(i\gamma^\mu\partial_\mu - m)\psi(x) = 0, \qquad \bar{\psi}(x)(i\gamma^\mu\partial_\mu - m) = 0. \tag{13.2}$$

These two equations are first order linear differential equation (while Klein-Gordon was of the second order) but present a nice propriety: each component of $\psi(x)$ satisfy the Klein-Gordon equation. This can be shown in the following way:

$$\begin{aligned}
0 &= (i\gamma^\mu\partial_\mu + m)(i\gamma^\mu\partial_\mu - m)\psi(x) = -(\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu - im\gamma^\mu\partial_\mu + im\gamma^\mu\partial_\mu + m^2)\psi(x)\\
&= -(\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu + m^2)\psi(x) = -\left(\frac{1}{2}\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu + \frac{1}{2}\gamma^\mu\gamma^\nu\partial_\nu\partial_\mu + m^2\right)\psi(x)\\
&= -\left(\frac{1}{2}\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu + \frac{1}{2}\gamma^\nu\gamma^\mu\partial_\mu\partial_\nu + m^2\right)\psi(x) = -\left(\frac{1}{2}\{\gamma^\mu,\gamma^\nu\}\partial_\mu\partial_\nu + m^2\right)\psi(x)\\
&= -(\eta^{\mu\nu}\partial_\mu\partial_\nu + m^2)\psi(x) = -(\Box + m^2)\psi(x).
\end{aligned}$$

Now, introducing the right and left projectors

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} -\mathbb{1}_{2\times2} & 0 \\ 0 & \mathbb{1}_{2\times2} \end{pmatrix}, \qquad P_{R/L} = \frac{1\pm\gamma^5}{2},$$

which project each spinor on the two irreducible Weyl's representation of the Lorentz group, we can turn the lagrangian (13.1) into the one which depends explicitly on those components. Observing that

$$\begin{aligned}
\bar\psi\psi_{R/L} &= \bar\psi\frac{1\pm\gamma^5}{2}\psi = \psi^\dagger\beta\left(\frac{1\pm\gamma^5}{2}\right)^2\psi = \psi^\dagger\frac{1\mp\gamma^5}{2}\beta\frac{1\pm\gamma^5}{2}\psi\\
&= \bar\psi_{L/R}\psi_{R/L},\\
\bar\psi\gamma^\mu\psi_{R/L} &= \bar\psi\gamma^\mu\frac{1\pm\gamma^5}{2}\psi = \psi^\dagger\beta\gamma^\mu\left(\frac{1\pm\gamma^5}{2}\right)^2\psi = \psi^\dagger\frac{1\pm\gamma^5}{2}\beta\gamma^\mu\frac{1\pm\gamma^5}{2}\psi\\
&= \bar\psi_{R/L}\gamma^\mu\psi_{R/L},
\end{aligned}$$

it is straightforward to obtain the lagrangian:

$$\mathcal{L} = \bar\psi_L i\gamma^\mu\partial_\mu\psi_L + \bar\psi_R i\gamma^\mu\partial_\mu\psi_R - m(\bar\psi_L\psi_R + \bar\psi_R\psi_L). \tag{13.3}$$

This lagrangian manifestly shows how right and left-handed components are mixed by the mass term, therefore only massless particles can be fully right or left-handed.
We can see that the right or left-handed nature of a particle is related to what we call **helicity**, the projection of the spin on the momentum of the particle. Notice that this quantity can be defined only for massless particles, since for the massive ones a boost could affect the value of this quantity. Studying the components of equation (13.2)

$$\begin{pmatrix} 0 & i\partial_t + i\vec\sigma\cdot\vec\nabla \\ i\partial_t - i\vec\sigma\cdot\vec\nabla & 0 \end{pmatrix}\begin{pmatrix}\psi_L^{(W)}\\ \psi_R^{(W)}\end{pmatrix} - m\begin{pmatrix}\psi_L^{(W)}\\ \psi_R^{(W)}\end{pmatrix} = 0,$$

in the case of a massless particle we get two differential equations, one for each component, those can be expressed in terms of operators:

$$\begin{cases} i\partial_t\psi_R^{(W)} = -i\vec\sigma\cdot\vec\nabla\psi_R^{(W)} \\ i\partial_t\psi_L^{(W)} = i\vec\sigma\cdot\vec\nabla\psi_L^{(W)} \end{cases} \qquad \Rightarrow \qquad \begin{cases} \hat{\mathcal{H}}\psi_R^{(W)} = \hat{\vec{S}}\cdot\hat{\vec{p}}\psi_R^{(W)} \\ \hat{\mathcal{H}}\psi_L^{(W)} = -\hat{\vec{S}}\cdot\hat{\vec{p}}\psi_L^{(W)} \end{cases}.$$

Considering now that, for massless particles $E = |\vec{p}|$ we can define the helicity operator $\hat{\vec{S}} \cdot \vec{n}$, where $\vec{n}$ is the versor of the 3-momentum, and thus we can see that the elicity is well-defined for right or left-handed particles:

$$\begin{cases} \hat{\vec{S}} \cdot \vec{n} \psi_R^{(W)} = \psi_R^{(W)} \\ \hat{\vec{S}} \cdot \vec{n} \psi_L^{(W)} = -\psi_L^{(W)} \end{cases} .$$

13.2.1 Solution of the Dirac's equation

We will now study the plane wave solution of the equation (13.2), of the form

$$\psi_\alpha(x) = u_\alpha(\vec{p}) e^{-ipx},$$

where $u_\alpha(\vec{p})$ is a 4-D spinor. By plugging it in the equation (13.2) we get that p must satisfy the following relation:

$$(\gamma^\mu p_\mu - m) = 0 \quad \Rightarrow \quad \left[\begin{pmatrix} 0 & \mathbb{1}_{2\times 2} \\ -\mathbb{1}_{2\times 2} & 0 \end{pmatrix} p^0 + \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} p_i - \begin{pmatrix} \mathbb{1}_{2\times 2} & 0 \\ 0 & \mathbb{1}_{2\times 2} \end{pmatrix} m \right] u(\vec{p}) = 0,$$

defining the two vector of matrices $\sigma^\mu = (\mathbb{1}_{2\times 2}, \sigma^i)$ and $\bar{\sigma}^\mu = (\mathbb{1}_{2\times 2}, -\sigma^i)$, we can turn the above relation in a set of two relations for the Weyl's components of the spinor u:

$$\begin{cases} (p_\mu \bar{\sigma}^\mu) u_L = m\, u_R, \\ (p_\mu \sigma^\mu) u_R = m\, u_L. \end{cases}$$

In order to proceed we need to notice that

$$(p_\mu \sigma^\mu)(p_n u \bar{\sigma}^\nu) = (p_0 + p_i \sigma^i)(p_0 - p_j \sigma^j) = p_0^2 - p_i p_j \sigma^i \sigma^j = p_0^2 - p_i p_j (\delta^{ij} + \varepsilon^{ijk}) = m^2.$$

It is now easy to check that $u_L = A(p_\mu \sigma^\mu)\chi$ with A an arbitrary constant and χ an arbitrary constant 2-d spinor, satisfy those relations, since, substituting in the first equation we get

$$(p_\mu \bar{\sigma}^\mu) u_L = A(p_\mu \bar{\sigma}^\mu)(p_\mu \sigma^\mu)\chi = Am^2 \chi = m\, u_R, \qquad \Rightarrow \qquad u_R = Am\chi,$$

and then, substituting this result in the second equation

$$(p_\mu \sigma^\mu) u_R = Am(p_\mu \sigma^\mu)\chi = m\, u_L.$$

Therefore, the general solution is of the form:

$$u(\vec{p}) = A \begin{pmatrix} p_\mu \sigma^\mu \\ m \end{pmatrix} \chi,$$

we can now choose A and χ in such a way that this solution becomes more symmetric, we will use (where ξ is a normalized constant spinor, such that $\xi^\dagger \xi = 1$)

$$\begin{cases} A = \frac{1}{m} \\ \chi = \sqrt{p_\mu \bar{\sigma}^\mu} \xi \end{cases} \quad \Rightarrow \begin{cases} u_L = \frac{1}{m} \sqrt{(p_\mu \sigma^\mu)^2} \sqrt{(p_\nu \bar{\sigma}^\nu)^2} \xi = \sqrt{(p_\nu \sigma^\nu)^2} \xi \\ u_R = \frac{m}{m} \chi = \sqrt{p_\mu \bar{\sigma}^\mu} \xi \end{cases}$$

in this way the solution reads:

$$u(\vec{p}) = \begin{pmatrix} \sqrt{p_\mu \sigma^\mu} \\ \sqrt{p_\mu \bar{\sigma}^\mu} \end{pmatrix} \xi$$

Notice that, if we now suppose that the solution is of the form $\psi(x) = v(\vec{p})e^{ipx}$, going through the same steps as before, we would get that

$$v(\vec{p}) = \begin{pmatrix} \sqrt{p_\mu \sigma^\mu} \\ -\sqrt{p_\mu \bar{\sigma}^\mu} \end{pmatrix} \eta,$$

with η constant normalized spinor.
The first type of solutions are the so-called positive energy solutions, while the last ones are the negative energy solutions.

Let's study further those in the rest reference frame of the particle (where $p^\mu = (E,0,0,0)$), in this reference frame the positive energy solution assume the form:

$$\psi(x) = \sqrt{m} \begin{pmatrix} \xi \\ \xi \end{pmatrix} e^{-ip^\mu x_\mu}.$$

Using Lorentz transformations we can now obtain solutions for moving particles.

Lastly, notice that ξ and η really behave like the spin of a particle. If we consider a rotation, we can obtain such transformation using the generators of the rotations in the Lorentz group

$$S^{ij} = -\frac{i}{2}\varepsilon^{ijk} \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}, \text{with } \omega_{ij} = -\varepsilon_{ijk}\theta^k,$$

in this way the transformation of the spinor is given by

$$e^{\frac{\omega_{ij}}{2}S^{ij}} = \begin{pmatrix} e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}} & 0 \\ 0 & e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}} \end{pmatrix} \quad \Rightarrow \quad \xi' = e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}}\xi,$$

and this shows that ξ and η behave exactly as the spinors that describe spin $\frac{1}{2}$ particles, and indeed they do.

13.3 Quantization of the Dirac's field

In order to quantize the field, we need to derive two useful formulae: the first one is some sort of inner product between two spinors, which is simply a consequence of normalization conditions

$$\xi^{r\dagger}\xi^s = \eta^{r\dagger}\eta^s = \delta^{rs} \qquad r,s \in \{1,2\}.$$

We can easily prove, with this relation, that:

$$
\begin{aligned}
u^{r\dagger}u^s &= (\xi^{r\dagger}\sqrt{p_\mu\sigma^\mu},\ \xi^{r\dagger}\sqrt{p_\mu\bar{\sigma}^\mu})\begin{pmatrix}\sqrt{p_\mu\sigma^\mu}\xi^s\\ \sqrt{p_\mu\bar{\sigma}^\mu}\xi^s\end{pmatrix}\\
&= \xi^{r\dagger}(p_\mu\sigma^\mu)\xi^s + \xi^{r\dagger}(p_\mu\bar{\sigma}^\mu)\xi^s\\
&= \xi^{r\dagger}p_0\xi^s + \xi^{r\dagger}(p_i\sigma^i)\xi^s + \xi^{r\dagger}p_0\xi^s - \xi^{r\dagger}(p_i\sigma^i)\xi^s\\
&= 2p_0\delta^{rs}
\end{aligned}
$$

In the same way ...

$$
\begin{aligned}
v^{r\dagger}v^s &= 2p^0\delta^{rs}\\
u^{r\dagger}v^s &= v(\vec{p})^{r\dagger}v(-\vec{p})^s = u(\vec{p})^{r\dagger}u(-\vec{p})^s = 0.
\end{aligned}
$$

The second relation that we will use is a sort of outer product:

$$
\begin{aligned}
\sum_{s=1}^{2}u^s(\vec{p})\bar{u}^s(\vec{p}) &= \sum_{s=1}^{2}\begin{pmatrix}\sqrt{p_\mu\sigma^\mu}\xi^s\\ \sqrt{p_\mu\bar{\sigma}^\mu}\xi^s\end{pmatrix}(\xi^{r\dagger}\sqrt{p_\mu\sigma^\mu},\ \xi^{r\dagger}\sqrt{p_\mu\bar{\sigma}^\mu})\gamma^0\\
&= \sum_{s=1}^{2}\begin{pmatrix}\sqrt{p\sigma}\xi^s\xi^{s\dagger}\sqrt{p\bar{\sigma}} & \sqrt{p\sigma}\xi^s\xi^{s\dagger}\sqrt{p\sigma}\\ \sqrt{p\bar{\sigma}}\xi^s\xi^{s\dagger}\sqrt{p\bar{\sigma}} & \sqrt{p\bar{\sigma}}\xi^s\xi^{s\dagger}\sqrt{p\sigma}\end{pmatrix}\\
&= \begin{pmatrix}\sqrt{p\sigma p\bar{\sigma}} & \sqrt{p\sigma p\sigma}\\ \sqrt{p\bar{\sigma}p\bar{\sigma}} & \sqrt{p\bar{\sigma}p\sigma}\end{pmatrix} = \begin{pmatrix}m & p^\mu\sigma_\mu\\ p^\mu\bar{\sigma}_\mu & m\end{pmatrix} = \gamma^\mu p_\mu + m\mathbb{1}_{4\times 4}.
\end{aligned}
$$

Repeating the same calculation for the negative energy solution we get:

$$
\sum_{s=1}^{2}v^s(\vec{p})\bar{v}^s(\vec{p}) = \gamma^\mu p_\mu - m\mathbb{1}_{4\times 4}.
$$

13.3.1 The wrong quantization of the Dirac's field

We will now attempt to quantize Dirac's theory in the same manner as for the Klein-Gordon one, this procedure will lead us to the conclusion that a different approach is necessary in order to do so.

To start we will impose the canonical quantization conditions

$$
\begin{cases}[\hat{\psi}_\alpha(\vec{x}),\hat{\psi}_\beta(\vec{y})] = [\hat{\pi}_\alpha(\vec{x}),\hat{\pi}_\beta(\vec{y})] = 0\\ [\hat{\psi}_\alpha(\vec{x}),\hat{\pi}_\beta(\vec{y})] = i\delta^{\alpha\beta}\delta^3(\vec{x}-\vec{y})\end{cases},
$$

where the conjugate field π is given by the lagrangian (13.1) by $\frac{\partial\mathcal{L}}{\partial\dot{\psi}} = i\bar{\psi}\gamma^0 = i\psi^\dagger$.

We will now suppose that, for each solution, with positive and negative energy and different spin, there exist some creation and annihilation operators, such that the fields operators assume the following form:

$$
\begin{cases}\hat{\psi}(\vec{x}) = \sum_{s=1}^{2}\int\frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left(\hat{b}^s_{\vec{p}}u^s(\vec{p})e^{i\vec{p}\cdot\vec{x}} + \hat{c}^{s\dagger}_{\vec{p}}v^s(\vec{p})e^{-i\vec{p}\cdot\vec{x}}\right)\\ \hat{\pi}(\vec{x}) = i\sum_{s=1}^{2}\int\frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left(\hat{b}^{s\dagger}_{\vec{p}}u^{s\dagger}(\vec{p})e^{-i\vec{p}\cdot\vec{x}} + \hat{c}^s_{\vec{p}}v^\dagger(\vec{p})e^{i\vec{p}\cdot\vec{x}}\right)\end{cases},
$$

we can easily prove that, in order to get the canonical commutation relation above, we need to define these operators with the following commutators

$$
\begin{aligned}
&[\hat{b}^s_{\vec{p}}, \hat{b}^{r\dagger}_{\vec{q}}] = (2\pi)^3 \delta^{sr} \delta^3(\vec{p}-\vec{q}),\\
&[\hat{c}^s_{\vec{p}}, \hat{c}^{r\dagger}_{\vec{q}}] = -(2\pi)^3 \delta^{sr} \delta^3(\vec{p}-\vec{q}),\\
&[\hat{b}^s_{\vec{p}}, \hat{b}^{r}_{\vec{q}}] = [\hat{c}^s_{\vec{p}}, \hat{c}^{r}_{\vec{q}}] = [\hat{b}^s_{\vec{p}}, \hat{c}^{r}_{\vec{q}}] = [\hat{b}^{s\dagger}_{\vec{p}}, \hat{c}^{r}_{\vec{q}}] = [\hat{b}^s_{\vec{p}}, \hat{c}^{r\dagger}_{\vec{q}}] = 0.
\end{aligned}
$$

In fact using these relations we get:

$$
\begin{aligned}
[\hat{\psi}_\alpha(\vec{x}), \hat{\pi}_\beta(\vec{y})] = \sum_{r,s} \int \frac{d^3p\, d^3q}{(2\pi)^6} \frac{1}{2\sqrt{E_{\vec{p}E_{\vec{q}}}}} \Big(&[\hat{b}^s_{\vec{p}}, \hat{b}^{r\dagger}_{\vec{q}}] u^s(\vec{p}) u^{r\dagger}(\vec{q}) e^{i(\vec{p}\cdot\vec{x}-\vec{q}\cdot\vec{y})} + \\
&+ [\hat{c}^s_{\vec{p}}, \hat{c}^{r\dagger}_{\vec{q}}] v^s(\vec{p}) v^{r\dagger}(\vec{q}) e^{-i(\vec{p}\cdot\vec{x}-\vec{q}\cdot\vec{y})} \Big)\\
= \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} &\left(u^s(\vec{p}) u^{s\dagger}(\vec{p}) e^{i\vec{p}\cdot(\vec{x}-\vec{y})} v^s(\vec{p}) v^{s\dagger}(\vec{p}) e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} \right)\\
= \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} &\left((\gamma^\mu p_\mu + m)\gamma^0 e^{i\vec{p}\cdot(\vec{x}-\vec{y})} + (\gamma^\mu p_\mu - m)\gamma^0 e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} \right)\\
= \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{\gamma^0 p_0 \gamma^0}{2E_{\vec{p}}} &e^{i\vec{p}\cdot(\vec{x}-\vec{y})} = \delta^3(\vec{x}-\vec{y}),
\end{aligned}
$$

where we have firstly used the commutation relations above, after the integration over q-space we have used the outer product that we calculated in the previous section, and lastly we have changed the variables $\vec{p} \to -\vec{p}$ in the last term of the integrand, in order to cancels out all spacial and mass terms.

Notice how strange these commutation relations are: instead of the normal ones, that come from the harmonic oscillator, here we have a minus sign in the commutator of $\hat{c}$, we will now show that this sign leads to negative probabilities or negative energies.
In order to do so we will compute the hamiltonian operator for this system: starting from the Legendre's transform of (13.1)

$$
\mathcal{H} = \pi\dot{\psi} - \mathcal{L} = i\psi^\dagger\psi - i\bar{\psi}\gamma^\mu\partial_\mu\psi + m\bar{\psi}\psi = i\bar{\psi}(-\gamma^i\partial_i + m)\psi,
$$

we can get the classical hamiltonian by integrating over configuration space this one, then we have to substitute the fields operators, starting with

$$
\begin{aligned}
(-\gamma^i\partial_i+m)\hat{\psi} &= \sum_{s=1}^{2}\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left(\hat{b}^s_{\vec{p}}(-i\gamma^i\partial_i+m)u^s(\vec{p})e^{i\vec{p}\cdot\vec{x}}+\hat{c}^{s\dagger}_{\vec{p}}(-i\gamma^i\partial_i+m)v^s(\vec{p})e^{-i\vec{p}\cdot\vec{x}}\right)\\
&= \sum_{s=1}^{2}\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left(\hat{b}^s_{\vec{p}}(\gamma^i p^i+m)u^s(\vec{p})e^{i\vec{p}\cdot\vec{x}}+\hat{c}^{s\dagger}_{\vec{p}}(-\gamma^i p^i+m)v^s(\vec{p})e^{-i\vec{p}\cdot\vec{x}}\right)\\
&= \sum_{s=1}^{2}\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left(\hat{b}^s_{\vec{p}}(-\gamma^i p_i+m)u^s(\vec{p})e^{i\vec{p}\cdot\vec{x}}+\hat{c}^{s\dagger}_{\vec{p}}(\gamma^i p_i+m)v^s(\vec{p})e^{-i\vec{p}\cdot\vec{x}}\right)\\
&= \sum_{s=1}^{2}\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left(\hat{b}^s_{\vec{p}}\gamma^0 p_0 u^s(\vec{p})e^{i\vec{p}\cdot\vec{x}}-\hat{c}^{s\dagger}_{\vec{p}}\gamma^0 p_0 v^s(\vec{p})e^{-i\vec{p}\cdot\vec{x}}\right)\\
&= \sum_{s=1}^{2}\int \frac{d^3p}{(2\pi)^3}\sqrt{\frac{E_{\vec{p}}}{2}}\gamma^0\left(\hat{b}^s_{\vec{p}}u^s(\vec{p})e^{i\vec{p}\cdot\vec{x}}-\hat{c}^{s\dagger}_{\vec{p}}v^s(\vec{p})e^{-i\vec{p}\cdot\vec{x}}\right),
\end{aligned}
$$

then we have to evaluate

$$
\begin{aligned}
\hat{\mathcal{H}} &= \int d^3x\ \hat{\bar{\psi}}(-\gamma^i\partial_i+m)\hat{\psi}\\
&= \sum_{r,s}\int \frac{d^3x\, d^3p\, d^3q}{(2\pi)^6}\frac{1}{2}\sqrt{\frac{E_{\vec{p}}}{E_{\vec{q}}}}\left(\hat{b}^{r\dagger}_{\vec{q}}u^{r\dagger}(\vec{q})e^{-i\vec{q}\cdot\vec{x}}+\hat{c}^r_{\vec{q}}v^{r\dagger}(\vec{q})e^{i\vec{q}\cdot\vec{x}}\right)\left(\hat{b}^s_{\vec{p}}u^s(\vec{p})e^{i\vec{p}\cdot\vec{x}}-\hat{c}^{s\dagger}_{\vec{p}}v^s(\vec{p})e^{-i\vec{p}\cdot\vec{x}}\right)\\
&= \sum_{r,s}\int \frac{d^3p}{2(2\pi)^3}\left(\hat{b}^{r\dagger}_{\vec{p}}\hat{b}^s_{\vec{p}}u^{r\dagger}(\vec{p})u^s(\vec{p})-\hat{c}^r_{\vec{p}}\hat{c}^{s\dagger}_{\vec{p}}v^{r\dagger}(\vec{p})v^s(\vec{p})-\hat{b}^{r\dagger}_{-\vec{p}}\hat{c}^{s\dagger}_{\vec{p}}u^{r\dagger}(-\vec{p})v^s(\vec{p})+\hat{c}^r_{-\vec{p}}\hat{b}^s_{\vec{p}}v^{r\dagger}(-\vec{p})u^s(\vec{p})\right),
\end{aligned}
$$

recalling that $u^{r\dagger}(-\vec{p})v^s(\vec{p})=v^{r\dagger}(-\vec{p})u^s(\vec{p})=0$ and $u^{r\dagger}(\vec{p})u^s(\vec{p})=v^{r\dagger}(\vec{p})v^s(\vec{p})=2p_0\delta^{rs}$ we get:

$$
\hat{\mathcal{H}} = \sum_{s=1}^{2}\int \frac{d^3p}{(2\pi)^3}E_{\vec{p}}(\hat{b}^{s\dagger}_{\vec{p}}\hat{b}^s_{\vec{p}}-\hat{c}^s_{\vec{p}}\hat{c}^{s\dagger}_{\vec{p}}).
$$

Lastly, using the commutator of $\hat{c}$ we can get the final expression for the hamiltonian operator:

$$
\hat{\mathcal{H}} = \sum_{s=1}^{2}\int \frac{d^3p}{(2\pi)^3}E_{\vec{p}}(\hat{b}^{s\dagger}_{\vec{p}}\hat{b}^s_{\vec{p}}-\hat{c}^{s\dagger}_{\vec{p}}\hat{c}^s_{\vec{p}}+(2\pi)^3\delta^3(0)),
$$

again we have obtained a divergence term (as in Klein-Gordon hamiltonian) and some number operators, one for each type of particle and spin. Notice that the ones for the c type particles have a minus sign, which come from the strange sign of the commutator. We will now show that this doesn't lead to negative energies but to negative probabilities.

Let's compute the commutators

$$
\begin{aligned}
[\hat{\mathcal{H}}, \hat{b}^{s\dagger}_{\vec{p}}] &= \sum_{r=1}^{2} \int \frac{d^3q}{(2\pi)^3} E_{\vec{q}}\, [\hat{b}^{r\dagger}_{\vec{q}} \hat{b}^{r}_{\vec{q}}, \hat{b}^{s\dagger}_{\vec{p}}] = \sum_{r=1}^{2} \int \frac{d^3q}{(2\pi)^3} E_{\vec{q}} \left(\hat{b}^{r\dagger}_{\vec{q}} [\hat{b}^{r}_{\vec{q}}, \hat{b}^{s\dagger}_{\vec{p}}] + [\hat{b}^{r\dagger}_{\vec{q}}, \hat{b}^{s\dagger}_{\vec{p}}] \hat{b}^{r}_{\vec{q}} \right) \\
&= \sum_{r=1}^{2} \int \frac{d^3q}{(2\pi)^3} E_{\vec{q}}\, (2\pi)^3 \delta^{rs} \delta^3(\vec{q}-\vec{p}) \hat{b}^{r\dagger}_{\vec{q}} = E_{\vec{p}} \hat{b}^{s\dagger}_{\vec{p}}, \\
[\hat{\mathcal{H}}, \hat{c}^{s\dagger}_{\vec{p}}] &= \sum_{r=1}^{2} \int \frac{d^3q}{(2\pi)^3} E_{\vec{q}}\, [\hat{c}^{r\dagger}_{\vec{q}} \hat{c}^{r}_{\vec{q}}, \hat{c}^{s\dagger}_{\vec{p}}] = \sum_{r=1}^{2} \int \frac{d^3q}{(2\pi)^3} E_{\vec{q}} \left(\hat{c}^{r\dagger}_{\vec{q}} [\hat{c}^{r}_{\vec{q}}, \hat{c}^{s\dagger}_{\vec{p}}] + [\hat{c}^{r\dagger}_{\vec{q}}, \hat{c}^{s\dagger}_{\vec{p}}] \hat{c}^{r}_{\vec{q}} \right) \\
&= \sum_{r=1}^{2} \int \frac{d^3q}{(2\pi)^3} E_{\vec{q}}\, (2\pi)^3 \delta^{rs} \delta^3(\vec{q}-\vec{p}) \hat{c}^{r\dagger}_{\vec{q}} = E_{\vec{p}} \hat{c}^{s\dagger}_{\vec{p}},
\end{aligned}
$$

therefore we can clearly see that each state created by those operators have positive energy:

$$
\begin{aligned}
&\hat{\mathcal{H}} \hat{b}^{s\dagger}_{\vec{p}} |0\rangle = (E_{\vec{p}} \hat{b}^{s\dagger}_{\vec{p}} + \hat{b}^{s\dagger}_{\vec{p}} \hat{b}^{s}_{\vec{p}}) |0\rangle = E_{\vec{p}} \hat{b}^{s\dagger}_{\vec{p}} |0\rangle, \\
&\hat{\mathcal{H}} \hat{c}^{s\dagger}_{\vec{p}} |0\rangle = (E_{\vec{p}} \hat{c}^{s\dagger}_{\vec{p}} + \hat{c}^{s\dagger}_{\vec{p}} \hat{c}^{s}_{\vec{p}}) |0\rangle = E_{\vec{p}} \hat{c}^{s\dagger}_{\vec{p}} |0\rangle.
\end{aligned}
$$

However, as previously mentioned, the probability associated to a certain state can be negative:

$$
\langle 0 | \hat{c}^{s}_{\vec{p}} \hat{c}^{r\dagger}_{\vec{q}} |0\rangle = \langle 0 | \left([\hat{c}^{s}_{\vec{p}} \hat{c}^{r\dagger}_{\vec{q}}] + \hat{c}^{r\dagger}_{\vec{q}} \hat{c}^{s}_{\vec{p}} \right) |0\rangle = -(2\pi)^3 \delta^{rs} \delta^3(\vec{p}-\vec{q}).
$$

One could try to solve this issue by interpreting $\hat{c}^{\dagger}$ as an annihilation operator while $\hat{c}$ as the creator one: this doesn't solve the problem since it would mean that $\hat{c}^{s}_{\vec{p}} |0\rangle$ would have negative energy.
For these reasons this quantization approach fails.

13.3.2 The right quantization of the Dirac's field

In order to obtain the right Dirac's theory we will need to impose anticommutation relations, as the following

$$
\{\hat{\psi}_{\alpha}(\vec{x}), \hat{\psi}_{\beta}(\vec{y})\} = \delta_{\alpha\beta} \delta^3(\vec{x}-\vec{y}), \qquad \text{all the others vanish.}
$$

Imposing these relations on the previous field operators we obtain that creation and annihilation operators should obey

$$
\{\hat{b}^{r}_{\vec{p}}, \hat{b}^{s\dagger}_{\vec{q}}\} = \{\hat{c}^{r}_{\vec{p}}, \hat{c}^{s\dagger}_{\vec{q}}\} = (2\pi)^3 \delta^3(\vec{p}-\vec{q}).
$$

In this way the hamiltonian operator reads:

$$
\hat{\mathcal{H}} = \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} E_{\vec{p}} (\hat{b}^{s\dagger}_{\vec{p}} \hat{b}^{s}_{\vec{p}} - \hat{c}^{s}_{\vec{p}} \hat{c}^{s\dagger}_{\vec{p}}) = \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} E_{\vec{p}} (\hat{b}^{s\dagger}_{\vec{p}} \hat{b}^{s}_{\vec{p}} + \hat{c}^{s\dagger}_{\vec{p}} \hat{c}^{s}_{\vec{p}} - (2\pi)^3 \delta^3(0)),
$$

here we can notice that now the "strange" minus sign is disappeared, but the energy of vacuum is now negative: this behavior is the key element of why supersymmetry theories could be a solution to the infinite energy of vacuum.

If we now evaluate again the commutator of the hamiltonian, we obtain:

$$[\hat{\mathcal{H}}, \hat{b}^{s\dagger}_{\vec{p}}] = E_{\vec{p}} \hat{b}^{s\dagger}_{\vec{p}}, \qquad [\hat{\mathcal{H}}, \hat{c}^{s\dagger}_{\vec{p}}] = E_{\vec{p}} \hat{c}^{s\dagger}_{\vec{p}},$$

and, by the same steps as before, we can prove that there aren't negative energy states. Furthermore, the probability is conserved, as the norm of states is positive:

$$\langle 0| \hat{c}^{s}_{\vec{p}} \hat{c}^{r\dagger}_{\vec{q}} |0\rangle = \langle 0| \left(\{\hat{c}^{s}_{\vec{p}}, \hat{c}^{r\dagger}_{\vec{q}}\} - \hat{c}^{r\dagger}_{\vec{q}} \hat{c}^{s}_{\vec{p}} \right) |0\rangle = (2\pi)^3 \delta^{rs} \delta^3(\vec{p} - \vec{q}).$$

13.4 Conserved charge

Directly from the steps that led us to the lagrangian of the Dirac equation (recall that we wanted fermionic bilinears) we can easily obtain that the system that we have quantized has a $U(1)$ symmetry. Using the Nother's theorem we can obtain the conserved current associated with this symmetry: first let's consider an infinitesimal transformation

$$\psi' = e^{iq\theta} \psi \approxeq (1 + iq\theta)\psi, \qquad \bar{\psi}' = \bar{\psi} e^{-iq\theta} \approxeq \bar{\psi}(1 + -q\theta),$$

given the lagrangian $\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi$ the theorem gives:

$$J^\mu = \frac{\partial \mathcal{L}}{\partial \partial_\mu \psi} \delta \psi = -q\theta \bar{\psi} \gamma^\mu \psi.$$

Again we can rescale (to give the right units) the associated charge, which is

$$Q = \int d^3x\, J^0 = q \int d^3x\, \bar{\psi} \gamma^0 \psi.$$

We can that proceed to quantize this charge inserting the operators of the fields, thus:

$$\begin{aligned}
\hat{Q} &= \int d^3x \hat{\bar{\psi}} \gamma^0 \hat{\psi} = \hat{Q} = \int d^3x \hat{\psi}^\dagger \hat{\psi} \\
&= q \sum_{s,r=1}^{2} \int \frac{d^3p\, d^3q\, d^3x}{(2\pi)^6} \frac{1}{2\sqrt{E_{\vec{p}} E_{\vec{q}}}} \left(\hat{b}^{r\dagger}_{\vec{q}} u^{r\dagger}(\vec{q}) e^{-i\vec{q}\cdot\vec{x}} + \hat{c}^{r}_{\vec{q}} v^{s\dagger}(\vec{q}) e^{i\vec{q}\cdot\vec{x}} \right) \times \\
&\qquad\qquad \times \left(\hat{b}^{s}_{\vec{p}} u^{s}(\vec{p}) e^{i\vec{p}\cdot\vec{x}} + \hat{c}^{s\dagger}_{\vec{p}} v^{s}(\vec{p}) e^{-i\vec{p}\cdot\vec{x}} \right) \\
&= q \sum_{s,r=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} \left(\hat{b}^{r\dagger}_{\vec{p}} \hat{b}^{s}_{\vec{p}} u^{r\dagger}(\vec{p}) u^{s}(\vec{p}) + \hat{b}^{r\dagger}_{\vec{p}} \hat{c}^{s\dagger}_{-\vec{p}} u^{r\dagger}(\vec{p}) v^{s}(-\vec{p}) \right. \\
&\qquad\qquad \left. + \hat{c}^{r\dagger}_{-\vec{p}} \hat{b}^{s}_{\vec{p}} v^{r\dagger}(-\vec{p}) u^{s}(\vec{p}) + \hat{c}^{r}_{\vec{p}} \hat{c}^{s\dagger}_{\vec{p}} v^{r\dagger}(\vec{p}) v^{s}(\vec{p}) \right) \\
&= q \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \left(\hat{b}^{s\dagger}_{\vec{p}} \hat{b}^{s}_{\vec{p}} - \hat{c}^{s\dagger}_{\vec{p}} \hat{c}^{s}_{\vec{p}} \right).
\end{aligned}$$

Now, we can better understand the meaning of the two operators $\hat{b}^\dagger$ and $\hat{c}$, in fact the charge eigenvalue of a state created by $\hat{b}^\dagger$ will have a positive charge

$$
\begin{aligned}
\hat{Q}\hat{b}^{r\dagger}_{\vec{q}}\,|0\rangle &= q\sum_{s=1}^{2}\int\frac{d^3p}{(2\pi)^3}\left(\hat{b}^{s\dagger}_{\vec{p}}\hat{b}^{s}_{\vec{p}}-\hat{c}^{s\dagger}_{\vec{p}}\hat{c}^{s}_{\vec{p}}\right)\hat{b}^{r\dagger}_{\vec{q}}\,|0\rangle\\
&= q\sum_{s=1}^{2}\int\frac{d^3p}{(2\pi)^3}\left(\{\hat{b}^{s\dagger}_{\vec{p}}\hat{b}^{s}_{\vec{p}},\hat{b}^{r\dagger}_{\vec{q}}\}-\hat{b}^{s\dagger}_{\vec{p}}\hat{b}^{r\dagger}_{\vec{q}}\hat{b}^{s}_{\vec{p}}+\hat{c}^{s\dagger}_{\vec{p}}\hat{b}^{r\dagger}_{\vec{q}}\hat{c}^{s}_{\vec{p}}\right)|0\rangle\\
&= q\sum_{s=1}^{2}\int\frac{d^3p}{(2\pi)^3}\left(\hat{b}^{s\dagger}_{\vec{p}}\{\hat{b}^{s}_{\vec{p}},\hat{b}^{r\dagger}_{\vec{q}}\}+\{\hat{b}^{s\dagger}_{\vec{p}},\hat{b}^{r\dagger}_{\vec{q}}\}\hat{b}^{s}_{\vec{p}}\right)|0\rangle\\
&= q\sum_{s=1}^{2}\int\frac{d^3p}{(2\pi)^3}\hat{b}^{s\dagger}_{\vec{p}}(2\pi)^3\delta^{rs}\delta^3(\vec{p}-\vec{q})\,|0\rangle = q\hat{b}^{s\dagger}_{\vec{p}}\,|0\rangle\,,
\end{aligned}
$$

while the same calculation, made with $\hat{a}^\dagger$, would have a minus sign, and thus a negative charge. In this way we have showed that $\hat{b}^\dagger$ creates particles, while $\hat{c}^\dagger$ antiparticles.

13.5 Statistic of fermions and Pauli exclusion principle

We will now discuss how the statistics followed by fermionic particles (which obeys Dirac's equation) emerges naturally from this quantization procedure.
Let's consider a two particle system, if those are identical it can be obtained by using twice the same operator with different momenta and spins

$$
\hat{b}^{s_1\dagger}_{\vec{p}_1}\hat{b}^{s_2\dagger}_{\vec{p}_2}\,|0\rangle = |\vec{p}_1,s_1,\vec{p}_2,s_2\rangle\,.
$$

If we exchange the two particle we would obtain a state created by the same operators but in the reverse order, this can be expressed as the previous one using anticommutation relations

$$
|\vec{p}_2,s_2,\vec{p}_1,s_1\rangle = \hat{b}^{s_2\dagger}_{\vec{p}_2}\hat{b}^{s_1\dagger}_{\vec{p}_1}\,|0\rangle = -\hat{b}^{s_1\dagger}_{\vec{p}_1}\hat{b}^{s_2\dagger}_{\vec{p}_2}\,|0\rangle = -\,|\vec{p}_1,s_1,\vec{p}_2,s_2\rangle\,.
$$

In this way we have showed that identical particle states are antisymmetric under exchange of particles. This naturally leads to the **Pauli exclusion principle**, which states that two fermionic particles cannot be in the same state. In fact, if the two operators have the same momenta and spins it would mean that after the exchange of particles the system should be the same as before, but due to the antisymmetry it then should vanish

$$
|\vec{p},s,\vec{p},s\rangle = -\,|\vec{p},s,\vec{p},s\rangle \quad\Rightarrow\quad |\vec{p},s,\vec{p},s\rangle = 0.
$$

13.6 Propagator and time evolution of the Dirac's field

To end our discussion on the quantization of the Dirac's field we will analyze the time evolution of it. In order to do so we should go to the Heisenberg picture, in the same way we have done for the Klein-Gordon theory. Doing so it is easy to show that, in the Heisenberg picture, the field should have the following form

$$
\hat{\psi}(x) = \sum_{r=1}^{2}\int\frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left(\hat{b}^{s}_{\vec{p}}u^s(\vec{p})e^{-ip^\mu x_\mu}+\hat{c}^{s\dagger}_{\vec{p}}v^s(\vec{p})e^{ip^\mu x_\mu}\right).
$$

We then have to check that this theory is consistent with causality prescription given by relativity: to do so we define the propagator

$$iS_{\alpha\beta}(x-y) = \{\hat{\psi}_\alpha, \hat{\bar{\psi}}_\beta\}.$$

Let's compute this propagator in terms of the propagator of Klein-Gordon theory

$$\begin{aligned}
iS(x-y) &= \sum_{r,s=1}^{2} \int \frac{d^3p\, d^3q}{(2\pi)^6} \frac{1}{2\sqrt{E_{\vec{p}}E_{\vec{q}}}} \left[\{\hat{b}^s_{\vec{p}}, \hat{b}^{r\dagger}_{\vec{q}}\} u^s(\vec{p})\bar{u}^r(\vec{q}) e^{-i(px-qy)} + \{\hat{c}^{s\dagger}_{\vec{p}}, \hat{c}^r_{\vec{q}}\} v^s(\vec{p})\bar{v}^r(\vec{q}) e^{i(px-qy)} \right] \\
&= \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^6} \frac{1}{2E_{\vec{p}}} \left[u^s(\vec{p})\bar{u}^s(\vec{p}) e^{-ip(x-y)} + v^s(\vec{p})\bar{v}^s(\vec{p}) e^{ip(x-y)} \right] \\
&= \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^6} \frac{1}{2E_{\vec{p}}} \left[(\gamma^\mu p_\mu + m) e^{-ip(x-y)} + (\gamma^\mu p_\mu - m) e^{ip(x-y)} \right] \\
&= (i\gamma^\mu \partial_\mu + m)(D(x-y) - D(y-x)), \qquad \text{with} \qquad D(x-y) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} e^{-ip_\mu(x^\mu - y^\mu)}.
\end{aligned}$$

In the discussion of the have already shown that $D(x-y) - D(y-x)$ vanishes outside the light cone, and thus it does the propagator for Dirac's field too. Now, since all the observables, for fermionic particles, are bilinears, the anticommutation of the fields are enough to make commute the observable, thus granting the causality of the theory.

Lastly, observe that the propagator satisfy the Dirac's equation itself

$$\begin{aligned}
&(i\gamma^\mu \partial^{(x)}_\mu - m) S(x-y) = 0 \\
&\frac{1}{i}(i\gamma^\nu \partial^{(x)}_\nu - m)(i\gamma^\mu \partial_\mu + m)(D(x-y) - D(y-x)) \\
&= -\frac{1}{i}(\gamma^\nu \partial^{(x)}_\nu \gamma^\mu \partial^{(x)}_\mu + m^2)(D(x-y) - D(y-x)) \\
&= i\left(\frac{1}{2}\{\gamma^\nu \gamma^\mu\} \partial^{(x)}_\nu \partial^{(x)}_\mu + m^2\right)(D(x-y) - D(y-x)) \\
&= i(\eta^{\mu\nu} \partial^{(x)}_\nu \partial^{(x)}_\mu + m^2)(D(x-y) - D(y-x)) \\
&= i(\Box^{(x)} + m^2)(D(x-y) - D(y-x)) \\
&= i \int \frac{d^3p}{(2\pi)^3} \left[(-p^\mu p_\mu + m^2) e^{-ip(x-i)} - (-p^\mu p_\mu + m^2) e^{ip(x-i)} \right] = 0,
\end{aligned}$$

where in the last line we have used the mass-shell condition.

14. Quantization of the EM field

14.1 The electromagnetic field

We are going to discuss the quantization of the electromagnetic field, in order to do so we will need to understand first how classical field theory describe such field. From electromagnetism, we know that we can describe all electromagnetic interactions using two vector fields, which are solutions of the Maxwell's equations. These two fields can be combined in a covariant tensor, called **electromagnetic tensor**

$$F_{\mu\nu}\begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}.$$

Introducing the 4-potential $A^\mu = (\phi, \vec{A})$, as the 4-vector containing the electric potential ϕ and the magnetic vector potential $\vec{A}$, from the classical definitions of these two we obtain

$$\begin{cases} \vec{E} = -\vec{\nabla}\phi - \partial_t \vec{A}, \\ \vec{B} = \vec{\nabla}\times\vec{A} \end{cases} \qquad \Rightarrow \qquad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

Given all of these quantities, we can now guess that the lagrangian of the electromagnetic field is a scalar quantity such as $F^{\mu\nu}F_{\mu\nu}$ (actually there exist only two possible scalars, but the other one would vanish in the action). We can test our guess finding the equations of motion and comparing them to the Maxwell's equations:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} \qquad \Rightarrow \qquad 0 = \frac{\partial\mathcal{L}}{\partial\partial_\mu A_\nu} = -\partial_\mu F^{\mu\nu},$$

from this one equation we can obtain two Maxwell's equations

$$\vec{\nabla}\cdot\vec{E}=0, \qquad \partial_t\vec{E}=\vec{\nabla}\times\vec{B},$$

the others can be obtained from the Bianchi's identity (which the electromagnetic tensor satisfies)

$$\partial_\lambda F^{\mu\nu}+\partial_\mu F^{\nu\lambda}+\partial_\nu F^{\lambda\mu}=0 \quad \Rightarrow \begin{cases}\vec{\nabla}\cdot\vec{B}=0,\\ \partial_t\vec{B}=-\vec{\nabla}\times\vec{E}\end{cases}.$$

Alternatively, we could describe the system using the **dual field strength**

$$\tilde{F}^{\mu\nu}=\varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma},$$

this one will introduce the so called *E-M duality*, swapping the roles of the electric and magnetic fields.

14.1.1 Gauge symmetries

Notice that our system is described by 6 scalar fields, but we have already introduced the 4-potential and defined all of those in terms of the components A^μ, thus 2 of the 6 d.o.f are actually not needed to describe our system. It turns out the d.o.f can be reduced even more, in fact we can observe that

$$F^{00}=\partial_0A_0-\partial_0A_0=0,$$

this implies that in the lagrangian there isn't any term containing A_0 or its derivatives and thus its isn't a physical degree of freedom of the system.

This particular structure can be used to add more constraints on the system, and thus reduce even more the degrees of freedom of the field. In fact, we can add the 4-gradient of an arbitrary function to the for potential and the electromagnetic tensor wouldn't change, and thus the lagrangian too

$$A^\mu\to A'^\mu=A^\mu+\partial_\mu\lambda \;\Rightarrow\; F'^{\mu\nu}=\partial_\mu A_\mu-\partial_\mu\partial_\nu\lambda-\partial_\nu A^\mu+\partial_\mu\partial_\nu\lambda=F^{\mu\nu}.$$

This symmetry is called **Gauge symmetry**, and it is an internal local symmetry, and let us impose others constraints con the equations of motion called gauges:

- the **Lorenz gauge** $\partial_\mu A^\mu=0$,
- the **Coulomb gauge** $\vec{\nabla}\cdot\vec{A}=0$.

We should notice that the Lorentz gauge is Lorentz invariant and for this reason it is the one that we will use during quantization.

14.2 Quantization of the EM field

We will now quantize the system, thus promote to operators the conjugate field and the hamiltonian, therefore we need to obtain the classical form of those

$$\pi^\mu = \frac{\partial\mathcal{L}}{\partial\dot{A}_\mu} = \begin{cases} \pi^0 = 0 & \text{Because } \mathcal{L} \text{ doesn't contain } \dot{A}_0, \\ \pi^i = \frac{\partial}{\partial\dot{A}_i}\left[-\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu)\right] \end{cases}$$
$$= \begin{cases} \pi^0 = 0 & \text{Because } \mathcal{L} \text{ doesn't contain } \dot{A}_0, \\ \pi^i = -\frac{1}{2}[\partial^0 A^i - \partial^i A^0 - \partial^i A^0 + \partial^0 A^i] = -\dot{A}^i + \partial^i A^0 = E^i \end{cases},$$
$$\mathcal{H} = \int d^3x(\pi^\mu \dot{A}_\mu - \mathcal{L})$$
$$= \int d^3x(E^i(-\partial_i A_0 - E_i) - \mathcal{L}) = \int d^3x\left(\vec{E}^2 + \vec{E}\cdot\vec{\nabla}A_0 - \frac{\vec{E}^2 - \vec{B}^2}{2}\right)$$
$$= \int d^3x\left(\vec{E}\cdot\vec{\nabla}A_0 + \frac{\vec{E}^2 + \vec{B}^2}{2}\right) = \int d^3x\left(-A_0\vec{\nabla}\cdot\vec{E} + \frac{\vec{E}^2 + \vec{B}^2}{2}\right),$$

where we have used that $E_i = -\partial_i A_0 - \partial_0 A^i$ and $\mathcal{L} = \frac{\vec{E}^2 - \vec{B}^2}{2}$.
We should now notice that here the fact that A_0 is non-physical is manifest, since it could be interpreted as a sort lagrange multiplier imposing the constraint $\nabla\cdot\vec{E} = 0$.

Lastly, before quantizing, we should impose the Lorenz gauge $\partial_\mu A^\mu = 0$, in this way the equations of motion reduces to a wave equation

$$\partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) = \Box A^\nu - \partial^\nu(\partial_\mu A^\mu) = \Box A^\nu = 0,$$

this help us a lot in the process of quantization, since this is actually a Klein-Gordon equation with mass zero, and thus we can quantize each component of the field as a Klein-Gordon field.
To impose this gauge we can add an appropriate (total derivative) term to the lagrangian

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2}(\partial_\mu A^\mu)^2,$$

in this way the equations of motion reads

$$0 = -\partial_\mu F^{\mu\nu} - \partial_\nu \eta^{\mu\nu}(\partial_\sigma A^\sigma) = -\partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu + \partial^\nu A^\mu) = \Box A^\nu.$$

In this way also the conjugate filed changes, resulting in

$$\pi^\mu = \partial^\mu A^0 - \partial^0 A^\mu - \delta_0^\nu \partial_\nu A^\nu.$$

We now need to impose commutation relations on the fields operators, such as in Klein-Gordon

$$[\hat{A}_\mu(\vec{x}), \hat{A}_\nu(\vec{y})] = [\hat{\pi}_\mu(\vec{x}), \hat{\pi}_\nu(\vec{y})] = 0, \qquad [\hat{A}_\mu(\vec{x}), \hat{\pi}_\nu(\vec{y})] = i\eta_{\mu\nu}\delta^3(\vec{x}-\vec{y}),$$

these relations are verified defining the following field

$$\hat{A}_\mu(\vec{x}) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2|\vec{p}|}}(\hat{\xi}_\mu(\vec{p})e^{i\vec{p}\cdot x} + \hat{\xi}_\mu^\dagger(\vec{p})e^{-i\vec{p}\cdot x}).$$

The expression of this field can be simplified by expanding $\hat{\vec{\xi}}$, which are some vectors of operators, on an orthonormal basis

$$\hat{A}_\mu(\vec{x}) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2|\vec{p}|}}\sum_{\lambda=0}^{3}\varepsilon_\mu^{(\lambda)}(\vec{p})(\hat{a}_{\vec{p}}^{(\lambda)}e^{i\vec{p}\cdot x}+\hat{a}_{\vec{p}}^{(\lambda)\dagger}e^{-i\vec{p}\cdot x}),$$

and the conjugate field operator

$$\hat{\pi}_\mu(\vec{x}) = i\int \frac{d^3p}{(2\pi)^3}\sqrt{\frac{|\vec{p}|}{2}}\sum_{\lambda=0}^{3}\varepsilon_\mu^{(\lambda)}(\vec{p})(\hat{a}_{\vec{p}}^{(\lambda)}e^{i\vec{p}\cdot x}-\hat{a}_{\vec{p}}^{(\lambda)\dagger}e^{-i\vec{p}\cdot x}).$$

We now check that this sort of ansatz is correct utilizing the canonical commutation relation for the creation annihilation operators

$$\begin{cases}[\hat{a}_{\vec{p}}^{(\lambda)},\hat{a}_{\vec{q}}^{(\lambda')}] = [\hat{a}_{\vec{p}}^{(\lambda)\dagger},\hat{a}_{\vec{q}}^{(\lambda')\dagger}] = 0,\\ [\hat{a}_{\vec{p}}^{(\lambda)},\hat{a}_{\vec{q}}^{(\lambda')\dagger}] = -\eta^{\lambda\lambda'}(2\pi)^3\delta^3(\vec{p}-\vec{q})\end{cases}$$

$$[\hat{A}_\mu(\vec{x}),\hat{\pi}_\nu(\vec{y})] = i\int\frac{d^3p\,d^3q}{2(2\pi)^6}\sqrt{\frac{|\vec{q}|}{|\vec{p}|}}\sum_{\lambda,\lambda'=0}^{3}\varepsilon_\mu^{(\lambda)}(\vec{p})\varepsilon_\mu^{(\lambda')}(\vec{p})\Big(-[\hat{a}_{\vec{p}}^{(\lambda)},\hat{a}_{\vec{q}}^{(\lambda')\dagger}]e^{i(\vec{p}\cdot\vec{x}\vec{q}\cdot\vec{y})}+[\hat{a}_{\vec{p}}^{(\lambda)\dagger},\hat{a}_{\vec{q}}^{(\lambda')}]e^{-i(\vec{p}\cdot\vec{x}\vec{q}\cdot\vec{y})}\Big)$$

$$i\int\frac{d^3p\,d^3q}{2(2\pi)^6}\sqrt{\frac{|\vec{q}|}{|\vec{p}|}}\sum_{\lambda,\lambda'=0}^{3}\varepsilon_\mu^{(\lambda)}(\vec{p})\varepsilon_\mu^{(\lambda')}(\vec{p})\Big(\eta^{\lambda\lambda'}\delta^3(\vec{p}-\vec{q})e^{i(\vec{p}\cdot\vec{x}\vec{q}\cdot\vec{y})}+\eta^{\lambda\lambda'}\delta^3(\vec{p}-\vec{q})e^{-i(\vec{p}\cdot\vec{x}\vec{q}\cdot\vec{y})}\Big)$$

$$= i\int\frac{d^3p}{(2\pi)^3}\sum_{\lambda,\lambda'=0}^{3}\varepsilon_\mu^{\lambda}(\vec{p})\varepsilon_\nu^{\lambda'}(\vec{p})\eta^{\lambda,\lambda'}e^{ip(\vec{x}-\vec{y})} = i\sum_{\lambda,\lambda'=0}^{3}\varepsilon_\mu^{\lambda}(\vec{p})\varepsilon_\nu^{\lambda'}(\vec{p})\eta^{\lambda,\lambda'}\delta^3(\vec{p}-\vec{q})$$

$$= i\eta_{\mu\nu}\delta^3(\vec{p}-\vec{q}),$$

where we used the orthonormality condition $\sum_{\lambda,\lambda'=0}^{3}\varepsilon_\mu^{\lambda}(\vec{p})\varepsilon_\nu^{\lambda'}(\vec{p})\eta^{\lambda,\lambda'} = \eta_{\mu\nu}$.

Notice that these commutation relations leads to a strange, and usually unwanted, behavior of the states: $\hat{s}_{\vec{p}}^{(0)\dagger}$ creates negative norm photons, called **ghosts**

$$\langle 0|\,\hat{a}_{\vec{q}}^{(0)}\hat{a}_{\vec{p}}^{(0)\dagger}\,|0\rangle = -\eta^{00}(2\pi)^3\delta^3(\vec{p}-\vec{q}).$$

To remove this strange behavior we will study how to impose the Lorenz gauge to the quantized field.

14.3 The Lorenz gauge for the quantized field

Imposing the Lorenz gauge to the quantized field is not a trivial task since now the 4-potential is an operator ant thus it is not clear anymore how to translate $\partial_\mu A^\mu = 0$.
This can be thought, at a first glance, as the condition

$$\partial_\mu\hat{A}^\mu = 0,$$

but some further inspections show that this cannot be the right condition since in this way the first component of the momentum field $\pi^0 = -\partial_\mu\hat{A}^\mu$ should be identically equals to zero. In this way we would contradict the canonical commutation relations.

A second approach is to impose the gauge condition on the states of the Fock's space: for example we could ask that

$$\partial_\mu \hat{A}^\mu \left|\psi\right\rangle = 0, \qquad \forall \left|\psi\right\rangle / \ \left\langle\psi|\psi\right\rangle \geq 0,$$

however, in this way, we mess up the vacuum state. In fact, $\hat{A}^\mu(\vec{x})$ contains both creations and annihilation operators and thus its 4-divergence too, it is clear though that this operator won't result in the null state. We now go to the Heisenberg picture and split the filed operator in

$$\hat{A}^{\mu+}(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2p}} \sum_{\lambda=0}^{3} \varepsilon_\mu^{(\lambda)} \hat{a}_{\vec{p}}^{(\lambda)} e^{-ip_\nu x^\nu},$$
$$\hat{A}^{\mu-}(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2p}} \sum_{\lambda=0}^{3} \varepsilon_\mu^{(\lambda)} \hat{a}_{\vec{p}}^{(\lambda)\dagger} e^{ip_\nu x^\nu},$$
$$\hat{A}^\mu(x) = \hat{A}^{\mu+}(x) + \hat{A}^{\mu-}(x),$$

in this way it is easy to see that, containing only annihilator operators, $\hat{A}^{\mu+}(x)$ will vanish acting on the vacuum state, while this is impossible for $\hat{A}^{\mu-}(x)$ to vanish in this way.
In order to preserve how the vacuum state behaves we will ask the stricter condition $\partial_\mu \hat{A}^{\mu+}(x) \left|\psi\right\rangle = 0 \ \forall \left|\psi\right\rangle$. Actually only apparently this is stricter, since

$$\partial_\mu \hat{A}^{\mu+}(x) \left|\psi\right\rangle = 0 \iff \left\langle\psi\right| \partial_\mu \hat{A}^{\mu-}(x) = 0,$$

since $(\partial_\mu \hat{A}^{\mu+}(x))^\dagger = \partial_\mu \hat{A}^{\mu-}(x)$, and thus we obtain the **Gupta-Bleuler condition**:

$$\left\langle\psi\right| \partial_\mu \hat{A}^\mu(x) \left|\psi\right\rangle = 0, \qquad \forall \left|\psi\right\rangle, \tag{14.1}$$

which is the general condition that the trace of this operator is zero.

We now want to study what implies this condition: first let's consider a field describing a classical EM wave propagating in the z direction, thus its 4-momentum is $p^\mu = (E,0,0,E)$. The *"plus"* field operator, once differentiated, becomes

$$\partial_\mu \hat{A}^{\mu+}(x) = -i \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2p}} \sum_{\lambda=0}^{3} \varepsilon_\mu^{(\lambda)} p^\mu \hat{a}_{\vec{p}}^{(\lambda)} e^{-ip_\nu x^\nu},$$

the gauge condition we imposed, choosing ε to be the canonical basis, now reads

$$E(\varepsilon_0^{(0)} \hat{a}_{\vec{p}}^{(0)} + \varepsilon_3^{(3)} \hat{a}_{\vec{p}}^{(3)}) \left|\psi\right\rangle = E(\hat{a}_{\vec{p}}^{(0)} - \hat{a}_{\vec{p}}^{(3)}) \left|\psi\right\rangle = 0, \quad \forall \left|\psi\right\rangle,$$

this can be used to obtain the following relation:

$$\left\langle\psi\right| \hat{a}_{\vec{p}}^{(0)\dagger} \hat{a}_{\vec{p}}^{(0)} \left|\psi\right\rangle = \left\langle\psi\right| \hat{a}_{\vec{p}}^{(3)\dagger} \hat{a}_{\vec{p}}^{(3)} \left|\psi\right\rangle,$$

which implies that the number of time-like and longitudinal photons of a state is always the same.

A second consequence of the Gupta-Bleuler condition 14.1 is that now the Fock's space has zero norm states. To see this we can introduce the following operators

$$\hat{b}_{\pm,\vec{p}} = \hat{a}^{(0)}_{\vec{p}} \pm \hat{a}^{(3)}_{\vec{p}}$$

which creates or annihilates states given by a linear combination of time-like and longitudinal photons, since we want to have states where the number of those photons is the same. We have now to deduce the commutation relations of these operators

$$\begin{aligned}
[\hat{b}_{\pm,\vec{p}}, \hat{b}^{\dagger}_{\pm,\vec{q}}] &= [\hat{a}^{(0)}_{\vec{p}}, \hat{a}^{(0)\dagger}_{\vec{q}}] + [\hat{a}^{(3)}_{\vec{p}}, \hat{a}^{(3)\dagger}_{\vec{q}}] = (-(2\pi)^3 + (2\pi)^3)\delta^3(\vec{p}-\vec{q}) = 0,\\
[\hat{b}_{\pm,\vec{p}}, \hat{b}^{\dagger}_{\mp,\vec{q}}] &= [\hat{a}^{(0)}_{\vec{p}}, \hat{a}^{(0)\dagger}_{\vec{q}}] - [\hat{a}^{(3)}_{\vec{p}}, \hat{a}^{(3)\dagger}_{\vec{q}}] = -2(2\pi)^3\delta^3(\vec{p}-\vec{q}),
\end{aligned}$$

in this way we can compute the norm of the states created by these two operators

$$\langle 0|\, \hat{b}_{\pm,\vec{p}} \hat{b}^{\dagger}_{\pm,\vec{p}}\, |0\rangle = \langle 0|\, [\hat{b}_{\pm,\vec{p}}, \hat{b}^{\dagger}_{\pm,\vec{p}}]\, |0\rangle = 0,$$

which is zero.

This could seem strange but it is actually intended to be like this: in fact we will use the Gupta-Bleuler condition to distinguish between physical and nonphysical states (those who are in our Fock's space but should be removed to obtain the 2 d.o.f of EM field)

$$\hat{b}_{-,\vec{p}}\, |\psi\rangle = 0,$$

those sates which satisfy this relation are physical, while the others must be discarded. This request is enough to remove longitudinal and time-like photons, since

$$\begin{aligned}
\hat{b}_{-,\vec{q}} \hat{a}^{(0)\dagger}_{\vec{p}}\, |0\rangle &= \frac{1}{2}\hat{b}_{-,\vec{q}}(\hat{b}^{\dagger}_{+,\vec{p}} + \hat{b}^{\dagger}_{-,\vec{p}})\, |0\rangle = \frac{1}{2}[\hat{b}_{-,\vec{q}}, \hat{b}^{\dagger}_{+,\vec{p}}]\, |0\rangle \neq 0,\\
\hat{b}_{-,\vec{q}} \hat{a}^{(3)\dagger}_{\vec{p}}\, |0\rangle &= \frac{1}{2}\hat{b}_{-,\vec{q}}(\hat{b}^{\dagger}_{+,\vec{p}} - \hat{b}^{\dagger}_{-,\vec{p}})\, |0\rangle = \frac{1}{2}[\hat{b}_{-,\vec{q}}, \hat{b}^{\dagger}_{+,\vec{p}}]\, |0\rangle \neq 0,
\end{aligned}$$

and also the states created by $\hat{b}^{\dagger}_{+,\vec{p}}$

$$\hat{b}_{-,\vec{q}} \hat{b}^{\dagger}_{+,\vec{p}}\, |0\rangle = [\hat{b}_{-,\vec{q}}, \hat{b}^{\dagger}_{+,\vec{p}}]\, |0\rangle \neq 0,$$

but the same doesn't happen for those created by $\hat{b}^{\dagger}_{-,\vec{p}}$

$$\hat{b}_{-,\vec{q}} \hat{b}^{\dagger}_{-,\vec{p}}\, |0\rangle = [\hat{b}_{-,\vec{q}}, \hat{b}^{\dagger}_{-,\vec{p}}]\, |0\rangle = 0,$$

but their norm is zero.
Lastly we should check the the transversal photons are physical, as they are premitted in the classical theory

$$\hat{b}_{-,\vec{q}} \hat{a}^{(1,2)\dagger}_{\vec{p}}\, |0\rangle = \hat{a}^{(1,2)\dagger}_{\vec{p}} \hat{b}_{-,\vec{q}}\, |0\rangle = 0,$$

thus they satisfy Gupta-Bleuler condition and they have positive norm.

14.4 The hamiltonian of the electromagnetic field

To end the quantization process of the field we need to quantize energy, and thus obtain the hamiltonian operator.
First we have to calculate the classical hamiltonian of the system

$$\begin{aligned}
\mathcal{L} &= -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2}\partial_\mu A^\mu \partial_\nu A^\nu \\
&= -\frac{1}{4}(\partial^\mu A^\nu - \partial^\nu A^\mu)(\partial_\mu A_\nu - \partial_\nu A_\mu) - \frac{1}{2}\partial_\mu A^\mu \partial_\nu A^\nu \\
&= -\frac{1}{2}(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu) - \frac{1}{2}\partial_\mu A^\mu \partial_\nu A^\nu \\
&\text{Integrating by parts} \\
&= -\frac{1}{2}\partial^\mu A^\nu \partial_\mu A_\nu - \frac{1}{2}(A^\nu \partial^\mu \partial_\nu A_\mu - A^\nu \partial^\nu \partial_\mu A_\mu) \\
&= -\frac{1}{2}\partial^\mu A^\nu \partial_\mu A_\nu, \\
\pi^\mu &= \frac{\partial \mathcal{L}}{\partial \dot{A}_\mu} = -\dot{A}^\mu, \\
\pi^\mu \dot{A}_\mu - \mathcal{L} &= -\pi^\mu \pi_\mu + \frac{1}{2}\dot{A}^\mu \dot{A}_\mu + \frac{1}{2}\partial^i A^\mu \partial_i A_\mu = \frac{1}{2}(\partial^i A^\mu \partial_i A_\mu - \pi^\mu \pi_\mu), \\
\mathcal{H} &= \int d^3x \, \frac{1}{2}(\partial^i A^\mu \partial_i A_\mu - \pi^\mu \pi_\mu).
\end{aligned}$$

We now quantize the hamiltonian by plugging the field operators in Schrödinger picture: evaluating each pice and the summing all up:

$$\begin{aligned}
&\int d^3x \, \partial^i \hat{A}^\mu \partial_i \hat{A}_\mu = \\
&= \int \frac{d^3p \, d^3q \, d^3x}{(2\pi)^6} \frac{\vec{p}\cdot\vec{q}}{2\sqrt{|\vec{p}||\vec{q}|}} \sum_{\lambda,\lambda'=0}^{3} \varepsilon^{\mu(\lambda)} \varepsilon^{(\lambda')}_\mu (\hat{a}^{(\lambda)}_{\vec{p}} e^{i\vec{p}\cdot\vec{x}} - \hat{a}^{(\lambda)\dagger}_{\vec{p}} e^{-i\vec{p}\cdot\vec{x}}) \times \\
&\qquad \times (\hat{a}^{(\lambda')}_{\vec{q}} e^{i\vec{q}\cdot\vec{x}} - \hat{a}^{(\lambda')\dagger}_{\vec{q}} e^{-i\vec{q}\cdot\vec{x}}) \\
&= \int \frac{d^3p}{(2\pi)^3} \frac{|\vec{p}|}{2} \sum_{\lambda,\lambda'=0}^{3} \eta^{\lambda\lambda'} (-\hat{a}^{(\lambda)}_{\vec{p}} \hat{a}^{(\lambda)}_{-\vec{p}} - \hat{a}^{(\lambda)}_{\vec{p}} \hat{a}^{(\lambda')\dagger}_{\vec{p}} - \hat{a}^{(\lambda)\dagger}_{\vec{p}} \hat{a}^{(\lambda')}_{\vec{p}} - \hat{a}^{(\lambda)\dagger}_{\vec{p}} \hat{a}^{(\lambda')\dagger}_{-\vec{p}}) \\
&\int d^3x \, \hat{\pi}^\mu \hat{\pi}_\mu = \\
&= -\int \frac{d^3p \, d^3q \, d^3x}{(2\pi)^6} \sqrt{\frac{|\vec{p}||\vec{q}|}{4}} \sum_{\lambda,\lambda'=0}^{3} \varepsilon^{\mu(\lambda)} \varepsilon^{(\lambda')}_\mu (\hat{a}^{(\lambda)}_{\vec{p}} e^{i\vec{p}\cdot\vec{x}} - \hat{a}^{(\lambda)\dagger}_{\vec{p}} e^{-i\vec{p}\cdot\vec{x}}) \times \\
&\qquad \times (\hat{a}^{(\lambda')}_{\vec{q}} e^{i\vec{q}\cdot\vec{x}} - \hat{a}^{(\lambda')\dagger}_{\vec{q}} e^{-i\vec{q}\cdot\vec{x}}) \\
&= -\int \frac{d^3p}{(2\pi)^3} \frac{|\vec{p}|}{2} \sum_{\lambda,\lambda'=0}^{3} \eta^{\lambda\lambda'} (\hat{a}^{(\lambda)}_{\vec{p}} \hat{a}^{(\lambda)}_{-\vec{p}} - \hat{a}^{(\lambda)}_{\vec{p}} \hat{a}^{(\lambda')\dagger}_{\vec{p}} - \hat{a}^{(\lambda)\dagger}_{\vec{p}} \hat{a}^{(\lambda')}_{\vec{p}} + \hat{a}^{(\lambda)\dagger}_{\vec{p}} \hat{a}^{(\lambda')\dagger}_{-\vec{p}})
\end{aligned}$$

thus the hamiltonian operator is

$$\begin{aligned}
\hat{\mathcal{H}} &= \int d^3x\, \frac{1}{2}(\partial^i \hat{A}^\mu \partial_i \hat{A}_\mu - \hat{\pi}^\mu \hat{\pi}_\mu) \\
&= -\frac{1}{2}\int \frac{d^3p}{(2\pi)^3}|\vec{p}| \sum_{\lambda,\lambda'=0}^{3} \eta^{\lambda\lambda'}(\hat{a}^{(\lambda)}_{\vec{p}}\hat{a}^{(\lambda')\dagger}_{\vec{p}} + \hat{a}^{(\lambda)\dagger}_{\vec{p}}\hat{a}^{(\lambda')}_{\vec{p}}) \\
&\quad \text{Using normal ordering} \\
&= -\int \frac{d^3p}{(2\pi)^3}|\vec{p}|(\hat{a}^{(0)}_{\vec{p}}\hat{a}^{(0)\dagger}_{\vec{p}} - \hat{a}^{(1)}_{\vec{p}}\hat{a}^{(1)\dagger}_{\vec{p}} - \hat{a}^{(2)}_{\vec{p}}\hat{a}^{(2)\dagger}_{\vec{p}} - \hat{a}^{(3)}_{\vec{p}}\hat{a}^{(3)\dagger}_{\vec{p}}) \\
&= \int \frac{d^3p}{(2\pi)^3}|\vec{p}|\left[\sum_{i=1}^{2}\hat{a}^{(i)\dagger}_{\vec{p}}\hat{a}^{(i)}_{\vec{p}} - (\hat{a}^{(0)}_{\vec{p}}\hat{a}^{(0)\dagger}_{\vec{p}} - \hat{a}^{(3)}_{\vec{p}}\hat{a}^{(3)\dagger}_{\vec{p}})\right].
\end{aligned}$$

Clearly we have obtained number operators of each type of photon multiplied by its own momentum, which is its energy; notice that we have obtained a term in which we have the difference between the time-like and transversal photon, however the Gupta-Bleuler condition (14.1) requires that those two numbers are the same, thus the presence of those photons would not be seen by the hamiltonian operator.