Notes on

INTRODUCTION TO QUANTUM FIELD THEORY

Roberto Bonciani[1]

Dipartimento di Fisica, Università di Roma "La Sapienza"
e INFN Sezione di Roma,
Piazzale Aldo Moro 2,
00185 Roma

Academic Years 2021-2022, 2022/2023

[1]Email: roberto.bonciani@roma1.infn.it

Indice

Capitolo 1

Necessity of a Theory of Fields

1.1 Introduction

Non-relativistic Quantum Mechanics (NRQM), developed from the beginning of last century untill $\sim$ 1926 is a theory devoted to the study of a single particle. To the particle is associated a wave function, $\psi(x,t)$, whose time evolution is determined by the wave equation:

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{x},t) = H\,\psi(\mathbf{x},t)\,, \tag{1.1}$$

that, in the case in which $H = \frac{p^2}{2m} + V$, represents the well known non relativistic Schrödinger's equation for a particle moving in a potential V. The modulus squared of the wave function, $|\psi(x,t)|^2$, is interpreted as the probability density of finding the particle in x at the time t. Such a Theory leaves the concept of classic determinism in favor of a treatment of the microscopic physical phenomena that is intrinsically statistical. However, this theory is not yet completely satisfactory, since it is not so general to include the possibility that the particle's speed is close to the speed of light. In other words, it does not include Special Relativity and it is therefore valid for velocities much smaller than c. Another crucial point is that, in NRQM it is possible to study only transition amplitudes that do not involve a different number of particles in the initial and final state (for instance a scattering process in which two particles that collide have an energy which is sufficient to produce particles in the final state that are different from the ones in the initial state). This is a characteristic that a relativistic theory must have, because of the correspondence energy-mass. We have, therefore, to find a theory which is much more flexible and general than NRQM, that include more known processes and that has, as a non relativistic limit, Schrödinger's theory.

The first attempt to include Special Relativity in quantum mechanics regarded the search of a relativistically "correct" evolution equation and it brought to the so-called Klein-Gordon equation (Schrödinger himself worked to such equation in the same years or even before to write his famous article on wave mechanics). If we consider the fact that Eq. (1.1) can be derived from the non-relativistic energy-momentum relation

$$E = \frac{p^2}{2m}\,, \tag{1.2}$$

with the correspondences

$$E \to i\hbar\frac{\partial}{\partial t}\,, \qquad \mathbf{p} \to -i\hbar\nabla\,, \tag{1.3}$$

we can try to include Special Relativity in quantum mechanics using the correct relativistic energy-momentum relation

$$\frac{E^2}{c^2} = p^2 + m^2c^2\,, \tag{1.4}$$

finding the following differential equation:

$$\left(\frac{\hbar^2}{c^2}\frac{\partial^2}{\partial t^2} - \hbar^2\nabla^2 + m^2c^2\right)\phi(\mathbf{x},t) = 0\,. \tag{1.5}$$

In the case in which one would like to interpret Eq. (1.5) as a wave equation, ("*à la Schrödinger*"), he would face many problems. As we will se in the next chapters, first of all the probability density connected to the field $\phi(\mathbf{x},t)$ is not positive definite. This put immediately in troubles the probabilistic interpretation. Moreover, Eq. (1.5) admits plane-wave solutions with positive energy $E = \sqrt{p^2+m^2}$ but also with negative energy, $E = -\sqrt{p^2+m^2}$. While classically this would not cause particular issues[1], from a quantum mechanical point of view, this would mean that a particle can jump from a positive-energy state to a negative-energy one emitting a photon (for instance) and then, since the spectrum is unbounded from below, it would keep on emitting and jumping to bigger and bigger negative energies.

These two issues made in such a way that the Klein-Gordon theory was temporarly abandoned.

A successful step forward was instead done by Dirac in 1927. Dirac realized that the non positivity of the probability density in the KG equation was due to the fact that the derivative with respect to the time is of the second order and postulated the following first order (in time and, as required by special relativity, also in space) differential equation describing the wave equation of an electron[2]:

$$i\frac{\partial}{\partial t}\psi(\mathbf{x},t) = (-i\alpha\cdot\nabla + \beta m)\,\psi(\mathbf{x},t)\,. \tag{1.6}$$

Eq. (4.248) is such that the probability density, $\rho = |\psi|^2$, is actually positive definite. However, Dirac's equation still admits positive as well as negative-energy plane-wave solutions, but if we can for some reason neglect the contribution of negative-energy solutions we can solve the equation for the Hydrogen atom, finding a spectrum in very good agreement with the experimental measurements. Finally, Dirac's equation includes the description of the spin, that emerges in a natural way from the theory and does not need an ad hoc contruction, and its non-relativistic limit is the Pauli equation, as we would expect.

In order to physically interpret negative-energy solutions, Dirac introduced the so-called "holes theory", that predict a particle with the same mass of the electron but with negative charge, the electron "antiparticle" (or "positron"). This explaination was again a success, since the positron was actually detected in cosmic rays by Anderson in 1932.

In this context is inserted also the problem of the quantization of the elecctromagnetic field, that in non relativistic Schrödinger's theory is still considered as a classical field. The quantization of this field would consists in finding a method to describe the corpuscular nature of the field, i.e. the photon. In this sense, there is an asymmetry between the treatment of the Dirac electron or the Klein-Gordon scalar particle and the electromagnetic field.

While for the former we look for an equation that starts already from the corpuscular nature of the field, which is evident from the classical limit, for the latter the problem is faced differently: we have the field equations (Maxwell's eqs) that involve a classical field and show the wave nature of the light and we want to quantize them to describe the microscopic quantum nature of the field.

Quantum field theory makes the procedure uniform. We consider the relativistic equations of Dirac and Klein-Gordon as "classic equations" for the relative fields. Then, we quantize them to describe the particle nature of those fields. Since this is the philosophy, the name "second quantization" is not appropriate anymore. The so-called "first quantization" is nothing else than the identification of the correct relativistic equation that the field has to satisfy.

[1]The two energy solutions are separated by a finite gap that classically will never be overcome. Moreover, classically one can always discart negative energy solutions on the basis of the fact that they are non-physical.

[2]We use natural units $\hbar = c = 1$.

1.2 Summary of the quantization procedure

On the basis of what so far discussed, we can summarize point-by-point the quantization procedure that we will follow in constructing the theory:

- Firstly we find the field equations. We will study the Klein-Gordon, Dirac and Maxwell's equations. All the equations will be considered as classical equations that the different fields have to satisfy.
- The field equations are the Euler-Lagrange equations derived from a Lagrangian density. We introduce the lagrangian formalism. We look for conserved quantities, via the Nöther's theorem, that will play the role of observables in the quantum theory.
- In order to formulate the procedure of quantization of a field, i.e. of a system with infinite degrees of freedom, we look at a peculiar system: the vibrating string. This system can be thought of as the continuum limit of a one-dimensional distribution of harmonic oscillators, that we know how to treat and quantize in the discrete.
- We find the momenta conjugated to the fields and then we move to the Hamiltonian description of the system. The fields are promoted to time-dependent operators (in the Heisenberg picture) that act on a Hilbert space[3]. We then impose the commutation relations among fields and conjugated momenta, performing the so-called canonical quantization.
- We apply canonical quantization first of all to the free (non interacting) fields. In order to construct a coherent picture, we will be able to include spin-statistics in the quantization, treating consistently particles that obey Bose-Einstein statistics with commutation relations, while the particles that obey Fermi-Dirac statistics have to be quantized using anticommutation relations.
- Probably the most interesting part, since it is the one that regards directly our experiments, is the treatment of interacting fields. To consider the interactions, we will have to find a suitable Lagrangian (for example by minimal substitution, in the case of electromagnetic interactions), and Hamiltonian and extend to this case the canonical quantization rules. the conserved current does not contain time derivatives of the fields and, therefore, for
- The formalism so developed, will give the possibility to study transition processes from n-particle initial states to m-particle final states ...

Before starting, in the next Chapter, with the introduction of the various fields treated in the theory, we lay the foundations for their definition.

First of all, we study the vibrating string in order to understand what can be intended as a field, just to have a concrete idea linked to a very well understood mechanical system). Then we will use this system to explain quantization.

The theory we are trying to build has as its main feature the invariance under Poincaré transformations, in the sense that *Physics* studied in two different inertial frames must be the same. Consequently the action will have to be invariant under Poincaré transformations and the various fields will have a well defined behavior under these transformations.

We will, therefore, briefly recap the main features of the Lorentz and Poincaré groups.

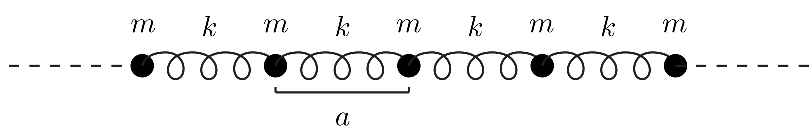
[3]Actually we will see that, since the numer of particles is not anymore a conserved quantity, we need a peculiar space, tensor product of a variable numer of Hilbert spaces, called the Fock space.

1.3 One-dimensional chain

One of the key points of Quantum Field Theory is the fact that we have to construct a formalism with infinite degrees of freedom, in order to be able to adjust the treatment of a system with a variable number of particles. Since our intuition is connected to the one-particle case (analytical mechanics, non relativistic quantum mechanics) we will start with a discontinuus system and we will define a sort of limit to the continuum to move from the one-particle description to the "many particles" description that will be connected to the field.

The field describes, in this treatment, the fluctuations with respect to a certain state (for instance the equilibrium state or the vacuum state) and the particles will be connected to the quanta of the modes (Fourier modes) of these fluctuations.

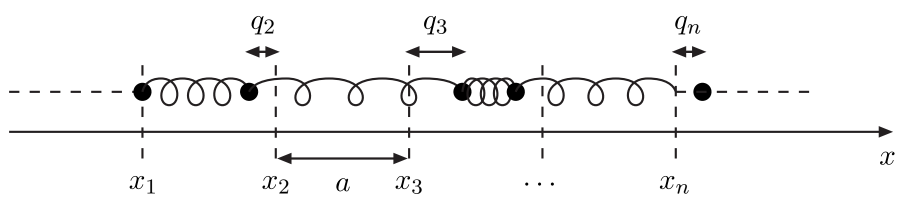
In order to understand the transition to the continuum and the quantization of a continuus system, we study the case of the linear chain of harmonic oscillators. Let us consider, then, a system of $(N+1)$ material points, all of them with mass m, interacting through an harmonic potential (springs with same constant k) as in the figure:



At the equilibrium, the particles are separated by a distance a and therefore the lenght of the chain is $L = aN$. Let us consider for simplicity N even (this is not a big constraint since then we want to take the limit $N \to \infty$).

Let us also consider $m = 1$.

In the excited situation the n-th particle oscillates around the equilibrium configuration of a quantity that we will call $q_n(t)$.



Let us make another assumption: the interaction of the n-th particle is limited to the nearest particles, in such a way that the potential energy of the system can be written as follows:

$$V = \frac{1}{2}\omega^2 \sum_{n=1}^{N} (q_n - q_{n+1})^2 \tag{1.7}$$

and therefore the equations of motions are

$$\ddot{q}_n = -\frac{\partial V}{\partial q_n} = -\omega^2 \left[(q_n - q_{n+1}) - (q_{n-1} - q_n)\right] = \omega^2 (q_{n+1} + q_{n-1} - 2q_n)\,. \tag{1.8}$$

This means that the n-th particle feels a force which is decomposed in two parts: the force of the spring on the left and the force of the spring on the right.

Since we are dealing with a chain, we have to choose what happens at the end of the chain, i.e. boundary conditions. We can impose two kinds of boundary conditions: *i)* chain with fixed end-points $q_1 = q_{N+1} = 0$, or *ii)* chain with periodic boundary conditions $q_n = q_{n+N}$. Since in the end we want to take also the limit $L \to \infty$, both cases bring to the same conclusions.

We will chose the case of periodic boundary conditions[4]

$$q_n = q_{n+N}\,, \tag{1.9}$$

studying basically the case of a ring of particles connected by springs. In this situation we have N particles and N springs, so the sums over n go from 1 to N.

The kinetic energy of the chain is

$$T = \frac{1}{2}\sum_{n=1}^{N} \dot{q}_n^2 \tag{1.10}$$

and therefore we can write down the lagrangian and the hamiltonian of the system as

$$L = T - V = \frac{1}{2}\sum_{n=1}^{N} \dot{q}_n^2 - \frac{1}{2}\omega^2 \sum_{n=1}^{N} (q_n - q_{n+1})^2\,, \tag{1.11}$$

$$H = T + V = \frac{1}{2}\sum_{n=1}^{N} \dot{q}_n^2 - \frac{1}{2}\omega^2 \sum_{n=1}^{N} (q_n - q_{n+1})^2\,. \tag{1.12}$$

We notice that the equations of motion (1.8) and the hamiltonian (4.107) are not diagonalized. We can diagonalize them moving to the normal modes, i.e. looking for a solution in Fourier series. In order to do that, we look for the following solution

$$q_n^{(j)}(t) = c_j(t)e^{ik_j x_n} = \Big| \text{ since } x_n = na \Big| = c_j(t)e^{ik_j na}\,, \tag{1.13}$$

anticipating what will come from the imposition of the boundary conditions, that k_j is indeed enumerable.

If we substitute (1.13) into the equations of motion (1.8) we find

$$\ddot{c}_j(t)e^{ik_j x_n} = \omega^2 \left[e^{ik_j(n+1)a} + e^{ik_j(n-1)a} - 2e^{ik_j na}\right] c_j(t)\,, \tag{1.14}$$

$$= \omega^2 \left[e^{ik_j a} + e^{-ik_j a} - 2\right] c_j(t)e^{ik_j x_n}\,, \tag{1.15}$$

$$= -\omega^2 \left[2 - 2\cos(k_j a)\right] c_j(t)e^{ik_j x_n}\,, \tag{1.16}$$

$$= -4\omega^2 \sin^2\left(\frac{k_j a}{2}\right) c_j(t)e^{ik_j x_n}\,. \tag{1.17}$$

Defining

$$\omega_j^2 = 4\omega^2 \sin^2\left(\frac{k_j a}{2}\right), \tag{1.18}$$

we find that $c_j(t)$ has to be the solution of the equation of an harmonic oscillator of frequency ω_j

$$\ddot{c}_j(t) + \omega_j^2 c(t) = 0\,. \tag{1.19}$$

Eq. (1.18) is the dispersion relation that link the frequency ω_j to the wave number k_j. The $c_j(t)$ are the normal modes, that decouple the system. The relation (1.18) is periodic. If we consider k_j and $k_j + \frac{2\pi}{a}m$, with $m \in \mathbb{Z}$, we get the same value for ω_j. Therefore, we can restrict our analysis to the so-called “first Brillouin zone”, i.e.

$$|k_j| \leq \frac{\pi}{a}\,. \tag{1.20}$$

Imposing the boundary conditions in Eq. (1.9), we find

$$e^{ik_j(n+N)a} = e^{ik_j na}\,, \tag{1.21}$$

[4]For fixed end-point conditions see for instance ....

or

$$k_j = \frac{2\pi}{aN} j = \frac{2\pi}{L} j \,. \tag{1.22}$$

In the first Brillouin zone we have to impose Eq. (1.20), therefore

$$\left| \frac{2\pi}{aN} j \right| \leq \frac{\pi}{a} \quad \Rightarrow |j| \leq \frac{N}{2} \,. \tag{1.23}$$

We then find $N+1$ modes. However, the solution $j=0$ gives $k_j = 0$, then $\omega_j = 0$ and finally q_n linear in time. This corresponds to a rigid translation of the chain, that we are not going to consider (we want to study the vibrations only). Therefore

$$k_j = \frac{2\pi}{L} j \,, \qquad j = \pm 1, \pm 2, ..., \pm \frac{N}{2} \,. \tag{1.24}$$

The general solution can be cast in the following form

$$q_n(t) = \sum_{j=-N/2}^{N/2} q_n^{(j)} = \sum_{j=-N/2}^{N/2} e^{ik_j an} \frac{Q_j(t)}{\sqrt{N}} = \sum_{j=-N/2}^{N/2} e^{i\frac{2\pi}{N}nj} \frac{Q_j(t)}{\sqrt{N}} \,, \tag{1.25}$$

where we put

$$Q_j(t) = \sqrt{N} c_j(t) \,, \tag{1.26}$$

that, again, are solutions of the following differential equation:

$$\ddot{Q}_j + \omega_j^2 Q_j = 0 \,. \tag{1.27}$$

Since $q_n(t)$ has to be a real quantity (it is the physical displacement of the n-th particle), we have to impose $q_n^* = q_n$ and then

$$\sum_{j=-N/2}^{N/2} e^{-i\frac{2\pi}{N}nj} \frac{Q_j^*(t)}{\sqrt{N}} = \sum_{j=-N/2}^{N/2} e^{i\frac{2\pi}{N}nj} \frac{Q_j(t)}{\sqrt{N}} \,. \tag{1.28}$$

Puting $j \to -j$ in the r.h.s. of Eq. (1.28), we find that the following relation must hold:

$$Q_{-j}(t) = Q_j^*(t) \,. \tag{1.29}$$

Using the following representation of the Kronecker delta[5]

$$\sum_{n=1}^{N} e^{i\frac{2\pi}{N}(j-j')n} = N\, \delta_{jj'} \,, \tag{1.31}$$

we can write the lagrangian in terms of the normal modes (for the moment we leave unexpressed the potential energy V):

$$L = \frac{1}{2}\sum_{n=1}^{N} \dot{q}_n^2 - V = \frac{1}{2}\sum_{n=1}^{N} \sum_{j,j'=-N/2}^{N/2} e^{i\frac{2\pi}{N}(j+j')n} \frac{\dot{Q}_j \dot{Q}_{j'}}{N} - V \,,$$

[5]This formula can be justified easily as follows. If $j = j'$, this is trivially N. If, instead, $j \neq j'$ we have

$$\begin{aligned} \sum_{n=1}^{N} e^{i\frac{2\pi}{N}(j-j')n} &= \sum_{n=0}^{N} \left[e^{i\frac{2\pi}{N}(j-j')} \right]^n - 1 = \frac{1 - e^{i\frac{2\pi}{N}(j-j')(N+1)}}{1 - e^{i\frac{2\pi}{N}(j-j')}} - 1 \,, \\ &= \frac{e^{i\frac{2\pi}{N}(j-j')} - e^{i2\pi(j-j')} e^{i\frac{2\pi}{N}(j-j')}}{1 - e^{i\frac{2\pi}{N}(j-j')}} = e^{i\frac{2\pi}{N}(j-j')} \frac{1 - e^{i2\pi(j-j')}}{1 - e^{i\frac{2\pi}{N}(j-j')}} = 0 \,. \end{aligned} \tag{1.30}$$

$$
\begin{aligned}
&= \left|\text{with} \quad j' \to -j'\right| = \frac{1}{2} \sum_{j,j'=-N/2}^{N/2} N\delta_{jj'} \frac{\dot{Q}_j \dot{Q}_{-j'}}{N} - V \,, \\
&= \frac{1}{2} \sum_{j=-N/2}^{N/2} \dot{Q}_j \dot{Q}_{-j} - V = \frac{1}{2} \sum_{j=-N/2}^{N/2} \dot{Q}_j^* \dot{Q}_j - V \,. \qquad (1.32)
\end{aligned}
$$

Moreover, since $\dot{Q}_j^* \dot{Q}_j = \dot{Q}_{-j} \dot{Q}_{-j}^*$ we have

$$
\frac{1}{2} \sum_{j=-N/2}^{N/2} \dot{Q}_j^* \dot{Q}_j = \sum_{j=1}^{N/2} \dot{Q}_j^* \dot{Q}_j = \sum_{j=1}^{N/2} \left|\dot{Q}_j\right|^2 . \qquad (1.33)
$$

Therefore, we can define the momenta conjugated to Q_j as

$$
P_j = \frac{\partial L}{\partial \dot{Q}_j} = \dot{Q}_j^* \,. \qquad (1.34)
$$

On the other hand, we have the momenta conjugated to q_n defined as follows

$$
\begin{aligned}
p_n(t) &= \frac{\partial L}{\partial \dot{q}_n(t)} = \dot{q}_n(t) = \sum_{j=-N/2}^{N/2} e^{i\frac{2\pi}{N}nj} \frac{\dot{Q}_j(t)}{\sqrt{N}} = |j \to -j| \,, \\
&= \sum_{j=-N/2}^{N/2} e^{-i\frac{2\pi}{N}nj} \frac{\dot{Q}_{-j}(t)}{\sqrt{N}} = \sum_{j=-N/2}^{N/2} e^{-i\frac{2\pi}{N}nj} \frac{\dot{Q}_j^*(t)}{\sqrt{N}} = \sum_{j=-N/2}^{N/2} e^{-i\frac{2\pi}{N}nj} \frac{P_j(t)}{\sqrt{N}} \,, \qquad (1.35)
\end{aligned}
$$

where we used the fact that also $p_n(t)$ must be real, and then $p_n^* = p_n$, that implies

$$
\dot{Q}_{-j}(t) = \dot{Q}_j^*(t) \,, \qquad (1.36)
$$

or, in terms of $P_j(t)$,

$$
P_{-j}(t) = P_j^*(t) \,. \qquad (1.37)
$$

For later use, we can write Q_j and P_j in terms of q_n and p_n, using the representation of the delta in Eq. (1.31). In fact

$$
\sum_{n=1}^{N} q_n(t) e^{-i\frac{2\pi}{N}jn} = \sum_{n=1}^{N} \sum_{j'=-N/2}^{N/2} e^{i\frac{2\pi}{N}(j'-j)n} \frac{Q_{j'}(t)}{\sqrt{N}} = \sum_{j'=-N/2}^{N/2} N\delta_{jj'} \frac{Q_{j'}(t)}{\sqrt{N}} = \sqrt{N} Q_j(t) \,. \qquad (1.38)
$$

Therefore

$$
Q_j(t) = \sum_{n=1}^{N} e^{-i\frac{2\pi}{N}jn} \frac{q_n(t)}{\sqrt{N}} \,. \qquad (1.39)
$$

In the same way, we find

$$
P_j(t) = \sum_{n=1}^{N} e^{i\frac{2\pi}{N}jn} \frac{p_n(t)}{\sqrt{N}} \,. \qquad (1.40)
$$

Now, let us write the hamiltonian of the system in terms of normal modes. We have

$$
H = \frac{1}{2} \sum_{n=1}^{N} \dot{q}_n^2 + \frac{1}{2}\omega^2 \sum_{n=1}^{N} (q_n - q_{n+1})^2 = \frac{1}{2} \sum_{n=1}^{N} p_n^2 + \frac{1}{2}\omega^2 \sum_{n=1}^{N} (q_n - q_{n+1})^2 \,,
$$

$$
\begin{aligned}
&= \frac{1}{2}\sum_{n=1}^{N}\left\{\sum_{j,j'=-N/2}^{N/2} e^{-i\frac{2\pi}{N}(j+j')n}\frac{P_jP_{j'}}{N}+\omega^2\left[\sum_{j=-N/2}^{N/2} e^{i\frac{2\pi}{N}jn}\frac{Q_j(t)}{\sqrt{N}}-\sum_{j=-N/2}^{N/2} e^{i\frac{2\pi}{N}j(n+1)}\frac{Q_j(t)}{\sqrt{N}}\right]^2\right\}, \\
&= \frac{1}{2}\sum_{n=1}^{N}\left\{\sum_{j,j'=-N/2}^{N/2} e^{-i\frac{2\pi}{N}(j+j')n}\frac{P_jP_{j'}}{N}+\omega^2\left[\sum_{j=-N/2}^{N/2}(1-e^{i\frac{2\pi}{N}j})e^{i\frac{2\pi}{N}jn}\frac{Q_j(t)}{\sqrt{N}}\right]^2\right\}, \\
&= \frac{1}{2}\sum_{n=1}^{N}\left\{\sum_{j,j'=-N/2}^{N/2} e^{-i\frac{2\pi}{N}(j+j')n}\frac{P_jP_{j'}}{N}+\omega^2\left[\sum_{j,j'=-N/2}^{N/2}(1-e^{i\frac{2\pi}{N}j})(1-e^{i\frac{2\pi}{N}j'})e^{i\frac{2\pi}{N}(j+j')n}\frac{Q_jQ_{j'}}{N}\right]\right\}, \\
&= \Big|\text{we put } j\to -j \text{ in the kinetic energy and } j'\to -j' \text{ in the potential energy}\Big|, \\
&= \frac{1}{2}\sum_{n=1}^{N}\left\{\sum_{j,j'=-N/2}^{N/2} e^{i\frac{2\pi}{N}(j-j')n}\frac{P_{-j}P_{j'}}{N}+\omega^2\left[\sum_{j,j'=-N/2}^{N/2}(1-e^{i\frac{2\pi}{N}j})(1-e^{-i\frac{2\pi}{N}j'})e^{i\frac{2\pi}{N}(j-j')n}\frac{Q_jQ_{-j'}}{N}\right]\right\}, \\
&= \Big|\text{using Eq. (1.31)}\Big|, \\
&= \frac{1}{2}\sum_{j,j'=-N/2}^{N/2}\left\{N\delta_{jj'}\frac{P_j^*P_{j'}}{N}+\omega^2\left[(1-e^{i\frac{2\pi}{N}j})(1-e^{-i\frac{2\pi}{N}j'})N\delta_{jj'}\frac{Q_jQ_{j'}^*}{N}\right]\right\}, \\
&= \frac{1}{2}\sum_{j=-N/2}^{N/2}\left\{|P_j|^2+\omega^2\left(2-e^{i\frac{2\pi}{N}j}-e^{-i\frac{2\pi}{N}j})\right)|Q_j|^2\right\}, \\
&= \Big|\text{since } \omega_j^2 = 4\omega^2\sin^2\left(\frac{\pi}{N}j\right)\Big|, \\
&= \frac{1}{2}\sum_{j=-N/2}^{N/2}\left\{|P_j|^2+\omega_j^2\,|Q_j|^2\right\}. \qquad (1.41)
\end{aligned}
$$

Since

$$
\begin{aligned}
|P_j|^2 &= P_j^*P_j = P_{-j}P_{-j}^* = |P_{-j}|^2\,, && (1.42)\\
|Q_j|^2 &= Q_j^*Q_j = Q_{-j}Q_{-j}^* = |Q_{-j}|^2\,, && (1.43)\\
\omega_j^2 &= \omega_{-j}^2\,, && (1.44)
\end{aligned}
$$

we can write the hamiltonian as follows:

$$
H = \sum_{j=1}^{N/2}\left\{|P_j|^2+\omega_j^2\,|Q_j|^2\right\}. \qquad (1.45)
$$

Looking at Eq. (1.45) is clear that the system is equivalent to a system composed by N decoupled harmonic oscillators (Q_j and P_j are complex quantities).

1.3.1 Limit to the continuum

In order to move to the continuum, let us write the displacement $q_n(t)$ as a funcion of x_n as follows:

$$
u(x_n,t) = q_n(t)\,. \qquad (1.46)
$$

In this way, $u(x_n,t)$ represents the displacement from the equilibrium of the n-th massive point-like particle (i.e. a function that for each x_n gives the displacement from x_n of that massive point). We can then rewrite the equations of motion (1.8) in terms of $u(x_n,t)$ geting

$$
\ddot{u}(x_n,t) = \omega^2\left[(u(x_{n+1},t)-u(x_n,t))-(u(x_n,t)-u(x_{n-1},t))\right]. \qquad (1.47)
$$

Let us consider the limit in which the massive points become denser and denser, $N \to \infty$ and $a \to 0$, keeping the lenght of the string finite, $Na = L = const$. We have also to consider the fact that puting more and more point in the chain we add masses. However, we want to keep the mass of the chain finite, therefore we may also keep the ratio $m/a = \mu$ constant (constant mass density over the string). Note that for simplicity we kept $m = 1$ from the beginning, so we will have to rescale the field by a factor $1/\sqrt{a}$ to keep the energy of the string finite.

We can interpret this limit considering the propagation of acoustic waves with a wave lenght $\lambda >> a$.

In this limit we have

$$\lim_{a\to 0} \frac{u(x_{n+1},t) - u(x_n,t)}{a} = u'(x,t)\,, \tag{1.48}$$

where $x \in [x_{n+1}, x_n]$ and where $x_{n+1} \to x$, $x_n \to x$. In the same way we have

$$\lim_{a\to 0} \frac{u'(x_1,t) - u'(x_2,t)}{a} = u''(x,t)\,, \tag{1.49}$$

where $x \in [x_1, x_2]$. In the end, Eq. (1.47) becomes

$$\ddot{u}(x,t) = \omega^2 a^2 u''(x,t)\,, \tag{1.50}$$

which is a wave equation and $\omega^2 a^2 = v^2$ is the velocity of propagation of the waves throught the string. It is clear that $\omega^2 a^2$ should be a constant since

$$\omega_j^2 = 4\omega^2 \sin^2\left(\frac{k_j a}{2}\right) \sim 4\omega^2 \left(\frac{k_j a}{2}\right)^2 = \omega^2 a^2\, k_j^2 \tag{1.51}$$

and therefore in the limit $a \to 0$, ω should go like $1/a$ to give a finite frequency and wave number of the j-th mode. Then we find

$$\ddot{u}(x,t) = v^2 u''(x,t)\,, \tag{1.52}$$

with the usual D'Alambert solution $u(x,t) = f(x+vt) + g(x-vt)$ ...

Let us see what happens to the hamiltonian. In terms of $u(x_n,t)$ we can write

$$H = \frac{1}{2}\sum_{n=1}^{N} \dot{(u)}^2(x_n,t) + \frac{1}{2}\omega^2 \sum_{n=1}^{N} \left[u(x_n,t) - u(x_{n-1},t)\right]^2 . \tag{1.53}$$

In the limit to the continuum, we have[6]

$$\sum_{n=1}^{N} \sim \frac{1}{a}\int_0^L dx \tag{1.55}$$

and therefore

$$H = \frac{1}{2a}\int_0^L \left[\dot{u}^2(x,t) + v^2 u'^2(x,t)\right] dx\,. \tag{1.56}$$

The fact that H has a factor $1/a$ depends on having set $m = 1$ at the beginning. So, in order to have a finite energy on the string we have to impose that the function $u(x,t)$ goes to 0 as $\sqrt{a}$. We then

[6] We can understand recalling the way we define the integral

$$\int_0^L f(x)\, dx = \lim_{\substack{N\to\infty \\ a\to 0}} \sum_{n=1}^{N} a\, f_n\,, \tag{1.54}$$

where f_n is the value of the function $f(x)$ in $x = x_n = na$.

redefine the field[7]

$$\phi(x,t) = \frac{u(x,t)}{\sqrt{a}}\,, \tag{1.61}$$

in terms of which we have

$$H = \frac{1}{2}\int_0^L \left[\dot\phi^2(x,t) + v^2\phi'^2(x,t)\right] dx \tag{1.62}$$

and the equations of motion

$$\ddot\phi(x,t) = v^2\phi''(x,t)\,. \tag{1.63}$$

In the limiting procedure, we do not change the boundary conditions. In fact the relation

$$u(x_{n+N},t) = u(x_n,t)\,, \tag{1.64}$$

corresponds to

$$u(x+L,t) = u(x,t)\,, \tag{1.65}$$

and this in any case gives rise to a wave number which is enumerable:

$$e^{ik_j(x+L)} = e^{ik_jx}\,, \tag{1.66}$$

tha implies

$$k_j = \frac{2\pi}{L}\,j\,, \qquad j\in\mathbb{Z}\,. \tag{1.67}$$

Now j can go from $-\infty$ to $+\infty$ (always except 0).

Since $aN = L$ we can define the field ϕ in normal modes

$$\phi(x,t) = \sum_{j=-\infty}^{\infty} e^{i\frac{2\pi}{L}jx}\frac{Q_j(t)}{\sqrt{L}} \tag{1.68}$$

and the dispersion relation becomes

$$\omega_j^2 \to 4\omega^2\left(\frac{k_ja}{2}\right)^2 = v^2k_j^2\,, \tag{1.69}$$

typical of a wave equation.

The normal modes Q_j still satisfy a single harmonic oscillator differential equation

$$\ddot Q_j(t) + v^2k_j^2\,Q_j(t) = 0 \tag{1.70}$$

[7] We can look at the same procedure keeping m and setting $m/a = \mu = const$ in the limit to the continuum. In these terms, we have

$$T = \frac{1}{2}m\sum_n \dot q_n^2 \to \frac{1}{2a}\int_0^L dx\, m\dot u^2(x,t) \to \frac{\mu}{2}\int_0^L dx\,\dot u^2(x,t)\,, \tag{1.57}$$

$$V = \frac{1}{2}k\sum_n (q_{n+1}-q_n)^2 \to \frac{1}{2a}\int_0^L dx\, ka^2\, u'^2(x,t) \to \frac{\tau}{2}\int_0^L dx\, u'^2(x,t)\,, \tag{1.58}$$

$$Na = L\,, \quad \frac{m}{a} = \mu\,, \quad ka = \tau\,, \tag{1.59}$$

where τ is the tension of the string and $\tau/\mu = v^2$. Then

$$L = \frac{\mu}{2}\int_0^L \left[\dot u^2(x,t) - v^2u'^2(x,t)\right] = \frac{1}{2}\int_0^L \left[\dot\phi^2(x,t) - v^2\phi'^2(x,t)\right]\,, \tag{1.60}$$

where we defined the field $\phi(x,t) = \sqrt{\mu}\,u(x,t)$.

and in terms of $Q_j(t)$ and $P_j(t)$ we can express the hamiltomnian as follows:

$$
\begin{aligned}
H &= \frac{1}{2}\int_0^L \left[\dot{\phi}^2(x,t) + v^2\phi'^2(x,t)\right] dx\,, \\
&= \frac{1}{2}\int_0^L dx \left[\sum_{jj'} e^{i\frac{2\pi}{L}(j+j')x}\frac{\dot{Q}_j\dot{Q}_{j'}}{L} - v^2 k_j k_{j'} \sum_{jj'} e^{i\frac{2\pi}{L}(j+j')x}\frac{Q_j Q_{j'}}{L}\right]\,, \\
&= \left| j' \to -j' \right|\,, \\
&= \frac{1}{2}\int_0^L dx \left[\sum_{jj'} e^{i\frac{2\pi}{L}(j-j')x}\frac{\dot{Q}_j\dot{Q}_{-j'}}{L} - v^2 k_j k_{-j'} \sum_{jj'} e^{i\frac{2\pi}{L}(j-j')x}\frac{Q_j Q_{-j'}}{L}\right]\,, \\
&= \left|\text{since} \quad \int_0^L e^{i\frac{2\pi}{L}(j-j')x} dx = L\delta_{jj'}\right|\,, \\
&= \frac{1}{2}\sum_{jj'}\left[L\delta_{jj'}\frac{\dot{Q}_j\dot{Q}_{-j'}}{L} - v^2 k_j k_{-j'} L\delta_{jj'}\frac{Q_j Q_{-j'}}{L}\right]\,, \\
&= \frac{1}{2}\sum_{j=-\infty}^{\infty}\left[\dot{Q}_j\dot{Q}_j^* + v^2 k_j^2 Q_j Q_j^*\right]\,, \\
&= \sum_{j=1}^{\infty}\left[\left|\dot{Q}\right|^2 + v^2 k_j^2 \left|Q\right|^2\right]\,, \qquad (1.71)
\end{aligned}
$$

since, as before, $\dot{Q}_j\dot{Q}_j^* = \dot{Q}_{-j}^*\dot{Q}_{-j}$, $Q_jQ_j^* = Q_{-j}^*Q_{-j}$ and $k_j^2 = k_{-j}^2$. Again, we find that the system is equivalent to an infinite sum of harmonic oscillators. The fact that the various frequencies are enumerable is due to the finite lenght of the ring and the boundary conditions imposed. In the case in which also the lenght L goes to infinity, we would have to deal with a Fourier transform instead of a series.

The quantization of the system is done by the quantization of these harmonic oscillators.

1.3.2 Quantization of the vibrating string

We are in a situation in which our continuous system, the vibrating string, is solved in terms of normal modes, diagonalizing the hamiltonian that can be written as an infinite sum of decoupled harmonic oscillators. This pattern gives rise to a "simple" procedure for the quantization of the system. This can be done through the canonical quantization of the harmonic oscillators.

In the discrete system we have

$$
\begin{aligned}
H &= \sum_{j=1}^{N/2}\left[|P_j|^2 + \omega_j^2 |Q_j|^2\right]\,, \qquad (1.72) \\
\omega_j^2 &= 4\omega^2 \sin^2\left(\frac{k_j a}{2}\right)\,, \qquad (1.73) \\
k_j &= \frac{2\pi}{L}j\,, \qquad |j| = 1, 2, ..., \frac{N}{2}\,. \qquad (1.74)
\end{aligned}
$$

In the continuum case

$$
\begin{aligned}
H &= \sum_{j=1}^{\infty}\left[|P_j|^2 + \omega_j^2 |Q_j|^2\right]\,, \qquad (1.75) \\
\omega_j^2 &= v^2 k_j^2\,, \qquad (1.76)
\end{aligned}
$$

$$k_j = \frac{2\pi}{L} j , \qquad j \in \mathbb{Z} . \tag{1.77}$$

In both cases, in order to quantize the single harmonic oscillators, we will have to promote Q_j and P_j to operators. Consequently, the relations (1.29,1.37) will become

$$\hat{Q}_j^\dagger = \hat{Q}_{-j} , \qquad \hat{P}_j^\dagger = \hat{P}_{-j} . \tag{1.78}$$

Let us start with the discrete case. We can introduce annihilation and creation operators[8], $\hat{a}_j$ and $\hat{a}_j^\dagger$, such that

$$\hat{a}_j = \sqrt{\frac{\omega_j}{2}} \hat{Q}_j + \frac{i}{\sqrt{2\omega_j}} \hat{P}_j^\dagger , \tag{1.81}$$

$$\hat{a}_j^\dagger = \sqrt{\frac{\omega_j}{2}} \hat{Q}_j^\dagger - \frac{i}{\sqrt{2\omega_j}} \hat{P}_j , \tag{1.82}$$

where $\omega_j = 2\omega |\sin(k_j a/2)|$. Starting from the quantization relations[9] on the hermitian operators $\hat{q}_n$ and the conjugated momenta $\hat{p}_n$

$$[\hat{q}_n, \hat{p}_m] = i\delta_{nm} , \qquad [\hat{q}_n, \hat{q}_m] = [\hat{p}_n, \hat{p}_m] = 0 , \tag{1.83}$$

we find that also the operators $\hat{Q}_j$ and $\hat{P}_j$ obey similar commutation relations[10]:

$$[\hat{Q}_j, \hat{P}_{j'}] = i\delta_{jj'} , \qquad [\hat{Q}_j, \hat{Q}_{j'}] = [\hat{P}_j, \hat{P}_{j'}] = 0 , \tag{1.87}$$

Finally, we find (we will omit from now on the hat for simplicity of notation, but we are speaking about operators)

$$[a_j, a_k^\dagger] = \frac{1}{\sqrt{4\omega_j\omega_k}} [\omega_j Q_j + iP_j^\dagger, \omega_k Q_k^\dagger - iP_k] , \tag{1.88}$$

$$= \frac{1}{\sqrt{4\omega_j\omega_k}} \left\{ -i\omega_j [Q_j, P_k] + i\omega_k [P_j^\dagger, Q_k^\dagger] \right\} , \tag{1.89}$$

$$= \frac{1}{\sqrt{4\omega_j\omega_k}} \left\{ -i\omega_j [Q_j, P_k] - i\omega_k [Q_{-k}, P_{-j}] \right\} , \tag{1.90}$$

[8] We can write $\hat{Q}_j$ and $\hat{P}_j$ in terms of $\hat{a}_j$ and $\hat{a}_j^\dagger$ as follows:

$$\hat{Q}_j = \frac{1}{\sqrt{2\omega_j}} \left(\hat{a}_j + \hat{a}_{-j}^\dagger \right) , \quad \hat{P}_j = -i\sqrt{\frac{\omega_j}{2}} \left(\hat{a}_{-j} - \hat{a}_j^\dagger \right) , \tag{1.79}$$

where

$$\hat{a}_{-j} = \sqrt{\frac{\omega_j}{2}} \hat{Q}_j^\dagger + \frac{i}{\sqrt{2\omega_j}} \hat{P}_j , \qquad \hat{a}_{-j}^\dagger = \sqrt{\frac{\omega_j}{2}} \hat{Q}_j - \frac{i}{\sqrt{2\omega_j}} \hat{P}_j^\dagger . \tag{1.80}$$

We have, then, $[a_{-j}, a_{-k}^\dagger] = -\delta_{jk}$ and $[a_{-j}, a_{-k}] = [a_{-j}^\dagger, a_{-k}^\dagger] = 0$.

[9] Remember that we are using the natural units $\hbar = c = 1$.

[10] We have

$$\hat{Q}_j = \sum_{n=1}^{N} e^{-ik_j an} \frac{\hat{q}_n}{\sqrt{N}} , \quad \hat{P}_j = \sum_{n=1}^{N} e^{ik_j an} \frac{\hat{p}_n}{\sqrt{N}} \tag{1.84}$$

and therefore

$$[\hat{Q}_j, \hat{P}_{j'}] = \hat{Q}_j \hat{P}_{j'} - \hat{P}_{j'} \hat{Q}_j = \sum_{n,m} e^{-i\frac{2\pi}{N} jn} e^{i\frac{2\pi}{N} j'm} \frac{1}{N} \hat{q}_n \hat{p}_m - \sum_{n,m} e^{-i\frac{2\pi}{N} jn} e^{i\frac{2\pi}{N} j'm} \frac{1}{N} \hat{p}_m \hat{q}_n , \tag{1.85}$$

$$= \left| \text{using (1.83)} \;\; \hat{p}_m \hat{q}_n = \hat{q}_n \hat{p}_m + i\delta_{nm} \right| = i\delta_{jj'} . \tag{1.86}$$

$$\begin{aligned} &= \delta_{jk}\,, && (1.91)\\ [a_j, a_k] &= [a_j^\dagger, a_k^\dagger] = 0\,. && (1.92) \end{aligned}$$

In terms of annihilation and creation operators, the hamiltonian becomes:

$$\begin{aligned} H &= \sum_{j=1}^{N/2} \left[|P_j|^2 + \omega_j^2 \, |Q_j|^2 \right] , && (1.93)\\ &= \sum_{j=1}^{N/2} \frac{\omega_j}{2} \left[a_j^\dagger a_j + a_j a_j^\dagger + a_{-j}^\dagger a_{-j} + a_{-j} a_{-j}^\dagger \right] . && (1.94) \end{aligned}$$

If we use the commutation relations (1.91,1.92), we find

$$\begin{aligned} H &= \sum_{j=1}^{N/2} \omega_j \left[a_j^\dagger a_j + a_{-j}^\dagger a_{-j} + 1 \right] , \\ &= \sum_{j=-N/2}^{N/2} \omega_j \left[a_j^\dagger a_j + \frac{1}{2} \right] , && (1.95) \end{aligned}$$

where $\omega_j = 2\omega |\sin{(k_j a/2)}|$, which has the known form of the sum of N independent harmonic oscillators. Note that H is independent of time (energy conservation) although in the relations (1.81,1.82) they do depend on time. The time dependence of the operators a_j and $a_j^\dagger$ are given by the Hamilton's equations

$$\begin{aligned} \dot{a}_j(t) &= i[H, a_j] = -i\omega_j a_j(t)\,, && (1.96)\\ \dot{a}_j^\dagger(t) &= i[H, a_j^\dagger] = i\omega_j a_j^\dagger(t)\,, && (1.97) \end{aligned}$$

since, by derect inspection we have $[H, a_j^\dagger] = \omega_j a_j^\dagger$ and $[H, a_j] = -\omega_j a_j$. Then[11]

$$\begin{aligned} a_j(t) &= e^{-i\omega_j t} a_j(0)\,, && (1.100)\\ a_j^\dagger(t) &= e^{i\omega_j t} a_j^\dagger(0)\,. && (1.101) \end{aligned}$$

We can express the displacements $q_n(t)$ in terms of annihilation and creation operators:

$$\begin{aligned} q_n(t) &= \sum_{j=-N/2}^{N/2} e^{ik_j an} \frac{Q_j(t)}{\sqrt{N}}\,, \\ &= \sum_{j=-N/2}^{N/2} \frac{1}{\sqrt{N\,2\omega_j}} e^{ik_j an} (a_j(t) + a_{-j}^\dagger(t))\,, \\ &= \sum_{j=-N/2}^{N/2} \frac{1}{\sqrt{N\,2\omega_j}} e^{ik_j an} e^{-i\omega_j t} a_j(0) + \sum_{j=-N/2}^{N/2} \frac{1}{\sqrt{N\,2\omega_j}} e^{ik_j an} e^{i\omega_j t} a_{-j}^\dagger(0)\,, \\ &= \left| j \to -j \text{ in the second piece} \right| , \\ &= \sum_{j=-N/2}^{N/2} \frac{1}{\sqrt{N\,2\omega_j}} \left(e^{-i\omega_j t + ik_j an} a_j(0) + e^{i\omega_j t - ik_j an} a_j^\dagger(0) \right) . && (1.102) \end{aligned}$$

[11]For a_{-j} and $a_{-j}^\dagger$ we also have similar relations but with an opposite sign in the exponent

$$\begin{aligned} a_{-j}(t) &= e^{i\omega_j t} a_{-j}(0)\,, && (1.98)\\ a_{-j}^\dagger(t) &= e^{-i\omega_j t} a_{-j}^\dagger(0)\,, && (1.99) \end{aligned}$$

due to their commutation relations with the hamiltonian.

1.3.3 Fock space and phonons

We can now study the spectrum of the hamiltonian and give an interpretation of what we find.

The state with lowest energy is determined by the condition

$$a_j|0\rangle = 0\,, \qquad \forall j\,. \tag{1.103}$$

The corresponding eigenvalue is

$$E_0 = \sum_j \frac{1}{2}\omega_j\,. \tag{1.104}$$

For the moment we are considering the discrete case, therefore (1.104) constitutes a finite energy. When we will move to the continuum, this term will become infinite and we will have to redefine the energy of the vacuum state in order to "reabsorb" this infinity.

The creation operators act on the vacuum state as follows

$$a_j^\dagger|0\rangle = |j\rangle\,, \tag{1.105}$$

where $|j\rangle$ is an eigenstate of the hamiltonian with a definite energy ω_j. This state is also an eigenstate of the momentum (as we will see) and therefore corresponds to a state with a definite energy and momentum. This quantum of excitation can be interpreted as a particle, that has a definite energy and a brings a definite momentum. Note: it has nothing to do with the particles connected with springs that we started with. Now we are speaking about the quantization of the vibrations of the chain (a pure quantum description).

We can act again, n times, with $a_j^\dagger$ on the vacuum state, finding the state

$$\frac{(a_j^\dagger)^n}{\sqrt{n!}}|0\rangle = |n_j\rangle\,, \tag{1.106}$$

which is a state with energy given by the sum of the single energies (so, in this case n times ω_j) and momentum given by the sum of the momenta. This state can be interpreted as a state in which we have n particles with energy ω_j and definite momentum. Since $a_j^\dagger$ commutes with the other $a_k^\dagger$, the general eigenstate of the hamiltonian che bi written as follows:

$$|n_1, n_2, ..., n_N\rangle = \frac{1}{\sqrt{n_1!n_2!...n_N!}}(a_1^\dagger)^{n_1}(a_2^\dagger)^{n_2}...(a_N^\dagger)^{n_N}|0\rangle\,. \tag{1.107}$$

This state represent a state in which we have n_1 particles with energy ω_1, n_2 particles with energy ω_2, ... n_N particles with energy ω_N.

This interpretation is corroborated by the analogous case of the electromagnetic radiation and the explaination of the Photoelectric effect, in which it was introduced the quantum of the electromagnetic radiation (the photon) with a given discrete energy $\hbar\omega$.

The space we have introduced with this contruction is a direct sum of a variable number of Hilbert spaces and it is calle the Fock space.

Note the flexibility of this point of view! We can deal with a variable number of particles in our state. We can excite a particle state with $a_j^\dagger$ from the vacuum in such a way to move, for instance from a n-particle state to a $(n+1)$-particle state. We can destroy a particle in our state, acting with a_j ... and so on.

1.3.4 Commutation relations in the continuum

Using the commutation relations for Q_j and P_j, Eq. (1.87), we can find the commutation relations for the fields, since

$$\phi(x,t) \quad = \quad \frac{1}{\sqrt{L}}\sum_j e^{i\frac{2\pi}{L}jx}Q_j(t)\,, \tag{1.108}$$

$$\pi(x,t) = \frac{1}{\sqrt{L}}\sum_j e^{-i\frac{2\pi}{L}jx}P_j(t)\,. \tag{1.109}$$

We have[12]

$$\begin{aligned}
[\phi(x,t),\pi(y,t)] &= \frac{1}{L}\sum_{j,k} e^{i\frac{2\pi}{L}(jx-ky)}Q_j(t)P_k(t) - \frac{1}{L}\sum_{j,k} e^{i\frac{2\pi}{L}(jx-ky)}P_k(t)Q_j(t)\,, && (1.111)\\
&= \Big|\ \text{since } [Q_j(t),P_k(t)] = i\delta_{jk}\ \Big|\\
&= \frac{1}{L}\sum_{j,k} e^{i\frac{2\pi}{L}(jx-ky)}i\delta_{jk}\,, && (1.112)\\
&= \frac{i}{L}\sum_{j} e^{i\frac{2\pi}{L}j(x-y)}\,, && (1.113)\\
&= i\,\delta(x-y)\,. && (1.114)
\end{aligned}$$

Equivalently, we find

$$[\phi(x,t),\phi(y,t)] = [\pi(x,t),\pi(y,t)] = 0\,. \tag{1.115}$$

It is important to note the relationship between quantization conditions (in this case given by the commutation relations) and statistics obeyed by the particles. Since

$$[a_j, a_{j'}] = [a_j^\dagger, a_{j'}^\dagger] = 0\,, \tag{1.116}$$

the two-particle state is such that

$$|i,j\rangle = a_i^\dagger a_j^\dagger|0\rangle = a_j^\dagger a_i^\dagger|0\rangle = |j,i\rangle\,, \tag{1.117}$$

therefore, totally symmetric under the exchange of the particles. This is the case also for multi-particle states. This means that we are describing bosons.

1.3.5 Normal ordering

As we already noticed, in the continuum case the energy of the vacuum state becomes infinite:

$$E_0 = \sum_{j=1}^{\infty}\omega_j\,. \tag{1.118}$$

However, we note that in general what matters is the energy of a state with respect to the energy of the vacuum (i.e. a difference in energies). The "absolute" energy of the vacuum state is not an observable. We can then "redefine" the energy of the vacuum in such a way that

$$H|0\rangle = 0\,, \tag{1.119}$$

[12] We use the following representation of the Dirac delta:

$$\frac{1}{L}\sum_{j,k} e^{i\frac{2\pi}{L}j(x-y)} = \delta(x-y)\,, \tag{1.110}$$

that can be "proved" looking at its action on a generic function $f(x) = \sum_k c_k e^{i\frac{2\pi}{L}kx}$:

$$\begin{aligned}
\int_0^L \frac{1}{L}\sum_j e^{i\frac{2\pi}{L}j(x-y)}f(x)dx &= \int_0^L \frac{1}{L}\sum_{j,k} e^{i\frac{2\pi}{L}j(x-y)}c_k e^{i\frac{2\pi}{L}kx} = \int_0^L \frac{1}{L}\sum_{j,k} e^{i\frac{2\pi}{L}x(j+k)}e^{-i\frac{2\pi}{L}yj}c_k = \Big|\ j\to -j\ \Big|,\\
&= \int_0^L \frac{1}{L}\sum_{j,k} e^{i\frac{2\pi}{L}x(k-j)}e^{i\frac{2\pi}{L}yj}c_k = \Big|\ \text{since } \frac{1}{L}\int_0^L e^{i\frac{2\pi}{L}x(k-j)}\delta_{jk}\ \Big| = \sum_k c_k e^{i\frac{2\pi}{L}ky} = f(y)\,.
\end{aligned}$$

removing (so to say) the infinite constant value and imposing that the energy of the vacuum is simply zero (this is also needed for Lorentz invariance).

Formally, this operation is achieved defining the "normal ordering" of the operator H. This is indicated with the singn $:H:$ and defined as follows:

$$:H:= H - \langle 0|H|0\rangle \,, \tag{1.120}$$

or, as a rule, puting all the creation operators in the expression on the left of the annihilation operators (respecting the statistics).

Capitolo 2

Special Relativity

2.1 Notes on Special Relativity

The concept of finding a class of physical frames in which one can write physics laws in a unique formal way goes back to Newtonian mechanics and it was introduced by Galileo with the Principle of Inertia. It can be refrased as: *In every inertial frame (IF), Physics is described by the same (in form) equation* $\mathbf{F} = m\mathbf{a}$.

This principle, with the additional constraint of the "universal time", brings to a class of transformations, the Galilean transformations (GT)

$$\mathbf{x}'(t') = \mathbf{x}(t) - \mathbf{v}_0\, t\,, \tag{2.1}$$

$$t' = t\,, \tag{2.2}$$

that leave unchanged the equations of motion, $\mathbf{F} = m\mathbf{a}$.

NB Galilean Relativity Principle (GRP) is adapted to Newtonian mechaniccs. It does not take into account Classical Electrodynamics. Maxwell's equations

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\phi = \rho\,, \tag{2.3}$$

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{A} = \frac{1}{c}\mathbf{j}\,, \tag{2.4}$$

are not invariant under GT (but under Lorentz transformations). The point lies on the fact that we see experimentally that the speed of light, $c = \frac{1}{\sqrt{\epsilon_0\mu_0}} \simeq 3 \cdot 10^8\ \mathrm{ms}^{-1}$, is a universal constant, with the same value in every inertial frame. In Maxwell's equations c appears explicitely! Therefore, they cannot be invariant under galilean transformations. The composition of velocities is totally different in the two cases.

At the beginning of XX^{th} century physicists have to understand which one, among the following three options, is the correct one:

1. It exists a "Relativity Principle" for Mechanics, but not for Electrodynamics, and Electrodynamics changes in every inertial frame (System of Eather ...).

2. It exists a unique "Relativity Principle" both for Mechanics and Electrodynamics, the GRP, and therefore Maxwell's equations are wrong.

3. It exists a unique "Relativity Principle" both for Mechanics and Electrodynamics, and GT are only a low-speed limit of more complex invariance transformations, and $\mathbf{F} = m\mathbf{a}$ is a low-speed limit of a formulation of Mechanics which is covariant under a new class of transformations, Lorentz transformations.

The third hypothesis revealed to be the correct one. Based on electromagnetism, mechanics was reformulated throught a redefinition of the concept of time.

The theory of Special Relativity, formalized by Einstein in 1905, is based on the following two postulates:

- Physical laws are the same in every *inertial frame*.
- The speed of light is the same in every inertial frame, and the relation $c = \frac{1}{\sqrt{\epsilon_0 \mu_0}}$ applies.

The second postulate brings to the criticism of the concept of *simultaneous events*, that now has to depend on the reference system. The *absolute time*, à la Newton, loses meaning and it emerges the necessity to consider time on the same ground as space coordinates.

2.1.1 Simultaneous events

In a given inertial frame, physical phenomena are analyzed in terms of *events*: the physical phenomenon "happens" in a certain point $\mathbf{x}$ at a certain time t. The event is indicated with the following vector: $(ct, \mathbf{x})$ in Minkowski space (see later).

Definition 2.1.1 *In a given IF we say that two events are simultaneous if they happen in two different space points and two light rays moving from each point in the direction of the other meet at half of the distance.*

It is clear that, if the speed of light is the same in every IF, a pair of events that are simultaneous in an IF cannot be the same in another IF.

Suppose that in a given IF, S, a light signal is emitted from $P_1 = (x_1, y_1, z_1)$ at t_1 and reaches $P_2 = (x_2, y_2, z_2)$ at t_2. Since the speed of light is c, we will have

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 + z_2)^2 = c^2(t_1 - t_2)^2 \,. \tag{2.5}$$

If S' is another IF in which at t_1 P_1 coincides with P'_1 and P_2 with P'_2, since c is the same in both IF we will have

$$(x'_1 - x'_2)^2 + (y'_1 - y'_2)^2 + (z'_1 + z'_2)^2 = c^2(t'_1 - t'_2)^2 \,. \tag{2.6}$$

Definition 2.1.2 *The expression*

$$\Delta s_{12}^2 = c^2(t_1 - t_2)^2 - (x_1 - x_2)^2 - (y_1 - y_2)^2 - (z_1 + z_2)^2 \,, \tag{2.7}$$

can be taken as the interval between the two events in S and it is a relativistic invariant.

If $\Delta s_{12}^2 = 0$ in S, we have $\Delta {s'_{12}}^2 = 0$ in S'.

Definition 2.1.3 *The infinitesimal interval will be*

$$ds^2 = c^2\, dt^2 - dx^2 - dy^2 - dz^2 \,. \tag{2.8}$$

Property 2.1.4 *ds^2 is a relativistic invariant.*

Since $ds^2 = 0$ implies $ds'^2 = 0$, it means that they are infinitesimals of the same order. We can put

$$ds^2 = a\, ds'^2 \,. \tag{2.9}$$

Since we want the space-time to be homogeneous and isotropic, a cannot depend on X^μ, and neither on the vector $\mathbf{v}$, relative velocity of the frame S' with respect to S. It could depend on $v = |\mathbf{v}|$, modulus of the relative velocity. However, let us consider three reference systems, S_1, S_2 and S_3. S_2 moves with respect to S_1 with velocity $\mathbf{v}_2$ and S_3 moves with respect to S_1 with velocity $\mathbf{v}_3$. S_3 will move with respect to S_2 with velocity $\mathbf{v}_{23}$. Therefore, we have

$$ds_1^2 = a(v_2)\, ds_2^2 \,, \qquad ds_1^2 = a(v_3)\, ds_3^2 \,, \qquad ds_2^2 = a(v_{23})\, ds_3^2 \,. \tag{2.10}$$

Taking the ratio of the first two equations we have that

$$ds_2^2 = \frac{a(v_3)}{a(v_2)}\, ds_3^2 \,, \tag{2.11}$$

but also (considering the third one)

$$ds_2^2 = a(v_{23})\, ds_3^2 \,. \tag{2.12}$$

Therefore

$$\frac{a(v_3)}{a(v_2)} = a(v_{23}) \,. \tag{2.13}$$

The r.h.s. depends on v_2, v_3 but also on the directions of $\mathbf{v}_2$ and $\mathbf{v}_3$ (v_{23} is the modulus of the relative velocity), while the l.h.s. doers not depend on the directions, it means that they should be constants and that therefore $a = 1$. In the end

$$ds^2 = ds'^2 \,. \tag{2.14}$$

NOTE: the interval Δs^2 is not positive definite (as it is instead in the Euclidean case) but it can be > 0, < 0 or $= 0$.

1. If there exists an inertial frame in which the two events happen at the same spatial point, but at subsequent times, in that frame we must have

   $$\Delta s'^2 = c^2 \Delta t'^2 > 0 \,. \tag{2.15}$$

   Since Δs^2 is a relativistic invariant, in another frame we will have, in any case,

   $$\Delta s^2 = c^2 \Delta t^2 - \Delta l^2 = \Delta s'^2 > 0 \,. \tag{2.16}$$

   We call this interval a **time-like interval**. In this case the two events can be connected by a causal-effect relationship.

2. If there exists an inertial frame in which the two events happen at two different spatial points, but at the same time, in that frame we must have

   $$\Delta s'^2 = -\Delta l'^2 < 0 \,. \tag{2.17}$$

   Since Δs^2 is a relativistic invariant, in another frame we will have, in any case,

   $$\Delta s^2 = c^2 \Delta t^2 - \Delta l^2 = \Delta s'^2 < 0 \,. \tag{2.18}$$

   We call this interval a **space-like interval**. In this case the two events cannot be causally connected.

3. Finally, an interval for which $\Delta s^2 = 0$ is called **light-like**.

The property to be time-like, space-like or light-like is a characteristic of the vector and does not depend on the inertial frame.

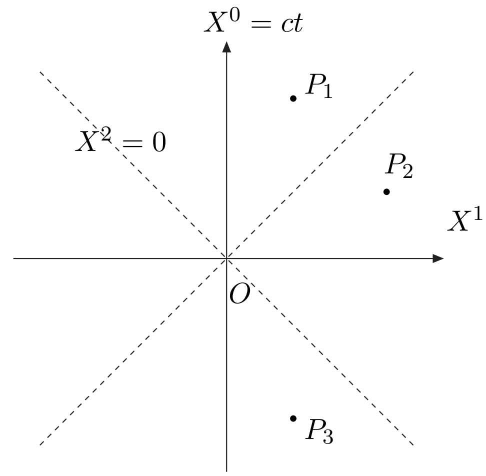


Figura 2.1: Minkowski space and the light cone.

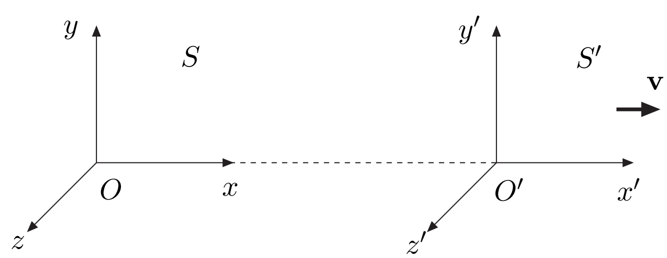
2.1.2 Causal structure of the Space-Time

Let us consider an event O in the space-time as the origin of our frame (for simplicity of representation we consider a 1+1 dimensional Minkowski space) as in Fig. 2.1. The events for which $c^2\Delta t^2 = \Delta l^2$, reported as $X^2 = 0$, are represented in the diagram as streight lines at 45° and they are characteristic of the propagation of light. This means that an event on one of these streight lines is connected to the origin O by a signal that travels at the speed of light. The region between the two lines at 45° is called the **light cone** ("cone" because in more dimensions is a cone). Events within the light cone, like P_1 or P_3, have time-like distance from O and, therefore, they can be in causal relationship with O. P_3 "happens" before O, while P_1 after. We can make a Lorentz transformation to a frame in which P_3 and O happens in the same spatial place but at two subsequent instants. In the same way we can find a frame in which O and P_1 happens in the same spatial place but at two subsequent instants. The region outside the light cone cannot be connected causally with events within the light cone. In fact, a signal from P_3, for instance, in order to reach O would have to travel at a speed bigger than the speed of light and this is not possible. We can find a frame in which O and P_3 happen simultaneously in two separate space points (Note: in the case of space-like separations, we can also find a frame in which the temporal succession of the two events is inverted).

2.1.3 Lorentz transformations: Boosts

Let us consider two inertial frames, S and S'. S', for instance, will move with respect to S with constant velocity $\mathbf{v}$. Knowing the coordinates of one event in S, say (ct, x, y, z), we would like to find the transformation laws that allow us to represent the same event in S', (ct', x', y', z').

Let us suppose for simplicity that S' moves with respect to S with a translation in the x direction



and that the axis x and x' coincide. If an event has coordinates X^μ in S, its coordinates X'^μ in S' will be:

$$\begin{cases} X'^0 &= \gamma X^0 - \beta\gamma X^1 \\ X'^1 &= -\beta\gamma X^0 + \gamma X^1 \\ X'^2 &= X^2 \\ X'^3 &= X^3 \end{cases} \tag{2.19}$$

where $\beta = \frac{v}{c}$ and $\gamma = \frac{1}{\sqrt{1-\beta^2}}$. Eq. (2.19) can be written in the following way:

$$X'^\mu = \Lambda^\mu_\nu X^\nu , \tag{2.20}$$

where we used the Lorentz transformation Λ^μ_ν, that can be written in matrix form as follows:

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{2.21}$$

Another way to parametrize Λ is through hyperbolic functions. In fact, we know that $\gamma^2(1-\beta^2) = 1$. Therefore, we can define an imaginary angle ϕ, such that

$$\gamma = \cosh\phi , \quad \text{e} \quad \gamma\beta = \sinh\phi . \tag{2.22}$$

Then, we can write:

$$\Lambda = \begin{pmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{2.23}$$

The inverse Lorentz transformation is the one that gives the coordinates X^μ in S, knowing X'^μ in S and can be found immediately inverting the velocity in (2.21)

$$\begin{cases} X^0 &= \gamma X'^0 + \beta\gamma X'^1 \\ X^1 &= \beta\gamma X^0 + \gamma X'^1 \\ X^2 &= X'^2 \\ X^3 &= X'^3 \end{cases} \tag{2.24}$$

and then

$$X^\nu = \Lambda_\mu^{\cdot\nu} X'^\mu , \tag{2.25}$$

where now $(\Lambda^{-1})^\nu_\mu = \Lambda_\mu^{\cdot\nu}$ is such that

$$\Lambda^{-1} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.26}$$

and

$$\Lambda_\rho^{\cdot\nu} \Lambda^\mu_\nu = \eta^\mu_\rho = \delta^\mu_\rho , \tag{2.27}$$

that can be checked multiplying Eq. (2.21) times Eq. (2.26) and using $\gamma^2 - \beta^2\gamma^2 = 1$:

$$\Lambda\Lambda^{-1} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{2.28}$$

Note that the boost defined in this way is particularly simple, in the sense that the translation with velocity v is done in the x direction only, $\mathbf{v} = v\,\hat{\mathbf{i}}$. In general we can have the velocity directed in a general direction (we will write the general boost below). Moreover, it can happen that the frame S' does not have the axis x', y' and z' parallel to x, y and z. In this case we can rotate S' to make in such a way that the axis become parallel to the axis of S. This corresponds to an isometry in the three-dimensional space. Since the time is not affected, only $d\mathbf{x}$ will change. However, since we are speaking about an isometry, we will have $|d\mathbf{x}'| = |d\mathbf{x}|$ and therefore, in the end, $ds'^2 = c^2dt^2 - dx'^2 = c^2dt^2 - dx^2 = ds^2$. This means that the spatial rotations are part of the Lorents transformations (they leave ds^2 unchanged). A Lorentz transformation is a composition of a rigid rotation and a boost in the $\mathbf{v}$ direction.

Non relativistic limit

If we consider the limit in which

$$\frac{v}{c} \ll 1\,, \tag{2.29}$$

we have

$$\frac{1}{\sqrt{1-\beta^2}} \simeq 1 + \frac{1}{2}\beta^2 + ... \tag{2.30}$$

and therefore at zeroth order in β we find

$$\begin{cases} t' & \simeq & t \\ x' & \simeq & x - vt \\ y' & = & y \\ z' & = & z \end{cases} \tag{2.31}$$

i.e. the Galilean transformations.

2.1.4 Boost in a general direction

The boost in the x direction shows the general feature that only the components in the direction of the velocity, and the time, are affected by the transformation. The components perpendicular to the direction of the velocity are not. We can write a boost in a general direction $\hat{\mathbf{v}}$, decomposing the vector $\mathbf{X}$ as the sum of two vectors: one parallel and the other perpendicular to $\hat{\mathbf{v}}$.

$$\mathbf{X} = \mathbf{X}_\parallel + \mathbf{X}_\perp\,, \qquad \mathbf{X}' = \mathbf{X}'_\parallel + \mathbf{X}'_\perp\,. \tag{2.32}$$

Then, the boost can be written as follows:

$$\begin{cases} X'^0 & = & \frac{X^0 - \beta\, X_\parallel}{\sqrt{1-\beta^2}} = \frac{X^0 - \beta\,\mathbf{X}\cdot\hat{\mathbf{v}}}{\sqrt{1-\beta^2}} \\ \mathbf{X}'_\parallel & = & \frac{\mathbf{X}_\parallel - \beta\, X^0\hat{\mathbf{v}}}{\sqrt{1-\beta^2}} \\ \mathbf{X}'_\perp & = & \mathbf{X}_\perp\,. \end{cases} \tag{2.33}$$

2.1.5 Transformation of the three-velocity

It is interesting to look at the composition of velocities in special relativity. We will demonstrate that, if $\mathbf{u}$ is the three-velocity of a material point that moves with respect to the observer in S and the observer in S' moves with respect to S with a velocity $\mathbf{v}$, the velocity $\mathbf{u}'$ of the point seen by S' is such that if $|\mathbf{u}'| \leq c$ and $|\mathbf{v}| \leq c$ then $|\mathbf{u} \leq c$. For simplicity we consider a boost in the x direction. We have

$$\begin{cases} t' &= \gamma\left(t - \frac{v}{c^2}x\right) \\ x' &= \gamma\left(x - vt\right) \\ y' &= y \\ z' &= z \end{cases} \tag{2.34}$$

and the inverse transformation given by

$$\begin{cases} t &= \gamma\left(t' + \frac{v}{c^2}x'\right) \\ x &= \gamma\left(x' + vt'\right) \\ y &= y' \\ z &= z' \end{cases} \tag{2.35}$$

Let us consider the three-velocity $\mathbf{u}$ of the material point in components. We have

$$u_x = \frac{dx}{dt} = \frac{dx}{dt'}\frac{dt'}{dt} = \frac{d}{dt'}\left(\frac{x' + vt'}{\sqrt{1-\beta^2}}\right)\frac{d}{dt}\left(\frac{t - \frac{v}{c^2}\,x}{\sqrt{1-\beta^2}}\right) = \frac{u_x' + v}{\sqrt{1-\beta^2}}\frac{1 - \frac{v}{c^2}\,u_x}{\sqrt{1-\beta^2}}\,. \tag{2.36}$$

From Eq. (2.36) we find

$$u_x(1-\beta^2) = u_x' - \frac{v}{c^2}u_x u_x' + v - \beta^2 u_x\,, \tag{2.37}$$

and therefore

$$u_x = \frac{u_x' + v}{1 + \frac{v}{c^2}\,u_x'}\,. \tag{2.38}$$

For the component in y we have

$$u_y = \frac{dy}{dt'}\frac{dt'}{dt} = \ldots = \frac{u_y'\sqrt{1-\beta^2}}{1 + \frac{v}{c^2}\,u_x'} \tag{2.39}$$

and for u_z

$$u_z = \frac{dz}{dt'}\frac{dt'}{dt} = \ldots = \frac{u_z'\sqrt{1-\beta^2}}{1 + \frac{v}{c^2}\,u_x'}\,. \tag{2.40}$$

In summary

$$\begin{cases} u_x &= \frac{u_x' + v}{1 + \frac{v}{c^2}\,u_x'}\,, \\ u_y &= \frac{u_y'\sqrt{1-\beta^2}}{1 + \frac{v}{c^2}\,u_x'}\,, \\ u_z &= \frac{u_z'\sqrt{1-\beta^2}}{1 + \frac{v}{c^2}\,u_x'}\,. \end{cases} \tag{2.41}$$

Note that if $c \to \infty$ (or better if we consider the limit $v/c \ll 1$) we find the "euclidean" composition of the velocities

$$\begin{cases} u_x &\simeq u_x' + v\,, \\ u_y &\simeq u_y'\,, \\ u_z &\simeq u_z'\,. \end{cases} \tag{2.42}$$

Let us assume that the point moves in the x direction only (for simplicity) and that $u'_x \le c$, $v \le c$. Then we have

$$(c - u'_x) \ge 0\,, \quad \text{and} \quad (c - v) \ge 0\,. \tag{2.43}$$

It follows that

$$(c - u'_x)(c - v) = c^2 - cv - u'_x c + v u'_x \ge 0 \tag{2.44}$$

and then (deviding by c^2 which is $\neq 0$ and positive)

$$1 + \frac{u'_x v}{c^2} \ge \frac{u'_x + v}{c}\,. \tag{2.45}$$

Looking at the x component in Eq. (2.41), we get

$$u'_x + v = u_x \left(1 + \frac{v}{c^2}\, u'_x\right) \ge u_x \left(\frac{u'_x + v}{c}\right) \tag{2.46}$$

and therefore

$$u_x \le c\,. \tag{2.47}$$

If $u'_x = c$, we find immediately that

$$u_x = \frac{u'_x + v}{1 + \frac{v}{c^2}\, u'_x} = \frac{c + v}{1 + \frac{v}{c}} = c\,. \tag{2.48}$$

2.2 Kinematics of the classical particle

We want now to describe the kinematics and the dynamics of a point-like massive particle in a covariant way. The goal is to be able to re-write the second principle of dynamics in a manifetly covariant way, using tensor relations, in such a way that for $v \ll c$ we can recover Newtonian mechanics.

2.2.1 Four-velocity and four-acceleration

In newtonian mechanics we introduce the velocity of the particle as $\mathbf{v} = \frac{d\mathbf{x}}{dt}$. An obvious relativistic generalization of the $d\mathbf{x}$ is dX^μ. However, dt is not a relativistic invariant, and therefore $\frac{dX^\mu}{dt}$ does not transform as a four-vector. We should find an invariant that can replace dt.

We know that

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 \tag{2.49}$$

is an invariant. Let us then perform a LT to an inertial frame in which $dx' = dy' = dz' = 0$. If we rename τ the time in that frame, i.e. the time in the frame in which the particle is at rest, we have

$$d\tau = dt\sqrt{1 - \beta^2} \tag{2.50}$$

and

$$ds^2 = c^2 d\tau^2 \tag{2.51}$$

then also $d\tau$ is an invariant. τ is called the *proper time* of the particle.

Let us consider now the following vector (with the dimensions of a velocity)

$$\mathcal{U}^\mu = (\mathcal{U}^0, \underline{\mathcal{U}}) = \frac{dX^\mu}{d\tau} = \gamma\, \frac{dX^\mu}{dt}\,. \tag{2.52}$$

$\mathcal{U}^\mu$ is indeed a four-vector, since dX^μ is a four-vector and $d\tau$ is an invariant. We have

$$\mathcal{U}^0 \;=\; \frac{1}{\sqrt{1 - \beta^2}} \frac{dX^0}{dt} = \frac{c}{\sqrt{1 - \beta^2}}\,, \tag{2.53}$$

$$\underline{\mathcal{U}} \quad = \quad \frac{1}{\sqrt{1-\beta^2}}\frac{d\mathbf{X}}{dt} = \frac{\mathbf{v}}{\sqrt{1-\beta^2}}\,. \tag{2.54}$$

$\mathcal{U}^\mu$ is a time-like vector, since

$$\mathcal{U}^\mu\mathcal{U}_\mu = \left(\frac{c}{\sqrt{1-\beta^2}}\right)^2 - \left(\frac{\mathbf{v}}{\sqrt{1-\beta^2}}\right)^2 = c^2 > 0\,. \tag{2.55}$$

Following on the same line, we can define the four-acceleration

$$\mathcal{A}^\mu = \frac{d\mathcal{U}^\mu}{d\tau} = \gamma\frac{d\mathcal{U}^\mu}{dt}\,. \tag{2.56}$$

The components of $\mathcal{A}^\mu$ are

$$\mathcal{A}^0 \quad = \quad \gamma\frac{d\mathcal{U}^0}{dt} = \dots = \frac{\mathbf{v}\cdot\mathbf{a}}{c(1-\beta^2)}\,, \tag{2.57}$$

$$\underline{\mathcal{A}} \quad = \quad \gamma\frac{d\underline{\mathcal{U}}}{dt} = \dots = \frac{\mathbf{a}}{(1-\beta^2)} + \frac{\mathbf{v}\cdot\mathbf{a}}{c^2(1-\beta^2)^2}\mathbf{v}\,, \tag{2.58}$$

where $\mathbf{v} = \frac{d\mathbf{x}}{dt}$ and $\mathbf{a} = \frac{d\mathbf{v}}{dt}$. Note that for $c \to \infty$ we have that the temporal component of $\mathcal{A}^\mu$ goes to zero, while the spatial component becomes $\mathbf{a}$, the usual non relativistic acceleration.

2.2.2 Four-momentum

In newtonian mechanics an important quantity is the momentum of the particle, which ia defined as $\mathbf{p} = m\mathbf{v}$. A covariant form of $\mathbf{p}$ can be constructed in the following way

$$P^\mu = m\mathcal{U}^\mu\,, \tag{2.59}$$

where m coincides with the inertial mass of the particle when $v \ll c$. We have

$$P^\mu = \left(\frac{mc}{\sqrt{1-\beta^2}}, \frac{m\mathbf{v}}{\sqrt{1-\beta^2}}\right)\,, \tag{2.60}$$

which is called the **energy-momentum** four-vector. P^μ is such that

$$P^2 = P^\mu P_\mu = \frac{m^2c^2}{1-\beta^2} - \frac{m^2v^2}{1-\beta^2} = m^2c^2 > 0\,. \tag{2.61}$$

It is a time-like vector. The relation $P^\mu P_\mu = m^2c^2$ is called the mass-shell relation. Since the lagrangian of the free particle is

$$L = -mc^2\sqrt{1-\beta^2}\,, \tag{2.62}$$

such that the three-momentum $\mathbf{p}$ is actually

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1-\beta^2}}\,, \tag{2.63}$$

we can look at the energy, performing a Legendre transformation to

$$E = \mathbf{p}\cdot\mathbf{v} - L = \frac{mc^2}{\sqrt{1-\beta^2}}\,. \tag{2.64}$$

Note that even if $\mathbf{v} = \mathbf{0}$ the energy of the free particle is not zero, but

$$E \overset{v\to 0}{\to} mc^2\,. \tag{2.65}$$

Using Eq. (2.64), we have

$$P^\mu = \left(\frac{mc}{\sqrt{1-\beta^2}}, \frac{m\mathbf{v}}{\sqrt{1-\beta^2}} \right) = \left(\frac{E}{c}, \mathbf{p} \right) \tag{2.66}$$

and from the mass-shell relation we have

$$\frac{E^2}{c^2} = |\mathbf{p}|^2 + m^2 c^2 \,. \tag{2.67}$$

2.3 Vectors and Tensors

After the introduction of Lorentz transformations, we now want to study how mathematical objects, that will be used to describe our Physics, transform under Lorentz transformations (LT). This is the subject of Tensor Analysis.

Let us start introducing a more general definition of vectors in a Euclidean space.

2.3.1 Vectors and Contravariant Components

In Special Relativity (SR) we have to deal with different kind of vectors. The fact that in Newtonian mechanics, for instance, we do need just the usual Euclidean definition is simply due to the fact that usually we use an ortonormal system of basis vectors for the vectorial space. In this situation the metric tensor reduces to a Kronecker delta function and it becomes impossible to appreciate the difference between different definitions of vectors.

Let us consider a vector space $\mathcal{V}$ on $\mathbb{R}$. Let $\{\mathbf{e}_i\}$ is a set of independent vectors which constitutes a basis for $\mathcal{V}$.

If $\mathbf{v} \in \mathcal{V}$, it can be expressed as a linear combination of the basis vectors

$$\mathbf{v} = v^i \mathbf{e}_i \,, \quad \text{with} \quad i = 1, \cdots, dim(\mathcal{V}) \,. \tag{2.68}$$

The real numbers v^i are called the *contravariant components* of $\mathbf{v}$. The place of the index i, as superscript is relevant. As we will see in a moment, components with an index as subscript describe a different kind of vector.

Let us consider now a different basis of $\mathcal{V}$, $\{\mathbf{e}'_i\}$ and let Λ be the transformation from the old to the new basis. We have

$$\boxed{\mathbf{e}'_i = \Lambda^j_i \mathbf{e}_j} \tag{2.69}$$

Note that the index j of $\mathbf{e}_j$ is contracted with the upper index of Λ. Under basis transformation, the components of $\mathbf{v}$ tranform accordingly. The transformation law is the following. Remember that the vector $\mathbf{v}$ is an absolute quantity, that can be represented using different basis. But $\mathbf{v}$ is always the same vector. Therefore, in the new basis we will have

$$\mathbf{v} = v'^i \mathbf{e}'_i = v'^i \Lambda^j_i \mathbf{e}_j \,, \tag{2.70}$$

but we can also write

$$\mathbf{v} = v^j \mathbf{e}_j \,, \tag{2.71}$$

and matching Eq. (2.70) and Eq. (2.71) we find

$$v^j = \Lambda^j_i v'^i \,. \tag{2.72}$$

Note that the index of v'^i is contracted with the lower index of Λ (it goes with the transposed). Multiplying Eq. (2.72) by $\left(\Lambda^{-1}\right)^l_j$ on the l.h and r.h.s, we have

$$\left(\Lambda^{-1}\right)^l_j v^j = \left(\Lambda^{-1}\right)^l_j \Lambda^j_i v'^i = \left(\Lambda^{-1}\Lambda\right)^l_i v'^i = \delta^l_i v'^i = v'^l \,. \tag{2.73}$$

Finally

$$\boxed{v'^l = \left(\Lambda^{-1}\right)^l_j v^j} \tag{2.74}$$

Therefore, if the basis transforms with Λ, the contravariant components of $\mathbf{v}$ transform with the inverse transposed of Λ, $\left(\Lambda^T\right)^{-1} = \left(\Lambda^{-1}\right)^T$.

In matrix notation

$$\mathbf{v} = \Lambda^T \mathbf{v}' , \tag{2.75}$$

$$\mathbf{v}' = \left(\Lambda^T\right)^{-1} \mathbf{v} = \left(\Lambda^{-1}\right)^T \mathbf{v} , \tag{2.76}$$

2.3.2 Dual vectors and covariant components

Once the vectorial space $\mathcal{V}$ is defined, it is automatically defined also the "*dual*" space, $\mathcal{V}^*$, which is the vectorial space of linear funtionals on $\mathcal{V}$:

$$\sigma : \mathcal{V} \to \mathbb{R} , \tag{2.77}$$

$$\mathbf{v} \to \sigma(\mathbf{v}) . \tag{2.78}$$

Since $\mathcal{V}^*$ is a vectorial space, we can find a basis $\{\mathbf{k}^i\}$ in which the functional σ can be represented in a unique way as

$$\sigma = \sigma_i \mathbf{k}^i . \tag{2.79}$$

the set σ_i are real numbers that represent the components of σ in this basis.

Although $\mathcal{V}$ and $\mathcal{V}^*$ are different spaces, they are connected. They have the same dimensionality and they are isomorfic, but *they are different*! If the basis changes in $\mathcal{V}$, this will imply a change of basis of $\mathcal{V}^*$. Therefore, we can ask how the components of σ behave under the basis transformation in Eq. (2.69). We labeled the components of σ in the $\{\mathbf{k}^i\}$ basis with a lower index because the properties of these components under a basis tranformation in $\mathcal{V}$ are different from those of the contravariant components of a vector in $\mathcal{V}$.

Using (2.79), we can write

$$\sigma(\mathbf{v}) = \sigma_i \mathbf{k}^i(\mathbf{v}) = \sigma_i \mathbf{k}^i(v^j \mathbf{e}_j) = \sigma_i v^j \mathbf{k}^i(\mathbf{e}_j) . \tag{2.80}$$

The number $\mathbf{k}^i(\mathbf{e}_j)$ tells how the components of the basis in the functional space $\mathcal{V}^*$ act on the components of the base in $\mathcal{V}$. We say that the two chosen basis are "*dual*" when we have

$$\mathbf{k}^i(\mathbf{e}_j) = \delta^i_j , \tag{2.81}$$

with δ the Kronecker delta $\delta^i_i = 1$, $\delta^i_j = 0$ if $i \neq j$. In this case the situation is much simpler and we have

$$\sigma(\mathbf{v}) = \sigma_i v^i . \tag{2.82}$$

Note that (2.82) is not a scalar product! It is the sum of the product of the corresponding components of σ and $\mathbf{v}$, vectors that belong to two different vector spaces.

Let us consider dual bases. If we apply the basis functionals $\{\mathbf{k}^i\}$ to the vector $\mathbf{v} \in \mathcal{V}$ we have

$$\mathbf{k}^i(\mathbf{v}) = \mathbf{k}^i(v^j \mathbf{e}_j) = v^j \, \mathbf{k}^i(\mathbf{e}_j) = v^i , \tag{2.83}$$

because of (2.81). Therefore, the action of $\mathbf{k}^i$ on $\mathbf{v}$ is to extract its contravariant component. On the other hand, we have

$$\sigma(\mathbf{e}_j) = \sigma_i \, \mathbf{k}^i(\mathbf{e}_j) = \sigma_j , \tag{2.84}$$

because of (2.79).

If we consider the change of basis (2.69), it will imply a change of basis in $\mathcal{V}^*$, say from $\mathbf{k}^i$ to $\mathbf{k'}^i$. In $\mathbf{k'}^i$ the expression of σ will be given by

$$\sigma = \sigma'_i \, \mathbf{k'}^i \,. \tag{2.85}$$

We have, because of (2.84)

$$\sigma'_i = \sigma(\mathbf{e'}_i) = \sigma_j \mathbf{k}^j(\mathbf{e'}_i) = \sigma_j \mathbf{k}^j(\Lambda^l_i \mathbf{e}_l) = \sigma_j \Lambda^l_i \mathbf{k}^j(\mathbf{e}_l) = \sigma_j \Lambda^l_i \delta^j_l = \sigma_j \Lambda^j_i \,. \tag{2.86}$$

In summary

$$\boxed{\sigma'_i = \Lambda^j_i \sigma_j} \tag{2.87}$$

and the components σ_i transform according to the transformation of basis (as in (2.69)). That is why they are called "*covariant*" components.

Scalar Product and Metric Tensor

Just to have in mind a practical example, let us introduce the *scalar product* and refrase what we just said in this case.

The *scalar product* between two vectors of $\mathcal{V}$ is an application of $\mathcal{V} \times \mathcal{V} \to \mathbb{R}$, which is bilinear, symmetric and not degenerate

$$\mathbf{v}, \mathbf{w} \, \in \mathcal{V} \to (\mathbf{v}, \mathbf{w}) \in \mathbb{R} \,. \tag{2.88}$$

The scalar product induces a norm on $\mathcal{V}$, that in turn induces a metric. Therefore, with a scalar product our vector space becomes a metric space.

Let us fix the first vector $\mathbf{v}$ and consider the scalar product with every other vector $\mathbf{w} \in \mathcal{V}$. In this case we defined a functional $\mathbf{f_v} = (\mathbf{v}, .)$ such that

$$\mathbf{w} \, \in \mathcal{V} \to \mathbf{f_v}(\mathbf{w}) = (\mathbf{v}, \mathbf{w}) \in \mathbb{R} \,. \tag{2.89}$$

$\mathbf{f_v}$ is formally a vector of the dual space of $\mathcal{V}$, $\mathcal{V}^*$. We can choose a basis in $\mathcal{V}^*$. Let us call it $\{\mathbf{k}^i\}$, with the index i as superscript. Therefore, $\mathbf{f_v}$ will be expressed in a unique way in this basis:

$$\mathbf{f} = f_i \mathbf{k}^i \,. \tag{2.90}$$

Because of the definition (2.89), in this case we have

$$f_i = \mathbf{f}(\mathbf{e_i}) = (\mathbf{v}, \mathbf{e}_i) = v_i \,. \tag{2.91}$$

We call them *covariant* components of $\mathbf{v}$.

We have, in particular

$$\mathbf{f_v}(\mathbf{w}) = (\mathbf{v}, \mathbf{w}) = v^i \, (\mathbf{e}_i, \mathbf{w}) = v^i w^j \, (\mathbf{e}_i, \mathbf{e}_j) = v^i w^j \, g_{ij} \,, \tag{2.92}$$

where we introduced the "*metric tensor*"

$$g_{ij} = (\mathbf{e}_i, \mathbf{e}_j) \,. \tag{2.93}$$

The metric tensor is symmetric (by construction). If it is also positive definite, there is a theorem that proves that with a change of basis we can find

$$g_{ij} = (\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij} \,. \tag{2.94}$$

In this case we see that covariant and contravariant components are exactly the same. If the metric tensor is not positive definite, then they are diferent and related by the metric tensor

$$v_i = g_{ij} v^j \,. \tag{2.95}$$

This is the case of Special Relativity.

Since the covariant components of $\mathbf{v}$ are defined as

$$v_i = (\mathbf{v}, \mathbf{e}_i) = (v^j \mathbf{e}_j, \mathbf{e}_i) = v^j (\mathbf{e}_j, \mathbf{e}_i) = g_{ij} v^j \,, \tag{2.96}$$

under basis transformation, as in Eq. (2.69), they change according to the following relation

$$v'_i = (\mathbf{v}, \mathbf{e'}_i) = (v^j \mathbf{e}_j, \Lambda_i^\rho \mathbf{e}_\rho) = v^j \Lambda_i^\rho (\mathbf{e}_j, \mathbf{e}_\rho) = v^j \Lambda_i^\rho g_{j\rho} = \Lambda_i^\rho v_\rho \,. \tag{2.97}$$

In summary

$$\boxed{v'_i = \Lambda_i^j v_j} \tag{2.98}$$

Now we can ask how the metric tensor transforms under (2.69)? We have

$$g'_{ij} = (\mathbf{e'}_i, \mathbf{e'}_j) = (\Lambda_i^\rho \mathbf{e}_\rho, \Lambda_j^\sigma \mathbf{e}_\sigma) = \Lambda_i^\rho \Lambda_j^\sigma (\mathbf{e}_\rho, \mathbf{e}_\sigma) = \Lambda_i^\rho \Lambda_j^\sigma g_{\rho\sigma} \,. \tag{2.99}$$

Therefore

$$\boxed{g'_{ij} = \Lambda_i^\rho \Lambda_j^\sigma g_{\rho,\sigma}} \,. \tag{2.100}$$

A two-indices object, g_{ij}, that transforms like in Eq. (2.100) is called a *covariant tensor of rank 2*.

Using Eq. (2.100) we can show that the scalar product is an absolute quantity, that does not depend on the chosen basis. In fact

$$\begin{aligned}
(\mathbf{u'}, \mathbf{v'}) &= g'_{\mu\nu} u'^\mu v'^\nu = \Lambda_\mu^\rho \Lambda_\nu^\sigma g_{\rho\sigma} u'^\mu v'^\nu = \Lambda_\mu^\rho \Lambda_\nu^\sigma g_{\rho\sigma} \left(\Lambda^{-1}\right)_\gamma^\mu \left(\Lambda^{-1}\right)_\delta^\nu u^\gamma v^\delta && (2.101)\\
&= \left(\Lambda\Lambda^{-1}\right)_\gamma^\rho \left(\Lambda\Lambda^{-1}\right)_\delta^\sigma u^\gamma v^\delta = \delta_\gamma^\rho \delta_\delta^\sigma u^\gamma v^\delta = g_{\gamma\delta} u^\gamma v^\delta && (2.102)\\
&= (\mathbf{u}, \mathbf{v}) \,. && (2.103)
\end{aligned}$$

We can also define the inverse of the matric tensor (the contravariant version of the metric tensor) $g^{\mu\nu}$ such that

$$g^{\mu\nu} g_{\nu\rho} = \delta_\rho^\mu = \delta_\mu^\rho = g_{\mu\nu} g^{\nu\rho} \tag{2.104}$$

and

$$(\mathbf{u}, \mathbf{v}) = g_{\mu\nu} u^\mu v^\nu = u_\nu v^\nu = u_\nu v_\mu g^{\mu\nu} \,. \tag{2.105}$$

Under basis transformation, $g^{\mu\nu}$ behaves as follows:

$$\begin{aligned}
(\mathbf{u}, \mathbf{v}) &= g^{\gamma\delta} u_\gamma v_\delta = (\mathbf{u'}, \mathbf{v'}) = g'^{\mu\nu} u'_\mu v'_\nu && (2.106)\\
&= g'^{\mu\nu} \Lambda_\mu^\gamma \Lambda_\nu^\delta u_\gamma v_\delta && (2.107)
\end{aligned}$$

and therefore

$$g^{\gamma\delta} = g'^{\mu\nu} \Lambda_\mu^\gamma \Lambda_\nu^\delta \,. \tag{2.108}$$

Multiplying on both sides by $\left(\Lambda^{-1}\right)_\gamma^l \left(\Lambda^{-1}\right)_\delta^m$, we have

$$\left(\Lambda^{-1}\right)_\gamma^l \left(\Lambda^{-1}\right)_\delta^m g^{\gamma\delta} = \left(\Lambda^{-1}\right)_\gamma^l \left(\Lambda^{-1}\right)_\delta^m \Lambda_\mu^\gamma \Lambda_\nu^\delta g'^{\mu\nu} = \left(\Lambda^{-1}\Lambda\right)_\mu^l \left(\Lambda^{-1}\Lambda\right)_\nu^m g'^{\mu\nu} \tag{2.109}$$

and finally

$$g'^{lm} = \left(\Lambda^{-1}\right)_\gamma^l \left(\Lambda^{-1}\right)_\delta^m g^{\gamma\delta} \,. \tag{2.110}$$

2.3.3 Vectors and Tensors in Differential Form

A more convenient (and general) way to define vectors and tensors is to use the apparatus of differential geometry. In this way, we use local definitions that are valid also for non linear spaces, like in General Relativity.

Contravariant Vectors

Let us suppose to work in a Euclidean space and let $(x^1, \cdots, x^n)$ be a system of euclidean coordinates. A curve in this space is given in parametric form as

$$\begin{cases} x^1 = x^1(t) \\ \quad \cdot \\ \quad \cdot \\ \quad \cdot \\ x^n = x^n(t) \end{cases} \tag{2.111}$$

with $t \in [a, b] \subset \mathbb{R}$. The *velocity* vector of the curve in the point $\mathbf{x}_0 = \mathbf{x}(t_0)$ is

$$\mathbf{v}_x = \left(\frac{dx^1}{dt}, \cdots, \frac{dx^n}{dt} \right) \Big|_{t=t_0}, \quad v_x^i = \frac{dx^i}{dt}. \tag{2.112}$$

Let us suppose that in a neighbourhood of $\mathbf{x}_0$ the new coordinates $(z^1, \cdots, z^n)$ are introduced, in such a way that

$$x^i = x^i(z^1, \cdots, z^n), \quad i = 1, \cdots, n \tag{2.113}$$

and such that in this neighbourhood we have

$$\det J = \det \left\{ \frac{\partial x^i}{\partial z^i} \right\} \neq 0. \tag{2.114}$$

In the new coordinates, the parametric equations of the curve are

$$\begin{cases} z^1 = z^1(t) \\ \quad \cdot \\ \quad \cdot \\ \quad \cdot \\ z^n = z^n(t) \end{cases} \tag{2.115}$$

and we can write

$$x^i(t) = x(\mathbf{z}(t)). \tag{2.116}$$

The velocity vector in the new coordinates is

$$\mathbf{v}_z = \left(\frac{dz^1}{dt}, \cdots, \frac{dz^n}{dt} \right) \Big|_{t=t_0}, \quad v_z^i = \frac{dz^i}{dt}, \tag{2.117}$$

In the transformation from x to z coordinates, the velvcity vector transforms as follows

$$v_x^i = \frac{dx^i}{dt} = \frac{\partial x^i}{\partial z^j} \frac{dz^j}{dt} = \frac{\partial x^i}{\partial z^j} v_z^j, \tag{2.118}$$

Therefore

$$\boxed{v_x^i = \frac{\partial x^i}{\partial z^j} v_z^j} \tag{2.119}$$

or, in matrix form

$$\mathbf{v}_x = J \, \mathbf{v}_z. \tag{2.120}$$

A vector whose components transform as in Eq. (2.119) is called a *contravariant* vector (or contravariant tensor of rank 1).

Covariant Vectors

Let us consider the gradient of a scalar function $f(x^1, \cdots, x^n)$:

$$\xi = \nabla f = \left(\frac{\partial f}{\partial x^1}, \cdots, \frac{\partial f}{\partial x^n} \right), \quad \xi_i = \frac{\partial f}{\partial x^i}. \tag{2.121}$$

If we introduce a new system of coordinates $(z^1, \cdots, z^n)$ such that $x^i = x^i(z^1, \cdots, z^n)$ and $\det J = \det \left\{ \frac{\partial x^i}{\partial z^i} \right\} \neq 0$, we define

$$\eta = \left(\frac{\partial f}{\partial z^1}, \cdots, \frac{\partial f}{\partial z^n} \right), \quad \eta_i = \frac{\partial f}{\partial z^i}. \tag{2.122}$$

Changing system of coordinates, the gradient transforms in the following way

$$\eta_i = \frac{\partial f}{\partial z^i} = \frac{\partial x^j}{\partial z^i} \frac{\partial f}{\partial x^j} = \frac{\partial x^j}{\partial z^i} \xi_j. \tag{2.123}$$

Therefore

$$\boxed{\eta_i = \frac{\partial x^j}{\partial z^i} \xi_j} \tag{2.124}$$

A vector whose components transform as in Eq. (2.124) is called a *covariant* vector (or covariant tensor of rank 1).

Summarizing, if the jacobian of the transformation is

$$J = \begin{pmatrix} \frac{\partial x^1}{\partial z^1} & \cdot & \cdot & \frac{\partial x^1}{\partial z^n} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \frac{\partial x^n}{\partial z^1} & \cdot & \cdot & \frac{\partial x^n}{\partial z^n} \end{pmatrix}, \tag{2.125}$$

we have in matrix form:

$$\text{contravariant} \qquad \xi = J\,\eta, \tag{2.126}$$

$$\text{covariant} \qquad \eta = J^t\,\xi, \qquad \Longrightarrow \qquad \xi = \left(J^t\right)^{-1} \eta. \tag{2.127}$$

Note: The transformations of the contravariant and covariant vector coincide in the case in which

$$J = \left(J^t\right)^{-1}, \qquad \Longrightarrow \qquad J J^t = 1, \tag{2.128}$$

therefore, if in every point the transformation is linear ($J = \mathit{const}$) and ortogonal.

Metric Tensor

Let us now introduce the scalar product of two vectors.

Let us suppose that the coordinate system $(x^1, \cdots, x^n)$ is euclidean, that ξ_1 and ξ_2 are two vectors with origin in $P_0 = (x_0^1, \cdots, x_0^n)$ and let us introduce in a neighbourhood of $(x_0^1, \cdots, x_0^n)$ another system of coordinates $(z^1, \cdots, z^n)$ such that $x^i = x^i(z^1, \cdots, z^n)$ and $\det J = \det \left\{ \frac{\partial x^i}{\partial z^i} \right\} \neq 0$, with $x_0^i = x^i(z_0^1, \cdots, z_0^n)$.

Knowing that

$$\xi_1^i = \left. \frac{\partial x^i}{\partial z^j} \right|_{P_0} \eta_1^j, \qquad \xi_2^i = \left. \frac{\partial x^i}{\partial z^j} \right|_{P_0} \eta_2^j, \tag{2.129}$$

we define the scalar product as

$$(\boldsymbol{\xi_1}, \boldsymbol{\xi_2}) = \xi_1^i \xi_2^i = \left. \frac{\partial x^i}{\partial z^j} \frac{\partial x^i}{\partial z^k} \right|_{P_0} \eta_1^j \eta_2^k = g_{jk} \eta_1^j \eta_2^k, \tag{2.130}$$

where we introduced the *metric tensor*

$$g_{jk} = \frac{\partial x^i}{\partial z^j}\frac{\partial x^i}{\partial z^k}\bigg|_{P_0} = J^i_j J^i_k = \delta_{rs} J^r_j J^s_k \,. \tag{2.131}$$

Let us see how the metric tensor transfoms under change of coordinates. If we introduce in a neighbourhood of P_0 a new system of coordinates $(y^1, \cdots, y^n)$ such that $z^i = z^i(y^1, \cdots, y^n)$ and $\det J \neq 0$, we will have

$$\eta^i_1 = \frac{\partial z^i}{\partial y^j}\bigg|_{P_0} \zeta^j_1 \,, \qquad \eta^i_2 = \frac{\partial z^i}{\partial y^j}\bigg|_{P_0} \zeta^j_2 \,, \tag{2.132}$$

Therefore

$$(\xi_{\mathbf{1}}, \xi_{\mathbf{2}}) = g_{ij}\eta^i_1\eta^j_2 = \frac{\partial z^i}{\partial y^k} g_{ij} \frac{\partial z^j}{\partial y^l}\bigg|_{P_0} \zeta^k_1\zeta^l_1 = g'_{kl}\zeta^k_1\zeta^l_1 \,. \tag{2.133}$$

The metric tensor transforms according to the following rule:

$$g'_{kl} = \frac{\partial z^i}{\partial y^k} g_{ij} \frac{\partial z^j}{\partial y^l}\bigg|_{P_0} = J^i_k J^j_l g_{ij} \,. \tag{2.134}$$

A tensor that transforms like in Eq. (2.134) is called a *covariant tensor* of rank 2.

Note that the metric tensor is a symmetric tensor

$$g_{ij} = g_{ji} \,, \tag{2.135}$$

because of the fact that the scalar product is symmetric. Moreover, in general

$$g_{ij} = g_{ij}(P_0) = g_{ij}(z^1_0, \cdots, z^n_0) \,, \tag{2.136}$$

then it is a function of the point in which ξ_1 and ξ_2 are defined.

Definition The metric $g_{ij}(z)$ is called *euclidean* if it exists a system of coordinates $(x^1, \cdots, x^n)$, with $x^i = x^i(z^1, \cdots, z^n)$ and $\det(J) \neq 0$, $g_{ij} = \frac{\partial x^k}{\partial z^i}\frac{\partial x^k}{\partial z^j}$, such that in these coordinates we have

$$g'_{ij} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{2.137}$$

Definition The metric $g_{ij}(z)$ is called *pseudo-euclidean* if it exists a system of coordinates $(x^1, \cdots, x^n)$, with $x^i = x^i(z^1, \cdots, z^n)$ and $\det(J) \neq 0$, $g_{ij} = \frac{\partial x^k}{\partial z^i}\frac{\partial x^k}{\partial z^j}$, such that in these coordinates we have

$$g'_{ij} = \delta_{ij} = \begin{cases} 1 & \text{for}\, i \leq p \;\, (i = j) \\ -1 & \text{for}\, p+1 \leq i \leq p+q = n \;\, (i = j) \\ 0 & i \neq j \end{cases} \tag{2.138}$$

The space where such a metric is defined is called *Pseudo Euclidean* and it is labeled with $\mathbb{R}^n_{p,q}$. We will call *Minkowski space*, a pseudo euclidean space $\mathbb{R}^4_{1,3}$, with metric

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \,. \tag{2.139}$$

In pseudo-euclidean coordinates we have

$$|\xi| = g_{ij}\xi^i\xi^j = (\xi^1)^2 + \cdots + (\xi^p)^2 - (\xi^{p+1})^2 - \cdots - (\xi^n)^2 \,. \tag{2.140}$$

We can extend the notion of metric tensor also to covariant vectors. We have

$$\xi_1 = \left(J^t\right)^{-1} \eta_1 , \tag{2.141}$$
$$\xi_2 = \left(J^t\right)^{-1} \eta_2 , \tag{2.142}$$

or, in components

$$\xi_{1,i} = \frac{\partial z^j}{\partial x^i} \eta_{1,j} , \tag{2.143}$$
$$\xi_{2,i} = \frac{\partial z^j}{\partial x^i} \eta_{2,j} , \tag{2.144}$$

where ξ_1 and ξ_2 are two vectors in the euclidean coordinate system $(x^1, \cdots, x^n)$, while η_1 and η_2 the same vectors in the system $(z^1, \cdots, z^n)$, such that $x^i = x^i(z^1, \cdots, z^n)$ and $\det(J) \neq 0$. Therefore,

$$(\xi_1, \xi_2) = \xi_{1,i}\xi_{1,i} = \frac{\partial z^j}{\partial x^i}\frac{\partial z^k}{\partial x^i} \eta_{1,j}\eta_{1,k} = g^{jk}\eta_{1,j}\eta_{1,k} . \tag{2.145}$$

In order to understand how g^{ij} transform under change of coordinate system, let us introduce another coordinate system $(y^1, \cdots, y^n)$ such that $z^i = z^i(y^1, \cdots, y^n)$ and $\det(J) \neq 0$. We have

$$\eta_{1,i} = \frac{\partial y^j}{\partial z^i} \zeta_{1,j} , \tag{2.146}$$
$$\eta_{2,i} = \frac{\partial y^j}{\partial z^i} \zeta_{2,j} , \tag{2.147}$$

and therefore

$$g^{jk}\eta_{1,j}\eta_{2,k} = \frac{\partial y^l}{\partial z^j} g^{jk} \frac{\partial y^r}{\partial z^k} \zeta_{1,l}\zeta_{2,r} , \tag{2.148}$$

and, finally

$$g'^{lr} = \frac{\partial y^l}{\partial z^j} g^{jk} \frac{\partial y^r}{\partial z^k} . \tag{2.149}$$

A quantity that transforms as in Eq. (2.149) is called *contravariant tensor of rank 2*.
Theorem We have $\left\{g^{ij}\right\} = \left\{g_{ij}\right\}^{-1}$.

In fact, let us look at the transformation rules in a matrix form. If we define the covariant metric tensor as g_c and the contravariant metric tensor as g^c, we have

$$g'_c = J^t g_c J , \tag{2.150}$$
$$g'^c = \left[\left(J^t\right)^{-1}\right]^t g^c \left(J^t\right)^{-1} . \tag{2.151}$$

However $\left[\left(J^t\right)^{-1}\right]^t = J^{-1}$ and therefore

$$g'^c = J^{-1} g^c \left(J^t\right)^{-1} . \tag{2.152}$$

From Eq. (2.150) we have

$$\left(g'_c\right)^{-1} = J^{-1} \left(g_c\right)^{-1} \left(J^t\right)^{-1} \tag{2.153}$$

and therefore we find

$$g^c = \left(g_c\right)^{-1} . \tag{2.154}$$

In components we have

$$g_{ij}g^{jk} = g_i^k = \delta_{ik} = g^{ij}g_{jk} = g_k^i , \tag{2.155}$$

where δ_{ik} is the Kronecker delta (and therefore we do not have to distinguish between upper or lower indices).

Mixed Tensors

Let us suppose now that in every point of our space, with coordinates $(x^1, \cdots, x^n)$, is defined a linear operator $A(\mathbf{x})$. If ξ is a vector in $\mathbf{x}$, we have

$$\eta^i = a^i_j(\mathbf{x})\, \xi^j \tag{2.156}$$

and for the covariant vectors

$$\eta_j = a^i_j(\mathbf{x})\, \xi_i\,. \tag{2.157}$$

If we introduce now, in the neighbourhood of $\mathbf{x}$ a new system of coordinates $(z^1, \cdots, z^n)$ such that $x^i = x^i(\mathbf{z})$, we will have

$$\eta^i = \frac{\partial x^i}{\partial z^j}\eta'^j\,, \qquad \xi^i = \frac{\partial x^i}{\partial z^j}\xi'^j \tag{2.158}$$

and

$$\eta_j = \frac{\partial z^i}{\partial x^j}\eta'_i\,, \qquad \xi^j = \frac{\partial z^i}{\partial x^j}\xi'_i\,. \tag{2.159}$$

Because of Eq. (2.156) we have

$$\frac{\partial x^i}{\partial z^k}\eta'^k = a^i_j\frac{\partial x^j}{\partial z^l}\xi'^l\,, \tag{2.160}$$

from which

$$\eta'^k = \frac{\partial z^k}{\partial x^i}\frac{\partial x^j}{\partial z^l}a^i_j\xi'^l = a'^k_l\xi'^l\,. \tag{2.161}$$

Therefore

$$a'^k_l = \frac{\partial z^k}{\partial x^i}\frac{\partial x^j}{\partial z^l}a^i_j\,. \tag{2.162}$$

A quantity that transforms according to Eq. (2.162) is called *mixed tensor* of rank 2.

General definition

In general, we define a tensor of (p,q) type, of rank $p+q$, on a n-dimensional vector space, a collection of n^{p+q} numbers, in a certain system of coordinates $(x^1, \cdots, x^n)$, whose numerical expression depends on the system of coordinate as follows: if $(z^1, \cdots, z^n)$ is another system of coordinates and $x^i = x^i(z^1, \cdots, z^n)$ we have

$$T^{i_1,\cdots,i_p}_{j_1,\cdots,j_q} = \frac{\partial x^{i_1}}{\partial z^{k_1}}\cdots\frac{\partial x^{i_p}}{\partial z^{k_p}},\frac{\partial z^{l_1}}{\partial x^{j_1}}\cdots\frac{\partial z^{l_q}}{\partial x^{j_q}}\,T'^{k_1,\cdots,k_p}_{l_1,\cdots,l_q}\,. \tag{2.163}$$

Since $det(J) \neq 0$, the relation (2.163) can be inverted:

$$T'^{k_1,\cdots,k_p}_{l_1,\cdots,l_q} = \frac{\partial z^{k_1}}{\partial x^{i_1}}\cdots\frac{\partial z^{k_p}}{\partial x^{i_p}},\frac{\partial x^{j_1}}{\partial z^{l_1}}\cdots\frac{\partial x^{j_q}}{\partial z^{l_q}}\,T^{i_1,\cdots,i_p}_{j_1,\cdots,j_q}\,. \tag{2.164}$$

In every point of the space, (p,q) tensors form a linear space.

2.4 Minkowski Space

The ideal space for the study of Special Relativity is a 4-dimensional vector space ($X^0 = ct$, $X^1 = x$, $X^2 = y$, $X^3 = z$), called *Minkowski space*, $\mathbb{M}^4$, with pseudo-euclidean metric $\eta^{\mu\nu}$, that in matrix form is given by the following expression:

$$\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\,. \tag{2.165}$$

We have

$$\eta^{\mu\nu} = \eta_{\mu\nu}\,, \tag{2.166}$$

such that

$$\eta^{\mu\nu}\eta_{\nu\rho} = \eta^{\mu}_{\rho} = \delta^{\mu}_{\rho}\,. \tag{2.167}$$

The scalar product in this space is defined as follows:

$$X \cdot X = X_\mu X^\mu = \eta_{\mu\nu} X^\mu X^\nu = \eta^{\mu\nu} X_\mu X_\nu = (X^0)^2 - (X^1)^2 - (X^2)^2 - (X^3)^2 \tag{2.168}$$

and it is not positive definite.

The vectors in $\mathbb{M}^4$ are called four-vectors. We have *contravariant vectors* (with contravariant indices)

$$V^\mu = (V^0, \mathbf{V})\,. \tag{2.169}$$

The covariant vector V_μ can be recovered by V^μ using the metric

$$V_\mu = \eta_{\mu\nu} V^\nu = (V^0, -\mathbf{V})\,, \tag{2.170}$$

2.5 Lorentz group

So far we have considered boosts. However, the Lorentz tranformations do not include only boosts. The specific requirement is that a Lorents transformation leave unchanged the quadratic form

$$(X^0)^2 - (X^1)^2 - (X^2)^2 - (X^2)^2\,. \tag{2.171}$$

A boost does exactly this. However, there is another transformation that can leave (2.171) unchanged. In fact, if we consider a rigid rotation in the euclidean 3-dim space, this will leave unchanged the quadratic form $(X^1)^2 + (X^2)^2 + (X^2)^2$, and therefore also (2.171), since the time is not included in the transformation. Finally, also transformations of the following matrix form

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}\,, \tag{2.172}$$

where R is ortogonal ($RR^t = 1$), are Lorentz tranformations.

Let us consider a generic Lorentz transformation, Λ^μ_ν, that can be a composition of a boost and a rigid rotation. In order Λ^μ_ν to be a Lorentz transformation, it must fulfill the following relation:

$$\eta_{\mu\nu}\Lambda^\mu_\sigma \Lambda^\nu_\rho = \eta_{\sigma\rho}\,. \tag{2.173}$$

The relation in Eq. (2.173) comes from the fact that Lorentz transformations preserve the metric and leave unchanged the lenght of the four-vector X^μ:

$$X'^2 = \eta_{\mu\nu} X'^\mu X'^\nu = \eta_{\mu\nu}\Lambda^\mu_\sigma X^\sigma \Lambda^\nu_\rho X^\rho = \eta_{\mu\nu}\Lambda^\mu_\sigma\Lambda^\nu_\rho X^\sigma X^\rho\,, \tag{2.174}$$

$$X^2 = \eta_{\sigma\rho} X^\sigma X^\rho \tag{2.175}$$

and from $X'^2 = X^2$ we find (2.173). Using

$$X'^2 = \eta^{\mu\nu} X'_\mu X'_\nu = \eta^{\mu\nu}\Lambda_\mu^{\cdot\sigma} X_\sigma \Lambda_\nu^{\cdot\rho} X_\rho = \eta^{\mu\nu}\Lambda_\mu^{\cdot\sigma}\Lambda_\nu^{\cdot\rho} X_\sigma X_\rho\,, \tag{2.176}$$

$$X^2 = \eta^{\sigma\rho} X_\sigma X_\rho\,, \tag{2.177}$$

we find also the following form

$$\eta^{\mu\nu}\Lambda_\mu^{\cdot\sigma}\Lambda_\nu^{\cdot\rho} = \eta^{\sigma\rho}\,. \tag{2.178}$$

Eq. (2.173) can be written in matrix form as $\Lambda^t \eta \Lambda = \eta$.

Property 2.5.1 *The Lorentz transformations, Λ^μ_ν, form a group, the Lorents Group (LG).*

In order to see that, we have to prove that: *i)* if $\Lambda_1 \in LG$ and $\Lambda_2 \in LG$, then $\Lambda_1\Lambda_2 \in LG$; *ii)* the identity is a Lorentz transformation; *iii)* $\exists!\, \Lambda^{-1} \in LG$ such that $\Lambda\Lambda^{-1} = \Lambda^{-1}\Lambda = 1$.

- Let us consider

$$X^\mu \overset{\Lambda_1}{\to} X'^\mu \overset{\Lambda_2}{\to} X''^\mu \,, \tag{2.179}$$

then we have

$$X''^\mu = (\Lambda_2)^\mu_\nu X'^\nu = (\Lambda_2)^\mu_\nu (\Lambda_1)^\nu_\rho X^\rho \,. \tag{2.180}$$

We have to prove that

$$\Lambda^\mu_\rho = (\Lambda_2)^\mu_\nu (\Lambda_1)^\nu_\rho \tag{2.181}$$

is indeed a Lorentz transformation (it satisfies Eq. (2.173)). In fact, we have

$$\eta_{\mu\nu}\left[(\Lambda_2)^\mu_\gamma(\Lambda_1)^\gamma_\sigma\right]\left[(\Lambda_2)^\nu_\delta(\Lambda_1)^\delta_\rho\right] = \eta_{\mu\nu}(\Lambda_2)^\mu_\gamma(\Lambda_2)^\nu_\delta\left[(\Lambda_1)^\gamma_\sigma(\Lambda_1)^\delta_\rho\right] = \eta_{\gamma\delta}(\Lambda_1)^\gamma_\sigma(\Lambda_1)^\delta_\rho = \eta_{\sigma\rho}\,, \tag{2.182}$$

where we used the fact that Λ_2 and Λ_1 are indeed Lorentz transformations.

- The identity transformation

$$\Lambda^\mu_\nu = \delta^\mu_\nu \tag{2.183}$$

trivially satisfies relation (2.173):

$$\eta_{\mu\nu}\delta^\mu_\sigma\delta^\nu_\rho = \eta_{\sigma\rho}\,. \tag{2.184}$$

- The inverse exists. In fact, since $\Lambda^t\eta\Lambda = \eta$, $(det(\Lambda))^2 = 1$ or $det(\Lambda) = \pm 1$ (and in particular $det(\Lambda) \neq 0$). Let us see how the inverse can be defined. Multiplying both sides of Eq. (2.173) by $\eta^{\rho\sigma'}$ we have

$$\eta^{\rho\sigma'}\eta_{\mu\nu}\Lambda^\mu_\sigma\Lambda^\nu_\rho = \eta^{\rho\sigma'}\eta_{\sigma\rho} = \eta^{\sigma'}_\sigma \tag{2.185}$$

This means that

$$\eta^{\rho\sigma'}\eta_{\mu\nu}\Lambda^\nu_\rho = \Lambda_\mu^{\cdot\sigma'} = (\Lambda^{-1})^{\sigma'}_\mu\,. \tag{2.186}$$

Let us check that, indeed, $(\Lambda^{-1})^{\sigma'}_\mu$ is a Lorentz transformation, i.e. that

$$\eta_{\mu\nu}(\Lambda^{-1})^\mu_\sigma(\Lambda^{-1})^\nu_\rho = \eta_{\rho\sigma}\,. \tag{2.187}$$

Using relation (2.186) we have to prove that

$$\eta_{\mu\nu}(\eta^{\mu\xi}\eta_{\sigma\omega}\Lambda^\omega_\xi)(\eta^{\nu\xi'}\eta_{\rho\omega'}\Lambda^{\omega'}_{\xi'}) = \eta_{\mu\nu}\eta^{\mu\xi}\eta^{\nu\xi'}\eta_{\sigma\omega}\eta_{\rho\omega'}\Lambda^\omega_\xi\Lambda^{\omega'}_{\xi'} = \eta^{\xi\xi'}(\eta_{\sigma\omega}\Lambda^\omega_\xi)(\eta_{\rho\omega'}\Lambda^{\omega'}_{\xi'})\,. \tag{2.188}$$

If we multiply the r.h.s. and l.h.s. of Eq. (2.173) by $\Lambda_{\nu'}^{\cdot\rho}$ we find

$$\eta_{\mu\nu}\Lambda^\mu_\sigma\Lambda^\nu_\rho\Lambda_{\nu'}^{\cdot\rho} = \eta_{\mu\nu'}\Lambda^\mu_\sigma = \eta_{\sigma\rho}\Lambda_{\nu'}^{\cdot\rho}\,. \tag{2.189}$$

Using Eq. (2.189) in Eq. (2.188), we find

$$\begin{aligned}\eta_{\mu\nu}(\eta^{\mu\xi}\eta_{\sigma\omega}\Lambda^\omega_\xi)(\eta^{\nu\xi'}\eta_{\rho\omega'}\Lambda^{\omega'}_{\xi'}) &= (\eta^{\xi\xi'}\eta_{\xi\delta})\Lambda_\sigma^{\cdot\delta}(\eta_{\rho\omega'}\Lambda^{\omega'}_{\xi'}) = \delta^{\xi'}_\delta\Lambda_\sigma^{\cdot\delta}\eta_{\rho\omega'}\Lambda^{\omega'}_{\xi'}\,, &(2.190)\\ &= \Lambda_\sigma^{\cdot\xi'}\eta_{\rho\omega'}\Lambda^{\omega'}_{\xi'} = \eta_{\rho\omega'}\delta^\sigma_{\omega'} = \eta_{\rho\sigma}\,. &(2.191)\end{aligned}$$

In summary, Lorentz transformations form a group that is called Lorentz Group. Since, as we noticed, $\Lambda \in LG$ is such that $(det(\Lambda))^2 = 1$, we have elements of the group with $det(\Lambda) = 1$ and elements with $det(\Lambda) = -1$. The identity has $det(\Lambda) = 1$. This means that only the subset with $det(\Lambda) = 1$ can form a subgroup of the LG.

Moreover, from (2.173) we have

$$1 = \eta_{00} = \eta_{\mu\nu}\Lambda^{\mu}_{0}\Lambda^{\nu}_{0} = (\Lambda^0_0)^2 - \sum_i (\Lambda^i_0)^2 \,, \tag{2.192}$$

from which we obtain

$$(\Lambda^0_0)^2 \geq 1 \qquad \Longrightarrow \qquad \Lambda^0_0 \geq 1 \quad \text{or} \quad \Lambda^0_0 \leq -1 \,. \tag{2.193}$$

In total the LG has four different subsets, listed in the following table:

Symbol	$\mathbf{\Lambda^0_0}$	**det**$\mathbf{\Lambda}$	**Name**
$L^{\uparrow}_{+}$	≥ 1	$+1$	Proper hortocronous
$L^{\downarrow}_{+}$	≤ -1	$+1$	Proper anticronous
$L^{\uparrow}_{-}$	≥ 1	-1	Improper hortocronous
$L^{\downarrow}_{-}$	≤ -1	-1	Improper anticronous

Only $L^{\uparrow}_{+}$ is a soubgroup of the LG (the identity is such that $\delta^0_0 = 1$) and its elements can be obtained from the identity with a continuous change of the parameters of the group (for instance the velocity of the boosts and the angles of the rigit rotation). A group that depends in a continuous and differentiable way on a set of parameters is called a Lie group.

The four subsets $L^{\uparrow}_{+}$, $L^{\downarrow}_{+}$, $L^{\uparrow}_{-}$ and $L^{\downarrow}_{-}$ can be connected only via the discontinuous transformations called *Parity* and *Time Reversal*.

- A *Parity* transformation acts only on the spatial part of the four-vector, inverting it:

$$(X^0, \mathbf{X}) \stackrel{\Lambda_P}{\to} (X^0, -\mathbf{X}) \,. \tag{2.194}$$

  In matrix form we have

$$\Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad (= \eta) \,. \tag{2.195}$$

  Parity belongs to the set $L^{\uparrow}_{-}$.

- A *Time Reversal* transformation acts only on the temporal part of the four-vector, inverting it:

$$(X^0, \mathbf{X}) \stackrel{\Lambda_T}{\to} (-X^0, \mathbf{X}) \,. \tag{2.196}$$

  In matrix form we have

$$\Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad (= -\eta) \,. \tag{2.197}$$

  Time Reversal belongs to the set $L^{\downarrow}_{-}$.

We can connect the four subsets using Parity and Time Reversal, as in figure 2.2. For instance, take an element of $L^{\uparrow}_{+}$, say $\Lambda^{\uparrow}_{+}$. The element $\Lambda_1 = \Lambda^{\uparrow}_{+}\Lambda_T$ is such that $det(\Lambda_1) = det(\Lambda^{\uparrow}_{+}\Lambda_T) = det(\Lambda^{\uparrow}_{+})det(\Lambda_T) = -1$ and $(\Lambda_1)^0_0 = -(\Lambda^{\uparrow}_{+})^0_0 \leq -1$. Therefore, the action of Λ_T was such that $\Lambda_1 \in L^{\downarrow}_{-}$. If we consider, instead, $\Lambda_2 = \Lambda^{\uparrow}_{+}\Lambda_P$, we find again $det(\Lambda_2) = -1$ but now the sign of $(\Lambda_2)^0_0$ is the same as the one of $(\Lambda^{\uparrow}_{+})^0_0$. Therefore $\Lambda_2 \in L^{\uparrow}_{-}$.

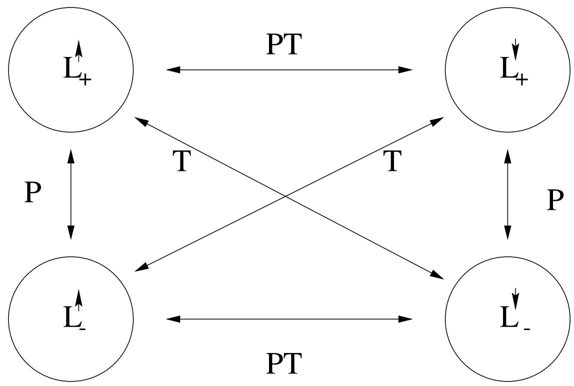


Figura 2.2: Connection of the four subsets of the LG.

2.6 Poincaré group

We consider so far only homogeneous transformations. However, in addition to boosts and rigid rotations, we have the freedom to redefine the origin of our inertial frame, adding a constant vector (rigid translation) and Physics must not to be affected by this operation. Such a transformation can be expressed as

$$X^\mu \to x'^\mu = \Lambda^\mu_\nu X^\nu + a^\mu \,. \tag{2.198}$$

Boosts with rotations and translations of the space axis are called *Poincaré transformations* (or inhomogeneous Lorents transformations). We indicate such a transformation as $T(\Lambda, a)$.

Poincaré transformations form a group, the Poincaré group (PG). In fact:

- the identity $T(1,0) \in PG$;
- the composition of two transformation is a Poincaré transformation, since

$$X''^\mu = (\Lambda_1)^\mu_\nu X'^\nu + a'^\mu = (\Lambda_1)^\mu_\nu \left[(\Lambda_2)^\nu_\rho X^\rho + a^\nu\right] + a'^\mu = (\Lambda_1)^\mu_\nu (\Lambda_2)^\nu_\rho\, X^\rho + (\Lambda_1)^\mu_\nu a^\nu + a'^\mu \,, \tag{2.199}$$

  and $(\Lambda_1)^\mu_\nu (\Lambda_2)^\nu_\rho$ is a Lorentz transformation, while $(\Lambda_1)^\mu_\nu a^\nu + a'^\mu$ is a constant vector. Therefore

$$T(\Lambda', a')T(\Lambda, a) = T(\Lambda'\Lambda, \Lambda' a + a') \,. \tag{2.200}$$

- The inverse exists, $T^{-1}(\Lambda, a) = T(\Lambda^{-1}, -\Lambda^{-1}a)$ such that

$$T(\Lambda, a)T^{-1}(\Lambda, a) = T(\Lambda, a)T(\Lambda^{-1}, -\Lambda^{-1}a) = T(\Lambda\Lambda^{-1}, -\Lambda\Lambda^{-1}a + a) = T(1, 0) \,. \tag{2.201}$$

2.7 Infinitesimal Transformations

Since $L^\uparrow_+$ is constituted by elements that can be connected smoothly to the identity, we can study the local properties of LG and PG using infinitesimal transformations. An infinitesimal transformation is a Lorentz (Poincaré) transformation in which the parameters go smoothly to zero. Therefore, for instance

$$\Lambda^\mu_\nu \simeq \delta^\mu_\nu + \epsilon^\mu_\nu \,, \tag{2.202}$$

at first order. Including the translations we will have

$$T(\Lambda^\mu_\nu, a^\mu) \simeq T(\delta^\mu_\nu + \epsilon^\mu_\nu, \delta a^\mu) \,, \tag{2.203}$$

where we considered

$$X'^{\mu} \simeq (\delta^{\mu}_{\nu} + \epsilon^{\mu}_{\nu})\, X^{\nu} + \delta a^{\mu} = X^{\mu} + \epsilon^{\mu}_{\nu}\, X^{\nu} + \delta a^{\mu}\,. \tag{2.204}$$

The infinitesimal Lorentz transformation $(\delta^{\mu}_{\nu} + \epsilon^{\mu}_{\nu})$ has to satisfy the usual relation (2.173). Therefore:

$$\eta_{\sigma\rho} = \eta_{\mu\nu}(\delta^{\mu}_{\sigma} + \epsilon^{\mu}_{\sigma})(\delta^{\mu}_{\rho} + \epsilon^{\mu}_{\rho})\,, \tag{2.205}$$

$$= \eta_{\mu\nu}\delta^{\mu}_{\sigma}\delta^{\mu}_{\rho} + \eta_{\mu\nu}\delta^{\mu}_{\sigma}\epsilon^{\mu}_{\rho} + \eta_{\mu\nu}\epsilon^{\mu}_{\sigma}\delta^{\mu}_{\rho} + \ldots\,, \tag{2.206}$$

$$= \eta_{\sigma\rho} + \epsilon_{\sigma\rho} + \epsilon_{\rho\sigma}\,. \tag{2.207}$$

This means that the ϵ tensor must be antisymmetric

$$\epsilon_{\sigma\rho} = -\epsilon_{\rho\sigma} \tag{2.208}$$

and therefore it has 6 independent elements. The LG then depends upon 6 independent parameters (three for the boosts and three for the rotations). Including the 4 parameters of the rigid translation, we have that the Poincaré group depends on 10 parameters.

NB The transformation Λ^{μ}_{ν} is already a "representation" of the Lorentz group, as we will see in the next chapter. It is a particular representation called the "fundamental" representation. We can find immediately from Eq. (2.202) the form of the generators of the group in this representation. In fact, we can write

$$\Lambda^{\mu}_{\nu} \simeq \delta^{\mu}_{\nu} + \epsilon^{\mu}_{\nu} = \delta^{\mu}_{\nu} - \frac{i}{2}(J^{\rho\sigma})^{\mu}_{\nu}\epsilon_{\rho\sigma}\,, \tag{2.209}$$

where

$$(J^{\rho\sigma})^{\mu}_{\nu} = i(\eta^{\rho\mu}\delta^{\sigma}_{\nu} - \eta^{\sigma\mu}\delta^{\rho}_{\nu}) \tag{2.210}$$

and we recover[1] Eq. (2.202). The six 4×4 matrices defined in Eq. (2.210) are the generators of the group in this representation.

2.8 Some notes on Group Theory

In this section we recall some concepts of group theory. This is important in order to fully characterize the Poincaré group and to understand the transformation properties of the quantities we will study in the rest of the course.

Definition 2.8.1 *A Group (G) is a collection of elements that are combined through a closed operation (product) such that*

$$\text{if } \ a, b \in G \quad \Longrightarrow a \cdot b \in G\,. \tag{2.211}$$

The product must obey the following properties:

1. $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (associative)
2. $\exists e \in G$ such that $a \cdot e = e \cdot a = a$, $\forall a \in G$ (identity: null element of the product)
3. $\forall a \in G$, $\exists a^{-1} \in G$, such that $a \cdot a^{-1} = a^{-1} \cdot a = e$ (inverse)

It can be demonstrated that the identity element e and the inverse a^{-1} are unique and that $(a^{-1})^{-1} = a$.

Definition 2.8.2 *If $a \cdot b = b \cdot a$, $\forall\, a, b \in G$, the group G is called Abelian.*

Definition 2.8.3 *We call subgroup of G, a subset H which is closed under the operation defined on G.*

[1] Remember that $\epsilon_{\rho\sigma}$ is antysymmetric.

Definition 2.8.4 *We call homomorphism between two groups, an application*

$$\phi\,:\; G_1 \to G_2\,, \tag{2.212}$$

such that $\forall g_1, g_2 \in G_1$ *we have*

$$\phi(g_1 \cdot g_2) = \phi(g_1) \circ \phi(g_2)\,, \tag{2.213}$$

where "$\circ$*" is the product defined in* G_2*. If* ϕ *is invertible is called isomorphism.*

2.8.1 Representations

The set of linear invertible transformations on a vector space V is a group which is called $GL(V)$.

Definition 2.8.5 *A Representation of a group* G *on a vector space* V *is an homomorphism*

$$D_R\,:\, G \to GL(V)\,. \tag{2.214}$$

Therefore, if $g \in G$ *it follows that* $D_R(g) \in GL(V)$ *and*

$$\begin{aligned} D_R(g_1 \cdot g_2) &= D_R(g_1) D_R(g_2)\,, && (2.215)\\ D_S(e) &= 1\,. && (2.216) \end{aligned}$$

So to say, the group is the abstract entity, while the representation is the realization of the group structure via operators on a vector space. $dim(V)$ is the dimension of the representation. If $dim(V) = n$ finite, we can immediately figure out a representation as a space of matrices acting on some finite dimensional vector space. In this case the product can be the usual product rows by columns.

Definition 2.8.6 *Two representations* D_1 *and* D_2 *of the same group* G *on two vector spaces* V *and* W *are called "equivalent" if there exists an invertible application between the two vector spaces*

$$T\,:\, V \to W \tag{2.217}$$

such that

$$T\, D_1(g)\, T^{-1} = D_2(g)\,, \quad \forall g \in G\,. \tag{2.218}$$

Irriducible representations

Let us now introduce the concept of reducible and irreducible representations.

Definition 2.8.7 *A subspace* S *of* V *is called "invariant" with respect to the representation* $D_S(g)$ *if* $\forall v \in S$ *and* $\forall g \in G$*, we have* $D_R(g)v \in S$*.*

Therefore

Definition 2.8.8 *A representation* D_R *on a vector space* V *is "irreducible" if* V *does not contain subspaces invariant under* D_R*. On the contrary, it is "reducible" if it contains invariant subspaces.*

In this case the representation D_R can be expressed as direct sum of irriducible representations

$$D_R(G) = \sum_{\bigoplus m} D_R^{(m)}(G) \tag{2.219}$$

and the operators that act on V (finite-dimensional, for instance) will appear, in a suitable basis, as block diagonal matrices.

2.8.2 Lie groups

A special role in Physics is played by the connected Lie groups

Definition 2.8.9 *A "Lie group" is a group whose elements g depend in a continuous and differentiable way on a set of real parameters θ^a, $a = 1, 2, ..., n$, $\theta^a \in \mathbb{R}$*

$$g = g(\theta^1, ..., \theta^n) \,. \tag{2.220}$$

Without loss of generality, we can chose θ^a such that for $\theta^a = 0$ we have $g(0) = e$. In this way, every element of the group is connected to the identity by a continuous path in $\mathbb{R}^n$

Lie Algebra

There is an important structure that is connected to the Lie group (and that, as the group itself, does not depend on the representation of the group) and is called the Lie Algebra. Although not necessary, in order to find out the algebra connected to the Lie group we consider a particular representation, $D_R(g(\theta))$. For $\theta \to 0$ we have to recover the identity (acting on the vector space) and since $D_R(g(\theta))$ is continuous and differentiable in θ, we can define the infinitesimal trasformation (at first order in θ) as

$$D_R(g(\theta)) \simeq 1 + i\theta_a T^a_R \,, \tag{2.221}$$

where we considered the fact the our Lie group can depend upon a set of parameters, θ_a, $a = 1, ..., n$, and where

$$T^a_R = -i \left. \frac{\partial D_R}{\partial \theta_a} \right|_{\theta=0} \,. \tag{2.222}$$

The operators T^a_R are called the *generators* of the Lie group, in the representation R.

In terms of the generators, we can write any transformation $D_R(g(\theta))$ in exponential form. As a simple example, let us consider the case of a Lie group depending on a single parameter θ and let us indicate with $D_R(\theta)$ the operator corresponding to $g(\theta)$ in a certain representation. Using the group properties and (2.221), we will have

$$D_R(\theta + d\theta) = D_R(\theta)\, D_R(d\theta) \simeq D_R(\theta)\ (1 + id\theta T_R) = D_R(\theta) + id\theta T_R\, D_R(\theta) \,. \tag{2.223}$$

Since

$$D_R(\theta + d\theta) - D_R(\theta) \simeq \frac{dD_R}{d\theta}\, d\theta = iT_R D_R(\theta)\, d\theta \,, \tag{2.224}$$

we then have

$$D_R(\theta) = e^{iT_R\theta} \,. \tag{2.225}$$

This formula can be proven to hold in general. If the Lie group depends upon a certain number of parameters, we can indeed write the operator D_R in exponential form (it is called the exponential map)

$$D_R(g) = e^{iT^a_R\theta_a} \,. \tag{2.226}$$

The generators T^a_R obey an algebra, that can be found as follows. Since $D(g)$ is a representation of our Lie group, it has tu fullfil the following relation

$$D_R(g_1)\, D_R(g_2) = D_S(g_1 g_2) = D_R(g_3) \,, \tag{2.227}$$

where $g_3 = g_1 g_2$. Using the exponential map, this means

$$e^{i\alpha_a T^a_R} e^{i\beta_a T^a_R} = e^{i\delta_a T^a_R} \,, \tag{2.228}$$

since $D_S(g_1 g_2) = D_R(g_3)$ and therefore it must be of the same form of $D_R(g_1)$ and $D_R(g_2)$, and

$$\delta_a = \delta_a(\alpha_a, \beta_a)\,. \tag{2.229}$$

However, in general we have

$$e^A e^B \neq e^{A+B}\,, \tag{2.230}$$

and therefore (in general) $\delta_a \neq \alpha_a + \beta_a$.

Let us consider infinitesimal transformations and let us take the logarithm of both sides of Eq. (2.228). We have

$$\log\left[\left(1 + i\alpha_a T_R^a + \frac{1}{2}(i\alpha_a T_R^a)^2 + ..\right)\left(1 + i\beta_a T_R^a + \frac{1}{2}(i\beta_a T_R^a)^2 + ...\right)\right] = i\delta_a T_R^a \tag{2.231}$$

or

$$\log\left[1 + i\alpha_a T_R^a + i\beta_a T_R^a - \frac{1}{2}(\alpha_a T_R^a)^2 - \frac{1}{2}(\beta_a T_R^a)^2 - \alpha_a\beta_b T_R^a T_R^b + ...\right] = i\delta_a T_R^a\,. \tag{2.232}$$

Expanding the log up to second order ($\log(1+x) \simeq x - x^2/2...$) we have

$$\begin{aligned} i\delta_a T_R^a \;&\simeq\; i\alpha_a T_R^a + i\beta_a T_R^a - \frac{1}{2}(\alpha_a T_R^a)^2 - \frac{1}{2}(\beta_a T_R^a)^2 - \alpha_a\beta_b T_R^a T_R^b + \frac{1}{2}\alpha_a\beta_b T_R^a T_R^b + \frac{1}{2}\alpha_a\beta_b T_R^b T_R^a \\ &\quad + \frac{1}{2}(\alpha_a T_R^a)^2 + \frac{1}{2}(\beta_a T_R^a)^2\,, \qquad (2.233) \\ &= i(\alpha_a + \beta_a) T_R^a - \frac{1}{2}\alpha_a\beta_b (T_R^a T_R^b - T_R^b T_R^a)\,, \qquad (2.234) \end{aligned}$$

or

$$\alpha_a\beta_b[T^a, T^b] = 2i(\alpha_c + \beta_c - \delta_c)T^c = \gamma_c T^c\,. \tag{2.235}$$

Since this relation must hold for every α_c and β_c, γ must be proportional to $\alpha_a\beta_b$:

$$\gamma_c = \alpha_a\beta_b\, f_c^{a\,b}\,. \tag{2.236}$$

The constants $f_c^{a\,b}$ are called *structure constants*. Finally we find

$$[T^a, T^b] = i f_c^{a\,b}\, T^c\,, \tag{2.237}$$

which is the Lie algebra that the generators have to fulfill.

The explicit form of the generators T^a depends on the specific representation. However, the algebra (2.237) is completely general and valid for every representation. We can prove that the structure constants are independent on the representation as well. They remain the same in every representation. Finally, we found the algebra imposing the group structure at second order in α and β. However, it can be proven that aty higher orders no further requirements occur. Knowing the structure constants and the generators is sufficient to know everything about the local structure of the group.

If $f_c^{a\,b} = 0$, we have $[T^a, T^b] = 0$ and the group is Abelian.

The idea behind the study of the algebra connected to the Lie group lies in the fact that

1. The representations of the algebra induce a corresponding representation on the Lie group;
2. It is easier to study the algebra than the group, since the generators form a vector space (and it is easier to deal with sums that with products).

Casimir operators

In the study of the representations an important role is played by the *Casimir operators.*

Definition 2.8.10 *A Casimir operator is an operator that commutes with all the generators T^a of the group.*

For instance we can think about the angular momentum, remembering that J^2 actually commutes with the three components of the angular momentum J^i (that are the generators for the rotations), $[J^2, J^i] = 0$. Casimir operators help in the study of the irreducible representations. They are linked to the first Schur's lemma

Lemma 2.8.11 (Schur's lemma). *If $U(G)$ is an irriducible representation of the group G on a vector space V and J^2 is a Casimir operator for that representation ($[J^2, U(g)] = 0$ for $\forall g \in G$), then J^2 is proportional to the identity.*

As an example we can consider again the angular momentum for which we have $J^2 = j(j+1)\mathbb{1}$.

If we consider abelian groups, since $f_c^{a\,b} = 0$, we have $[T^a, T^b] = 0$ and therefore every generator is a Casimir. It follows that every irreducible representation of an abelian group will be constituted by operators that, since they all commute with the generators, are proportional to the identity. Therefore the irreducible representations will have dimension one.

Unitary representations

Of particular importance in Physics are the unitary representations of the Lie groups. To have a unitary representation means that the generators are hermitian and therefore they can be identified with observables. In order to have finite-dimensional unitary representations of a Lie group we need the group to be compact. This means that the parameters, which the group depends on, should range in a closed interval of the reals. This is the case, for instance, of the two- and three-dimensional rotations (rotations in the Euclidean space), for which the angles are defined in closed intervals. This is, instead, not the case for the Lorentz group. Although the part that regards rotations is compact, the boosts are not. The parameter that defines the boosts are the components of the velocity $\mathbf{v}$ of the inertial frame S' with respect to the inertial frame S. The modulus of $\mathbf{v}$ is such that $0 \leq \frac{v}{c} < 1$. So, v can never reach the speed of light c ($v = c$ is a singular point for Lorentz transformations). This fact makes in such a way that the Lorents group is not compact.

There is a theorem that states: "Non compact groups have no finite-dimensional unitary representations".

Therefore, finite-dimensional representations of the Lorentz group cannot be unitary. However, we can find unitary infinite-dimensional representations and this is what matters for our physical descriptions, since quantum states live in infinite-dimensional spaces (Hilbert space ...).

2.8.3 A simple example: the (abelian) group $SO(2)$ and $U(1)$

As a first simple example let us consider the group of rotations in two dimensions. This is a Lie group depending on a single real parameter, the angle of rotation ϕ.

$R(\phi)$ is a rotation of angle ϕ acting on a certain vector space, $\mathbf{v} \in V$. Under $R(\phi)$ the modulous of $\mathbf{v}$ has to remain unchanged. If we have

$$\mathbf{v} \to \mathbf{v}' = R(\phi)\mathbf{v}\,, \tag{2.238}$$

then

$$|\mathbf{v}'|^2 = \mathbf{v}^t R(\phi)^t R(\phi)\mathbf{v} \equiv \mathbf{v}^t\mathbf{v} = |\mathbf{v}|^2 \qquad \Longrightarrow \qquad R(\phi)^t R(\phi) = \mathbb{1} = R(\phi)R(\phi)^t\,. \tag{2.239}$$

This defines the *orthogonal transofrmations*. Moreover, from $R(\phi)^t R(\phi) = \mathbb{1}$ it follows that $\det R(\phi) = \pm 1$. The subset with $\det R(\phi) = 1$ forms a subgroup (it is the one containing the identity) which is known as $SO(2)$, *special ortogonal group*.

$SO(2)$ is an Abelian group, since

$$R(\phi_1)R(\phi_2) = R(\phi_1 + \phi_2) = R(\phi_2)R(\phi_1)\,. \tag{2.240}$$

It has the global property

$$R(\phi) = R(\phi \pm 2\pi) \tag{2.241}$$

and this property is not related to the infinitesimal transformations (local structure) but to the global structure of the group.

Let us find the generators of $SO(2)$. For the infinitesimal transformations we have

$$R(d\phi) \simeq \mathbb{1} + K\,d\phi\,. \tag{2.242}$$

Since $R(\phi)^t R(\phi) = \mathbb{1}$ we have

$$(\mathbb{1} + K\,d\phi)^t(\mathbb{1} + K\,d\phi) = \mathbb{1}\,, \tag{2.243}$$

therefore at first order in $d\phi$ we have

$$\mathbb{1} + (K + K^t)\,d\phi = \mathbb{1}\,, \tag{2.244}$$

or

$$K = -K^t\,. \tag{2.245}$$

We can put $K = -iJ$ with J hermitian ($J^\dagger = J$) and then

$$R(d\phi) \simeq \mathbb{1} - iJ\,d\phi\,. \tag{2.246}$$

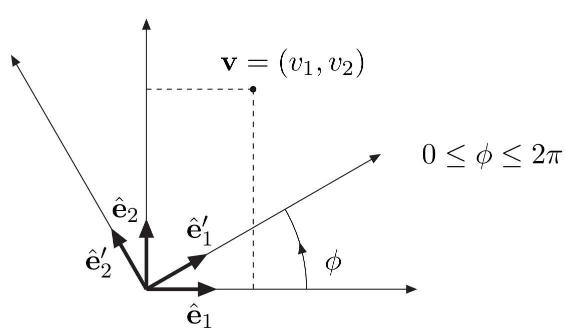
Exponentiating we find

$$R(\phi) = e^{-i\phi J}\,. \tag{2.247}$$

J is the generator of the group.

The representation on the Euclidean 2-dim vector space

We can consider a matrix representation of $SO(2)$ as rotations on a 2-dim Euclidean vector space.



We can write the transformation (that rotates the basis) as

$$\hat{\mathbf{e}}'_i = D(\phi)_i^j \hat{\mathbf{e}}_j\,, \tag{2.248}$$

where the matrix $D(\phi)$ is

$$D(\phi) = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}\,. \tag{2.249}$$

If $\mathbf{v} \in V$ it can be written in components as

$$\mathbf{v} = v^i \hat{\mathbf{e}}_i = v'^i \hat{\mathbf{e}}'_i \,. \tag{2.250}$$

The components transform with $D^{-1} = D^t$ and in this case

$$v'^i = D^i_j v^j \,, \tag{2.251}$$

that in matrix form means

$$\begin{pmatrix} v'^1 \\ v'^2 \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} \,. \tag{2.252}$$

In order to find the expression of the generator in this representation, let us find the infinitesimal transformation expanding for a small parameter the expression in Eq. (2.249)

$$D(d\phi) = \begin{pmatrix} 1 & d\phi \\ -d\phi & 1 \end{pmatrix} = 1 - i d\phi J \,. \tag{2.253}$$

Therefore

$$J = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \,. \tag{2.254}$$

In fact

$$e^{-i\phi J} = \mathbb{1} - i\phi J + \frac{1}{2}(-i\phi J)^2 + ... \,, \tag{2.255}$$

$$= |\,\text{since} \quad J^2 = \mathbb{1}\,, \quad J^3 = J\,, \quad J^4 = \mathbb{1} ... \,|$$

$$= \mathbb{1} - i\phi J - \frac{\phi^2}{2}\mathbb{1} - iJ\left(-\frac{\phi^3}{3!}\right) + ... \,, \tag{2.256}$$

$$= \mathbb{1}\cos\phi - iJ\sin\phi \,, \tag{2.257}$$

$$= \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} = D(\phi) \,. \tag{2.258}$$

Knowing J and ϕ we can have all the elements of the group.

The study of the algebra (then the generators) gives many pieces of information about the group. In particular, it gives the "local" properties. However, from the exponential form we cannot extract the global relation $D(\phi) = D(\phi + 2\pi)$, which we have to impose separately. This global relation is important in the study of the irreducible representations of the group.

Irreducible representations

Let us consider a representation (labeled by R) of the group on a finite-dimensional vector space. We have

$$D_R(\phi) = e^{-i\phi J} \,. \tag{2.259}$$

Moreover, we have

$$D_R(\phi_1) D_S(\phi_2) = D_S(\phi_1 + \phi_2) \,, \tag{2.260}$$

$$D_R(\phi) = D_R(\phi + 2\pi) \,, \tag{2.261}$$

and these relations have to be satisfied by every possible representation.

J is the generator 'in the representation R" and it is an hermitian operator on V. Because of that, it is diagonalizable, it has real einvalues and the eigenvectors form an ortonormal basis for V.

Since $D_R(\phi) = f(J)$ (it is a function of J), we have

$$[D_R(\phi), J] = 0 \,. \tag{2.262}$$

Therefore, they have a common basis of eigenvectors, say $|\alpha\rangle$, on which

$$J|\alpha\rangle = \alpha|\alpha\rangle\,, \tag{2.263}$$
$$D_R(\phi)|\alpha\rangle = e^{-i\phi\alpha}|\alpha\rangle\,. \tag{2.264}$$

From these relations α can be whatever. However, we still have to impose the global relation

$$D_R(\phi) = D_R(\phi + 2\pi)\,. \tag{2.265}$$

If we do that, we find that $\alpha \equiv m \in \mathbb{Z}$. Therefore

$$J|m\rangle = m|m\rangle\,, \tag{2.266}$$
$$D_R(\phi)|m\rangle = e^{-i\phi m}|m\rangle\,. \tag{2.267}$$

These are all one-dimensional representations, as it was expected from the fact that the group is abelian.

The representation is diagonal in the basis of eigenvectors $|m\rangle$ and has $e^{-i\phi m}$ as eigenvalues. Therefore, the representation is completely reducible in irreducible one-dimensional representations.

Every $|m\rangle$ is invariant under $D_R(\phi)$ and therefore we can express $D_R(\phi)$ as a direct sum of irriducible representations $D_m(\phi) = e^{-i\phi m}$

$$D_R(\phi) = \sum_{\oplus} D_m(\phi)\,. \tag{2.268}$$

Diagonalizing J, the generator of the group in the representation R, we found the irreducible representations.

Let us have a closer look to these irreducible representations.

1. If $m = 0$ we have $D_0(\phi) = \mathbb{1}$, therefore the trivial representation (the identity)
2. If $m = 1$ we have $D_1(\phi) = e^{-i\phi}$. This is an isomorphism between elements of $SO(2)$ and numbers on the unit circle in the complex plane. When ϕ ranges in the closed interval $[0, 2\pi]$, $D_1(\phi) = e^{-i\phi}$ covers the unit circle clockwise.
3. If $m = -1$ we have the same as above, but anti-clockwise.
4. $m = \pm 2$ covers the unit circle twice.
5. ... etc ...

Among these representations, only $m = \pm 1$ are faithful (one-to-one).

If we now go back to the representation on the Euclidean 2-dim vector space (2-dim representation) we understand that it has to be reducible to two 1-dim irreducible representations. It is indeed equivalent to the direct sum of $m = \pm 1$ representations. In fact

$$J = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \tag{2.269}$$

has two eigenvalues, ± 1. The corresponding eigenvectors are

$$\hat{\mathbf{e}}_\pm = \frac{\hat{\mathbf{e}}_1 \pm i\hat{\mathbf{e}}_2}{\sqrt{2}}\,. \tag{2.270}$$

With respect to the new basis we have

$$J\hat{\mathbf{e}}_\pm = \pm\hat{\mathbf{e}}_\pm\,, \tag{2.271}$$
$$D_R(\phi)\hat{\mathbf{e}}_\pm = e^{\pm i\phi}\hat{\mathbf{e}}_\pm\,. \tag{2.272}$$

2.9 The generators of the Poincaré group and the algebra

A general Lorentz transformation can be found as the product of a transformation of the proper hortocronous group and parity and/or time reversal. Therefore, the study of the Lorentz group can be "reduced" to the study of the proper hortocronous group and, separately, of P and T. We start with $L_+^\uparrow$. Parity and Time reversal will follow.

For the study of the generators of the Poincaré group we will consider infinitesimal transformations. We will have:

$$T(\Lambda, a) \simeq T(\delta^{\mu\nu} + \delta\omega^{\mu\nu}, \delta a^\mu)\,, \tag{2.273}$$

where $\delta\omega^{\mu\nu}$ is an antisymmetric tensor of rank 2,

$$\delta\omega^{\mu\nu} = -\delta\omega^{\nu\mu}\,, \tag{2.274}$$

which depends on the parameters of the infinitesimal proper Lorentz transformation and on δa^μ, an infinitesimal four-vector.

The infinitesimal transformation (2.273) will be composed by a translation:

$$T(1, \delta a^\mu) = 1 - i\delta a_\mu P^\mu\,, \tag{2.275}$$

and by a proper Lorentz transformation:

$$T(\delta^{\mu\nu} + \delta\omega^{\mu\nu}, 0) = 1 - \frac{i}{2}\delta\omega_{\mu\nu} J^{\mu\nu}\,. \tag{2.276}$$

The four-vector P^μ is the generator of the translations, while the antisymmetric tensor or rank 2 $J^{\mu\nu} = -J^{\nu\mu}$ id the generator of the proper Lorentz transformations (boosts and tri-dimensional rotations)[2]. In total, $J^{\mu\nu}$ and P^μ constitute a system of 10 operators: 4 components of P^μ, together with the 6 components of the antisymmetric tensor $J^{\mu\nu}$.

The exponentiation of Eq. (2.275) and (2.276) gives the following relations for the finite transformations

$$T(1, a) = e^{-i\,P^\mu a_\mu}\,, \quad T(\Lambda, 0) = e^{-\frac{i}{2}\,J^{\mu\nu}\omega_{\mu\nu}}\,. \tag{2.278}$$

Let us find, now, the behaviour of the generators under Poincaré transformations.

We label with $T(\Lambda, b)$ the generic finite transformation, and with $T(1 + \delta\omega, \delta a)$ the infinitesimal transformation, on which we will focus our attention. The trasformation of $T(1+\delta\omega, \delta a)$ under $T(\Lambda, b)$ is given by::

$$\begin{aligned} T(\Lambda, b)T(1 + \delta\omega, \delta a)T^{-1}(\Lambda, b) &= T(\Lambda[1 + \delta\omega], \Lambda\delta a + b)\,T^{-1}(\Lambda, b) = \\ &= T(\Lambda[1 + \delta\omega], \Lambda\delta a + b)\,T(\Lambda^{-1}, -\Lambda^{-1}b) = \\ &= T(\Lambda[1 + \delta\omega]\Lambda^{-1}, -[\Lambda(1 - \delta\omega)]\Lambda^{-1}b + \Lambda\delta a + b) = \\ &= T(\Lambda[1 + \delta\omega]\Lambda^{-1}, \Lambda\delta a - \Lambda\delta\omega\Lambda^{-1}b)\,. \end{aligned} \tag{2.279}$$

Expanding at first order in the parameters the infinitesimal transformation in Eq. (2.279), we find:

$$\begin{aligned} T(\Lambda, b)\left(-\frac{i}{2}\delta\omega_{\mu\nu}J^{\mu\nu} - i\delta a_\mu P^\mu\right)T^{-1}(\Lambda, b) &= \\ = -\frac{i}{2}\left(\Lambda\delta\omega\Lambda^{-1}\right)_{\mu\nu} J^{\mu\nu} - i\left(\Lambda\delta a - \Lambda\delta\omega\Lambda^{-1}b\right)_\mu P^\mu\,, \end{aligned} \tag{2.280}$$

---

[2]In quantum mechanics, both P^μ and $J^{\mu\nu}$ are operators. If we look for unitary representations of the group, $T(\delta^{\mu\nu} + \omega^{\mu\nu}, \epsilon^\mu)$ is unitary and, therefore, the generators are hermitian:

$$J^{\mu\nu\dagger} = J^{\mu\nu}\,, \qquad P^{\mu\dagger} = P^\mu\,. \tag{2.277}$$

where we have

$$\left(\Lambda\delta\omega\Lambda^{-1}\right)_{\mu\nu} = \Lambda_{\mu\sigma}\delta\omega^{\sigma\rho}(\Lambda^{-1})_{\rho\nu} = \delta\omega^{\sigma\rho}\Lambda_{\mu}^{\cdot\alpha}\eta_{\alpha\sigma}\eta_{\rho\beta}(\Lambda^{-1})^{\beta}_{\nu} = \delta\omega^{\sigma\rho}\eta_{\alpha\sigma}\eta_{\rho\beta}\Lambda_{\mu}^{\cdot\alpha}\Lambda_{\nu}^{\cdot\beta}\,, \quad (2.281)$$

$$= \delta\omega_{\alpha\beta}\Lambda_{\mu}^{\cdot\alpha}\Lambda_{\nu}^{\cdot\beta} \quad (2.282)$$

and

$$\left(\Lambda\delta\omega\Lambda^{-1}b\right)_{\mu}P^{\mu} = \delta\omega_{\rho\sigma}\Lambda_{\alpha}^{\cdot\rho}\Lambda_{\mu}^{\cdot\sigma}b^{\alpha}P^{\mu}\,, \quad (2.283)$$

$$= \frac{1}{2}\delta\omega_{\rho\sigma}\left(\Lambda_{\alpha}^{\cdot\rho}\Lambda_{\mu}^{\cdot\sigma}b^{\alpha}P^{\mu} - \Lambda_{\alpha}^{\cdot\sigma}\Lambda_{\mu}^{\cdot\rho}b^{\alpha}P^{\mu}\right)\,, \quad (2.284)$$

$$= \frac{1}{2}\delta\omega_{\rho\sigma}\Lambda_{\alpha}^{\cdot\rho}\Lambda_{\mu}^{\cdot\sigma}\left(b^{\alpha}P^{\mu} - b^{\mu}P^{\alpha}\right)\,. \quad (2.285)$$

Eq. (2.280), then, gives the following transformation laws:

$$T(\Lambda,b)J^{\mu\nu}T^{-1}(\Lambda,b) = \Lambda_{\rho}^{\cdot\mu}\Lambda_{\sigma}^{\cdot\nu}\left(J^{\rho\sigma} - (b^{\rho}P^{\sigma} - b^{\sigma}P^{\rho})\right)\,, \quad (2.286)$$

$$T(\Lambda,b)P^{\mu}T^{-1}(\Lambda,b) = \Lambda_{\nu}^{\cdot\mu}P^{\nu}\,, \quad (2.287)$$

i.e. P^{μ} and $J^{\mu\nu}$ transform as a four-vector and a rank-2 tensor, respectively.

If now we also consider $T(\Lambda,b)$ as an infinitesimal transformation and we expand it at first order, Eq. (2.280) gives the following commutation rules for the generators of the Poincaré algebra:

$$[P^{\mu},P^{\nu}] = 0\,, \quad (2.288)$$

$$[P^{\mu},J^{\lambda\sigma}] = i\left(P^{\lambda}\eta^{\mu\sigma} - P^{\sigma}\eta^{\mu\lambda}\right)\,, \quad (2.289)$$

$$[J^{\mu\nu},J^{\rho\sigma}] = i\left(J^{\nu\sigma}\eta^{\mu\rho} + J^{\rho\nu}\eta^{\sigma\mu} - J^{\mu\sigma}\eta^{\nu\rho} - J^{\rho\mu}\eta^{\sigma\nu}\right)\,. \quad (2.290)$$

The generator of the translations, P^{μ}, can be identified with the four-momentum operator; the time component, P^0 is the energy of the system (hamiltonian), while the vector $\mathbf{P} = (P^1,P^2,P^3)$ is the tri-momentum. The tensor $J^{\mu\nu}$ is connected with the angular momentum and the boosts. In particular, angular momentum is given by the three components

$$\mathbf{J} = (J^{23},J^{31},J^{12})\,, \quad (2.291)$$

while the other components are the generators of the boosts:

$$\mathbf{K} = (J^{10},J^{20},J^{30})\,. \quad (2.292)$$

For these quantities, we can find the following commutation rules, that come from Eqs. (2.290):

$$[J_i,J_j] = i\epsilon_{ijk}J_k\,, \quad (2.293)$$

$$[J_i,P^0] = [P_i,P^0] = 0\,, \quad (2.294)$$

$$[J_i,P_j] = i\epsilon_{ijk}P_k\,, \quad (2.295)$$

$$[K_i,K_j] = -i\epsilon_{ijk}J_k\,, \quad (2.296)$$

$$[J_i,K_j] = i\epsilon_{ijk}K_k\,. \quad (2.297)$$

It has to be noticed that, while the components of the angular momentum constitutes a closed algebra (see (2.293)), the components of the boosts do not (see (2.296).

2.10 Finite-dimensional irreducible representations of the Poincaré group

The "most interesting part" of the study of the irreducible representations of the Poincaré group lies in the fact that the fields transform under Poincaré transformations according to a certain finite-dimensional representations of the group.

A field is a function of the point in the events space M (Minkowski space) with values in $\mathcal{C}^n$:

$$\phi(X) \quad : \quad M \longrightarrow \mathcal{C}^n \tag{2.298}$$
$$\phi(X) \quad = \quad \big(\phi_1(X), \phi_2(X), ..., \phi_n(X)\big)\,, \tag{2.299}$$

i.e. in general a function of X^μ, with a certain number of complex components. If we act on the space-time point with a Poincaré tranformation $X^\mu \to X'^\mu = \Lambda^\mu_\nu X^\nu + a^\mu$, i.e. a change of reference system, the field $\phi(X)$ will transform according to the following linear homogeneous transformation

$$\phi(X) \to \phi'(X') \,=\, S(\Lambda)\,\phi(X)\,, \tag{2.300}$$

where $S(\Lambda)$ is a representation of the Lorentz group on the vector space where the vector $\phi(X)$ lives[3].

We are interested on finite-dimensional representations, i.e. linear operators that act on vectors $\phi(X)$ with a finite number of components.

2.10.1 Tensor fields. Integer spin representations

We call *tensor field* an object with m contravariant and n covariant indices, $T^{\mu_1....\mu_m}_{\alpha_1....\alpha_n}(X)$, such that for $X^\mu \to X'^\mu \,=\, \Lambda^\mu_\nu X^\nu$ we have the following transformation:

$$T^{\mu_1....\mu_m}_{\alpha_1....\alpha_n}(X) \stackrel{X\to X'}{\longrightarrow} T'^{\mu_1....\mu_m}_{\alpha_1....\alpha_n}(X') \,=\, \Lambda^{\mu_1}_{\nu_1}..\Lambda^{\mu_m}_{\nu_m}\,\Lambda_{\alpha_1}^{\cdot\beta_1}..\Lambda_{\alpha_n}^{\cdot\beta_n}\,T^{\nu_1....\nu_m}_{\beta_1....\beta_n}(X)\,, \tag{2.302}$$

where $\Lambda_\alpha^{\cdot\beta} \,=\, (\Lambda^{-1})^\alpha_\beta$.

We will indicate the general tensor $T^{\mu_1....\mu_m}_{\alpha_1....\alpha_n}$ of rank m, n with the symbol (m, n). The space of tensors (m, n) is a linear space, i.e. the linear combinations of (m, n) tensors is an (m, n) tensor:

$$(m, n) \,\oplus\, (m, n) \,\longrightarrow\, (m, n)\,. \tag{2.303}$$

Moreover, we can define different operations as:

- il *tensor product*, $\otimes$, such that:

$$(m, n) \,\otimes\, (m', n') \,\longrightarrow\, (m + m', n + n')\,, \tag{2.304}$$

  i.e. such that the product of two tensors of ranks (m, n) and (m', n') is a tensor of rank $(m + m', n + n')$;

- la *contraction* of two indices. If we match a contravariant and a covariant indices, the rank of the tensor is lowered by 1 unit in m and 1 in n. For instance

$$\begin{array}{ccc} T^{\alpha\beta\gamma}_\delta & \longrightarrow & T^{\alpha\beta\gamma}_\beta = K^{\alpha\gamma} \\ (3,1) & \longrightarrow & (2,0) \end{array} \tag{2.305}$$

- *raising* or *lowering* of indices. Through the matric tensor $\eta_{\mu\nu}$ we can make a contravariant index covariant, and viceversa:

$$T^{\alpha\beta\gamma}_\delta \;=\; \eta_{\delta\rho}\, T^{\alpha\beta\gamma\rho}\,. \tag{2.306}$$

[3]The reason why we said that, indeed, the field transforms with a representation of the Lorentz group, i.e. the homogeneous part of the Poincaré group, is because every field is a scalar under rigid translations. For every field (scalar, vector, spinor ...) we have

$$\phi'(X^\mu + a^\mu) = \phi(X^\mu)\,. \tag{2.301}$$

Scalar Field. Trivial representation of the Lorentz group

The first field that we will consider is the *Scalar Field*. This is a "tensor" of rank $(0,0)$ and therefore it transforms, under Lorentz transformation, according to the trivial representation of the group $S(\Lambda) = 1$:

$$\phi(X) \longrightarrow \phi'(X') = \phi(X)\,. \tag{2.307}$$

Vector Field. Four-dimensional representation of the Lorentz group

Let us consider now a tensor field of rank $(1,0)$:

$$V^\mu(X) = \left(V^0(X), \mathbf{V}(X)\right), \tag{2.308}$$

that we will call *contravariant four-vector*. Under a Lorentz transformation, the field V^μ transforms as the space-time point, X^μ, i.e.:

$$V^\mu(X) \longrightarrow V'^\mu(X') = \Lambda^\mu_\nu\, V^\nu(X)\,. \tag{2.309}$$

An example of such a field is the electromagnetic four-potential $A^\mu(X)$.

We can consider a tensor field of rank $(0,1)$, $U_\mu(X)$, and it will transform as follows::

$$U_\mu(X) \longrightarrow U'_\mu(X') = \Lambda_\mu^{\cdot\nu}\, U_\nu(X)\,, \tag{2.310}$$

i.e. $U_\mu(X)$ transforms with the inverse of Λ. U_μ is a *covariant four-vector*. An example of such a type of vectors is the gradient of a scalar field, $\partial_\mu\phi(X) = \frac{\partial\phi(X)}{\partial X^\mu}$.

Il prodotto tensoriale fra un vettore controvariante ed uno covariante, con indici saturati, $V^\mu U_\mu$, è uno scalare di Lorentz. Infatti:

$$\begin{aligned} V^\mu(X)\,U_\mu(X) \longrightarrow V'^\mu(X)\,U'_\mu(X) &= \Lambda^\mu_\nu \Lambda_\mu^{\cdot\rho}\, V^\nu(X)\,U_\rho(X) = & (2.311)\\ &= \delta^\rho_\nu\, V^\nu(X)\,U_\rho(X) = & (2.312)\\ &= V^\nu(X)\,U_\nu(X) & (2.313) \end{aligned}$$

In particolare la norma di un quadrivettore $V^\mu V_\mu = \|V\|^2$ è un invariante di Lorentz.

Allora, nel caso di campo vettoriale si ha $S(\Lambda) = \Lambda^\mu_\nu$ che agisce sulle quattro componenti del quadrivettore V^μ. Abbiamo trovato una rappresentazione quaridimensionale del Gruppo di Lorentz. Siccome questo è costituito da boosts e dalle rotazioni in tre dimensioni, potremo considerare 6 matrici: $(\tilde\Lambda^\mu_\nu)_x$, $(\tilde\Lambda^\mu_\nu)_y$, $(\tilde\Lambda^\mu_\nu)_z$, per i boosts lungo gli assi cartesiani e R_x, R_y e R_z per le rotazioni intorno agli stessi. Per esempio si ha:

$$(\tilde\Lambda^\mu_\nu)_x = \begin{pmatrix} \cosh\phi_x & \sinh\phi_x & 0 & 0\\ \sinh\phi_x & \cosh\phi_x & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{e} \quad R_z = \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & \cos\theta_z & \sin\theta_z & 0\\ 0 & -\sin\theta_z & \cos\theta_z & 0\\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.314}$$

dove ϕ_x coinvolge la velocità di traslazione lungo l'asse delle x e θ_x è l'angolo di rotazione attorno allo stesso asse.

Per trovare nella stessa rappresentazione matriciale i generatori del Gruppo, basta ricordare che per il mapping esponenziale si ha:

$$K_x = -i\frac{d}{d\phi_x}(\tilde\Lambda^\mu_\nu)_x\bigg|_{\phi_x=0} = -i\begin{pmatrix} 0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.315}$$

$$K_y = -i\frac{d}{d\phi_y}(\tilde{\Lambda}^\mu_\nu)_y\Big|_{\phi_y=0} = -i\begin{pmatrix}0&0&1&0\\0&0&0&0\\1&0&0&0\\0&0&0&0\end{pmatrix} \quad (2.316)$$

$$K_z = -i\frac{d}{d\phi_z}(\tilde{\Lambda}^\mu_\nu)_z\Big|_{\phi_z=0} = -i\begin{pmatrix}0&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&0\end{pmatrix} \quad (2.317)$$

e i generatori delle rotazioni sono semplicemente dati da quelli trovati in tre dimensioni euclidee con l'aggiunta di una riga ed una colonna di zeri per la parte temporale:

$$J_x = -i\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&1\\0&0&-1&0\end{pmatrix} \quad (2.318)$$

$$J_y = -i\begin{pmatrix}0&0&0&0\\0&0&0&-1\\0&0&0&0\\0&1&0&0\end{pmatrix} \quad (2.319)$$

$$J_z = -i\begin{pmatrix}0&0&0&0\\0&0&1&0\\0&-1&0&0\\0&0&0&0\end{pmatrix}\,. \quad (2.320)$$

È da notare che, mentre le matrici J_i sono hermitiane e quindi la loro esponenziazione porta ad una matrice unitaria, lo stesso non si verifica per le K_i, che non sono hermitiane. Questo è dovuto al fatto che il gruppo $SO(3)$ è un gruppo compatto e quindi ammette rappresentazioni unitarie, mentre l'intero Gruppo di Lorentz (con i boosts) non è compatto e quindi non ammette rappresentazioni unitarie.

I generatori del gruppo formano l'algebra di Lie associata ed infatti si può verificare che, nella rappresentazione matriciale quadridimensionale appena data, valgono le (2.293), (2.296) e (2.297). Mentre le espressioni matriciali sono peculiari della rappresentazione cercata, le regole che definiscono l'algebra associata hanno carattere universale. Per ogni rappresentazione del Gruppo di Lorentz, i generatori del gruppo devono soddisfare le (2.293), (2.296) e (2.297).

La rappresentazione (2.317, 2.320) del Gruppo di Lorentz viene denotata con $SO(3,1)$.

N.B. L'importanza dello studio delle proprietà di trasformazione delle quantità con cui abbiamo a che fare nella costruzione della teoria e lo sviluppo delle notazioni tensoriali dipendono da quanto segue. Supponiamo che nel sistema di riferimento S una legge fisica sia espressa da un'uguaglianza tensoriale:

$$T^\alpha_\beta = U^\alpha_\beta\,. \quad (2.321)$$

Dei due tensori che esprimono la legge (2.321) sappiamo esattamente le proprietà di trasformazione sotto il Gruppo di Lorentz. Allora in un'altro sistema di riferimento S', si avrà:

$$T'^\alpha_\beta = \Lambda^\alpha_\gamma \Lambda_\beta^{\cdot\delta}\, T^\gamma_\delta = \quad (2.322)$$

$$= \Lambda^\alpha_\gamma \Lambda_\beta^{\cdot\delta}\, U^\gamma_\delta = \quad (2.323)$$

$$= U'^\alpha_\beta\,, \quad (2.324)$$

dove per passare dalla (2.322) alla (2.323) abbiamo sfruttato la (2.321). La (2.324) ci dice che una relazione fra tensori rimane invariata in forma sotto trasformazioni di Lorentz: in S' vale la stessa relazione per i tensori trasformati.

Si dice che la legge (2.321) è covariante a vista.

2.10.2 Campi spinoriali. Spinori di Dirac

Cerchiamo, adesso, una rappresentazione bidimensionale del Gruppo di Lorentz.

Per far questo, prendiamo spunto dall'omomorfismo fra il gruppo speciale delle rotazioni in tre dimensioni, $SO(3)$, ed il suo ricoprimento universale, $SU(2)$, gruppo delle trasformazioni unitarie speciali su uno spazio bidimensionale.

$SO(3)$ dipende da tre parametri (per esempio i tre angoli di Eulero) e ne possiamo dare una rappresentazione in termini di matrici 3×3 reali relative alle tre rappresentazioni lungo gli assi cartesiani: $R_x(\theta_x)$, $R_y(\theta_y)$ e $R_z(\theta_z)$. Più in generale, se $\mathbf{n}$ individua una direzione nello spazio euclideo, una rotazione di un angolo θ eseguita intorno all'asse $\mathbf{n}$ secondo la regola della mano destra può essere scritta come segue:

$$R_{\mathbf{n}}(\theta) \,=\, e^{i\mathbf{J}\cdot\boldsymbol{\theta}}\,, \tag{2.325}$$

dove abbiamo posto $\boldsymbol{\theta} = \theta\mathbf{n}$ e dove le matrici J_i sono i generatori del gruppo $SO(3)$ e soddisfano le regole dell'algebra di Lie associata:

$$[J_i, J_j] \,=\, i\epsilon_{ijk}J_k\,. \tag{2.326}$$

Prima di tutto mostriamo come anche $SU(2)$ dipenda da tre parametri reali. Possiamo dare una rappresentazione del gruppo in termini di matrici 2×2 complesse:

$$U \,=\, \begin{pmatrix} a & b \\ c & d \end{pmatrix}\,, \tag{2.327}$$

con $a, b, c, d \in \mathcal{C}$. Siccome la U deve essere unitaria ($U^\dagger U = 1$) e a determinante $detU = +1$, gli 8 parametri indipendenti che sembrano definire la U nella (2.327) si riducono a tre. Infatti, la prima richiesta porta a 4 relazioni reali e la seconda ad un'altra. Sotto queste imposizioni la U diventa:

$$U \,=\, \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}\,. \tag{2.328}$$

Una matrice siffatta agisce su uno spazio vettoriale complesso a due dimensioni, detto spazio degli spinori a due componenti:

$$\xi \,=\, \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \qquad \xi \longrightarrow \xi' \,=\, U\,\xi \tag{2.329}$$

I generatori del gruppo $SU(2)$ sono le matrici di Pauli σ_i:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \;;\quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \;;\quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,, \tag{2.330}$$

che obbediscono alle seguenti regole di commutazione (dell'algebra associata):

$$[\frac{1}{2}\sigma_i, \frac{1}{2}\sigma_j] \,=\, i\epsilon_{ijk}\,\frac{1}{2}\sigma_k\,. \tag{2.331}$$

Una trasformazione di $SU(2)$ preserva la forma quadratica $(x^2 + y^2 + z^2)$. Infatti, consideriamo il raggio vettore $\mathbf{r} = (x, y, z)$ e la matrice

$$h \,=\, \sigma \cdot \mathbf{r} \,=\, \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix}\,, \tag{2.332}$$

hermitiana a traccia nulla.

Trasformiamo la h tramite la trasformazione di similitudine:

$$h \longrightarrow h' \,=\, U\,h\,U^\dagger\,. \tag{2.333}$$

La (2.333) conserva l'hermiticità e la traccia nulla. Infatti, si ha:

$$h'h'^{\dagger} \,=\, UhU^{\dagger}\,(UhU^{\dagger})^{\dagger} \,=\, UhU^{\dagger}Uh^{\dagger}U^{\dagger} \,=\, Uhh^{\dagger}U^{\dagger} \,=\, 1\,, \tag{2.334}$$

dove abbiamo utilizzato $U^{\dagger}U = 1$ e $hh^{\dagger} = 1$; e:

$$tr\,h' \,=\, tr\,(UhU^{\dagger}) \,=\, tr\,(U^{\dagger}Uh) \,=\, tr\,h\,, \tag{2.335}$$

per la proprietà ciclica della traccia.

Inoltre, la (2.333) conserva anche il determinante, che non è altro che la forma quadratica $\|\mathbf{r}\|^2$ con un segno cambiato:

$$x'^2 + y'^2 + z'^2 \,=\, -det\,h' \,=\, -det\,h \,=\, x^2 + y^2 + z^2\,. \tag{2.336}$$

Per questo è logico che $SU(2)$ sia legato in qualche modo al gruppo delle rotazioni $SO(3)$. In particolare si ha una corrispondenza 2 a 1 fra gli elementi di $SU(2)$ e gli elementi di $SO(3)$ e le rappresentazioni irriducibili di $SO(3)$ sono contenute in quelle di $SU(2)$ (si dice che $SU(2)$ è il ricoprimento universale di $SO(3)$).

La corrispondenza fra i due gruppi si può riassumere in:

$$U \,=\, e^{\frac{i}{2}\sigma\cdot\theta} \quad\Longleftrightarrow\quad R \,=\, e^{i\mathbf{J}\cdot\theta}\,, \tag{2.337}$$

$$\frac{1}{2}\sigma_i \quad\Longleftrightarrow\quad J_i\,. \tag{2.338}$$

Cerchiamo, allora, una rappresentazione bidimensionale del Gruppo di Lorentz.

Per le rotazioni potremo utilizzare la (2.338).

Per trovare i generatori bidimensionali dei boosts, $\mathbf{K}^{(2)}$, sfruttiamo le regole di commutazione dell'algebra, che sono verificate per qualunque rappresentazione:

$$[K_i^{(2)}, K_j^{(2)}] \,=\, -i\epsilon_{ijk}\,\frac{1}{2}\sigma_k \tag{2.339}$$

$$[\frac{1}{2}\sigma_i, K_j^{(2)}] \,=\, i\epsilon_{ijk}\,K_k^{(2)} \tag{2.340}$$

$$[\frac{1}{2}\sigma_i, \frac{1}{2}\sigma_j] \,=\, i\epsilon_{ijk}\,\frac{1}{2}\sigma_k\,. \tag{2.341}$$

Si trova:

$$K_i^{(2)} \,=\, \pm\frac{i}{2}\sigma_i\,, \tag{2.342}$$

cioè abbiamo due possibili rappresentazioni, una con $K_i^{(2)} = \frac{i}{2}\sigma_i$ e l'altra con $K_i^{(2)} = -\frac{i}{2}\sigma_i$. Corrispondentemente si avrammo due tipi di spinori, che sotto boost si trasformano in maniera diversa:

$$\phi_R \;\;\rightarrow\;\; \phi'_R \,=\, \exp\Big\{\frac{i}{2}\sigma\cdot(\theta - i\phi)\Big\}\,\phi_R\,, \tag{2.343}$$

$$\phi_L \;\;\rightarrow\;\; \phi'_L \,=\, \exp\Big\{\frac{i}{2}\sigma\cdot(\theta + i\phi)\Big\}\,\phi_L\,, \tag{2.344}$$

dove θ sono i parametri della rotazione e ϕ quelli dei boosts.

È da notare che ϕ_R e ϕ_L (spinori *destri* e *sinistri*) si trasformano allo stesso modo sotto rotazioni. Infatti, se consideriamo soltanto rappresentazioni irriducibili del gruppo delle rotazioni in teoria non relativistica abbiamo un solo tipo di spinori: quelli di Pauli. L'introduzione delle trasformazioni di Lorentz, invece, distingue fra due tipi di componenti.

Questa rappresentazione bidimensionale del Gruppo di Lorentz si indica con $SL(2, C)$.

Si può vedere la corrispondenza fra le due rappresentazioni $SO(3,1)$ e $SL(2,C)$ esattamente come abbiamo stabilito la corrispondenza fra le trasformazioni di $SO(3)$ e di $SU(2)$. Consideriamo il quadrivettore $\sigma^\mu = (\sigma^0 = 1, \sigma)$ e saturiamolo con l'evento X^μ:

$$X = \sigma_\mu X^\mu = \begin{pmatrix} X^0 - X^3 & -X^1 + iX^2 \\ -X^1 - iX^2 & X^0 + X^3 \end{pmatrix}, \tag{2.345}$$

matrice hermitiana.

Se $A \in SL(2,C)$, facciamo la trasformazione su X:

$$X \longrightarrow X' = AXA^\dagger, \tag{2.346}$$

mediante la quale X' è ancora una matrice hermitiana. La (2.346) preserva il determinante di X, che non è altro che la forma quadratica

$$X^\mu X_\mu = (X^0)^2 - (X^1)^2 - (X^2)^2 - (X^3)^2, \tag{2.347}$$

cioè la norma del quadrivettore X^μ. In altre parole è una trasformazione di Lorentz.

Siccome per parità i due spinori, destro e sinistro, si trasformano l'uno nell'altro, se vogliamo considerare il Gruppo di Lorentz completo non ha più senso distinguere fra ϕ_R e ϕ_L.

Introdurremo, allora, lo *spinore di Dirac*, costituito come segue:

$$\psi = \begin{pmatrix} \phi_R \\ \phi_L \end{pmatrix} \tag{2.348}$$

tale che:

$$\psi \longrightarrow \psi' = \begin{pmatrix} \exp\left\{\frac{i}{2}\sigma \cdot (\theta - i\phi)\right\} & 0 \\ 0 & \exp\left\{\frac{i}{2}\sigma \cdot (\theta + i\phi)\right\} \end{pmatrix} \psi. \tag{2.349}$$

Gli spinori di Dirac costituiscono lo spazio vettoriale per la rappresentazione irriducibile bidimensionale del Gruppo di Lorentz.

2.11 Infinite dimensional representations of the Poincaré group: particle states

Capitolo 3

Conservation Laws

3.1 Lagrangian formalism

Following the formalization of analytical mechanics we will study now the system in the Lagrangian formalism. This formalism is optimal for the study of the symmetries of the theory. We will use the related Hamiltonian formalism when we will quantize the system, making a correspondence between the Poisson brackets and the commutators of the operators that describe the physical observables. This will be the subject of the co-called "canonical quantization".

3.1.1 Relativistic free particle

In order to recall the basic principles of Lagrangian mechanics, let us concentrate on a simple example: the relativistic free particle. In classical Physics, we find the equations of motion from Least Action Principle, or Hamilton's Principle.

A physical system can be described by a function called the "Lagrangian" L (given by the difference between the kinetic and the potential energy), which depends on the coordinates (collectively labeled with q), the velocities ($\dot{q}$ and (at most) the time[1]:

$$L = L(q, \dot{q}, t)\,. \tag{3.1}$$

The motion of the classical system is the function of time $q(t)$ that minimizes the following functional (the Action)

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt\,, \tag{3.2}$$

with respect to path variation with fixed end points

$$\delta S = 0\,. \tag{3.3}$$

Eq. (3.3) gives a system of second order differential equations, called the Euler-Lagrange equations, that for Newtonian mechanics are the generalization of the second principle $\mathbf{F} = m\mathbf{a}$.

Let us consider now a relativistic system. Our study should be independent on the inertial frame where the observer lives. In other words, our description of the system should be invariant under Lorentz transformations and therefore the Action should be invariant, in such a way that the Euler-Lagrange differential equations are unchanged in form in every inertial frame. Considering Eq. (3.2), this means that $L\,dt$ must be a Lorentz scalar. We already pointed out that dt is not a Lorentz scalar and, moreover, time derivation as dX^μ/dt does not transform as a four-vector under Lorentz transformations. Let us consider, then, instead of the time, the "proper time" of the particle, τ for which we know that $d\tau$ is actually invariant.

[1]If the system is closed the explicit dependence on time in the Lagrangian is not present

We can write the product $L\,dt$ in a manifestly invariant way as follows:

$$S = \int_{\tau_1}^{\tau_2} L(X^\mu, \dot{X}^\mu, \tau)\, d\tau\,, \tag{3.4}$$

where $\dot{X}^\mu = dX^\mu/d\tau$. Since $d\tau$ is invariant, so should be the function $L(X^\mu, \dot{X}^\mu, \tau)$[2].

Let us now consider a generic variation of the "path" $X^\mu(\tau)$. Let us consider a reparametrization $\tau \to \tau'$ and a variation of $X^\mu(\tau)$ and $\dot{X}^\mu(\tau)$ as follows:

$$\tau \;\to\; \tau' = \tau + \Delta\tau(\tau)\,, \tag{3.6}$$

$$X^\mu(\tau) \;\to\; X'^\mu(\tau') = X^\mu(\tau) + \Delta X^\mu(\tau)\,, \tag{3.7}$$

$$\dot{X}^\mu(\tau) \;\to\; \dot{X}'^\mu(\tau') = \dot{X}^\mu(\tau) + \Delta\dot{X}^\mu(\tau)\,. \tag{3.8}$$

Note that $\Delta\tau$ is a function of τ. Moreover, the variation of the path we are considering is a global variation, such that

$$\Delta X^\mu(\tau) = X'^\mu(\tau') - X^\mu(\tau)\,, \tag{3.9}$$

that include a variation for the change in the parametrization and a variation in form of the function $X^\mu(\tau)$ given a certain τ. Finally, since we are speaking about global variations, in the last equation we have to consider that $\Delta\dot{X}^\mu(\tau) \neq d/d\tau\, \Delta X^\mu(\tau)$.

Let us, moreover, consider a lagrangian that does not change in form under this reparametrization. In principle, we could have

$$L'(X'^\mu, \dot{X}'^\mu, \tau') - L(X'^\mu, \dot{X}'^\mu, \tau') = \delta L\,. \tag{3.10}$$

In order not to affect the equations of motion, δL should be the total derivative of a function that vanishes at the end points. However, let us consider $\delta L = 0$.

To the first order in the variation we have

$$L(X'^\mu, \dot{X}'^\mu, \tau') \simeq L(X^\mu, \dot{X}^\mu, \tau) + \frac{\partial L}{\partial \tau}\Delta\tau + \frac{\partial L}{\partial X^\mu}\Delta X^\mu + \frac{\partial L}{\partial \dot{X}^\mu}\Delta\dot{X}^\mu\,. \tag{3.11}$$

Since the total derivative of L with respect of τ is

$$\frac{dL}{d\tau} = \frac{\partial L}{\partial \tau} + \frac{\partial L}{\partial X^\mu}\frac{\partial X^\mu}{\partial \tau}\Delta\tau + \frac{\partial L}{\partial \dot{X}^\mu}\frac{\partial \dot{X}^\mu}{\partial \tau}\Delta\tau\,, \tag{3.12}$$

we can extract $\frac{\partial L}{\partial \tau}$ from the previous equation and substitute it in Eq. (3.11) getting

$$L(X'^\mu, \dot{X}'^\mu, \tau') \simeq L(X^\mu, \dot{X}^\mu, \tau) + \frac{dL}{d\tau}\Delta\tau + \left[\frac{\partial L}{\partial X^\mu}(\Delta X^\mu - \dot{X}^\mu\Delta\tau) + \frac{\partial L}{\partial \dot{X}^\mu}(\Delta\dot{X}^\mu - \ddot{X}^\mu\Delta\tau)\right], \tag{3.13}$$

which is written in terms of a total proper time derivative (we will use this in a while, in order to do the integration).

Consider also that

$$d\tau' = d\tau + \frac{d\Delta\tau}{d\tau}\, d\tau = \left(1 + \frac{d\Delta\tau}{d\tau}\right) d\tau\,. \tag{3.14}$$

[2] In particular, if under a Lorentz transformation the lagrangian becomes $L'(X'^\mu, \dot{X}'^\mu, \tau)$ (τ does not change under LT) we must have

$$L'(X'^\mu, \dot{X}'^\mu, \tau) = L(X'^\mu, \dot{X}'^\mu, \tau)\,. \tag{3.5}$$

This means that the form of the function should be the same (when X^μ becomes X'^μ ... etc.). Moreover, in practice we will never consider lagrangians that depend explicitely on time.

At first order we can therefore write:

$$
\begin{aligned}
\delta S &= \int_{\tau_1'}^{\tau_2'} L(X'^\mu, \dot{X}'^\mu, \tau')\, d\tau' - \int_{\tau_1}^{\tau_2} L(X'^\mu, \dot{X}^\mu, \tau)\, d\tau\,, && (3.15)\\
&\simeq \int_{\tau_1}^{\tau_2} d\tau \left(1 + \frac{d\Delta\tau}{d\tau}\right) \left\{L + \frac{dL}{d\tau}\Delta\tau + \left[\frac{\partial L}{\partial X^\mu}(\Delta X^\mu - \dot{X}^\mu\Delta\tau) + \frac{\partial L}{\partial \dot{X}^\mu}(\Delta \dot{X}^\mu - \ddot{X}^\mu\Delta\tau)\right]\right\}\\
&\quad - \int_{\tau_1}^{\tau_2} d\tau\, L\,, && (3.16)\\
&\simeq \int_{\tau_1}^{\tau_2} d\tau \left[\frac{d}{d\tau}(L\Delta\tau) + \frac{\partial L}{\partial X^\mu}(\Delta X^\mu - \dot{X}^\mu\Delta\tau) + \frac{\partial L}{\partial \dot{X}^\mu}(\Delta \dot{X}^\mu - \ddot{X}^\mu\Delta\tau)\right], && (3.17)
\end{aligned}
$$

where we discarted terms of second and higher order in the variations.

Let us now specify a bit better the variation ΔX^μ etc.

We can write

$$
\begin{aligned}
\Delta X^\mu &= X'^\mu(\tau') - X^\mu(\tau)\,, && (3.18)\\
&= X'^\mu(\tau') - X'^\mu(\tau) + X'^\mu(\tau) - X^\mu(\tau)\,, && (3.19)\\
&\simeq \dot{X}'^\mu(\tau)\,\Delta\tau + X'^\mu(\tau) - X^\mu(\tau)\,, && (3.20)\\
&\simeq \dot{X}^\mu(\tau)\,\Delta\tau + \delta X^\mu\,, && (3.21)
\end{aligned}
$$

where we neglected terms of higher order in the variations and we introduced a variation in form of the function $X^\mu(\tau)$ which consists on a variation at a fixed parameter τ. Therefore the total variation of $X^/mu(\tau)$ is represented as the sum of a variation that depends on the fact that τ varies (and therefore the derivative is involved) and a variation in form of the function, at fixed parametert τ. Note the fact that, then

$$
\delta X^\mu = (\Delta X^\mu - \dot{X}^\mu\Delta\tau)\,. \qquad (3.22)
$$

Similarly, we have

$$
\begin{aligned}
\Delta \dot{X}^\mu &= \dot{X}'^\mu(\tau') - \dot{X}^\mu(\tau)\,, && (3.23)\\
&= \dot{X}'^\mu(\tau') - \dot{X}'^\mu(\tau) + \dot{X}'^\mu(\tau) - \dot{X}^\mu(\tau)\,, && (3.24)\\
&\simeq \ddot{X}'^\mu(\tau)\,\Delta\tau + \dot{X}'^\mu(\tau) - \dot{X}^\mu(\tau)\,, && (3.25)\\
&\simeq \ddot{X}^\mu(\tau)\,\Delta\tau + \delta\dot{X}^\mu\,, && (3.26)\\
&= \ddot{X}^\mu(\tau)\,\Delta\tau + \frac{d}{d\tau}\,\delta X^\mu\,, && (3.27)
\end{aligned}
$$

where we used the fact that

$$
\delta\dot{X}^\mu = \frac{d}{d\tau}\,\delta X^\mu\,, \qquad (3.28)
$$

since the variation δ is taken at equal τ.

Substituting in Eq. (3.17) and integrating by parts we find

$$
\begin{aligned}
\delta S &\simeq \int_{\tau_1}^{\tau_2} d\tau \left[\frac{d}{d\tau}(L\Delta\tau) + \frac{\partial L}{\partial X^\mu}\delta X^\mu + \frac{\partial L}{\partial \dot{X}^\mu}\frac{d}{d\tau}\,\delta X^\mu\right], && (3.29)\\
&= \int_{\tau_1}^{\tau_2} d\tau \left[\frac{d}{d\tau}(L\Delta\tau) + \frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{X}^\mu}\delta X^\mu\right) + \left(\frac{\partial L}{\partial X^\mu} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{X}^\mu}\right)\delta X^\mu\right], && (3.30)\\
&= \int_{\tau_1}^{\tau_2} d\tau \left[\frac{d}{d\tau}\left(L\Delta\tau + \frac{\partial L}{\partial \dot{X}^\mu}\delta X^\mu\right) + \left(\frac{\partial L}{\partial X^\mu} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{X}^\mu}\right)\delta X^\mu\right], && (3.31)\\
&= \left|\text{using again } \delta X^\mu = \Delta X^\mu - \dot{X}^\mu\,\Delta\tau\right|
\end{aligned}
$$

$$= \int_{\tau_1}^{\tau_2} d\tau \left(\frac{\partial L}{\partial X^\mu} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{X}^\mu} \right) \delta X^\mu + \left(L - \frac{\partial L}{\partial \dot{X}^\mu}\dot{X}^\mu \right) \Delta\tau \Big|_{\tau_1}^{\tau_2} + \frac{\partial L}{\partial X^\mu}\Delta X^\mu \Big|_{\tau_1}^{\tau_2} . \tag{3.32}$$

Let us impose now

$$\delta S = 0 . \tag{3.33}$$

If we consider just functional variations (i.e. $\Delta\tau = 0$ and $\Delta X^\mu = \delta X^\mu$ with $\delta X^\mu(\tau_1) = \delta X^\mu(\tau_2) = 0$) that vanish in τ_1 and τ_2, we find the Hamilton's principle and therefore the Euler-Lagrange equations of motion:

$$\frac{\partial L}{\partial X^\mu} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{X}^\mu} = 0 . \tag{3.34}$$

Let us suppose now that the path is such that the Euler-Lagrange equations are satisfied. If we require that the theroy is invariant under a slightly more general variation, with a reparametrization of the curve

$$\Delta X^\mu(\tau) = 0 , \tag{3.35}$$

the third term in Eq. (3.32) vanishes and the second gives rise to the following equation

$$L = \frac{\partial L}{\partial \dot{X}^\mu}\dot{X}^\mu , \tag{3.36}$$

which means[3] that the lagrangian must be an homogeneous functions of degree one in $\dot{X}^\mu$.

In the case of the free particle we can chose[4]

$$L = \alpha\sqrt{\dot{X}^\mu \dot{X}_\mu} , \tag{3.39}$$

where α is a constant that can be fixed imposing the correct behaviour for velocities small with respect to the speed of light.

$$S = \alpha \int_{\tau_1}^{\tau_2} \sqrt{\dot{X}^\mu \dot{X}_\mu}\, d\tau = \alpha \int_{\tau_1}^{\tau_2} \sqrt{dX^\mu dX_\mu} = \alpha \int_{\tau_1}^{\tau_2} ds = \alpha c \int_{\tau_1}^{\tau_2} d\tau = \alpha c \int_{t_1}^{t_2} \sqrt{1-\beta^2}\, dt , \tag{3.40}$$

where we moved back to a non-manifestly-invariant form. In the case $\beta << 1$ we have

$$L \simeq \alpha c \left(1 - \frac{v^2}{2c^2} + ... \right) \tag{3.41}$$

and we must impose $\alpha = -mc$, in such a way that

$$L \simeq -mc^2 + \frac{mv^2}{2} + ... \tag{3.42}$$

that reproduces the correct Newtonian kinetic energy (up to a constant).

Finally

$$L = -mc\sqrt{1 - \frac{v^2}{c^2}} . \tag{3.43}$$

[3]Eq. (3.36) has also another meaning: if we write the four-momentum of the particle as

$$P_\mu = \frac{\partial L}{\partial \dot{X}^\mu} , \tag{3.37}$$

and we suppose that P_μ is the canonical conjugated momentum to X^μ, the canonical Hamiltonian is given by

$$H = \frac{\partial L}{\partial \dot{X}^\mu}\dot{X}^\mu - L = 0 . \tag{3.38}$$

Therefore we find that the Hamiltonian is identically zero. This is due to the fact that P_μ actually does not have all the components independent, but they are constrained by the mass-shell relation, $P^\mu P_\mu = m^2$. This relation makes in such a way that the trasformation from $(X^\mu, \dot{X}^\mu)$ to (X^μ, P^μ) is indeed non canonical.

[4]This ammounts to take as action the integral of the ds.

3.1.2 Euler-Lagrange Equations

Since

$$L = -mc\sqrt{\dot{X}^\mu \dot{X}_\mu}\,, \tag{3.44}$$

the variable X^μ is cyclic, $\partial L/\partial x^\mu = 0$, the equations of motion are

$$\frac{d}{d\tau}\frac{\partial L}{\partial \dot{X}^\mu} = 0\,, \tag{3.45}$$

or

$$m\dot{X}_\mu = P_\mu = const\,. \tag{3.46}$$

Remenbering the components of the four-momentum, we have

$$P_\mu = \left(\frac{mc}{\sqrt{1-\beta^2}}, -\frac{m\mathbf{v}}{\sqrt{1-\beta^2}}\right) = \left(\frac{E}{c}, -\mathbf{p}\right) = const\,. \tag{3.47}$$

Finally:

$$\begin{cases} \frac{E}{c} = const \\ \mathbf{p} = const \end{cases} \tag{3.48}$$

3.1.3 Conservation Laws

Let us consider now the invariance of the Lagrangian under Poincaré transformations

$$X^\mu \to X'^\mu = \Lambda^\mu_\nu X^\nu + a^\mu\,. \tag{3.49}$$

Note that the transformation leaves unchanged the proper time, τ. We will find that, as a consequence of this invariance, we get some conservation laws.

If we consider an infinitesimal transformation, we have

$$X^\mu \to X'^\mu = X^\mu + \delta X^\mu \tag{3.50}$$

and since

$$\Lambda^\mu_\nu \simeq \delta^\mu_\nu + \epsilon^\mu_\nu\,, \tag{3.51}$$

we will have

$$\delta X^\mu = X'^\mu - X^\mu \simeq \epsilon^\mu_\nu X^\nu + \delta a^\mu\,. \tag{3.52}$$

The Lagrangian will change accordingly:

$$L(X'^\mu, \dot{X}'^\mu, \tau) \simeq L(X^\mu, \dot{X}^\mu, \tau) + \frac{\partial L}{\partial X^\mu}\delta X^\mu + \frac{\partial L}{\partial \dot{X}^\mu}\delta \dot{X}^\mu\,, \tag{3.53}$$

$$= L(X^\mu, \dot{X}^\mu, \tau) + \frac{\partial L}{\partial X^\mu}\delta X^\mu + \frac{\partial L}{\partial \dot{X}^\mu}\frac{d}{d\tau}\delta X^\mu\,, \tag{3.54}$$

$$= L(X^\mu, \dot{X}^\mu, \tau) + \frac{\partial L}{\partial X^\mu}\delta X^\mu + \frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{X}^\mu}\delta X^\mu\right) - \delta X^\mu \frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{X}^\mu}\right)\,, \tag{3.55}$$

$$= L(X^\mu, \dot{X}^\mu, \tau) + \left(\frac{\partial L}{\partial X^\mu} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{X}^\mu}\right)\delta X^\mu + \frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{X}^\mu}\delta X^\mu\right)\,, \tag{3.56}$$

where we used the fact that the variation δX^μ is indeed a local variation ($\delta\tau = 0$) and therefore

$$\delta \dot{X}^\mu = \frac{d}{d\tau}\delta X^\mu\,. \tag{3.57}$$

If now we impose $L(X'^{\mu}, \dot{X}'^{\mu}, \tau) = L(X^{\mu}, \dot{X}^{\mu}, \tau)$, we have

$$\left(\frac{\partial L}{\partial X^{\mu}} - \frac{d}{d\tau} \frac{\partial L}{\partial \dot{X}^{\mu}} \right) \delta X^{\mu} + \frac{d}{d\tau} \left(\frac{\partial L}{\partial \dot{X}^{\mu}} \delta X^{\mu} \right) = 0 \tag{3.58}$$

and on the solution of the equations of motion finally

$$\frac{d}{d\tau} \left(\frac{\partial L}{\partial \dot{X}^{\mu}} \delta X^{\mu} \right) = 0 \,, \tag{3.59}$$

or

$$\frac{\partial L}{\partial \dot{X}^{\mu}} \delta X^{\mu} = const \,. \tag{3.60}$$

Substituting Eq. (3.52) and remembering that

$$\frac{\partial L}{\partial \dot{X}^{\mu}} = P_{\mu} \,, \tag{3.61}$$

we get the following conservation laws

$$P_{\mu} \epsilon^{\mu}_{\nu} X^{\nu} = \epsilon^{\mu\nu} P_{\mu} X_{\nu} = const \,, \tag{3.62}$$

$$P_{\mu} \delta a^{\mu} = const \,. \tag{3.63}$$

Since $\epsilon^{\mu\nu}$ is antisymmetric in the exchange of the two indices, the part of the tensor $P_{\mu} X_{\nu}$ that survives is only the antisymmetric part

$$\epsilon^{\mu\nu} P_{\mu} X_{\nu} = \epsilon^{\mu\nu} \left[\frac{1}{2} \left(P_{\mu} X_{\nu} + P_{\nu} X_{\mu} \right) + \frac{1}{2} \left(P_{\mu} X_{\nu} - P_{\nu} X_{\mu} \right) \right] = \frac{1}{2} \epsilon^{\mu\nu} \left(P_{\mu} X_{\nu} - P_{\nu} X_{\mu} \right) \,. \tag{3.64}$$

Finally, since $\epsilon^{\mu\nu}$ and δa^{μ} are constants, we have

$$M_{\mu\nu} = P_{\mu} X_{\nu} - P_{\nu} X_{\mu} = const \,, \tag{3.65}$$

$$P_{\mu} = const \,, \tag{3.66}$$

conservation of the generalized angular momentum and of the mopmentum. Note that $M_{\mu\nu}$ id an anti-symmetric tensor and therefore it has 6 independent quantities, while P_{μ} has 4. In total, the invariance under Poincaré transformations (that depend on 10 parameters) gives 10 coonserved quantites.

3.2 Lagrangian formalism for the vibrating string

Let us now consider the case of the vibrating string and study the lagrangian approach in the continuum case.

We have

$$L = T - V \,, \tag{3.67}$$

where

$$T = \frac{1}{2} \sum_{n=1}^{N} p_n^2 \to \frac{1}{2} \int_0^L \dot{\phi}^2(x,t) \, dx \,, \tag{3.68}$$

$$V = \frac{1}{2} \omega^2 \sum_{n=1}^{N} (q_n - q_{n+1})^2 \to \frac{v^2}{2} \int_0^L \phi'^2(x,t) \, dx \,. \tag{3.69}$$

In total:

$$L = \int_0^L dx \left\{ \frac{1}{2} \left[\dot{\phi}^2(x,t) - v^2 \phi'^2(x,t) \right] \right\} \,. \tag{3.70}$$

We see that we can write the lagrangian as a space integral of a "lagrangian density"

$$\mathcal{L} = \frac{1}{2}\left[\dot{\phi}^2(x,t) - v^2\phi'^2(x,t)\right], \tag{3.71}$$

$$L = \int_0^L dx\mathcal{L}. \tag{3.72}$$

The same can be found for the hamiltonian:

$$\mathcal{H} = \frac{1}{2}\left[\dot{\phi}^2(x,t) + v^2\phi'^2(x,t)\right], \tag{3.73}$$

$$H = \int_0^L dx\mathcal{H}. \tag{3.74}$$

We want not derive the vibrating string equations of motion from the Hamilton's Principle, as in the case of the point like particle. Then we will consider the following case:

1. Our system is described by a Lagrangian density, $\mathcal{L}$, local function of the fields;
2. $\mathcal{L}$ depends upon the fields, their first derivative (and at most on space point and time)

$$\mathcal{L} = \mathcal{L}\left(\phi, \dot{\phi}, \phi', x, t\right). \tag{3.75}$$

We can define the action as

$$S = \int_{t_1}^{t_2} L\, dt = \int_{t_1}^{t_2}\int_0^L \mathcal{L}\, dx\, dt \tag{3.76}$$

and we require that the equations of motion derive from the imposition of

$$\delta S = 0. \tag{3.77}$$

The variation of the fields has to be imposed to vanish on the boundary of integration:

$$\delta\phi(x,t_1) = \delta\phi(x,t_2) = 0, \quad \forall x \in [0,L], \tag{3.78}$$

$$\delta\phi(0,t) = \delta\phi(L,t) = 0, \quad \forall t \in [t_1,t_2]. \tag{3.79}$$

Moreover, note that $\delta\phi$ is a variation in form of the field, at a given point x and t. This means that

$$\delta\dot{\phi} = \frac{\partial}{\partial t}\delta\phi, \quad \text{and} \quad \delta\phi' = \frac{\partial}{\partial x}\delta\phi. \tag{3.80}$$

We have

$$0 = \delta S = \int dt\, dx\, \delta\mathcal{L} = \int dt\, dx\left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial\dot{\phi}}\delta\dot{\phi} + \frac{\partial\mathcal{L}}{\partial\phi'}\delta\phi'\right], \tag{3.81}$$

$$= \int dt\, dx\left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\delta\phi\right) - \left(\frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\right)\delta\phi + \frac{\partial}{\partial x}\left(\frac{\partial\mathcal{L}}{\partial\phi'}\delta\phi\right) - \left(\frac{\partial}{\partial x}\frac{\partial\mathcal{L}}{\partial\phi'}\right)\delta\phi\right], \tag{3.82}$$

$$= \int dx\left(\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\delta\phi\right)\Bigg|_{t_1}^{t_2} + \int dt\left(\frac{\partial\mathcal{L}}{\partial\phi'}\delta\phi\right)\Bigg|_L^0 + \int dt\, dx\left[\frac{\partial\mathcal{L}}{\partial\phi} - \frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\dot{\phi}} - \frac{\partial}{\partial x}\frac{\partial\mathcal{L}}{\partial\phi'}\right]\delta\phi. \tag{3.83}$$

Using (3.78,3.79), for the arbitraryness of $\delta\phi$, we have the Euler-Lagrange equation of motion:

$$\frac{\partial\mathcal{L}}{\partial\phi} - \frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\dot{\phi}} - \frac{\partial}{\partial x}\frac{\partial\mathcal{L}}{\partial\phi'} = 0. \tag{3.84}$$

If we consider the lagrangian density of the vibrating string, Eq. (3.71), we find the wave equation

$$\frac{1}{v^2}\frac{\partial^2\phi}{\partial t^2} - \frac{\partial^2\phi}{\partial x^2} = 0\,. \tag{3.85}$$

Knowing the lagrangian density, we can perform a Legendre transformation to get the hamiltonian density. We define the momentum conjugate to the field ϕ

$$\pi(x,t) = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} \tag{3.86}$$

and then

$$\mathcal{H} = \pi\dot{\phi} - \mathcal{L} = \frac{1}{2}\left[\pi^2 + v^2\pi'^2\right]\,. \tag{3.87}$$

3.3 Lagrangian formalism: relativistic fields

We will consider now the field as a function of the space-time point X^μ, $\phi(X)$, and we will let our system be described by a lagrangian, local funcion of the field (or fields, if we have more than one), of its derivatives and, at most, of the space-time point (as we will see, we cannot have an explicit dependence on x^μ for theories that have to be Poincaré invariant). The request of locality for the lagrangian is connected to necessity that physical quantities are observable (causality principle).

Our goal, is to include in our quantum description of microscopic phenomena special relativity. Therefore, we will require that the action, $S = \int dt\, L$, is invariant under Poincaré transformations (i.e. the action is a scalar). In fact, Physics must be independent on the inertial frame in which we describe it.

If the lagrangian is $L = L(\phi_i(X), \partial_\mu\phi_i(X), ..., \partial_\mu^{(n)}\phi_i(X), X_\mu)$, we define the lagrangian density $\mathcal{L}$, such that:

$$L = \int_V \mathcal{L}(\phi_i(X), \partial_\mu\phi_i(X), ..., \partial_\mu^{(n)}\phi_i(X) X_\mu)\, d^3X\,. \tag{3.88}$$

The action, $S(V)$, will be given by the following expression:

$$S(V) \;=\; \int_{t_1}^{t_2} Ldt = \int_{t_1}^{t_2}\int_V \mathcal{L}(\phi_i(X), \partial_\mu\phi_i(X), ..., \partial_\mu^{(n)}\phi_i(X) X_\mu)\, d^4X\,. \tag{3.89}$$

We impose that S is invariant under proper Poincaré transformations (disconituous transformations, as the parity for instance, have to be studied apart). Since the volume element d^4X is actually invariant

$$d^4X' = |\mathrm{det}\Lambda|\, d^4X = d^4X\,, \tag{3.90}$$

that follows from the fact that the proper Lorentz transformation has determinant +1, we have to impose that the lagrangian density is invariant under proper Poincaré transformations. This means, for instance, that $\mathcal{L}$ cannot depend explicitely on the space-time point.

Apart from locality and Poincaré invariance, we can constrain the lagrangian density with additional requirements: *i)* The action (and then the lagrangian) should be a real functional to avoid problems in the probabilistic interpretation of the theory; *ii)* In order to have equations of motion that are at most second order differential equations, the lagrangian density can depend upon up to first order derivative of the fields; *iii)* We can require that the lagrangian is invariant under other transformations; for instance including internal symmetries, gauge transformations and so on ...

In general, therefore, we will have to deal with lagrangian densities of the following kind:

$$\mathcal{L} \;=\; \mathcal{L}(\phi_i(X), \partial_\mu\phi_i(X))\,, \tag{3.91}$$

where the label i of the fields ϕ_i can be a collective index or a Lorentz index, depending on what we are considering.

Once the lagrangian density is defined, we can define the conjugated momenta to the fields

$$\pi_i(X) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_i} \tag{3.92}$$

and then the hamiltonian density, via a Legendre transformation

$$\mathcal{H} = \sum_i \pi_i \dot{\phi}_i - \mathcal{L}\,, \tag{3.93}$$

that coincides with the energy density of the system.

3.4 Hamilton's principle and the equations of motion

Once the lagrangian density (and the action) is specified, we can find the equations of motion for the fields ϕ_i from the Hamilton's principle. We ask then that

$$\delta S = 0\,, \tag{3.94}$$

i.e. that the action is stationary on the variations $\delta\phi_i(X)$, that will have to be the analogous of the fixed-endpoints variations of the analytical mechanics of the massive point particle.

In our case we have to deal with an integration over the space volume V and one over the time, between t_1 andt_2 (that can also be $\pm\infty$). Then, if Σ is the surface that delimits the integration volume, $\delta\phi_i(X)$ should be such that:

$$\delta\phi(\mathbf{x},t) = 0 \qquad \text{if} \quad \mathbf{x} \in \Sigma \tag{3.95}$$

$$\delta\phi(\mathbf{x},t_1) = \delta\phi(\mathbf{x},t_2) = 0 \quad \forall \quad \mathbf{x} \in V\,. \tag{3.96}$$

We will have, then

$$0 = \delta S = \int d^4X \left[\frac{\partial \mathcal{L}}{\partial \phi_i}\,\delta\phi_i + \frac{\partial \mathcal{L}}{\partial \phi_{i,\mu}}\,\delta\phi_{i,\mu}\right] = \tag{3.97}$$

$$= \int d^4X \left[\frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu\Big(\frac{\partial \mathcal{L}}{\partial \phi_{i,\mu}}\Big)\right] \delta\phi_i + \int d^4X\, \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial \phi_i}\delta\phi_i\right] = \tag{3.98}$$

$$= \int d^4X \left[\frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu\Big(\frac{\partial \mathcal{L}}{\partial \phi_{i,\mu}}\Big)\right] \delta\phi_i\,, \tag{3.99}$$

where, in order to move from (3.97) to (3.98) we integrated by parts and from (3.98) to (3.99) we used the vanishing of the field variations on the boundary of the integration domain. Since $\delta\phi_i$ is arbitrary, Eq. (3.99) gives rise to the Euler-Lagrange equations for the fields:

$$\frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \phi_{i,\mu}}\right) = 0\,. \tag{3.100}$$

We must notice that $\mathcal{L}$ is determined up to a total derivative. In fact, if $\mathcal{L}$ gives rise to the equations of motion (3.100), also $\mathcal{L}' = \mathcal{L} + \partial_\mu \Lambda^\mu(X)$ gives the same equations, provided that $\Lambda^\mu(X)$ vanishes on the boundary of the integration domain.

3.5 Global continuous symmetries and Nöether's theorem

Abbiamo visto come, nel formalismo lagrangiano, si facciano derivare le equazioni del moto dal Principio variazionale di Hamilton. Supporremo quindi che il nostro sistema fisico sia descritto da una densità di lagrangiana, funzione locale dei campi e al massimo delle loro derivate prime. Aggiungeremo l'ipotesi che $\mathcal{L}$ dipenda anche esplicitamente dal punto dello spazio-tempo X^μ, anche se in realtà poi avremo a che fare con lagrangiane indipendenti da X^μ. Questo per necessità di formulare in maniera generale il teorema di Nöether.

Supponiamo di operare sul sistema una generica trasformazione. A livello matematico ciò si tradurrà in una trasformazione sull'azione, $S(V)$, che coinvolga X^μ, $\phi_i(X)$ e $\mathcal{L}$.

Hanno particolare interesse le trasformazioni che lasciano invariata la "fisica" del problema, cioè che permettano di avere le stesse ampiezze di transizione e quindi, in ultima analisi, le stesse equazioni del moto. Trasformazioni di questo genere vengono dette **simmetrie** del sistema e generalmente hanno struttura di gruppo.

Se scriviamo una trasformazione generica come segue:

$$\begin{cases} X^\mu & \longrightarrow \quad X'^\mu \,=\, X^\mu + \delta X^\mu \\ \phi_i(X) & \longrightarrow \quad \tilde{\phi}_i(X') \,=\, \phi_i(X) + \Delta\phi_i(X) \\ \mathcal{L} & \longrightarrow \quad \tilde{\mathcal{L}}\left(\tilde{\phi}_i(X'), \tilde{\phi}_{i,\mu}(X'), X'\right) = \mathcal{L}\left(\phi_i(X), \phi_{i,\mu}(X), X\right) + \Delta\mathcal{L}\left(\phi_i(X), \phi_{i,\mu}(X), X\right) \end{cases} \tag{3.101}$$

si avrà corrispondentemente:

$$S(\mathcal{V}) \longrightarrow S'(\mathcal{V}') \,=\, \int_{\mathcal{V}'} d^4X'\, \tilde{\mathcal{L}}\left(\tilde{\phi}_i(X'), \tilde{\phi}_{i,\mu}(X'), X'\right)\,, \tag{3.102}$$

dove $\mathcal{V}$ è il volume quadridimensionale d'integrazione.

Le Eqs. (3.101) costituiscono una simmetria del sistema se si ha:

$$S'(\mathcal{V}') \,=\, S(\mathcal{V})\,. \tag{3.103}$$

L'importanza del teorema di Nöether sta nel fatto che questo asserisce che ad ogni simmetria continua del sistema viene associata una legge di conservazione locale, ovvero una quantità conservata, che possiamo identificare quantisticamente come un'osservabile. Il numero delle quantità conservate è pari al numero di parametri indipendenti da cui dipende la trasformazione (3.101). Quindi lo studio delle simmetrie del sistema ci permette di fare un salto nella trattazione del problema e di individuare subito un certo numero di osservabili.

È da notare che la richiesta (3.103) rappresenta la simmetria più generale possibile: non è detto che non esistano delle simmetrie più limitate. Per esempio un certa trasformazione può lasciare invariata la lagrangiana o la densità di lagrangiana e queste implicano a loro volta la (3.103). Consideriamo quindi il caso generale e poi ci limiteremo ad alcuni casi più restrittivi.

Cominciamo col puntualizzare alcune cose a proposito delle (3.101).

Le trasformazioni che considereremo in questo paragrafo sono tutte trasformazioni infinitesime, alle quali ci limitiamo perché stiamo considerando trasformazioni continue, il cui comportamento è deducibile da quello nell'intorno dell'identità.

Queste trasformazioni possono agire sullo spazio-tempo, $X^\mu \to X'^\mu$, ed indurre quindi una corrispondente variazione sul campo ϕ_i, $\phi_i(X) \to \tilde{\phi}_i(X')$ (*simmetrie geometriche*), ma possono anche agire soltanto sulla forma funzionale del campo ϕ_i, indipendentemente dal punto in cui essa è valutata (*simmetrie interne*). Quindi, la variazione del campo $\phi_i(X)$ comprende genericamente le due possibilità. Per esempio, una trasformazione di Lorentz sullo spazio-tempo, cioè il passaggio da un sistema di riferimento inerziale ad un altro nello studio della fisica di un problema, indurrà una conseguente trasformazione sui campi dovuta alla diversa natura di questi: se si ha un campo scalare si avrà

$\tilde{\phi}(X') = \phi(X)$, mentre per un campo tensoriale o spinoriale la trasformazione $X'^{\mu} = \Lambda^{\mu}_{\nu} X^{\nu}$ determinerà la trasformazione $\phi'(X') = S(\Lambda)\phi(X)$ nelle rispettive rappresentazioni del gruppo. Oppure, senza trasformazioni dello spazio-tempo, potremo pensare ad una simmetria sotto la ridefinizione dei campi ϕ_i.

Definiamo genericamente la *variazione totale* di $\phi_i(X)$ e $\mathcal{L}$ come segue:

$$\begin{aligned} \Delta\phi_i(X) &= \tilde{\phi}_i(X') - \phi_i(X) = \tilde{\phi}_i(X') - \tilde{\phi}_i(X) + \tilde{\phi}_i(X) - \phi_i(X)\,, && (3.104)\\ &\simeq \partial_\mu \tilde{\phi}_i(X)\, \delta X^{\mu} + \tilde{\phi}_i(X) - \phi_i(X)\,, && (3.105)\\ &\simeq \partial_\mu \phi_i(X)\, \delta X^{\mu} + \delta\phi_i(X)\,, && (3.106) \end{aligned}$$

dove abbiamo posto $\delta\phi_i(X) = \tilde{\phi}_i(X) - \phi_i(X)$, variazione in forma di ϕ_i e dove abbiamo sostituito $\tilde{\phi}_i$ con ϕ_i all'interno della derivazione fra (3.105) e (3.106), a meno di termini di ordine superiore al primo. Inoltre:

$$\begin{aligned} \Delta\mathcal{L} &= \tilde{\mathcal{L}}(\tilde{\phi}_i(X')...) - \mathcal{L}(\phi_i(X), ...)\,, && (3.107)\\ &= \delta\mathcal{L}(\phi_i(X)...) + \frac{\partial\mathcal{L}}{\partial\phi_i}\delta\phi_i + \frac{\partial\mathcal{L}}{\partial\phi_{i,\mu}}\delta\phi_{i,\mu} + \partial_\mu\mathcal{L}\delta X^{\mu}\,, && (3.108) \end{aligned}$$

dove $\delta\mathcal{L}$ è la variazione in forma della densità di lagrangiana ed il resto deriva dall'aver considerato $\phi_i(X)$, $\phi_{i,\mu}(X)$ e X^{μ} come variabili indipendenti in $\mathcal{L}$ e $\partial_\mu\mathcal{L}$ è la derivata totale[5] di $\mathcal{L}$ rispetto ad X^{μ} $(\partial_\mu\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi_i}\frac{\partial\phi_i}{\partial X^{\mu}} + ...)$.

Considerando le trasformazioni infinitesime, imporre la (3.103) equivale ad imporre:

$$0 = \delta S = \int_{\mathcal{V}'} d^4X'\, \tilde{\mathcal{L}} - \int_{\mathcal{V}} d^4X\, \mathcal{L}\,. \qquad (3.112)$$

Quindi, per poter procedere nel calcolo, dovremo riportare i due integrali allo stesso dominio d'integrazione. Trasformando $\int_{V'} d^4X'$ in $\int_V d^4X$ dovremo tener conto dello jacobiano della trasformazione

$$X^{\mu} \to X'^{\mu} = X^{\mu} + \delta X^{\mu}\,, \qquad (3.113)$$

ovvero di:

$$det(J) = \left|\frac{\partial X'^{\nu}}{\partial X^{\mu}}\right| = det\left(\delta^{\nu}_{\mu} + \partial_\mu \delta X^{\nu}\right) \simeq 1 + \partial_\mu \delta X^{\mu}\,, \qquad (3.114)$$

dove abbiamo usato la relazione $det(1+\epsilon) \simeq 1 + tr(\epsilon)$.

Sostituendo nell'Eq. (3.112) e sviluppando al primo ordine, si ottiene:

$$\begin{aligned} 0 &= \int_{\mathcal{V}} d^4X \left\{(1 + \partial_\mu\delta X^{\mu})\,\tilde{\mathcal{L}} - \mathcal{L}\right\}\,, && (3.115)\\ &\simeq \int_{\mathcal{V}} d^4X \left\{\delta\mathcal{L} + \frac{\partial\mathcal{L}}{\partial\phi_i}\delta\phi_i + \frac{\partial\mathcal{L}}{\partial\phi_{i,\mu}}\delta\phi_{i,\mu} + \partial_\mu\mathcal{L}\delta X^{\mu} + \partial_\mu\delta X^{\mu}\tilde{\mathcal{L}}\right\}\,, && (3.116)\\ &\simeq \int_{\mathcal{V}} d^4X \left\{\delta\mathcal{L} + \frac{\partial\mathcal{L}}{\partial\phi_i}\delta\phi_i + \partial_\mu\Big[\frac{\partial\mathcal{L}}{\partial\phi_{i,\mu}}\delta\phi_i\Big] - \Big[\partial_\mu\frac{\partial\mathcal{L}}{\partial\phi_{i,\mu}}\Big]\delta\phi_i + \partial_\mu\mathcal{L}\delta X^{\mu} + \partial_\mu\delta X^{\mu}\mathcal{L}\right\}, && (3.117) \end{aligned}$$

[5] We can rewrite the total difference as follows:

$$\begin{aligned} \Delta\mathcal{L} &= \tilde{\mathcal{L}}(\tilde{\phi}_i(X')...) - \mathcal{L}(\phi_i(X), ...)\,, && (3.109)\\ &= \tilde{\mathcal{L}}(\tilde{\phi}_i(X')...) - \mathcal{L}(\tilde{\phi}_i(X')...) + \mathcal{L}(\tilde{\phi}_i(X')...) - \mathcal{L}(\phi_i(X')...) + \mathcal{L}(\phi_i(X')...) - \mathcal{L}(\phi_i(X), ...)\,, && (3.110)\\ &\simeq \delta\mathcal{L} + \frac{\partial\mathcal{L}}{\partial\phi_i}\delta\phi_i + \frac{\partial\mathcal{L}}{\partial\phi_{i,\mu}}\delta\phi_{i,\mu} + \partial_\mu\mathcal{L}\delta X^{\mu}\,, && (3.111) \end{aligned}$$

where we identified the functional variation of the lagrangian density, $\delta\mathcal{L} \simeq \tilde{\mathcal{L}}(\tilde{\phi}_i(X')...) - \mathcal{L}(\tilde{\phi}_i(X')...)$, its derivative with respect to the variation in form of the fields, $\mathcal{L}(\tilde{\phi}_i(X')...) - \mathcal{L}(\phi_i(X')...) \simeq \frac{\partial\mathcal{L}}{\partial\phi_i}\delta\phi_i + \frac{\partial\mathcal{L}}{\partial\phi_{i,\mu}}\delta\phi_{i,\mu}$ and the total derivative with respect to X^{μ}, $\mathcal{L}(\phi_i(X')...) - \mathcal{L}(\phi_i(X), ...) \simeq \partial_\mu\mathcal{L}\delta X^{\mu}$.

$$= \int_{\mathcal{V}} d^4X \left\{ \delta\mathcal{L} + \partial_\mu \Big[\frac{\partial \mathcal{L}}{\partial \phi_{i,\mu}}\, \delta\phi_i + \mathcal{L}\,\delta X^\mu \Big] \right\} , \tag{3.118}$$

dove per passare da (3.116) a (3.117) abbiamo integrato per parti e sostituito, a meno di infinitesimi superiori al primo, $\tilde{\mathcal{L}}$ con $\mathcal{L}$, e per passare da (3.117) a (3.118) abbiamo sfruttato le equazioni del moto.

Per l'arbitrarietà del d^4X, la (3.118) dà la seguente equazione:

$$\partial_\mu \Big[\frac{\partial \mathcal{L}}{\partial \phi_{i,\mu}}\, \delta\phi_i + \mathcal{L}\,\delta X^\mu \Big] \;=\; -\delta\mathcal{L} . \tag{3.119}$$

Consideriamo il termine $\delta\mathcal{L}$.

Se la trasformazione è una simmetria, come abbiamo imposto, la variazione in forma della densità di lagrangiana non può essere qualunque. Infatti, dovendo rimanere invariate le equazioni di moto, $\delta\mathcal{L}$ potrà al massimo essere la quadridivergenza di una certa funzione $\delta\Omega^\mu$:

$$\delta\mathcal{L} \;=\; \partial_\mu\, \delta\Omega^\mu , \tag{3.120}$$

con $\delta\Omega^\mu$ che si annulla sulla frontiera del dominio d'integrazione.

L'Eq. (3.119) diventa, allora, semplicemente un'equazione di continuità:

$$\partial_\mu\, J^\mu \;=\; 0 , \tag{3.121}$$

dove abbiamo definito la seguente quadricorrente:

$$J^\mu \;=\; \left(\frac{\partial \mathcal{L}}{\partial \phi_{i,\mu}}\, \delta\phi_i + \mathcal{L}\,\delta X^\mu + \delta\Omega^\mu \right) . \tag{3.122}$$

Se i campi ϕ_i e la funzione arbitraria $\delta\Omega^\mu$ si annullano all'infinito, la conservazione della corrente J^μ, espressa dall'Eq. (3.121), porta alla conservazione della *carica*:

$$Q \;=\; \int_{\mathcal{V}} d^3X\, J^0 . \tag{3.123}$$

Infatti, si ha:

$$\frac{dQ}{dt} \;=\; \partial_0 \int_V d^3X\, J^0 \;=\; \int_{\partial V} d\Sigma\, \mathbf{J}\cdot\mathbf{n} \;=\; 0 , \tag{3.124}$$

che implica:

$$Q \;=\; cost . \tag{3.125}$$

È chiaro che, a seconda della trasformazione (o meglio a seconda di quanti parametri indipendenti contiene la trasformazione) (3.101), avremo più correnti conservate e quindi più cariche conservate. Il numero di queste dipende proprio dal numero di parametri indipendenti della trasformazione.

È da notare, inoltre, che se le simmetrie "di Nöether" formano un gruppo, l'algebra di questo gruppo induce sulle cariche conservate la stessa algebra. In altre parole le cariche sono i generatori del gruppo di trasformazioni considerato.

Andiamo, adesso, a vedere alcuni esempi.

3.5.1 Simmetrie geometriche. Trasformazioni di Lorentz

Consideriamo il caso in cui $\delta\Omega^\mu = 0$, cioè in cui la densità di lagrangiana viene lasciata invariata dalla trasformazione, e operiamo una trasformazione di Lorentz infinitesima:

$$X'^\mu = X^\mu + \epsilon^{\mu\nu} X_\nu , \tag{3.126}$$

dove il tensore del secondo ordine $\epsilon_{\mu\nu}$ è antisimmetrico. Infatti, siccome X^2 è un'invariante di Lorentz, si ha:

$$X^2 = X'^2 \tag{3.127}$$

e siccome per la trasformazione infinitesima $X' = X + \delta X$, elevando al quadrato si trova

$$X'^2 = (X + \delta X)^2 \simeq X^2 + X \cdot \delta X \tag{3.128}$$

che, per la (3.127), dà:

$$X \cdot \delta X = 0\,. \tag{3.129}$$

Ma siccome, ancora, $\delta X^\mu = \epsilon^{\mu\nu} X_\nu$, si ha infine:

$$X_\mu X_\nu \epsilon^{\mu\nu} = 0\,, \tag{3.130}$$

che è vera solo se $\epsilon^{\mu\nu}$ è antisimmetrico, essendo $X_\mu X_\nu$ simmetrico.

Questo vuol dire che $\epsilon^{\mu\nu}$ ha $6 = \frac{n(n-1)}{2}$ parametri indipendenti: 3 per le rotazioni e 3 per i boosts, lungo i tre assi coordinati.

Consideriamo l'indice "i" del campo ϕ_i come un indice di Lorentz, ovvero consideriamo il caso di un unico campo che si trasformi sotto la (3.126) secondo una certa rappresentazione del Gruppo di Lorentz. Allora si avrà:

$$\begin{aligned} \phi^i(X) \quad &\rightarrow \quad S(\Lambda)^i_j \phi^j(X) \simeq \left[1 - \frac{1}{2}\Sigma_{\nu\rho}\epsilon^{\nu\rho}\right]^i_j \phi^j(X) = \\ &= \quad \phi^i(X) - \frac{1}{2}\left(\Sigma_{\nu\rho}\epsilon^{\nu\rho}\right)^i_j \phi^j(X)\,. \end{aligned} \tag{3.131}$$

Le $\Sigma_{\nu\rho}$ sono i generatori del Gruppo di Lorentz, o meglio una loro rappresentazione nella base dei campi (rappresentazione tensoriale o spinoriale), mentre $\epsilon^{\mu\nu}$ rappresenta gli "angoli" di rotazione.

In totale, quindi:

$$\begin{cases} \delta X^\mu & = \quad \epsilon^{\mu\nu} X_\nu \\ \Delta\phi^i(X) & = \quad -\frac{1}{2}\left(\Sigma_{\nu\rho}\epsilon^{\nu\rho}\right)^i_j \phi^j(X) \end{cases}\,. \tag{3.132}$$

Siccome

$$J^\mu \;=\; \frac{\partial \mathcal{L}}{\partial \phi^i_{,\mu}}\delta\phi^i + \mathcal{L}\delta X^\mu \;=\; \frac{\partial \mathcal{L}}{\partial \phi^i_{,\mu}}\left[\Delta\phi^i - \partial_\mu\phi^i\delta X^\mu\right] + \mathcal{L}\delta X^\mu\,, \tag{3.133}$$

la quadricorrente conservata è data dalla seguente relazione:

$$J^\mu \;=\; \frac{\partial \mathcal{L}}{\partial \phi^i_{,\mu}}\left[-\frac{1}{2}\Big(\Sigma_{\nu\rho}\epsilon^{\nu\rho}\Big)^i_j \phi^j - \partial_\rho\phi^i\delta X^\rho\right] + \mathcal{L}\delta X^\mu \;= \tag{3.134}$$

$$=\; -\frac{1}{2}\frac{\partial \mathcal{L}}{\partial \phi^i_{,\mu}}\left(\Sigma_{\nu\rho}\epsilon^{\nu\rho}\right)^i_j \phi^j - \frac{\partial \mathcal{L}}{\partial \phi^i_{,\mu}}\partial_\nu\phi^i\epsilon^{\nu\rho}X_\rho + g^\mu_\nu\epsilon^{\nu\rho}X_\rho\mathcal{L} \;= \tag{3.135}$$

$$=\; -\frac{1}{2}\frac{\partial \mathcal{L}}{\partial \phi^i_{,\mu}}\left(\Sigma_{\nu\rho}\epsilon^{\nu\rho}\right)^i_j \phi^j - \epsilon^{\nu\rho}X_\rho\left[\frac{\partial \mathcal{L}}{\partial \phi^i_{,\mu}}\phi^i_{,\nu} - g^\mu_\nu\mathcal{L}\right] \;= \tag{3.136}$$

$$=\; -\frac{1}{2}\frac{\partial \mathcal{L}}{\partial \phi^i_{,\mu}}\left(\Sigma_{\nu\rho}\epsilon^{\nu\rho}\right)^i_j \phi^j - \epsilon^{\nu\rho}X_\rho T^\mu_\nu\,, \tag{3.137}$$

dove abbiamo posto:

$$T^\mu_\nu = \frac{\partial \mathcal{L}}{\partial \phi^i_{,\mu}}\phi^i_{,\nu} - g^\mu_\nu\mathcal{L}\,. \tag{3.138}$$

Siccome, inoltre, $\epsilon^{\mu\nu}$ è antisimmetrico nello scambio dei due indici, l'unico contributo non nullo di $\epsilon^{\nu\rho}X_\rho T^\mu_\nu$ deriva dalla parte antisimmetrica di $X_\rho T^\mu_\nu$ (in ν e ρ):

$$\frac{1}{2}\left(X_\rho T^\mu_\nu - X_\nu T^\mu_\rho\right)\,. \tag{3.139}$$

Per cui, infine, si ha:

$$J^\mu = \frac{1}{2}\epsilon^{\nu\rho}\left[-\frac{\partial\mathcal{L}}{\partial\phi^i_{,\mu}}\Big(\Sigma_{\nu\rho}\Big)^i_j\phi^j - \Big(X_\rho T^\mu_\nu - X_\nu T^\mu_\rho\Big)\right] = \tag{3.140}$$

$$= \frac{1}{2}\epsilon^{\nu\rho}\mathcal{M}^\mu_{\nu\rho}\,, \tag{3.141}$$

dove abbiamo definito il tensore:

$$\mathcal{M}^\mu_{\nu\rho} = \left(X_\nu T^\mu_\rho - X_\rho T^\mu_\nu\right) - \frac{\partial\mathcal{L}}{\partial\phi^i_{,\mu}}\left(\Sigma_{\nu\rho}\right)^i_j\phi^j\,, \tag{3.142}$$

che ha 24 componenti indipendenti (4 in μ e $6 = \frac{n(n-1)}{2}$ in $\rho\nu$).

Il tensore $\mathcal{M}^\mu_{\rho\nu}$ è una generalizzazione del momento angolare.

È formato da un momento "orbitale" $X_\rho T^\mu_\nu - X_\nu T^\mu_\rho$, che infatti ha la struttura di un prodotto vettoriale e da un momento "intrinseco" (momento di spin) $-\frac{\partial\mathcal{L}}{\partial\phi^i_{,\mu}}\left(\Sigma_{\rho\nu}\right)^i_j\phi^j$. Il primo momento angolare deriva dall'azione del Gruppo di Lorentz sulle coordinate spazio-temporali; lo spin dall'azione dello stesso sulle coordinate spinoriali del campo.

La conservazione della quadricorrente, $\partial_\mu J^\mu = 0$, essendo $\epsilon^{\rho\nu}$ una costante (sono gli angoli di rotazione e non dipendono da X), porta alla seguente equazione per il tensore $\mathcal{M}$:

$$\partial_\mu\mathcal{M}^\mu_{\rho\nu} = 0\,. \tag{3.143}$$

La (3.143) costituisce in realtà 6 correnti conservate, che sono le 6 componenti indipendenti in ρ e ν di $\mathcal{M}^\mu_{\rho\nu}$.

Posto

$$M_{\rho\nu} = \int d^3X\ \mathcal{M}^o_{\rho\nu}\,, \tag{3.144}$$

se i campi vanno a zero all'infinito, si ha la conservazione delle 6 cariche:

$$\dot{M}_{\rho\nu} = 0\,. \tag{3.145}$$

3.5.2 Campo scalare e conservazione del quadriimpulso e del momento angolare orbitale

Se ci riduciamo al caso particolare di un campo scalare, avremo

$$\Delta\phi(X) = \phi'(X') - \phi(X) = 0\,. \tag{3.146}$$

Consideriamo prima di tutto una traslazione spazio-temporale di un quadrivettore a^μ costante:

$$\begin{cases}\delta X^\mu = a^\mu\\ \Delta\phi = 0\end{cases} \tag{3.147}$$

cosicché si abbia:

$$\delta\phi(X) = -\partial_\mu\phi(X)\,\delta X^\mu = -\partial_\mu\phi(X)\,a^\mu\,. \tag{3.148}$$

Allora, si può ricavare facilmente la conservazione del quadriimpulso. Infatti, si ha:

$$J^\mu = \left[\mathcal{L}g^\mu_\nu - \frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\partial_\nu\right]a^\nu = \tag{3.149}$$

$$= -T^\mu_\nu\,a^\nu\,, \tag{3.150}$$

dove T^μ_ν è il tensore energia-impulso del sistema.

Siccome la traslazione a^μ è costante, la legge di conservazione della corrente J^μ implica:

$$\partial_\mu T^\mu_\nu \,=\, 0\,, \tag{3.151}$$

che sono quattro leggi di conservazione locale.

Definiamo il quadriimpulso del sistema come segue:

$$P_\nu \,=\, \int d^3X\, T^0_\nu\,. \tag{3.152}$$

Allora la (3.151) porta alla

$$\dot{P}_\nu \,=\, 0\,. \tag{3.153}$$

Infatti, le

$$\begin{cases} \partial_\mu T^\mu_0 & = & 0 \\ \partial_\mu T^\mu_1 & = & 0 \\ \partial_\mu T^\mu_2 & = & 0 \\ \partial_\mu T^\mu_3 & = & 0 \end{cases} \tag{3.154}$$

implicano

$$\begin{cases} \partial_0 T^0_0 & = & \partial_i T^i_0 \\ \partial_0 T^0_1 & = & \partial_i T^i_1 \\ \partial_0 T^0_2 & = & \partial_i T^i_2 \\ \partial_0 T^0_3 & = & \partial_i T^i_3 \end{cases} \tag{3.155}$$

e integrando in d^3X, supposto che i campi vadano a zero all'infinito, si ottiene la (3.153) componente per componente:

$$\begin{cases} \partial_0 P_0 & = & \partial_i \int d^3X\, T^i_0 & \to 0 \\ . & = & . & . \\ . & = & . & . \\ \partial_0 P_3 & = & \partial_i \int d^3X\, T^i_3 & \to 0 \end{cases} \tag{3.156}$$

Se invece delle traslazioni consideriamo le trasformazioni proprie di Lorentz, avremo:

$$J^\mu \,=\, \frac{1}{2}\epsilon^{\rho\nu}\left[X_\rho T^\mu_\nu - X_\nu T^\mu_\rho\right] \,=\, \frac{1}{2}\epsilon^{\rho\nu}\,\mathrm{M}^\mu_{\rho\nu}\,. \tag{3.157}$$

La conservazione della corrente J^μ implica:

$$\partial_\mu \mathrm{M}^\mu_{\rho\nu} \,=\, 0\,, \tag{3.158}$$

ovvero:

$$\partial_0 \mathrm{M}^0_{\rho\nu} \,=\, \partial_i \mathrm{M}^i_{\rho\nu}\,. \tag{3.159}$$

Consideriamo le componenti M^0_{ij}. Si ha:

$$\mathrm{M}^0_{ij} \,=\, \left[X_i T^0_j - X_j T^0_i\right] \,=\, \left[X_i \mathcal{P}_j - X_j \mathcal{P}_i\right]\,, \tag{3.160}$$

dove abbiamo introdotto $\mathcal{P}_i$ densità spaziale d'impulso. Allora:

$$\mathrm{M}^0_{ij} \,=\, \epsilon_{ijk}\, L_k \,=\, \begin{pmatrix} 0 & L_3 & -L_2 \\ -L_3 & 0 & L_1 \\ L_2 & -L_1 & 0 \end{pmatrix} \tag{3.161}$$

dove $\mathbf{L} = \mathbf{r} \wedge \mathcal{P}$ è la densità spaziale di momento angolare. Integrando la (3.159) in d^3X si ottiene la conservazione del momento angolare orbitale:

$$\dot{\mathrm{L}} \,=\, 0\,, \tag{3.162}$$

dove

$$\mathrm{L}_i \,=\, \int d^3X\, L_i\,. \tag{3.163}$$

3.5.3 Simmetrie interne globali

Come abbiamo già accennato, l'altro esempio di trasformazione (3.101) da considerare è quello di una variazione che coinvolga soltanto una ridefinizione in forma dei campi, ma non un cambiamento di sistema di riferimento.

Genericamente avremo:

$$\begin{cases} \delta X^\mu & = & 0 \\ \Delta\phi_i & = & \delta\phi_i \neq 0 \end{cases} \tag{3.164}$$

da cui scaturisce la legge di conservazione locale $\partial_\mu J^\mu = 0$ con:

$$J^\mu = \frac{\partial\mathcal{L}}{\partial\phi^i_{,\mu}}\,\delta\phi^i\,. \tag{3.165}$$

Se i campi vanno a zero all'infinito, si conserva la carica:

$$Q = \int d^3X\,J^0 = \int d^3X\,\frac{\partial\mathcal{L}}{\partial\dot{\phi}^i}\delta\phi^i\,. \tag{3.166}$$

Campo scalare carico

Il tipico esempio di simmetria interna è l'invarianza della lagrangiana del campo scalare carico,

$$\mathcal{L} = \partial_\mu\phi^\dagger\partial^\mu\phi - m^2\phi^\dagger\phi\,, \tag{3.167}$$

sotto trasformazioni di fase globali:

$$\begin{cases} \phi & \to & \phi' = e^{i\alpha}\phi \\ \phi^\dagger & \to & \phi'^\dagger = \phi^\dagger e^{-i\alpha} \end{cases}\,. \tag{3.168}$$

Quest'invarianza determina la conservazione della corrente:

$$j^\mu = \frac{J^\mu}{\alpha} = i\left[\left(\partial_\mu\phi^\dagger\right)\phi - \left(\partial_\mu\phi\right)\phi^\dagger\right] \tag{3.169}$$

e della carica:

$$Q = i\int d^3X\,\left(\dot{\phi}^\dagger\phi - \dot{\phi}\phi^\dagger\right)\,, \tag{3.170}$$

che può essere vista nel modello interagente come carica elettrica delle particelle e antiparticelle scalari ϕ.

Campo di Dirac

Le trasformazioni di fase globali lasciano invariata anche un'altra lagrangiana: quella del campo di Dirac libero:

$$\mathcal{L} = \overline{\psi}\,(i\,\partial\!\!\!/ - m)\,\psi\,. \tag{3.171}$$

Riscriviamo le (3.168) per il campo ψ:

$$\begin{cases} \psi & \to & \psi' = e^{-i\alpha}\psi \\ \overline{\psi} & \to & \overline{\psi}' = \overline{\psi}\,e^{i\alpha} \end{cases}\,. \tag{3.172}$$

Allora, avremo una corrente conservata:

$$j^\mu = \frac{J^\mu}{\alpha} = \overline{\psi}\,\gamma^\mu\,\psi \tag{3.173}$$

ed una carica conservata:

$$Q = \int d^3X\,\left(\overline{\psi}\,\gamma^0\,\psi\right) = \int d^3X\,\psi^\dagger\psi\,. \tag{3.174}$$

Capitolo 4

Free Fields

In this chapter we will study the non-interacting fields, from the classical viewpoint to their canonical quantization.

4.1 The Klein-Gordon Field (classical field)

We introduced different finite-dimensional representation of the Lorentz group. According to them, we can classify our fields. We start with the simplest representation, the trivial one[1], and we consider then a scalar field (real or complex) $\phi(X)$ which under a Poincaré transformation $X^\mu \to X'^\mu = \Lambda^\mu_\nu X^\nu + a^\mu$ transforms as

$$\phi(X) \to \phi'(X') = \phi(X)\,. \tag{4.1}$$

4.1.1 The Klein-Gordon equation

The Klein-Gordon field will satisfy a differential equation that can be found using the relativistic dispertion relation (energy-momentum relation or mass-shell condition)

$$E^2 = p^2 + m^2\,, \tag{4.2}$$

replacing the energy and the momentum with the correspondence principle

$$\begin{cases} E & \to & i\frac{\partial}{\partial t}\,, \\ p & \to & -i\nabla\,. \end{cases} \tag{4.3}$$

We find

$$-\frac{\partial^2}{\partial t^2}\phi(X) = (-\nabla^2 + m^2)\phi(X)\,, \tag{4.4}$$

that, remembering the covariant form $\partial_\mu\partial^\mu = \partial_0^2 - \partial_i^2$, can be written in manifestly covariant way as follows

$$(\partial_\mu\partial^\mu + m^2)\phi(X) \;=\; 0\,. \tag{4.5}$$

Eq. (4.5) is invariant under Poincaré transformations. In fact we can check easily that

$$(\partial'_\mu\partial'^\mu + m^2)\phi'(X') = (\Lambda_\mu^{\ \cdot\nu}\Lambda^\mu_\rho\,\partial_\nu\partial^\rho + m^2)\phi'(X') = (\delta^\nu_\rho\partial_\nu\partial^\rho + m^2)\phi'(X') = (\partial_\nu\partial^\nu + m^2)\phi(X) = 0\,. \tag{4.6}$$

Eq. (4.5) has to be considered as a classical equation for the classical field $\phi(X)$. Then we will quantize our system. In this sense there is no "second quantization", but only the quantization of the classical field (we will quantize once!).

[1] In this representation the generators of the group are zero, $J^{\mu\nu} = 0$.

If, as was the case when the equation was proposed around 1926, we would like to interpret Eq. (4.5) as a wave equation (so to say "à la Schrödinger"), we would face many issues. The main ones can be summarized as follows:

- First of all, the fact that we have a differential equation which is second-order in time seams to be in constrast with the basic laws of quantum mechanics according to which we can determine the time evolution of the wave function knowing just the function at a certain time t_0. In order to solve a second-order differential equation, we must, instead, provide the initial values of the field and its time derivative.

- The fact that we have a differential equation which is second-order in time makes in such a way that the probabilistic interpretation of the theory is at risk. What we would like to interpret as "probability density" is in fact non positive definite. We can see that considering the differential equation for the complex-conjugated field (which is the same as Eq. (4.5) since the differential operator is real):
$$(\partial_\mu\partial^\mu + m^2)\phi^*(X) \,=\, 0\,. \tag{4.7}$$
If we multiply Eq. (4.5) by ϕ^* and we subtract Eq. (4.7) multiplied by ϕ, we have
$$\begin{aligned} 0 \;&=\; \phi^*(\partial_\mu\partial^\mu + m^2)\phi - \phi(\partial_\mu\partial^\mu + m^2)\phi^* \;= \\ &=\; \phi^*\frac{\partial^2}{\partial t^2}\phi - \phi\frac{\partial^2}{\partial t^2}\phi^* + \phi\nabla^2\phi^* - \phi^*\nabla^2\phi \;= \\ &=\; \frac{\partial}{\partial t}(\phi^*\frac{\partial}{\partial t}\phi - \phi\frac{\partial}{\partial t}\phi^*) + \nabla\cdot(\phi\nabla\phi^* - \phi^*\nabla\phi)\,, \end{aligned} \tag{4.8}$$
which is a continuity equation in which the probability density should be given by the following expression[2]:
$$\rho = i\left(\phi^*\frac{\partial}{\partial t}\phi - \phi\frac{\partial}{\partial t}\phi^*\right) = i\phi^*\overleftrightarrow{\partial_0}\,\phi\,, \tag{4.9}$$
such that
$$\frac{d}{dt}\int d^3X\, i\phi^*\overleftrightarrow{\partial_0}\,\phi = 0\,. \tag{4.10}$$
Eq. (4.10) defines the correct scalar product (in the Hilbert space of $\phi(X)$), which is conserved (time independent):
$$(\phi_1, \phi_2) = \int d^3X\, i\phi_1^*\overleftrightarrow{\partial_0}\,\phi_2\,. \tag{4.11}$$
However, ρ it is not a positive definite expression and, therefore, the connection with the probabilistic interpretation of the theory fails.

- Finally, an even more serious problem arises from the plane wave solutions of the Klein-Gordon equation. As we will see in the next section.

4.1.2 Plane wave solutions of the Klein-Gordon equation

We look for a solution of Eq. (4.5) as a plane wave solution
$$\phi(X) \,=\, A\,e^{-iP_\mu X^\mu}\,. \tag{4.12}$$
Substituting Eq. (4.12) in Eq. (4.5) we find that the plane wave is a solution prodived that
$$(\partial_\mu\partial^\mu + m^2)Ae^{-iP_\mu X^\mu} = (-P_\mu P^\mu + m^2)Ae^{-iP_\mu X^\mu} = 0\,, \tag{4.13}$$

[2]We put an "i" in order to have a real ρ

i.e.

$$P_\mu P^\mu = E^2 - p^2 = m^2\,. \tag{4.14}$$

Eq. (4.14) gives two possible solutions for the energy E:

$$E_+ = \sqrt{p^2+m^2} = \omega_p\,, \tag{4.15}$$
$$E_- = -\sqrt{p^2+m^2} = -\omega_p\,. \tag{4.16}$$

We therefore have two different solutions:

$$f_p^+(X) = A\,e^{-iE_+t+i\mathbf{p}\cdot\mathbf{x}} = A\,e^{-i\omega_p t+i\mathbf{p}\cdot\mathbf{x}}\,, \tag{4.17}$$
$$f_p^-(X) = A\,e^{-iE_-t+i\mathbf{p}\cdot\mathbf{x}} = A\,e^{i\omega_p t+i\mathbf{p}\cdot\mathbf{x}}\,. \tag{4.18}$$

Solution (4.18) is of difficult interpretation within a theory such as wave mechanics, "à la Schrödinger". The issue can be solved moving to a field theory. Eq. (4.5) should be interpreted not as a wave equation, but as the differential equation that the classical field ϕ has to fulfill.

The general solution will be a superposition of f^+ and f^-:

$$\phi(X) = \int d^3p \left(\alpha(p)\,A\,e^{-i\omega_p t+i\mathbf{p}\cdot\mathbf{x}} + \beta(p)\,A\,e^{i\omega_p t+i\mathbf{p}\cdot\mathbf{x}}\right)\,. \tag{4.19}$$

Let us normalize our functions with respect to the scalar product (4.11). We have

$$(f_p^+(X), f_{p'}^+(X)) = i|A|^2 \int d^3X \left[e^{iP_\mu X^\mu}\partial_0 e^{-iP'_\mu X^\mu} - e^{-iP'_\mu X^\mu}\partial_0 e^{iP_\mu X^\mu}\right]\,, \tag{4.20}$$
$$= i|A|^2 \int d^3X \left[-i\omega_{p'} e^{i(P-P')_\mu X^\mu} - i\omega_p e^{-i(P-P')_\mu X^\mu}\right]\,, \tag{4.21}$$
$$= |A|^2 \int d^3X (\omega_{p'}+\omega_p) e^{i(P-P')_\mu X^\mu}\,, \tag{4.22}$$
$$= |A|^2(\omega_{p'}+\omega_p) e^{i(\omega_p-\omega_{p'})t} \int d^3X e^{-i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}}\,, \tag{4.23}$$
$$= (2\pi)^3|A|^2(\omega_{p'}+\omega_p) e^{i(\omega_p-\omega_{p'})t}\delta(\mathbf{p}-\mathbf{p}')\,, \tag{4.24}$$
$$= (2\pi)^3|A|^2 2\omega_p \delta(\mathbf{p}-\mathbf{p}')\,. \tag{4.25}$$

Imposing

$$(f_p^+(X), f_{p'}^+(X)) = \delta(\mathbf{p}-\mathbf{p}')\,, \tag{4.26}$$

we find[3]

$$A = \frac{1}{(2\pi)^{\frac{3}{2}}\sqrt{2\omega_p}}\,. \tag{4.27}$$

Finally

$$f_p^+(X) = \frac{e^{-iP_\mu X^\mu}}{(2\pi)^{\frac{3}{2}}\sqrt{2\omega_p}}\,. \tag{4.28}$$

In the same way we can normalize $f_p^-(X)$, that has negative norm (a remark of the fact that negative energy solutions cannot be linked to usual wave mechanics solutions). We have

$$(f_p^-(X), f_{p'}^-(X)) = i|A|^2 \int d^3X \left[e^{-i\omega_p t-i\mathbf{p}\cdot\mathbf{x}}\partial_0 e^{i\omega_{p'}t-i\mathbf{p}'\cdot\mathbf{x}} - e^{i\omega_{p'}t-i\mathbf{p}'\cdot\mathbf{x}}\partial_0 e^{-i\omega_p t-i\mathbf{p}\cdot\mathbf{x}}\right]\,, \tag{4.29}$$
$$= i|A|^2 \int d^3X \left[i\omega_{p'} e^{-i(\omega_p-\omega_{p'})t-i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}} + i\omega_p e^{i(\omega_p-\omega_{p'})t-i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}}\right]\,, \tag{4.30}$$

[3] We choose A real.

$$= \quad -|A|^2(\omega_{p'}+\omega_p)e^{-i(\omega_p-\omega_{p'})t}\int d^3X e^{-i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}}\,, \tag{4.31}$$

$$= \quad -(2\pi)^3|A|^2 2\omega_p\delta(\mathbf{p}-\mathbf{p}')\,. \tag{4.32}$$

Imposing

$$(f_p^-(X), f_{p'}^-(X)) = -\delta(\mathbf{p}-\mathbf{p}')\,, \tag{4.33}$$

we find the same expression for the normalization factor:

$$A = \frac{1}{(2\pi)^{\frac{3}{2}}\sqrt{2\omega_p}}\,. \tag{4.34}$$

Therefore

$$f_p^-(X) = \frac{e^{i\omega_p t + i\mathbf{p}\cdot\mathbf{x}}}{(2\pi)^{\frac{3}{2}}\sqrt{2\omega_p}}\,. \tag{4.35}$$

We can prove that f^+ and f^- are ortogonal, as follows

$$(f_p^-(X), f_{p'}^+(X)) \quad = \quad i|A|^2\int d^3X\left[e^{-i\omega_p t - i\mathbf{p}\cdot\mathbf{x}}\partial_0 e^{-i\omega_{p'}t + i\mathbf{p}'\cdot\mathbf{x}} - e^{-i\omega_{p'}t + i\mathbf{p}'\cdot\mathbf{x}}\partial_0 e^{-i\omega_p t - i\mathbf{p}\cdot\mathbf{x}}\right], \tag{4.36}$$

$$= \quad i|A|^2\int d^3X\left[-i\omega_{p'}e^{-i(\omega_p+\omega_{p'})t - i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}} + i\omega_p e^{-i(\omega_p+\omega_{p'})t - i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}}\right], \tag{4.37}$$

$$= \quad |A|^2(\omega_{p'}-\omega_p)e^{-i(\omega_p+\omega_{p'})t}\int d^3X e^{-i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}}\,, \tag{4.38}$$

$$= \quad 0\,. \tag{4.39}$$

We express the classical solution of the KG equation in terms of plane waves as the following combination:

$$\phi(X) \quad = \quad \int d^3p\left(\alpha(p)\, f_p^+(X) + \beta(p)\, f_p^-(X)\right)\,, \tag{4.40}$$

$$= \quad \int \frac{d^3p}{(2\pi)^{\frac{3}{2}}\sqrt{2\omega_p}}\left(\alpha(p)\, e^{-i\omega_p t + i\mathbf{p}\cdot\mathbf{x}} + \beta(p)\, e^{i\omega_p t + i\mathbf{p}\cdot\mathbf{x}}\right)\,. \tag{4.41}$$

Since we are integrating in the whole domain of $\mathbf{p}$, we can change $\mathbf{p}$ with $-\mathbf{p}$ in the second integral finding

$$\phi(X) \quad = \quad \int \frac{d^3p}{(2\pi)^{\frac{3}{2}}\sqrt{2\omega_p}}\left(\alpha(p)\, e^{-iP_\mu X^\mu} + \tilde{\beta}(p)\, e^{iP_\mu X^\mu}\right)\,, \tag{4.42}$$

$$= \quad \int d^3p\left(\alpha(p)\, f_p^+(X) + \tilde{\beta}(p)\, f_p^-(X)\right)\,, \tag{4.43}$$

where now

$$f_p^-(X) = (f_p^+(X))^* = \frac{e^{iP_\mu X^\mu}}{(2\pi)^{\frac{3}{2}}\sqrt{2\omega_p}}\,. \tag{4.44}$$

If we consider a real field, then we have to impose $\phi^*(X) = \phi(X)$:

$$\phi^*(X) = \int d^3p\left(\alpha^*(p)\,(f_p^+(X))^* + \tilde{\beta}^*(p)\, f_p^+(X)\right) = \phi(X)\,, \tag{4.45}$$

that means $\tilde{\beta}(p) = \alpha^*(p)$.

The final expression for the classical real Klein-Gordon field (in terms of normal modes) is the following

$$\phi(X) = \int \frac{d^3p}{(2\pi)^{\frac{3}{2}}\sqrt{2\omega_p}} \left(\alpha(p)\, e^{-iP_\mu X^\mu} + \alpha^*(p)\, e^{iP_\mu X^\mu}\right) . \tag{4.46}$$

NOTE: The expression (4.44) for f^- has all the characteristics of the previous one, i.e. negative norm and ortogonality with f^+. The fact that negative-energy solutions are related to the positive-energy ones by the transformation $(E, \mathbf{p}) \to (-E, -\mathbf{p})$ has a nice meaning in terms of the Feynman-Stuckelberg interpretation.

We can use the scalar product to extract the coefficients $\alpha(p)$ and $\alpha^*(p)$:

$$\alpha(p) \;=\; (f_p^+, \phi) = i\int d^3X\, (f^+)^* \overleftrightarrow{\partial_0}\, \phi\,, \tag{4.47}$$

$$\alpha^*(p) \;=\; -(f_p^-, \phi) = -i\int d^3X\, (f^-)^* \overleftrightarrow{\partial_0}\, \phi\,. \tag{4.48}$$

4.1.3 Lagrangian density of the Klein-Gordon real field

We now want to find the Lagrangian density of the Klein-Gordon real field, i.e. the functional $\mathcal{L}$ such that through the Euler-Lagrange equations we can obtain Eq. (4.5). In order to do that, we use the Hamilton's principle following the reverse procedure. If $\delta\phi$ is the variation of the field $\phi(X)$, that vanishes on the boundary of the integration domain, we multiply Eq.(4.5) by that variation and we integrate by parts. We have

$$0 \;=\; \int d^4X\, \left(\partial_\mu\partial^\mu\phi + m^2\phi\right)\, \delta\phi \;= \tag{4.49}$$

$$=\; \int d^4X\, \left[\partial_\mu\left(\partial^\mu\phi\, \delta\phi\right) - \partial^\mu\phi\, \partial_\mu\delta\phi + \frac{m^2}{2}\delta\phi^2\right] \;= \tag{4.50}$$

$$=\; -\int d^4X\, \left[\partial^\mu\phi\, \delta\left(\partial_\mu\phi\right) - \frac{m^2}{2}\delta\phi^2\right] \;= \tag{4.51}$$

$$=\; -\delta\left[\int d^4X\, \frac{1}{2}\left(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2\right)\right], \tag{4.52}$$

where moving from (4.49) to (4.50) and from (4.51) to (4.52) we integrated by parts and where the first term of Eq. (4.50) gives a vanishing integral, since it is the total derivative and the variation of the field annihilates on the boundary of integration.

The lagrangian density is:

$$\mathcal{L} \;=\; \frac{1}{2}\left(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2\right) . \tag{4.53}$$

We note that an overall sugn in the definition of $\mathcal{L}$ does not affect the equations of motion. However, on the sign of $\mathcal{L}$ depends the sign of the hamiltonian density $\mathcal{H}$. We chose the sign of (4.53) in such a way to have an hamiltonian density definite positive.

Conserved quantities

From the invariance under Poincaré transformations of the lagrangian (4.53), we have a relation for the Nöther's charges. Translation invariance gives the conservation of the energy-momentum tensor

$$T^{\mu\nu} = \frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\phi^{,\nu} - \eta^{\mu\nu}\mathcal{L} = \partial^\mu\phi\partial^\nu\phi - \eta^{\mu\nu}\mathcal{L}\,, \tag{4.54}$$

such that $\partial_\mu T^{\mu\nu} = 0$. The four conserved charges are the energy density

$$\mathcal{H} = T^{00} = (\partial_0\phi)^2 - \mathcal{L} = \frac{1}{2}\left[\dot{\phi}^2 + (\nabla\phi)^2 + m^2\phi^2\right] \tag{4.55}$$

and the momentum density

$$P_i = T_i^0 = \dot{\phi}\partial_i\phi\,. \tag{4.56}$$

From Lorentz invariance, instead, we find the conservation of the 6 charges

$$M^0_{\mu\nu} = \int d^3X\,\left(X_\mu T^0_\nu - X_\nu T^0_\mu\right)\,, \tag{4.57}$$

among which for instance the orbital angular momentum

$$M^0_{ij} = \int d^3X\,\left(X_i T^0_j - X_j T^0_i\right) = \int d^3X\,\left(X_i\partial_0\phi\partial_j\phi - X_j\partial_0\phi\partial_i\phi\right)\,, \tag{4.58}$$

$$= \int d^3X\,\partial_0\phi\left(X_i\partial_j - X_j\partial_i\right)\phi = -i\int d^3X\,\partial_0\phi L_{ij}\phi\,, \tag{4.59}$$

where

$$L_{ij} = i\left(X_i\partial_j - X_j\partial_i\right) \tag{4.60}$$

is the angular momentum operator

4.1.4 Hamiltonian

The hamiltonian density is in Eq. (4.55)[4]. It can be obtained also with a Legendre transformation of the lagrangian density. We define the conjugated momentum to the field

$$\pi(X) = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \dot{\phi}\,. \tag{4.61}$$

Then we find

$$\mathcal{H} = \frac{1}{2}\left[\pi^2 + (\nabla\phi)^2 + m^2\phi^2\right]\,. \tag{4.62}$$

Recalling that

$$\phi(X) = \int\frac{d^3p}{(2\pi)^{\frac{3}{2}}\sqrt{2\omega_p}}\left(\alpha(p)\,e^{-iP_\mu X^\mu} + \alpha^*(p)\,e^{iP_\mu X^\mu}\right)\,, \tag{4.63}$$

$$\pi(X) = -i\int\frac{d^3p}{(2\pi)^{\frac{3}{2}}\sqrt{2\omega_p}}\omega_p\left(\alpha(p)\,e^{-iP_\mu X^\mu} - \alpha^*(p)\,e^{iP_\mu X^\mu}\right)\,, \tag{4.64}$$

$$\nabla\phi(X) = i\int\frac{d^3p}{(2\pi)^{\frac{3}{2}}\sqrt{2\omega_p}}\mathbf{p}\left(\alpha(p)\,e^{-iP_\mu X^\mu} - \alpha^*(p)\,e^{iP_\mu X^\mu}\right)\,, \tag{4.65}$$

we can find the hamiltonian as follows

$$H = \int d^3X\,\mathcal{H} = \int d^3X\,\frac{1}{2}\left[\pi^2 + (\nabla\phi)^2 + m^2\phi^2\right]\,, \tag{4.66}$$

$$= \frac{1}{2}\int\frac{d^3p\,d^3p'}{(2\pi)^3\sqrt{4\omega_p\omega_{p'}}}\int d^3X\left\{\right.$$

[4]Actuaslly, since in Eq. (4.55) there is the energy density, it is written in terms of ϕ and $\dot{\phi}$. If we want the hamiltonian we have to intend Eq. (4.61) as an equation from which to extract $\dot{\phi}$ in terms of π.

$$
\begin{aligned}
&-\omega_p\omega_{p'}\left[\alpha(p)\,e^{-iP_\mu X^\mu}-\alpha^*(p)\,e^{iP_\mu X^\mu}\right]\left[\alpha(p')\,e^{-iP'_\mu X^\mu}-\alpha^*(p')\,e^{iP'_\mu X^\mu}\right]\\
&-\mathbf{p}\cdot\mathbf{p}'\left[\alpha(p)\,e^{-iP_\mu X^\mu}-\alpha^*(p)\,e^{iP_\mu X^\mu}\right]\left[\alpha(p')\,e^{-iP'_\mu X^\mu}-\alpha^*(p')\,e^{iP'_\mu X^\mu}\right]\\
&+m^2\left[\alpha(p)\,e^{-iP_\mu X^\mu}+\alpha^*(p)\,e^{iP_\mu X^\mu}\right]\left[\alpha(p')\,e^{-iP'_\mu X^\mu}+\alpha^*(p')\,e^{iP'_\mu X^\mu}\right]\Bigg\}, \qquad (4.67)
\end{aligned}
$$

$$
\begin{aligned}
= \;& \frac{1}{2}\int\frac{d^3p\,d^3p'}{(2\pi)^3\sqrt{4\omega_p\omega_{p'}}}\int d^3X\,\Bigg\{\\
&\alpha(p)\alpha(p')e^{-i(P+P')_\mu X^\mu}(-\omega_p\omega_{p'}-\mathbf{p}\cdot\mathbf{p}'+m^2)\\
&+\alpha^*(p)\alpha^*(p')e^{i(P+P')_\mu X^\mu}(-\omega_p\omega_{p'}-\mathbf{p}\cdot\mathbf{p}'+m^2)\\
&+\alpha(p)\alpha^*(p')e^{-i(P-P')_\mu X^\mu}(\omega_p\omega_{p'}+\mathbf{p}\cdot\mathbf{p}'+m^2)\\
&+\alpha^*(p)\alpha(p')e^{i(P-P')_\mu X^\mu}(\omega_p\omega_{p'}+\mathbf{p}\cdot\mathbf{p}'+m^2)\Bigg\}, \qquad (4.68)
\end{aligned}
$$

$$
\begin{aligned}
= \;& |\,\text{integrating in } d^3X\,|\\
= \;& \frac{1}{2}\int\frac{d^3p\,d^3p'}{\sqrt{4\omega_p\omega_{p'}}}\Bigg\{\\
&\left[\alpha(p)\alpha(p')e^{-i(\omega_p+\omega_{p'})t}+\alpha^*(p)\alpha^*(p')e^{i(\omega_p+\omega_{p'})t}\right](-\omega_p\omega_{p'}-\mathbf{p}\cdot\mathbf{p}'+m^2)\,\delta(\mathbf{p}+\mathbf{p}')\\
&+\left[\alpha(p)\alpha^*(p')e^{-i(\omega_p-\omega_{p'})t}+\alpha^*(p)\alpha(p')e^{i(\omega_p-\omega_{p'})t}\right](\omega_p\omega_{p'}+\mathbf{p}\cdot\mathbf{p}'+m^2)\,\delta(\mathbf{p}-\mathbf{p}')\Bigg\} \qquad (4.69)
\end{aligned}
$$

$$
\begin{aligned}
= \;& |\,\text{integrating in } d^3p',\text{ since } -\omega_p^2+p^2+m^2=0\,|\\
= \;& \int d^3p\,\frac{1}{2}\omega_p\left[\alpha(p)\alpha^*(p)+\alpha^*(p)\alpha(p)\right], \qquad (4.70)
\end{aligned}
$$

where we considered $\alpha^*(p)\alpha(p)\neq\alpha(p)\alpha^*(p)$, although classically this does not have any meaning.

As in the case of the vibrating string, we find that the hamiltonian is the "sum" of an infinite number of hamiltonians of harmonic oscillator of frequency ω_p.

4.1.5 Complex scalar field and the charge

So far we treated the real field case. Let us consider now the possibility to have a complex scalar field. We can treat the case, starting with two real fields, ϕ_1 and ϕ_2 degenerate in mass, for which the lagrangian density is

$$
\mathcal{L}=\frac{1}{2}\left[\partial_\mu\phi_1\partial^\mu\phi_1+\partial_\mu\phi_2\partial^\mu\phi_2-m^2(\phi_1^2+\phi_2^2)\right], \qquad (4.71)
$$

rotating in the complex plane to ϕ and ϕ^* defined such that

$$
\phi = \frac{\phi_1+i\phi_2}{\sqrt{2}}, \qquad (4.72)
$$

$$
\phi^* = \frac{\phi_1-i\phi_2}{\sqrt{2}}, \qquad (4.73)
$$

or

$$
\phi_1 = \frac{\phi+\phi^*}{\sqrt{2}}, \qquad (4.74)
$$

$$
\phi_2 = \frac{\phi-\phi^*}{i\sqrt{2}}, \qquad (4.75)
$$

Substituting (4.74,4.75) in Eq. (4.71) we find

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi\,. \tag{4.76}$$

The lagrangian (4.76) has a global internal symmetry, under $U(1)$ phase transformations:

$$\begin{aligned} \phi(X) &\rightarrow \phi'(X) = e^{-i\theta}\phi(X)\,, && (4.77)\\ \phi^*(X) &\rightarrow \phi'^*(X) = e^{i\theta}\phi^*(X)\,, && (4.78) \end{aligned}$$

with $\theta \in \mathbb{R}$. It is easy to check that indeed

$$\mathcal{L}' = \partial_\mu \phi'^* \partial^\mu \phi' - m^2 \phi'^* \phi' = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi = \mathcal{L}\,. \tag{4.79}$$

The Nöther's current associated to this symmetry is

$$J^\mu = \frac{\partial \mathcal{L}}{\partial \phi_{i,\mu}} \delta\phi_i = \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \delta\phi + \frac{\partial \mathcal{L}}{\partial \phi^*_{,\mu}} \delta\phi^*\,. \tag{4.80}$$

Note that the piece proportional to the lagrangian density is not present, since for this internal symmetry we do not have any change in the space-time point, $\delta X^\mu = 0$.

The infinitesimal transformation can be found expanding (4.77,4.78):

$$\begin{aligned} \delta\phi &= -i\theta\phi\,, && (4.81)\\ \delta\phi^* &= i\theta\phi^*\,, && (4.82) \end{aligned}$$

from which

$$\partial_\mu \left[\frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \delta\phi + \frac{\partial \mathcal{L}}{\partial \phi^*_{,\mu}} \delta\phi^* \right] = \partial_\mu \left[-i\theta\phi\partial^\mu\phi^* + i\theta\phi^*\partial^\mu\phi \right] = \theta\, \partial_\mu \left[i\phi^* \overleftrightarrow{\partial^\mu} \phi \right] = 0\,. \tag{4.83}$$

Since relation (4.83) holds for any θ, we define the current as

$$J^\mu = i\phi^* \overleftrightarrow{\partial^\mu} \phi \tag{4.84}$$

and the conserved current is

$$Q = \int d^3X i\phi^* \overleftrightarrow{\partial^0} \phi\,. \tag{4.85}$$

4.1.6 Non relativistic limit

In the $\beta \ll 1$ limit, the energy can be expanded as well finding

$$E \sim m + \frac{p^2}{2m} + \ldots \tag{4.86}$$

We have a big constant, the mass m, which comes from relativity and is not present in non relativistic newtonian mechanics, and a small term which is indeed the non relativistic kinetic energy. Therefore, we consider the limit in which the momenta and energies are small with respect to the big term m. We define

$$E' = E - m\,, \tag{4.87}$$

and therefore we have $E' \ll m$.

The positive-energy solutions will oscillate with a term that is as big as m, $\sim e^{-imt}$, and a slightly varying term $\sim e^{-iE't}$. In order to study the latter, we have to factorize the former. We put

$$\phi = \varphi\, e^{-imt}\,, \tag{4.88}$$

in such a way that

$$\left| i\frac{\partial \varphi}{\partial t} \right| \sim E'\varphi \ll m\varphi \,. \tag{4.89}$$

We have, at first order in E':

$$\begin{aligned}
\frac{\partial \phi}{\partial t} &= \left(\frac{\partial \varphi}{\partial t} - im\varphi \right) e^{-imt} \,, & (4.90)\\
\frac{\partial^2 \phi}{\partial t^2} &= \left(\frac{\partial^2 \varphi}{\partial t^2} - im\frac{\partial \varphi}{\partial t} \right) e^{-imt} - im \left(\frac{\partial \varphi}{\partial t} - im\varphi \right) e^{-imt} \,, & (4.91)\\
&\simeq \left(-i\, 2m\frac{\partial \varphi}{\partial t} - m^2\varphi \right) e^{-imt} \,. & (4.92)
\end{aligned}$$

Substituting in the Klein-Gordon equation we have

$$\frac{\partial^2 \phi}{\partial t^2} = \nabla^2\phi - m^2\phi \tag{4.93}$$

and, therefore

$$\left(-i\, 2m\frac{\partial \varphi}{\partial t} - m^2\varphi \right) e^{-imt} = \nabla^2\varphi e^{-imt} - m^2\varphi e^{-imt} \,. \tag{4.94}$$

Finally

$$i\frac{\partial \varphi}{\partial t} = -\frac{1}{2m}\nabla^2\varphi \,, \tag{4.95}$$

which is the Schródinger equation for a free spinless particle.

4.1.7 The two-component form

The Klein-Gordon equation is second order in time. Every second-order differential equation is equivalent to a system of first-order differential equations. We can then cast the Klein-Gordon equation into a form which is "similar" to the Schrödinger equation, loosing the manifest covariance and getting an equation that involves a two-component field and a matrix 2x2 that plays the role of the hamiltonian in the Schródinger equation (however, in this two-component description the Schrödinger-like hamiltonian is not even hermitian).

We consider then ϕ and $\partial_0\phi$ as independent fields, defining two fields, ϕ_+ and ϕ_-, as follows

$$\begin{aligned}
\phi_+ &= \frac{1}{\sqrt{2m}} \left(i\partial_0\phi + m\phi \right) \,, & (4.96)\\
\phi_- &= \frac{1}{\sqrt{2m}} \left(-i\partial_0\phi + m\phi \right) \,, & (4.97)
\end{aligned}$$

such that

$$\begin{aligned}
\phi &= \frac{1}{\sqrt{2m}} \left(\phi_+ + \phi_- \right) \,, & (4.98)\\
i\partial_0\phi &= \sqrt{\frac{m}{2}} \left(\phi_+ - \phi_- \right) \,. & (4.99)
\end{aligned}$$

(For the moment we do not include interactions).

Let us take the derivatives with respect to time of $\phi_\pm$:

$$i\partial_0\phi_\pm = \frac{1}{\sqrt{2m}} \left(\mp\frac{\partial^2}{\partial t^2}\phi + mi\partial_0\phi \right) \,, \tag{4.100}$$

$$
\begin{aligned}
&= \ |\,\text{using the Klein-Gordon equation}\,| \\
&= \frac{1}{\sqrt{2m}}\left[\mp(-p^2-m^2)\phi + mi\partial_0\phi\right] , && (4.101)\\
&= \frac{1}{\sqrt{2m}}\left[\mp(-p^2-m^2)\frac{1}{\sqrt{2m}}(\phi_+ + \phi_-) + m\sqrt{\frac{m}{2}}(\phi_+ - \phi_-)\right] , && (4.102)\\
&= \pm\left(\frac{p^2}{2m}+\frac{m}{2}\right)(\phi_+ + \phi_-) + \frac{m}{2}(\phi_+ - \phi_-) . && (4.103)
\end{aligned}
$$

Finally

$$
\begin{cases} i\partial_0\phi_+ = \left(\frac{p^2}{2m}+m\right)\phi_+ + \frac{p^2}{2m}\phi_- , \\ i\partial_0\phi_- = -\frac{p^2}{2m}\phi_+ - \left(\frac{p^2}{2m}+m\right)\phi_- . \end{cases} \tag{4.104}
$$

If we arrange $\phi_\pm$ in a two-component vector

$$
\Psi = \begin{pmatrix}\phi_+\\ \phi_-\end{pmatrix} \tag{4.105}
$$

we find the following form for the KG equation

$$
(i\partial_0 - H)\Psi = 0 . \tag{4.106}
$$

In Eq. (4.106), we defined the following matrix

$$
H = \begin{pmatrix} \frac{p^2}{2m}+m & \frac{p^2}{2m} \\ -\frac{p^2}{2m} & -\frac{p^2}{2m}-m \end{pmatrix} = \left(\frac{p^2}{2m}+m\right)\tau^3 + \frac{p^2}{2m}\,i\tau^2 , \tag{4.107}
$$

where

$$
\tau^1 = \begin{pmatrix}0&1\\1&0\end{pmatrix} , \quad \tau^2 = \begin{pmatrix}0&-i\\i&0\end{pmatrix} , \quad \tau^3 = \begin{pmatrix}1&0\\0&-1\end{pmatrix} , \tag{4.108}
$$

are the Pauli matrices. Note that although we found a natural representation in terms of the Pauli matrices, we are not speaking about spin. The particle we are describing with the KG equation are spinless particles. Here we are considering an $SU(2)$ rotation, but in another space (the one identified by the two-component vectors Ψ).

Since

$$
(i\partial_0 + H)(i\partial_0 - H) = -\frac{\partial^2}{\partial t^2} - H^2 \tag{4.109}
$$

and

$$
H^2 = \left(\frac{p^2}{2m}+m\right)^2(\tau^3)^2 - \frac{p^4}{4m^2}(\tau^2)^2 + i\frac{p^2}{2m}\left(\frac{p^2}{2m}+m\right)[\tau^3,\tau^2]_+ = (p^2+m^2)\mathbb{1} , \tag{4.110}
$$

we find that since Eq. (4.106) holds, we have

$$
(i\partial_0 + H)(i\partial_0 - H)\Psi = 0 , \tag{4.111}
$$

and therefore, because of Eq. (4.109):

$$
\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2\right)\Psi = 0 . \tag{4.112}
$$

Both ϕ_+ and ϕ_- satisfy the KG equation.

The operator H is not hermitian, $H^\dagger \neq H$. However, it is hermitian "in τ^3 metric", i.e.

$$\tau^3 H^\dagger \tau^3 = H\,, \tag{4.113}$$

as can be checked by direct inspection.

We can also plug in the new form the conserved charge

$$\begin{aligned} Q &= i\int d^3X\,\phi^*\overleftrightarrow{\partial^0}\phi = \frac{i}{2m}\int d^3X\left(\phi_+^* + \phi_-^*\right)\overleftrightarrow{\partial^0}\left(\phi_+ + \phi_-\right)\,, && (4.114)\\ &= \ldots\,, \\ &= \int d^3X\,\left(|\phi_+|^2 - |\phi_-|^2\right)\,, && (4.115)\\ &= \int d^3X\,\Psi^\dagger\tau^3\Psi\,. && (4.116)\end{aligned}$$

We can introduce the "Klein-Gordon" adjoint

$$\overline{\Psi} = \Psi^\dagger\tau^3 \tag{4.117}$$

and write then

$$Q = \int d^3X\,\overline{\Psi}\Psi\,. \tag{4.118}$$

Plane wave solutions

We can study the solutions of the KG equation in the two-component form. We look for a solution of the following kind

$$\Psi \sim A\begin{pmatrix}\phi_+\\ \phi_-\end{pmatrix}e^{-iEt+i\mathbf{p}\cdot\mathbf{x}}\,. \tag{4.119}$$

Substituting in Eq. (4.106) we find the system

$$\begin{pmatrix} E - m - \frac{p^2}{2m} & -\frac{p^2}{2m}\\ \frac{p^2}{2m} & E + m + \frac{p^2}{2m}\end{pmatrix}\begin{pmatrix}\phi_+\\ \phi_-\end{pmatrix} = 0\,, \tag{4.120}$$

that has a non-trivial solution if

$$\det\begin{pmatrix} E - m - \frac{p^2}{2m} & -\frac{p^2}{2m}\\ \frac{p^2}{2m} & E + m + \frac{p^2}{2m}\end{pmatrix} = E^2 - p^2 - m^2 = 0\,. \tag{4.121}$$

Eq. (4.121) require positive and negative energy solutions

$$E = \pm\sqrt{p^2+m^2} = \pm\omega_p \tag{4.122}$$

and the system becomes

$$\begin{cases} \left(\pm\omega_p - m - \frac{p^2}{2m}\right)\phi_+ = \frac{p^2}{2m}\phi_- \\ \frac{p^2}{2m}\phi_+ = -\left(\pm\omega_p + m + \frac{p^2}{2m}\right)\phi_- \end{cases} \tag{4.123}$$

Positive-energy solutions

The system is Eq. (4.123) in which we consider $+\omega_p$:

$$\begin{aligned} \left(\omega_p - m - \tfrac{p^2}{2m}\right)\phi_+^{(+)} &= \tfrac{p^2}{2m}\phi_-^{(+)} \\ \tfrac{p^2}{2m}\phi_+^{(+)} &= -\left(\omega_p + m + \tfrac{p^2}{2m}\right)\phi_-^{(+)} \end{aligned} \tag{4.124}$$

that gives

$$\frac{\phi_+^{(+)}}{\phi_-^{(+)}} = \frac{\frac{p^2}{2m}}{\omega_p - m - \frac{p^2}{2m}} \tag{4.125}$$

and since

$$\omega_p - m - \frac{p^2}{2m} = \frac{(\omega_p - m)^2}{2m}\,, \tag{4.126}$$

$$\omega_p + m + \frac{p^2}{2m} = \frac{(\omega_p + m)^2}{2m}\,, \tag{4.127}$$

$$p^2 = (\omega_p + m)(\omega_p - m)\,, \tag{4.128}$$

we find

$$\frac{\phi_+^{(+)}}{\phi_-^{(+)}} = \frac{\frac{p^2}{2m}}{\omega_p - m - \frac{p^2}{2m}} = -\frac{\omega_p + m}{\omega_p - m}\,. \tag{4.129}$$

We choose the positive-energy solution as

$$\Psi^{(+)} = A^{(+)}\begin{pmatrix}\omega_p + m \\ \omega_p - m\end{pmatrix} e^{-i\omega_p t + i\mathbf{p}\cdot\mathbf{x}}\,. \tag{4.130}$$

The normalization factor $A^{(+)}$ will be discussed below.

Negative-energy solutions

The system is Eq. (4.123) in which we consider $-\omega_p$:

$$\begin{aligned} \left(-\omega_p - m - \tfrac{p^2}{2m}\right)\phi_+^{(-)} &= \tfrac{p^2}{2m}\phi_-^{(-)} \\ \tfrac{p^2}{2m}\phi_+^{(-)} &= -\left(-\omega_p + m + \tfrac{p^2}{2m}\right)\phi_-^{(-)} \end{aligned} \tag{4.131}$$

that gives

$$\frac{\phi_+^{(-)}}{\phi_-^{(-)}} = -\frac{\frac{p^2}{2m}}{\omega_p + m + \frac{p^2}{2m}} = -\frac{\omega_p - m}{\omega_p + m}\,. \tag{4.132}$$

Then

$$\Psi^{(-)} = A^{(-)}\begin{pmatrix}\omega_p - m \\ \omega_p + m\end{pmatrix} e^{i\omega_p t + i\mathbf{p}\cdot\mathbf{x}}\,. \tag{4.133}$$

The normalization factor $A^{(-)}$ will be discussed below.

Normalization

For the normalization of the two solutions we impose

$$\int d^3X\,\bar{\Psi}_p^{(+)}\Psi_{p'}^{(+)} = \delta(\mathbf{p} - \mathbf{p}')\,, \tag{4.134}$$

$$\int d^3X\,\bar{\Psi}^{(-)}_p\Psi^{(-)}_{p'} \;=\; -\delta(\mathbf{p}-\mathbf{p}')\,. \tag{4.135}$$

We have

$$\begin{aligned}\int d^3X\,\bar{\Psi}^{(+)}_p\Psi^{(+)}_{p'} &= \int d^3X\,\Psi^{(+)\dagger}_p\tau^3\Psi^{(+)}_{p'}\,, && (4.136)\\ &= \left|A^{(+)}\right|^2 e^{-i(\omega_{p'}-\omega_p)t}\left[(m+\omega_p)(m+\omega_{p'})-(\omega_p-m)(\omega_{p'}-m)\right]\int d^3X\,e^{i(\mathbf{p}'-\mathbf{p})\cdot\mathbf{x}}\,,\\ &= \left|A^{(+)}\right|^2 (2\pi)^3\,4m\omega_p\,\delta(\mathbf{p}-\mathbf{p}') && (4.137)\end{aligned}$$

and therefore[5]

$$A^{(+)} = \frac{1}{\sqrt{(2\pi)^3\,4m\omega_p}}\,. \tag{4.138}$$

For $A^{(-)}$ we find the same expression

$$A^{(-)} = \frac{1}{\sqrt{(2\pi)^3\,4m\omega_p}}\,. \tag{4.139}$$

Finally:

$$\begin{aligned}\Psi^{(+)} &= \frac{1}{\sqrt{(2\pi)^3\,2\omega_p}}\frac{m+\omega_p}{\sqrt{2m}}\begin{pmatrix}1\\ \frac{\omega_p-m}{\omega_p+m}\end{pmatrix}e^{-i\omega_p t+i\mathbf{p}\cdot\mathbf{x}}\,, && (4.140)\\ &= \frac{1}{\sqrt{(2\pi)^3\,2\omega_p}}\frac{m+\omega_p}{\sqrt{2m}}\begin{pmatrix}1\\ \frac{p^2}{(\omega_p+m)^2}\end{pmatrix}e^{-i\omega_p t+i\mathbf{p}\cdot\mathbf{x}}\,, && (4.141)\\ \Psi^{(-)} &= \frac{1}{\sqrt{(2\pi)^3\,2\omega_p}}\frac{m+\omega_p}{\sqrt{2m}}\begin{pmatrix}\frac{\omega_p-m}{\omega_p+m}\\ 1\end{pmatrix}e^{i\omega_p t+i\mathbf{p}\cdot\mathbf{x}}\,, && (4.142)\\ &= \frac{1}{\sqrt{(2\pi)^3\,2\omega_p}}\frac{m+\omega_p}{\sqrt{2m}}\begin{pmatrix}\frac{p^2}{(\omega_p+m)^2}\\ 1\end{pmatrix}e^{i\omega_p t+i\mathbf{p}\cdot\mathbf{x}}\,. && (4.143)\end{aligned}$$

Charge conjugation

The Klein-Gordon equation, as the other covariant equations, has a symmetry related to the existence of both positive and negative energy solutions. These solutions can be transformed into each other by *charge conjugation*:

$$\phi \to \phi^C = \tau_1\phi^*\,, \tag{4.144}$$

such that $\left(\phi^C\right)^C = \tau_1\left(\tau^1\phi^*\right)^* = \phi$.

ϕ^C satisfies the same free[6] equation as ϕ. In fact, taking the complex conjugate of Eq. (4.106), with Eq. (4.107), we find

$$-i\frac{\partial}{\partial t}\phi^* = \left\{\left(\frac{p^2}{2m}+m\right)\tau^3 + \frac{p^2}{2m}(-i)(-\tau^2)\right\}\phi^*\,. \tag{4.145}$$

If now we multiply on the l.h.s. by τ^1, recalling the fact that τ^1 anti-commutes with τ^2 and τ^3, we have

$$i\frac{\partial}{\partial t}\phi^C = \left\{\left(\frac{p^2}{2m}+m\right)\tau^3 + \frac{p^2}{2m}i)\tau^2\right\}\phi^C\,. \tag{4.146}$$

[5]We choose $A^{(+)}$ real.

[6]When we will introduce the electromagnetic interaction, we will see that ϕ^C is a solution of the equation in which the electric charge changes sign.

The charge conjugated field ϕ^C has a charge which is opposite to the charge of ϕ. In fact

$$\phi^C = \tau_1 \phi^* , \quad \left(\phi^C\right)^\dagger = (\phi^*)^\dagger \, (\tau^1)^\dagger = \phi^t \tau^1 . \tag{4.147}$$

Therefore

$$Q' = \int d^3X \, \left(\phi^C\right)^\dagger \tau^3 \phi^C = \int d^3X \, \phi^t \tau^1 \tau^3 \tau^1 \phi^* , \tag{4.148}$$

$$= - \int d^3X \, \left(|\phi_+|^2 - |\phi_-|^2\right) = -Q . \tag{4.149}$$

Concerning the plane-wave solutions, we find that charge conjugation connects the negative-energy solution to the positive-energy one in the following way:

$$\Psi^{(-)\,C}_{-p} = \tau^1 \frac{1}{\sqrt{(2\pi)^3 \, 2\omega_p}} \frac{m+\omega_p}{\sqrt{2m}} \begin{pmatrix} \frac{p^2}{(\omega_p+m)^2} \\ 1 \end{pmatrix} e^{-i\omega_p t + i \mathbf{p}\cdot\mathbf{x}} , \tag{4.150}$$

$$= \frac{1}{\sqrt{(2\pi)^3 \, 2\omega_p}} \frac{m+\omega_p}{\sqrt{2m}} \begin{pmatrix} 1 \\ \frac{p^2}{(\omega_p+m)^2} \end{pmatrix} e^{i\omega_p t + i \mathbf{p}\cdot\mathbf{x}} = \Psi^{(+)}_p . \tag{4.151}$$

Therefore, charge conjugation turns a negative-energy state of momentum $-p$ into a positive-energy one of opposite charge and momentum $+p$. If we "connect" ϕ with a "particle" state, ϕ^C will be connected with an anti-particle state.

4.2 Quantization of the Klein-Gordon field

We worked out the theory of the classical $\phi(X)$ field and we analysed the system using the lagrangian and the hamiltonian formalism. The first gives us the possibility to find the conserved quantities of the physical system, as the charges of the symmetries of the corresponding action. The second is the correct framework for the canonical quantization.

4.2.1 Real field

Just to recap, the real field, $\phi(X)$, has to satisfy the Klein-Gordon equation

$$(\partial^2 + m^2)\phi(X) = 0 , \tag{4.152}$$

which is the Euler-Lagrange equation of the following lagrangian density

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2\right) . \tag{4.153}$$

The lagrangian density is invariant under Poincaré transformations and, therefore, it follows that the four-momentum and the generalized angular momentum (in particular the orbital angular momentum, since the real scalar field does not have spin) are conserved.

The energy density coming from Nöther's theorem is

$$\mathcal{H} = T^0_0 = \frac{1}{2}\left(\dot{\phi}^2 + (\nabla\phi)^2 + m^2\phi^2\right) = \frac{1}{2}\left(\pi^2 + (\nabla\phi)^2 + m^2\phi^2\right) , \tag{4.154}$$

since

$$\pi(X) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \dot{\phi} \tag{4.155}$$

and therefore it coincides with the hamiltonian density that can be found via a Legendre transformation

$$\mathcal{H} = \pi\dot{\phi} - \mathcal{L} . \tag{4.156}$$

Note that $\mathcal{H}$ is positive definite. The hamiltonian can be diagonalized in terms of normal modes if we find a plane-wave solution of Eq. (4.152)

$$\phi(X) = \int \frac{d^3p}{\sqrt{(2\pi)^3\, 2\omega_p}} \left[a(p)e^{-ip_\mu X^\mu} + a^*(p)e^{ip_\mu X^\mu}\right] . \tag{4.157}$$

We have

$$H = \int d^3p\, \frac{\omega_p}{2} \left[a(p)a^*(p) + a^*(p)a(p)\right] , \tag{4.158}$$

that has the form of an infinite sum of harmonic oscillators. This suggests the right way for the quantization of this system. We will have to promote the field (4.157) from a classical function to an operator. In order to do that, we can only interpret the coefficients $a(p)$ and $a^*(p)$ as operators

$$a(p) \to \hat{a}(p)\,, \qquad a^*(p) \to \hat{a}^\dagger(p) \tag{4.159}$$

and we will have to check that these operators are indeed creation-annihilation operators, as we can understand from the form of the hamiltonian.

Considering the correspondence with the non-relativistic point-like particle, the degree of freedom that was describing the position at a certain time in that framework, corresponds now to the field in a certain point at the time t:

$$\mathbf{q}(t) \to \phi(\mathbf{x}, t)\,. \tag{4.160}$$

The conjugated momentum, $\mathbf{p}(t)$, corresponds to the conjugated momentum $\pi(\mathbf{x}, t)$.

Note that the description in terms of the fields is given naturally in a time-dependent framework. When we promote the field to an operator, this operator will have to be treated in Heisenberg picture.

To the canonical quantization relations

$$[q_i(t), p_j(t)] = i\,\delta_{ij}\,, \qquad [q_i(t), q_j(t)] = [p_i(t), p_j(t)] = 0\,, \tag{4.161}$$

will have to involve the fields. We will have to impose the following equal-time quantization relations

$$[\phi(\mathbf{x}, t), \pi(\mathbf{y}, t)] = i\,\delta(\mathbf{x} - \mathbf{y})\,, \tag{4.162}$$

$$[\phi(\mathbf{x}, t), \phi(\mathbf{y}, t)] = [\pi(\mathbf{x}, t), \pi(\mathbf{y}, t)] = 0\,, \tag{4.163}$$

where now the field is

$$\int \frac{d^3p}{\sqrt{(2\pi)^3\, 2\omega_p}} \left[\hat{a}(p)e^{-iP_\mu X^\mu} + \hat{a}^\dagger(p)e^{iP_\mu X^\mu}\right] \tag{4.164}$$

and $\pi(X) = \dot{\phi}(X)$.

We have to check that the quantization relations (4.162,4.163) imply that $\hat{a}^\dagger$ and $\hat{a}$ are indeed creation-annihilation operators, and the hamiltonian can be written in terms of the number operator. In order to to that, we remember that[7]

$$\hat{a}(p) = (f_p^+, \phi) = i \int d^3X\, \left(f_p^+\right)^* \overleftrightarrow{\partial_0}\, \phi\,, \tag{4.168}$$

[7]This can be checked by direct inspection, using the form of f^+ and the field in normal modes. In fact

$$(f_p^+, \phi) = i \int d^3X\, \left(f_p^+\right)^* \overleftrightarrow{\partial_0}\, \phi\,, \tag{4.165}$$

$$= i \int d^3X \int \frac{d^3p'}{(2\pi)^3\sqrt{4\omega_p\omega_{p'}}} e^{iP_\mu X^\mu} \overleftrightarrow{\partial_0} \left[\hat{a}(p')e^{-iP'_\mu X^\mu} + \hat{a}^\dagger(p')e^{iP'_\mu X^\mu}\right] , \tag{4.166}$$

$$= .. = \hat{a}(p)\,. \tag{4.167}$$

$$\hat{a}^\dagger(p) = -\left(f_p^-, \phi\right) = -i\int d^3X \left(f_p^-\right)^* \overleftrightarrow{\partial_0}\phi\,. \tag{4.169}$$

Therefore

$$\begin{aligned}
\hat{a}(p) &= i\int d^3X \frac{e^{iP_\mu X^\mu}}{\sqrt{(2\pi)^3\,2\omega_p}}\overleftrightarrow{\partial_0}\phi\,, &(4.170)\\
&= i\int \frac{d^3X}{\sqrt{(2\pi)^3\,2\omega_p}}\left(e^{iP_\mu X^\mu}\dot{\phi} - i\omega_p e^{iP_\mu X^\mu}\phi\right)\,, &(4.171)\\
&= \int \frac{d^3X}{\sqrt{(2\pi)^3\,2\omega_p}}\left(\omega_p\phi + i\dot{\phi}\right)e^{iP_\mu X^\mu}\,, &(4.172)\\
\hat{a}^\dagger(p) &= \int \frac{d^3X}{\sqrt{(2\pi)^3\,2\omega_p}}\left(\omega_p\phi - i\dot{\phi}\right)e^{-iP_\mu X^\mu}\,. &(4.173)
\end{aligned}$$

With these expressions we can construct the commutator $[a(p), a^\dagger(p')]$ and, using the quantization relations for the fields, prove that $[a(p), a^\dagger(p')] = \delta(\mathbf{p}-\mathbf{p}')$. We have

$$\begin{aligned}
[\hat{a}(p), \hat{a}^\dagger(p')] &= \int \frac{d^3X\, d^3Y}{(2\pi)^3\sqrt{4\omega_p\omega_{p'}}}\Big[\left(\omega_p\phi(X) + i\dot{\phi}(X)\right)e^{iP_\mu X^\mu}\left(\omega_{p'}\phi(Y) - i\dot{\phi}(Y)\right)e^{-iP'_\mu Y^\mu}\,,\\
&\qquad - \left(\omega_{p'}\phi(Y) - i\dot{\phi}(Y)\right)e^{-iP'_\mu Y^\mu}\left(\omega_p\phi(X) + i\dot{\phi}(X)\right)e^{iP_\mu X^\mu}\Big]\,, &(4.174)\\
&= |\,\text{where we have to remember that } X^0 = Y^0 = t\,|\\
&= \int \frac{d^3X\, d^3Y}{(2\pi)^3\sqrt{4\omega_p\omega_{p'}}}e^{-iP'_\mu Y^\mu + iP_\mu X^\mu}\Big[\left(\omega_p\phi(X) + i\dot{\phi}(X)\right)\left(\omega_{p'}\phi(Y) - i\dot{\phi}(Y)\right)\,,\\
&\qquad - \left(\omega_{p'}\phi(Y) - i\dot{\phi}(Y)\right)\left(\omega_p\phi(X) + i\dot{\phi}(X)\right)\Big]\,, &(4.175)\\
&= \int \frac{d^3X\, d^3Y}{(2\pi)^3\sqrt{4\omega_p\omega_{p'}}}e^{-i(\omega_{p'}-\omega_p)t}e^{i\mathbf{p}'\cdot\mathbf{y} - i\mathbf{p}\cdot\mathbf{x}}\Big(\omega_p\omega_{p'}[\phi(\mathbf{x},t), \phi(\mathbf{y},t)] - i\omega_p[\phi(\mathbf{x},t), \dot{\phi}(\mathbf{y},t)]\,,\\
&\qquad + i\omega_{p'}[\dot{\phi}(\mathbf{x},t), \phi(\mathbf{y},t)] + [\dot{\phi}(\mathbf{x},t), \dot{\phi}(\mathbf{y},t)]\Big)\,, &(4.176)\\
&= \int \frac{d^3X\, d^3Y}{(2\pi)^3\sqrt{4\omega_p\omega_{p'}}}e^{-i(\omega_{p'}-\omega_p)t}e^{i\mathbf{p}'\cdot\mathbf{y} - i\mathbf{p}\cdot\mathbf{x}}\,(\omega_p + \omega_{p'})\delta(\mathbf{x}-\mathbf{y})\,, &(4.177)\\
&= \int \frac{d^3X}{(2\pi)^3\sqrt{4\omega_p\omega_{p'}}}e^{-i(\omega_{p'}-\omega_p)t}e^{-i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}}\,(\omega_p + \omega_{p'})\,, &(4.178)\\
&= \frac{1}{\sqrt{4\omega_p\omega_{p'}}}e^{-i(\omega_{p'}-\omega_p)t}\,(\omega_p + \omega_{p'})\,\delta(\mathbf{p}-\mathbf{p}')\,, &(4.179)\\
&= \delta(\mathbf{p}-\mathbf{p}')\,. &(4.180)
\end{aligned}$$

In the same way we find

$$[\hat{a}(p), \hat{a}(p')] = [\hat{a}^\dagger(p), \hat{a}^\dagger(p')] = 0\,. \tag{4.181}$$

Moreover, using the expression of the fields we find

$$\begin{aligned}
H &= \int d^3X\,\frac{1}{2}\left(\dot{\phi}^2 + (\nabla\phi)^2 + m^2\phi^2\right)\,, &(4.182)\\
&= \int d^3p\frac{\omega_p}{2}\left(\hat{a}(p)\hat{a}^\dagger(p) + \hat{a}^\dagger(p)\hat{a}(p)\right)\,, &(4.183)
\end{aligned}$$

that in normal ordering gives

$$:H:= \int d^3p\,\omega_p\,\hat{a}^\dagger(p)\hat{a}(p)\,. \tag{4.184}$$

The operators $\hat{a}(p)$ and $\hat{a}^\dagger(p)$ are therefore, indeed, annihilation and creation operators. They act on the Fock space, in such a way that

$$\hat{a}(p)|0\rangle = 0\,. \tag{4.185}$$

In this way, the energy of the vacuum is 0:

$$:H:|0\rangle = \int d^3p\omega_p\,\hat{a}^\dagger(p)\hat{a}(p)\,|0\rangle = 0\,. \tag{4.186}$$

If we act once with $\hat{a}^\dagger(p)$ on the vacuum we find a one-particle state with definite energy (and momentum):

$$\hat{a}^\dagger(p)|0\rangle = |p\rangle\,, \tag{4.187}$$

such that

$$\begin{aligned}
:H:\hat{a}^\dagger(p)|0\rangle &= \int d^3p'\omega_{p'}\,\hat{a}^\dagger(p')\hat{a}(p')\,\hat{a}^\dagger(p)|0\rangle\,, & (4.188)\\
&= \int d^3p'\omega_{p'}\,\hat{a}^\dagger(p')\delta(\mathbf{p}-\mathbf{p}')|0\rangle + \int d^3p'\omega_{p'}\,\hat{a}^\dagger(p')\hat{a}^\dagger(p)\hat{a}(p')|0\rangle\,, & (4.189)\\
&= \omega_p\,\hat{a}^\dagger(p)|0\rangle\,. & (4.190)
\end{aligned}$$

Therefore, $\hat{a}^\dagger(p)|0\rangle$ is an eigenstate of the hamiltonian with energy ω_p. If we consider, for example, the state

$$|p_1,p_2\rangle = \hat{a}^\dagger(p_1)\hat{a}^\dagger(p_2)|0\rangle\,, \tag{4.191}$$

we find:

$$\begin{aligned}
:H:\hat{a}^\dagger(p_1)\hat{a}^\dagger(p_2)|0\rangle &= \int d^3p\omega_p\,\hat{a}^\dagger(p)\hat{a}(p)\,\hat{a}^\dagger(p_1)\hat{a}^\dagger(p_2)|0\rangle\,, & (4.192)\\
&= \int d^3p\omega_p\,\hat{a}^\dagger(p)\left(\hat{a}^\dagger(p_1)\hat{a}(p)+\delta(\mathbf{p}-\mathbf{p}_1)\right)\hat{a}^\dagger(p_2)|0\rangle\,, & (4.193)\\
&= \int d^3p\omega_p\,\hat{a}^\dagger(p)\hat{a}^\dagger(p_1)\hat{a}(p)\hat{a}^\dagger(p_2)|0\rangle + \omega_{p_1}\hat{a}^\dagger(p_1)\hat{a}^\dagger(p_2)|0\rangle\,, & (4.194)\\
&= \int d^3p\omega_p\,\hat{a}^\dagger(p)\hat{a}^\dagger(p_1)\left(\hat{a}^\dagger(p_2)\hat{a}(p)+\delta(\mathbf{p}-\mathbf{p}_2)\right)|0\rangle \\
&\qquad +\omega_{p_1}\hat{a}^\dagger(p_1)\hat{a}^\dagger(p_2)|0\rangle\,, & (4.195)\\
&= (\omega_{p_1}+\omega_{p_2})\,\hat{a}^\dagger(p_1)\hat{a}^\dagger(p_2)|0\rangle\,. & (4.196)
\end{aligned}$$

Therefore, $\hat{a}^\dagger(p_1)\hat{a}^\dagger(p_2)|0\rangle$ is again an eigenstate of $:H:$ with energy $(\omega_{p_1}+\omega_{p_2})$. Etc ...

Let us look at the momentum of these states. From Nöther's theorem we have[8]

$$\begin{aligned}
P^i &= \int d^3X\,T^{0i} = \int d^3X\,\dot{\phi}\,\partial^i\phi\,, & (4.197)\\
&= \int d^3X \int \frac{d^3p\,d^3p'}{(2\pi)^3\,\sqrt{4\omega_p\omega_{p'}}}(-i\omega_p)\left(a(p)e^{-iP_\mu X^\mu}-a^\dagger(p)e^{iP_\mu X^\mu}\right)(-ip'^i)\Big(a(p')e^{-iP'_\mu X^\mu} \\
&\quad -a^\dagger(p')e^{iP'_\mu X^\mu}\Big)\,, & (4.198)\\
&= -\int \frac{d^3p\,d^3p'}{(2\pi)^3\,\sqrt{4\omega_p\omega_{p'}}}\,\omega_p p'^i \int d^3X\,\big[(a(p)a(p')e^{-i(P_\mu+P'_\mu)X^\mu}+a^\dagger(p)a^\dagger(p')e^{i(P_\mu+P'_\mu)X^\mu}) \\
&\qquad -(a(p)a^\dagger(p')e^{-i(P_\mu-P'_\mu)X^\mu}+a^\dagger(p)a(p')e^{i(P_\mu-P'_\mu)X^\mu})\big]\,, & (4.199)\\
&= \big|\,\text{integrating in } d^3X\,\big|
\end{aligned}$$

[8]from now on we will omit the "hat" on the creation/annihilation operators.

$$= \int \frac{d^3p\, d^3p'}{\sqrt{4\omega_p\omega_{p'}}}\, \omega_p p'^i \left(a(p)a^\dagger(p') + a^\dagger(p)a(p')\right)\delta(\mathbf{p} - \mathbf{p}')\,, \tag{4.200}$$

$$= \left|\,\text{integrating in } d^3p'\,\right|$$

$$= \int d^3p\, \frac{1}{2} p^i \left(a(p)a^\dagger(p) + a^\dagger(p)a(p)\right), \tag{4.201}$$

where, from (4.199) to (4.200), we considered that

$$\int \frac{d^3p\, d^3p'}{(2\pi)^3\sqrt{4\omega_p\omega_{p'}}}\, \omega_p\, p'^i a(p)a(p')e^{-i(\omega_p+\omega_{p'})t}\delta(\mathbf{p} + \mathbf{p}') \;=\; -\int d^3p\, \frac{1}{2}p^i\, a(p)a(-p)\, e^{-2i\omega_p}\,, \tag{4.202}$$

but

$$\int d^3p\, \frac{1}{2}p^i\, a(p)a(-p)\, e^{-2i\omega_p} \;=\; \int d^3p\, \frac{1}{2}p^i\, a(-p)a(p)\, e^{-2i\omega_p}\,, \tag{4.203}$$

$$= \left|\,\text{since } [a(p), a(-p)] = 0\,\right|$$

$$= -\int d^3p\, \frac{1}{2}p^i\, a(p)a(-p)\, e^{-2i\omega_p}\,, \tag{4.204}$$

$$= \left|\,\text{where we changed } p \to -p\,\right|. \tag{4.205}$$

Therefore

$$\int d^3p\, \frac{1}{2}p^i\, a(p)a(-p)\, e^{-2i\omega_p} = 0 \tag{4.206}$$

and, finally

$$\int \frac{d^3p\, d^3p'}{(2\pi)^3\sqrt{4\omega_p\omega_{p'}}}\, \omega_p\, p'^i a(p)a(p')e^{-i(\omega_p+\omega_{p'})t}\delta(\mathbf{p} + \mathbf{p}') = 0\,. \tag{4.207}$$

The same is true for

$$\int \frac{d^3p\, d^3p'}{(2\pi)^3\sqrt{4\omega_p\omega_{p'}}}\, \omega_p\, p'^i a^\dagger(p)a^\dagger(p')e^{i(\omega_p+\omega_{p'})t}\delta(\mathbf{p} + \mathbf{p}') = 0\,. \tag{4.208}$$

In normal ordering, then, we have

$$: P^i := \int d^3p\, p^i\, a^\dagger(p)a(p)\,, \tag{4.209}$$

that commutes with the hamiltonian (as it should) and such that

$$: P^i : a^\dagger(p)|0\rangle \;=\; p^i a^\dagger(p)|0\rangle\,, \tag{4.210}$$

$$: P^i : a^\dagger(p_1)a^\dagger(p_2)|0\rangle \;=\; (p_1^i + p_2^i)a^\dagger(p_1)a^\dagger(p_2)|0\rangle\,, \tag{4.211}$$

etc ...

These results corroborate the interpretation of the state $|p_1, p_2\rangle = a^\dagger(p_1)a^\dagger(p_2)|0\rangle$ as a two-particle state, with definite energy, which is the sum of the energies of the one-particle state $|p_1\rangle$ and the one-particle state $|p_2\rangle$, and with definite momentum, which is the vectorial sum of the momenta $\mathbf{p}_1$ and $\mathbf{p}_2$.

One- and two-particle states. Bosons

The state $|p\rangle = a^\dagger(p)|0\rangle$ is a plane wave, a state with definite energy and momentum and therefore totally delocalized. A one-particle state will be described as a superposition of plane waves

$$|\psi\rangle = \int d^3p\, \psi(p)a^\dagger(p)|0\rangle\,. \tag{4.212}$$

The function $\psi(p)$ is the actual wave function in p-representation. In fact

$$\begin{aligned}
\langle p'|\psi\rangle &= \langle 0|a(p')\int d^3p\,\psi(p)a^\dagger(p)|0\rangle\,, && (4.213)\\
&= \int d^3p\,\psi(p)\langle 0|a(p')a^\dagger(p)|0\rangle\,, && (4.214)\\
&= \int d^3p\,\psi(p)\delta(\mathbf{p}-\mathbf{p}') = \psi(p') && (4.215)
\end{aligned}$$

gives the probability amplitude to have a particle with momentum p'. We have

$$\begin{aligned}
\langle\psi|\psi\rangle &= \int d^3p\,\psi^*\langle 0|a(p)\int d^3q\psi(q)a^\dagger(q)|0\rangle\,, && (4.216)\\
&= \int d^3pd^3q\,\psi^*\psi(q)\delta(\mathbf{p}-\mathbf{q})\,, && (4.217)\\
&= \int d^3p\,|\psi(p)|^2\,. && (4.218)
\end{aligned}$$

Therefore, $\psi(p)$ should be normalizable and we can choose $\langle\psi|\psi\rangle = 1$.

Let us consider now a two-particle state. As in the previous case, we can write

$$|\psi\rangle = \int d^3p_1\,d^3p_2\,\psi(p_1,p_2)a^\dagger(p_1)a^\dagger(p_2)|0\rangle\,. \qquad (4.219)$$

Due to the commutation relations we have

$$|p_1,p_2\rangle = a^\dagger(p_1)a^\dagger(p_2)|0\rangle = a^\dagger(p_2)a^\dagger(p_1)|0\rangle = |p_2,p_1\rangle\,. \qquad (4.220)$$

This means that in the integral (4.219) the function $\psi(p_1,p_2)$ should be symmetric in the exchange $1\leftrightarrow 2$ (or, in other words, only the symmetric part of $\psi(p_1,p_2)$ gives an ingtegral different from zero). It follows that the commutation relations that we used for the quantization of the KG field give rise to bosonic particles.

4.2.2 Complex field

Let us now discuss the quantization of the complex field $\phi(X)$ with lagrangian density

$$\mathcal{L} = \partial_\mu\phi^*\partial^\mu\phi - m^2\phi^*\phi\,, \qquad (4.221)$$

that can be found from the lagrangian of two real fields, ϕ_1 and ϕ_2, degenerate in mass, rotated in the complex plane as

$$\begin{aligned}
\phi &= \frac{\phi_1 + i\phi_2}{\sqrt{2}}\,, && (4.222)\\
\phi^* &= \frac{\phi_1 - i\phi_2}{\sqrt{2}}\,. && (4.223)
\end{aligned}$$

We promote the fields to operators, and then we will have ϕ and $\phi^\dagger$. We can find the expression of the fields in terms of creation-annihilation operators using Eqs. (4.222,4.223) in which we substitute the expressions for the two real fields ϕ_1 and ϕ_2. We find

$$\phi(X) = \int d^3p\left(f_p^+\,\frac{a_1(p)+ia_2(p)}{\sqrt{2}} + f_p^-\,\frac{a_1^\dagger(p)+ia_2^\dagger(p)}{\sqrt{2}}\right)\,,$$

$$= \int d^3p \left(f_p^+ \, a(p) + f_p^- \, b^\dagger(p) \right) , \tag{4.224}$$

$$\phi^\dagger(X) = \int d^3p \left(f_p^+ \frac{a_1(p) - i a_2(p)}{\sqrt{2}} + f_p^- \frac{a_1^\dagger(p) - i a_2^\dagger(p)}{\sqrt{2}} \right) ,$$

$$= \int d^3p \left(f_p^+ \, b(p) + f_p^- \, a^\dagger(p) \right) , \tag{4.225}$$

where we defined

$$a(p) = \frac{a_1(p) + i a_2(p)}{\sqrt{2}} , \tag{4.226}$$

$$b(p) = \frac{a_1(p) - i a_2(p)}{\sqrt{2}} . \tag{4.227}$$

Note that, trivially, $b^\dagger(p) \neq a^\dagger(p)$, as it should since the field is not anymore hermitian.

Interpreting $a(p)$, $a^\dagger(p)$ and $b(p)$, $b^\dagger(p)$ as creation-annihilation operators, we find a spectrum constituted by two kind of particles: "type a" and "type b" particles. Let us study their quantum numbers.

For the quantization of the system we have to find the conjugated momenta to ϕ and $\phi^\dagger$:

$$\pi_\phi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \dot{\phi}^\dagger , \tag{4.228}$$

$$\pi_{\phi^\dagger} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}^\dagger} = \dot{\phi} , \tag{4.229}$$

and impose the commutation relations at equal time:

$$[\phi(\mathbf{x},t), \dot{\phi}^\dagger(\mathbf{y},t)] = [\phi^\dagger(\mathbf{x},t), \dot{\phi}(\mathbf{y},t)] = i\delta(\mathbf{x}-\mathbf{y}) , \tag{4.230}$$

$$[\phi(\mathbf{x},t), \phi(\mathbf{y},t)] = [\phi^\dagger(\mathbf{x},t), \phi^\dagger(\mathbf{y},t)] = [\dot{\phi}(\mathbf{x},t), \dot{\phi}^\dagger(\mathbf{y},t)] = ... = 0 , \tag{4.231}$$

where the dots mean "all other combinations". These quantization rules induce analogous commutation relations among the operators $a(p)$, $a^\dagger(p)$ and $b(p)$, $b^\dagger(p)$. In fact we find:

$$[a(p), a^\dagger(p')] = [b(p), b^\dagger(p')] = \delta(\mathbf{p} - \mathbf{p}') , \tag{4.232}$$

and all the other combinations give zero commutator. It turnes out that $a(p)$, $a^\dagger(p)$ and $b(p)$, $b^\dagger(p)$ are indeed creation-annihilation operators and the conserved quantities such that the hamiltonian and the momentum can be written in terms of them.

We have[9]

$$H = \int d^3X \, \left(\dot{\phi}^\dagger \dot{\phi} + \nabla\phi^\dagger \cdot \nabla\phi + m^2 \phi^\dagger \phi \right) . \tag{4.233}$$

Now, using Eqs. (4.224,4.225), we find

$$H = \int d^3p \, \omega_p \, \frac{1}{2} \left(a^\dagger(p) a(p) + a(p) a^\dagger(p) + b^\dagger(p) b(p) + b(p) b^\dagger(p) \right) , \tag{4.234}$$

or

$$: H := \int d^3p \, \omega_p \left(a^\dagger(p) a(p) + b^\dagger(p) b(p) \right) . \tag{4.235}$$

For the momentum we find an analogous expression:

$$: P^i := \int d^3p \, p^i \left(a^\dagger(p) a(p) + b^\dagger(p) b(p) \right) . \tag{4.236}$$

[9]We can find this relation taking the hamiltonian as a sum of the two hamiltonians of the real fields ϕ_1 and ϕ_2 and then rotate with (4.222,4.223).

Fock space

Since we have two kinds of creation-annihilation operators we have

$$a(p)|0\rangle = b(p)|0\rangle = 0\,, \tag{4.237}$$

and then we have, for instance, one-particle states of kind a, $a^\dagger(p)|0\rangle$, and b, $b^\dagger(p)|0\rangle$, with definite energy and momentum. Let us see:

$$:H:\, a^\dagger(p)|0\rangle = \int d^3p'\,\omega_{p'}\left(a^\dagger(p')a(p') + b^\dagger(p')b(p')\right) a^\dagger(p)|0\rangle = \omega_p\, a^\dagger(p)|0\rangle\,, \tag{4.238}$$

since a and b operators commute. Then, we conclude that $a^\dagger(p)|0\rangle$ is an eigenstate of the hamiltonian with energy ω_p. However, we also have

$$:H:\, b^\dagger(p)|0\rangle = \int d^3p'\,\omega_{p'}\left(a^\dagger(p')a(p') + b^\dagger(p')b(p')\right) b^\dagger(p)|0\rangle = \omega_p\, b^\dagger(p)|0\rangle\,. \tag{4.239}$$

Therefore, also $b^\dagger(p)|0\rangle$ is an eigenstate of the hamiltonian with the same energy ω_p.

The same is true for the momentum. We have:

$$:P^i:\, a^\dagger(p)|0\rangle = p^i\, a^\dagger(p)|0\rangle\,, \tag{4.240}$$

$$:P^i:\, b^\dagger(p)|0\rangle = p^i\, b^\dagger(p)|0\rangle\,. \tag{4.241}$$

Then the states $a^\dagger(p)|0\rangle$ and $b^\dagger(p)|0\rangle$ have the same energy and momentum. They are degenerate with respect to these quantum numbers.

However, note that in the complex-field case there is another conserved quantity, which is the charge (that will be interpreted as the actual electric charge once we will introduce electromagnetic interactions).

From the Nöther's theorem we have

$$Q = i\int d^3X\,\phi^\dagger\overleftrightarrow{\partial_0}\phi\,, \tag{4.242}$$

and substituting the fields in terms of creation-annihilation operators and performing the integrations, we find

$$:Q:= \int d^3p\,\left(a^\dagger(p)a(p) - b^\dagger(p)b(p)\right)\,. \tag{4.243}$$

Therefore, this operator (that commutes with the hamiltonian) resolves the degeneracy, distinguishing between particles of type a and particles of type b.

In fact, now, we have

$$:Q:\, a^\dagger(p)|0\rangle = a^\dagger(p)|0\rangle\,, \tag{4.244}$$

$$:Q:\, b^\dagger(p)|0\rangle = -b^\dagger(p)|0\rangle\,. \tag{4.245}$$

States of type a are eigenstates of the charge with eigenvalue $+1$, while states of type b have opposite charge (-1).

Then, the spectrum is constructed using the following operators: $a^\dagger(p)$ creates a particle state of type a with energy ω_p, momentum $\mathbf{p}$ and charge $+1$, while $a(p)$ annihilates such state; $b^\dagger(p)$ creates a particle state of type b with energy ω_p, momentum $\mathbf{p}$ and charge -1, while $b(p)$ annihilates such state. We say that particles of type b are the "anti-particcles" of the particles of type a. In the real case, we have $a(p) = b(p)$ and therefore the particle is its own anti-particle.

States a and b appear in the theory in a totally symmetric way. Therefore, the names "particle" and "anti-particle" are totally interchangable.

Note that the field operator $\phi(X)$ is a linear combination of annihilation operators $a(p)$ and creation operators $b^\dagger(p)$ (and viceversa for $\phi^\dagger(X)$). This suggests a sort of "equivalence" between the creation of a charge $+1$ and the annihilation of a charge -1.

4.2.3 Locality and causality in QFT

4.3 The Dirac Field (classical field)

Let us proceed with the study of the finite-dimensional representations of the Lorentz group. We consider now a field that transforms, under Poincaré transformations, as a spinor in the $(\frac{1}{2},0)\oplus(0,\frac{1}{2})$ representation:

$$\psi(X)\to\psi'(X')=S(\Lambda)\,\psi(X)\,, \tag{4.246}$$

where, as we will se below

$$S(\Lambda)=e^{\frac{1}{8}[\gamma_\mu,\gamma_\nu]\epsilon^{\mu\nu}}\,. \tag{4.247}$$

4.3.1 The Dirac equation

Historically, we can look at the Dirac equation as an attempt to overcome the difficulties emerged by the Klein-Gordon equation. In particular, we refer to the failure of the probabilistic interpretation of the theory, due to the fact that what should be interpreted as a probability density is not positive definite. This is connected to the fact that the time derivative in the KG equation is of second order. Therefore, we look for a covariant equation of the kind

$$i\frac{\partial}{\partial t}\psi(X)=H\psi(X)\,, \tag{4.248}$$

in which, for the covariance, in the hamiltonian the space derivatives have to be as well of the first order:

$$H=\alpha\cdot\mathbf{p}+\beta m\,. \tag{4.249}$$

α and β are four matrices that we will define according to some obvious constraints. Firstly, since the hamiltonian has to the hermitean[10], we have to have

$$(\alpha^i)^\dagger=\alpha^i\,,\quad \beta^\dagger=\beta\,. \tag{4.255}$$

We require that:

1. If $\psi(X)$ is a solution of (4.248), it has to be also a solution of the KG equation, since it has to fulfil the correct relativistic energy-momentum relation ($E^2=p^2+m^2$).

[10] We just recall what an hermitean operator is. Given a scalar product defined on the Hilbert space under study, (ψ,ϕ), we say that the operator $\hat{A}$ is hermitean if

$$(\psi,\hat{A}\phi)=(\psi\hat{A},\phi)=(\phi,\hat{A}\psi)^*\,. \tag{4.250}$$

Let us consider, for instance, the operator momentum $\hat{p}$, in x representation, acting on L_2:

$$\hat{p}=-i\frac{d}{dx}\,. \tag{4.251}$$

We have

$$(\psi,\phi)=\int_{-\infty}^{\infty}\psi^*(x)\phi(x)\,dx \tag{4.252}$$

and, then

$$(\psi,\hat{p}\phi) = \int_{-\infty}^{\infty}\psi^*(x)\left(-i\frac{d}{dx}\phi(x)\right)dx=\int_{-\infty}^{\infty}-i\frac{d}{dx}\left(\psi^*(x)\phi(x)\right)dx-\int_{-\infty}^{\infty}\left(-i\frac{d}{dx}\psi^*(x)\right)\phi(x)\,dx\,, \tag{4.253}$$

$$= \left\{\int_{-\infty}^{\infty}\phi^*(x)\left(-i\frac{d}{dx}\psi(x)\right)dx\right\}^*=(\phi,\hat{A}\psi)^*\,, \tag{4.254}$$

where we used the fact that, if the fields go to zero rapidly at infinity, the first integral of the next-to-the last line vanishes. The same happens with the hamiltonian $H=-i\alpha\cdot\nabla+\beta m$.

2. The equation should be relativistically covariant.
3. The equation has to give rise to a conserved current, j^μ, that has to transform as a four-vector and such that j^0 is positive definite.

Using (4.249) and the correspondence principle we find the following form for the equation:

$$i\frac{\partial}{\partial t}\psi(X) = (-i\alpha\cdot\nabla + \beta\, m)\psi(X)\,, \tag{4.256}$$

4.3.2 α^i and β matrices

Applying twice the operator $i\frac{\partial}{\partial t}$ we should recover the KG equation. We have

$$-\frac{\partial^2}{\partial t^2}\psi(X) = (-i\alpha\cdot\nabla + \beta\, m)^2\psi(X)\,, \tag{4.257}$$

where, in components

$$\begin{aligned}(-i\alpha\cdot\nabla + \beta\, m)^2 &= -\alpha^i\alpha^j\partial_i\partial_j - im(\alpha^i\beta + \beta\alpha^i)\partial_i + \beta^2 m^2\,, && (4.258)\\ &= -\frac{1}{2}(\alpha^i\alpha^j + \alpha^j\alpha^i)\partial_i\partial_j - im(\alpha^i\beta + \beta\alpha^i)\partial_i + \beta^2 m^2\,, && (4.259)\end{aligned}$$

since $\partial_i\partial_j$ is totally symmetric in the exchange $i \leftrightarrow j$ and therefore only the symmetric part of $\alpha^i\alpha^j$ survives in the sum.

The Klein-Gordon equation is given by

$$\frac{\partial^2}{\partial t^2}\psi = (\nabla^2 - m^2)\psi\,. \tag{4.260}$$

In order for the Eq. (4.257) to reproduce Eq. (4.260), we must have

$$\begin{aligned}\beta^2 &= 1\,, && (4.261)\\ \alpha^i\beta + \beta\alpha^i &= 0\,, && (4.262)\\ \frac{1}{2}(\alpha^i\alpha^j + \alpha^j\alpha^i) &= \delta^{ij}\,. && (4.263)\end{aligned}$$

Relations (4.261), (4.262), (4.263) can be written in a more compact way as follows:

$$\begin{aligned}&[\alpha^i, \alpha^j]_+ = 2\delta^{ij}\,, && (4.264)\\ &[\alpha^i, \beta]_+ = 0\,, && (4.265)\\ &\beta^2 = 1\,. && (4.266)\end{aligned}$$

These relations imply the following properties for α^i and β:
First of all, from Eq. (4.264) it follows that, also for α^i we have

$$(\alpha^i)^2 = 1\,. \tag{4.267}$$

Therefore, α^i and β have real eigenvalues (because they are hermitean) and they have to be ± 1.

Another property is that $\mathrm{tr}\alpha^i = \mathrm{tr}\beta = 0$. In fact, using Eqs. (4.266,4.265) we have:

$$\alpha^i = \beta^2\alpha^i = -\beta\alpha^i\beta \tag{4.268}$$

and, for the cyclicity of the trace

$$\mathrm{tr}\,\alpha^i = \mathrm{tr}\,(\beta^2\alpha^i) = -\mathrm{tr}\,(\beta\alpha^i\beta) = -\mathrm{tr}\,(\beta^2\alpha^i) = -\mathrm{tr}\,\alpha^i\,, \tag{4.269}$$

from which

$$\mathrm{tr}\alpha^i = 0\,. \tag{4.270}$$

The same happens for β.

Since the trace is zero and the eigenvalues are ± 1, α^i and β should have even dimensionality. They cannot be matrices 2×2, since we cannot accomodate, in that space, four anticommuting matrices. This can be done, instead, using 4×4 matrices.

A possible representation for α^i and β is the so-called Dirac representation:

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,, \quad \alpha^i = \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix}\,, \tag{4.271}$$

where σ^i are the Pauli matrices (2×2), generators of the $SU(2)$ group:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\,, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,, \tag{4.272}$$

such that

$$[\sigma^i, \sigma^j] = 2i\epsilon^{ijk}\sigma^k\,, \tag{4.273}$$

$$[\sigma^i, \sigma^j]_+ = 2\delta^{ij}\,. \tag{4.274}$$

It is simple to check, by direct inspection, that the matrices in Eq. (4.271) satisfy Eqs. (4.264,4.265,4.266,4.267).

4.3.3 Covariance of the Dirac equation

The equation in (4.256) is not written in a manifestly covariant form. We introduce the following matrices (the Dirac matrices)

$$\gamma^0 = \beta\,, \quad \gamma^i = \beta\alpha^i\,, \tag{4.275}$$

such that we can define $\gamma^\mu = (\gamma^0, \gamma^1, \gamma^2, \gamma^3)$ and write Eq. (4.256) in the following form[11]

$$(i\gamma^\mu\partial_\mu - m)\psi(X) = 0\,. \tag{4.276}$$

Using the "slash" notation for a four-vector, $\not{a} = \gamma_\mu a^\mu$, we can also write

$$(i\not{\partial} - m)\psi(X) = 0\,. \tag{4.277}$$

Consistently with the representation (4.271), we have

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}\,. \tag{4.278}$$

Moreover, relations (4.264,4.265,4.266,4.267) can be summarized by the (Clifford) algebra

$$[\gamma_\mu, \gamma_\nu]_+ = 2\eta_{\mu\nu}\,, \tag{4.279}$$

with

$$(\gamma^0)^2 = 1\,, \quad (\gamma^i)^2 = -1\,, \quad (\gamma^0)^\dagger = \gamma^0\,, \quad (\gamma^i)^\dagger = -\gamma^i\,, \quad (\gamma^0)^\dagger = \left(\gamma^0\right)^{-1}\,, \quad (\gamma^i)^\dagger = \left(\gamma^i\right)^{-1}\,. \tag{4.280}$$

So γ^0 is hermitian and unitary, while γ^i are anti-hermitian and unitary.

[11] We multiply Eq. (4.256) on the l.h.s. by β and we use the definition of the gamma's.

We now want to find the representation $S(\Lambda)$ of the Lorentz group such that, if in the inertial frame S our system is described by Eq. (4.277), in the inertial frame S' it will be described by

$$(i\,\partial\!\!\!/' - m)\psi'(X') = 0\,, \tag{4.281}$$

where

$$\partial\!\!\!/' = \gamma^\mu \partial'_\mu \tag{4.282}$$

and

$$\psi'(X') = S(\Lambda)\psi(X)\,. \tag{4.283}$$

We note that in Eq. (4.282) the γ^μ is the same as in Eq. (4.277). In fact, the representation of the gamma matrices can change by a unitary transformation[12], that, however, does not affect the physical description of our system. We can then decide to use the same representation in the two inertial frames and use the same gammas.

$S(\Lambda)$ has to be a representation of the Lorentz group and therefore it must fulfil the following relations:

$$S(\Lambda_1)\,S(\Lambda_2) = S(\Lambda_1\Lambda_2)\,, \tag{4.289}$$

for any Λ_1 and Λ_2 Lorents transformations, and

$$S^{-1}(\Lambda) = S(\Lambda^{-1})\,, \tag{4.290}$$

since $S(\Lambda)\,S(\Lambda^{-1}) = S(\Lambda\Lambda^{-1}) = 1 = S(\Lambda)\,S^{-1}(\Lambda)$.

We have

$$0 = (i\gamma^\mu\partial_\mu - m)\psi(X) = (i\gamma^\mu\partial_\mu - m)S^{-1}(\Lambda)\psi'(X')\,. \tag{4.291}$$

Moreover

$$\partial_\mu = \frac{\partial}{\partial X^\mu} = \frac{\partial}{\partial X'^\nu}\frac{\partial X'^\nu}{\partial X^\mu} = \frac{\partial}{\partial X'^\nu}\Lambda^\nu_\mu\,. \tag{4.292}$$

Multiplying Eq. (4.291) by $S(\Lambda)$ on the left, and using (4.292), we find

$$\begin{aligned} 0 &= S(\Lambda)(i\gamma^\mu\partial_\mu - m)S^{-1}(\Lambda)\psi'(X')\,, && (4.293)\\ &= S(\Lambda)(i\gamma^\mu\partial'_\nu\Lambda^\nu_\mu - m)S^{-1}(\Lambda)\psi'(X')\,, && (4.294)\\ &= (iS(\Lambda)\Lambda^\nu_\mu\gamma^\mu S^{-1}(\Lambda)\partial'_\nu - m)\psi'(X')\,. && (4.295)\end{aligned}$$

[12] Moving from the equation in the inertial frame S to the inertial frame S' we should consider

$$(i\tilde{\gamma}^\mu\partial'_\mu - m)\psi'(X') = 0\,, \tag{4.284}$$

where the $\tilde{\gamma}^\mu$ have to satisfy the Clifford algebra. This means that the $\tilde{\gamma}^\mu$ are related to the γ^μ by a similarity transformation:

$$\tilde{\gamma}^\mu = S\gamma^\mu S^{-1}\,, \tag{4.285}$$

such that we have

$$[\tilde{\gamma}^\mu, \tilde{\gamma}^\nu]_+ = S\gamma^\mu S^{-1}S\gamma^\nu S^{-1} + S\gamma^\nu S^{-1}S\gamma^\mu S^{-1} = S[\gamma^\mu,\gamma^\nu]_+S^{-1} = S\,2\eta^{\mu\nu}S^{-1} = 2\eta^{\mu\nu}\,. \tag{4.286}$$

If we want to preserve also the hermiticity or anti-hermiticity and the unitarity of the gamma's, then we can choose the similarity transformation to be unitary:

$$S = U\,, \quad \text{such that} \quad U^\dagger U = 1\,. \tag{4.287}$$

In so doing, we can choose a representation for the gamma's or another representation, without affecting the physics:

$$0 = (i\gamma^\mu\partial_\mu - m)\psi(X) = (i\tilde{\gamma}^\mu\partial_\mu - m)\,U\psi(X)\,, \tag{4.288}$$

but $U\psi(X)$ represents the same physics than $\psi(X)$.

In the end, the two inetrial frames "can agree" on a given representation and use the same gamma's.

In order to reproduce Eq. (4.282) we have to impose

$$S(\Lambda)\Lambda^{\nu}_{\mu}\gamma^{\mu}S^{-1}(\Lambda) = \gamma^{\nu}\,, \tag{4.296}$$

and therefore, multiplying on the l.h.s. by $S^{-1}(\Lambda)$ and on the r.h.s. by $S(\Lambda)$

$$S^{-1}(\Lambda)\gamma^{\nu}S(\Lambda) = \Lambda^{\nu}_{\mu}\gamma^{\mu}\,. \tag{4.297}$$

Let us find an explicit form for $S(\Lambda)$. If we consider an infinitesimal transformation

$$\Lambda^{\mu}_{\nu} \simeq \delta^{\mu}_{\nu} + \epsilon^{\mu}_{\nu}\,, \tag{4.298}$$

where $\epsilon_{\mu\nu} = -\epsilon_{\nu\mu}$, we have

$$S(\Lambda) \simeq 1 - \frac{i}{4}\sigma_{\mu\nu}\epsilon^{\mu\nu}\,, \tag{4.299}$$

$$S^{-1}(\Lambda) \simeq 1 + \frac{i}{4}\sigma_{\mu\nu}\epsilon^{\mu\nu}\,. \tag{4.300}$$

Substituting in Eq. (4.297), we find (at first order)

$$\left(1 + \frac{i}{4}\sigma_{\mu\nu}\epsilon^{\mu\nu}\right)\gamma_{\rho}\left(1 - \frac{i}{4}\sigma_{\alpha\beta}\epsilon^{\alpha\beta}\right) = \gamma_{\rho} + \epsilon_{\rho\alpha}\gamma^{\alpha} \tag{4.301}$$

and neglecting higher order terms, we find

$$\frac{i}{4}[\sigma_{\mu\nu}, \gamma_{\rho}]\epsilon^{\mu\nu} = \gamma_{\nu}\eta_{\rho\mu}\epsilon^{\mu\nu} = \frac{1}{2}\left(\gamma_{\nu}\eta_{\rho\mu} - \gamma_{\mu}\eta_{\rho\nu}\right)\epsilon^{\mu\nu}\,, \tag{4.302}$$

where, due to the fact that $\epsilon^{\mu\nu}$ is anti-symmetric, we anti-symmetrized the tensor $\gamma_{\nu}\eta_{\rho\mu}$. In the end, we have

$$[\sigma_{\mu\nu}, \gamma_{\rho}] = -2i(\gamma_{\nu}\eta_{\rho\mu} - \gamma_{\mu}\eta_{\rho\nu})\,. \tag{4.303}$$

This relation is satisfied by

$$\sigma_{\mu\nu} = \frac{i}{2}[\gamma_{\mu}, \gamma_{\nu}]\,, \tag{4.304}$$

as can be checked by direct inspection:

$$\sigma_{\mu\nu} = \frac{i}{2}(\gamma_{\mu}\gamma_{\nu} - \gamma_{\nu}\gamma_{\mu}) = \frac{i}{2}(2\eta_{\mu\nu} - 2\gamma_{\nu}\gamma_{\mu})\,, \tag{4.305}$$

$$[\sigma_{\mu\nu}, \gamma_{\rho}] = \frac{i}{2}[(2\eta_{\mu\nu} - 2\gamma_{\nu}\gamma_{\mu}), \gamma_{\rho}] = -i[\gamma_{\nu}\gamma_{\mu}, \gamma_{\rho}] = -i\left(\gamma_{\nu}\gamma_{\mu}\gamma_{\rho} - \gamma_{\rho}\gamma_{\nu}\gamma_{\mu}\right)\,, \tag{4.306}$$

$$= -i\left(\gamma_{\nu}\gamma_{\mu}\gamma_{\rho} + \gamma_{\nu}\gamma_{\rho}\gamma_{\mu} - 2\eta_{\rho\nu}\gamma_{\mu}\right) = -i\left(\gamma_{\nu}2\eta_{\rho\mu} - 2\eta_{\rho\nu}\gamma_{\mu}\right)\,, \tag{4.307}$$

$$= -2i(\gamma_{\nu}\eta_{\rho\mu} - \gamma_{\mu}\eta_{\rho\nu})\,. \tag{4.308}$$

Therefore

$$S(\Lambda) \simeq 1 + \frac{1}{8}[\gamma_{\mu}, \gamma_{\nu}]\epsilon^{\mu\nu}\,. \tag{4.309}$$

Exponentiating, we get

$$S(\Lambda) = e^{\frac{1}{8}[\gamma_{\mu}, \gamma_{\nu}]\epsilon^{\mu\nu}}\,. \tag{4.310}$$

$\sigma_{\mu\nu}$ are the generators of the Lorentz group in this representation. Using the Dirac form of the gamma matrices, we find explicitely

$$\sigma_{00} = \sigma_{ii} = 0\,, \tag{4.311}$$

$$\sigma_{0i} = -\sigma_{i0} = \frac{i}{2}[\gamma_0, \gamma_i] = -i\begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix}\,, \tag{4.312}$$

$$\sigma_{ij} = \frac{i}{2}[\gamma_i, \gamma_j] = \epsilon_{ijk}\begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix} \tag{4.313}$$

and we see, once more, that the σ_{ij} are the generators of the rotations and are hermitian, while σ_{0i} are the generators of the boosts and are anti-hermitian.

4.3.4 Unitarity and Dirac adjoint

The operator $S(\Lambda)$ is not unitary and this is due to the fact that the Lorentz group is not compact and therefore we cannot find finite-dimensional unitary representations. In fact we have

$$\sigma^\dagger_{\mu\nu} = -\frac{i}{2}(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu)^\dagger = ... = \frac{i}{2}(\gamma^\dagger_\mu\gamma^\dagger_\nu - \gamma^\dagger_\nu\gamma^\dagger_\mu) = \frac{i}{2}[\gamma^\dagger_\mu, \gamma^\dagger_\nu] \neq \sigma_{\mu\nu}\,, \tag{4.314}$$

since $\gamma^\dagger_0 = \gamma_0$ but $\gamma^\dagger_i = -\gamma_i$. This implies that

$$S^\dagger(\Lambda) = e^{\frac{i}{4}\sigma^\dagger_{\mu\nu}\epsilon^{\mu\nu}} \neq e^{\frac{i}{4}\sigma_{\mu\nu}\epsilon^{\mu\nu}} = S^{-1}(\Lambda)\,. \tag{4.315}$$

However, we can prove (by direct inspection) that

$$\gamma_0\sigma^\dagger_{\mu\nu}\gamma_0 = \sigma_{\mu\nu} \tag{4.316}$$

and therefore

$$\gamma_0 S^\dagger(\Lambda)\gamma_0 = S^{-1}(\Lambda)\,. \tag{4.317}$$

$S(\Lambda)$ is not unitary but is "unitary with respect to the metric γ_0". A consequence of this behaviour is that a bilinear in the fields like $\psi^\dagger\psi$ is not a scalar under Lorentz transformations, but it transforms as the temporal component of a four-vector. If we want to construct a scalar (and this is important because then we would like to find the lagrangian density for the Dirac field and it must be a scalar) we have to consider the so-called Dirac adjoint:

$$\overline{\psi} = \psi^\dagger\gamma_0\,. \tag{4.318}$$

With the Dirac adjoint of ψ, we can construct a scalar: $\overline{\psi}\psi$. In fact, under a Lorentz transformation we have:

$$\overline{\psi}'\psi' = \psi'^\dagger\gamma_0\psi' = (S(\Lambda)\psi)^\dagger\gamma_0 S(\Lambda)\psi = \psi^\dagger S^\dagger(\Lambda)\gamma_0 S(\Lambda)\psi = \overline{\psi}S^{-1}(\Lambda)S(\Lambda)\psi = \overline{\psi}\psi\,. \tag{4.319}$$

We have still to verify that $S(\Lambda)$ satisfy Eq. (4.289), and Eq. (4.290) follows directly. Let us consider the first Lorentz transformation Λ_1. We have

$$S^{-1}(\Lambda_1)\gamma^\mu S(\Lambda_1) = (\Lambda_1)^\mu_\nu\gamma^\nu\,. \tag{4.320}$$

Multiplying on the l.h.s. by $(\Lambda_1^{-1})^\rho_\mu$, we find

$$(\Lambda_1^{-1})^\rho_\mu S^{-1}(\Lambda_1)\gamma^\sigma S(\Lambda_1) = (\Lambda_1^{-1})^\rho_\mu(\Lambda_1)^\mu_\nu\gamma^\nu = \delta^\rho_\nu\gamma^\nu = \gamma^\rho\,. \tag{4.321}$$

Let us consider now the second transformation, Λ_2. We have

$$S^{-1}(\Lambda_2)\gamma^\rho S(\Lambda_2) = (\Lambda_2)^\rho_\alpha\gamma^\alpha\,. \tag{4.322}$$

We now substitute in the expression above the γ^ρ with the analogous expression in (4.321):

$$\begin{aligned}(\Lambda_2)^\rho_\alpha\gamma^\alpha &= S^{-1}(\Lambda_2)(\Lambda_1^{-1})^\rho_\mu S^{-1}(\Lambda_1)\gamma^\mu S(\Lambda_1)S(\Lambda_2)\,, && (4.323)\\ &= (\Lambda_1^{-1})^\rho_\mu S^{-1}(\Lambda_2)S^{-1}(\Lambda_1)\gamma^\mu S(\Lambda_1)S(\Lambda_2)\,. && (4.324)\end{aligned}$$

Multiplying on the left by $(\Lambda_1)^\sigma_\rho$ we find

$$S^{-1}(\Lambda_2)S^{-1}(\Lambda_1)\gamma^\mu S(\Lambda_1)S(\Lambda_2) = (\Lambda_1)^\sigma_\rho(\Lambda_2)^\rho_\alpha\gamma^\alpha = (\Lambda_1\Lambda_2)^\sigma_\alpha\gamma^\alpha\,. \tag{4.325}$$

Since $(\Lambda_1\Lambda_2)$ is indeed a Lorentz transformation, we can also write

$$S^{-1}(\Lambda_1\Lambda_2)\gamma^\sigma S(\Lambda_1\Lambda_2) = (\Lambda_1\Lambda_2)^\sigma_\alpha\gamma^\alpha\,. \tag{4.326}$$

Therefore, it follows the statement:

$$S(\Lambda_1)S(\Lambda_2) = S(\Lambda_1\Lambda_2)\,. \tag{4.327}$$

$S(\Lambda)$ is indeed a representation of the Lorentz group.

Example

As an example, we write the explicit form of $S(\Lambda)$, acting on the spinorial field $\psi(X)$, when Λ is a boost. Let us choose for simplicity a boost in the x direction. We will have

$$X^\mu \to X'^\mu = \Lambda^\mu_\nu X^\nu\,, \tag{4.328}$$

where, in matrix form

$$\Lambda^\mu_\nu = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.329}$$

or, using the hyperbolic parametrization

$$\beta = \frac{v}{c} = \tanh(\theta)\,, \quad \gamma = \frac{1}{\sqrt{1-\beta^2}} = \cosh(\theta)\,, \tag{4.330}$$

$$\Lambda^\mu_\nu = \begin{pmatrix} \cosh(\theta) & -\sinh(\theta) & 0 & 0 \\ -\sinh(\theta) & \cosh(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.331}$$

For the infinitesimal transformation ($\theta \ll 1$)

$$\Lambda^\mu_\nu \simeq \delta^\mu_\nu + \epsilon^\mu_\nu = \begin{pmatrix} 1 & -\theta & 0 & 0 \\ -\theta & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -\theta & 0 & 0 \\ -\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\,. \tag{4.332}$$

We have

$$\epsilon_{\mu\nu} = \eta_{\mu\rho}\epsilon^\rho_\nu \tag{4.333}$$

and therefore we find

$$\epsilon_{10} = -\epsilon_{01} = \theta \tag{4.334}$$

and the other components are zero. Then

$$S(\Lambda) \simeq 1 - \frac{i}{4}\epsilon_{\mu\nu}\sigma^{\mu\nu} = 1 - \frac{i}{4}\left(\epsilon_{10}\sigma^{10} + \epsilon_{01}\sigma^{01}\right) = 1 - \frac{i}{4}\left(2\epsilon_{10}\sigma^{10}\right) = 1 - \frac{i}{2}\theta\sigma^{10}\,, \tag{4.335}$$

where

$$\sigma^{10} = \frac{i}{2}[\gamma^1,\gamma^0] = \frac{i}{2}\left(\gamma^1\gamma^0 - \gamma^0\gamma^1\right) = -i\gamma^0\gamma^1 = -i\alpha^1 = -i\begin{pmatrix} 0 & \sigma^1 \\ \sigma^1 & 0 \end{pmatrix} = -i\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}\,. \tag{4.336}$$

Exponentiating, we find

$$S(\Lambda) = e^{-\frac{\theta}{2}\alpha^1} = 1 - \frac{\theta}{2}\alpha^1 + \frac{1}{2}\left(-\frac{\theta}{2}\alpha^1\right)^2 + \frac{1}{6}\left(-\frac{\theta}{2}\alpha^1\right)^3 + \ldots\,, \tag{4.337}$$

$$= \left|\text{recalling that } (\alpha^1)^2 = 1\right|$$

$$= \sum_{k=0}^{\infty}\frac{1}{(2k)!}\left(-\frac{\theta}{2}\right)^{2k} + \alpha^1\sum_{k=0}^{\infty}\frac{1}{(2k+1)!}\left(-\frac{\theta}{2}\right)^{2k+1}\,, \tag{4.338}$$

$$= \cosh\left(\frac{\theta}{2}\right) - \alpha^1 \sinh\left(\frac{\theta}{2}\right) , \tag{4.339}$$

$$= \begin{pmatrix} \cosh\left(\frac{\theta}{2}\right) & -\sigma^1 \sinh\left(\frac{\theta}{2}\right) \\ -\sigma^1 \sinh\left(\frac{\theta}{2}\right) & \cosh\left(\frac{\theta}{2}\right) \end{pmatrix} . \tag{4.340}$$

Since

$$\cosh^2\left(\frac{\theta}{2}\right) = \frac{\cosh(\theta)+1}{2} , \quad \sinh^2\left(\frac{\theta}{2}\right) = \frac{\cosh(\theta)-1}{2} , \tag{4.341}$$

recalling the hyperbolic parametrization in terms of β and γ we find

$$S(\Lambda) = \sqrt{\frac{\gamma+1}{2}} \begin{pmatrix} 1 & -\frac{\beta\gamma}{\gamma+1}\sigma^1 \\ -\frac{\beta\gamma}{\gamma+1}\sigma^1 & 1 \end{pmatrix} . \tag{4.342}$$

This form can be generalized to a boost in the $\mathbf{n}$ direction as

$$S(\Lambda) = \sqrt{\frac{\gamma+1}{2}} \begin{pmatrix} 1 & -\frac{\beta\gamma}{\gamma+1}\sigma\cdot\mathbf{n} \\ -\frac{\beta\gamma}{\gamma+1}\sigma\cdot\mathbf{n} & 1 \end{pmatrix} . \tag{4.343}$$

4.3.5 Probability density

One of the main problems in the "first-quantization" interpretation of the Klein-Gordon equation was the failure of the probabilistic interpretation due to the non-positivity of the probability density. This can be linked to the fact that the KG equation is second order in the time derivative. Let us see what happens in the case of the Dirac equation.

Firstly, let us write the equation for the daggered field, $\psi^\dagger$. We have

$$-i\frac{\partial}{\partial t}\psi^\dagger = i(\nabla\psi^\dagger)\cdot\alpha + m\psi^\dagger\beta , \tag{4.344}$$

where we used the fact that $(\alpha^i)^\dagger = \alpha^i$ and $\beta^\dagger = \beta$. Multiplying the Direc equation for ψ by $\psi^\dagger$ on the left and subtracting Eq. (4.344) multiplied by ψ on the right, we find

$$i\psi^\dagger\left(\frac{\partial}{\partial t}\psi\right) + i\left(\frac{\partial}{\partial t}\psi^\dagger\right)\psi = \psi^\dagger(i\alpha\cdot\nabla\psi + \beta m\psi) - (i\nabla\psi^\dagger\cdot\alpha + m\psi^\dagger\beta)\psi , \tag{4.345}$$

or

$$i\frac{\partial}{\partial t}(\psi^\dagger\psi) = -i\psi^\dagger\,\alpha\cdot\nabla\psi - i(\nabla\psi^\dagger)\cdot\alpha\,\psi = -i\nabla\cdot(\psi^\dagger\alpha\psi) . \tag{4.346}$$

If we define the vector

$$j^\mu = (\psi^\dagger\psi, \psi^\dagger\alpha\psi) = \overline{\psi}\gamma^\mu\psi , \tag{4.347}$$

Eq. (4.369) becomes

$$\partial_\mu j^\mu = 0 . \tag{4.348}$$

Eq. (4.348) implies the conservation of the "charge"

$$Q = \int d^3X\,\psi^\dagger\psi , \tag{4.349}$$

and since $\psi^\dagger\psi$ is a positive definite quantity, it can be interpreted as a probability density (and then Q is the total probability to find the particle in all the space, therefore $Q = 1$).

The vector j^μ defined in Eq. (4.347) transforms indeed as a four-vector under Lorentz transformations. In fact

$$j'^\mu(X') = \overline{\psi'}(X')\gamma^\mu\psi'(X') = (\psi'^\dagger(X')\gamma^0)\gamma^\mu\psi'(X') , \tag{4.350}$$

$$= (S(\Lambda)\psi(X))^\dagger\,\gamma^0\gamma^\mu S(\Lambda)\psi(X) = \psi^\dagger S^\dagger(\Lambda)\gamma^0\gamma^\mu S(\Lambda)\psi(X) , \tag{4.351}$$

$$= \overline{\psi}(X)\gamma^0 S^\dagger(\Lambda)\gamma^0\gamma^\mu S(\Lambda)\psi(X) = \overline{\psi}(X)S^{-1}(\Lambda)\gamma^\mu S(\Lambda)\psi(X) , \tag{4.352}$$

$$= \Lambda^\mu_\nu\overline{\psi}(X)\gamma^\nu\psi(X) , \tag{4.353}$$

$$= \Lambda^\mu_\nu j^\nu(X) . \tag{4.354}$$

4.3.6 Lagrangian and Hamiltonian densities

The Dirac field is a spinorial complex field. Then, we will consider ψ and $\overline{\psi}$ as independent fields. While ψ obeys the equation

$$(i \not\partial - m)\psi(X) = 0 \,, \tag{4.355}$$

the equation for the adjoint field can be found taking the dagger of (4.355)

$$-i\partial_\mu \psi^\dagger(X)\,(\gamma^\mu)^\dagger - m\psi^\dagger(X) = 0 \tag{4.356}$$

and multiplying by γ^0 on the r.h.s.

$$-i\partial_\mu \psi^\dagger(X)\,\gamma^0\gamma^0(\gamma^\mu)^\dagger\gamma^0 - m\overline{\psi}(X) = -i\partial_\mu\overline{\psi}(X)\gamma^\mu - m\overline{\psi}(X) = 0 \tag{4.357}$$

or, better

$$\overline{\psi}(X)\,(i \overleftarrow{\not\partial} + m) = 0 \,. \tag{4.358}$$

Using the same approach as in the Klein-Gordon case, we can recover the lagrangian density starting from the Euler-Lagrange equations (4.355,4.358) mulptiplied by the variation of the fields and making in such a way to find the Hamilton Principle

$$0 = \delta S = \int \left\{ \delta\overline{\psi}(i \not\partial - m)\psi + \overline{\psi}\,(i \overleftarrow{\not\partial} + m)\delta\psi \right\} \,, \tag{4.359}$$

$$= \delta \int d^3X\, \overline{\psi}(i \not\partial - m)\psi \,. \tag{4.360}$$

The lagrangian density is therefore

$$\mathcal{L} = \overline{\psi}(i \not\partial - m)\psi \,. \tag{4.361}$$

It is easy to check that the Euler-Lagrange equations of the lagrangian density (4.361) are indeed Eqs. (4.355,4.358). In fact, in order to get the equations for ψ we have

$$0 = \frac{\partial \mathcal{L}}{\partial \overline{\psi}} - \partial_\mu \frac{\partial \mathcal{L}}{\partial \overline{\psi}_{,\mu}} = (i \not\partial - m)\psi(X) \,, \tag{4.362}$$

since $\mathcal{L}$ does not involve derivatives of the field $\overline{\psi}$ and therefore

$$\frac{\partial \mathcal{L}}{\partial \overline{\psi}_{,\mu}} = 0 \,. \tag{4.363}$$

For $\overline{\psi}$ wi have

$$0 = \frac{\partial \mathcal{L}}{\partial \psi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial \psi_{,\mu}} = -m\overline{\psi} - (i\gamma_\mu\partial_\mu\overline{\psi}) \tag{4.364}$$

and therefore Eq. (4.358).

The lagrangian density (4.361) has a problem. It is a singular lagrangian, in the sense that the momentum conjugate to $\overline{\psi}$ is zero:

$$\pi_\psi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\psi^\dagger \,, \tag{4.365}$$

$$\pi_{\psi^\dagger} = \frac{\partial \mathcal{L}}{\partial \dot{\psi}^\dagger} = 0 \,. \tag{4.366}$$

This is due to the fact that $\mathcal{L}$ does not involve derivatives of $\psi^\dagger$ or, which is the same, the canonical momenta do not depend on velocities. The canonical formalism rely on momenta that are the time derivative of the conjugated degree of freedom. In this case, then, in principle we cannot proceed with a Legendre transformation getting the hamiltonian (the energy) of the system. The problem was solved by Dirac himself, that proposed a procedure to arrive to the hamiltonian. This procedure coincides, in this case, with the naive formula

$$\mathcal{H} = \pi_\psi \dot{\psi} - \mathcal{L} \tag{4.367}$$

and, considering the configurations of the field that satisfy Dirac's equation, we get

$$\mathcal{H} = i\psi^\dagger \partial_0 \psi \quad (= \psi^\dagger(-i\alpha \cdot \nabla + \beta m)\psi)\,. \tag{4.368}$$

Contrarly to the KG field, this expression is not positive definite. However, we will see that when we will move to the quantized version of $\mathcal{H}$, as an operator acting on the Fock space, it will be positive definite.

The expression (4.368) can be recovered also using Nöther's theorem.

4.3.7 Conserved quantities

The lagrangian density (4.361) is Poincaré invariant. This imply that, according to Nöther's theorem, there are some quantities that are conserved.

If we consider the non homogeneous part of the Poincaré group (rigid translations), we get the relation

$$\partial_\mu T^\mu_\nu = 0\,, \tag{4.369}$$

where the tensor T^μ_ν is the so-called "energy-momentum" tensor

$$T^\mu_\nu = \frac{\partial \mathcal{L}}{\partial \psi_{,\mu}} \psi_{,\nu} + \frac{\partial \mathcal{L}}{\partial \psi^\dagger_{,\mu}} \psi^\dagger_{,\nu} - \eta^\mu_\nu \mathcal{L} \tag{4.370}$$

and , considering the configurations of the field that satisfy Dirac's equation, we get

$$T^\mu_\nu = i\overline{\psi}\gamma^\mu \psi_{,\nu}\,. \tag{4.371}$$

It is easy to check that the form given in Eq. (4.371) satisfies Eq. (4.369).

The conserved four-vector is

$$P_\nu = \int d^3X\, T^0_\nu = \int d^3X\, i\psi^\dagger \partial_\nu \psi\,. \tag{4.372}$$

Therefore

$$H = \int d^3X\, T^0_0 = \int d^3X\, i\psi^\dagger \partial_0 \psi\,, \tag{4.373}$$

$$\mathbf{P} = \int d^3X\, T^{0i} = -\int d^3X\, i\psi^\dagger \nabla \psi\,. \tag{4.374}$$

If we consider instead the Lorentz group, we get

$$\partial_\mu \mathcal{M}^\mu_{rho\sigma} = 0\,, \tag{4.375}$$

where

$$\mathcal{M}^\mu_{\rho\sigma} = i\overline{\psi}\gamma^\mu \left(X_\rho \partial_\sigma - X_\sigma \partial_\rho + \frac{1}{4}[\gamma_\rho, \gamma_\sigma] \right) \psi\,. \tag{4.376}$$

The conserved charges are the following 6 charges:

$$M_{\rho\sigma} = \int d^3X\, \mathcal{M}^0_{\rho\sigma}\,. \tag{4.377}$$

The angular momentum is

$$\mathbf{J} = (M^{23}, M^{31}, M^{12}) = \int d^3X\, \psi^\dagger(-i\mathbf{x} \wedge \nabla + \frac{1}{2}\begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix})\, \psi\,, \tag{4.378}$$

where we can recognize an orbital part and a spin part.

Global phase invariance

The lagrangian density (4.361) is invariant under the following transformation

$$\psi(X) \quad \rightarrow \quad \psi'(X) = e^{-i\theta}\psi(X)\,, \tag{4.379}$$
$$\overline{\psi}(X) \quad \rightarrow \quad \overline{\psi'}(X) = e^{i\theta}\overline{\psi}(X)\,, \tag{4.380}$$

where $\theta \in \mathbb{R}$. This is a continuous transformation. The infinitesimal tranformation is

$$\delta\psi \quad = \quad -i\theta\psi\,, \tag{4.381}$$
$$\delta\overline{\psi} \quad = \quad i\theta\overline{\psi}\,. \tag{4.382}$$

This symmetry gives rise to a four-vector

$$J^\mu = \frac{\partial \mathcal{L}}{\partial \psi_{,\mu}}\delta\psi = \theta\,\overline{\psi}\gamma^\mu \tag{4.383}$$

such that

$$\partial_\mu J^\mu = 0\,. \tag{4.384}$$

Then, we can define the current as in Eq. (4.347) such that the conserved charge is the one in Eq. (4.349). Once we introduce the interaction of the Dirac field with the electromagnetic field the charge will be correctly interpreted in QFT as the electric charge (and not connected with the probability density of a tentative "first quantization" interpretation of the theory).

4.3.8 The matrix γ_5

The matrix

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{4.385}$$

plays un important role in the Clifford algebra of the γ matrices. It has the following properties:

1. γ_5 is hermitian:

$$(\gamma^5)^\dagger \quad = \quad (i\gamma^0\gamma^1\gamma^2\gamma^3)^\dagger = -i(\gamma^3)^\dagger(\gamma^2)^\dagger(\gamma^1)^\dagger(\gamma^0)^\dagger\,, \tag{4.386}$$
$$= \quad i\gamma^3\gamma^2\gamma^1\gamma^0 = i\gamma^0\gamma^1\gamma^2\gamma^3\,, \tag{4.387}$$
$$= \quad \gamma_5\,. \tag{4.388}$$

2. γ_5 anticommutes with all the γ^μ:

$$[\gamma_5, \gamma^0]_+ \quad = \quad i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^0 + i\gamma^0\gamma^0\gamma^1\gamma^2\gamma^3 = -i\gamma^1\gamma^2\gamma^3 + i\gamma^1\gamma^2\gamma^3 = 0\,, \tag{4.389}$$
$$[\gamma_5, \gamma^i]_+ \quad = \quad i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^i + i\gamma^i\gamma^0\gamma^1\gamma^2\gamma^3 = ... = 0\,. \tag{4.390}$$

The representation for γ_5 follows the representation of the γ^μ. In the Pauli representation we have

$$\gamma_5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,. \tag{4.391}$$

We can find a more "covariant" form of γ_5 noting that the expression

$$i\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma \quad \text{with } \mu \neq \nu \neq \rho \neq \sigma \tag{4.392}$$

gives $\pm\gamma_5$. Actually, if $\mu\nu\rho\sigma$ is an even permutation of 0123, we have $+\gamma_5$; if $\mu\nu\rho\sigma$ is an odd permutation of 0123, we have $-\gamma_5$. Using the totally antisymmetric tensor $\epsilon_{\mu\nu\rho\sigma}$ we have

$$\epsilon_{\mu\nu\rho\sigma}(i\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) = \sum_{\text{even}\mathcal{P}(0123)} (+1)(+\gamma_5) + \sum_{\text{odd}\mathcal{P}(0123)} (-1)(-\gamma_5) = 24\,\gamma_5\,. \tag{4.393}$$

Therefore

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \frac{i}{24}\epsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma \,. \tag{4.394}$$

In order to find how γ_5 transforms under Lorentz transformations $S(\Lambda)$, consider that

$$\det\Lambda = \epsilon_{\mu\nu\rho\sigma}\Lambda^\mu_0\Lambda^\nu_0\Lambda^\rho_0\Lambda^\sigma_0 \,, \tag{4.395}$$

from which we can write

$$\epsilon_{\mu\nu\rho\sigma}\Lambda^\mu_\alpha\Lambda^\nu_\beta\Lambda^\rho_\delta\Lambda^\sigma_\gamma = \det\Lambda\,\epsilon_{\alpha\beta\delta\gamma} \,. \tag{4.396}$$

Therefore

$$\begin{aligned}
S^{-1}(\Lambda)\gamma_5 S(\Lambda) &= \frac{i}{24}\epsilon_{\mu\nu\rho\sigma}S^{-1}(\Lambda)\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma S(\Lambda)\,, && (4.397)\\
&= \frac{i}{24}\epsilon_{\mu\nu\rho\sigma}S^{-1}\gamma^\mu S S^{-1}\gamma^\nu S S^{-1}\gamma^\rho S S^{-1}\gamma^\sigma S\,, && (4.398)\\
&= \frac{i}{24}\epsilon_{\mu\nu\rho\sigma}\Lambda^\mu_\alpha\gamma^\alpha\,\Lambda^\nu_\beta\gamma^\beta\,\Lambda^\rho_\delta\gamma^\delta\,\Lambda^\sigma_\gamma\gamma^\gamma\,, && (4.399)\\
&= \frac{i}{24}\epsilon_{\mu\nu\rho\sigma}\Lambda^\mu_\alpha\Lambda^\nu_\beta\Lambda^\rho_\delta\Lambda^\sigma_\gamma\,\gamma^\alpha\gamma^\beta\gamma^\delta\gamma^\gamma\,, && (4.400)\\
&= \det\Lambda\,\frac{i}{24}\epsilon_{\alpha\beta\delta\gamma}\,\gamma^\alpha\gamma^\beta\gamma^\delta\gamma^\gamma\,, && (4.401)\\
&= \det\Lambda\,\gamma_5\,. && (4.402)
\end{aligned}$$

4.3.9 Bilinear covariants

The space of Dirac matrices is a 16-dim space. We can prove that a basis for such space is constituted by the following 16 matrices:

$$\Gamma = \{\mathbb{1}, \gamma^\mu, \gamma_5, \sigma^{\mu\nu}, \gamma^\mu\gamma_5\} \,. \tag{4.403}$$

With this choice, it is very easy to understand the transformation behaviour of bilinears in the fields, like $\overline{\psi}\Gamma\psi$, under Lorentz transformations.

In fact, we already proved that

$$\overline{\psi}\,\mathbb{1}\,\psi = \overline{\psi}\psi \tag{4.404}$$

transforms as a scalar under Lorentz transformations. Moreover,

$$\overline{\psi}\gamma^\mu\psi \tag{4.405}$$

transforms as a four-vector.

For the other possible bilinears we have:

$$\begin{aligned}
\overline{\psi'}(X')\gamma_5\psi'(X') &= \psi^\dagger(X)S^\dagger(\Lambda)\gamma^0\gamma_5 S(\Lambda)\psi(X)\,, && (4.406)\\
&= \overline{\psi}(X)\,S^{-1}(\Lambda)\gamma_5 S(\Lambda)\psi(X)\,, && (4.407)\\
&= \det\Lambda\,\overline{\psi}(X)\gamma_5\psi(X)\,. && (4.408)
\end{aligned}$$

We say that $\overline{\psi}(X)\gamma_5\psi(X)$ is a pseudo-scalar.

$$\begin{aligned}
\overline{\psi'}(X')\gamma^\mu\gamma_5\psi'(X') &= \psi^\dagger(X)S^\dagger(\Lambda)\gamma^0\gamma^\mu\gamma_5 S(\Lambda)\psi(X)\,, && (4.409)\\
&= \overline{\psi}(X)\,S^{-1}(\Lambda)\gamma^\mu S(\Lambda)S^{-1}(\Lambda)\gamma_5 S(\Lambda)\psi(X)\,, && (4.410)\\
&= \det\Lambda\,\Lambda^\mu_\nu\overline{\psi}(X)\gamma^\nu\gamma_5\psi(X)\,. && (4.411)
\end{aligned}$$

We say that $\overline{\psi}(X)\gamma^\mu\gamma_5\psi(X)$ is a pseudo-vector.

Since $\sigma^{\mu\nu} = \frac{i}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\nu)$, we concentrate on

$$\begin{aligned} \overline{\psi'}(X')\gamma^\mu\gamma^\nu\psi'(X') &= \psi^\dagger(X)S^\dagger(\Lambda)\gamma^0\gamma^\mu\gamma^\nu S(\Lambda)\psi(X)\,, & (4.412)\\ &= \overline{\psi}(X)\,S^{-1}(\Lambda)\gamma^\mu S(\Lambda)S^{-1}(\Lambda)\gamma^\nu S(\Lambda)\psi(X)\,, & (4.413)\\ &= \Lambda^\mu_\alpha\Lambda^\nu_\beta\,\overline{\psi}(X)\gamma^\alpha\gamma^\beta\psi(X)\,. & (4.414) \end{aligned}$$

Therefore, $\overline{\psi}(X)\gamma^\mu\gamma^\nu\psi(X)$ transforms as a rank-2 tensor.

4.3.10 Algebra of the γ^μ matrices and γ_5

It is important, for future applications, to introduce some rules for the calculation of traces with γ matrices. We consider the Minkowski space with $4 = 3 + 1$ dimensions. We recall the algebra of the γ's

$$[\gamma_\mu, \gamma_\nu]_+ = 2\eta_{\mu\nu}\,. \tag{4.415}$$

We have

- $\boxed{\gamma_\mu\gamma^\mu = 4\,\mathbb{1}}$ In fact

$$\gamma_\mu\gamma^\mu = (\gamma^0)^2 - (\gamma^1)^2 - (\gamma^2)^2 - (\gamma^2)^2 = 4\,\mathbb{1}\,. \tag{4.416}$$

- $\boxed{\gamma_\mu\gamma^\nu\gamma^\mu = -2\gamma^\nu}$ In fact

$$\gamma_\mu\gamma^\nu\gamma^\mu = \gamma_\mu(-\gamma^\mu\gamma^\nu + 2\eta^{\mu\nu}) = -\gamma_\mu\gamma^\mu\gamma^\nu + 2\gamma^\nu = -4\gamma^\nu + 2\gamma^\nu = -2\gamma^\nu\,. \tag{4.417}$$

- $\boxed{\gamma_\mu\gamma^\lambda\gamma^\nu\gamma^\mu = 4\eta^{\lambda\nu}}$ In fact

$$\begin{aligned} \gamma_\mu\gamma^\lambda\gamma^\nu\gamma^\mu &= \gamma_\mu\gamma^\lambda(-\gamma^\mu\gamma^\nu + 2\eta^{\mu\nu}) = -\gamma_\mu\gamma^\lambda\gamma^\mu\gamma^\nu + 2\gamma^\nu\gamma^\lambda = 2\gamma^\lambda\gamma^\nu + 2\gamma^\nu\gamma^\lambda\,,\\ &= 2[\gamma^\lambda, \gamma^\nu]_+ = 4\eta^{\lambda\nu}\,. & (4.418) \end{aligned}$$

And, saturating with vectors, recalling the "slash" notation $\not a = a_\mu\gamma^\mu$, we have

- $\boxed{\not a\,\not a = a^2}$ In fact

$$\not a\,\not a = a^\mu a^\nu\gamma_\mu\gamma_\nu = a^\mu a^\nu(-\gamma_\mu\gamma_\nu + 2\eta_{\mu\nu}) = -\not a\,\not a + 2a^2\,, \tag{4.419}$$

  and therefore $\not a\,\not a = a^2$.

- $\boxed{\not a\,\not b + \not b\,\not a = 2a\cdot b}$

- $\boxed{\gamma_\mu\,\not a\gamma^\mu = -2\,\not a}$

- $\boxed{\gamma_\mu\,\not a\,\not b\gamma^\mu = 4a\cdot b}$

Concerning the traces of the γ's, we have:

- $\boxed{\mathrm{tr}\gamma^\mu = 0}$

- $\boxed{\mathrm{tr}(\not a\,\not b) = 4\,a\cdot b}$ In fact

$$\mathrm{tr}(\not a\,\not b) = a_\mu b_\nu \mathrm{tr}(\gamma^\mu\gamma^\nu) = a_\mu b_\nu\frac{1}{2}\mathrm{tr}(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu) = a_\mu b_\nu\frac{1}{2}\mathrm{tr}(2\eta^{\mu\nu}\mathbb{1}) = 4\,a\cdot b\,, \tag{4.420}$$

  where we used the cyclicity of the trace $\mathrm{tr}(\gamma^\mu\gamma^\nu) = \mathrm{tr}(\gamma^\nu\gamma^\mu)$ and therefore

$$\mathrm{tr}(\gamma^\mu\gamma^\nu) = \frac{1}{2}(\mathrm{tr}(\gamma^\mu\gamma^\nu) + \mathrm{tr}(\gamma^\nu\gamma^\mu))\,. \tag{4.421}$$

- $\boxed{\mathrm{tr}(\not a\,\not b\,\not c) = 0}$ In fact

$$\begin{aligned}
\mathrm{tr}(\not a\,\not b\,\not c) &= \mathrm{tr}(\not a\,\not b\,\not c\gamma_5\gamma_5) = |\,\text{cyclicity}\,| = \mathrm{tr}(\gamma_5\,\not a\,\not b\,\not c\gamma_5) = |\,\text{anti-commuting the }\ \gamma_5\,|\,, \\
&= -\mathrm{tr}(\gamma_5\gamma_5\,\not a\,\not b\,\not c) = -\mathrm{tr}(\not a\,\not b\,\not c)\,. \qquad (4.422)
\end{aligned}$$

- $\boxed{\mathrm{tr}(\not a\,\not b\,\not c\,\not d) = 4(a\cdot b)(c\cdot d) + 4(a\cdot d)(b\cdot c) - 4(a\cdot c)(b\cdot d)}$ In fact

$$\begin{aligned}
\mathrm{tr}(\not a\,\not b\,\not c\,\not d) &= \mathrm{tr}[(-2\,\not b\,\not a + 2(a\cdot b))\,\not c\,\not d] = 2(a\cdot b)\mathrm{tr}(\not c\,\not d) - \mathrm{tr}(\not b\,\not a\,\not c\,\not d)\,, \\
&= 8(a\cdot b)(c\cdot d) - \mathrm{tr}[\not b(-\,\not c\,\not a + 2a\cdot c)\,\not d] = 8(a\cdot b)(c\cdot d) - 8(a\cdot c)(b\cdot d) \\
&\quad + \mathrm{tr}(\not b\,\not c\,\not a\,\not d)\,, \\
&= 8(a\cdot b)(c\cdot d) - 8(a\cdot c)(b\cdot d) + \mathrm{tr}[\not b\,\not c(-\,\not d\,\not a + 2(a\cdot d)))]\,, \qquad (4.423) \\
&= 8(a\cdot b)(c\cdot d) - 8(a\cdot c)(b\cdot d) + 8(a\cdot d)(b\cdot c) - \mathrm{tr}(\not b\,\not c\,\not d\,\not a)\,, \\
&= 8(a\cdot b)(c\cdot d) - 8(a\cdot c)(b\cdot d) + \mathrm{tr}[\not b\,\not c(-\,\not d\,\not a + 2(a\cdot d)))]\,, \qquad (4.424) \\
&= 8(a\cdot b)(c\cdot d) - 8(a\cdot c)(b\cdot d) + 8(a\cdot d)(b\cdot c) - \mathrm{tr}(\not a\,\not b\,\not c\,\not d)\,, \qquad (4.425)
\end{aligned}$$

from which $\mathrm{tr}(\not a\,\not b\,\not c\,\not d) = 4(a\cdot b)(c\cdot d) + 4(a\cdot d)(b\cdot c) - 4(a\cdot c)(b\cdot d)$.

- In general

$$\mathrm{tr}(\not a_1\,\not a_2....\,\not a_n) = 0\,, \qquad \text{if } n \text{ is odd}\,, \qquad (4.426)$$

$$\begin{aligned}
\mathrm{tr}(\not a_1\,\not a_2....\,\not a_n) &= (a_1\cdot a_2)\mathrm{tr}(\not a_3\,\not a_4....\,\not a_n) - (a_1\cdot a_3)\mathrm{tr}(\not a_2\,\not a_4....\,\not a_n) + ... \\
&\quad + (a_1\cdot a_n)\mathrm{tr}(\not a_2\,\not a_3....\,\not a_{n-1})\,, \qquad \text{if } n \text{ is even}\,. \qquad (4.427)
\end{aligned}$$

And, including γ_5:

- $\boxed{\mathrm{tr}\gamma_5 = 0}$
- $\boxed{\mathrm{tr}(\gamma_5\,\not a) = 0}$
- $\boxed{\mathrm{tr}(\gamma_5\,\not a\,\not b) = 0}$
- $\boxed{\mathrm{tr}(\gamma_5\,\not a\,\not b\,\not c) = 0}$
- $\boxed{\mathrm{tr}(\gamma_5\,\not a\,\not b\,\not c\,\not d) = 4i\epsilon_{\mu\nu\rho\sigma}a^\mu b^\nu c^\rho d^\sigma}$
- $\boxed{\mathrm{tr}(\gamma_5\,\not a_1\,\not a_2....\,\not a_n) = 0}$ if n is odd
- $\boxed{\mathrm{tr}(\gamma_5\,\not a_1\,\not a_2....\,\not a_n) \neq 0}$ if n is even, $n > 4$.

4.3.11 Plane wave solutions

In this section we consider the plane wave solutions of the Dirac equation. We assume

$$\psi(X) = u(P)e^{-iP_\mu X^\mu}\,, \qquad (4.428)$$

where $u(p)$ is a spinor. Substituting into the Dirac equation, we get

$$(i\gamma^\mu\partial_\mu - m)u(P)e^{-iP_\mu X^\mu} = (i\gamma^\mu(-iP_\mu) - m)u(P)e^{-iP_\mu X^\mu} = (\not P - m)u(P)e^{-iP_\mu X^\mu} = 0\,. \qquad (4.429)$$

This leads to the following equation for the spinor $u(P)$:

$$(\not P - m)u(P) = 0\,, \qquad (4.430)$$

or, in matrix form, using a two-component spinor

$$\begin{pmatrix} P^0 - m & -\sigma \cdot \mathbf{p} \\ \sigma \cdot \mathbf{p} & -P^0 - m \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} . \tag{4.431}$$

The system has non-trivial solutions only if

$$\det \begin{pmatrix} P^0 - m & -\sigma \cdot \mathbf{p} \\ \sigma \cdot \mathbf{p} & -P^0 - m \end{pmatrix} = m^2 - (P^0)^2 + (\sigma \cdot \mathbf{p})^2 = m^2 - (P^0)^2 + p^2 = 0 , \tag{4.432}$$

where we used the fact that

$$(\sigma \cdot \mathbf{p})^2 = \sigma_i \sigma_j p^i p^j = \frac{1}{2}([\sigma_i, \sigma_j] + [\sigma_i, \sigma_j]_+) p^i p^j = (\delta_{ij} + \epsilon_{ijk}\sigma_k) p^i p^j = p^2 . \tag{4.433}$$

Therefore, as in the Klein-Gordon case, we find again two kind of solutions

$$P^0 = \pm\sqrt{p^2 + m^2} = \pm\omega_p . \tag{4.434}$$

We have two different plane waves, with positive and with negative frequency, that we wiill name

$$\psi^{(+)} = u(P) e^{-iP_\mu X^\mu} , \tag{4.435}$$
$$\psi^{(-)} = v(P) e^{iP_\mu X^\mu} . \tag{4.436}$$

Substituting into the Dirac equation, therefore, we find that the spinors $u(p)$ and $v(p)$ are solutions of the following equations

$$(\not{P} - m) u(P) = 0 , \tag{4.437}$$
$$(\not{P} + m) v(P) = 0 . \tag{4.438}$$

In order to solve the system (4.437,4.438) it is convenient to move in the frame in which the particle is at rest, i.e. in the frame in which $P^\mu = (m, \mathbf{0})$. In this frame, $\not{P} = \gamma^0 m$ and therefore we get

$$(\gamma^0 m - m) u(m, \mathbf{0}) = 0 , \tag{4.439}$$
$$(\gamma^0 m + m) v(m, \mathbf{0}) = 0 , \tag{4.440}$$

or, since $m \neq 0$

$$(\gamma^0 - \mathbb{1}) u(m, \mathbf{0}) = 0 , \tag{4.441}$$
$$(\gamma^0 + \mathbb{1}) v(m, \mathbf{0}) = 0 , \tag{4.442}$$

If we define the general spinors

$$u(m, \mathbf{0}) = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} , \qquad v(m, \mathbf{0}) = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} , \tag{4.443}$$

and we consider for instance the Pauli representation for the gamma matrices, Eqs. (4.441,4.442) have the following solution

$$u_3 = u_4 = v_1 = v_2 = 0 . \tag{4.444}$$

We do not find any constraint on the other components, and therefore the general solution is the following linear combination

$$u(m, \mathbf{0}) = \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \alpha u^{(1)}(m, \mathbf{0}) + \beta u^{(2)}(m, \mathbf{0}) , \tag{4.445}$$

$$v(m,\mathbf{0}) \;=\; \alpha' \begin{pmatrix}0\\0\\1\\0\end{pmatrix} + \beta' \begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \alpha' v^{(1)}(m,\mathbf{0}) + \beta' v^{(2)}(m,\mathbf{0})\,, \tag{4.446}$$

where we defined

$$u^{(1)}(m,\mathbf{0}) = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad u^{(2)}(m,\mathbf{0}) = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \quad v^{(1)}(m,\mathbf{0}) = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad v^{(2)}(m,\mathbf{0}) = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}, \tag{4.447}$$

and where the degeneracy is due to the spin. In the rest frame we have: positive-energy solutions, with spin up or spin down, and negative-energy solutions with spin up or spin down.

The spinors $u^{(1)}(m,\mathbf{0})$, $u^{(2)}(m,\mathbf{0})$, $v^{(1)}(m,\mathbf{0})$, $v^{(2)}(m,\mathbf{0})$, are eigenvectors of the third component of the spin, with eigenvalues $\pm 1/2$.

In order to find the general solution, $u(P)$, $v(P)$ in a frame in which $P^\mu = (E,\mathbf{p})$ we can boost our solutions, found in the rest frame, using Eq. (4.343).

If

$$\gamma = \frac{E}{m}\,, \quad \beta\gamma = \frac{p}{m}\,, \quad \mathbf{p} = p\hat{\mathbf{n}}\,, \tag{4.448}$$

we have

$$S(\Lambda) = \sqrt{\frac{\gamma+1}{2}} \begin{pmatrix} 1 & -\frac{\beta\gamma}{\gamma+1}\sigma\cdot\mathbf{n} \\ -\frac{\beta\gamma}{\gamma+1}\sigma\cdot\mathbf{n} & 1 \end{pmatrix} = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 & -\frac{\sigma\cdot\mathbf{p}}{E+m} \\ -\frac{\sigma\cdot\mathbf{p}}{E+m} & 1 \end{pmatrix}. \tag{4.449}$$

Therefore, using the generic two-component spinor in the rest frame

$$u^{(\alpha)}(m,\mathbf{0}) = \begin{pmatrix}\phi^{(\alpha)}\\0\end{pmatrix}, \quad v^{(\alpha)}(m,\mathbf{0}) = \begin{pmatrix}0\\\chi^{(\alpha)}\end{pmatrix}, \tag{4.450}$$

where $\alpha = 1,2$, $\phi^1 \propto u^{(1)}(m,\mathbf{0})$, $\phi^2 \propto u^{(2)}(m,\mathbf{0})$, ... etc, we have

$$u^{(\alpha)}(P) = S^{-1}(\Lambda)u(m,\mathbf{0}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 & \frac{\sigma\cdot\mathbf{p}}{E+m} \\ \frac{\sigma\cdot\mathbf{p}}{E+m} & 1 \end{pmatrix} \begin{pmatrix}\phi^{(\alpha)}\\0\end{pmatrix} = \begin{pmatrix} \sqrt{\frac{E+m}{2m}}\,\phi^{(\alpha)} \\ \frac{\sigma\cdot\mathbf{p}}{\sqrt{2m(E+m)}}\phi^{(\alpha)} \end{pmatrix}. \tag{4.451}$$

The same for the spinor $v(P)$:

$$v^{(\alpha)}(P) = S^{-1}(\Lambda)v(m,\mathbf{0}) = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 & \frac{\sigma\cdot\mathbf{p}}{E+m} \\ \frac{\sigma\cdot\mathbf{p}}{E+m} & 1 \end{pmatrix} \begin{pmatrix}0\\\chi^{(\alpha)}\end{pmatrix} = \begin{pmatrix} \frac{\sigma\cdot\mathbf{p}}{\sqrt{2m(E+m)}}\chi^{(\alpha)} \\ \sqrt{\frac{E+m}{2m}}\,\chi^{(\alpha)} \end{pmatrix}. \tag{4.452}$$

The same result can be found noting that

$$(\not{P} - m)(\not{P} + m) = P^2 - m^2 = 0\,. \tag{4.453}$$

Therefore, if we define

$$u^{(\alpha)}(P) \;=\; C_\alpha(\not{P} + m)u^{(\alpha)}(m,\mathbf{0})\,, \tag{4.454}$$

$$v^{(\alpha)}(P) \;=\; D_\alpha(-\not{P} + m)v^{(\alpha)}(m,\mathbf{0})\,, \tag{4.455}$$

where C_α and D_α are normalization factors, we find immediately

$$(\not{P} - m)u^{(\alpha)}(P) \;=\; 0\,, \tag{4.456}$$

$$(\not{P}+m)v^{(\alpha)}(P) \;=\; 0\,. \tag{4.457}$$

In order to find the normalization factors C_α and D_α, we note that (by direct inspection) the following relations hold in the rest frame:

$$\overline{u}^{(\alpha)}(m,\mathbf{0})u^{(\beta)}(m,\mathbf{0}) \;=\; \delta^{\alpha\beta}\,, \tag{4.458}$$
$$\overline{v}^{(\alpha)}(m,\mathbf{0})v^{(\beta)}(m,\mathbf{0}) \;=\; -\delta^{\alpha\beta}\,, \tag{4.459}$$
$$\overline{u}^{(\alpha)}(m,\mathbf{0})v^{(\beta)}(m,\mathbf{0}) \;=\; 0\,. \tag{4.460}$$
$$\tag{4.461}$$

These relations are already cast in scalar form, in the sense that in a generic frame it must hold

$$\overline{u}^{(\alpha)}(P)u^{(\beta)}(P) \;=\; \delta^{\alpha\beta}\,, \tag{4.462}$$
$$\overline{v}^{(\alpha)}(P)v^{(\beta)}(P) \;=\; -\delta^{\alpha\beta}\,, \tag{4.463}$$
$$\overline{u}^{(\alpha)}(P)v^{(\beta)}(P) \;=\; 0\,, \tag{4.464}$$
$$\tag{4.465}$$

that can be used to impose the normalization of the spinors:

$$\delta^{\alpha\beta} \;=\; \overline{u}^{(\alpha)}(P)u^{(\beta)}(P) = C^*_\alpha(u^{(\alpha)}(m,\mathbf{0}))^\dagger(\not{P}+m)^\dagger\gamma^0\,C_\beta(\not{P}+m)u^{(\beta)}(m,\mathbf{0})\,, \tag{4.466}$$
$$=\; C^*_\alpha C_\beta\,\overline{u}^{(\alpha)}(m,\mathbf{0})(\not{P}+m)^2u^{(\beta)}(m,\mathbf{0})\,, \tag{4.467}$$
$$=\; C^*_\alpha C_\beta\,\overline{u}^{(\alpha)}(m,\mathbf{0})(2m\;\not{P}+2m^2)u^{(\beta)}(m,\mathbf{0})\,, \tag{4.468}$$
$$=\; |\,\text{since } \overline{u}^{(\alpha)}(m,\mathbf{0})\;\not{P}u^{(\beta)}(m,\mathbf{0}) = P_0\overline{u}^{(\alpha)}(m,\mathbf{0})\gamma^0u^{(\beta)}(m,\mathbf{0}) = E\overline{u}^{(\alpha)}(m,\mathbf{0})u^{(\beta)}(m,\mathbf{0})\,|$$
$$=\; C^*_\alpha C_\beta\,2m(E+m)\overline{u}^{(\alpha)}(m,\mathbf{0})u^{(\beta)}(m,\mathbf{0})\,, \tag{4.469}$$
$$=\; |C_\alpha|^2\,2m(E+m)\,\delta^{\alpha\beta}\,. \tag{4.470}$$

This gives (apart from a phase that we choose to be equal to zero)

$$C_\alpha = \frac{1}{\sqrt{2m(E+m)}}\,. \tag{4.471}$$

The same expression we find for D_α, using Eq. (4.463). In the end

$$u^{(\alpha)}(P) \;=\; \frac{\not{P}+m}{\sqrt{2m(E+m)}}\,u^{(\alpha)}(m,\mathbf{0})\,, \tag{4.472}$$
$$v^{(\alpha)}(P) \;=\; \frac{-\not{P}+m}{\sqrt{2m(E+m)}}\,v^{(\alpha)}(m,\mathbf{0})\,. \tag{4.473}$$

It is easy to check that these spinors are indeed ortogonal (they satisfy Eq. (4.464)):

$$\overline{u}^{(\alpha)}(P)\,v^{(\beta)}(P) = \frac{\overline{u}^{(\alpha)}(m,\mathbf{0})(\not{P}+m)(-\not{P}+m)v^{(\beta)}(m,\mathbf{0})}{2m(E+m)} = 0\,. \tag{4.474}$$

Using the two-component expression for $u(P)$ and $v(P)$

$$u^{(\alpha)}(P) = \begin{pmatrix}\phi^{(\alpha)}\\ 0\end{pmatrix}\,, \qquad v^{(\alpha)}(P) = \begin{pmatrix}0\\ \chi^{(\alpha)}\end{pmatrix} \tag{4.475}$$

and explicitely expressing $(\not{P}+m)$ and $(-\not{P}+m)$ in matrix notation, we get the expressions of Eqs. (4.451,refvsp):

$$u^{(\alpha)}(P) \;=\; \begin{pmatrix}\sqrt{\frac{E+m}{2m}}\,\phi^{(\alpha)}\\ \frac{\sigma\cdot\mathbf{p}}{\sqrt{2m(E+m)}}\phi^{(\alpha)}\end{pmatrix}\,, \tag{4.476}$$

$$v^{(\alpha)}(P) \;=\; \begin{pmatrix} \frac{\sigma\cdot\mathbf{p}}{\sqrt{2m(E+m)}}\chi^{(\alpha)} \\ \sqrt{\frac{E+m}{2m}}\,\chi^{(\alpha)} \end{pmatrix}. \tag{4.477}$$

In this way, we found the spinors normalized in the sense of Eqs. (4.462,4.463). However, the scalar product for our fields is defined in terms of $\psi^\dagger$ and not $\overline{\psi}$. What we would like to impose is the normalization of the charge, which is given by

$$Q = \int d^3X\,\psi^\dagger\psi\,. \tag{4.478}$$

Let us note that

$$\begin{aligned}
(u^{(\alpha)}(P))^\dagger u^{(\alpha)}(P) &= \begin{pmatrix} \sqrt{\frac{E+m}{2m}}\,(\phi^{(\alpha)})^\dagger & \frac{\sigma\cdot\mathbf{p}}{\sqrt{2m(E+m)}}(\phi^{(\alpha)})^\dagger \end{pmatrix} \begin{pmatrix} \sqrt{\frac{E+m}{2m}}\,\phi^{(\alpha)} \\ \frac{\sigma\cdot\mathbf{p}}{\sqrt{2m(E+m)}}\phi^{(\alpha)} \end{pmatrix}, && (4.479)\\
&= \frac{E+m}{2m}(\phi^{(\alpha)})^\dagger\phi^{(\alpha)} + \frac{(\sigma\cdot\mathbf{p})^2}{2m(E+m)}(\phi^{(\alpha)})^\dagger\phi^{(\alpha)}\,, && (4.480)\\
&= |\,\text{since } (\phi^{(\alpha)})^\dagger\phi^{(\alpha)} = 1 \quad \text{and } (\sigma\cdot\mathbf{p})^2 = p^2 = E^2 - m^2\,| \\
&= \frac{E}{m}\,, && (4.481)
\end{aligned}$$

while $(u^{(\alpha)}(P))^\dagger u^{(\beta)}(P) = 0$ when $\alpha \neq \beta$. For the spinor $v(P)$ we get

$$\begin{aligned}
(v^{(\alpha)}(P))^\dagger v^{(\alpha)}(P) &= \begin{pmatrix} \frac{\sigma\cdot\mathbf{p}}{\sqrt{2m(E+m)}}(\chi^{(\alpha)})^\dagger & \sqrt{\frac{E+m}{2m}}\,(\chi^{(\alpha)})^\dagger \end{pmatrix} \begin{pmatrix} \frac{\sigma\cdot\mathbf{p}}{\sqrt{2m(E+m)}}\chi^{(\alpha)} \\ \sqrt{\frac{E+m}{2m}}\,\chi^{(\alpha)} \end{pmatrix}, && (4.482)\\
&= \frac{p^2}{2m(E+m)}(\chi^{(\alpha)})^\dagger\chi^{(\alpha)} + \frac{E+m}{2m}(\chi^{(\alpha)})^\dagger\chi^{(\alpha)}\,, && (4.483)\\
&= \frac{E}{m}\,. && (4.484)
\end{aligned}$$

In order to normalize, using the correct scalar product, the positive and negative energy solutions, then, we have to consider

$$\psi^{(+)}_{(\alpha)}(X) = N\,u^{(\alpha)}(P)\sqrt{\frac{m}{E}}e^{-iP_\mu X^\mu}\,, \tag{4.485}$$

$$\psi^{(-)}_{(\alpha)}(X) = N\,v^{(\alpha)}(P)\sqrt{\frac{m}{E}}e^{iP_\mu X^\mu}\,, \tag{4.486}$$

such that

$$(\psi^{(+)}_{(\alpha)}(X))^\dagger\psi^{(+)}_{(\beta)}(X) = \delta_{\alpha\beta}\,, \quad (\psi^{(-)}_{(\alpha)}(X))^\dagger\psi^{(-)}_{(\beta)}(X) = \delta_{\alpha\beta} \quad (\psi^{(+)}_{(\alpha)}(X))^\dagger\psi^{(-)}_{(\beta)}(X) = 0\,. \tag{4.487}$$

In Eqs. (4.485,4.486), N is a normalization factor.

Recalling the scalar product

$$(\psi_1, \psi_2) = \int d^3X\,\psi_1^\dagger\psi_2\,, \tag{4.488}$$

we now want to normalize the fields to the delta:

$$\begin{aligned}
(\psi^{(+)}_{(\alpha)}(X), \psi^{(+)}_{(\beta)}(X)) &= |N|^2\,\frac{m}{E}\int d^3X\,(u^{(\alpha)}(P))^\dagger u^{(\beta)}(Q)e^{i(P-Q)_\mu X^\mu}\,, \\
&= |N|^2\,\frac{m}{E}\,\frac{E}{m}\delta_{\alpha\beta}\int d^3X\,e^{i(P-Q)_\mu X^\mu} = |N|^2\,\delta_{\alpha\beta}\,(2\pi)^3\,\delta(\mathbf{p}-\mathbf{q})\,. && (4.489)
\end{aligned}$$

This implies (we choose N real):

$$N = \frac{1}{\sqrt{(2\pi)^3}} \,. \tag{4.490}$$

Finally, the full expression of the Dirac field in normal modes is

$$\psi(X) = \sum_{\alpha=1}^{2} \int \frac{d^3p}{(2\pi)^{\frac{3}{2}}} \sqrt{\frac{m}{E}} \left[b_{(\alpha)}(p)\, u^{(\alpha)}(P) e^{-iP_\mu X^\mu} + d^*_{(\alpha)}(p)\, v^{(\alpha)}(P) e^{iP_\mu X^\mu} \right] , \tag{4.491}$$

with $E = \sqrt{p^2 + m^2} > 0$ and where, for the classic field, $b_{(\alpha)}(P)$ and $d^*_{(\alpha)}(P)$ are the coefficients of the linear combination.

4.3.12 Energy projectors and polarization sum

It is convenient to introduce the projectors for positive and negative-energy, spin up and spin down solutions, in such a way that from a generic solution we could project out four independent solutions (positive-energy spin-up, positive-energy spin-down, negative-energy spin-up, negative-energy spin-down solutions).

Considering that

$$(\not{P} + m)(\not{P} + m) = 2m(\not{P} + m) \,, \qquad (\not{P} + m)(-\not{P} + m) = 0 \,, \tag{4.492}$$

let us write the following operators

$$\Lambda_\pm = \frac{\pm \not{P} + m}{2m} \,. \tag{4.493}$$

These are indeed the projectors we were looking for. In fact, if

$$\psi(X) \sim \alpha\, u(P) + \beta\, v(P) \,, \tag{4.494}$$

we have

$$\Lambda_+ \psi(X) = \alpha\, u(P) \,, \qquad \text{and} \quad \Lambda_- \psi(X) = \beta\, v(P) \,. \tag{4.495}$$

The operators $\Lambda_\pm$ are projectors. In fact

$$\Lambda_\pm^2 = \frac{1}{4m^2}(\pm \not{P} + m)(\pm \not{P} + m) = \frac{\pm \not{P} + m}{2m} = \Lambda_\pm \,, \tag{4.496}$$

$$\Lambda_+\Lambda_- = \frac{1}{4m^2}(\not{P} + m)(-\not{P} + m) = 0 \,, \tag{4.497}$$

$$\Lambda_+ + \Lambda_- = \frac{1}{2m}[\not{P} + m + (-\not{P} + m)] = \mathbb{1} \,. \tag{4.498}$$

The projectors $\Lambda_\pm$ can be written in terms of the polarization sum of the spinors as follows. We have

$$\sum_{\alpha=1}^{2} u_{(\alpha)}(P)\overline{u}_{(\alpha)}(P) = \sum_{\alpha=1}^{2} u_{(\alpha)}(P)(u_{(\alpha)}(P))^\dagger \gamma^0 \,, \tag{4.499}$$

$$= \frac{1}{2m(E+m)} \sum_{\alpha=1}^{2} (\not{P} + m) u_{(\alpha)}(m, \mathbf{0})(u_{(\alpha)}(m, \mathbf{0}))^\dagger (\not{P} + m)^\dagger \gamma^0 \,, \tag{4.500}$$

$$= \frac{1}{2m(E+m)} (\not{P} + m) \sum_{\alpha=1}^{2} u_{(\alpha)}(m, \mathbf{0})\overline{u}_{(\alpha)}(m, \mathbf{0})(\not{P} + m) \,, \tag{4.501}$$

$$= \left|\, \text{since} \sum_{\alpha=1}^{2} u_{(\alpha)}(m, \mathbf{0})\overline{u}_{(\alpha)}(m, \mathbf{0}) = \frac{1+\gamma^0}{2} \,\right|$$

$$
\begin{aligned}
&= \frac{1}{2m(E+m)}(\not{P}+m)\frac{1+\gamma^0}{2}(\not{P}+m)\,, &(4.502)\\
&= \frac{1}{4m(E+m)}\left[2m(\not{P}+m)+(\not{P}+m)(\gamma^0\gamma^\nu P_\nu+m\gamma^0)\right]\,, &(4.503)\\
&= |\,\text{since } \gamma^0\gamma^\nu=-\gamma^\nu\gamma^0+2\eta^{0\nu}\,|\\
&= \frac{1}{4m(E+m)}\left[2m(\not{P}+m)+(\not{P}+m)(2E+(-\not{P}+m)\gamma^0)\right]\,, &(4.504)\\
&= \frac{2(E+m)}{4m(E+m)}(\not{P}+m)\,, &(4.505)\\
&= \frac{(\not{P}+m)}{2m}=\Lambda_+\,. &(4.506)
\end{aligned}
$$

Analogously we find

$$\sum_{\alpha=1}^{2} v_{(\alpha)}(P)\overline{v}_{(\alpha)}(P)=\frac{(\not{P}-m)}{2m}=-\Lambda_-\,. \tag{4.507}$$

Therefore

$$\sum_{\alpha=1}^{2} u_{(\alpha)}(P)\overline{u}_{(\alpha)}(P)-\sum_{\alpha=1}^{2} v_{(\alpha)}(P)\overline{v}_{(\alpha)}(P)=\mathbb{1}\,. \tag{4.508}$$

4.3.13 Spin projectors

The positive and negative energy solutions are still doubly degenerate. It is possible to remove such deceneracy selecting a spin state through spin projectors.

Let us consider the solution of the Dirac equation in the rest frame. The spinors $u^{(1)}(m,\mathbf{0})$ and $u^{(2)}(m,\mathbf{0})$ are eigenstates of

$$\sigma_{12}=\frac{i}{2}[\gamma_1,\gamma_2]=\begin{pmatrix}\sigma^3 & 0\\ 0 & \sigma^3\end{pmatrix} \tag{4.509}$$

with eigenvalues $+1$ and -1, respectively. The same is true for $v^{(1)}(m,\mathbf{0})$ and $v^{(2)}(m,\mathbf{0})$. Therefore, a projector for eigenstates of spin up (in the $\hat{\mathbf{z}}$ direction) can be looked for in the following expression

$$\tilde{\Sigma}(\hat{\mathbf{z}})=\frac{1+\sigma_{12}}{2}\,, \tag{4.510}$$

such that

$$
\begin{aligned}
\tilde{\Sigma}(\hat{\mathbf{z}})u^{(1)}(m,\mathbf{0}) &= \frac{1+\sigma_{12}}{2}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}=u^{(1)}(m,\mathbf{0})\,, &(4.511)\\
\tilde{\Sigma}(\hat{\mathbf{z}})u^{(2)}(m,\mathbf{0}) &= \frac{1+\sigma_{12}}{2}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}=0\,, &(4.512)\\
\tilde{\Sigma}(\hat{\mathbf{z}})v^{(1)}(m,\mathbf{0}) &= \frac{1+\sigma_{12}}{2}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}=v^{(1)}(m,\mathbf{0})\,, &(4.513)\\
\tilde{\Sigma}(\hat{\mathbf{z}})v^{(2)}(m,\mathbf{0}) &= \frac{1+\sigma_{12}}{2}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}=0\,. &(4.514)
\end{aligned}
$$

We also notice that

$$\sigma_{12} = \frac{i}{2}[\gamma_1, \gamma_2] = i\gamma^1\gamma^2 = -\gamma^0\gamma_5\gamma^3 = \gamma_5\gamma_3\gamma^0\,, \tag{4.515}$$

and that

$$\tilde{\Sigma}(\hat{\mathbf{z}}) = \frac{1+\sigma_{12}}{2} = \frac{1+\sigma_{12}\hat{n}_R^3}{2}\,, \tag{4.516}$$

where $\hat{n}_R^3$ is the third spatial component of the space-like vector

$$\hat{n}_R^\mu = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \qquad \left(\hat{n}_R^2 = -1\,, \quad \hat{n}_R^\mu P^\mu = 0\right)\,, \tag{4.517}$$

in the rest frame. We can therefore write

$$\tilde{\Sigma}(\hat{\mathbf{z}}) = \frac{1+\gamma_5\gamma_3\gamma^0}{2} = \frac{1+\gamma_5\gamma_3\hat{n}^3\gamma^0}{2} = \frac{1+\gamma_5\ \hat{\not{n}}_R\gamma^0}{2} = \tilde{\Sigma}(\hat{n}_R)\,, \tag{4.518}$$

where, in the rest frame

$$\hat{\not{n}}_R = \gamma_0\hat{n}_R^0 + \gamma_i\hat{n}_R^i = \gamma_3\hat{n}_R^3\,. \tag{4.519}$$

The expression of $\Sigma(\hat{n}_R)$ in Eq. (4.518) is "almost" generalizable to a generic inertial frame. The problem is the presence of γ^0, that does not allow to use the same expression in another frame. If we could drop the γ^0 from Eq. (4.518), we would have reached our goal.

Projectors $\tilde{\Sigma}(\pm\hat{n}_R) = \frac{1\pm\gamma_5\hat{\not{n}}_R\gamma^0}{2}$ behave as follows

$$\tilde{\Sigma}(+\hat{n}_R)\begin{pmatrix}\alpha\\ \beta\\ \gamma\\ \delta\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1+\sigma^3 & 0\\ 0 & 1+\sigma^3\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\\ \gamma\\ \delta\end{pmatrix} = \begin{pmatrix}\alpha\\ 0\\ \gamma\\ 0\end{pmatrix}\,, \tag{4.520}$$

$$\tilde{\Sigma}(-\hat{n}_R)\begin{pmatrix}\alpha\\ \beta\\ \gamma\\ \delta\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1-\sigma^3 & 0\\ 0 & 1-\sigma^3\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\\ \gamma\\ \delta\end{pmatrix} = \begin{pmatrix}0\\ \beta\\ 0\\ \delta\end{pmatrix}\,. \tag{4.521}$$

Let us see to which projectors correspond instead the

$$\Sigma(\pm\hat{n}_R) = \frac{1\pm\gamma_5\ \hat{\not{n}}_R}{2}\,, \tag{4.522}$$

without the γ^0 in their expression. We have

$$\Sigma(+\hat{n}_R)\begin{pmatrix}\alpha\\ \beta\\ \gamma\\ \delta\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1+\sigma^3 & 0\\ 0 & 1-\sigma^3\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\\ \gamma\\ \delta\end{pmatrix} = \begin{pmatrix}\alpha\\ 0\\ 0\\ \delta\end{pmatrix}\,, \tag{4.523}$$

$$\Sigma(-\hat{n}_R)\begin{pmatrix}\alpha\\ \beta\\ \gamma\\ \delta\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1-\sigma^3 & 0\\ 0 & 1+\sigma^3\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\\ \gamma\\ \delta\end{pmatrix} = \begin{pmatrix}0\\ \beta\\ \gamma\\ 0\end{pmatrix}\,. \tag{4.524}$$

Therefore, in the rest frame $\Sigma(+\hat{n}_R)$ projects positive-energy spin-up and negative-energy spin-down solutions, while $\Sigma(-\hat{n}_R)$ projects positive-energy spin-down and negative-energy spin-up solutions.

The expression of the spin projectors in a general frame is, then

$$\Sigma(\pm\hat{n}) = \frac{1 \pm \gamma_5 \not{\hat{n}}}{2}, \tag{4.525}$$

in which now $\hat{n}^\mu$ is the boosted unit space-like vector. In fact, if Λ is the boost, according to which

$$P^\mu = \Lambda^\mu_\nu P^\nu_R, \tag{4.526}$$

where $P^\mu_R = (m, \mathbf{0})$, we have

$$\begin{aligned} \Sigma(\pm\hat{n}) &= S^{-1}(\Lambda)\Sigma(\pm\hat{n}_R)S(\Lambda) = \frac{1 \pm S^{-1}(\Lambda)\gamma_5 S(\Lambda) S^{-1}(\Lambda)\gamma_\mu S(\Lambda)\hat{n}^\mu_R}{2}, &(4.527)\\ &= \frac{1 \pm \gamma_5 \gamma^\nu \Lambda_{\mu\nu}\hat{n}^\mu_R}{2} = \frac{1 \pm \gamma_5 \not{\hat{n}}}{2}. &(4.528)\end{aligned}$$

The vector $\hat{n}^\mu$ is still space-like, $\hat{n}^2 = -1$, and since in the rest frame we have $\hat{n}^\mu_R P_{R\mu} = 0$, in the boosted frame we still have $\hat{n}^\mu P_\mu = 0$. The operator $\frac{1}{2}\gamma_5 \not{\hat{n}}$ is called the Pauli-Lubanski operator and it is the relativistic generalization of what in the rest frame is the projection of the spin $\sigma/2$ in the direction of $\hat{n}_R$. If in the rest frame we have

$$\begin{aligned} \Sigma(\pm\hat{n}_R)u^{(\alpha)}(m, \mathbf{0}) &= u^{(\alpha)}(m, \mathbf{0}), &(4.529)\\ \Sigma(\pm\hat{n}_R)v^{(\alpha)}(m, \mathbf{0}) &= v^{(\alpha)}(m, \mathbf{0}), &(4.530)\end{aligned}$$

in the boosted frame we still have

$$\begin{aligned} \Sigma(\pm\hat{n})u^{(\alpha)}(P) &= \Sigma(\pm\hat{n})S^{-1}(\Lambda)u^{(\alpha)}(m, \mathbf{0}) = S^{-1}(\Lambda)\Sigma(\pm\hat{n}_R)u^{(\alpha)}(m, \mathbf{0}) = S^{-1}(\Lambda)u^{(\alpha)}(m, \mathbf{0}),\\ &= u^{(\alpha)}(P), &(4.531)\\ \Sigma(\pm\hat{n})v^{(\alpha)}(P) &= \Sigma(\pm\hat{n})S^{-1}(\Lambda)v^{(\alpha)}(m, \mathbf{0}) = S^{-1}(\Lambda)\Sigma(\pm\hat{n}_R)v^{(\alpha)}(m, \mathbf{0}) = S^{-1}(\Lambda)v^{(\alpha)}(m, \mathbf{0}),\\ &= v^{(\alpha)}(P). &(4.532)\end{aligned}$$

$\Sigma(\pm\hat{n})$ project out positive energy solutions with spin projection in the $\hat{n}$ direction of $\pm\frac{1}{2}$ and negative energy solutions with spin projection $\mp\frac{1}{2}$.

$\Sigma(\pm\hat{n})$ are actually projectors, then they satisfy the following properties:

$$\begin{aligned} \Sigma^2(\pm\hat{n}) &= \frac{1}{4}(1 \pm \gamma_5 \not{\hat{n}})^2 = \frac{1}{4}(1 \pm 2\gamma_5 \not{\hat{n}} + \gamma_5 \not{\hat{n}}\gamma_5 \not{\hat{n}}) = \frac{1}{4}(2 \pm 2\gamma_5 \not{\hat{n}}),\\ &= \Sigma(\pm\hat{n}), &(4.533)\\ \Sigma(+\hat{n}) + \Sigma(-\hat{n}) &= \frac{1}{2}(1 + \gamma_5 \not{\hat{n}} + 1 - \gamma_5 \not{\hat{n}}) = \mathbb{1}, &(4.534)\\ \Sigma(+\hat{n})\Sigma(-\hat{n}) &= \left(\frac{1 + \gamma_5 \not{\hat{n}}}{2}\right)\left(\frac{1 - \gamma_5 \not{\hat{n}}}{2}\right) = 0. &(4.535)\end{aligned}$$

We have

$$[\Lambda_\pm, \Sigma(\pm\hat{n})] = 0, \quad \text{for every } \hat{n} \text{ such that } \hat{n}^\mu P_\mu = 0. \tag{4.536}$$

In fact

$$\begin{aligned} \not{P}\gamma_5 \not{\hat{n}} &= P^\mu \hat{n}^\nu \gamma_\mu \gamma_5 \gamma_\nu = -P^\mu \hat{n}^\nu \gamma_5 \gamma_\mu \gamma_\nu = -P^\mu \hat{n}^\nu \gamma_5(-\gamma_\nu\gamma_\mu + 2\eta_{\mu\nu}), &(4.537)\\ &= \gamma_5 \not{\hat{n}} \not{P} + 2\gamma_5 P_\nu \hat{n}^\nu, &(4.538)\\ &= |\,\text{since } \hat{n}^\mu P_\mu = 0\,|\\ &= \gamma_5 \not{\hat{n}} \not{P}. &(4.539)\end{aligned}$$

and therefore, if $P_\nu \hat{n}^\nu = 0$, we have

$$\left(\frac{\pm \not{P} + m}{2m}\right)\left(\frac{1 \pm \gamma_5 \not{\hat{n}}}{2}\right) = \left(\frac{1 \pm \gamma_5 \not{\hat{n}}}{2}\right)\left(\frac{\pm \not{P} + m}{2m}\right). \tag{4.540}$$

Using $\Lambda_\pm$ and $\Sigma(\pm\hat{n})$ we can compose projectors for definite energy and spin

$$P_1 = \Lambda_+ \Sigma(+\hat{n}), \tag{4.541}$$
$$P_2 = \Lambda_+ \Sigma(-\hat{n}), \tag{4.542}$$
$$P_3 = \Lambda_- \Sigma(+\hat{n}), \tag{4.543}$$
$$P_4 = \Lambda_- \Sigma(-\hat{n}), \tag{4.544}$$

such that

$$\sum_{i=1}^{4} P_i = \mathbb{1}, \quad P_i P_j = \delta_{ij}, \quad \mathrm{tr} P_i = 1. \tag{4.545}$$

4.3.14 Non relativistic limit of the Dirac's equation

We consider in this section the case in which the particle that we would like to describe using Dirac's equation moves with a speed much smaller than the speed of light, $v \ll c$, and it is in interaction with an electromagnetic field.

In order to describe the interaction, we perform the so-called "minimal substitution" in the Dirac's eqution. This ammounts to

$$\partial^\mu \to \partial^\mu + ieA^\mu, \tag{4.546}$$

where e is the electric charge of the electron (negative, so $e = -|e|$) and A^μ is the electromagnetic four-potential $A^\mu = (\phi, \mathbf{A})$. Under the substitution (5.19) the free Dirac's equation becomes

$$(i\not{\partial} - e\not{A} - m)\psi(X) = 0. \tag{4.547}$$

In components we have

$$(i\gamma_0\partial^0 - e\gamma_0 A^0 - m)\psi(X) + \gamma_i(i\partial^i - eA^i)\psi(X) = 0. \tag{4.548}$$

We would like that Eq. (4.548) would provide an accurate description of the behaviour of an electron (positve-energy state) in an electromagnetic field for small velocities. Its energy will be

$$E = \sqrt{p^2 + m^2} \sim m + \frac{p^2}{2m} + ... \tag{4.549}$$

where $\frac{p^2}{2m} \ll m$. In this situation the term $e^{-iP_\mu X^\mu}$ is dominated by e^{-imt} that oscillates much faster than any other term. It is then convenient to isolate such fast varying term redefining our positive-energy solution as

$$\psi(X) = \tilde{\psi}(X)e^{-imt}, \tag{4.550}$$

where now $\tilde{\psi}(X)$ oscillates much slower, $\sim e^{-iE't}$ where $E' = E - m \ll m$. Substituting in Eq. (4.548), we find an equation for $\tilde{\psi}(X)$:

$$\gamma_0(i\partial^0 - eA^0 + m)\tilde{\psi}(X) - m\tilde{\psi}(X) + \gamma_i(i\partial^i - eA^i)\tilde{\psi}(X) = 0. \tag{4.551}$$

If we express $\tilde{\psi}(X)$ with two two-component spinors

$$\tilde{\psi}(X) = \begin{pmatrix} \tilde{\phi} \\ \tilde{\chi} \end{pmatrix}, \tag{4.552}$$

we find

$$\begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} (i\partial^0 - eA^0 + m) \begin{pmatrix} \tilde{\phi} \\ \tilde{\chi} \end{pmatrix} - m \begin{pmatrix} \tilde{\phi} \\ \tilde{\chi} \end{pmatrix} + \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix} (i\partial^i - eA^i) \begin{pmatrix} \tilde{\phi} \\ \tilde{\chi} \end{pmatrix} = 0\,, \tag{4.553}$$

or, the following system:

$$\begin{cases} (i\partial^0 - eA^0)\tilde{\phi} = \sigma \cdot (\mathbf{p} - e\mathbf{A})\tilde{\chi} \\ (i\partial^0 - eA^0 + 2m)\tilde{\chi} = \sigma \cdot (\mathbf{p} - e\mathbf{A})\tilde{\phi} \end{cases} \tag{4.554}$$

In the second equation we can neglect the terms $i\partial^0\tilde{\chi}$ and $-eA^0\tilde{\chi}$ with respect to $2m\tilde{\chi}$ and we can therefore solve for $\tilde{\chi}$ as follows:

$$\tilde{\chi} = \frac{\sigma \cdot (\mathbf{p} - e\mathbf{A})}{2m}\tilde{\phi}\,. \tag{4.555}$$

Eq. (4.555) tells us that $\tilde{\chi}$ is "small" with respect to $\tilde{\phi}$ (of order of p/m). So, the spinor is basically described, in this limit, by the two-component spinor $\tilde{\phi}$. Substituting (4.555) in the first equation of (4.554), we find an equation for $\tilde{\phi}$:

$$(i\partial^0 - eA^0)\tilde{\phi} = \frac{[\sigma \cdot (\mathbf{p} - e\mathbf{A})]^2}{2m}\tilde{\phi}\,. \tag{4.556}$$

We have

$$\begin{aligned} [\sigma \cdot (\mathbf{p} - e\mathbf{A})]^2 &= \sigma^i\sigma^j(p^i - eA^i)(p^j - eA^j)\,, && (4.557) \\ &= |\,\text{since } \sigma^i\sigma^j = \frac{1}{2}[\sigma^i,\sigma^j] + \frac{1}{2}[\sigma^i,\sigma^j]_+\,| \\ &= (\delta^{ij} + i\epsilon^{ijk}\sigma^k)(p^i - eA^i)(p^j - eA^j)\,, && (4.558) \\ &= (\mathbf{p} - e\mathbf{A})^2 + i\epsilon^{ijk}\sigma^k(p^ip^j - ep^iA^j - eA^ip^j + e^2A^iA^j)\,. && (4.559) \end{aligned}$$

The two terms p^ip^j and A^iA^j are totally symmetric in ij, therefore, when we saturate with the epsilon-tensor they vanish. Then we have to remember that p^j and A^j do not commute, since $p^i = i\partial^i$ and $A^j = A^j(X)$. Therefore we have

$$p^iA^j = i\partial^iA^j + A^jp^i \tag{4.560}$$

and

$$\epsilon^{ijk}(-ep^iA^j - eA^ip^j) = \epsilon^{ijk}[-ei\partial^iA^j - e(A^ip^j + A^jp^i)] = -e\sigma^k(\nabla \wedge \mathbf{A})^k = -e\sigma \cdot \mathbf{B}\,, \tag{4.561}$$

since $\nabla \wedge \mathbf{A} = \mathbf{B}$ is the magnetic field.

Finally

$$[\sigma \cdot (\mathbf{p} - e\mathbf{A})]^2 = (\mathbf{p} - e\mathbf{A})^2 - e\sigma \cdot \mathbf{B} \tag{4.562}$$

and therefore

$$i\frac{\partial}{\partial t}\tilde{\phi} = \left(eA^0 + \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} - \frac{e}{m}\frac{\sigma}{2} \cdot \mathbf{B}\right)\tilde{\phi} = H\tilde{\phi}\,. \tag{4.563}$$

Eq. (4.563) is the Schrödinger equation of a spin-1/2 particle in an electromagnetic field. In particular, Dirac's equation describes the correct magnetic dipole moment of the electron

$$\mu = -\frac{e}{m}\mathbf{s} = -g\,\frac{e}{2m}\mathbf{s}\,, \tag{4.564}$$

where the factor $g = 2$ was introduced phenomenologically ad hoc to describe the anomalous Zeeman effect. Now, this is a prediction of the Dirac equation.

If we consider the system in a weak static magnetic field, $\mathbf{B} = B\hat{\mathbf{k}}$, in the z direction.

We have

$$\mathbf{A} = \frac{1}{2}\mathbf{B} \wedge \mathbf{r} = \frac{B}{2}\begin{pmatrix} -y \\ x \\ 0 \end{pmatrix} \tag{4.565}$$

Since B is a weak field, we neglect the term A^2 in Eq. (4.563) and find finally

$$H = eA^0 + \frac{1}{2m}\left[p^2 - e(\mathbf{p}\cdot\mathbf{A} + \mathbf{A}\cdot\mathbf{p})\right] - \frac{e}{m}\frac{\sigma}{2}\cdot\mathbf{B}\,. \tag{4.566}$$

In the case at hand we have that $[p^i, A^i] = 0$ and therefore

$$(\mathbf{p}\cdot\mathbf{A} + \mathbf{A}\cdot\mathbf{p}) = 2\mathbf{A}\cdot\mathbf{p} = B(xp_y - yp_x) = \mathbf{L}\cdot\mathbf{B}\,, \tag{4.567}$$

where $\mathbf{L}$ is the orbital angular momentum. Finally

$$H = \frac{p^2}{2m} + eA^0 - \frac{e}{2m}\left[\mathbf{L} + 2\mathbf{s}\right]\cdot\mathbf{B}\,, \tag{4.568}$$

that gives a good description of the Zeeman effect.

The fine structure of the hydrogen atom

Let us now consider the case of a central potential (Hydrogen atom) such that

$$\mathbf{A} = \mathbf{0}\,, e \quad A^0 = V(r) = -\frac{\alpha}{r}\,, \tag{4.569}$$

where $\alpha \sim \frac{1}{137}$ is the fine structure constant. Eqs. (4.554) become

$$\begin{cases} (E - V(r))\tilde{\phi} = \sigma\cdot\mathbf{p}\tilde{\chi} \\ (E - V(r) + 2m)\tilde{\chi} = \sigma\cdot\mathbf{p}\tilde{\phi} \end{cases} \tag{4.570}$$

Moreover, let us expand in the non relativistic regime keeping consistently terms of the order p^2/m^2 correcting the energy $p^2/(2m)$ and V. We then keep up to terms in p^4/m^3 and p^2V/m^2. This will give rise to the "relativistic corrections" to the non relativistic treatment of the hydrogen atom. The equation for $\tilde{\chi}$ now becomes

$$\tilde{\chi} = \frac{\sigma\cdot\mathbf{p}}{(E - V(r) + 2m)}\tilde{\phi} \simeq \frac{1}{2m}\left(1 - \frac{E - V(r)}{2m}\right)\sigma\cdot\mathbf{p}\,\tilde{\phi}\,. \tag{4.571}$$

There is another correction to take into account (see Maggiore) according to which the wave function is corrected by a factor

$$\tilde{\phi} = \left(1 - \frac{p^2}{8m^2}\right)\psi\,. \tag{4.572}$$

Finally we have

$$\tilde{\chi} = \frac{\sigma\cdot\mathbf{p}}{(E - V(r) + 2m)}\tilde{\phi} \simeq \frac{1}{2m}\left(1 - \frac{E - V(r)}{2m}\right)\sigma\cdot\mathbf{p}\left(1 - \frac{p^2}{8m^2}\right)\psi\,, \tag{4.573}$$

$$\simeq \frac{1}{2m}\left[\sigma\cdot\mathbf{p}\left(1 - \frac{p^2}{8m^2}\right) + \frac{E - V(r)}{2m}\sigma\cdot\mathbf{p}\right]\psi\,. \tag{4.574}$$

Substituting Eq. (4.572) and Eq. (4.574) in the first equation of (4.570), we find

$$(E - V(r))\left(1 - \frac{p^2}{8m^2}\right)\psi = \sigma\cdot\mathbf{p}\frac{1}{2m}\left[\sigma\cdot\mathbf{p}\left(1 - \frac{p^2}{8m^2}\right) + \frac{E - V(r)}{2m}\sigma\cdot\mathbf{p}\right]\psi\,, \tag{4.575}$$

$$= \left[\frac{p^2}{2m}\left(1-\frac{p^2}{8m^2}\right)+\frac{Ep^2}{4m^2}-\frac{\sigma\cdot\mathbf{p}\,V(r)\,\sigma\cdot\mathbf{p}}{4m^2}\right]\psi\,. \tag{4.576}$$

On the left-hand side we have

$$(E-V(r))\left(1-\frac{p^2}{8m^2}\right)\psi = \left((E-V(r))-\frac{Ep^2}{8m^2}+\frac{Vp^2}{8m^2}\right)\psi \tag{4.577}$$

and, neglecting terms of order Ep^4/m^4, we can write

$$\frac{Ep^2}{8m^2}\psi \simeq \frac{p^2}{8m^2}\left(\frac{p^2}{2m}+V(r)\right)\psi\,. \tag{4.578}$$

Finally, we get

$$i\frac{\partial}{\partial t}\psi = H\psi\,, \tag{4.579}$$

where

$$H = \left(\frac{p^2}{2m}+V(r)\right)-\frac{p^4}{8m^3}+\frac{1}{4m^2}\left[\sigma\cdot\mathbf{p}\,V(r)\,\sigma\cdot\mathbf{p}-\frac{1}{2}\left(p^2V(r)+V(r)p^2\right)\right]. \tag{4.580}$$

Let us analyse the two terms in the square brackets. We have

$$\begin{aligned}
\sigma\cdot\mathbf{p}\,V(r)\,\sigma\cdot\mathbf{p} &= \sigma^i\sigma^j\,p^iV(r)p^j = \sigma^i\sigma^j\,\left(i\partial^iV(r)p^j+V(r)p^ip^j\right), && (4.581)\\
&= \sigma^i\sigma^j\,\left(ieE^ip^j+V(r)p^ip^j\right), && (4.582)\\
&= |\,\text{since } \sigma^i\sigma^j=\delta^{ij}+i\epsilon^{ijk}\sigma^k\,| \\
&= ie\mathbf{E}\cdot\mathbf{p}+V(r)p^2-e\,\sigma\cdot(\mathbf{E}\wedge\mathbf{p})\,, && (4.583)
\end{aligned}$$

where we introduced the electric field $\partial^iV(r)=eE^i$ and we used the fact that $i\epsilon^{ijk}\sigma^kV(r)p^ip^j=0$ for the antisymmetry of the epsilon tensor. Moreover, we have

$$\begin{aligned}
p^2V(r)+V(r)p^2 &= p^ip^iV(r)+V(r)p^2 = p^i\left(ieE^i+Vp^i\right)+V(r)p^2\,, && (4.584)\\
&= ie\mathbf{p}\cdot\mathbf{E}+ie\mathbf{E}\cdot\mathbf{p}+2V(r)p^2\,, && (4.585)\\
&= e\nabla\cdot\mathbf{E}+2ie\mathbf{E}\cdot\mathbf{p}+2V(r)p^2\,. && (4.586)
\end{aligned}$$

Finally

$$\frac{1}{4m^2}\left[\sigma\cdot\mathbf{p}\,V(r)\,\sigma\cdot\mathbf{p}-\frac{1}{2}\left(p^2V(r)+V(r)p^2\right)\right] = \frac{1}{4m^2}\left[-\frac{e}{2}\nabla\cdot\mathbf{E}-e\,\sigma\cdot(\mathbf{E}\wedge\mathbf{p})\right]. \tag{4.587}$$

Since

$$e\mathbf{E} = -\nabla V(r) = -\mathbf{r}\left(\frac{1}{r}\frac{dV(r)}{dr}\right), \tag{4.588}$$

therefore

$$-\frac{e}{4m^2}\,\sigma\cdot(\mathbf{E}\wedge\mathbf{p}) = -\frac{1}{2m^2}\left(\frac{1}{r}\frac{dV(r)}{dr}\right)\frac{\sigma}{2}\cdot(-\mathbf{r}\wedge\mathbf{p}) = \frac{1}{2m^2}\left(\frac{1}{r}\frac{dV(r)}{dr}\right)\mathbf{s}\cdot\mathbf{L}\,, \tag{4.589}$$

where $\mathbf{L}=\mathbf{r}\wedge\mathbf{p}$ is the orbital angular momentum.

The resulting hamiltonian is

$$\begin{aligned}
H &= \left(\frac{p^2}{2m}+V(r)\right)-\frac{p^4}{8m^3}+\frac{1}{2m^2}\left(\frac{1}{r}\frac{dV(r)}{dr}\right)\mathbf{s}\cdot\mathbf{L}-\frac{e}{8m^2}(\nabla\cdot\mathbf{E})\,, && (4.590)\\
&= H_0+H_{pert}\,, && (4.591)
\end{aligned}$$

where

$$H_0 = \left(\frac{p^2}{2m} + V(r) \right) \tag{4.592}$$

is the central-potential hamiltonian of a spinless particle, with energy levels

$$E_n = -\frac{m\alpha^2}{2n^2} \, , \tag{4.593}$$

and eigenfunctions

$$\psi_{nlm} = R_{nl}(r)\, Y_l^m(\theta, \phi) \, . \tag{4.594}$$

The term

$$H_{pert} = -\frac{p^4}{8m^3} + \frac{1}{2m^2} \left(\frac{1}{r} \frac{dV(r)}{dr} \right) \mathbf{s} \cdot \mathbf{L} - \frac{e}{8m^2} (\nabla \cdot \mathbf{E}) \tag{4.595}$$

can be treated in perturbation theory and it is constituted by the so-called "relativistic correction"

$$H_r = -\frac{p^4}{8m^3} \, , \tag{4.596}$$

the spin-orbit interaction

$$H_{SO} = \frac{1}{2m^2} \left(\frac{1}{r} \frac{dV(r)}{dr} \right) \mathbf{s} \cdot \mathbf{L} \, , \tag{4.597}$$

and the Darwin term

$$H_D = -\frac{e}{8m^2} (\nabla \cdot \mathbf{E}) \, . \tag{4.598}$$

The hamiltonian (4.595) does not resolve completely the degeneracy of the energy levels of the hydrogen atom[13]. In particular the two levels $2S_{\frac{1}{2}}$ and $2P_{\frac{1}{2}}$ are still degenerate, while in Nature we register a small difference, of about 1000 MHz (Lamb shift). This difference can be accounted for treating correctly the system in quantum field theory, calculating higher-order QED quantum corrections.

4.3.15 Parity

So far we considered proper Lorentz transformations. In this section we will see how discontinuous transformations, as Parity or Time Reversal, act on the field.

Parity is a Lorentz tranformation. Moreover, it can be represented via a unitary operator. On the space-time point, Parity acts as follows:

$$\begin{cases} \mathbf{x} & \to & -\mathbf{x} \\ t & \to & t \end{cases} \, . \tag{4.599}$$

Therefore, in matrix notation we have

$$\Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} , \tag{4.600}$$

which is basically the metric.

If we want that Dirac's equation is invariant under Parity transformations, we have to require that Eq. (4.297) holds for $S(\Lambda_P)$ as well:

$$S^{-1}(\Lambda_P)\gamma_\nu S(\Lambda_P) = \Lambda^\mu_{P\,\nu} \gamma^\nu \, , \tag{4.601}$$

[13] This is the case also for the complete Dirac's equation. It is not a problem of the fact that we afforded the calculation perturbatively.

or, in components, multiplying on the left by $S(\Lambda_P)$ and bringing both terms of the equation on the l.h.s.

$$[\gamma^0, S(\Lambda_P)] = 0 \quad \text{e} \quad [\gamma^i, S(\Lambda_P)]_+ = 0\,. \tag{4.602}$$

Eqs. (4.602) are satisfied by the following choice:

$$S(\Lambda_P) = \eta_P \gamma^0\,, \tag{4.603}$$

where η_P is a constant to be determined. Note that we have to have

$$[S(\Lambda_P)]^2 = 1\,, \tag{4.604}$$

since applying twice Parity we would like to find the identity operator. Therefore Eq. (4.604) implies

$$\eta_P = \pm 1\,, \tag{4.605}$$

i.e. η_P is a phase, that for the moment we put $= 1$:

$$S(\Lambda_P) = \gamma^0\,. \tag{4.606}$$

We note that such choice respects the fact that $S(\Lambda_P)$ must be unitary. In fact

$$[S(\Lambda_P)]^\dagger = \gamma^{0\,\dagger} = \gamma^0 = S(\Lambda_P)\,. \tag{4.607}$$

The interacting Dirac's equation is indeed covariant under the Parity transformation. In fact we have

$$\psi'(X') = S(\Lambda_P)\psi(X)\,, \tag{4.608}$$

where $X'^\mu = (t, -\mathbf{x})$, and

$$0 = (i\,\partial\!\!\!/ - e\,A\!\!\!/ - m)\psi(X) = (i\,\partial\!\!\!/ - e\,A\!\!\!/ - m)S^{-1}(\Lambda_P)\psi'(X')\,. \tag{4.609}$$

Multiplying by $S(\Lambda_P)$ on the left we have

$$S(\Lambda_P) i\,\partial\!\!\!/ S^{-1}(\Lambda_P) = i\,\partial\!\!\!/'\,, \tag{4.610}$$

$$S(\Lambda_P)\,A\!\!\!/ S^{-1}(\Lambda_P) = A\!\!\!/'\,, \tag{4.611}$$

since A^0 does not change under parity but $\mathbf{A}$ changes sign and $\partial'^0 = \partial^0$, $\partial'^i = -\partial^i$.

Finally we have

$$(i\,\partial\!\!\!/' - e\,A\!\!\!/' - m)\psi'(X') = 0\,. \tag{4.612}$$

4.3.16 Time Reversal

Time Reversal invariance means that if we have a sequence of observations made on a state described by a certain wave function and we invert the temporal order of the sequence, we still find a physically realizable sequence of observations.

The action of Time Reversal on the space-time point is

$$\begin{cases} \mathbf{x} & \to & \mathbf{x} \\ t & \to & -t \end{cases}\,, \tag{4.613}$$

such that in matrix notation we have

$$\Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\,, \tag{4.614}$$

which is $-\eta^{\mu\nu}$.

We know that Time Reversal has to be represented by an anti-unitary anti-linear operator, therefore via an operator $\tilde{U}$ such that

$$\begin{cases} \tilde{U}^\dagger \tilde{U} = 1 \\ \tilde{U}(\alpha|\phi\rangle + \beta|\psi\rangle) = \alpha^* \tilde{U}|\phi\rangle + \beta^* \tilde{U}|\psi\rangle \end{cases} . \tag{4.615}$$

Therefore $\tilde{U}$ is defined through

$$\langle \phi | \tilde{U}^\dagger \psi \rangle = \langle U\phi|\psi\rangle^* = \langle \psi | U\phi\rangle \,. \tag{4.616}$$

Such operator can be constructed as a product of a unitary operator U times the operation of complex conjugation $\mathcal{K}$:

$$\tilde{U} = U\,\mathcal{K} \,. \tag{4.617}$$

In fact, we have

$$\tilde{U}(\alpha|\phi\rangle + \beta|\psi\rangle) = U\,\mathcal{K}(\alpha|\phi\rangle+\beta|\psi\rangle) = U\,(\alpha^*\mathcal{K}|\phi\rangle+\beta^*\mathcal{K}|\psi\rangle) = \alpha^* U\,\mathcal{K}|\phi\rangle+\beta^* U\,\mathcal{K}|\psi\rangle = \alpha^*\tilde{U}|\phi\rangle+\beta^*\tilde{U}|\psi\rangle \,. \tag{4.618}$$

If

$$|\phi'\rangle = \tilde{U}|\phi\rangle \,, \qquad \text{and} \quad |\psi'\rangle = \tilde{U}|\psi\rangle \,, \tag{4.619}$$

then we have

$$\begin{aligned}
\langle \phi'|\psi'\rangle &= \langle \phi'|\big(U\mathcal{K}\sum_\beta |\beta\rangle\langle\beta|\psi\rangle\big) \,, && (4.620)\\
&= \langle \phi'|\sum_\beta \langle\psi|\beta\rangle\, U\mathcal{K}|\beta\rangle \,, && (4.621)\\
&= \sum_{\beta\beta'} \langle\beta'|\phi\rangle\langle\beta'|U^\dagger U|\beta\rangle\langle\psi|\beta\rangle \,, && (4.622)\\
&= \sum_{\beta\beta'} \langle\psi|\beta\rangle\,\langle\beta'|\beta\rangle\,\langle\beta'|\phi\rangle \,, && (4.623)\\
&= \langle\psi|\phi\rangle = \langle\phi|\psi\rangle^* \,, && (4.624)
\end{aligned}$$

as it should be.

The representation of Time Reversal on the Dirac field can be found imposing the invariance of Dirac's equation. Let us consider Dirac's equation in the original form (replacing nevertheless α^i and β matrices with the gamma's)

$$i\frac{\partial}{\partial t}\psi = H\,\psi \,, \tag{4.625}$$

where, including electromagnetic interactions, we have

$$H = eA^0 + \gamma^0\gamma^i(-i\partial_i - eA_i) + \gamma^0 m \,. \tag{4.626}$$

We define the Time Reversal operator $K = T\mathcal{K}$, where $T^\dagger T = \mathbb{1}$, such that

$$\psi'(X') = K\,\psi(X) \,, \qquad \psi(X) = K^{-1}\psi'(X') \tag{4.627}$$

where $X' = (-t, \mathbf{x})$. Then

$$i\frac{\partial}{\partial t}K^{-1}\psi'(X') = -i\frac{\partial}{\partial t'}K^{-1}\psi'(X') = H\,K^{-1}\psi'(X') \,. \tag{4.628}$$

Multiplying on the left by K and remembering the anti-linear nature of K we get

$$\frac{\partial}{\partial t'}K(-i)K^{-1}\psi'(X') = i\frac{\partial}{\partial t'}\psi'(X') = K\,H\,K^{-1}\psi'(X')\,. \tag{4.629}$$

We now should have

$$K\,H\,K^{-1} = T\,H^*(t)\,T^{-1} = H(t')\,. \tag{4.630}$$

We have

$$H^*(t) = eA^0(t) + (\gamma^0\gamma^i)^*(i\partial_i - eA_i(t)) + \gamma^0 m\,. \tag{4.631}$$

Therefore, since

$$A^0(t) = A^0(t')\,, \quad \mathbf{A}(t) = -\mathbf{A}(t')\,, \tag{4.632}$$

because A^0 is generated by a static charge distribution, while $\mathbf{A}$ is generated by a current (and therefore when we invert the sign of time the current flows in the opposite direction and changes sign to the vector potential), we must have

$$T\,H^*(t)\,T^{-1} = eA^0(t') + T(\gamma^0\gamma^{i\,*})T^{-1}(i\partial_i + eA_i(t')) + T\gamma^0T^{-1}m\,. \tag{4.633}$$

This gives the two conditions

$$T\gamma^0T^{-1} = \gamma^0\,, \tag{4.634}$$

$$T(\gamma^0\gamma^{i\,*})T^{-1} = -\gamma^0\gamma^i\,. \tag{4.635}$$

Using the first equation into the second

$$T(\gamma^0\gamma^{i\,*})T^{-1} = T\gamma^0T^{-1}(T\gamma^{i\,*}T^{-1}) = \gamma^0(T\gamma^{i\,*}T^{-1}) = -\gamma^0\gamma^i\,, \tag{4.636}$$

we find that we have to impose

$$T\gamma^{i\,*}T^{-1} = -\gamma^i\,. \tag{4.637}$$

Eq. (4.637) is satisfied by

$$T = i\gamma^1\gamma^3\,, \tag{4.638}$$

which is an operator such that

$$T^\dagger T = \mathbb{1}\,, \quad T^2 = \mathbb{1}\,, \tag{4.639}$$

as it should.

We could have used directly the relation (4.297) in order to find $K = S(\Lambda_T)$. However, we have to remember that in order to find relation (4.297) we already assumed the operator $S(\Lambda)$ to be unitary and linear. In fact, we commuted without a sign the "i" that multiplies the gamma's. For Time Reversal, we should use the proper relation

$$S^{-1}(\Lambda_T)i\gamma^\nu S(\Lambda_T) = i(\Lambda_T)^\nu_\mu\gamma^\mu\,. \tag{4.640}$$

This means

$$-iT^{-1}\gamma^{\nu\,*}T = i(\Lambda_T)^\nu_\mu\gamma^\mu \tag{4.641}$$

and therefore

$$T^{-1}\gamma^0T = \gamma^0\,, \tag{4.642}$$

$$T^{-1}\gamma^{i\,*}T = -\gamma^i\,. \tag{4.643}$$

The solution of Eqs. (4.642,4.643) is again Eq. (4.638), since $\gamma^{0\,*} = \gamma^0$, $\gamma^{1\,*} = \gamma^1$, $\gamma^{2\,*} = -\gamma^2$, $\gamma^{3\,*} = \gamma^3$ and since $T^2 = \mathbb{1}$.

4.3.17 Charge Conjugation

There is another discontinuous symmetry which plays a big role in QFT: Charge Conjugation. So to say, it is the symmetry that relates positive energy with negative energy solutions (or particles to anti-particles, when we speak about states).

Let us start with the interacting Dirac equation

$$(i \not\partial - e \not A - m)\, \psi(X) \,=\, 0\,. \tag{4.644}$$

We look for a transformation $\mathcal{C}$ such that the solution of Eq. (4.644), $\psi(X)$, is transformed in $\psi_C(X)$, which will have to represent a fermion with the same mass as $\psi(X)$, but with opposite electric charge:

$$\psi(X) \longrightarrow \psi_C(X) \,=\, \mathcal{C}\, (\psi(X))\,. \tag{4.645}$$

The field $\psi_C(X)$ will have to satisfy the Dirac equation for a field in which $-e$ became $+e$:

$$(i \not\partial + e \not A - m)\, \psi_C(X) \,=\, 0\,. \tag{4.646}$$

We require that this transformation is such that if it acts twice, it brings the field $\psi(X)$ back to its original configuration, apart from a possible phase

$$\mathcal{C}\, (\mathcal{C}\, (\psi(X))) = \eta_C\, \psi(X)\,. \tag{4.647}$$

In order to transform Eq. (4.644) into Eq. (4.646), we need to change the relative sign between the terms $i \not\partial$ and $-e \not A$. This can be done, taking the adjoint of Eq. (4.644):

$$\psi^\dagger \left(i \not\partial^\dagger + e \not A^\dagger + m\right) \,=\, 0\,. \tag{4.648}$$

Multiplying on the r.h.s. by γ_0, and remembering that $\gamma^{\mu\,\dagger}\gamma_0 \,=\, \gamma_0\gamma^\mu$, we find:

$$\bar\psi\, (i \not\partial + e \not A + m) \,=\, 0\,. \tag{4.649}$$

If we take the transposed of Eq. (4.649), we find:

$$\left(i \not\partial^t + e \not A^t + m\right)\, \bar\psi^t \,=\, 0\,. \tag{4.650}$$

If we now would find a transformation C (a 4×4 matrix acting in the space of psinors), such that:

$$C\, \gamma_\mu^t\, C^{-1} \,=\, -\,\gamma_\mu\,, \tag{4.651}$$

we could multiply on the l.h.s. Eq. (4.650) by C and, changing an overall sign, we would obtain:

$$(i \not\partial + e \not A - m)\, C\, \bar\psi^t \,=\, (i \not\partial + e \not A - m)\, \psi_C(X) \,=\, 0\,, \tag{4.652}$$

where

$$\psi_C(X) \,=\, \eta_C\, C\, \bar\psi^t\,, \tag{4.653}$$

with η_C a phase factor.

Let us look for C in our representation for the γ matrices, in which γ^0 is diagonal:

$$\gamma_0^t \,=\, \gamma_0\,; \quad \gamma^{1\,t} \,=\, -\gamma^1\,; \quad \gamma^{2\,t} \,=\, \gamma^2\,; \quad \gamma^{3\,t} \,=\, -\gamma^3\,. \tag{4.654}$$

Since Eq. (4.651) should be valid, C has to anti-commute with γ_0 and γ^2 and to commute with γ^1 and γ^3. This means that we have to have

$$C \,=\, i\,\gamma^2\gamma^0 \,=\, \begin{pmatrix} 0 & -i\sigma^2 \\ -i\sigma^2 & 0 \end{pmatrix}\,. \tag{4.655}$$

In fact, using Eq. (4.655) we have

$$C\,\gamma^\mu \;=\; \begin{cases} -\gamma^\mu\, C & \text{se } \mu = 0, 2\ ; \\ \ \ \gamma^\mu\, C & \text{se } \mu = 1, 3\ . \end{cases} \tag{4.656}$$

The matrix C fullfils the following properties:

$$C^\dagger \;=\; C^t \;=\; C^{-1} \;=\; -C\,. \tag{4.657}$$

Therefore, we found the transformation $\mathcal{C}$ such that, if $\psi(X)$ is solution of the Dirac equation for the "electron", then

$$\psi_C(X) = \mathcal{C}\ (\psi(X)) = \eta_C\, C\, \bar{\psi}^t = \eta_C\, i\gamma^2\psi^* \tag{4.658}$$

is a solution for the Dirac equation of the "positron" (and vice versa).

Let us see how $\mathcal{C}$ acts on a field with a given energy and spin:

$$\psi'(X) \;=\; \left(\frac{\pm\, \not{P} + m}{2m}\right)\left(\frac{1+\gamma_5\ \not{n}}{2}\right)\psi(X)\,. \tag{4.659}$$

We have

$$\psi'_C(X) \;=\; C\,\bar{\psi}'^t(X) \;=\; \eta_C\, C\,\gamma_0 \left(\frac{\pm\, \not{P}^* + m}{2m}\right)\left(\frac{1+\gamma_5\ \not{n}^*}{2}\right)\psi^*(X)\,. \tag{4.660}$$

Since, moreover, the following equations hold

$$\gamma_0\gamma^{\mu\,*} \;=\; \gamma^{\mu\,t}\gamma_0\,;\quad \gamma_0\gamma_5 \;=\; -\gamma_5\gamma_0\,;\quad [C,\gamma_5] \;=\; 0\,, \tag{4.661}$$

we get

$$\psi'_C(X) \;=\; \eta_C\, C \left(\frac{\pm\, \not{P}^t + m}{2m}\right)\left(\frac{1-\gamma_5\ \not{n}^t}{2}\right)\gamma_0\psi^*(X) \;= \tag{4.662}$$

$$=\; \eta_C \left(\frac{\mp\, \not{P} + m}{2m}\right)\left(\frac{1+\gamma_5\ \not{n}}{2}\right) C\,\bar{\psi}^t(X) \;= \tag{4.663}$$

$$=\; \eta_C \left(\frac{\mp\, \not{P} + m}{2m}\right)\left(\frac{1+\gamma_5\ \not{n}}{2}\right)\psi_C(X)\,. \tag{4.664}$$

We see that $\psi'_C(X)$ is described by the same P^μ and n^μ of $\psi'(X)$, but the sign of the energy is reversed. This means that also the spin is reversed. In particular, let us consider a plane wave solution, in the rest frame of the particle, with negative energy and spin down:

$$\psi(X) \;=\; u^{(4)}(m,\mathbf{0})\, e^{imt} \;=\; e^{imt}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}\,. \tag{4.665}$$

Its charge conjugated solution will be

$$\psi_C(X) \;=\; i\eta_C\gamma^2\,\psi^* \;=\; \eta_C e^{-imt}\begin{pmatrix}0&0&0&1\\0&0&-1&0\\0&-1&0&0\\1&0&0&0\end{pmatrix}\begin{pmatrix}0\\0\\0\\1\end{pmatrix} \;=\; \eta_C\, e^{-imt}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\,, \tag{4.666}$$

i.e. a solution with positive energy and spin up.

The transformation $\mathcal{C}$, shows a symmetry of Dirac equation. In fact, if we consider the electromagnetic field $A^\mu(X)$, we can understand how, under charge conjucation, it simply gets a minus sign[14]. Therefore, the following transformation on the Dirac equation

$$\begin{cases} \psi(X) \xrightarrow{\mathcal{C}} \psi_C(X) \\ A^\mu(X) \longrightarrow -A^\mu(X) \end{cases} \tag{4.667}$$

gives a formal invariance of Eq. (4.644).

4.3.18 $\mathcal{PCT}$ transformation

Although Parity, Charge Conjugation and Time Reversal are not basic symmetries for Physics, it can be demonstrated that for a local Quantum Field Theory the product of the three, $\mathcal{PCT}$, is indeed a symmetry of Physics.

In the case of electromagnetic interactions, the three transformations are individually a symmetry for the the Lagrangian density (and therefore of the action). Therefore, it is not surprising that $\mathcal{PCT}$ is a symmetry. However, this holds also in cases in which they are individually broken.

Let us find the action of $\mathcal{CPT} = \Theta$ on our field $\psi(X)$. We have (neglecting the possible phases)

$$\begin{aligned} \psi'(X') &= \Theta\psi(X) = \mathcal{PCT}\psi(X) = \mathcal{PC}i\gamma^1\gamma^3\psi^*(X)\,, & (4.668)\\ &= \mathcal{P}i\gamma^2\gamma^0\gamma^0\left(i\gamma^1\gamma^3\psi^*(X)\right)^* = \mathcal{P}i\gamma^2(-i)\gamma^{1*}\gamma^{3*}\psi(X) = \mathcal{P}\gamma^2\gamma^1\gamma^3\psi(X)\,, & (4.669)\\ &= \gamma^0\gamma^2\gamma^1\gamma^3\psi(X) = i\gamma_5\psi(X)\,. & (4.670) \end{aligned}$$

Moreover, under $\mathcal{CPT}$ the point of the Minkowski space goes in $X^\mu \to -X^\mu$ and the electromagnetic field transforms as $A^\mu(X) \to A^{\mu\prime}(X') = -A^\mu(X)$. Therefore, the Dirac equation does not change in form after the transformation Θ:

$$\begin{aligned} 0 &= (i\,\partial\!\!\!/ - e\,A\!\!\!/ - m)\psi(X) = (i\,\partial\!\!\!/ - e\,A\!\!\!/ - m)(-i\gamma_5)\psi'(X')\,, & (4.671)\\ &= (-i\gamma_5)(i\,\partial\!\!\!/' - e\,A\!\!\!/'(X') - m)\psi'(X') & (4.672) \end{aligned}$$

and multiplying by $i\gamma_5$ we get $(i\,\partial\!\!\!/' - e\,A\!\!\!/'(X') - m)\psi'(X') = 0$.

Consequences of the CPT theorem ...

4.3.19 Massless fermionic field: the neutrino

Oltre ai fermioni di massa $m \neq 0$, nel Modello Standard delle interazioni fondamentali sono previsti anche fermioni che sperimentalmente sembrano avere massa nulla: i neutrini. Per essi, l'equazione di Dirac si riduce alla:

$$i\,\partial\!\!\!/\,\psi_\nu(X) = 0\,, \tag{4.673}$$

dove il pedice ν sta per *neutrino*.

Il fatto che il termine di massa non sia presente nella (**??**) permette di disaccoppiare i due spinori a due componenti $\phi_R(X)$ e $\phi_L(X)$, mediante i quali avevamo costruito lo spinore di Dirac[15] $\psi_\nu(X)$. Inoltre, come abbiamo già accennato nel primo capitolo, la massa nulla del campo fermionico in considerazione fa sì che le sue polarizzazioni possibili siano date dagli autovalori dell'elicità: $\pm\frac{1}{2}$ lungo la direzione del moto. È conveniente, allora, non utilizzare per le γ^μ la rappresentazione (**??**), ma

[14]This is indeed the case, because charge conjugation flips the sign of the charges. Therefore, since A^0 comes from a static distribution of charge, if the sign of this charge changes, we have to have that $A^0 \to -A^0$. The same happens for $\mathbf{A}$. It comes from a current and if we change the sign to the charges that generate the current, this changes sign to the vector potential generated by the current, $\mathbf{A} \to -\mathbf{A}$.

[15]Cfr. Capitolo 1.

introdurre la *rappresentazione di Weyl* o *rappresentazione chirale*, nella quale è diagonale la γ_5 (legata, come vedremo, all'elicità):

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad \gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.674}$$

Se poniamo, allora:

$$\psi_\nu(X) = \begin{pmatrix} \phi_L(X) \\ \phi_R(X) \end{pmatrix}, \tag{4.675}$$

la (4.673) si divide in due equazioni differenziali disaccoppiate per la componente destra, $\phi_R(X)$, e sinistra, $\phi_L(X)$, di $\psi_\nu(X)$:

$$i\partial^0 \phi_L(X) - i\,\sigma \cdot \nabla \phi_L(X) = 0, \tag{4.676}$$

$$i\partial^0 \phi_R(X) + i\,\sigma \cdot \nabla \phi_R(X) = 0, \tag{4.677}$$

ovvero, ponendo $\mathbf{p} = -i\nabla$:

$$i\partial^0 \phi_L(X) + \sigma \cdot \mathbf{p}\,\phi_L(X) = 0, \tag{4.678}$$

$$i\partial^0 \phi_R(X) - \sigma \cdot \mathbf{p}\,\phi_R(X) = 0. \tag{4.679}$$

L'operatore $\hat{h} = \frac{\sigma\cdot\mathbf{p}}{2\|\mathbf{p}\|}$ è detto *elicità* del neutrino e, come si vede, rappresenta in pratica la componente dello spin lungo la direzione del moto ($\frac{\mathbf{p}}{\|\mathbf{p}\|}$).

Siccome il neutrino ha massa nulla, avrà un quadrivettore energia-impulso di tipo luce, $P^2 = 0$, da cui risulta che:

$$E = \pm\|\mathbf{p}\|. \tag{4.680}$$

Consideriamo l'Eq. (4.678) in cui $\phi_L(X)$ sia un'onda piana ad energia positiva $E = \|\mathbf{p}\|$ (negativa $E = -\|\mathbf{p}\|$):

$$\phi_L(X) = \phi_L^0\, e^{\mp iP_\mu X^\mu}, \tag{4.681}$$

cioè quello che identificheremo con una "particella" ("antiparticella"). Sostituendo (4.681) in (4.678) si ottiene:

$$\hat{h}\,\phi_L(X) = \mp\frac{1}{2}\,\phi_L(X). \tag{4.682}$$

Questo vuol dire che l'Eq. (4.678) descrive neutrini ad elicità $-\frac{1}{2}$ (*neutrini left-handed*) e neutrini ad energia negativa ed elicità $\frac{1}{2}$. Consistentemente con la quantizzazione del campo di Dirac, che affronteremo nel prossimo capitolo, la seconda ipotesi è analoga ad asserire che l'Eq. (4.678) descrive anche antineutrini (cioè antiparticelle ad energia positiva) con elicità $\frac{1}{2}$ (*antineutrini right-handed*).

Se facciamo lo stesso ragionamento per l'Eq. (4.679), troviamo che questa descrive antineutrini ad elicità $-\frac{1}{2}$ (*antineutrini left-handed*) e neutrini ad elicità $\frac{1}{2}$ (*neutrini right-handed*).

A questo punto abbiamo a che fare con due funzioni d'onda perfettamente analoghe da un punto di vista teorico. Sperimentalmente, però, si può vedere che in natura sono presenti soltanto neutrini left-handed ed antineutrini right-handed. Inoltre, siccome il neutrino interviene soltanto nelle interazioni deboli e i due stati adesso menzionati non si possono connettere attraverso una trasformazione di parità, le interazioni deboli violano la parità.

Introduciamo i due seguenti prioettori:

$$P_L = \frac{(1-\gamma_5)}{2}, \qquad P_R = \frac{(1+\gamma_5)}{2}. \tag{4.683}$$

Come i può verificare facilmente, P_L e P_R godono di tutte le proprietà peculiari di un proiettore:

$$P_{L,R}^2 = P_{L,R}, \quad P_L + P_R = 1, \quad P_L P_R = 0, \tag{4.684}$$

dove abbiamo usato la proprietà $\gamma_5^2 = 1$.

P_L e P_R proiettano rispettivamente su $\phi_L(X)$ e $\phi_R(X)$:

$$P_L\,\psi(X) = \frac{(1-\gamma_5)}{2}\,\psi(X) = \frac{1}{2}\begin{pmatrix}2 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}\phi_L\\ \phi_R\end{pmatrix} = \begin{pmatrix}\phi_L\\ 0\end{pmatrix}, \tag{4.685}$$

$$P_R\,\psi(X) = \frac{(1+\gamma_5)}{2}\,\psi(X) = \frac{1}{2}\begin{pmatrix}0 & 0\\ 0 & 2\end{pmatrix}\begin{pmatrix}\phi_L\\ \phi_R\end{pmatrix} = \begin{pmatrix}0\\ \phi_R\end{pmatrix}, \tag{4.686}$$

per cui, nelle interazioni deboli compariranno soltanto le espressioni:

$$\frac{(1-\gamma_5)}{2}\,\psi_\nu(X)\,, \quad \text{e} \quad \bar{\psi}_\nu(X)\,\frac{(1+\gamma_5)}{2}\,. \tag{4.687}$$

La prima distrugge neutrini left-handed e crea antineutrini right-handed, mentre la seconda crea neutrini left-handed e distrugge antineutrini right-handed.

La teoria così viluppata si chiama *teoria del neutrino a due componenti* e fu proposta da Weyl nel 1929 e ripresa solo nel 1957, quando evidenze sperimentali confermarono che le interazioni deboli violano la parità.

A questo punto bisogna puntualizzare alcune cose.

- Da un punto di vista di teoria dei gruppi, il fatto che sia possibile descrivere il neutrino con uno spinore a due componenti deriva dal fatto che la rappresentazione spinoriale del Gruppo di Poincaré per massa nulla è riducibile nelle due rappresentazioni irriducibili ϕ_R e ϕ_L del Gruppo di Lorentz che avevamo incontrato nel primo capitolo. Se nella lagrangiana $\mathcal{L}$ è presente, invece, un termine di massa $m\bar{\psi}\psi$, questo mescola le due componenti ϕ_R e ϕ_L in un termine misto e $\mathcal{L}$ non è più invariante separatamente sotto i due tipi di trasformazioni di Lorentz.

- L'operatore γ_5 è detto *chiralità*. Per campi a massa nulla, la chiralità e l'elicità coincidono, ma questo non è vero per campi massivi. Si riottiene l'uguaglianza nel caso di alte energie, cioè quando la massa della particella è trascurabile in confronto alla sua energia. Questa osservazione fa comodo poiché in questo caso si può far ricorso alla simmetria chirale approssimata anche se stiamo trattando particelle massive, come l'elettrone, o i quarks e ricavare importanti relazioni fra gli elementi della matrice S ("regole di somma" in QCD).

- La teoria del neutrino a due componenti è invariante sotto trasformazioni chirali:

$$\psi_\nu(X) \to e^{i\gamma_5\Lambda}\,\psi_\nu(X)\,, \tag{4.688}$$

$$\bar{\psi}_\nu(X) \to \bar{\psi}_\nu(X)\,e^{i\gamma_5\Lambda}\,. \tag{4.689}$$

  La simmetria è violata da termini di massa.

4.4 Quantization of the Dirac Field

The quantization of the Dirac field should follow some basic principles, as in the case of the Klein-Gordon field. Firstly, we have to find a procedure that can accomodate the description of the particle and anti-particle states, both with positive energy. Moreover, we are dealing with fermionic states. Therefore, we would like to have a theory that incorporates directly the Pauli exclusion principle, or better the fact that fermions should obey Fermi-Dirac statistics.

The expression of the Dirac field in normal modes is the following:

$$\psi(X) = \sum_{\pm n}\int\frac{d^3p}{(2\pi)^{\frac{3}{2}}}\sqrt{\frac{m}{E}}\left[b(p,n)u(P,n)e^{-iP_\mu X^\mu} + d^*(p,n)v(P,n)e^{iP_\mu X^\mu}\right]. \tag{4.690}$$

If we want to quantize the field, we should promote to operators the coefficients $b(p,n)$, $d(p,n)$, $b^*(p,n)$, $d^*(p,n)$ (in particular $b^*(p,n)$, $d^*(p,n)$ will become $b^\dagger(p,n)$, $d^\dagger(p,n)$). However, we have to understand how they can act on a possible Fock space. In order to do that, let us see which is the expression of the hamiltonian

$$H = \int d^3X\, \mathcal{H} = i\int d^3X\, \psi^\dagger \frac{\partial}{\partial t}\psi\,, \tag{4.691}$$

in terms of $b(p,n)$, $d(p,n)$ and $b^*(p,n)$ $d^*(p,n)$.

We need to remember the ortogonality and completeness relations for the spinors. We have

$$\overline{u}(P,n)u(P,n') = -\overline{v}(P,n)v(P,n) = \delta_{nn'}\,, \tag{4.692}$$

$$u^\dagger(P,n)u(P,n') = -v^\dagger(P,n)v(P,n) = \frac{E}{m}\delta_{nn'}\,, \tag{4.693}$$

$$\overline{v}(P,n)u(P,n') = v^\dagger(P,n)u(\tilde{P},n') = u^\dagger(P,n)v(\tilde{P},n') = 0\,, \tag{4.694}$$

where $P^\mu = (E,\mathbf{p})$ and $\tilde{P}^\mu = (E,-\mathbf{p})$. We already demonstrated the first equations, we should demonstrate the last two. We have

$$\begin{aligned}
u^\dagger(P,n)v(\tilde{P},n') &= \frac{u^\dagger_{(\alpha)}(m,\mathbf{0})(\not{P}+m)^\dagger(-\tilde{\not{P}}+m)v_{(\beta)}(m,\mathbf{0})}{2m(E+m)}\,, &(4.695)\\
&= \frac{\overline{u}_{(\alpha)}(m,\mathbf{0})\gamma^0(\not{P}+m)^\dagger\gamma^0\gamma^0(-\tilde{\not{P}}+m)v_{(\beta)}(m,\mathbf{0})}{2m(E+m)}\,, &(4.696)\\
&= \frac{\overline{u}_{(\alpha)}(m,\mathbf{0})(\not{P}+m)(-\not{P}+m)\gamma^0 v_{(\beta)}(m,\mathbf{0})}{2m(E+m)}\,, &(4.697)\\
&= 0\,. &(4.698)
\end{aligned}$$

Moreover

$$\begin{aligned}
v^\dagger(P,n)u(\tilde{P},n') &= \frac{v^\dagger_{(\alpha)}(m,\mathbf{0})(-\not{P}+m)^\dagger(\tilde{\not{P}}+m)u_{(\beta)}(m,\mathbf{0})}{2m(E+m)}\,, &(4.699)\\
&= \frac{\overline{v}_{(\alpha)}(m,\mathbf{0})\gamma^0(-\not{P}+m)^\dagger\gamma^0\gamma^0(\tilde{\not{P}}+m)u_{(\beta)}(m,\mathbf{0})}{2m(E+m)}\,, &(4.700)\\
&= \frac{\overline{v}_{(\alpha)}(m,\mathbf{0})(-\not{P}+m)(\not{P}+m)\gamma^0 u_{(\beta)}(m,\mathbf{0})}{2m(E+m)}\,, &(4.701)\\
&= 0\,. &(4.702)
\end{aligned}$$

Now let us substitute Eq. (4.690) in Eq. (4.691). We have

$$\begin{aligned}
H &= i\int d^3X \sum_{\pm n,\pm n'}\int \frac{d^3p}{(2\pi)^{\frac{3}{2}}}\sqrt{\frac{m}{E}}\frac{d^3p'}{(2\pi)^{\frac{3}{2}}}\sqrt{\frac{m}{E'}}\Big[\\
&\qquad \Big(d(p,n)v^\dagger(P,n)e^{-iP_\mu X^\mu} + b^\dagger(p,n)u^\dagger(P,n)e^{iP_\mu X^\mu}\Big)\times\\
&\qquad \times(-iE')\Big(b(p',n')u(P',n')e^{-iP'_\mu X^\mu} - d^\dagger(p',n')v(P',n')e^{iP'_\mu X^\mu}\Big)\Big] &(4.703)\\
&= \sum_{\pm n,\pm n'}\int \frac{d^3p\, d^3p'}{(2\pi)^3}\frac{mE'}{\sqrt{EE'}}\int d^3X\Big[d(p,n)v^\dagger(P,n)b(p',n')u(P',n')e^{-i(P_\mu+P'_\mu)X^\mu}\\
&\qquad -d(p,n)v^\dagger(P,n)d^\dagger(p',n')v(P',n')e^{-i(P_\mu-P'_\mu)X^\mu}\\
&\qquad +b^\dagger(p,n)u^\dagger(P,n)b(p',n')u(P',n')e^{i(P_\mu-P'_\mu)X^\mu}\\
&\qquad -b^\dagger(p,n)u^\dagger(P,n)d^\dagger(p',n')v(P',n')e^{i(P_\mu+P'_\mu)X^\mu}\Big]\,, &(4.704)
\end{aligned}$$

$$
\begin{aligned}
&= \left| \text{the integral in } d^3X \text{ gives delta functions} \right| \\
&= \sum_{\pm n, \pm n'} \int d^3p\, d^3p' \frac{mE'}{\sqrt{EE'}} \Big[v^\dagger(P,n) u(\tilde{P}',n')\, d(p,n) b(p',n') e^{-i(E+E')t} \delta(\mathbf{p}+\mathbf{p}') \\
&\qquad - v^\dagger(P,n) v(P',n')\, d(p,n) d^\dagger(p',n') e^{-i(E-E')t} \delta(\mathbf{p}-\mathbf{p}') \\
&\qquad + u^\dagger(P,n) u(P',n')\, b^\dagger(p,n) b(p',n') e^{i(E-E')t} \delta(\mathbf{p}-\mathbf{p}') \\
&\qquad - u^\dagger(P,n) v(\tilde{P}',n')\, b^\dagger(p,n) d^\dagger(p',n') e^{i(E+E')t} \delta(\mathbf{p}+\mathbf{p}') \Big] , \qquad (4.705) \\
&= \left| \text{using the ortogonality relations and the delta's for the integration in } d^3p' \right| \\
&= \sum_{\pm n} \int d^3p E \left[b^\dagger(p,n) b(p,n) - d(p,n) d^\dagger(p,n) \right] . \qquad (4.706)
\end{aligned}
$$

Then as in the case of the charged KG field, we have to kinds of particle states. The peculiarity of the Dirac case, however, lies in the fact that there is the minus sign between the term with "particles" of kind "b" and those of kind "d". If we would impose commutation relations among the creation-annihilation operators, we would produce a state with negative energy. Moreover, commutation relations give rise, as we already noticed in the KG case, to symmetric wave functions and instead we would like to have anti-symmetric wave fuinctions, to satisfy Fermi-Dirac statistics.

Therefore, we will impose the following quantization rules:

$$
[b(p,n), b^\dagger(p',n')]_+ = [d(p,n), d^\dagger(p',n')]_+ = \delta_{nn'} \delta(\mathbf{p}-\mathbf{p}') , \qquad (4.707)
$$

$$
[b(p,n), b(p',n')]_+ = [d(p,n), d(p',n')]_+ = ... = 0 . \qquad (4.708)
$$

Using anticommutators (instead of commutators) we can write the energy in normal ordering as follows

$$
: H := \sum_{\pm n} \int d^3p E \left[b^\dagger(p,n) b(p,n) + d^\dagger(p,n) d(p,n) \right] . \qquad (4.709)
$$

The momentum operator has the same structure as the hamiltonian and then we find

$$
: P^i := \sum_{\pm n} \int d^3p p^i \left[b^\dagger(p,n) b(p,n) + d^\dagger(p,n) d(p,n) \right] . \qquad (4.710)
$$

The spectrum is recovered defining the action of $b(p,n)$ and $d(p,n)$ on the vacuum

$$
b(p,n)|0\rangle = 0 , \quad d(p,n)|0\rangle = 0 , \qquad (4.711)
$$

while the creation operators $b^\dagger(p,n)$ and $d^\dagger(p,n)$ create one-particle stated with definite energy and momentum

$$
b^\dagger(p,n)|0\rangle = |p\rangle , \quad d^\dagger(p,n)|0\rangle = |p\rangle , \qquad (4.712)
$$

If we refer only to H and $\mathbf{P}$, states of kind "b" and states of kind "d" are degenerate

$$
\begin{aligned}
: H : b^\dagger(p,n)|0\rangle &= E b^\dagger(p,n)|0\rangle , \qquad (4.713) \\
: H : d^\dagger(p,n)|0\rangle &= E d^\dagger(p,n)|0\rangle , \qquad (4.714) \\
: \mathbf{P} : b^\dagger(p,n)|0\rangle &= \mathbf{p} b^\dagger(p,n)|0\rangle , \qquad (4.715) \\
: \mathbf{P} : d^\dagger(p,n)|0\rangle &= \mathbf{p} d^\dagger(p,n)|0\rangle . \qquad (4.716)
\end{aligned}
$$

However, the Dirac lagrangian is invariant under global phase transformations and the conserved quantity, for the Nöther's theorem, is the "charge"

$$
Q = \int d^3X\, \psi^\dagger \psi . \qquad (4.717)
$$

If we substitute the expression of the field in normal modes in Eq. (4.717) and we integrate as in the case of the hamiltonian, we find

$$\int d^3X\, \psi^\dagger\psi = \sum_{\pm n}\int d^3p \left[b^\dagger(P,n)b(P,n) + d(P,n)d^\dagger(P,n)\right] \tag{4.718}$$

and therefore, in normal ordering

$$: Q := \sum_{\pm n}\int d^3p \left[b^\dagger(p,n)b(p,n) - d^\dagger(p,n)d(p,n)\right] . \tag{4.719}$$

The charge operator is able to distinguish between states of kind "b" and states of kind "d".

Let's remember that the current that gives rise to the conserved charge of the Nöther's theorem is

$$j^\mu = \overline{\psi}\gamma^\mu\psi \tag{4.720}$$

and that the interacting term of the Dirac's field with the electromagnetic field in the lagrangian density is

$$\mathcal{L}_{int} = -\mathcal{H}_{int} = -e\overline{\psi}\gamma^\mu\psi A_\mu\,. \tag{4.721}$$

Therefore

$$J^\mu = e\overline{\psi}\gamma^\mu\psi \tag{4.722}$$

can be interpreted as the electric current, while

$$: Q := \sum_{\pm n}\int d^3pe \left[b^\dagger(p,n)b(p,n) - d^\dagger(p,n)d(p,n)\right] \tag{4.723}$$

will be interpreted as the electric charge of the Dirac state.

This means that $b^\dagger(p,n)$ will create a particle with energy E, momentum $\mathbf{p}$ and charge $e = -|e|$ (the electron), while $d^\dagger(p,n)$ will create the anti-particle, with energy E, momentum $\mathbf{p}$ and charge $-e$ (the positron).

Two-particle states. Fermions

If we now consider a two-particle state, for instance a state with two electrons, since we have

$$b^\dagger(p_1,n)b^\dagger(p_2,n')|0\rangle = -b^\dagger(p_2,n')b^\dagger(p_1,n)|0\rangle \tag{4.724}$$

(and the same happens for antiparticle states) we will have totally antisymmetric states in the exchange of the two particles.

Anti-commutation rules for the fields

We quantized the Dirac's field imposing anti-commutation rules on the creation-annihilation operators in order to have a physical insight of what we were doing. These anti-commutations rules induce on the fields analogous anti-commutations rules. We have

$$\begin{aligned}
[\psi_\alpha(\mathbf{x},t),\psi^\dagger_\beta(\mathbf{y},t)]_+ \;\; = \;\; & \sum_{\pm n,\pm n'}\int \frac{d^3p\,d^3p'}{(2\pi)^3}\frac{m}{\sqrt{EE'}}\Big\{\left[b(p,n)u_\alpha(P,n)e^{-iP_\mu X^\mu} + d^\dagger(p,n)v_\alpha(P,n)e^{iP_\mu X^\mu}\right]\times \\
& \times\left[p',n')v^\dagger_\beta(P',n')e^{-iP'_\mu Y^\mu} + b^\dagger(p',n')u^\dagger_\beta(P',n')e^{iP'_\mu Y^\mu}\right] \\
& +\left[d(p',n')v^\dagger_\beta(P',n')e^{-iP'_\mu Y^\mu} + b^\dagger(p',n')u^\dagger_\beta(P',n')e^{iP'_\mu Y^\mu}\right]\times
\end{aligned}$$

$$\times\Big[b(p,n)u_\alpha(P,n)e^{-iP_\mu X^\mu}+d^\dagger(p,n)v_\alpha(P,n)e^{iP_\mu X^\mu}\Big]\Big\}_{(X^0=Y^0)}, \qquad (4.725)$$

$$= \sum_{\pm n,\pm n'}\int\frac{d^3p\,d^3p'}{(2\pi)^3}\frac{m}{\sqrt{EE'}}\Big\{$$
$$[b(p,n),d(p',n')]_+\,u_\alpha(P,n)v^\dagger_\beta(P',n')e^{-iP_\mu X^\mu}e^{-iP'_\mu Y^\mu}$$
$$+[b(p,n),b^\dagger(p',n')]_+\,u_\alpha(P,n)u^\dagger_\beta(P',n')e^{-iP_\mu X^\mu}e^{iP'_\mu Y^\mu}$$
$$+[d^\dagger(p,n),d(p',n')]_+\,v_\alpha(P,n)v^\dagger_\beta(P',n')e^{iP_\mu X^\mu}e^{-iP'_\mu Y^\mu}$$
$$+[d^\dagger(p,n),b^\dagger(p',n')]_+\,v_\alpha(P,n)u^\dagger_\beta(P',n')e^{iP_\mu X^\mu}e^{iP'_\mu Y^\mu}\Big\}_{(X^0=Y^0)}, \qquad (4.726)$$

$= \,|\,$using anti-commutation relations$\,|$

$$= \sum_{\pm n,\pm n'}\int\frac{d^3p\,d^3p'}{(2\pi)^3}\frac{m}{\sqrt{EE'}}\Big\{$$
$$+u_\alpha(P,n)u^\dagger_\beta(P',n')e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}\delta_{nn'}\delta(\mathbf{p}-\mathbf{p}')$$
$$+v_\alpha(P,n)v^\dagger_\beta(P',n')e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}\delta_{nn'}\delta(\mathbf{p}-\mathbf{p}')\Big\}_{(X^0=Y^0)}, \qquad (4.727)$$

$$= \sum_{\pm n}\int\frac{d^3p}{(2\pi)^3}\frac{m}{E}\Big\{u_\alpha(P,n)u^\dagger_\beta(P,n)+v_\alpha(\tilde{P},n)v^\dagger_\beta(\tilde{P},n)\Big\}e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}. \qquad (4.728)$$

Remembering the expression of the sum over polarizations, we have

$$\sum_{\pm n}u_\alpha(P,n)u^\dagger_\beta(P,n)=\left\{\left(\frac{\not{P}+m}{2m}\right)\gamma^0\right\}_{\alpha\beta},\quad \sum_{\pm n}v_\alpha(\tilde{P},n)v^\dagger_\beta(\tilde{P},n)=\left\{\left(\frac{\tilde{\not{P}}-m}{2m}\right)\gamma^0\right\}_{\alpha\beta}. \qquad (4.729)$$

Therefore

$$\left\{\left(\frac{\not{P}+m}{2m}\right)\gamma^0\right\}_{\alpha\beta}+\left\{\left(\frac{\tilde{\not{P}}-m}{2m}\right)\gamma^0\right\}_{\alpha\beta}=\left\{\left(\frac{\not{P}+\tilde{\not{P}}}{2m}\right)\gamma^0\right\}_{\alpha\beta}=\frac{E}{m}\delta_{\alpha\beta}. \qquad (4.730)$$

Substituting in the previous expression we have

$$[\psi_\alpha(\mathbf{x},t),\psi^\dagger_\beta(\mathbf{y},t)]_+=\int\frac{d^3p}{(2\pi)^3}\delta_{\alpha\beta}e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}=\delta_{\alpha\beta}\delta(\mathbf{x}-\mathbf{y}). \qquad (4.731)$$

Analogously we find

$$[\psi_\alpha(\mathbf{x},t),\psi_\beta(\mathbf{y},t)]_+=[\psi^\dagger_\alpha(\mathbf{x},t),\psi^\dagger_\beta(\mathbf{y},t)]_+=0. \qquad (4.732)$$

4.4.1 Microcausality and Dirac fields

4.5 The Electromagnetic Field (classical field)

In this section we will consider the case of a vector field, the electromagnetic field.

Maxwell's equations, in the Heaviside-Lorentz system, have the following form:

$$\nabla\cdot\mathbf{E}=\rho, \qquad (4.733)$$
$$\nabla\cdot\mathbf{H}=0, \qquad (4.734)$$
$$\nabla\wedge\mathbf{E}+\frac{1}{c}\frac{\partial\mathbf{H}}{\partial t}=0, \qquad (4.735)$$

$$\nabla \wedge \mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \mathbf{j}\,. \tag{4.736}$$

Taking the divergence of Eq. (4.736), we find the continuity equation

$$\frac{1}{c}\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0\,, \tag{4.737}$$

conservation of the electric charge.

Since the divergence of the magnetic field $\mathbf{H}$ is identically zero, we can introduce a vectorial function $\mathbf{A}(\mathbf{x}, t)$ such that:

$$\mathbf{H} = \nabla \wedge \mathbf{A}\,. \tag{4.738}$$

$\mathbf{A}$ is called *vector potential*.

Substituting Eq. (4.738) into Eq. (4.735) we obtain:

$$\nabla \wedge \mathbf{E} + \frac{1}{c}\frac{\partial}{\partial t}\nabla \wedge \mathbf{A} = \nabla \wedge \left(\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\right) = 0\,. \tag{4.739}$$

Eq. (4.739) inplies the existence of a scalar function, $\phi(\mathbf{x}, t)$, such that:

$$\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} = -\nabla \phi\,. \tag{4.740}$$

The electric field $\mathbf{E}$ can then be expressed as follows:

$$\mathbf{E} = -\nabla \phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\,. \tag{4.741}$$

The scalar function ϕ is called *scalar potential*.

The Maxwell's equations can be written in terms of the potentials. In this way we find that only two equations survive and the other two are identically satisfied.

In fact, Eq. (4.734) is identically satisfied. Eq. (4.733), with the electric field defined in (4.741), becomes

$$\nabla^2 \phi + \frac{1}{c}\frac{\partial}{\partial t}\nabla \cdot \mathbf{A} = -\rho\,. \tag{4.742}$$

Eq. (4.735) is identically satisfied. Finally, Eq. (4.738) becomes

$$\nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{A} = -\mathbf{j} + \nabla\left(\nabla \cdot \mathbf{A} + \frac{1}{c}\frac{\partial}{\partial t}\phi\right)\,. \tag{4.743}$$

In total, therefore, the four Eqs.(4.733, 4.734, 4.735, 4.736) are reduced to the following two:

$$\nabla^2 \phi + \frac{1}{c}\frac{\partial}{\partial t}\nabla \cdot \mathbf{A} = -\rho\,, \tag{4.744}$$

$$\nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{A} = -\mathbf{j} + \nabla\left(\nabla \cdot \mathbf{A} + \frac{1}{c}\frac{\partial}{\partial t}\phi\right)\,. \tag{4.745}$$

Eqs. (4.744 , 4.745) exhibit an important invariance under the following redefinition of the potentials:

$$\begin{cases} \mathbf{A} & \rightarrow \quad \mathbf{A}' = \mathbf{A} + \nabla \psi \\ \phi & \rightarrow \quad \phi' = \phi - \frac{1}{c}\frac{\partial \psi}{\partial t} \end{cases} \tag{4.746}$$

where $\psi(\mathbf{x}, t)$ is a generic function C^2 of its arguments.

This invariance is called *gauge invariance*. We find that the fields $\mathbf{E}$ and $\mathbf{H}$ are gauge-invariant quantities.

We can use gauge invariance in order to simplify Eqs. (4.744, 4.745). In fact, if we perform a transformation (4.746) of the potentials $(\phi, \mathbf{A})$ to the potentials $(\phi', \mathbf{A}')$, with ψ such that

$$\nabla^2\psi - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\psi = -\nabla\cdot\mathbf{A} - \frac{1}{c}\frac{\partial\phi}{\partial t}, \tag{4.747}$$

in the *new gauge* we will have

$$\nabla\cdot\mathbf{A}' + \frac{1}{c}\frac{\partial}{\partial t}\phi' = 0 \tag{4.748}$$

and Eqs. (4.744,4.745) will be simplified as follows:

$$\nabla^2\phi' - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\phi' = -\rho, \tag{4.749}$$

$$\nabla^2\mathbf{A}' - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{A}' = -\mathbf{j}, \tag{4.750}$$

(where we should remember that Eq. (4.748) holds). This choice of the gauge is called *Lorentz gauge.*

We have to notice that this choice of the function ψ does not determine in a univoque way the potentials ϕ' and $\mathbf{A}'$. It is possible to make another gauge transformation, staying in the Lorentz gauge. In fact, if

$$\nabla^2\chi - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\chi = 0, \tag{4.751}$$

the transformation (4.746) with χ at the place of ψ gives two new potentials $(\phi'', \mathbf{A}'')$ for which a relation like the one in Eq. (4.748) holds:

$$\begin{aligned}\nabla\cdot\mathbf{A}'' + \frac{1}{c}\frac{\partial}{\partial t}\phi'' &= \nabla\cdot\mathbf{A}' + \frac{1}{c}\frac{\partial}{\partial t}\phi' + \nabla^2\chi - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\chi = \\ &\equiv 0,\end{aligned} \tag{4.752}$$

where we used Eq. (4.748) and Eq. (4.751).

Gauge invariance tells us that not all the four components of the potentials are independent. In fact, Eq. (4.748) and Eq. (4.752) constitute two constraints for the four components of $(\phi, \mathbf{A})$. In total, therefore, only two components are independent, as we will see explicitely below.

4.5.1 Covariant form of Maxwell's equations

The charge density ρ and the current $\mathbf{j}$ transform, under Lorentz transformations, as the temporal and spatial parts of a four-vector

$$J^\mu = (\rho, \mathbf{j}). \tag{4.753}$$

The continuity equation, then, becomes simply forma:

$$\partial_\mu J^\mu = 0. \tag{4.754}$$

The differential D'Alambert operator

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2, \tag{4.755}$$

can be expressed in covariant form as follows

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 = \partial_\mu\partial^\mu = \partial^2, \tag{4.756}$$

that manifestly shows the fact that it is a Lorentz scalar. Finally, the scalar, ϕ, and vector, $\mathbf{A}$, potentials transform, again, as temporal and spatial parts of a four-vector

$$A^\mu = (\phi, \mathbf{A}) . \tag{4.757}$$

We can then write Eqs. (4.744, 4.745) in a manifestly covarian form as:

$$\partial^2 A^\mu - \partial^\mu (\partial_\nu A^\nu) = J^\mu , \tag{4.758}$$

which are invariant also under a gauge transformation

$$A^\mu \longrightarrow A'^\mu = A^\mu + \partial^\mu \psi . \tag{4.759}$$

In the Lorentz gauge, we will have

$$\begin{cases} \partial^2 A^\mu &= J^\mu \\ \partial_\mu A^\mu &= 0 \end{cases} \tag{4.760}$$

and in the free-field case

$$\begin{cases} \partial^2 A^\mu &= 0 \\ \partial_\mu A^\mu &= 0 \end{cases} \tag{4.761}$$

4.5.2 Electromagnetic tensor

We can write a manifestly covariant form of Maxwell's equations, introducing the electromagnetic tensor, that has, as components, the components of the electric and magnetic fields, $\mathbf{E}$ and $\mathbf{H}$, that are dirtectly gauge invariant.

Let us define the following anti-symmetric rank-2 tensor:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu = -F^{\nu\mu} . \tag{4.762}$$

Since we have

$$\begin{cases} \mathbf{E} &= -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \\ \mathbf{H} &= \nabla \wedge \mathbf{A} \end{cases} \Longrightarrow \begin{cases} E^i &= \partial^i A^0 - \partial^0 A^i \\ H^i &= \epsilon_{ijk}\partial^j A^k \end{cases} \tag{4.763}$$

we get immediately that $F^{\mu\nu}$ can be represented in form of a matrix 4×4 as follows:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -H^3 & H^2 \\ E^2 & H^3 & 0 & -H^1 \\ E^3 & -H^2 & H^1 & 0 \end{pmatrix} . \tag{4.764}$$

Using $F^{\mu\nu}$ we can express the two Maxwell's equations with sources as

$$\partial_\mu F^{\mu\nu} = J^\nu . \tag{4.765}$$

The second pair of equations can be obtained by the Bianchi identities for $F^{\mu\nu}$:

$$\partial^\mu F^{\nu\sigma} + \partial^\sigma F^{\mu\nu} + \partial^\nu F^{\sigma\mu} = 0 , \tag{4.766}$$

or we can introduce the dual of $F^{\mu\nu}$, $\mathcal{F}^{\mu\nu}$, via the following definition:

$$\mathcal{F}^{\alpha\beta} = \frac{1}{2}\epsilon^{\alpha\beta\mu\nu} F_{\mu\nu} , \tag{4.767}$$

where $\epsilon^{\alpha\beta\mu\nu}$ is the Levi-Civita tensor. The tensor $\mathcal{F}^{\mu\nu}$ has the following matrix representation:

$$\mathcal{F}^{\mu\nu} = \begin{pmatrix} 0 & -H^1 & -H^2 & -H^3 \\ H^1 & 0 & E^3 & -E^2 \\ H^2 & -E^3 & 0 & E^1 \\ H^3 & E^2 & -E^1 & 0 \end{pmatrix} , \tag{4.768}$$

and therefore it gives the opportunity to write the second pair of Maxwell's equation in the following form

$$\partial_\mu \mathcal{F}^{\mu\nu} = 0 . \tag{4.769}$$

4.5.3 Lagrangian density of the elettromagnetic field

Let us look for the Lagrangian density for the field A^μ in the vacuum, i.e. with $J^\mu = 0$. $\mathcal{L}$ should be a Lorentz scalar (invariant under proper Lorentz transformations), gauge invariant and, since the equations of motion are linear in the fields, $\mathcal{L}$ should contain quadratic terms. We have, at our disposal, the four-vector A^μ and the tensor $F^{\mu\nu}$ with which we can construct scalars like

$$F_{\mu\nu}F^{\mu\nu} ; \quad \mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu} ; \quad F_{\mu\nu}\mathcal{F}^{\mu\nu} . \tag{4.770}$$

Other terms like $F_{\mu\nu}A^\mu A^\nu$, $F_{\mu\nu}\, dX^\mu\, dX^\nu$, etc... are identically zero for the anti-symmetry of $F_{\mu\nu}$. Moreover, we have to discart also terms like $A_\mu A^\mu$. In fact, although it is a Lorentz scalar, it is not gauge invariant.

Of the three terms in Eq. (4.770) only one survives. In fact,

$$F_{\mu\nu}F^{\mu\nu} = 2\,(H^2 - E^2) , \tag{4.771}$$

while the second term gives

$$\mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu} = 2\,(E^2 - H^2) = -F_{\mu\nu}F^{\mu\nu} , \tag{4.772}$$

i.e. analogous to the first. The third is a pseudo-scalar and therefore it has to be discarted.

In total we have

$$\mathcal{L} = a\, F_{\mu\nu}F^{\mu\nu} , \tag{4.773}$$

where a is a proportionality constant that has to be found.

We find the correct equations of motion imposing $a = -\frac{1}{4}$. Finally

$$\mathcal{L} = -\frac{1}{4}\, F_{\mu\nu}F^{\mu\nu} . \tag{4.774}$$

Note: this "constructive" way to the Lagrangian density would have allowed also the presence of other terms. One could add for instance a term which is a Lorentz scalar and also gauge invariant, like $F_{\mu\lambda}F^{\lambda}_{\sigma}F^{\sigma\mu}$. This term, however, is an operator of dimension 6 and it is not renormalizable. This criterion will be clear when we will introduce radiative corrections.

We can find the Lagrangian density for the electromagnetic field in a more standard way, using Hamilton's principle. Considering a variation of the field, δA^ν, that vanishes on the boundary of the integration volume, we have

$$0 = \delta S = \int d^4X\, \left[\partial^2 A_\nu - \partial_\nu(\partial^\mu A_\mu)\right]\, \delta A^\nu = \int d^4X\, \left[\partial_\mu\partial^\mu A_\nu\, \delta A^\nu - \partial_\nu(\partial^\mu A_\mu)\, \delta A^\nu\right] , \tag{4.775}$$

$$= \left|\,\text{integrating by parts}\,\right|$$

$$= \int d^4X\, \left[\partial^\mu(\partial_\mu A_\nu\, \delta A^\nu) - \partial_\mu A_\nu\, \delta(\partial^\mu A^\nu) - \partial^\mu(\partial_\nu A_\mu\, \delta A^\nu) + \partial_\nu A_\mu\, \delta(\partial^\mu A^\nu)\right] , \tag{4.776}$$

$$
\begin{aligned}
&= \text{| the surface terms integrate to zero |} \\
&= \int d^4X \, \left[-\partial_\mu A_\nu \, \delta(\partial^\mu A^\nu) + \partial_\nu A_\mu \, \delta(\partial^\mu A^\nu)\right] , && (4.777)\\
&= \int d^4X \left[-\frac{1}{2}\partial_\mu A_\nu \, \delta(\partial^\mu A^\nu) - \frac{1}{2}\partial_\nu A_\mu \, \delta(\partial^\nu A^\mu) + \frac{1}{2}\partial_\nu A_\mu \, \delta(\partial^\mu A^\nu) + \frac{1}{2}\partial_\mu A_\nu \, \delta(\partial^\nu A^\mu)\right], && (4.778)\\
&= \text{| since } F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \text{ |} \\
&= \int d^4X \left[-\frac{1}{2}F_{\mu\nu} \, \delta F^{\mu\nu}\right] , && (4.779)\\
&= \delta \int d^4X \left[-\frac{1}{4}F_{\mu\nu} F^{\mu\nu}\right] . && (4.780)
\end{aligned}
$$

Therefore, we have

$$
\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} = -\frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu + \frac{1}{2}\partial_\mu A_\nu \partial^\nu A^\mu \, , \tag{4.781}
$$

which is Lorentz invariant, gauge invariant and local. The overall sign is important in order to have an energy density which is positive definite.

We can check that with this Lagrangian density we get the correct Maxwell's equation as Euler-Lagrange equations

$$
\frac{\partial \mathcal{L}}{\partial A_\mu} - \partial_\nu \, \frac{\partial \mathcal{L}}{\partial A_{\mu,\nu}} = 0 \, . \tag{4.782}
$$

In fact, we have

$$
\begin{cases} \frac{\partial \mathcal{L}}{\partial A_\mu} = 0 \, , \\ \frac{\partial \mathcal{L}}{\partial A_{\mu,\nu}} = \quad \partial^\mu A^\nu - \partial^\nu A^\mu \, , \end{cases} \tag{4.783}
$$

and therefore

$$
\partial^2 A^\mu - \partial^\mu(\partial_\nu A^\nu) = 0 \, . \tag{4.784}
$$

We can express the Lagrangian density in terms of the electric and magnetic fields. We have

$$
\begin{aligned}
&F^{00} = F^{ii} = 0 \, , && (4.785)\\
&F^{0i} = \partial^0 A^i - \partial^i A^0 = \partial^0 A^i + \partial_i A^0 = -E^i = -F^{i0} = -F_{0i} \, , && (4.786)\\
&F^{ij} = \partial^i A^j - \partial^j A^i = -\partial_i A^j + \partial_j A^i = -\epsilon^{ijk}(\nabla \wedge \mathbf{A})^k = -\epsilon^{ijk} H^k \, , && (4.787)
\end{aligned}
$$

and therefore

$$
F_{\mu\nu}F^{\mu\nu} = F_{0i}F^{0i} + F_{i0}F^{i0} + F_{ij}F^{ij} + F_{ji}F^{ji} = -2|\mathbf{E}|^2 + 2|\mathbf{H}|^2 \, . \tag{4.788}
$$

In total

$$
\mathcal{L} = \frac{|\mathbf{E}|^2 - |\mathbf{H}|^2}{2} \, . \tag{4.789}
$$

4.5.4 Energy-Momentum tensor

From Nöther's theorem we have

$$
\begin{aligned}
T^{\mu\nu} &= \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\alpha)} \partial^\nu A_\alpha - \eta^{\mu\nu} \mathcal{L} \, , && (4.790)\\
&= (-\partial^\mu A^\alpha + \partial^\alpha A^\mu) \, \partial^\nu A_\alpha + \frac{1}{4}\eta^{\mu\nu} F_{\alpha\beta} F^{\alpha\beta} \, , && (4.791)\\
&= -F^{\mu\alpha} \, \partial^\nu A_\alpha + \frac{1}{4}\eta^{\mu\nu} F_{\alpha\beta} F^{\alpha\beta} \, , && (4.792)
\end{aligned}
$$

such that

$$
\partial_\mu T^{\mu\nu} = 0 \, . \tag{4.793}
$$

This form, (4.792), is not symmetric in the exchange $\mu \leftrightarrow \nu$ and not even gauge invariant. In fact, under a gauge transformation

$$A^\mu \to A^\mu + \partial^\mu \chi \,, \tag{4.794}$$

we have

$$T^{\mu\nu} \to T'^{\mu\nu} = -F^{\mu\alpha}\,\partial^\nu (A_\alpha + \partial_\alpha \chi) + \frac{1}{4}\eta^{\mu\nu} F_{\alpha\beta}F^{\alpha\beta} = T^{\mu\nu} - F^{\mu\alpha}\,\partial^\nu \partial_\alpha \chi \,. \tag{4.795}$$

However, the conserved charges are gauge invariant (this is the important thing!). In fact, we have

$$F^{\mu\alpha}\,\partial^\nu \partial_\alpha \chi = \partial_\alpha (F^{\mu\alpha}\,\partial^\nu \chi) - (\partial_\alpha F^{\mu\alpha})\,\partial^\nu \chi = \partial_\alpha (F^{\mu\alpha}\,\partial^\nu \chi) \,, \tag{4.796}$$

since for the equations of motion, $\partial_\alpha F^{\mu\alpha} = 0$. Therefore, under the gauge transformation (4.794) we have

$$\begin{aligned} P^\nu &= \int d^3X\, T^{0\nu} \to \\ &\to P'^\nu = \int d^3X\, T'^{0\nu} = \int d^3X\, T^{0\nu} - \int d^3X\, F^{0\alpha}\,\partial^\nu \partial_\alpha \chi \,, && (4.797) \\ &= P^\nu - \int d^3X\, \partial_\alpha (F^{0\alpha}\,\partial^\nu \chi) \,, && (4.798) \\ &= P^\nu - \int d^3X\, \partial_i (F^{0i}\,\partial^\nu \chi) = P^\nu \,, && (4.799) \end{aligned}$$

since $F^{00} = 0$ and since the last integral gives a surface term that is zero in the limit of infinite volume (we understand always the fact that the fields go to zero sufficiently rapidly at infinity).

We can define a gauge invariant energy-momentum tensor adding to the form in Eq. (4.792) the following term

$$C^{\mu\nu} = \partial_\alpha (F^{\mu\alpha} A^\nu) \,, \tag{4.800}$$

which satisfies

$$\partial_\mu C^{\mu\nu} = 0 \,, \tag{4.801}$$

and it is such that

$$\int d^3X\, C^{0\nu} = \int d^3X\, \partial_i F^{0i}\, A^\nu = 0 \,. \tag{4.802}$$

The new energy-momentum tensor (symmetric in $\mu \leftrightarrow \nu$ and gauge-invariant) is

$$\tilde{T}^{\mu\nu} = T^{\mu\nu} + C^{\mu\nu} = F^{\mu\alpha} F^\nu_\alpha + \frac{1}{4}\eta^{\mu\nu} F_{\alpha\beta}F^{\alpha\beta} \,. \tag{4.803}$$

Using (4.803) we get the usual expressions for the energy density and the momentum:

$$\begin{aligned} \mathcal{H} &= \tilde{T}^{00} = \frac{|\mathbf{E}|^2 + |\mathbf{H}|^2}{2} \,, && (4.804) \\ \mathcal{P}^i &= \tilde{T}^{0i} = (\mathbf{E} \wedge \mathbf{H})^i \,, && (4.805) \end{aligned}$$

which is the Poynting vector.

The two expressions $T^{\mu\nu}$ and $\tilde{T}^{\mu\nu}$ are physically equivalent. The additional term is a total derivative in the Lagrangian density and, therefore, it does not affect the equations of motion. It is interesting to notice that such a piece changes the currents, while the charges are always the same.

4.5.5 Number of degrees of freedom

We describe the electromagnetic field with the four-vector A^μ. However, due to gauge invariance, the physical degrees of fredom are not 4 (as the fact that we use an object with four components would suggest) but 2. We can perform the calculation of the actual degrees of freedom in a covariant gauge or in a physical gauge, like the Coulomb gauge.

Covariant gauge

We can show that the field A_μ has only two degrees of freedom using the equations of motion and gauge invariance.

Consider the Fourier tranform of the field

$$A_\mu(X) = \int d^4K e^{iK_\nu X^\nu} A_\mu(K)\,. \tag{4.806}$$

Substituting into the equations of motion we get

$$-K^2\, A_\mu(K) + K_\mu(K^\nu A_\nu(K)) = 0\,. \tag{4.807}$$

Now, let us write the field $A_\nu(K)$ as a combination of 4 vectors of a basis for the Minkowski space. We can choose the following vectors:

$$K^\mu = (E, \mathbf{k})\,, \quad \tilde{K}^\mu = (E, -\mathbf{k})\,, \quad \epsilon^{(\lambda)\,\mu}(K)\,, \quad \lambda = 1, 2\,, \tag{4.808}$$

with

$$K^\mu \epsilon^{(\lambda)}_\mu(K) = 0\,. \tag{4.809}$$

We can write

$$A_\mu(K) = a_{(\lambda)}(K)\epsilon^{(\lambda)}_\mu(K) + b(K)K_\mu + c(K)\tilde{K}_\mu\,. \tag{4.810}$$

Substituting in Eq. (4.807) we get

$$\begin{aligned} 0 &= -K^2\left[a_{(\lambda)}(K)\epsilon^{(\lambda)}_\mu(K) + b(K)K_\mu + c(K)\tilde{K}_\mu\right] + K_\mu\left\{K^\nu\left[a_{(\lambda)}(K)\epsilon^{(\lambda)}_\nu(K) + b(K)K_\nu + c(K)\tilde{K}_\nu\right]\right\}\,, \\ &= -K^2\, a_{(\lambda)}(K)\epsilon^{(\lambda)}_\mu(K) - K^2 c(K)\tilde{K}_\mu + (K^\nu\tilde{K}_\nu)c(K)K_\mu\,. \end{aligned} \tag{4.811}$$

Since the 4 vectors form a basis, we have to have

$$K^2\, a_{(\lambda)}(K) = K^2\, c(K) = (K^\nu\tilde{K}_\nu)c(K) = 0 \tag{4.812}$$

and therefore, since $(K^\nu\tilde{K}_\nu) \neq 0$ we have $c(K) = 0$ and since we want $a_{(\lambda)}(K) \neq 0$ we have to have $K^0 = 0$. The coefficient $b(K)$ is indeterminate and we can choose it in such a way to be 0. This fact is connected to gauge invariance. In fact, if

$$A_\mu(X) \to A_\mu(X) + \partial_\mu\chi(X)\,, \tag{4.813}$$

the Fourier transform is such that

$$A_\mu(K) \to A_\mu(K) + iK_\mu\chi(K)\,, \tag{4.814}$$

where

$$\chi(X) = \int d^4K e^{iK_\nu X^\nu}\chi(K)\,. \tag{4.815}$$

Under (4.814) we have

$$\begin{aligned} A_\mu(K) &= a_{(\lambda)}(K)\epsilon^{(\lambda)}_\mu(K) + b(K)K_\mu + c(K)\tilde{K}_\mu \to \\ &\to A'_\mu(K) = a'_{(\lambda)}(K)\epsilon^{(\lambda)}_\mu(K) + b(K)K_\mu + i\chi(K)K_\mu + c'(K)\tilde{K}_\mu\,, \qquad (4.816) \\ &\qquad = a'_{(\lambda)}(K)\epsilon^{(\lambda)}_\mu(K) + c'(K)\tilde{K}_\mu\,, \qquad (4.817) \end{aligned}$$

where

$$a'_{(\lambda)}(K) = a_{(\lambda)}(K)\,, \quad b'(K) = b(K) + i\chi(K) = 0\,, \quad c'(K) = c(K)\,. \tag{4.818}$$

Therefore, choosing a gauge transformation such that $\chi(K) = ib(K)$ we can always remove the term proportional to K_μ. We remain, then with two degrees of freedom (since $b(K) = c(K) = 0$):

$$A_\mu(K) = a_{(\lambda)}(K)\epsilon_\mu^{(\lambda)}(K)\,, \tag{4.819}$$

with

$$K^\mu A_\mu(K) = a_{(\lambda)}(K)K^\mu\epsilon_\mu^{(\lambda)}(K) = 0\,, \tag{4.820}$$

which in coordinate space can be written as

$$\partial^\mu A_\mu(K) = 0\,, \tag{4.821}$$

i.e. the Lorentz gauge.

Coulomb gauge

In Coulomb gauge the number of degrees of freedom is even more evident. In fact, in the vacuum we can always choose a gauge such that

$$A^0 = \nabla\cdot\mathbf{A} = 0\,, \tag{4.822}$$

that are two constraints on the four components of A_μ (therefore two degrees of freedom left).

Let us show that this is possible. Let us make a gauge transformation as in Eq. (4.813) with

$$\chi(X) = -\int_0^t A^0(\mathbf{x},t')\,dt'\,, \tag{4.823}$$

in such a way that

$$A_\mu \to A'_\mu = A_\mu - \partial_\mu\int_0^t A^0(\mathbf{x},t')\,dt'\,. \tag{4.824}$$

Clearly we have

$$A'_0 = A_0 - \partial_0\int_0^t A^0(\mathbf{x},t')\,dt' = 0\,. \tag{4.825}$$

Let us perform now an additional gauge transformation

$$A'_\mu \to A''_\mu = A'_\mu + \partial_\mu\tilde{\chi}(X)\,, \tag{4.826}$$

such that $\nabla\cdot\mathbf{A}'' = 0$. To this end we choose $\tilde{\chi}(X)$ such that

$$\nabla\cdot\mathbf{A}'' = \nabla\cdot\mathbf{A}' - \nabla^2\tilde{\chi}(X) = 0\,, \tag{4.827}$$

or

$$\nabla^2\tilde{\chi}(X) = \nabla\cdot\mathbf{A}'\,. \tag{4.828}$$

A solution for this equation is[16]

$$\tilde{\chi}(X) = -\frac{1}{4\pi}\int d^3X'\,\frac{\nabla'\cdot\mathbf{A}'(\mathbf{x}',t)}{|\mathbf{x}-\mathbf{x}'|}\,. \tag{4.831}$$

[16] In fact we have

$$\nabla^2\frac{1}{|\mathbf{x}|} = -4\pi\delta^3(\mathbf{x})\,, \tag{4.829}$$

and therefore

$$\nabla^2\chi(X) = -\frac{1}{4\pi}\int d^3X'\,\nabla'\cdot\mathbf{A}'(\mathbf{x}',t)\,\nabla^2\frac{1}{|\mathbf{x}-\mathbf{x}'|} = \frac{1}{4\pi}\int d^3X'\,\nabla'\cdot\mathbf{A}'(\mathbf{x}',t)\,4\pi\delta^3(\mathbf{x}-\mathbf{x}') = \nabla\cdot A'(X)\,. \tag{4.830}$$

for which we have

$$\partial_0 \tilde{\chi}(X) = -\frac{1}{4\pi}\int d^3X' \, \frac{\nabla' \cdot \dot{\mathbf{A}}'(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|} = 0\,, \tag{4.832}$$

since for the Gauss equation

$$0 = \nabla \cdot \mathbf{E} = \nabla \cdot \left(\nabla \dot{A}'^0 - \partial^0 \mathbf{A}'\right) = -\nabla \cdot A'(X)\,. \tag{4.833}$$

In the end, in the gauge in which we defined A''_μ we have

$$A''_0 = A'_0 + \partial_0 \tilde{\chi} = 0 + 0 = 0\,, \quad \text{and} \quad \nabla \cdot \mathbf{A}'' = 0\,, \tag{4.834}$$

as we wanted to show.

4.6 Quantization of the Electromagnetic Field

We are now ready to consider the quantization of the electromagnetic field. We would like to mantain the general covariance of the theory and therefore we require to find non-trivial commutation relations among all the components of A^μ and the conjugated momentum

$$\Pi^\mu = \frac{\partial \mathcal{L}}{\partial \dot{A}_\mu}\,. \tag{4.835}$$

We would impose the following equal time commutation relations

$$\begin{aligned} &[A_\mu(\mathbf{x}, t), \Pi_\nu(\mathbf{y}, t)] = i\eta_{\mu\nu}\delta^3(\mathbf{x} - \mathbf{y})\,, \\ &[A_\mu(\mathbf{x}, t), A_\nu(\mathbf{y}, t)] = [\Pi_\mu(\mathbf{x}, t), \Pi_\nu(\mathbf{y}, t)] = 0\,. \end{aligned} \tag{4.836}$$

In order to evaluate the conjugated momentum, we refer to Eq. (4.781). We find

$$\Pi^\mu = \frac{\partial \mathcal{L}}{\partial \dot{A}_\mu} = -\partial^0 A^\mu + \partial^\mu A^0 = F^{\mu 0}\,. \tag{4.837}$$

Since $F^{\mu\nu}$ is antisymmetric, we have $\Pi^0 = 0$, at the operator level, and we are not able to impose the commutation relation

$$[A_0(\mathbf{x}, t), \Pi_0(\mathbf{y}, t)] = i\delta^3(\mathbf{x} - \mathbf{y})\,. \tag{4.838}$$

This is a problem that emerges from our requirement to mantain a manifestly covariant form of the quantization, while we know already that the time-degree of freedom is not physical. A possible solution is to get rid of the general covariance and to quantize only the two transverse degrees of freedom. This could be done, for instance, using a physical gauge, like the Coulomb gauge, in which we reduce from the beginning only to the two transverse degrees of freedome. However, in such approach we loose covariance, that is quite important in computations. We therefore choose to quantize the electromagnetic field preserving general covariance and renouncing to explicit gauge invariance (although we will recover gauge invariance checking that two computations in two different gauges give rise to the same result). This approach was introduced by Gupta and Bleuler.

The idea is to renounce to gauge invariance in order to cure the relation $\Pi^0 = 0$ in such a way that this does not hold at the operator level, but only when we evaluate the operator on a physical state.

Let us choose a lagrangian density that gives the correct equations of motion (Maxwell's equations) but only in Lorentz gauge:

$$\partial^2 A^\mu = 0\,. \tag{4.839}$$

These equations come from the lagrangian density

$$\mathcal{L} = -\frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu\,, \tag{4.840}$$

as can be easyly checked. The difference between the lagrangian density given in Eq. (4.781) and the one in Eq. (4.840) is

$$\begin{aligned}
\mathcal{L}_{GF} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu \,, && (4.841)\\
&= -\frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu - \frac{1}{2}\partial_\mu A_\nu \partial^\nu A^\mu + \frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu \,, && (4.842)\\
&= -\frac{1}{2}\partial_\mu A_\nu \partial^\nu A^\mu \,, && (4.843)\\
&= \partial^\nu\left(\frac{1}{2}(\partial_\mu A_\nu)A^\mu\right) - \frac{1}{2}(\partial^\nu\partial_\mu A_\nu)A^\mu \,, && (4.844)\\
&= \partial^\nu\left(\frac{1}{2}(\partial_\mu A_\nu)A^\mu\right) - \partial_\mu\left(\frac{1}{2}(\partial^\nu A_\nu)A^\mu\right) + \frac{1}{2}(\partial^\nu A_\nu)^2 \,, && (4.845)\\
&= |\,\text{up to total derivatives that do not affect the eqs of motion}\,| \\
&= \frac{1}{2}(\partial^\nu A_\nu)^2 \,. && (4.846)
\end{aligned}$$

Therefore, we quantive the lagrangian density

$$\mathcal{L} = \mathcal{L}_{EM} + \mathcal{L}_{GF} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}(\partial^\nu A_\nu)^2 \,. \tag{4.847}$$

The lagrangian $\mathcal{L}_{GF}$ is called "gauge fixing" lagrangian.

If we look for the Euler-Lagrange equations for the lagrangian density in Eq. (4.847) we find

$$\frac{\partial\mathcal{L}}{\partial A_\mu} = 0 \,, \quad \frac{\partial\mathcal{L}}{\partial A_{\mu,\nu}} = -A^{\mu,\nu} + A^{\nu,\mu} - \eta^{\mu\nu}(\partial^\lambda A_\lambda) \tag{4.848}$$

and therefore

$$0 = -\partial_\nu\frac{\partial\mathcal{L}}{\partial A_{\mu,\nu}} = \partial^2 A^\mu - \partial^\mu(\partial^\nu A_\nu) + \partial^\mu(\partial^\lambda A_\lambda) = \partial^2 A^\mu \,, \tag{4.849}$$

that are the Maxwell's equations in the Lorentz gauge[17].

Using the lagrangian density (4.847) we can recompute the momentum conjugated to A^μ finding

$$\Pi^\mu = F^{\mu 0} - \eta^{\mu 0}(\partial^\nu A_\nu) \,. \tag{4.852}$$

Now the temporal component is not anymore identically equal to zero. We have

$$\Pi^0 = -(\partial^\nu A_\nu) \,. \tag{4.853}$$

It is clear thet the Lorentz gauge gives $\partial^\nu A_\nu = 0$; however, we are now speaking about operators. We can require that in general

$$\partial^\nu A_\nu \neq 0 \,, \tag{4.854}$$

but it is zero only when evaluated between two physical states

$$\langle phys|\partial^\nu A_\nu|phys\rangle = 0 \,. \tag{4.855}$$

[17] We can in general use the lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{\lambda}{2}(\partial^\nu A_\nu)^2 \,, \tag{4.850}$$

with λ a constant (actually a Lagrange multiplier). The equations of motion would then be

$$\partial^2 A^\mu - (1-\lambda)\partial^\mu(\partial^\nu A_\nu) = 0 \,, \tag{4.851}$$

that give $\partial^2 A^\mu = 0$ when $\lambda = 1$. The case $\lambda = 1$ is called "Lorentz-Feynman gauge".

The condition (4.855) defines the physical states. Imposing (4.854) at the operator level, with (4.855) on the physical states, means that we enlarged the Fock space. We have states that are physical and non-physical states on which, in general, $\langle\phi|\partial^\nu A_\nu|\phi\rangle \neq 0$. The enlargement of the Fock space is the price to pay for the covariant quantization. The states corresponding to temporal and longitudinal photons will be non physical, while the transverse polarization states will be the physical ones.

We will comment more closely on Eq. (4.855) in a while.

Since now at the operator level we have $\Pi^0 \neq 0$, we can proceed imposing the quantization relations (4.836) that can be simplyfied as follows. We have

$$\begin{aligned} \Pi_0 &= -\partial^0 A_0 - \partial_i A^i \,, && (4.856)\\ \Pi_i &= -\partial_0 A_i + \partial_i A_0 \,. && (4.857) \end{aligned}$$

Therefore, for the temporal component

$$i\delta^3(\mathbf{x}-\mathbf{y}) = [A_0(\mathbf{x},t), \Pi_0(\mathbf{y},t)] = [A_0(\mathbf{x},t), -\partial^0 A_0(\mathbf{y},t) - \partial_i A^i(\mathbf{y},t)] = [A_0(\mathbf{x},t), -\dot{A}_0(\mathbf{y},t)]\,, \quad (4.858)$$

since

$$\begin{aligned} [A_0(\mathbf{x},t), -\partial_i A^i(\mathbf{y},t)] &= -A_0(\mathbf{x},t)\left(\frac{\partial}{\partial y_i}A^i(\mathbf{y},t)\right) + \left(\frac{\partial}{\partial y_i}A^i(\mathbf{y},t)\right)A_0(\mathbf{x},t)\,, && (4.859)\\ &= -\frac{\partial}{\partial y_i}[A_0(\mathbf{x},t), A^i(\mathbf{y},t)] = 0\,. && (4.860) \end{aligned}$$

For the spatial part we have

$$-i\delta^3(\mathbf{x}-\mathbf{y}) = [A_i(\mathbf{x},t), \Pi_i(\mathbf{y},t)] = [A_0(\mathbf{x},t), -\partial_0 A_i(\mathbf{y},t) + \partial_i A_0(\mathbf{y},t)] = [A_0(\mathbf{x},t), -\dot{A}_i(\mathbf{y},t)]\,, \quad (4.861)$$

since, again, we have

$$[A_0(\mathbf{x},t), \partial_i A_0(\mathbf{y},t)] = 0\,. \quad (4.862)$$

Finally we have

$$\begin{aligned} &[A_\mu(\mathbf{x},t), \dot{A}_\nu(\mathbf{y},t)] = -i\eta_{\mu\nu}\delta^3(\mathbf{x}-\mathbf{y})\,,\\ &[A_\mu(\mathbf{x},t), A_\nu(\mathbf{y},t)] = [\dot{A}_\mu(\mathbf{x},t), \dot{A}_\nu(\mathbf{y},t)] = 0\,. \end{aligned} \quad (4.863)$$

4.6.1 Plane wave solutions

In order to get the quanta (photons) we need to express the field in normal modes (plane wave solutions). We have to express A^μ in a basis of the Minkowski space. We do not have the opportunity to move to the rest frame, since the photons travel at the speed of light. However, in the frame in which the momentum is $P^\mu = (p, 0, 0, p)$, we choose the following 4 vectors:

1. The unit time-like vector (that defines the time axis)

$$n^\mu(p) = (1, 0, 0, 0) = \epsilon^{(0)\,\mu}(p) \quad (4.864)$$

   such that

$$\epsilon^{(0)\,\mu}(p)\epsilon^{(0)}_\mu(p) = 1\,. \quad (4.865)$$

2. The two transverse space-like vectors

$$\epsilon^{(\lambda)\,\mu}(p)\,, \quad \lambda = 1, 2\,, \quad (4.866)$$

   such that

$$\begin{aligned} &\epsilon^{(\lambda)}_\mu(p)\epsilon^{(0)\,\mu}(p) = \epsilon^{(\lambda)}_\mu(p)P^\mu = 0\,, && (4.867)\\ &\epsilon^{(\lambda)}_\mu(p)\epsilon^{(\lambda')\,\mu}(p) = -\delta^{\lambda\lambda'}\,. && (4.868) \end{aligned}$$

   We have

$$\epsilon^{(1)\,\mu}(p) = (0, 1, 0, 0)\,, \quad \epsilon^{(2)\,\mu}(p) = (0, 0, 1, 0)\,. \quad (4.869)$$

3. A fourth space-like vector
$$\epsilon^{(3)\,\mu}(p)\,, \tag{4.870}$$
such that
$$\epsilon^{(3)}_\mu(p)\epsilon^{(0)\,\mu}(p)=\epsilon^{(3)}_\mu(p)\epsilon^{(\lambda)\,\mu}(p)=0\,, \tag{4.871}$$
$$\epsilon^{(3)}_\mu p)\epsilon^{(3)\,\mu}(p)=-1\,. \tag{4.872}$$
For instance we can choose
$$\epsilon^{(3)\,\mu}(p)=\frac{P^\mu-(\epsilon^{(0)}_\nu(p)P^\nu)\epsilon^{(0)\,\mu}(p)}{(\epsilon^{(0)}_\nu(p)P^\nu)}=(0,0,0,1)\,. \tag{4.873}$$

These 4 vectors are orthonormal in the Minkowski space and satisfy completeness relations:
$$\epsilon^{(\lambda)}_\mu(p)\epsilon^{(\lambda')\,\mu}(p)=\eta^{\lambda\lambda'}\,,\quad \lambda=0,1,2,3\,, \tag{4.874}$$
$$\epsilon^{(\lambda)}_\mu(p)\epsilon^{(\lambda')}_\nu(p)\eta_{\lambda\lambda'}=\eta_{\mu\nu}\,. \tag{4.875}$$

We can prove the relation (4.875) in the frame in which $P^\mu=(p,0,0,p)$ noting that $\epsilon^{(\lambda)}_\mu=\delta^\lambda_\mu$ and since it is a covariant equation it holds unchanged in form in any other frame.

The expansion of A^μ in plane waves is therefore[18]
$$A_\mu(X)=\int\frac{d^3p}{\sqrt{(2\pi)^3 2E}}\sum_{\lambda=0}^{3}\epsilon^{(\lambda)}_\mu(p)\left[a_{(\lambda)}(p)e^{-iP_\mu X^\mu}+a^\dagger_{(\lambda)}(p)e^{iP_\mu X^\mu}\right]\,, \tag{4.877}$$
where we considered the fact thet the field A_μ is a real field and where we normalized already the expression, because for every $\mu=0,1,2,3$ we find a Klein-Gordon field
$$A_0(X)\;=\;\int\frac{d^3p}{\sqrt{(2\pi)^3 2E}}\left[\tilde a_0(p)e^{-iP_\mu X^\mu}+\tilde a^\dagger_0(p)e^{iP_\mu X^\mu}\right]\,, \tag{4.878}$$
$$A_1(X)\;=\;\int\frac{d^3p}{\sqrt{(2\pi)^3 2E}}\left[\tilde a_1(p)e^{-iP_\mu X^\mu}+\tilde a^\dagger_1(p)e^{iP_\mu X^\mu}\right]\,, \tag{4.879}$$
$$\vdots\qquad\vdots$$
where $\tilde a_0(p)=\sum_{\lambda=0}^3\epsilon^{(\lambda)}_0(p)a_{(\lambda)}(p)$, $\tilde a_1(p)=\sum_{\lambda=0}^3\epsilon^{(\lambda)}_1(p)a_{(\lambda)}(p)$ ... etc.

Remembering the form of $f^{(+)}_p(X)$ and $f^{(+)*}_p(X)$ we can write
$$A_\mu(X)=\int d^3p\sum_{\lambda=0}^{3}\epsilon^{(\lambda)}_\mu(p)\left[a_{(\lambda)}(p)f^{(+)}_p(X)+a^\dagger_{(\lambda)}(p)f^{(+)*}_p(X)\right]\,, \tag{4.880}$$

We would like to check that, imposing the quantization relations on the fields, the opertators $a_{(\lambda)}(P)$ and $a^\dagger_{(\lambda)}(P)$ are actually annihilation/creation operators (they obay the correct commutation relations). We can project out the operators $a_{(\lambda)}(P)$ and $a^\dagger_{(\lambda)}(P)$ in terms of the fields and then use

[18] If we consider circular polarization we have to introduce two complex vectors for the transverse states. Therefore, we have
$$A_\mu(X)=\int\frac{d^3p}{\sqrt{(2\pi)^3 2E}}\sum_{\lambda=0}^{3}\left[\epsilon^{(\lambda)}_\mu(p)a_{(\lambda)}(p)e^{-iP_\mu X^\mu}+\epsilon^{(\lambda)\,*}_\mu(p)a^\dagger_{(\lambda)}(p)e^{iP_\mu X^\mu}\right]\,. \tag{4.876}$$

the quantization relations for the fields and check that these relations induce the correct commutation relations of the annihilation/creation operators. We have

$$\sum_{\lambda=0}^{3}\epsilon_\mu^{(\lambda)}(p)a_{(\lambda)}(p)=i\int d^3X f_p^{(+)*}(X)\overleftrightarrow{\partial_0}A_\mu(X)\,. \tag{4.881}$$

If now we multiply on the left by $\epsilon^{(\lambda')\,\mu}(p)$ we find

$$\epsilon^{(\lambda')\,\mu}(p)\sum_{\lambda=0}^{3}\epsilon_\mu^{(\lambda)}(p)a_{(\lambda)}(p)=\eta^{\lambda'\lambda}a_{(\lambda)}(p)=i\int d^3X\epsilon^{(\lambda')\,\mu}f_p^{(+)*}(X)\overleftrightarrow{\partial_0}A_\mu(X)\,. \tag{4.882}$$

Then, multiplying on the left by $\eta_{\sigma\lambda'}$ we find

$$\begin{aligned} a_{(\sigma)}(p) &= i\eta_{\sigma\lambda'}\int d^3X\epsilon^{(\lambda')\,\mu}f_p^{(+)*}(X)\overleftrightarrow{\partial_0}A_\mu(X)\,, &(4.883)\\ &= i\eta_{\sigma\lambda'}\int d^3X\epsilon^{(\lambda')\,\mu}\left[f_p^{(+)*}(X)\dot{A}_\mu(X)-\left(\partial_0 f_p^{(+)*}(X)\right)A_\mu(X)\right]\,. &(4.884)\end{aligned}$$

Analogously we find

$$\begin{aligned} a^\dagger_{(\sigma)}(p) &= -i\eta_{\sigma\lambda'}\int d^3X\epsilon^{(\lambda')\,\mu}f_p^{(+)}(X)\overleftrightarrow{\partial_0}A_\mu(X)=i\eta_{\sigma\lambda'}\int d^3X\epsilon^{(\lambda')\,\mu}A_\mu(X)\overleftrightarrow{\partial_0}f_p^{(+)}(X)\,, &(4.885)\\ &= i\eta_{\sigma\lambda'}\int d^3X\epsilon^{(\lambda')\,\mu}\left[\left(\partial_0 f_p^{(+)}(X)\right)A_\mu(X)-f_p^{(+)}(X)\dot{A}_\mu(X)\right]\,. &(4.886)\end{aligned}$$

With these expressions we find, for instance

$$\begin{aligned} [a_{(\lambda)}(p),a^\dagger_{(\lambda')}(p')] &= a_{(\lambda)}(p)a^\dagger_{(\lambda')}(p')-a^\dagger_{(\lambda')}(p')a_{(\lambda)}(p)\,, &(4.887)\\ &= -\eta_{\lambda\delta}\eta_{\lambda'\delta'}\int d^3X\,d^3Y\,\epsilon^{(\delta)\,\mu}(p)\epsilon^{(\delta')\,\nu}(p')\Big\{\\ &\quad \left[f_p^{(+)*}(X)\dot{A}_\mu(X)-\left(\partial_0 f_p^{(+)*}(X)\right)A_\mu(X)\right]\times\\ &\quad \times\left[\left(\partial_0 f_{p'}^{(+)}(Y)\right)A_\nu(Y)-f_{p'}^{(+)}(Y)\dot{A}_\nu(Y)\right]\\ &\quad -\left[\left(\partial_0 f_{p'}^{(+)}(Y)\right)A_\nu(Y)-f_{p'}^{(+)}(Y)\dot{A}_\nu(Y)\right]\times\\ &\quad \times\left[f_p^{(+)*}(X)\dot{A}_\mu(X)-\left(\partial_0 f_p^{(+)*}(X)\right)A_\mu(X)\right]\Big\}\Big|_{X^0=Y^0}\,,\\ &= -\eta_{\lambda\delta}\eta_{\lambda'\delta'}\int d^3X\,d^3Y\,\epsilon^{(\delta)\,\mu}(p)\epsilon^{(\delta')\,\nu}(p')\Big\{\\ &\quad f_p^{(+)*}(X)\left(\partial_0 f_{p'}^{(+)}(Y)\right)\dot{A}_\mu(X)A_\nu(Y)\\ &\quad -f_p^{(+)*}(X)f_{p'}^{(+)}(Y)\,\dot{A}_\mu(X)\dot{A}_\nu(Y)\\ &\quad -\left(\partial_0 f_p^{(+)*}(X)\right)\left(\partial_0 f_{p'}^{(+)}(Y)\right)A_\mu(X)A_\nu(Y)\\ &\quad +\left(\partial_0 f_p^{(+)*}(X)\right)f_{p'}^{(+)}(Y)\,A_\mu(X)\dot{A}_\nu(Y)\\ &\quad -\left(\partial_0 f_{p'}^{(+)}(Y)\right)f_p^{(+)*}(X)\,A_\nu(Y)\dot{A}_\mu(X)\\ &\quad +\left(\partial_0 f_{p'}^{(+)}(Y)\right)\left(\partial_0 f_p^{(+)*}(X)\right)A_\nu(Y)A_\mu(X)\\ &\quad +f_{p'}^{(+)}(Y)f_p^{(+)*}(X)\,\dot{A}_\nu(Y)\dot{A}_\mu(X)\\ &\quad -f_{p'}^{(+)}(Y)\left(\partial_0 f_p^{(+)*}(X)\right)\dot{A}_\nu(Y)A_\mu(X)\Big\}\Big|_{X^0=Y^0}\,,\end{aligned}$$

$$
\begin{aligned}
&= -\eta_{\lambda\delta}\eta_{\lambda'\delta'} \int d^3X\, d^3Y\, \epsilon^{(\delta)\,\mu}(p)\epsilon^{(\delta')\,\nu}(p')\Big\{ \\
&\quad f_p^{(+)*}(X)(\partial_0 f_{p'}^{(+)}(Y))\, [\dot{A}_\mu(X), A_\nu(Y)]_{X^0=Y^0} \\
&\quad - f_p^{(+)*}(X) f_{p'}^{(+)}(Y)\, [\dot{A}_\mu(X), \dot{A}_\nu(Y)]_{X^0=Y^0} \\
&\quad -(\partial_0 f_p^{(+)*}(X))(\partial_0 f_{p'}^{(+)}(Y))\, [A_\mu(X), A_\nu(Y)]_{X^0=Y^0} \\
&\quad +(\partial_0 f_p^{(+)*}(X)) f_{p'}^{(+)}(Y)\, [A_\mu(X), \dot{A}_\nu(Y)]_{X^0=Y^0} \Big\}\,, \\
&= -\eta_{\lambda\delta}\eta_{\lambda'\delta'} \int d^3X\, d^3Y\, \epsilon^{(\delta)\,\mu}(p)\epsilon^{(\delta')\,\nu}(p')\Big\{ \\
&\quad f_p^{(+)*}(X)(\partial_0 f_{p'}^{(+)}(Y))\, i\eta_{\mu\nu}\delta^3(\mathbf{x}-\mathbf{y}) \\
&\quad +(\partial_0 f_p^{(+)*}(X)) f_{p'}^{(+)}(Y)\, (-i\eta_{\mu\nu})\delta^3(\mathbf{x}-\mathbf{y}) \Big\}\,, \\
&= -\eta_{\lambda\delta}\eta_{\lambda'\delta'}\epsilon^{(\delta)\,\mu}(p)\epsilon_\mu^{(\delta')}(p')\, i \int d^3X\, f_p^{(+)*}(X) \overleftrightarrow{\partial_0} f_{p'}^{(+)}(X)\,, \\
&= -\eta_{\lambda\delta}\eta_{\lambda'\delta'}\eta^{\delta\delta'}\delta^3(\mathbf{p}-\mathbf{p}') \\
&= -\eta_{\lambda\lambda'}\delta^3(\mathbf{p}-\mathbf{p}')
\end{aligned}
\tag{4.888}
$$

and, in the same way we find

$$
[a_{(\lambda)}(p), a_{(\lambda')}(p')] = [a^\dagger_{(\lambda)}(p), a^\dagger_{(\lambda')}(p')] = 0\,. \tag{4.889}
$$

Finally, in summary:

$$
[a_{(\lambda)}(p), a^\dagger_{(\lambda')}(p')] = -\eta_{\lambda\lambda'}\delta^3(\mathbf{p}-\mathbf{p}')\,, \tag{4.890}
$$

$$
[a_{(\lambda)}(p), a_{(\lambda')}(p')] = [a^\dagger_{(\lambda)}(p), a^\dagger_{(\lambda')}(p')] = 0\,. \tag{4.891}
$$

Note the "wrong" sign in Eq. (4.890) for the component 00! This has an important consequence. In fact, if we define a one-particle state as

$$
|1,\lambda\rangle = \int d^3p\, f(p)\, a^\dagger_\lambda(p)|0\rangle\,, \tag{4.892}
$$

its norm comes out to be negative in the case $\lambda = 0$. In fact

$$
\langle 1,\lambda|1,\lambda'\rangle = \int d^3p\, d^3p'\, f^*(p)f(p')\langle 0|a_\lambda(p)a^\dagger_{\lambda'}(p')|0\rangle\,, \tag{4.893}
$$

$$
= \int d^3p\, d^3p'\, f^*(p)f(p')\langle 0|[a_\lambda(p), a^\dagger_{\lambda'}(p')]|0\rangle\,, \tag{4.894}
$$

$$
= -\eta_{\lambda\lambda'}\int d^3p |f(p)|^2\,. \tag{4.895}
$$

For $\lambda = \lambda' = 0$ we find a state with negative norm and therefore it is not physical.

4.6.2 Physical states

Let us consider again the condition (4.855). We want to impose a linear condition on the operators acting on a physical state, such that (4.855) is fulfilled. The field $A_\mu(X)$ has two components

$$
A_\mu(X) = A_\mu^{(+)}(X) + A_\mu^{(-)}(X) = \int \frac{d^3p}{\sqrt{(2\pi)^3 2E}} \sum_{\lambda=0}^{3} \epsilon_\mu^{(\lambda)}(p) a_{(\lambda)}(p) e^{-iP_\mu X^\mu}\,,
$$

$$+\int \frac{d^3p}{\sqrt{(2\pi)^3 2E}} \sum_{\lambda=0}^{3} \epsilon_\mu^{(\lambda)}(p) a^\dagger_{(\lambda)}(p) e^{iP_\mu X^\mu} \,. \tag{4.896}$$

If we impose that a physical state satisfies the following condition

$$\partial^\mu A_\mu^{(+)} |phys\rangle = 0 \,, \tag{4.897}$$

since $A_\mu^{(-)} = (A_\mu^{(+)})^\dagger$ we have that (4.855) is automatically satisfied. In fact

$$0 = \langle phys|(\partial^\mu A_\mu^{(+)})^\dagger = \langle phys|\partial^\mu A_\mu^{(-)} \tag{4.898}$$

and therefore

$$0 = \langle phys|\partial^\mu A_\mu^{(-)} + \partial^\mu A_\mu^{(+)}|phys\rangle = \langle phys|(\partial^\mu A_\mu)|phys\rangle \,. \tag{4.899}$$

Eq. (4.897) is the Gupta-Bleuler condition. Let us see what is $\partial^\mu A_\mu^{(+)}$ in terms of creation/annihilation operators. We have

$$\begin{aligned} \partial^\mu A_\mu^{(+)} &= -i \int \frac{d^3p}{\sqrt{(2\pi)^3 2E}} e^{-iP_\mu X^\mu} P^\mu \sum_{\lambda=0}^{3} \epsilon_\mu^{(\lambda)}(p) a_{(\lambda)}(p) \,, & (4.900) \\ &= |\, \text{since}\, P^\mu \epsilon_\mu^{(1)}(p) = P^\mu \epsilon_\mu^{(2)}(p) = 0 \quad \text{and} \quad P^\mu \epsilon_\mu^{(3)}(p) = -P^\mu \epsilon_\mu^{(0)}(p)\,| \\ &= -i \int \frac{d^3p}{\sqrt{(2\pi)^3 2E}} e^{-iP_\mu X^\mu} (P^\mu \epsilon_\mu^{(0)}(p)) \left[a_{(0)}(p) - a_{(3)}(p)\right] \,. & (4.901) \end{aligned}$$

Therefore, $\partial^\mu A_\mu^{(+)}|phys\rangle = 0$ implies the condition

$$\left[a_{(0)}(p) - a_{(3)}(p)\right] |phys\rangle = 0 \,. \tag{4.902}$$

This means that $|phys\rangle$ are constructed as follows:

$$|phys\rangle = |n_{(0)}, n_{(3)}\rangle = \frac{\left[a^\dagger_{(0)}(p) - a^\dagger_{(3)}(p)\right]^n}{n!} |0\rangle \,. \tag{4.903}$$

In fact we have

$$[a_{(0)}(p') - a_{(3)}(p'), a^\dagger_{(0)}(p) - a^\dagger_{(3)}(p)] = -\delta^3(\mathbf{p} - \mathbf{p}') + \delta^3(\mathbf{p} - \mathbf{p}') = 0 \tag{4.904}$$

and then

$$\begin{aligned} \left[a_{(0)}(p) - a_{(3)}(p)\right] |n_{(0)}, n_{(3)}\rangle &= \left[a_{(0)}(p) - a_{(3)}(p)\right] \frac{\left[a^\dagger_{(0)}(p) - a^\dagger_{(3)}(p)\right]^n}{n!} |0\rangle \,, & (4.905) \\ &= \frac{\left[a^\dagger_{(0)}(p) - a^\dagger_{(3)}(p)\right]^n}{n!} \left[a_{(0)}(p) - a_{(3)}(p)\right] |0\rangle = 0 \,. & (4.906) \end{aligned}$$

$|n_{(0)}, n_{(3)}\rangle$ is the state with n temporal photons and n longitudinal photons. Note that a state $|n_{(0)}, n_{(3)}\rangle$ is the vacuum state for transverse photons

$$a_{(1)}(p)|n_{(0)}, n_{(3)}\rangle = a_{(2)}(p)|n_{(0)}, n_{(3)}\rangle = 0 \,. \tag{4.907}$$

The requirement for a physical state is that it contains the same number of temporal and longitudinal photons but there is no constraint on the number of transverse photons. We then have the following combination

$$|phys\rangle = |\psi_T\rangle + \delta|\phi\rangle \,, \tag{4.908}$$

where

$$|\psi_T\rangle = \alpha\, a^\dagger_{(1)}(p_1)|0\rangle + \beta\, a^\dagger_{(2)}(p_2)|0\rangle\,, \quad |\phi\rangle = |n_{(0)}, n_{(3)}\rangle\,. \tag{4.909}$$

Other states with $a^\dagger_{(0)}(p)|0\rangle$ and $a^\dagger_{(3)}(p)|0\rangle$ not in the combination $|n_{(0)}, n_{(3)}\rangle$ are not physical.

The state vector $|\phi\rangle$ is quite peculiar. It has zero norm. In fact

$$\begin{aligned}
\langle\phi|\phi\rangle &= \langle 0|(a_{(0)}(p) - a_{(3)}(p))(a^\dagger_{(0)}(p) - a^\dagger_{(3)}(p))|0\rangle\,, && (4.910)\\
&= \langle 0|(a_{(0)}(p)a^\dagger_{(0)}(p) + a_{(3)}(p)a^\dagger_{(3)}(p))|0\rangle\,, && (4.911)\\
&= |\,\text{since}\quad [a_{(0)}(p), a^\dagger_{(3)}(p)] = [a_{(3)}(p), a^\dagger_{(0)}(p)] = 0\,| \\
&= \langle 0|([a_{(0)}(p), a^\dagger_{(0)}(p)] + [a_{(3)}(p), a^\dagger_{(3)}(p)])|0\rangle\,, && (4.912)\\
&= |\,\text{since}\quad [a_{(0)}(p), a^\dagger_{(0)}(p)] = -[a_{(3)}(p), a^\dagger_{(3)}(p)]\,| \\
&= 0\,. && (4.913)
\end{aligned}$$

Moreover, $|\phi\rangle$ is ortogonal to $|\psi_T\rangle = \alpha\, a^\dagger_{(1)}(p_1)|0\rangle + \beta\, a^\dagger_{(2)}(p_2)|0\rangle$:

$$\langle\psi_T|\phi\rangle = \langle 0|\left[\alpha^*\, a_{(1)}(p_1) + \beta^*\, a_{(2)}(p_2)\right]\left[a^\dagger_{(0)}(p) - a^\dagger_{(3)}(p)\right]|0\rangle = 0\,. \tag{4.914}$$

This means that any scalar product between physical states are only given by scalar products between the transverse states.

4.6.3 Energy and momentum

Let us look for the expression of the energy and the momentum operators in terms of creation-annihilation operators. Let us note that we are considering the following lagrangian density

$$\mathcal{L} = -\frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu = -\frac{1}{2}\partial_\mu A_0 \partial^\mu A_0 + \frac{1}{2}\partial_\mu A^i \partial^\mu A^i\,, \tag{4.915}$$

which is the sum of three Lagrangian densities of the real fields A^i minus the lagrangian density of the real field A_0. We can therefore immediately understand that we have

$$:H: = \int d^3X\, \mathcal{H} = \int d^3X\, \left[\Pi^\mu \dot{A}_\mu - \mathcal{L}\right] = \int d^3p\, E\, \left[-a^\dagger_{(0)}(p)a_{(0)}(p) + \sum_{\lambda=1}^{3} a^\dagger_{(\lambda)}(p)a_{(\lambda)}(p)\right]\,. \tag{4.916}$$

The same expression holds for the momentum

$$:P^i: = \int d^3p\, p^i\, \left[-a^\dagger_{(0)}(p)a_{(0)}(p) + \sum_{\lambda=1}^{3} a^\dagger_{(\lambda)}(p)a_{(\lambda)}(p)\right]\,. \tag{4.917}$$

If we now evaluate the energy or the momentum of a physical state, we see that they get contributions only from the transverse states. In fact, we have

$$\left[a_{(0)}(p) - a_{(3)}(p)\right]|phys\rangle = 0 \tag{4.918}$$

and

$$\langle phys|\left[-a^\dagger_{(0)}(p)a_{(0)}(p) + a^\dagger_{(3)}(p)a_{(3)}(p)\right]|phys\rangle = \langle phys|\left[-a^\dagger_{(0)}(p) + a^\dagger_{(3)}(p)\right]a_{(0)}(p)|phys\rangle = 0\,. \tag{4.919}$$

Therefore

$$\langle phys|:H:|phys\rangle = \int d^3p\, E\, \langle phys|\sum_{\lambda=1}^{2} a^\dagger_{(\lambda)}(p)a_{(\lambda)}(p)|phys\rangle\,, \tag{4.920}$$

and

$$\langle phys| :P^i: |phys\rangle = \int d^3p\, p^i \, \langle phys| \sum_{\lambda=1}^{2} a^\dagger_{(\lambda)}(p) a_{(\lambda)}(p) |phys\rangle \,, \tag{4.921}$$

We can conclude that the physical state is determined only by the transverse modes. $|\psi_T\rangle$ and $|\psi_T\rangle + c|\phi\rangle$ are physically equivalent. They have the same energy, momentum, angular momentum ... they are physically indistinguishable. They represent the photon.

4.7 Propagator of the Klein-Gordon field

We studied so far the equations of motion of the scalar field without sources. Let us now consider the case in which we are in presence of a source, $j(X)$, which can be for instance a known function of the space-time point. The differential equation fullfiled by the field is

$$(\partial^2 + m^2)\phi(X) = j(X) \,, \tag{4.922}$$

to be intended as a classical equation. The solution of Eq. (4.922) can be obtained calculating the Green function, which is the solution of Eq. (4.922) in presence of a point-like source[19]

$$(\partial^2 + m^2) G(X - X') = \delta^4(X - X') \,, \tag{4.923}$$

such that

$$\phi(X) = \phi^0(X) + \int d^4X' \, G(X - X') j(X') \,, \tag{4.924}$$

where $\phi^0(X)$ is a solution of the homogeneous equation, respecting the given boundary conditions. It is easy to verify that (4.924) satisfies (4.922):

$$(\partial^2 + m^2) \left[\phi^0(X) + \int d^4X' \, G(X - X') j(X') \right] =$$

$$= (\partial^2 + m^2)\phi^0(X) + \int d^4X' \, (\partial^2 + m^2) G(X - X') j(X') \,, \tag{4.925}$$

$$= \int d^4X' \, \delta^4(X - X') j(X') = j(X) \,. \tag{4.926}$$

The problem now is to calculate the Green function. In order to do that, we Fourier transform:

$$G(X - X') \quad = \quad \int \frac{d^4p}{(2\pi)^4} e^{-iP_\mu (X-X')^\mu} \tilde{G}(P) \,, \tag{4.927}$$

$$\delta^4(X - X') \quad = \quad \int \frac{d^4p}{(2\pi)^4} e^{-iP_\mu (X-X')^\mu} \,. \tag{4.928}$$

Substituting in Eq. (4.923) we find

$$(-P^2 + m^2) \, \tilde{G}(P) = 1 \,, \tag{4.929}$$

and therefore

$$\tilde{G}(P) = -\frac{1}{P^2 - m^2} \,. \tag{4.930}$$

Finally

$$G(X - X') = -\int \frac{d^4P}{(2\pi)^4} e^{-iP_\mu (X-X')^\mu} \frac{1}{P^2 - m^2} \,. \tag{4.931}$$

[19] For translationally invariant systems the Green function is a function of $(X - X')$

The calculation of Eq. (4.931) has to be done in the complex plane, puting attention to the fact that the integrand has poles in the domain of integration. We can for instance integrate in dP^0. In this case we have two single poles on the real axis at

$$P^2 - m^2 = 0 \Longrightarrow P^0 = \pm\sqrt{p^2 + m^2} = \pm\omega\,. \tag{4.932}$$

Therefore, in order to perform the integration we have different choices, according to how we avoid the singularity in integrating in dP^0. The integration will be done in principal value and the infinitesimal arc with which to circumvent the poles can be chosen in the upper or in the lower complex semi-plane. The difference will be a residue, i.e. the solution of the homogeneous equation. The coice on the path depends on the boundary conditions.

4.7.1 Closed paths and residues

Let us consider the integration in P^0 on a closed path around one of the poles.

If C^+ is a closed positive path around $P^0 = \sqrt{p^2 + m^2} = \omega$, we can apply the residue theorem finding

$$\Delta^+ = -i\int \frac{d^3P}{(2\pi)^3}\int_{C^+} \frac{dP^0}{(2\pi)} \frac{e^{-iP_\mu X^\mu}}{(P^0-\omega)(P^0+\omega)} \tag{4.933}$$

$$= \int \frac{d^3P}{(2\pi)^3 2\omega} e^{-i(\omega t - \mathbf{p}\cdot\mathbf{x})}\,. \tag{4.934}$$

If C^- is a closed positive path around $P^0 = -\sqrt{p^2 + m^2} = -\omega$, instead, we get

$$\Delta^- = -i\int \frac{d^3P}{(2\pi)^3}\int_{C^-} \frac{dP^0}{(2\pi)} \frac{e^{-iP_\mu X^\mu}}{(P^0-\omega)(P^0+\omega)}\,, \tag{4.935}$$

$$= -\int \frac{d^3P}{(2\pi)^3 2\omega} e^{i(\omega t + \mathbf{p}\cdot\mathbf{x})}\,, \tag{4.936}$$

$$= |\,\text{transforming} \quad \mathbf{p} \to -\mathbf{p}\,|$$

$$= -\int \frac{d^3P}{(2\pi)^3 2\omega} e^{i(\omega t - \mathbf{p}\cdot\mathbf{x})}\,. \tag{4.937}$$

Both $\Delta^\pm$ are solution of the homogeneous equation. In fact

$$(\partial^2 + m^2)\Delta^\pm = -i\int \frac{d^3P}{(2\pi)^3}\int_{C^\pm} \frac{dP^0}{(2\pi)} (\partial^2 + m^2) \frac{e^{-iP_\mu X^\mu}}{(P^2 - m^2)}\,, \tag{4.938}$$

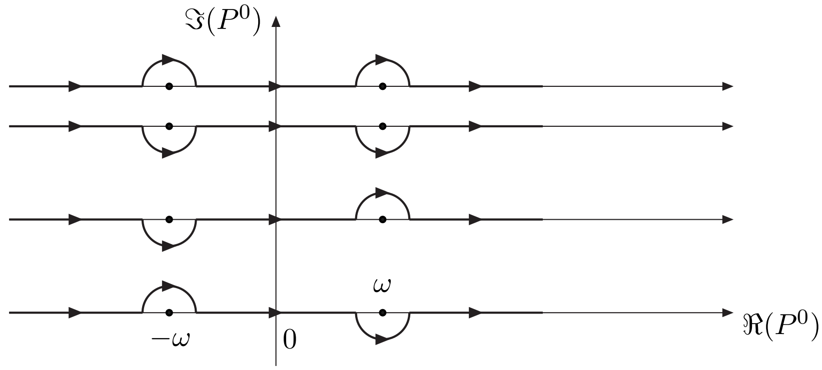
$$= i\int \frac{d^3P}{(2\pi)^3}\int_{C^\pm} \frac{dP^0}{(2\pi)} (P^2 - m^2) \frac{e^{-iP_\mu X^\mu}}{(P^2 - m^2)}\,, \tag{4.939}$$

$$= i\int \frac{d^3P}{(2\pi)^3}\int_{C^\pm} \frac{dP^0}{(2\pi)} e^{-iP_\mu X^\mu} = 0\,, \tag{4.940}$$

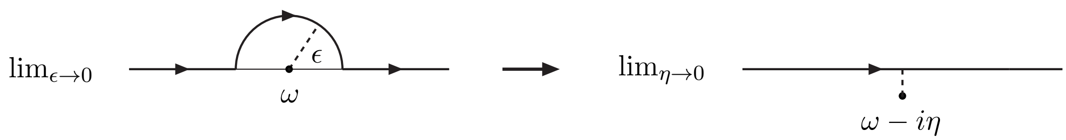
for Cauchy's theorem.

4.7.2 Open paths

The integration on an open path gives the solution for the non-homogeneous differential equation. Of particular interest are the so-called "retarded" and "advanced" Green functions. These provide the correct solutions for a classical field that preserves causality. They depend on how we regularize the two singularities on the real P^0 axis, that occur at $P^0 = \pm\omega$, where $\omega = \sqrt{p^2 + m^2}$. In principle, we have four possibilities to perform the integral: we can get around both singularities with a vanishing circle on the upper complex half-plane, or both with a vanishing circle on the lower complex half-plane, or we can use a circle on the upper half-plane and one in the lower one (with two evident configurations).



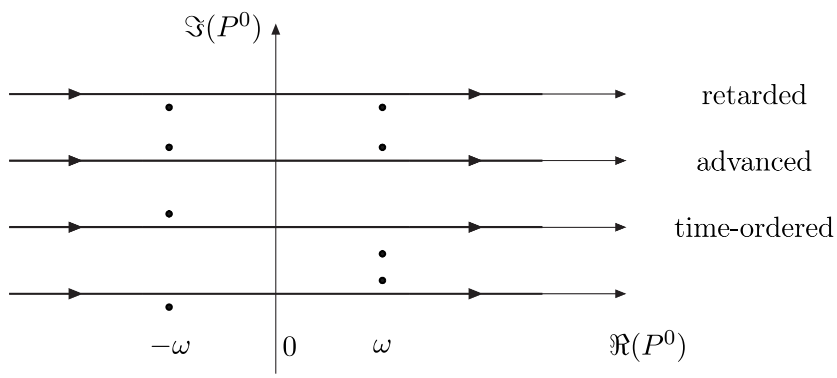
Using a different language (but same result), instead of considering a vanishing circle around the pole, we can displace the pole (using a vanishing imaginary part) keeping the integration on the real axis, as in the figure



such that

$$\lim_{\epsilon\to 0}\int_{\gamma_\epsilon}\frac{dP^0}{(2\pi)}\frac{f(P^0)}{P^0-\omega}=\lim_{\eta\to 0}\int\frac{dP^0}{(2\pi)}\frac{f(P^0)}{P^0-\omega+i\eta} \tag{4.941}$$

and usually the "limit" procedure is understood. Therefore, the situation becomes as follows:

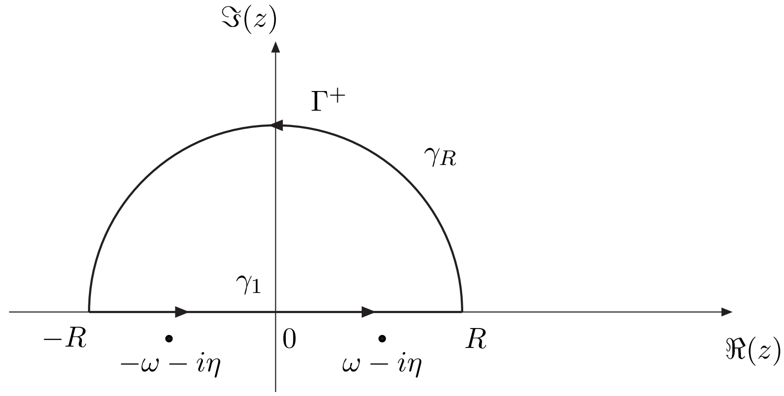


Retarded Green functions

The first case to be considered is the one in which the two poles are both displaced below the real axis. In this way, the Green function vanishes for $t < t'$. In fact, we define

$$\tilde{G}_{ret}(P) = -\frac{1}{(P^0+i\eta)^2-p^2-m^2} = -\frac{1}{(P^0-\omega+i\eta)(P^0+\omega+i\eta)}\,. \tag{4.942}$$

If we close the integration contour in the upper half-plane, for the case $t < t'$,



and we let $R \to \infty$. For Cauchy's theorem we have

$$0 = -\lim_{R\to\infty}\int d^3P\int_{\Gamma^+} dP^0 \frac{e^{-iP_\mu(X-X')^\mu}}{(P^0-\omega+i\eta)(P^0+\omega+i\eta)}\,, \tag{4.943}$$

$$= -\lim_{R\to\infty}\int d^3P\int_{-R}^{R} dP^0 \frac{e^{-iP_\mu(X-X')^\mu}}{(P^0-\omega+i\eta)(P^0+\omega+i\eta)}$$
$$-\lim_{R\to\infty}\int d^3P\int_{\gamma_R} dP^0 \frac{e^{-iP_\mu(X-X')^\mu}}{(P^0-\omega+i\eta)(P^0+\omega+i\eta)}\,, \tag{4.944}$$

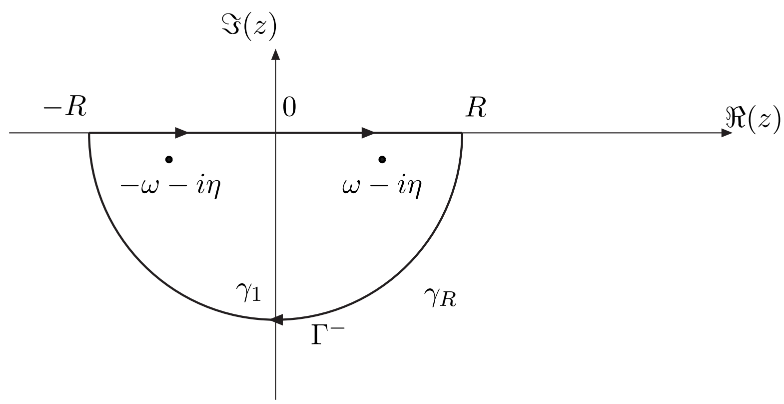
and for Jordan's lemma, we have that (for $t-t'<0$)

$$\lim_{R\to\infty}\int d^3P\int_{\gamma_R} dP^0 \frac{e^{-iP_\mu(X-X')^\mu}}{(P^0-\omega+i\eta)(P^0+\omega+i\eta)} = 0\,. \tag{4.945}$$

Therefore:

$$G_{ret}(X-X') = 0\,, \qquad \text{for} \quad t-t'<0\,. \tag{4.946}$$

If $t-t'>0$, instead, we have to close the integration contour in the lower half P^0 plane, in order to use Jordan's lemma. This means that we are including in the contour the two poles.



Now the residues theorem gives us (remember we are closing the contour clock-wise):

$$-2\pi i\sum \text{Res}(f,\pm\omega) = -\lim_{R\to\infty}\int_{\Gamma^-} dP^0 \frac{e^{-iP_\mu(X-X')^\mu}}{(P^0-\omega+i\eta)(P^0+\omega+i\eta)}\,, \tag{4.947}$$

$$= -\lim_{R\to\infty}\int_{-R}^{R} dP^0 \frac{e^{-iP_\mu(X-X')^\mu}}{(P^0-\omega+i\eta)(P^0+\omega+i\eta)}\,. \tag{4.948}$$

Finally

$$G_{ret}(X-X') = -\frac{\theta(X^0-X'^0)}{(2\pi)^4}\int d^4P \frac{e^{-iP_\mu(X-X')^\mu}}{(P^0+i\eta)^2-p^2-m^2}\,, \tag{4.949}$$

$$= -\frac{\theta(X^0 - X'^0)}{(2\pi)^4} \int \frac{d^4P}{2\omega} e^{-iP_\mu(X-X')^\mu} \left[\frac{1}{(P^0 - \omega + i\eta)} - \frac{1}{(P^0 + \omega + i\eta)}\right], \quad (4.950)$$

$$= -\frac{\theta(X^0 - X'^0)}{(2\pi)^4} \int \frac{d^3P}{2\omega} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} \int dP^0\, e^{-iP_0(X^0-X'^0)} \left[\frac{1}{(P^0 - \omega + i\eta)} - \frac{1}{(P^0 + \omega + i\eta)}\right], \quad (4.951)$$

$$= \left| \text{for the residues theorem Res} \left(\frac{e^{-iP_0(X^0-X'^0)}}{(P^0 \pm \omega + i\eta)}, \mp\omega\right) = -2\pi i e^{\pm i\omega(X^0-X'^0)} \right|$$

$$= i\frac{\theta(X^0 - X'^0)}{(2\pi)^3} \int \frac{d^3P}{2\omega} \left[e^{-i\omega(X^0-X'^0)+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} - e^{i\omega(X^0-X'^0)+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}\right] \quad (4.952)$$

$$= |\, \mathbf{p} \to -\mathbf{p} \text{ in the second integral} \,|$$

$$= i\frac{\theta(X^0 - X'^0)}{(2\pi)^3} \int \frac{d^3P}{2\omega} \left[e^{-i\omega(X^0-X'^0)+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} - e^{i\omega(X^0-X'^0)-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}\right] \quad (4.953)$$

$$= \theta(X^0 - X'^0)\left(i\Delta^+ + i\Delta^-\right). \quad (4.954)$$

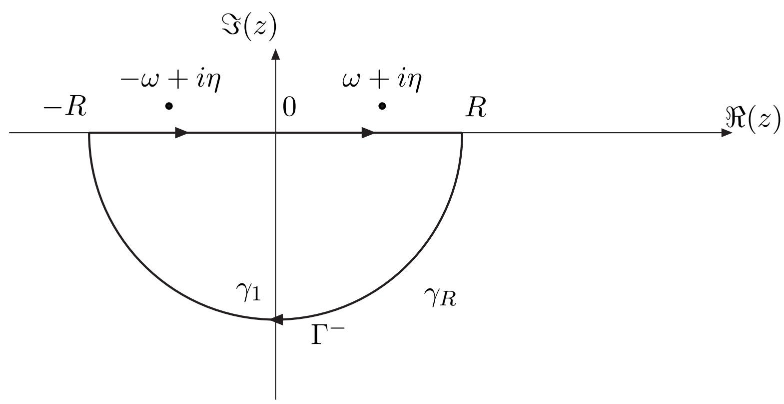
The Green function $G_{ret}(X - X')$ is real and transports in the future both solutions, with positive or negative frequency. It has to be used in problems in which we have the boundary at a certain t' and we ask what happens in consequence of that, for $t > t'$. It is a causal Green function in the sense that it is different from zero in the future light-cone of X'. For space-like separations, $(X - X')^2 < 0$, since it is invariant under proper Lorentz transformations, it vanishes. In fact, if $(X - X')^2 < 0$, we can find a frame in which $t < t'$, for which then $G_{ret}(X - X') = 0$ and it remains zero in every frame.

Advanced Green functions

The second case is constituted by the advanced Green function, which is defined to vanish for $t - t' > 0$. We define

$$\tilde{G}_{adv}(P) = -\frac{1}{(P^0 - i\eta)^2 - p^2 - m^2} = -\frac{1}{(P^0 - \omega - i\eta)(P^0 + \omega - i\eta)}. \quad (4.955)$$

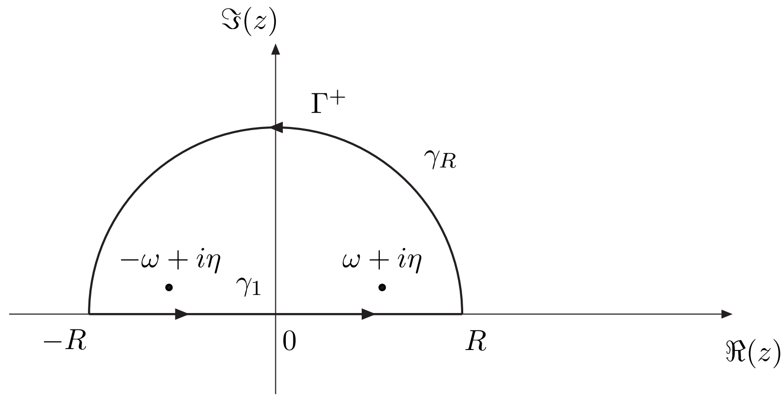
If we close the integration contour in the lower half-plane, for the case $t > t'$,



and we find

$$G_{adv}(X - X') = 0, \qquad \text{for} \quad t - t' > 0. \quad (4.956)$$

If $t - t' < 0$, instead, we have to close the integration contour in the upper half P^0 plane, in order to use Jordan's lemma. This means that we are including in the contour the two poles.

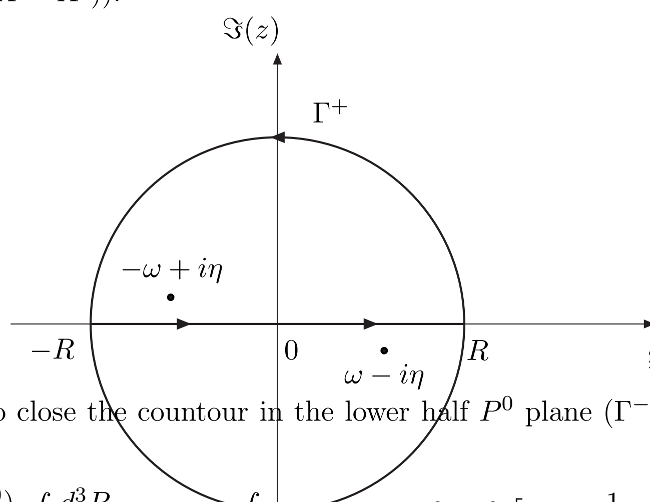


For the residues theorem we have

$$
\begin{aligned}
G_{adv}(X-X') &= -\frac{\theta(X'^0-X^0)}{(2\pi)^4}\int d^4P\frac{e^{-iP_\mu(X-X')^\mu}}{(P^0-i\eta)^2-p^2-m^2}\,, &(4.957)\\
&= -\frac{\theta(X'^0-X^0)}{(2\pi)^4}\int \frac{d^4P}{2\omega}e^{-iP_\mu(X-X')^\mu}\left[\frac{1}{(P^0-\omega-i\eta)}-\frac{1}{(P^0+\omega-i\eta)}\right], &(4.958)\\
&= \left|\text{for the residues theorem}\,\mathrm{Res}\left(\frac{e^{-iP_0(X^0-X'^0)}}{(P^0\pm\omega-i\eta)},\mp\omega\right)=2\pi i e^{\pm i\omega(X^0-X'^0)}\right|\\
&= -i\frac{\theta(X'^0-X^0)}{(2\pi)^3}\int\frac{d^3P}{2\omega}\left[e^{-i\omega(X^0-X'^0)+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}-e^{i\omega(X^0-X'^0)+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}\right] &(4.959)\\
&= |\,\mathbf{p}\to-\mathbf{p}\text{ in the second integral}\,|\\
&= -i\frac{\theta(X^0-X'^0)}{(2\pi)^3}\int\frac{d^3P}{2\omega}\left[e^{-i\omega(X^0-X'^0)+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}-e^{i\omega(X^0-X'^0)-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}\right] &(4.960)\\
&= -\theta(X'^0-X^0)\left(i\Delta^++i\Delta^-\right)\,. &(4.961)
\end{aligned}
$$

Also $G_{adv}(X-X')$ is real and it transports in the past both solutions, with positive or negative frequency. It has to be used in problems in which we have the boundary at a certain t' in the future and we ask what happens in the present in order to cause this boundary in the future. It is not obvious, but it is possible. It is a causal Green function in the sense that it is different from zero in the past light-cone of X' and for space-like separations, $(X-X')^2<0$, since it is invariant under proper Lorentz transformations, it vanishes (like $G_{ret}(X-X')$).

Feynman propagator



Quantum-mechanically the correct propagator, that propagates "particle" and "anti-particle" states in the future, is the Feynman propagator. It is defined giving a vanishing positive immaginary part to the pole in $-\omega$ and a vanishing negative immaginary part to the pole in ω (see figure).

This time, for $t-t'>0$ we have to close the countour in the lower half P^0 plane (Γ^-), while for $t-t'<0$ in the upper (Γ^+). We have

$$
\begin{aligned}
D_F(X-X') &= -\frac{\theta(X^0-X'^0)}{(2\pi)^4}\int\frac{d^3P}{2\omega}e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}\int_{\Gamma^-}dP^0\,e^{-iP_0(X^0-X'^0)}\Big[\frac{1}{(P^0-\omega+i\eta)}\\
&\qquad-\frac{1}{(P^0+\omega-i\eta)}\Big]\,, &(4.962)\\
&\quad-\frac{\theta(X'^0-X^0)}{(2\pi)^4}\int\frac{d^3P}{2\omega}e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}\int_{\Gamma^+}dP^0\,e^{-iP_0(X^0-X'^0)}\Big[\frac{1}{(P^0-\omega+i\eta)}
\end{aligned}
$$

$$-\frac{1}{(P^0+\omega-i\eta)}\Bigg], \tag{4.963}$$

$$= i\theta(X^0-X'^0)\int\frac{d^3P}{(2\pi)^3 2\omega}e^{-i\omega(t-t')+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}$$

$$+i\theta(X'^0-X^0)\int\frac{d^3P}{(2\pi)^3 2\omega}e^{i\omega(t-t')+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}, \tag{4.964}$$

$$= i\theta(X^0-X'^0)\int\frac{d^3P}{(2\pi)^3 2\omega}e^{-i\omega(t-t')+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}$$

$$+i\theta(X'^0-X^0)\int\frac{d^3P}{(2\pi)^3 2\omega}e^{i\omega(t-t')-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}, \tag{4.965}$$

$$= \theta(X^0-X'^0)\,i\Delta^+(X-X')-\theta(X'^0-X^0)\,i\Delta^-(X-X'). \tag{4.966}$$

Now, the Green function is complex and it propagates in the future the positive frequency solutions and in the past the negative frequency ones. This is consistent with Dirac's "hole theory" interpretation of particle and anti-particle states propagating both in the future.

We can have a physical interpretation of the meaning of the propagator considering the following simple "quantum process" associated to the charged KG field $\phi(X)$. Let us consider the creation of a particle (one-particle state) at the time t in $\mathbf{y}$. We will have

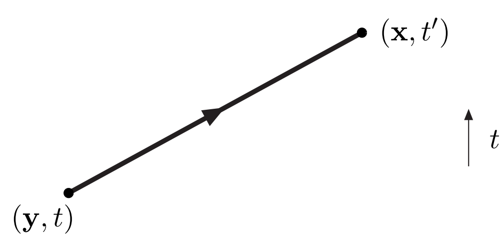
$$|\psi(\mathbf{y},t)\rangle = \phi^\dagger(Y)|0\rangle = \int d^3P\left[f_p^{(+)}b(p)|0\rangle + f_p^{(+)*}a^\dagger(p)|0\rangle\right], \tag{4.967}$$

$$= \int\frac{d^3P}{\sqrt{(2\pi)^3 2\omega}}e^{iP_\mu Y^\mu}a^\dagger(p)|0\rangle. \tag{4.968}$$

The probability amplitude of finding the particle in $\mathbf{x}$ at $t'>t$ is given by

$$\theta(t'-t)\langle\psi(\mathbf{x},t')|\psi(\mathbf{y},t)\rangle = \theta(t'-t)\langle 0|\phi(X)\phi^\dagger(Y)|0\rangle, \tag{4.969}$$

that can be interpreted as the creation of a particle of charge $q=+1$ in $(\mathbf{y},t)$ by $\phi^\dagger(Y)$, its propagation up to $(\mathbf{x},t')$ and its annihilation in this point by $\phi(X)$.



Such a relation enters a scattering process, where two nucleons (a proton and a neutron) exchange a charged pion (see Fig. 4.1 (a)). The same "effect" can be recovered creating a negative charge in $(\mathbf{x},t')$, that then propagates up to $(\mathbf{y},t)$ and is annihilated in this point, with $t>t'$. Therefore we have to consider also the amplitude

$$\theta(t-t')\,\langle 0|\phi^\dagger(Y)\phi(X)|0\rangle, \tag{4.970}$$

that enters for instance the diagram in Fig. 4.1 (b). The complete amplitude will be the sum of the two amplitudes:

$$\langle 0|T(\phi(X)\phi^\dagger(Y))|0\rangle = \langle 0|T(\phi^\dagger(Y)\phi(X))|0\rangle \tag{4.971}$$

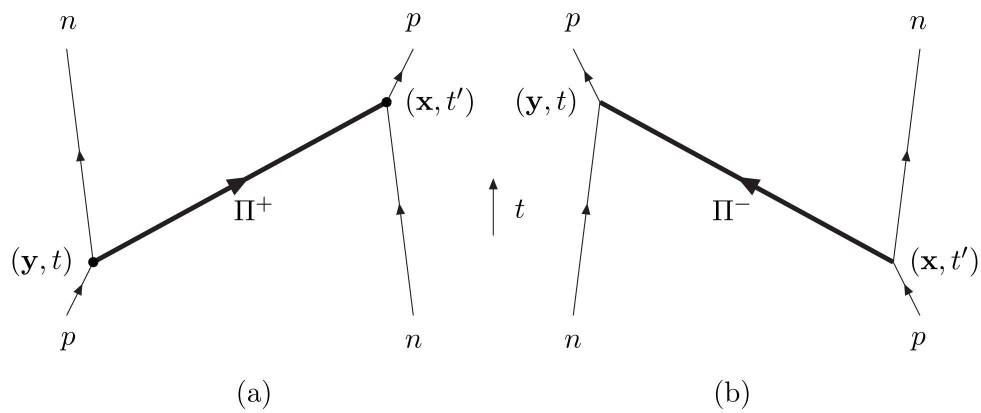


Figura 4.1: Scattering of a proton and a neutron

$$\begin{aligned}
&= \theta(X^0-Y^0)\langle 0|\phi(X)\phi^\dagger(Y)|0\rangle + \theta(Y^0-X^0)\langle 0|\phi^\dagger(Y)\phi(X)|0\rangle & (4.972)\\
&= \int \frac{d^3P\, d^3P'}{(2\pi)^3\sqrt{4\omega\omega'}}\Big[\theta(X^0-Y^0)e^{i(P_\mu Y^\mu - P'_\mu X^\mu)}\langle 0|a(p')a^\dagger(p)|0\rangle \\
&\qquad + \theta(Y^0-X^0)e^{-i(P_\mu Y^\mu - P'_\mu X^\mu)}\langle 0|b(p)b^\dagger(p')|0\rangle\Big]\,, & (4.973)\\
&= |\,\text{since } \langle 0|a(p')a^\dagger(p)|0\rangle = \delta(\mathbf{p}-\mathbf{p}')...\,| \\
&= \int \frac{d^3P}{(2\pi)^3 2\omega}\Big[\theta(X^0-Y^0)e^{-iP_\mu(X-Y)^\mu} + \theta(Y^0-X^0)e^{iP_\mu(X-Y)^\mu}\Big] & (4.974)\\
&= -iD_F(X-Y)\,. & (4.975)
\end{aligned}$$

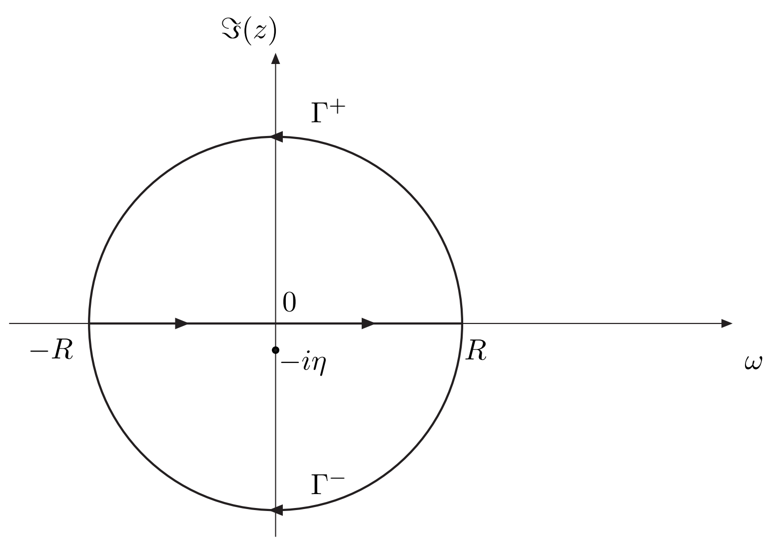
Where we defined the *time-ordered product* of the two bosonic fields $\phi(X)$ and $\phi^\dagger(Y)$) as follows:

$$T(\phi(X)\phi^\dagger(Y)) = T(\phi^\dagger(Y)\phi(X)) = \theta(X^0-Y^0)\,\phi(X)\phi^\dagger(Y) + \theta(Y^0-X^0)\,\phi^\dagger(Y)\phi(X)\,. \quad (4.976)$$

A more convenient way to write the Feynman propagator is using the following integral representation for the step function:

$$\theta(t) = \lim_{\eta\to 0^+}\frac{i}{2\pi}\int_{-\infty}^{\infty} d\omega\, \frac{e^{-i\omega t}}{\omega + i\eta}\,. \quad (4.977)$$

In fact,



for $t < 0$ the integral contour has to be chosen to be Γ^+ in such a way that we can apply Cauchy's theorem and Jordan's lemma, getting

$$0 = \theta(t) + \lim_{R\to\infty} \int_{\gamma_{R^+}} \frac{i}{2\pi} d\omega \, \frac{e^{-i\omega t}}{\omega + i\eta} = \theta(t) \, . \tag{4.978}$$

For $t > 0$, instead, we close the integral contour in the lower complex plane (Γ^-) getting

$$-2\pi i \, \mathrm{Res}(\theta, -i\eta) = \theta(t) + \lim_{R\to\infty} \int_{\gamma_{R^-}} \frac{i}{2\pi} d\omega \, \frac{e^{-i\omega t}}{\omega + i\eta} = \theta(t) \, , \tag{4.979}$$

where

$$-2\pi i \, \mathrm{Res}(\theta, -i\eta) = -2\pi i \lim_{\eta\to 0^+} \frac{i}{2\pi} e^{-\eta t} = 1 \, . \tag{4.980}$$

Including the integral representation of the Heaviside θ in Eq. (4.975) we find

$$\begin{aligned}
-iD_F(X-Y) &= \int \frac{d^3P}{(2\pi)^3 2\omega_p} \Big[\theta(X^0 - Y^0) e^{-iP_\mu (X-Y)^\mu} + \theta(Y^0 - X^0) e^{iP_\mu (X-Y)^\mu} \Big] \, , && (4.981) \\
&= i \int \frac{d^3P}{(2\pi)^4} \int \frac{d\omega}{2\omega_p} \Big[\frac{e^{-i\omega(X^0-Y^0)}}{\omega + i\eta} e^{-iP_\mu (X-Y)^\mu} + \frac{e^{i\omega(X^0-Y^0)}}{\omega + i\eta} e^{iP_\mu (X-Y)^\mu} \Big] , && (4.982) \\
&= | \text{we substitute } P^0 = \omega + \omega_p \, , \text{ such that } \omega = P^0 - \omega_p \text{ and } d\omega = dP^0 \, | \\
&= i \int \frac{d^4P}{(2\pi)^4} \frac{1}{2\omega_p} \Big[\frac{e^{-iP^0(X^0-Y^0) + i\omega_p (X^0-Y^0)} \, e^{-i\omega_p (X^0-Y^0) + i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}}{(P^0 - \omega_p + i\eta)} \\
&\quad + \frac{e^{iP^0(X^0-Y^0) - i\omega_p (X^0-Y^0)} \, e^{i\omega_p (X^0-Y^0) - i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}}{(P^0 - \omega_p + i\eta)} \Big] \, , && (4.983) \\
&= i \int \frac{d^4P}{(2\pi)^4} \frac{1}{2\omega_p} \Big[\frac{e^{-iP_\mu (X-Y)^\mu}}{(P^0 - \omega_p + i\eta)} + \frac{e^{iP_\mu (X-Y)^\mu}}{(P^0 - \omega_p + i\eta)} \Big] \, , && (4.984) \\
&= | \text{substituting in the second integral } P^\mu \to -P^\mu \, | \\
&= i \int \frac{d^4P}{(2\pi)^4} \frac{1}{2\omega_p} e^{-iP_\mu (X-Y)^\mu} \Big[\frac{1}{(P^0 - \omega_p + i\eta)} - \frac{1}{(P^0 + \omega_p - i\eta)} \Big] \, , && (4.985) \\
&= i \int \frac{d^4P}{(2\pi)^4} \frac{e^{-iP_\mu (X-Y)^\mu}}{P^2 - m^2 + i\eta} \, . && (4.986)
\end{aligned}$$

Finally

$$D_F(X-Y) = -\int \frac{d^4P}{(2\pi)^4} \frac{e^{-iP_\mu (X-Y)^\mu}}{P^2 - m^2 + i\eta} \, . \tag{4.987}$$

We can check again that D_F is a Green function for the KG operator:

$$\begin{aligned}
(\partial^2 + m^2)_X D_F(X-y) &= -\lim_{\eta\to 0} \int \frac{d^4P}{(2\pi)^4} (\partial^2 + m^2)_X \frac{e^{-iP_\mu (X-Y)^\mu}}{P^2 - m^2 + i\eta} \, , && (4.988) \\
&= -\lim_{\eta\to 0} \int \frac{d^4P}{(2\pi)^4} (-P^2 + m^2) \frac{e^{-iP_\mu (X-Y)^\mu}}{P^2 - m^2 + i\eta} \, , && (4.989) \\
&= \int \frac{d^4P}{(2\pi)^4} e^{-iP_\mu (X-Y)^\mu} = \delta^4(X-Y) \, . && (4.990)
\end{aligned}$$

The same result can be found, acting with $(\partial^2 + m^2)_X$ (derivatives with respect to X) directly on $i\langle 0|T(\phi(X)\phi^\dagger(Y))|0\rangle$. In fact:

$$(\partial^2 + m^2)_X \, i\langle 0|T(\phi(X)\phi^\dagger(Y))|0\rangle = \partial_0^2 \, i\langle 0|T(\phi(X)\phi^\dagger(Y))|0\rangle$$

$$
\begin{aligned}
& \qquad +i\langle 0|T((-\nabla+m^2)_X\phi(X)\phi^\dagger(Y))|0\rangle\,, && (4.991)\\
= \;& \partial_0\, i\langle 0|\delta(X^0-Y^0)[\phi(X),\phi^\dagger(Y)]|0\rangle + \partial_0\, i\langle 0|T(\dot\phi(X)\phi^\dagger(Y))|0\rangle \\
& \qquad +i\langle 0|T((-\nabla+m^2)_X\phi(X)\phi^\dagger(Y))|0\rangle\,, && (4.992)\\
= \;& \partial_0\, i\langle 0|T(\dot\phi(X)\phi^\dagger(Y))|0\rangle \\
& \qquad +i\langle 0|T((-\nabla+m^2)_X\phi(X)\phi^\dagger(Y))|0\rangle\,, && (4.993)\\
= \;& i\langle 0|\delta(X^0-Y^0)[\dot\phi(X),\phi^\dagger(Y)]|0\rangle + i\langle 0|T(\ddot\phi(X)\phi^\dagger(Y))|0\rangle \\
& \qquad +i\langle 0|T((-\nabla+m^2)_X\phi(X)\phi^\dagger(Y))|0\rangle\,, && (4.994)\\
= \;& \delta^4(X-Y)\,, && (4.995)
\end{aligned}
$$

where we used the fact that[20]

$$
\partial_0\theta(X^0-Y^0)=\delta(X^0-Y^0)\,,\quad \text{and}\quad \partial_0\theta(Y^0-X^0)=-\delta(X^0-Y^0) \qquad (4.998)
$$

and we used the commutation relations of the fields.

The propagator for the real KG field is

$$
i\langle 0|T(\phi(X)\phi(Y))|0\rangle = D_F(X-Y)\,. \qquad (4.999)
$$

Propagators and commutators

4.8 Propagator of the Dirac field

In the case of the Dirac field we define the propagator as in the case of KG field

$$
S_F(X-Y)_{\alpha\beta} = -i\langle 0|T(\psi_\alpha(X)\overline{\psi}_\beta(Y))|0\rangle\,, \qquad (4.1000)
$$

but now, since we are dealing with fermions, the T-ordered product is defined as follows:

$$
T(\psi_\alpha(X)\overline{\psi}_\beta(Y)) = -T(\overline{\psi}_\beta(Y)\psi_\alpha(X)) = \theta(X^0-Y^0)\psi_\alpha(X)\overline{\psi}_\beta(Y)-\theta(Y^0-X^0)\overline{\psi}_\beta(Y)\psi_\alpha(X)\,. \qquad (4.1001)
$$

$S_F(X-Y)$ is indeed a Green's function for the Dirac equation. In fact

$$
\begin{aligned}
(i\,\partial\!\!\!/_X - m)_{\alpha\beta}\langle 0|T(\psi_\beta(X)\overline{\psi}_\gamma(Y))|0\rangle \;=\;& \langle 0|i\gamma^0_{\alpha\beta}\delta(X^0-Y^0)[\psi_\beta(X),\psi^\dagger_\delta(Y)]_+\gamma^0_{\delta\gamma}|0\rangle\,,\\
& +\langle 0|T(i\gamma^0_{\alpha\beta}\partial_0\psi_\beta(X)\overline{\psi}_\gamma(Y))|0\rangle\,,\\
& +\langle 0|T([i\gamma^i_{\alpha\beta}\partial_i - m]\psi_\beta(X)\overline{\psi}_\gamma(Y))|0\rangle\,, && (4.1002)\\
=\;& |\,\text{since } [\psi_\beta(\mathbf{x},t),\psi^\dagger_\delta(\mathbf{y},t)]_+ = \delta_{\beta\delta}\delta^3(\mathbf{x}-\mathbf{y})\,|\\
=\;& \langle 0|i\gamma^0_{\alpha\beta}\delta(X^0-Y^0)\delta_{\beta\delta}\delta^3(\mathbf{x}-\mathbf{y})\gamma^0_{\delta\gamma}|0\rangle\,,\\
& +\langle 0|T([i\gamma^\mu_{\alpha\beta}\partial_\mu - m]\psi_\beta(X)\overline{\psi}_\gamma(Y))|0\rangle\,, && (4.1003)\\
=\;& i\delta_{\alpha\gamma}\delta^4(X-Y)\,. && (4.1004)
\end{aligned}
$$

Therefore

$$
(i\,\partial\!\!\!/_X - m)_{\alpha\beta}S_F(X-Y)_{\beta\gamma} = \delta_{\alpha\gamma}\delta^4(X-Y)\,. \qquad (4.1005)
$$

[20] These relations can be demonstrated using the integral representation for the Heaviside function. We have

$$
\begin{aligned}
\frac{\partial}{\partial t}\theta(t) \;=\;& \lim_{\eta\to 0^+}\frac{\partial}{\partial t}\frac{i}{2\pi}\int_{-\infty}^{\infty} d\omega\,\frac{e^{-i\omega t}}{\omega+i\eta} = \lim_{\eta\to 0^+}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\frac{\omega e^{-i\omega t}}{\omega+i\eta} = \int_{-\infty}^{\infty}\frac{d\omega}{2\pi}e^{-i\omega t} = \delta(t)\,, && (4.996)\\
\frac{\partial}{\partial t}\theta(-t) \;=\;& \lim_{\eta\to 0^+}\frac{\partial}{\partial t}\frac{i}{2\pi}\int_{-\infty}^{\infty} d\omega\,\frac{e^{i\omega t}}{\omega+i\eta} = -\lim_{\eta\to 0^+}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\frac{\omega e^{i\omega t}}{\omega+i\eta}\\
=\;& |\,\omega\to-\omega\,| = -\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}e^{-i\omega t} = -\delta(t) && (4.997)
\end{aligned}
$$

The Green's function $S_F(X-Y)_{\beta\gamma}$ can be expressed in terms of $D_F(X-Y)$ as follows

$$S_F(X-Y)_{\beta\gamma} = -(i\,\partial\!\!\!/_X + m)_{\beta\gamma} D_F(X-Y)\,. \tag{4.1006}$$

In fact, we have

$$\begin{aligned}(i\,\partial\!\!\!/_X - m)_{\alpha\beta} S_F(X-Y)_{\beta\gamma} &= -(i\,\partial\!\!\!/_X - m)_{\alpha\beta}(i\,\partial\!\!\!/_X + m)_{\beta\gamma} D_F(X-Y)\,,\\ &= \delta_{\alpha\gamma}(\partial^2 + m^2) D_F(X-Y) = \delta_{\alpha\gamma}\delta^4(X-Y)\,. \end{aligned}\tag{4.1007}$$

Therefore

$$S_F(X-Y)_{\beta\gamma} = -(i\,\partial\!\!\!/_X + m)_{\beta\gamma}\int \frac{d^4P}{(2\pi)^4}\left[-\frac{e^{-iP_\mu(X-Y)^\mu}}{P^2-m^2+i\eta}\right]\,, \tag{4.1008}$$

$$= \int \frac{d^4P}{(2\pi)^4} e^{-iP_\mu(X-Y)^\mu}\frac{(P\!\!\!/ + m)_{\beta\gamma}}{P^2-m^2+i\eta}\,. \tag{4.1009}$$

4.9 Propagator of the Electromagnetic field

Finally, the propagator for the electromagnetic field will be given by the following expression:

$$\langle 0|T\left(A_\mu(X)A_\nu(Y)\right)|0\rangle = i\eta_{\mu\nu}D(X-Y) = \int \frac{d^4P}{(2\pi)^4} e^{-iP_\mu X^\mu}\frac{-i\eta_{\mu\nu}}{P^2+i\eta}\,. \tag{4.1010}$$

The field $A_\mu(X)$ is a bosonic field and the T-ordered product has to be defined as in the case of the KG field:

$$T\left(A_\mu(X)A_\nu(Y)\right) = T\left(A_\nu(Y)A_\mu(X)\right) = \theta(X^0-Y^0)A_\mu(X)A_\nu(Y)+\theta(Y^0-X^0)A_\nu(Y)A_\mu(X)\,. \tag{4.1011}$$

We can check that the expression in Eq. (4.1010) is indeed a Green function for the equation of motion[21]

$$\partial^2 A_\mu(X) = J^\mu(X)\,, \tag{4.1015}$$

i.e. a function $G_{\mu\nu}(X-Y)$ such that

$$\partial_X^2 G_{\mu\nu}(X-Y) = \eta_{\mu\nu}\delta^4(X-Y)\,. \tag{4.1016}$$

In fact we have

$$\begin{aligned}\partial_X^2\left(-i\langle 0|T\left(A_\mu(X)A_\nu(Y)\right)|0\rangle\right) &= -i\,\partial_0\langle 0|\delta(X^0-Y^0)[A_\mu(X),A_\nu(Y)]|0\rangle\\ &\quad -i\partial_0\langle 0|T\left(\dot{A}_\mu(X)A_\nu(Y)\right)|0\rangle\\ &\quad +i\langle 0|T\left(\nabla_X^2 A_\mu(X)A_\nu(Y)\right)|0\rangle\,, \end{aligned}\tag{4.1017}$$

$$= -i\,\langle 0|\delta(X^0-Y^0)[\dot{A}_\mu(X),A_\nu(Y)]|0\rangle\,, \tag{4.1018}$$

$$= \delta^4(X-Y)\,. \tag{4.1019}$$

[21] We can construct the propagator in the general case in which the lagrangian density is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{\lambda}{2}(\partial_\alpha A^\alpha)^2\,, \tag{4.1012}$$

i.e. the equations of motion are

$$\partial^2 A_\mu(X) - (1-\lambda)\partial^\mu(\partial_\alpha A^\alpha) = J^\mu(X)\,. \tag{4.1013}$$

We find

$$\langle 0|T\left(A_\mu(X)A_\nu(Y)\right)|0\rangle = \int \frac{d^4P}{(2\pi)^4}e^{-iP_\mu X^\mu}\left[\frac{-i\eta_{\mu\nu}}{P^2+i\eta} - i\frac{1-\lambda}{\lambda}\frac{P_\mu P_\nu}{(P^2+i\eta)^2}\right]\,, \tag{4.1014}$$

that for $\lambda = 1$ gives back the propagator in the so-called Feynman gauge. The physical quantities in the end should be independent of λ.

Capitolo 5

Interactions among fields

In this chapter we will consider interactions among fields. As usual, we will start our study considering the classical lagrangian density and then, afterwards, we will quantize the theory, considering interactions among particles.

5.1 Possible interaction terms

In general, each fundamental constituent is represented by a separate quantum field. If we have free fields, the lagrangian density will be sum of the free lagrangian densities of the various fields. The interaction is done adding an additional term to the lagrangian such that

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{int}\,, \tag{5.1}$$

where

$$\mathcal{L}_0 = \sum_i \mathcal{L}_{0\,i} \tag{5.2}$$

and $\mathcal{L}_{int}$ will contain products of fields with relative coupling constants.

The form of $\mathcal{L}_{int}$ has to fulfill some requirements, that are usually connected to symmetries that we believe the system has to obay. The more general requirements are

1. $\mathcal{L}_{int}$ has to be Poincaré invariant.
2. $\mathcal{L}_{int}$ must be hermitian
3. $\mathcal{L}_{int}$ should be local, meaning with this statement that it should be given in terms of products of fields evaluated in the same space-time point.

On top of these very general requirements, one can add specific requirements for instance driven by phenomenology or by additional theoretical constraints. For instance, as a guiding principle, one can choose "simplicity", requiring that the interaction terms should depend on "as few free parameters as possible", or that we should get a renormalizable theory (and this forces the interacting term to have a peculiar form).

An example to understand how to construct $\mathcal{L}_{int}$ can be represented by a "Yukawa" interaction, that couples a fermionic field with a scalar field in the following way:

$$\mathcal{L}_{int} = g\,\overline{\psi}(X)\psi(X)\phi(X)\,. \tag{5.3}$$

This lagrangian density obays to all the requirements we listed above. It is invariant under Poincaré transformations. It is hermitian, provided that $\phi(X)$ is a scalar real field and g is a real number. It is a local term, since all the fields depend on the same space-time point. The number g is called the "coupling" and it measures the strenght of the interaction.

Another possible interaction is the one between an isospin doublet (neutron and proton) with a meson isotopic vector ϕ_i with three components: ϕ (pseudo-scalar neutral field, corresponding to the π^0 meson) and $\phi^\pm$ (charged fields corresponding to the $\pi^\pm$ mesons):

$$\mathcal{L}_{int} = g \sum_i \overline{\psi}(X)\gamma_5\tau_i\psi(X)\phi_i(X)\,, \tag{5.4}$$

where τ_i are the three Pauli matrices.

Other interesting lagrangians are self interacting terms of scalar fields, as

$$\mathcal{L}_{int} = \lambda\,(\phi(X))^4\,, \tag{5.5}$$

(but we can also think about $g(\phi(X))^3$), with ϕ a real field.

Very important are the so-called "Gauge Theories", as Quantum Electrodynamics for which the interaction lagrangian is

$$\mathcal{L}_{int} = -e\overline{\psi}(X)\gamma^\mu\psi(X)A_\mu(X) \tag{5.6}$$

and we will se below how it is found.

Mass dimension

One of the possible constraints we can impose to $\mathcal{L}_{int}$ is the fact that it has to be a term that does not spoil the renormalizability of the lagrangian density. It can be proved that this feature is linked to the dimensions of the couplings (and operators) present in the lagrangian density. If in our theory we have terms with couplings that have nevative mass dimensions, in a 4-dimensional Minkowski space, renormalizability is lost. Therefore, it is important to know the dimensions of different ingredients that enter our lagrangians.

In our system of units[1] we put $\hbar = 1$. This means that the action is dimensionless and therefore, since

$$S = \int d^4X\,\mathcal{L}\,, \qquad \text{with} \quad [S] = [m]^0\,, \tag{5.8}$$

we have

$$[d^4X] = [m]^{-4}\,, \qquad \text{and therefore} \quad [\mathcal{L}] = [m]^4 \tag{5.9}$$

(the lagrangian density must have dimension 4 in mass). In this way we can find the dimensions of the different fields and coupling constants present in the lagrangian density.

Let us consider an n-dimensional space time. We have, for instance

$$[\overline{\psi}\ \partial\!\!\!/\psi] \equiv [m]^n\,. \tag{5.10}$$

Therefore, since $[\partial^\mu] = [m]$, we have for the fermionic field

$$[\psi] = [m]^{\frac{n-1}{2}}\,. \tag{5.11}$$

In 4 dimensions $[\psi] = [m]^{3/2}$.

The other kinetic term, the one for the electromagnetic field, gives the dimensions of the components of A^μ. In fact

$$[\partial_\mu A_\nu \partial^\mu A^\nu] \equiv [m]^n\,. \tag{5.12}$$

[1]We have that the dimensions of a lenght are equivalent to the dimensions of an inverse mass or an inverse energy

$$[X] = \frac{1}{[E]} = \frac{1}{[m]}\,. \tag{5.7}$$

Therefore, since $[\partial^\mu] = [m]$, we have

$$[A^\mu] = [m]^{\frac{n-2}{2}} \tag{5.13}$$

and this is the case also for the scalar field.

Studying the massive term, we have

$$[m\overline{\psi}\psi] \equiv [m]^n \,. \tag{5.14}$$

Therefore, since $[\psi] = [m]^{\frac{n-1}{2}}$, we have that $[m] = [m]$, as it should! The parameter m has indeed the dimensions of a mass.

For the electric charge, instead, we have

$$[e\overline{\psi}\ \not{A}\psi] \equiv [m]^n \tag{5.15}$$

and since $[\psi] = [m]^{\frac{n-1}{2}}$ and $[A^\mu] = [m]^{\frac{n-2}{2}}$, we find

$$[e] = [m]^{\frac{4-n}{2}} \,. \tag{5.16}$$

The electric charge is dimensionless in 4 dimensions. It follows that QED is a renormalizable theory.

This is also the case for the self-interacting terms $g\phi^3$ and $\lambda\phi^4$. In fact

$$[g] = [m]^{n-\frac{3(n-2)}{2}} = [m]^{3-\frac{n}{2}} \tag{5.17}$$

and in 4 dimensions $[g] = [m]$. Moreover

$$[\lambda] = [m]^{n-2(n-2)} = [m]^{4-n} \tag{5.18}$$

and in 4 dimensions λ is dimensionless.

5.2 Classical interaction of a point-like charged particle with the electromagnetic field. Minimal substitution

Let us consider the electromagnetic interaction of a classical point-like charged particle with the electromagnetic field. Electromagnetic interactions are usually introduced using the so-called "minimal substitution" that prescribes the following change in the momentum of the particle

$$P^\mu \to P^\mu - eA^\mu \tag{5.19}$$

Eq. (5.19) comes from the dynamics of the charged particle subjected to the electromagnetic field and, therefore, to the Lorentz force

$$\mathbf{F} = e\left(\mathbf{E} + \frac{1}{c}\mathbf{v} \wedge \mathbf{H}\right) \,, \tag{5.20}$$

where $\mathbf{v}$ is the velocity of the particle. The relativistic equation of motion

$$\frac{d}{dt}\mathbf{q} = \frac{d}{dt}\frac{m\mathbf{v}}{\sqrt{1-\beta^2}} = e\mathbf{E} + \frac{e}{c}\mathbf{v} \wedge \mathbf{H} \,, \tag{5.21}$$

can be found as Euler-Lagrange equations from the lagrangian

$$L = -mc^2\sqrt{1-\beta^2} - e\phi + \frac{e}{c}\mathbf{v} \cdot \mathbf{A} \,, \tag{5.22}$$

where

$$L_0 = -mc^2\sqrt{1-\beta^2} \tag{5.23}$$

is the free Lagrangian and

$$L_{int} = -e\phi + \frac{e}{c}\mathbf{v} \cdot \mathbf{A} \tag{5.24}$$

is the interacting one. In fact, we have

$$\frac{\partial L}{\partial \mathbf{x}} - \frac{d}{dt}\frac{\partial L}{\partial \mathbf{v}} = 0\,, \tag{5.25}$$

where

$$\frac{\partial L}{\partial \mathbf{x}} = -e\nabla\phi + \frac{e}{c}\nabla(\mathbf{v} \cdot \mathbf{A})\,, \tag{5.26}$$

$$\frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1-\beta^2}} + \frac{e}{c}\mathbf{A}\,. \tag{5.27}$$

Remembering that

$$\frac{e}{c}\mathbf{v} \wedge \mathbf{H} = \frac{e}{c}\mathbf{v} \wedge (\nabla \wedge \mathbf{A}) = \frac{e}{c}\left(\nabla(\mathbf{v} \cdot \mathbf{A}) - (\nabla \cdot \mathbf{v})\mathbf{A}\right) = \nabla(\mathbf{v} \cdot \mathbf{A})\,, \tag{5.28}$$

since $(\nabla \cdot \mathbf{v}) = 0$, we find

$$\frac{d}{dt}\frac{m\mathbf{v}}{\sqrt{1-\beta^2}} = e\left(-\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\right) + \frac{e}{c}\mathbf{v} \wedge (\nabla \wedge \mathbf{A}) = e\mathbf{E} + \frac{e}{c}\mathbf{v} \wedge \mathbf{H}\,. \tag{5.29}$$

The momentum and energy of the interacting particle are

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1-\beta^2}} + \frac{e}{c}\mathbf{A}\,, \tag{5.30}$$

$$H = \mathbf{p} \cdot \mathbf{v} - L = \frac{mv^2}{\sqrt{1-\beta^2}} + \frac{e}{c}\mathbf{v} \cdot \mathbf{A} - \left(-mc^2\sqrt{1-\beta^2} - e\phi + \frac{e}{c}\mathbf{v} \cdot \mathbf{A}\right)\,,$$

$$= \frac{mc^2}{\sqrt{1-\beta^2}} + e\phi\,. \tag{5.31}$$

In the free particle case we have

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1-\beta^2}}\,, \tag{5.32}$$

$$H = \frac{mc^2}{\sqrt{1-\beta^2}}\,, \tag{5.33}$$

in such a way that the on-shell condition holds:

$$\frac{H^2}{c^2} - p^2 = \frac{m^2c^2}{1-\beta^2} - \frac{m^2v^2}{1-\beta^2} = m^2c^2\,. \tag{5.34}$$

In the interacting case we have

$$\left(\frac{H}{c} - \frac{e}{c}\phi\right)^2 - \left(p - \frac{e}{c}\mathbf{A}\right)^2 = m^2c^2\,. \tag{5.35}$$

This relation explains the “minimal substitution”.

5.3 Electromagnetic Interaction of the Dirac field

Let us consider now a fermionic field, $\psi(X)$, in interaction with the electromagnetic field, A^μ. The free lagrangian density of the system is the following

$$\mathcal{L}_0 = \bar{\psi}\,(i\,\partial\!\!\!/ - m)\,\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,. \tag{5.36}$$

In order to introduce the interaction lagrangian, we use the minimal substitution

$$\partial^\mu \to \partial^\mu + ieA^\mu\,. \tag{5.37}$$

and therefore we find:

$$\begin{aligned}\mathcal{L} &= \bar{\psi}\,(i\,\partial\!\!\!/ - e\,A\!\!\!/ - m)\,\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} = \bar{\psi}\,(i\,\partial\!\!\!/ - m)\,\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - e\bar{\psi}\gamma_\mu\psi\,A^\mu\,, && (5.38)\\ &= \mathcal{L}_0 + \mathcal{L}_{int}\,. && (5.39)\end{aligned}$$

The term

$$\mathcal{L}_{int} = -e\bar{\psi}A\!\!\!/\psi = -e\bar{\psi}\gamma_\mu\psi\,A^\mu \tag{5.40}$$

constitutes the interaction lagrangian.

The lagrangian density (5.39) is symmetric under Poincaré transformations (including parity and time reversal). Moreover, as in the case of the free lagrangian, $\mathcal{L}$ is invariant under global phase transformations

$$\psi(X) \to \psi'(X) = e^{-i\theta}\psi(X)\,, \qquad \bar{\psi}(X) \to \bar{\psi}'(X) = e^{-i\theta}\bar{\psi}(X)\,. \tag{5.41}$$

The conserved current is again $j^\mu = e\bar{\psi}\gamma^\mu\psi$, keeping the form of the free case. Therefore, we see that the electromagnetic interaction is the coupling of the conserved fermionic current with the field A^μ.

Although the free lagrangian (5.36) is invariant under gauge tranformations

$$A^\mu \to A'^\mu = A^\mu + \partial^\mu\Lambda\,, \tag{5.42}$$

the interacting lagrangian (5.39) is not. In fact, we have:

$$\mathcal{L} \to \mathcal{L}' = \mathcal{L}_0 - e\bar{\psi}\gamma_\mu\psi\,A'^\mu = \mathcal{L}_0 - e\bar{\psi}\gamma_\mu\psi\,A^\mu - e\bar{\psi}\gamma_\mu\psi\partial^\mu\Lambda = \mathcal{L} - e\bar{\psi}\gamma_\mu\psi\,\partial^\mu\Lambda\,. \tag{5.43}$$

In the transformed lagrangian, the following additional term appears:

$$-e\bar{\psi}\gamma_\mu\psi\,\partial^\mu\Lambda\,. \tag{5.44}$$

We can actually cancel this additional term, restoring a symmetry of the lagrangian, if we consider, together to the gauge transformation of the electromagnetic field, a "local" phase transformation of the fermionic field:

$$\psi(X) \to e^{-ie\Lambda(X)}\psi(X)\,. \tag{5.45}$$

The action of (5.45) on the lagrangian density is:

$$\begin{aligned}\mathcal{L} \;\to\; \mathcal{L}' &= \bar{\psi}e^{ie\Lambda(X)}(i\,\partial\!\!\!/ - m)(e^{-ie\Lambda(X)}\psi) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - e\bar{\psi}\gamma_\mu\psi\,A^\mu\,, && (5.46)\\ &= \bar{\psi}(i\,\partial\!\!\!/ - m)\psi + \bar{\psi}e^{ie\Lambda(X)}i(-ie\,\partial\!\!\!/\Lambda(X))e^{-ie\Lambda(X)}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - e\bar{\psi}\gamma_\mu\psi\,A^\mu\,, && (5.47)\\ &= \mathcal{L} + e\bar{\psi}\gamma_\mu\psi\partial^\mu\Lambda\,. && (5.48)\end{aligned}$$

Note that the interaction term, since it does not contain derivatives of the fields, remains unchanged under the local phase transformation.

Finally, under the following tramsformation

$$\begin{aligned}\partial^\mu &\to \partial'^\mu = \partial^\mu + ieA^\mu && (5.49)\\ \psi &\to \psi' = e^{-ie\Lambda(X)}\psi && (5.50)\\ \overline{\psi} &\to \overline{\psi'} = \overline{\psi}e^{ie\Lambda(X)}\,, && (5.51)\end{aligned}$$

that will be called *gauge transformation*, the lagrangian density (5.39) is invariant.

We can introduce a formal derivative, called *covariant derivative*, defined as follows:

$$D^\mu = \partial^\mu + ieA^\mu\,, \tag{5.52}$$

that under gauge transformations transforms as the field

$$(D^\mu\psi)' = e^{-ie\Lambda(X)}(D^\mu\psi)\,, \tag{5.53}$$

with wich we can write the QED lagrangian density in a more compact way:

$$\mathcal{L} = \bar{\psi}\left(i\,\not{D} - m\right)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,, \tag{5.54}$$

form which is manifestly invariant under gauge transformations.

Eq. (5.53) can be easily proven. In fact, under gauge transformations we have

$$\begin{aligned}D^\mu\psi \to (D^\mu\psi)' &= (\partial^\mu + ieA^\mu + ie\partial^\mu\Lambda)(e^{-ie\Lambda}\psi)\,, && (5.55)\\ &= (\partial^\mu e^{-ie\Lambda})\psi + e^{-ie\Lambda}\partial^\mu\psi + ieA^\mu e^{-ie\Lambda}\psi + ie\partial^\mu\Lambda e^{-ie\Lambda}\psi\,, && (5.56)\\ &= -ie\partial^\mu\Lambda e^{-ie\Lambda}\psi + e^{-ie\Lambda}\partial^\mu\psi + ieA^\mu e^{-ie\Lambda}\psi + ie\partial^\mu\Lambda e^{-ie\Lambda}\psi\,, && (5.57)\\ &= e^{-ie\Lambda}(\partial^\mu + ieA^\mu)\psi\,, && (5.58)\\ &= e^{-ie\Lambda}D^\mu\psi\,. && (5.59)\end{aligned}$$

The Gauge Principle

We understood the theoretical structure of the electromagnetic lagrangian starting from phenomenologically proven elements (Maxwell's theory). We found gauge invariance. We may try, now, to invert the view point and use gauge invariance as an instrument to find the "correct" interacting lagrangian starting from the free lagrangians (the kinetic terms that have to be in any case there). In order to do that, we should go through the following steps:

1. We start with the free matter lagrangian $\mathcal{L}_0 = \bar{\psi}\left(i\,\not{\partial} - m\right)\psi$ and we look for a "generalization" that is invariant under local phase transformations $\psi' = e^{-ie\Lambda(X)}\psi$.

2. In order to do that, we introduce a "gauge field" A^μ, that has to transform as $A'^\mu = A^\mu + \partial^\mu\Lambda$ and replace the derivative with a covariant derivative $D^\mu = \partial^\mu + ieA^\mu$.

3. The kinetic term of the gauge field has to be part of the lagrangian density.

The local phase transformations, $U = e^{-ie\Lambda(X)}$, form a Lie group with one parameter $U(1)$, which is abelian.

We can thing about generalizations of the gauge group to non abelian groups (that therefore give more complicated lagrangians).

5.3.1 Non-Abelian Gauge theories. Quantum Chromodynamics (QCD)

Let us suppose to consider a non abelian gauge group, as $SU(N)$. This is a Lie group, that depends on (N^2-1) parameters (and therefore has (N^2-1) generators). The general transformation

$$U = e^{-ig\theta^a t^a}\,, \qquad a = 1, ..., N^2-1\,, \tag{5.60}$$

will act on the fields ψ

$$\psi'(X) = U\psi(X) = e^{-ig\theta^a t^a}\psi(X)\,. \tag{5.61}$$

The field will have to be in a certain representation of the $SU(N)$ group. Accordingly, the generators t^a will be linear hermitian operators on that representation. They obay the usual Lie algebra

$$[t^a, t^b] = if^{abc}t^c\,, \tag{5.62}$$

where f^{abc} are the structure constants of $SU(N)$, totally anti-symmetric in a, b, c.

The spinor fields will be defined in the fundamental representation of $SU(N)$ and therefore are "vectors" with N components in this space: $\psi_i(X)$, with $i = 1, ..., N$. The generators t^a are therefore $N \times N$ matrices, in this representation, acting on the N-tuple $\psi_i(X)$. In components: t^a_{ij}, with $i, j = 1...N$. The structure constants belong to the adjoint representation. We can define

$$T^a_{bc} = -if^{abc}\,, \tag{5.63}$$

that are (N^2-1) matrices of dimensionality $(N^2-1)\times(N^2-1)$ that obey the Lie algebra:

$$[T^a, T^b] = if^{abc}T^c\,, \tag{5.64}$$

They act on the gauge field, defined in the adjoint representation, A^a_μ, with (N^2-1) components (each one carrying a Lorentz index μ).

The structure of $SU(N)$ is linked to the fact that in order to explain certain phenomenological properties it was necessary to introduce a new quantum number, the "color". Experimentally, we see that phenomenology can be explained using three colors, i.e. $N = 3$. Therefore, the gauge group of QCD is $SU(3)$.

Using the gauge principle we now construct a lagrangian density which is invariant under $SU(3)$ gauge transformations.

We start with the free matter lagrangian density

$$\mathcal{L} = \bar{\psi}^i_q(i\,\partial\!\!\!/ - m_q)\psi^i_q\,, \tag{5.65}$$

where the subscript q labels the flavor (different kind of quarks of the three families) and the superscript i labels the colors. The lagrangian in Eq. (5.65) is invariant under global $SU(3)$ transformations

$$U = e^{-ig\theta^a t^a} \tag{5.66}$$

with $\theta^a \in \mathbb{R}$.

We impose invariance under local transformations, considering now $\theta^a = \theta^a(X)$. We have, therefore, to introduce a gauge field A^a_μ and a covariant derivative

$$D_\mu = \partial_\mu - igt^a A^a_\mu\,, \tag{5.67}$$

or, in components in the color space, $(D_\mu)_{ij} = \delta_{ij}\partial_\mu - igt^a_{ij}A^a_\mu$. The constant g is the coupling constant of the interaction and it is unique for every color component.

We now impose that the gauge field, A^a_μ, transform under gauge transformations in such a way to let

$$\mathcal{L} = \bar{\psi}^i_q(i\,D\!\!\!\!/_{ij} - m_q\delta_{ij})\psi^j_q \tag{5.68}$$

invariant. This means that $D_\mu\psi$ has to transform as ψ (Eq. (5.61)) itself:

$$(D_\mu\psi)' \equiv U\,(D_\mu\psi) = U\,(\partial_\mu\psi - igt^a A^a_\mu\psi)\,. \tag{5.69}$$

Since we have

$$(D_\mu\psi)' = (\partial_\mu - igt^a A'^a_\mu)\psi' = (\partial_\mu U)\psi + U(\partial_\mu\psi) - igt^a A'^a_\mu\, U\psi\,, \tag{5.70}$$

we have to impose

$$(\partial_\mu U) - igt^a A'^a_\mu\, U = -igU\, t^a A^a_\mu\,. \tag{5.71}$$

This gives the following transformation rule for $t^a A'^a_\mu$

$$t^a A'^a_\mu = U\left(t^a A^a_\mu - \frac{i}{g}U^{-1}(\partial_\mu U)\right)U^{-1}\,. \tag{5.72}$$

Now we need the kinetic term for the gauge field A^a_μ, which is invariant under Eq. (5.72). Taking inspiration from QED, one could choose

$$\mathcal{L} = -\frac{1}{4}F^a_{\mu\nu}F^{a\,\mu\nu} = -\frac{1}{4}(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu)(\partial^\mu A^{a\,\nu} - \partial^\nu A^{a\,\mu})\,. \tag{5.73}$$

However, this term is not invariant under (5.72). In order to demonstrate this, we use infinitesimal transformations

$$U = e^{-ig\theta^a t^a} \simeq 1 - ig\theta^a t^a\,. \tag{5.74}$$

Eq. (5.72) becomes

$$t^a A'^a_\mu \simeq ... = t^a\left(A^a_\mu + gf^{abc}\theta^b A^c_\mu - \partial_\mu\theta^a\right)\,. \tag{5.75}$$

Therefore

$$A'^a_\mu - A^a_\mu = \delta A^a_\mu = g\, f^{abc}\theta^b A^c_\mu - \partial_\mu\theta^a\,. \tag{5.76}$$

With this transformation we find

$$F'^a_{\mu\nu} = ... = (\delta^{ac} + g\, f^{abc}\theta^b)F^c_{\mu\nu} + g\, f^{abc}\left[(\partial_\mu\theta^b)A^c_\nu - (\partial_\nu\theta^b)A^c_\mu\right] \tag{5.77}$$

and we see that, defined as in Eq. (5.73), the term $F^a_{\mu\nu}F^{a\,\mu\nu}$ is not invariant.

Let us start from another property of the QED covariant derivative. In QED, we have

$$[D_\mu, D_\nu] = (\partial_\mu + ieA_\mu)(\partial_\nu + ieA_\nu) - (\partial_\nu + ieA_\nu)(\partial_\mu + ieA_\mu) = ieF_{\mu\nu}\,. \tag{5.78}$$

Let us impose that in QCD

$$[D_\mu, D_\nu] \equiv -ig\, t^a F^a_{\mu\nu}\,. \tag{5.79}$$

We have

$$\begin{aligned}
[D_\mu, D_\nu] &= (\partial_\mu - igt^a A^a_\mu)(\partial_\nu - igt^b A^b_\nu) - (\partial_\nu - igt^b A^b_\nu)(\partial_\mu - igt^a A^a_\mu) = ... && (5.80)\\
&= -igt^a\left[\partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g\, f^{abc}A^b_\mu A^c_\nu\right]\,. && (5.81)
\end{aligned}$$

We can then define

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g\, f^{abc}A^b_\mu A^c_\nu\,. \tag{5.82}$$

With this definition we can check that indeed $F'^a_{\mu\nu}F'^{a\,\mu\nu} = F^a_{\mu\nu}F^{a\,\mu\nu}$.

Finally, we find the following lagrangian density

$$\mathcal{L} = \bar{\psi}^i_q(i\,\not{D}_{ij} - m_q\delta_{ij})\psi^j_q - \frac{1}{4}F^a_{\mu\nu}F^{a\,\mu\nu}\,, \tag{5.83}$$

with $F^a_{\mu\nu}$ defined in Eq. (5.82) and D_μ defined in Eq. (5.67). The lagrangian (5.83) is invariant under Poincaré transformations (other than parity, time reversal and charge conjugation) and under gauge trasformations:

$$\psi'(X) = e^{-ig\theta^a(X)t^a}\psi(X)\,, \tag{5.84}$$
$$\bar{\psi}'(X) = e^{ig\theta^a(X)t^a}\bar{\psi}(X)\,, \tag{5.85}$$
$$t^a A'^a_\mu = U\left(t^a A^a_\mu - \frac{i}{g}U^{-1}(\partial_\mu U)\right)U^{-1}\,. \tag{5.86}$$

The lagrangian has a unique coupling, g. In 4 dimensions, we can see that g is dimensionless. QCD is a renormalizable theory. In the interacting QCD lagrangian we have vertices, like in QED, that couple $\bar{\psi}\psi A^a_\mu$, but, unlike QED, now we have also gluon (gauge field) self-interacting vertices (three and four gluons coupled). These difference comes from the fact that the photon is not charged, while the gluon carries color.

5.3.2 Quantization of the electromagnetic Lagrangian

The canonical quantization of the lagrangian (5.39) follows the same procedure we outlined for the free lagrangians. Commutation or anticommutation rules have to be imposed between the field and the conjugated momentum (the only difference lies in the fact that now we have interacting fields and therefore we cannot in general express them in terms of annihilation/creation operators). If we generally refer to the field present in the lagrangian as $\phi^i(X)$, the conjugated momentum is defined as usual:

$$\Pi^i = \frac{\partial\mathcal{L}}{\partial\dot{\phi}^i} = \frac{\partial\mathcal{L}_0}{\partial\dot{\phi}^i} + \frac{\partial\mathcal{L}_{int}}{\partial\dot{\phi}^i}\,. \tag{5.87}$$

In the case of the electromagnetic interactions of the Dirac field, Eq. (5.39), the interaction lagrangian does not contain derivatives of the fields. Therefore

$$\frac{\partial\mathcal{L}_{int}}{\partial\dot{\phi}^i} = 0\,. \tag{5.88}$$

This property implies that the conjugated momentum has the same expression as in the free field case

$$\Pi^i = \frac{\partial\mathcal{L}}{\partial\dot{\phi}^i} = \frac{\partial\mathcal{L}_0}{\partial\dot{\phi}^i}\,. \tag{5.89}$$

We say that we are in presence of a "non-derivative interaction". For the Dirac field, in fact, we have

$$\Pi_\psi = \frac{\partial\mathcal{L}}{\partial\dot{\psi}} = i\psi^\dagger\,. \tag{5.90}$$

The same happens for A^μ. For the canonical quantization of the gauge field we have to add the gauge fixing lagrangian and then

$$\Pi^\mu = \frac{\partial\mathcal{L}}{\partial\dot{A}^\mu} = \frac{\partial\mathcal{L}_0}{\partial\dot{A}^\mu} = F^{\mu 0} - \eta^{\mu 0}(\partial_\alpha A^\alpha)\,. \tag{5.91}$$

We can quantize imposing the usual equal time rules

$$[\psi(\mathbf{x},t),\psi^\dagger(\mathbf{y},t)]_+ = \delta^3(\mathbf{x}-\mathbf{y})\,, \qquad [A_\mu(\mathbf{x},t),\dot{A}_\nu(\mathbf{y},t)] = -i\eta_{\mu\nu}\delta^3(\mathbf{x}-\mathbf{y}) \tag{5.92}$$

and the other commutators and anticommutators that have to vanish.

Non-derivative interactions have an additional feature. Since the conjugated momenta are the same as in the free case, the interaction hamiltonian is minus the interaction lagrangian. We have

$$\mathcal{H} = \Pi\dot{\phi} - \mathcal{L} = \Pi\dot{\phi} - \mathcal{L}_0 - \mathcal{L}_{int}\,, \tag{5.93}$$
$$= \mathcal{H}_0 - \mathcal{L}_{int} \tag{5.94}$$

and therefore:

$$\mathcal{H}_{int} = -\mathcal{L}_{int}\,. \tag{5.95}$$

Scalar QED

The non-derivative interaction simplifies the structure of the theory, but of course not all the interactions are non derivative. Just to remain the the case of electromagnetic interactions, a different behaviour is plaied by the scalar field. Let us consider a charged Klein-Gordon field in interaction with A^μ. We have:

$$\begin{aligned}
\mathcal{L} &= (D_\mu\phi)^\dagger D^\mu\phi - m^2\phi^\dagger\phi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,, && (5.96)\\
&= [(\partial_\mu + ieA_\mu)\phi]^\dagger(\partial^\mu + ieA^\mu)\phi - m^2\phi^\dagger\phi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,, && (5.97)\\
&= (\partial_\mu - ieA_\mu)\phi^\dagger\,(\partial^\mu + ieA^\mu)\phi - m^2\phi^\dagger\phi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,, && (5.98)\\
&= \mathcal{L}_0 - ie[\phi^\dagger(\partial^\mu\phi) - (\partial^\mu\phi^\dagger)\phi]\,A_\mu + e^2A^2\phi^\dagger\phi\,. && (5.99)
\end{aligned}$$

Therefore

$$\mathcal{L}_{int} = -ie[\phi^\dagger(\partial^\mu\phi) - (\partial^\mu\phi^\dagger)\phi]\,A_\mu + e^2A^2\phi^\dagger\phi \tag{5.100}$$

and it contains derivatives of the scalar field! Also in this case $\mathcal{L}$ is manifestly gauge invariant (by construction). However, while for the Dirac field the gauge-invariant current is the same as the one in the free field case, $J_\mu^{Dirac} = -e\bar{\psi}\gamma_\mu\psi$, and the interaction term comes from the interaction of this current with A^μ, for the KG field this is not the case. In fact, in the free field case we had

$$j^\mu \;=\; ie\,[\phi^\dagger\partial^\mu\phi - (\partial^\mu\phi^\dagger)\phi]\,. \tag{5.101}$$

In the interacting case, the conserved current, in the sense of Nöther's theorem, has an additional piece:

$$\begin{aligned}
J^\mu &= \frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\delta\phi + \frac{\partial\mathcal{L}}{\partial\phi^\dagger_{,\mu}}\delta\phi^\dagger\,, && (5.102)\\
&= ie\,[\phi^\dagger\partial^\mu\phi - (\partial^\mu\phi^\dagger)\phi] - 2e^2A^\mu\phi^\dagger\phi\,, && (5.103)\\
&= j^\mu - 2e^2A^\mu\phi^\dagger\phi\,, && (5.104)
\end{aligned}$$

that comes from gauge invariance. In fact we have:

$$J^\mu = ie[\phi^\dagger(D^\mu\phi) - [(D_\mu\phi)^\dagger]\phi]\,, \tag{5.105}$$

therefore the free case current in which we replace the derivatives with the covariant derivatives. The additional term, $-2e^2A^\mu\phi^\dagger\phi$, that preserves gauge invariance, gives rise to interection vertices with four fields, which are not present in spinor-QED.

Concerning the quantization, conjugated momenta to ϕ and $\phi^\dagger$ are not the same as in the free field case. In fact we have:

$$\begin{aligned}
\Pi_\phi &= \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \dot{\phi}^\dagger - ie\phi^\dagger A^o\,, && (5.106)\\
\Pi_{\phi^\dagger} &= \frac{\partial\mathcal{L}}{\partial\dot{\phi}^\dagger} = \dot{\phi} + ie\phi A^o\,. && (5.107)
\end{aligned}$$

Therefore, they can change the commutation rules. In this specific case, the presence of A^0 does not change the rules just because, considering independent the degrees of freedom relative to different fields, we have to impose:

$$[\phi(\mathbf{x},t), \phi^\dagger(\mathbf{x},t)] = [\phi(\mathbf{x},t), A^0(\mathbf{x},t)] \;=\; 0\,. \tag{5.108}$$

Therefore, in the end

$$\begin{aligned} i\delta^3(\mathbf{x}-\mathbf{y}) &= [\phi(\mathbf{x},t), \Pi_\phi(\mathbf{y},t)]\,, &(5.109)\\ &= [\phi(\mathbf{x},t), \dot{\phi}^\dagger(\mathbf{y},t) - ie\phi^\dagger(\mathbf{y},t)A^0(\mathbf{y},t)] = \\ &= [\phi(\mathbf{x},t), \dot{\phi}^\dagger(\mathbf{y},t)] + [\phi(\mathbf{x},t), -ie\phi^\dagger(\mathbf{y},t)A^0(\mathbf{y},t)]\,, &(5.110)\\ &= [\phi(\mathbf{x},t), \dot{\phi}^\dagger(\mathbf{y},t)] - ie[\phi(\mathbf{x},t), \phi^\dagger(\mathbf{y},t)]A^0(\mathbf{y},t)\\ &\quad -ie\phi^\dagger(\mathbf{y},t)[\phi(\mathbf{x},t), A^0(\mathbf{y},t)]\,, &(5.111)\\ &= [\phi(\mathbf{x},t), \dot{\phi}^\dagger(\mathbf{y},t)]\,, &(5.112) \end{aligned}$$

and the same for $[\phi^\dagger(\mathbf{x},t), \Pi_{\phi^\dagger}(\mathbf{y},t)]$. Also in this case we can limit ourselves to impose the quantization rules between fields as in the free case. We have, however, differences in the hamiltonian density.

5.3.3 Quantization of the electromagnetic Lagrangian and gauge invariance

The quantum QED lagrangian density is

$$\mathcal{L} = \bar{\psi}\,(i\,\not{D} - m)\,\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{\lambda}{2}(\partial_\alpha A^\alpha)^2\,, \tag{5.113}$$

where we added the gauge-fixing lagrangian for the quantization of the gauge field (the so-called "Feynman gauge" is recovered with the choice $\lambda = 1$). Although we started with a classical lagrangian invariant under gauge transformations, the lagrangian defined in Eq. (5.113) is not anymore gauge invariant. The gauge fixing term spoils gauge invariance. Let us consider infinitesimal gauge transformations, in which $\Lambda(x) = \epsilon\,\omega(x)$ with ϵ a small real parameter and $\omega(X)$ a scalar real field. At first order in ϵ we have

$$\begin{aligned} \psi(X) &\to \psi'(X) \simeq (1 - ie\epsilon\omega(X))\psi(X)\,, &(5.114)\\ A_\mu(X) &\to A'_\mu(X) = A_\mu(X) + \epsilon\,\partial_\mu\omega(X)\,. &(5.115) \end{aligned}$$

The gauge fixing lagrangian, under this infinitesimal gauge transformations behaves as follows:

$$\begin{aligned} \mathcal{L}_{GF} &\to \mathcal{L}'_{GF} = -\frac{\lambda}{2}(\partial_\alpha A^\alpha + \epsilon\partial^2\omega(X))^2\,, &(5.116)\\ &\simeq \mathcal{L}_{GF} - \lambda\epsilon(\partial_\alpha A^\alpha)\partial^2\omega(X)\,, &(5.117) \end{aligned}$$

where we kept only first-order terms in ϵ.

In order to "cure" the additional term, let us consider the field $\omega(X)$ as a dynamical field, adding to the lagrangian density a kinetic term

$$\mathcal{L}_\omega = -\frac{1}{2}\partial_\mu\omega(X)\partial^\mu\omega(X)\,, \tag{5.118}$$

and considering the following transformation for the field $\omega(X)$

$$\omega(X) \to \omega'(X) \simeq \omega(X) + \epsilon\lambda(\partial_\alpha A^\alpha)\,. \tag{5.119}$$

In so doing, le lagrangian density $\mathcal{L}_\omega$ changes as follows

$$\begin{aligned} \mathcal{L}_\omega \to \mathcal{L}'_\omega &= -\frac{1}{2}\partial_\mu\omega'(X)\partial^\mu\omega'(X)\,, &(5.120)\\ &\simeq -\frac{1}{2}\partial_\mu\omega(X)\partial^\mu\omega(X) - \epsilon\lambda\partial^\mu\omega(X)\partial_\mu(\partial_\alpha A^\alpha)\,, &(5.121)\\ &\simeq -\frac{1}{2}\partial_\mu\omega(X)\partial^\mu\omega(X) - \epsilon\lambda\left\{\partial_\mu\left[\partial^\mu\omega(X)(\partial_\alpha A^\alpha)\right] - \partial^2\omega(X)(\partial_\alpha A^\alpha)\right\}\,, &(5.122) \end{aligned}$$

$$\simeq -\frac{1}{2}\partial_\mu\omega(X)\partial^\mu\omega(X) + \epsilon\lambda\partial^2\omega(X)(\partial_\alpha A^\alpha)\,, \tag{5.123}$$

where we droped a total divergence.

In total, the lagrangian density

$$\mathcal{L} = \bar{\psi}\,(i\,\not{D} - m)\,\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{\lambda}{2}(\partial_\alpha A^\alpha)^2 - \frac{1}{2}\partial_\mu\omega(X)\partial^\mu\omega(X) \tag{5.124}$$

is invariant under the following generalized gauge transformations

$$\psi(X) \quad\rightarrow\quad \psi'(X) \simeq (1 - ie\epsilon\omega(X))\psi(X)\,, \tag{5.125}$$

$$A_\mu(X) \quad\rightarrow\quad A'_\mu(X) = A_\mu(X) + \epsilon\,\partial_\mu\omega(X)\,, \tag{5.126}$$

$$\omega(X) \quad\rightarrow\quad \omega'(X) \simeq \omega(X) + \epsilon\lambda(\partial_\alpha A^\alpha)\,, \tag{5.127}$$

that are called BRST transformations (from Becchi-Rouet-Stora-Tyutin).

It is important to note that in the case of QED the field $\omega(X)$ is a free field. It does not interact with $\psi(X)$ and $A_\mu(X)$. Therefore, it does not change the dynamics of the system under study; it is totally decoupled. This is not, for instance, the case in QCD, where the BRST fields are called "ghosts" and couple to the gauge fields.

The BRST transformations are the extension of gauge transformations to the quantized lagrangian.

5.4 The Scattering (S) Matrix

The typical measurement done in Particle Physics regards a scettering process: a beam of particles, prepared in a certain state of momentum and polarization, collide on a target (fixed target experiments) or against another beam (for instance in a collider machine). After the collision, the final state is studied and the cross section (that is connected essentially to the probability of a certain transition) is measured. The cross section can also be predicted in Quantum Field Theory, using perturbation theory (as we will see). It depends upon the interaction between fields. The comparison between the theoretical prediction and the result of the measurement gives the possibility to confirm of nullify the theoretical description of the scattering (and, therefore, of the nature of the interaction itself).

The theoretical description of the scattering process goes as follows. In the initial state the particles that will collide are so far apart from each other that they can be considered as non interacting. Actually, a detailed description has to take into account the fact that also in absence of the colliding particle, the free particle is never "actually free", since it interacts with its own field. We can think about the particle as surrounded by a cloud of virtual particles that are created and reabsorbed continuously. This phenomenon will be described through the renormalization of the external field, that we will not consider in our "tree-level" description. After the collision has taken place, the scattered particles are again so far apart that they can be considered as free particles. The scattering take place effectively in a very limited space, for a very limited time, such that we can consider the initial time back in the past ($t = -\infty$), while the final time forward in the future ($t = +\infty$).

The time evolution of the state in the scattering is governed by an operator, S, that transforms the initial state $|\phi_i\rangle = |\phi(t = -\infty)\rangle$ to a possible final state $|\phi(t = +\infty)\rangle$:

$$|\phi(t = +\infty)\rangle = S\,|\phi(t = -\infty)\rangle\,. \tag{5.128}$$

We will be interested to calculate the probability amplitude that the initial state, after the scattering gives rise to a particular final state $|\phi_f\rangle$, therefore the scalar product of the transformed state $|\phi(t = +\infty)\rangle$ with the particular final state $|\phi_f\rangle$:

$$\langle\phi_f|\phi(t = +\infty)\rangle = \langle\phi_f|S\,|\phi(t = -\infty)\rangle = \langle\phi_f|S\,|\phi_i\rangle = S_{fi}\,. \tag{5.129}$$

The modulous squared of this matrix element will give the probability and will be instrumental for the calculation of the cross section (which is an observable). The S matrix describes the interaction and, therefore, depends on the Hamiltonian of the system.

5.4.1 Schrödinger, Heisenberg pictures

We consider a system with an Hamiltonian of the following form

$$H = H_0 + H_{int} \,, \tag{5.130}$$

where H_0 is the free Hamiltonian (containing the kinetic and, possibly mass, terms of the different fields, but not interaction terms) and H_{int} is the interaction Hamiltonian (and of course they do not commute). For a closed system, H does not depend on time (it is a constant of motion).

We also consider the situation in which H_{int} can be considered a perturbation of the free Hamiltonian H_0 (see later on).

In the so-called Schrödinger picture, the system is described by a state vector, $|\psi_S(t)\rangle$, that depends on time (while the operators do not) and whose time evolution is given by the following relation

$$|\psi_S(t)\rangle = e^{-iH(t-t_0)}|\psi_S(t_0)\rangle \,. \tag{5.131}$$

Another way to say the same thing is that $|\psi(t)\rangle$ has to obey the following differential equation in time

$$i\frac{\partial}{\partial t}|\psi_S(t)\rangle = H\,|\psi_S(t)\rangle \,, \tag{5.132}$$

with certain initial conditions.

An operator O_S, which correspond to a certain observable, in this picture does not depend on time and the dependence on time of the expectation value of O_S on physical states

$$\langle O\rangle(t) = \langle\psi'_S(t)|O_S|\psi_S(t)\rangle \tag{5.133}$$

goes through the state vector.

Using Eq. (5.131), we can also write

$$\langle O\rangle(t) = \langle\psi'_S(t_0)|e^{iH(t-t_0)}O_S e^{-iH(t-t_0)}|\psi_S(t_0)\rangle = \langle\psi'(t_0)|O_H(t)|\psi(t_0)\rangle \,, \tag{5.134}$$

and express the expectation value $\langle O\rangle(t)$ using state vectors that are now frozen at the time t_0 (i.e. they do not depend anymore on time) and the operator

$$O_H(t) = e^{iH(t-t_0)}O_S e^{-iH(t-t_0)} \,, \tag{5.135}$$

that depends on time through the evolution operator $e^{-iH(t-t_0)}$ and is called the operator O in the Heisenberg picture. We have

$$|\psi_H\rangle = e^{iH(t-t_0)}|\psi_S(t)\rangle = |\psi_S(t_0)\rangle \,, \tag{5.136}$$

$$O_H(t) = e^{iH(t-t_0)}O_S e^{-iH(t-t_0)} \tag{5.137}$$

and

$$\langle O\rangle(t) = \langle\psi'_S(t)|O_S|\psi_S(t)\rangle = \langle\psi'_H|O_H(t)|\psi_H\rangle \,. \tag{5.138}$$

In the Heisenberg picture the operators depend on time and therefore they obey the Hamilton's equation

$$i\frac{d}{dt}O_H(t) = [O_H(t), H] \,. \tag{5.139}$$

In the case in which the interaction is absent ("switched off"), $H_{int} = 0$, in Schródinger picture the state evolves with the free Hamiltonian (free state) and, equivalently, in the Heisenberg picture the state is time-independent while the operators evolve with the free Hamiltonian (they are free operators)

$$i\frac{d}{dt}O_H(t) = [O_H(t), H_0] \,. \tag{5.140}$$

5.4.2 Interaction picture

Both Schrödinger and Heisenberg pictures are not ideal in the case of interactions. We introduce another picture, which is called Interaction Picture, such that the time evolution is spread out both on the states and on the operators, but such that the states evolve with the interaction Hamiltonian (in Interaction Picture) and the operators evolve with the free Hamiltonian. This means that the states are constant if the interaction is switched off, and the operators are free operators and cn be written in terms of creation/annihilation operators. This is crucial for the computation of transition amplitudes.

If $|\psi_S(t)\rangle$ is the state in Schrödinger representation, we define

$$|\psi_I(t)\rangle = e^{iH_0 t}|\psi_S(t)\rangle \,. \tag{5.141}$$

Analogously, an operator in this representation can be written, with respect to the time-independent operator in Schrödinger representation, as

$$O_I(t) = e^{iH_0 t} O_S e^{-iH_0 t} \,. \tag{5.142}$$

Note that the operator $e^{iH_0 t}$ "anti-evolves" (so to say) the state, but only using the free Hamiltonian (not the full Hamiltonian). The consequence of that is that if we compute the time evolution of $|\psi_I(t)\rangle$ we find

$$\begin{aligned}
i\frac{\partial}{\partial t}|\psi_I(t)\rangle &= i\frac{\partial}{\partial t}\left(e^{iH_0 t}|\psi_S(t)\rangle\right) = -H_0 e^{iH_0 t}|\psi_S(t)\rangle + e^{iH_0 t} i\frac{\partial}{\partial t}|\psi_S(t)\rangle \,, && (5.143)\\
&= -H_0 e^{iH_0 t}|\psi_S(t)\rangle + e^{iH_0 t}\left(H_0 + H_{int}\right)|\psi_S(t)\rangle \,, && (5.144)\\
&= -H_0 e^{iH_0 t}|\psi_S(t)\rangle + H_0 e^{iH_0 t}|\psi_S(t)\rangle + e^{iH_0 t} H_{int}|\psi_S(t)\rangle \,, && (5.145)\\
&= e^{iH_0 t} H_{int} e^{-iH_0 t}|\psi_I(t)\rangle \,. && (5.146)
\end{aligned}$$

The operator

$$e^{iH_0 t} H_{int} e^{-iH_0 t} = H_{int,I} \tag{5.147}$$

is the interaction Hamiltonian in Interaction Representation:

$$i\frac{\partial}{\partial t}|\psi_I(t)\rangle = H_{int,I}|\psi_I(t)\rangle \,. \tag{5.148}$$

On the other hand, the time evolution of the operator $O_I(t)$ is given by the free Hamilton equation

$$i\frac{d}{dt}O_I(t) = [O_I(t), H_0] \tag{5.149}$$

and we will be able to write $O_I(t)$ in terms of free fields and, therefore, in terms of creation/annihilation operators.

We can write, then, the expectation value of the S operator between the initial and final states in Interaction Picture (and it will be equal to the same expectation value in another picture):

$$S_{fi} = \langle\psi_{If}(t=+\infty)|S_I|\psi_{Ii}(t=-\infty)\rangle \,, \tag{5.150}$$

In the limit at $t \to -\infty$ and at $t \to \infty$ after the scattering, the interaction is supposed to vanish and, therefore, the states $|\psi_{Ii}(t=+\infty)\rangle$ and $|\psi_{Ii}(t=-\infty)\rangle$ are free states, independent of t. We consider eigenstates of the energy and momentum (plane waves). The S operator in Interaction Picture is constructed, then, with free field operators and can be expressed in terms of creation/annihilation operators.

5.4.3 Dyson formula

In order to find an expression for S, let us consider Eq. (5.148). We can write it in an integral form as follows

$$|\psi_I(t)\rangle = |\psi_I(t=-\infty)\rangle - i\int_{-\infty}^{t} H_{int,I}(t_1)|\psi_I(t_1)\rangle\, dt_1\,. \tag{5.151}$$

For the sake of simplicity in the notation, let us drop the subscript "I" everywhere (from now on we understand that we are in Interaction Picture). Eq. (5.151) is totally equivalent to Eq. (5.148), but it can be used for a recursive solution that is "justified" by the use of perturbation theory. In fact, using a second time Eq. (5.151), we have

$$\begin{aligned} |\psi(t)\rangle &= |\psi_{-\infty}\rangle - i\int_{-\infty}^{t} dt_1\, H_{int}(t_1)\left(|\psi_{-\infty}\rangle - i\int_{-\infty}^{t_1} H_{int}(t_2)|\psi_I(t_2)\rangle\, dt_2\right), && (5.152)\\ &= |\psi_{-\infty}\rangle - i\int_{-\infty}^{t} dt_1\, H_{int}(t_1)|\psi_{-\infty}\rangle \\ &\quad +(-i)^2\int_{-\infty}^{t}\int_{-\infty}^{t_1} dt_1\, dt_2\, H_{int}(t_1)H_{int}(t_2)|\psi_I(t_2)\rangle\,, && (5.153)\end{aligned}$$

where $t \geq t_1 \geq t_2$. Substituting recursively Eq. (5.151) we have, after N iterations

$$\begin{aligned} |\psi(t)\rangle &= |\psi_{-\infty}\rangle + \sum_{n=1}^{N}(-i)^n \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2 ... \int_{-\infty}^{t_{n-1}} dt_n\, H_{int}(t_1)H_{int}(t_2)...H_{int}(t_n)|\psi_{-\infty}\rangle \\ &\quad +(-i)^{N+1}\int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2\, ... \int_{-\infty}^{t_n} dt_{n+1}\, H_{int}(t_1)H_{int}(t_2)...H_{int}(t_{n+1})|\psi_I(t_2)\rangle\,, \quad (5.154)\end{aligned}$$

where, again, $t \geq t_1 \geq t_2 \geq ... \geq t_n$. Note that Eq. (5.154) is constituted by an operator which is applied to $|\psi_{-\infty}\rangle$ and a "rest" with an additional power of the interaction Hamiltonian that, if we consider H_{int} as a perturbation of H_0, can be considered small with respect to the first term. We assume[2] that the rest goes to zero when $N \to \infty$ and we take this limit. In so doing we find

$$|\psi(t)\rangle = \sum_{n=0}^{\infty}(-i)^n \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2 ... \int_{-\infty}^{t_{n-1}} dt_n\, H_{int}(t_1)H_{int}(t_2)...H_{int}(t_n)|\psi_{-\infty}\rangle \tag{5.155}$$

Now we want to uncorrelate the integrations, in such a way that every integral goes from $t=-\infty$ to t. If we do that, we are introducing also integrations in which the time is not strictly ordered and it can be that $t_i < t_{i+n}$ although at the beginning we had to have $t_i > t_{i+n}$. In order to correct for this, we have to introduce the "Time Ordering" of the product of Hamiltonians evaluated at different times and we have to devide by the number of permutations of the n terms:

$$|\psi(t)\rangle = \sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\int_{-\infty}^{t} dt_1 \int_{-\infty}^{t} dt_2 ... \int_{-\infty}^{t} dt_n\, T\left(H_{int}(t_1)H_{int}(t_2)...H_{int}(t_n)\right)\, |\psi_{-\infty}\rangle\,. \tag{5.156}$$

If we take the limit $t \to \infty$, we can identify the S operator as

$$S = \sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 ... \int_{-\infty}^{\infty} dt_n\, T\left(H_{int}(t_1)H_{int}(t_2)...H_{int}(t_n)\right)\,, \tag{5.157}$$

than can also be written in a formal way as follows:

$$S = T\left(e^{-i\int_{-\infty}^{\infty} H_{int}(t)\, dt}\right) = T\left(e^{-i\int d^4X\, \mathcal{H}_{int}(X)}\right)\,. \tag{5.158}$$

We have to remember that $\mathcal{H}_{int}$ is the interaction Hamiltonian density in Interaction Picture, and therefore it is written in terms of free fields.

[2]This assumption has to be considered with great care.

Capitolo 6

Cross Section and Decay Rate

6.1 From transition amplitude to probability

The transition amplitude has the following form

$$S_{fi} = \delta_{fi} + (2\pi)^4 \delta^4(\sum_i P_i - \sum_f P_f) \prod_{ferm} \sqrt{\frac{m}{VE}} \prod_{bos} \sqrt{\frac{1}{V2E}} \mathcal{M} , \tag{6.1}$$

where δ_{fi} represents the absence of scattering (since we want $i \neq f$ this term is zero), the δ^4 represents the conservation of the total four-momentum, then we have normalization factors for the fermions and for the bosons involved in the scattering and, finally, the matrix element $\mathcal{M}$ that contains the external fields, the interaction vertices and the propagators.

The transition amplitude is not an observable. In order to define a measurable quantity we have, firstly, to move to a probability, taking the modulus squared of S_{fi}:

$$|S_{fi}|^2 = \left|(2\pi)^4 \delta^4(\sum_i P_i - \sum_f P_f)\right|^2 \prod_{ferm} \frac{m}{VE} \prod_{bos} \frac{1}{V2E} \; |\mathcal{M}|^2 \; . \tag{6.2}$$

Let us analyse first the modulus squared of the delta function. In order to do that, it is more convenient to write the delta using its Fourier transform:

$$(2\pi)^4 \delta^4(P_f - P_i) = \lim_{T\to\infty, V\to\infty} \int_V d^3X \int_{-\frac{T}{2}}^{\frac{T}{2}} dt \, e^{i(P_f - P_i)_\mu X^\mu} \, . \tag{6.3}$$

We focus on $\delta(E_f - E_i)$ (for the spatial part we obtain the same result). We have

$$(2\pi)\delta(E_f - E_i) = \lim_{T\to\infty} \int_{-\frac{T}{2}}^{\frac{T}{2}} dt \, e^{i\Delta E\, t} = \lim_{T\to\infty} \frac{2\sin\left(\frac{\Delta E\, T}{2}\right)}{\Delta E} \, . \tag{6.4}$$

Therefore

$$\left|(2\pi)\delta(E_f - E_i)\right| = \lim_{T\to\infty} \frac{4\sin^2\left(\frac{\Delta E\, T}{2}\right)}{\Delta E^2} = \lim_{T\to\infty} 2\pi\, T\, \delta(E_f - E_i) \, . \tag{6.5}$$

The same happens for the spatial part and in the end we obtain

$$\left|(2\pi)^4 \delta^4(\sum_i P_i - \sum_f P_f)\right|^2 \lim_{T\to\infty, V\to\infty} VT\, (2\pi)^4 \delta^4(\sum_i P_i - \sum_f P_f) \, . \tag{6.6}$$

We define the probability density per unit time, or probability density rate, as

$$w_{fi} \;\; = \;\; \frac{|S_{fi}|^2}{T} = V\, (2\pi)^4 \delta^4(\sum_i P_i - \sum_f P_f) \prod_{ferm} \frac{m}{VE} \prod_{bos} \frac{1}{V2E} \; |\mathcal{M}|^2 \; , \tag{6.7}$$

$$= \quad V\,(2\pi)^4 \delta^4(\sum_i P_i - \sum_f P_f) \prod_{ext} \frac{1}{V2E} \prod_{ferm} (2m)\ |\mathcal{M}|^2\ . \tag{6.8}$$

w_{fi} is the probability density per unit time to have the final state "f", with momenta $\mathbf{p}_f$, starting with the initial state "i". However, from an experimental point of view, it is not possible to measure an exact $\mathbf{p}_f$ and one would like to have an interval of momentum, between $\mathbf{p}_f$ and $\mathbf{p}_f + d\mathbf{p}_f$. In this interval, we have a certain number of states, equally probable, on which we have to sum our probability density. If we quantize in the box, the momentum is discrete

$$\mathbf{p} = \frac{2\pi}{L}\mathbf{n}\,, \tag{6.9}$$

where $\mathbf{n}$ is an integer vector. Therefore, the number of states with momentum between $\mathbf{p}_f$ and $\mathbf{p}_f + d\mathbf{p}_f$ is

$$d^3n = \frac{L^3}{(2\pi)^3} d^3p_f = \frac{V}{(2\pi)^3} d^3p_f\,. \tag{6.10}$$

We then have

$$dw_{fi} = V\,(2\pi)^4 \delta^4(\sum_i P_i - \sum_f P_f) \prod_{ext} \frac{1}{V2E} \prod_{ferm} (2m)\ |\mathcal{M}|^2 \prod_f \frac{V}{(2\pi)^3} d^3p_f\,. \tag{6.11}$$

6.2 Cross Section

We define now the observable for scattering processes, which is called Cross Section. We have in mind a process in which a monochromatic beam of particles prepared at $t = -\infty$ collides on a target in which we have a certain density of scatter centers (if we perform a boost in the incoming momentum direction, we can move to the center of mass frame, in which we see the two particles that take part to the scattering that move one against the other). Let us suppose that the beam has a certain section S. If n is the number of incoming particles per unit time and unit surface (N_i the total number of incoming particles in a time interval Δt, $N_i = nS\Delta t$), and N the number of scattered particles per unit time and diffusing particle (N_d the total number of particles scattered per unit diffusing center in Δt, $N_d = N\,\Delta t$) we define the cross section as

$$\sigma = \frac{N}{n}\,. \tag{6.12}$$

The cross section has the dimensions of a surface. In fact

$$\sigma = \frac{N}{n} = \frac{N_d}{\Delta t}\frac{S\Delta t}{N_i} = \frac{N_d}{N_i} S \tag{6.13}$$

and therefore $[\sigma] = l^2$.

n is the incoming flux and can be expressed as the product of the density of incoming particles times the relative velocity of those particles with respect to the diffusing center. In fact we have

$$n = \frac{N_i}{S\Delta t}\frac{L}{L} = \frac{N_i}{V}\frac{L}{\Delta t} = \rho\,|\mathbf{v}_{rel}|\,, \tag{6.14}$$

where L is the linear dimension traveled in Δt by the incoming particles (they all have the same velocity).

If we consider the rate of scattered particles in a certain small region of the phase space, we can define the differential cross section as

$$d\sigma = \frac{dN}{n}\,. \tag{6.15}$$

The dN is exactly the dw_{fi}. Therefore

$$d\sigma = \frac{dN}{n} = \frac{dw_{fi}}{n} = \frac{1}{\rho\,|\mathbf{v}_{rel}|} V(2\pi)^4\delta^4(\sum_i P_i - \sum_f P_f) \prod_{ext} \frac{1}{V2E} \prod_{ferm} (2m)\ |\mathcal{M}|^2 \prod_f \frac{L^3}{(2\pi)^3} d^3p_f \quad (6.16)$$

We consider the situation in which one particle at a time scatters on a diffusing center. In the volume V we will have one incoming particle and therefore

$$\rho = \frac{1}{V}\,. \quad (6.17)$$

Moreover, we have a $2 \to n$ scattering and therefore

$$\begin{aligned} d\sigma &= \frac{V^2}{|\mathbf{v}_{rel}|}(2\pi)^4\delta^4(\sum_i P_i - \sum_f P_f) \prod_{ext} \frac{1}{V2E} \prod_{ferm} (2m)\ |\mathcal{M}|^2 \prod_f \frac{V}{(2\pi)^3} d^3p_f\,, \\ &= (2\pi)^4\delta^4(\sum_i P_i - \sum_f P_f) \frac{1}{4E_1E_2|\mathbf{v}_{rel}|} \prod_{ferm} (2m)\ |\mathcal{M}|^2 \prod_f \frac{d^3p_f}{(2\pi)^3 2E_f}\,, \end{aligned} \quad (6.18)$$

where all the volumes cancel.

The cross section is a Lorentz scalar. In Eq. (6.18) everything is manifestly Lorentz invariant except the flux term, that we have to specify. In fact

$$E_1E_2|\mathbf{v}_{rel}| = E_1E_2|\mathbf{v}_1 - \mathbf{v}_2| = E_1E_2\left|\frac{\mathbf{p}_1}{E_1} - \frac{\mathbf{p}_2}{E_2}\right|\,. \quad (6.19)$$

In the frame in which particle 2 is at rest we have $\mathbf{p}_2 = \mathbf{0}$, $E_2 = m_2$ and therefore

$$E_1E_2|\mathbf{v}_{rel}| = E_1m_2\frac{|\mathbf{p}_1|}{E_1} = m_2|\mathbf{p}_1| = m_2\sqrt{E_1^2 - m_1^2} = \sqrt{m_2^2E_1^2 - m_1^2m_2^2} = \sqrt{(P_{1\,\mu}P_2^\mu)^2 - m_1^2m_2^2}\,, \quad (6.20)$$

which is now written in a manifestly covariant way.

Finally

$$d\sigma = (2\pi)^4\delta^4(\sum_i P_i - \sum_f P_f) \frac{1}{4\sqrt{(P_{1\,\mu}P_2^\mu)^2 - m_1^2m_2^2}} \prod_{ferm} (2m)\ |\mathcal{M}|^2 \prod_f \frac{d^3p_f}{(2\pi)^3 2E_f}\,. \quad (6.21)$$

6.3 Decay Rate

The interaction can cause the decay of a particle, that in the free theory would be stable. This can happen if the kinematic constraints are fulfilled. The process to consider is now a process $1 \to n$ and the initial state is constituted by one particle.

We define the decay rate as the probability density per unit time to have a certain final state starting with the initial state constituted by the particle that decays (the decaying particle has momentum $P^\mu = (E, \mathbf{p})$):

$$d\Gamma = dw_{fi} = (2\pi)^4\delta^4(P - \sum_f P_f) \frac{1}{2E} \prod_{ferm} (2m)\ |\mathcal{M}|^2 \prod_f \frac{d^3p_f}{(2\pi)^3 2E_f}\,. \quad (6.22)$$

Some comments are in order:

- The decay rate is, again, independent on the volume.

- The decay rate in Eq. (6.22) is the "partial decay rate" of the decaying particle in a certain final state. It is governed by the matrix element $|\mathcal{M}|^2$. To understand better, consider the decay of a Z boson in the Standard Model. The Z can decay in different final states. We can for instance compute the decay rate of $Z \to e^+e^-$. This would be

$$d\Gamma_{Z\to e^+e^-} = (2\pi)^4\delta^4(P - P_{e^-} - P_{e^+})\,\frac{1}{2E}\prod_{ferm}(2m)\,|\mathcal{M}_{Z\to e^+e^-}|^2\,\frac{d^3p_{e^-}}{(2\pi)^3 2E_{e^-}}\frac{d^3p_{e^+}}{(2\pi)^3 2E_{e^+}}\,. \quad (6.23)$$

To have the decay rate in this channel, we have to integrate all over the phase space

$$\Gamma_{Z\to e^+e^-} = \int d\Gamma_{Z\to e^+e^-}\,, \quad (6.24)$$

which is a "partial decay rate" because it involves a single channel. The Z boson can decay also into other leptonic pairs or quark pairs. Therefore, if we sum over all the possibilities that the interaction that we are considering allows, we have the total decay rate

$$\Gamma = \sum_f \Gamma_f\,. \quad (6.25)$$

The ratio

$$B_f = \frac{\Gamma_f}{\Gamma} \quad (6.26)$$

is called the "branching ratio" and it gives the probability of finding the state f among the possible decay products.

- All the pieces of formula (6.22) are Lorentz invariant except the term $\frac{1}{2E}$. In fact Γ is not a Lorentz scalar, but it transforms as the inverse of the temporal component of a four-vector. In the frame in which the decaying particle is at rest, this factor becomes $\frac{1}{2M}$, where M is the mass of the particle. In a generic frame in which the decaying particle has velocity β we have

$$E = \gamma M = \frac{M}{\sqrt{1-\beta^2}} \quad (6.27)$$

and therefore the rate in that frame is smaller than the one in the rest frame of the decaying particle by a factor $1/\gamma$

$$\Gamma_E = \frac{1}{\gamma}\Gamma_M\,. \quad (6.28)$$

The lifetime of the particle, which is the inverse of the total rate $\tau = \frac{1}{\Gamma}$, therefore is bigger in the frame in which the particle has velocity β (dilatation of time).

- In the rest frame we have

$$\delta^4(P - \sum_f P_f) \to \delta^4(M - \sum_f P_f) = \delta(M - \sum_f E_f)\delta^3(\sum_f \mathbf{p}_f)\,. \quad (6.29)$$

Therefore

$$M = \sum_f E_f = \sum_f \sqrt{p_f^2 + m_f^2}\,. \quad (6.30)$$

This means that, in the case in which the decay products are massive we have to have $M \geq \sum_f m_f$. The energy at disposal for the decay products is at most M. In the limiting case in which also the decay products are produced at rest, we have $p_f = 0$ and $M = \sum_f m_f$, otherwise the energy M has to go partly into the masses of the particles produced and partly into their momenta.

6.3.1 Two-body phase space

The part of both the cross section and the decay width formulas that has to do with the differentials in the final state momenta is called n-body phase space:

$$d\Phi^{(n)} = (2\pi)^4 \delta^4(\sum_i P_i - \sum_f P_f) \prod_f \frac{d^3 p_f}{(2\pi)^3 2E_f} \,. \tag{6.31}$$

Of particular importance is the two-body phase space. If we consider $f = 3, 4$ and $\sum_i P_i = P$, then

$$d\Phi^{(2)} = (2\pi)^4 \delta^4(P - P_3 - P_4) \frac{d^3 p_3}{(2\pi)^3 2E_3} \frac{d^3 p_4}{(2\pi)^3 2E_4} \,. \tag{6.32}$$

Let us consider, for instance, the case in which a particle of mass M decays into two particles of masses m_3 and m_4. We can calculate $d\Phi^{(2)}$ in the rest frame of the decaying particle. We have

$$d\Phi^{(2)} = (2\pi)^4 \delta^4(M - P_3 - P_4) \frac{d^3 p_3}{(2\pi)^3 2E_3} \frac{d^3 p_4}{(2\pi)^3 2E_4} \,, \tag{6.33}$$

$$= \delta(M - E_3 - E_4) \delta^3(\mathbf{p}_3 + \mathbf{p}_4) \frac{d^3 p_3 \, d^3 p_4}{(2\pi)^2 4E_3 E_4} \,. \tag{6.34}$$

Now suppose we have to integrate all over the phase space. We can integrate first of all in $\mathbf{p}_4$ using the delta ($\mathbf{p}_4 = -\mathbf{p}_3$

$$d\Phi^{(2)} = \frac{1}{(2\pi)^2} \delta(M - E_3 - E_4) \frac{d^3 p_3}{4E_3 E_4} \,, \tag{6.35}$$

where now

$$E_3 = \sqrt{p_3^2 + m_3^2} \,, \quad \text{and} \quad E_4 = \sqrt{p_3^2 + m_4^2} \,, \tag{6.36}$$

since we have to replace everywhere $\mathbf{p}_4$ with $-\mathbf{p}_3$.

Now we have to integrate in $d^3 p_3$. We can write

$$d^3 p_3 = p_3^2 dp_3 d\Omega = p_3^2 dp_3 d\phi d\cos\theta \,, \tag{6.37}$$

where $d\Omega = d\phi d\cos\theta$ is the solid angle and where $p_3 > 0$ is the modulus of $\mathbf{p}_3$. We can integrate in dp_3 as follows

$$d\Phi^{(2)} = \frac{d\Omega}{(2\pi)^2} \int_0^\infty dp_3 \, \frac{p_3^2}{4\sqrt{p_3^2 + m_3^2}\sqrt{p_3^2 + m_4^2}} \delta\left(M - \sqrt{p_3^2 + m_3^2} - \sqrt{p_3^2 + m_4^2}\right) \,. \tag{6.38}$$

For the properties of the delta function

$$\delta(f(x)) = \frac{1}{|f'(x_0)|} \delta(x - x_0) \,, \tag{6.39}$$

where x_0 is a zero of $f(x)$. In the phase space $p_3 \geq 0$ we have a single zero which is

$$p_3 = \frac{1}{2M} \sqrt{M^4 + m_3^4 + m_4^4 - 2M^2 m_3^2 - 2M^2 m_4^2 - 2m_3^2 m_4^2} = \frac{1}{2M} \sqrt{\lambda(M^2, m_3^2, m_4^2)} \,. \tag{6.40}$$

Therefore, in the end we have (substituting the root in Eq. (6.40) into the square roots and simplifying)

$$d\Phi^{(2)} = \frac{d\Omega}{32\pi^2 M^2} \sqrt{\lambda(M^2, m_3^2, m_4^2)} \,. \tag{6.41}$$

In the case in which $m_3 = m_4 = m$ the formula simplifies considerably:

$$d\Phi^{(2)} = \frac{d\Omega}{32\pi^2}\sqrt{1 - \frac{4m^2}{M^2}}\,. \tag{6.42}$$

The same formula holds, mutatis mutandis, for the scattering $2 \to 2$, in which we calculate the cross section in the c.m. frame. If $P_1 + P_2 \to P_3 + P_4$ in the c.m. frame we will have

$$P_1^\mu = (E_1, \mathbf{p})\,, \quad P_2^\mu = (E_2, -\mathbf{p})\,, \tag{6.43}$$

where $E_1 = \sqrt{p^2 + m_1^2}$ and $E_2 = \sqrt{p^2 + m_2^2}$. If we define

$$S = (P_1 + P_2)^2 = (E_1 + E_2)^2\,, \tag{6.44}$$

the energy at disposal for the reaction is $\sqrt{S}$ (which corresponds to M in the case of decay of one particle. For the four-momenta 3 and 4 we will have

$$P_3^\mu = (E_3, \mathbf{p}_3)\,, \quad P_4^\mu = (E_4, -\mathbf{p}_3)\,, \tag{6.45}$$

where $E_3 = \sqrt{p_3^2 + m_3^2}$ and $E_4 = \sqrt{p_3^2 + m_4^2}$. Therefore, with respect to the case of a decay we just have to substitute M with $\sqrt{S}$:

$$d\Phi^{(2)} = \frac{d\Omega}{32\pi^2 S}\sqrt{\lambda(S, m_3^2, m_4^2)}\,, \tag{6.46}$$

in which we have to remember that $\mathbf{p}_4 = -\mathbf{p}_3$ and

$$p_3 = \frac{1}{2\sqrt{S}}\sqrt{S^2 + m_3^4 + m_4^4 - 2Sm_3^2 - 2Sm_4^2 - 2m_3^2 m_4^2}\,. \tag{6.47}$$

6.4 The process $e^+ + e^- \to \mu^+ + \mu^-$

In this section we consider the process $e^+ + e^- \to \mu^+ + \mu^-$. The cross section is given by

$$d\sigma_{e^+e^-\to\mu^+\mu^-} = (2\pi)^4\delta^4(p_1 + p_2 - p_3 - p_4)\frac{(2m_e)^2(2m_\mu)^2}{4\sqrt{(p_1\cdot p_2)^2 - m_1^2 m_2^2}}\frac{d^3p_3}{(2\pi)^3 2E_3}\frac{d^3p_4}{(2\pi)^3 2E_4}\,|\mathcal{M}|^2\,, \tag{6.48}$$

where m_e and m_μ are the masses of the electron and of the muon respectively. Let us now concentrate on the different pieces of the calculation starting from the modulus squared of the transition amplitude.

6.4.1 Modulus Squared of the Transition Amplitude

In the Standard Model there are three families of leptons; in this section we consider only the first two: electron and muons. They differ by the mass: $m_e \sim 0.5$ MeV, $m_\mu \sim 105$ MeV, but they have the same electric charge.

The interaction Lagrangian density is

$$\mathcal{L}_{int} = -e : \left(\bar\psi_e \not{A}\psi_e + \bar\psi_\mu \not{A}\psi_\mu\right) : \tag{6.49}$$

We consider the second order expansion of the S matrix. The T-product inside the integral reads:

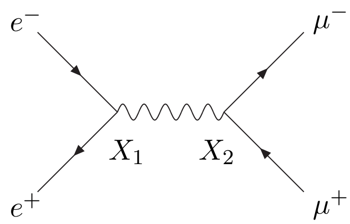
$$\frac{(-i\,e)^2}{2}T\left(: \left(\bar\psi_e \not{A}\psi_e + \bar\psi_\mu \not{A}\psi_\mu\right)_{X_1} :: \left(\bar\psi_e \not{A}\psi_e + \bar\psi_\mu \not{A}\psi_\mu\right)_{X_2} :\right) =$$
$$\frac{(-i\,e)^2}{2}\left\{T\left(: \left(\bar\psi_e \not{A}\psi_e\right)_{X_1} :: \left(\bar\psi_e \not{A}\psi_e\right)_{X_2} :\right) + T\left(: \left(\bar\psi_\mu \not{A}\psi_\mu\right)_{X_1} :: \left(\bar\psi_\mu \not{A}\psi_\mu\right)_{X_2} :\right)\right.$$

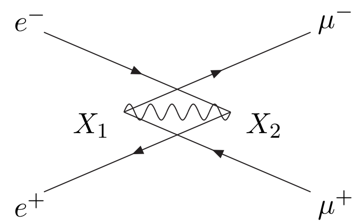
$$+T\left(: (\bar\psi_e \not{A}\psi_e)_{X_1} :: (\bar\psi_\mu \not{A}\psi_\mu)_{X_2} : \right) + T\left(: (\bar\psi_\mu \not{A}\psi_\mu)_{X_1} :: (\bar\psi_e \not{A}\psi_e)_{X_2} : \right)\Big\} \, . \tag{6.50}$$

In order to evaluate the corresponding matrix elements, we apply the Wick's theorem. Operators belonging to different fields cannot be contracted. Moreover, we do not have to consider contractions between two operators evaluated in the same point. Therefore, the only possibility consists in contracting the photon field. The four terms above have matrix elements different from zero for different initial and final states. The first and the second terms in Eq. (6.50) represent electron-positron to electron-positron and muon-anti muon to muon-anti muon scattering processes, respectively. We are interested, instead, in electron-positron to muon-anti muon scattering, that is represented by the third and fourth terms of Eq. (6.50). Considering as initial state $|e^+e^-\rangle$ and final state $|\mu^+\mu^-\rangle$, these two terms give the following contributions:

$$\frac{(-i\,e)^2}{2}\Big\{T\left(: (\bar\psi_e \not{A}\psi_e)_{X_1} :: (\bar\psi_\mu \not{A}\psi_\mu)_{X_2} : \right) + T\left(: (\bar\psi_\mu \not{A}\psi_\mu)_{X_1} :: (\bar\psi_e \not{A}\psi_e)_{X_2} : \right)\Big\} =$$
$$\frac{(-i\,e)^2}{2}\left(:(\bar\psi_e \overbrace{\not{A}\psi_e)_{X_1} : :(\bar\psi_\mu \not{A}}\psi_\mu)_{X_2} : + :(\bar\psi_\mu \overbrace{\not{A}\psi_\mu)_{X_1} : :(\bar\psi_e \not{A}}\psi_e)_{X_2} : \right) . \tag{6.51}$$

We have to select, in the first contribution, the annihilation of an electron and a positron in X_1 and the creation of a muon and an anti-muon in X_2, while, in the second contribution, the annihilation of an electron and a positron in X_2 and the creation of a muon and an anti-muon in X_1. These two contributions can be represented by the following Feynman diagrams (in X space):

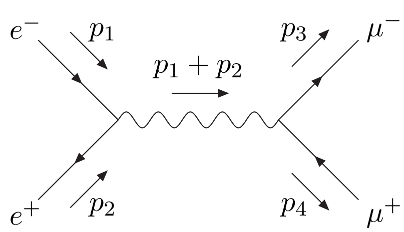




When we integrate in X_1 and X_2, if we exchange X_1 with X_2 in the second term, we find the same contribution coming from the first term, that therefore has to be considered twice:

$$S^{(2)} = (-ie)^2 \int d^4X_1\, d^4X_2 \left(:(\bar\psi_e \overbrace{\not{A}\psi_e)_{X_1} : :(\bar\psi_\mu \not{A}}\psi_\mu)_{X_2} : \right) . \tag{6.52}$$

Moving to momentum space we have then to consider the following Feynman diagram:



The matrix element $\mathcal{M}$ and its complex conjugated are given by

$$\begin{aligned}\mathcal{M} &= \bar u(p_3,n_3)_i\big(-ie\gamma_\nu\big)_{ij} v(p_4,n_4)_j \,\frac{-i\eta^{\mu\nu}}{(p_1+p_2)^2}\,\bar v(p_2,n_2)_k\big(-ie\gamma_\mu\big)_{kl} u(p_1,n_1)_l \, , &(6.53)\\ &= ie^2 \bar u_{3_i}(\gamma_\nu)_{ij} v_{4_j}\,\frac{1}{(p_1+p_2)^2}\,\bar v_{2_k}(\gamma^\nu)_{kl} u_{1_l} \, , &(6.54)\end{aligned}$$

$$\mathcal{M}^* = -ie^2 \bar{v}_{4_{j'}} (\gamma_\rho)_{j'i'} u_{3_{i'}} \frac{1}{(p_1+p_2)^2} \bar{u}_{1_{l'}} (\gamma^\rho)_{l'k'} v_{2_{k'}} , \tag{6.55}$$

where we wrote the products like $\bar{u}_3 \gamma_\nu v_4$ making the components explicit. Finally,

$$|\mathcal{M}|^2 = \frac{e^4}{(p_1+p_2)^4} \left(u_{3_{i'}} \bar{u}_{3_i} (\gamma_\nu)_{ij} v_{4_j} \bar{v}_{4_{j'}} (\gamma_\rho)_{j'i'} \right) \left(v_{2_{k'}} \bar{v}_{2_k} (\gamma^\nu)_{kl} u_{1_l} \bar{u}_{1_{l'}} (\gamma^\rho)_{l'k'} \right) , \tag{6.56}$$

where we grouped together the spinors that refer to the same external momentum.

The expression (6.56) has to be evaluated according to what we intend to measure experimentally. Very often we are interested to unpolarized cross sections. Since we admit an undefined spin state of the final state, quantum mechanically we have to sum over the final state spins. We can reach the same final state both with a certain spin configuration of particle 1 and of particle 2. Therefore, we can sum over the initial state provided that we devide by the different spin states available. In the case of two fermions in the initial state, we have to consider 2 states for each particle and therefore a 1/4 factor overall:

$$|\mathcal{M}|^2 \Longrightarrow \frac{1}{4} \sum_{n,n'} |\mathcal{M}|^2 \; ; \tag{6.57}$$

this means: "sum over the final state spins" and "average over the initial state spins".

Since we have a $\sum_{n,n'}$, in Eq. (6.56) we can recognize the polarization sums:

$$\sum_n u(p,n)\bar{u}(p,n) = \frac{\not{p}+m}{2m} , \quad \sum_n v(p,n)\bar{v}(p,n) = \frac{\not{p}-m}{2m} . \tag{6.58}$$

We have

$$\sum_{n,n'} \left(u_{3_{i'}} \bar{u}_{3_i} \gamma_{\nu_{ij}} v_{4_j} \bar{v}_{4_{j'}} \gamma_{\rho_{j'i'}} \right) \left(v_{2_{k'}} \bar{v}_{2_k} \gamma^{\nu_{kl}} u_{1_l} \bar{u}_{1_{l'}} \gamma^{\rho_{l'k'}} \right) =$$

$$\left(\frac{\not{p}_3 + m_\mu}{2m_\mu} \right)_{i'i} (\gamma_\nu)_{ij} \left(\frac{\not{p}_4 - m_\mu}{2m_\mu} \right)_{jj'} (\gamma_\rho)_{j'i'} \left(\frac{\not{p}_2 - m_e}{2m_e} \right)_{k'k} (\gamma^\nu)_{kl} \left(\frac{\not{p}_1 + m_e}{2m_e} \right)_{ll'} (\gamma^\rho)_{l'k'} , \tag{6.59}$$

$$= tr\Big(\frac{\not{p}_3 + m_\mu}{2m_\mu} \gamma_\nu \frac{\not{p}_4 - m_\mu}{2m_\mu} \gamma_\rho \Big) \, tr\Big(\frac{\not{p}_2 - m_e}{2m_e} \gamma^\nu \frac{\not{p}_1 + m_e}{2m_e} \gamma^\rho \Big) , \tag{6.60}$$

which is the product of two traces on the Dirac indices. Therefore, we have:

$$\frac{1}{4} \sum_{n,n'} |\mathcal{M}|^2 = \overline{|\mathcal{M}|^2} , \tag{6.61}$$

$$= \frac{e^4}{4(p_1+p_2)^4} \, tr\Big(\frac{\not{p}_3 + m_\mu}{2m_\mu} \gamma_\nu \frac{\not{p}_4 - m_\mu}{2m_\mu} \gamma_\rho \Big) \, tr\Big(\frac{\not{p}_2 - m_e}{2m_e} \gamma^\nu \frac{\not{p}_1 + m_e}{2m_e} \gamma^\rho \Big) , \tag{6.62}$$

$$= \frac{e^4}{64 m_e^2 m_\mu^2 (p_1+p_2)^4} tr\Big((\not{p}_3 + m_\mu) \gamma_\nu (\not{p}_4 - m_\mu) \gamma_\rho \Big) \, tr\Big((\not{p}_1 + m_e) \gamma^\rho (\not{p}_2 - m_e) \gamma^\nu \Big) . \tag{6.63}$$

Let us evaluate the two traces, remembering that the trace of an odd number of γ matrices is zero.

$$Tr_1 = tr\Big((\not{p}_3 + m_\mu) \gamma_\nu (\not{p}_4 - m_\mu) \gamma_\rho \Big) = tr\Big(\not{p}_3 \gamma_\nu \not{p}_4 \gamma_\rho - m_\mu^2 \gamma_\nu \gamma_\rho \Big) , \tag{6.64}$$

$$= 4 p_{3\nu} p_{4\rho} + 4 p_{3\rho} p_{4\nu} - 4 (p_3 \cdot p_4) g_{\nu\rho} - 4 m_\mu^2 g_{\nu\rho} , \tag{6.65}$$

$$Tr_2 = tr\Big((\not{p}_1 + m_e) \gamma^\rho (\not{p}_2 - m_e) \gamma^\nu \Big) = ... = 4 p_{1\rho} p_{2\nu} + 4 p_{1\nu} p_{2\rho} - 4 (p_1 \cdot p_2) g_{\nu\rho} - 4 m_e^2 g_{\nu\rho} . \tag{6.66}$$

Therefore, the product of the two traces is

$$Tr_1 \, Tr_2 = 16 \big[(p_1 \cdot p_4)(p_2 \cdot p_3) + (p_1 \cdot p_3)(p_2 \cdot p_4) - (p_1 \cdot p_2)(p_3 \cdot p_4) \big] - 16 m_e^2 (p_3 \cdot p_4)$$

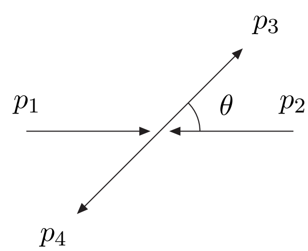
$$
\begin{aligned}
&+16\big[(p_1\cdot p_3)(p_2\cdot p_4)+(p_1\cdot p_4)(p_2\cdot p_3)-(p_1\cdot p_2)(p_3\cdot p_4)\big]-16m_e^2(p_3\cdot p_4)\\
&-16\big[(p_1\cdot p_2)(p_3\cdot p_4)+(p_1\cdot p_2)(p_3\cdot p_4)-4(p_1\cdot p_2)(p_3\cdot p_4)\big]+64m_e^2(p_3\cdot p_4)\\
&-16m_\mu^2(p_1\cdot p_2)-16m_\mu^2(p_1\cdot p_2)+64m_\mu^2(p_1\cdot p_2)+64m_e^2m_\mu^2\,, \qquad (6.67)\\
=\;&32(p_1\cdot p_3)(p_2\cdot p_4)+32(p_1\cdot p_4)(p_2\cdot p_3)+32m_e^2(p_3\cdot p_4)+32m_\mu^2(p_1\cdot p_2)\\
&+64m_e^2m_\mu^2\,. \qquad (6.68)
\end{aligned}
$$

In total, we have:

$$
\overline{|\mathcal{M}|^2}=\frac{e^4}{2m_e^2m_\mu^2(p_1+p_2)^4}\big[(p_1\cdot p_3)(p_2\cdot p_4)+(p_1\cdot p_4)(p_2\cdot p_3)+m_e^2(p_3\cdot p_4)+m_\mu^2(p_1\cdot p_2)+2m_e^2m_\mu^2\big]\,. \qquad (6.69)
$$

6.4.2 Kinematics

In order to express the scalar products, we choose a reference system. Since $|\mathcal{M}|^2$ is invariant it is convenient to calculate it in the center of mass (c.m.) frame. In this frame we have the following situation:



where θ is the so-called scattering angle. Therefore, we have:

$$
p_1^\nu=(E_1,\mathbf{p})\,,\qquad p_2^\nu=(E_2,-\mathbf{p})\,. \qquad (6.70)
$$

Since $p_1^2=m_e^2=p_2^2$, it follows that $E_1=E_2=E$, then

$$
p_1^\nu=(E,\mathbf{p})\,,\qquad p_2^\nu=(E,-\mathbf{p}) \qquad (6.71)
$$

and

$$
(p_1+p_2)^2=4E^2\,. \qquad (6.72)
$$

Also $\mathbf{p}_3$ and $\mathbf{p}_4$ are back-to-back and therefore if we call

$$
p_3^\nu=(E_3,\mathbf{p}')\,,\qquad p_4^\nu=(E_4,-\mathbf{p}')\,, \qquad (6.73)
$$

with $p_3^2=m_\mu^2=p_4^2$, we have to have $E_4=E_3$. Moreover, since $p_1+p_2=p_3+p_4$, we also have

$$
p_1+p_2=2E=p_3+p_4=2E_3\Longrightarrow E_3=E\,. \qquad (6.74)
$$

Finally

$$
p_3^\nu=(E,\mathbf{p}')\,,\qquad p_4^\nu=(E,-\mathbf{p}')\,. \qquad (6.75)
$$

The various scalar products can be expressed in terms of E, p, p' and the scattering angle θ:

$$
\begin{aligned}
p_1\cdot p_2 &= E^2+p^2\,, && (6.76)\\
p_3\cdot p_4 &= E^2+p'^2\,, && (6.77)\\
p_1\cdot p_3 &= E^2-\mathbf{p}\cdot\mathbf{p}'=E^2-pp'\cos\theta=p_2\cdot p_4\,, && (6.78)\\
p_1\cdot p_4 &= E^2+\mathbf{p}\cdot\mathbf{p}'=E^2+pp'\cos\theta=p_2\cdot p_3\,, && (6.79)
\end{aligned}
$$

$$ (6.80) $$

therefore we find:

$$
\begin{aligned}
\overline{|\mathcal{M}|^2} &= \frac{e^4}{2m_e^2 m_\mu^2 16E^4}\big[(E^2 - pp'\cos\theta)^2 + (E^2 + pp'\cos\theta)^2 + m_e^2(E^2 + p'^2) \\
&\qquad + m_\mu^2(E^2 + p^2) + 2m_e^2 m_\mu^2\big]\,, \qquad (6.81)\\
&= \frac{e^4}{2m_e^2 m_\mu^2 16E^4}\big[2E^4 + 2p^2p'^2\cos^2\theta + (m_e^2 + m_\mu^2)E^2 + m_e^2 p'^2 + m_\mu^2 p^2 + 2m_e^2 m_\mu^2\big]. \qquad (6.82)
\end{aligned}
$$

In the cross section, the term $\overline{|\mathcal{M}|^2}$ is multiplied by a factor $(2m_e)^2(2m_\mu)^2$ and then we have

$$
(2m_e)^2(2m_\mu)^2\overline{|\mathcal{M}|^2} = \frac{e^4}{2E^4}\big[2E^4 + 2p^2p'^2\cos^2\theta + (m_e^2 + m_\mu^2)E^2 + m_e^2 p'^2 + m_\mu^2 p^2 + 2m_e^2 m_\mu^2\big]\,. \qquad (6.83)
$$

This means that in the cross section we do not have mass terms in the denominator. Since $m_e \ll m_\mu$ (we also have $m_e^2 \ll E^2$ and $m_\mu^2 \ll E^2$), in Eq. (6.83) we can neglect terms proportional to m_e^2, finding a simpler formula:

$$
\begin{aligned}
(2m_e)^2(2m_\mu)^2\overline{|\mathcal{M}|^2} &\approx \frac{e^4}{2E^4}\big[2E^4 + 2p^2p'^2\cos^2\theta + m_\mu^2 E^2 + m_\mu^2 p^2\big] \\
&= \frac{e^4}{2E^4}\big[2(E^4 + E^2p'^2\cos^2\theta + m_\mu^2 E^2)\big]\,, \\
&= \frac{e^4}{E^2}\big[E^2 + p'^2\cos^2\theta + m_\mu^2\big]\,, \qquad (6.84)
\end{aligned}
$$

since for $m_e \sim 0$ we have $m_e^2 = 0 = E^2 - p^2$ and therefore $p^2 = E^2$.

6.4.3 Flux Factor

The calculation of the flux factor in our case gives the following result:

$$
4\sqrt{(p_1 \cdot p_2)^2 - m_e^2 m_\mu^2} \approx 4\sqrt{(p_1 \cdot p_2)^2} = 8E^2\,. \qquad (6.85)
$$

6.4.4 Cross Section

Finally, merging the various pieces together we find:

$$
\begin{aligned}
d\sigma &= (2\pi)^4\delta^4(p_1 + p_2 - p_3 - p_4)\frac{1}{8E^2}\frac{e^4}{E^2}\big[E^2 + p'^2\cos^2\theta + m_\mu^2\big]\frac{d^3p_3}{(2\pi)^3 2E_3}\frac{d^3p_4}{(2\pi)^3 2E_4}\,, \qquad (6.86)\\
&= \delta^4(p_1 + p_2 - p_3 - p_4)\frac{e^4}{128\pi^2 E^6}\big[E^2 + p'^2\cos^2\theta + m_\mu^2\big]d^3p_3\, d^3p_4\,, \qquad (6.87)\\
&= \delta^4(p_1 + p_2 - p_3 - p_4)\frac{\alpha^2}{8E^6}\big[E^2 + p'^2\cos^2\theta + m_\mu^2\big]d^3p_3\, d^3p_4\,, \qquad (6.88)
\end{aligned}
$$

where we introduced the fine structure constant $\alpha = e^2/(4\pi)$.

If we use the $\delta^4(p_1 + p_2 - p_3 - p_4)$ in the d^3p_4 integration, we find

$$
\begin{aligned}
d\sigma &= \delta(E_1 + E_2 - E_3 - E_4)\,\frac{\alpha^2}{8E^6}\,\big[E^2 + p'^2\cos^2\theta + m_\mu^2\big]\,d^3p'\,, \qquad (6.89)\\
&= \delta^4(2(E - E'))\,\frac{\alpha^2}{8E^6}\,\big[E^2 + p'^2\cos^2\theta + m_\mu^2\big]\,d^3p'\,, \qquad (6.90)
\end{aligned}
$$

$$= \frac{1}{2}\delta(E-E')\,\frac{\alpha^2}{8E^6}\left[E^2+p'^2\cos^2\theta+m_\mu^2\right]d^3p'\,. \tag{6.91}$$

We can express d^3p' in terms of the solid angle

$$d^3p' = p'^2dp'd\Omega(=p'^2dp'd\cos\theta d\phi) \tag{6.92}$$

and calculate the differential cross section, with respect to the solid angle $d\Omega$

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{16E^6}\int\delta(E-E')\left[E^2+p'^2\cos^2\theta+m_\mu^2\right]p'^2\,dp'\,. \tag{6.93}$$

Since $p'^2 = E'^2 - m_\mu^2$, we have

$$\frac{dp'}{dE'} = \frac{E'}{\sqrt{E'^2-m_\mu^2}} = \frac{E'}{p'}\,. \tag{6.94}$$

Finally,

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{16E^6}\int\delta(E-E')\left[E^2+p'^2\cos^2\theta+m_\mu^2\right]p'E'\,dE'\,, \tag{6.95}$$

$$= \frac{\alpha^2}{16E^6}\left[E^2+p'^2\cos^2\theta+m_\mu^2\right]p'E\,, \tag{6.96}$$

$$= \frac{\alpha^2}{16E^4}\frac{p'}{E}\left[E^2+p'^2\cos^2\theta+m_\mu^2\right], \tag{6.97}$$

where we used the fact that now $p' = \sqrt{E^2-m_\mu^2}$.

To find the total cross section we must integrate in $d\Omega$. For simplicity let us consider the ultra-relativistic limit, $E^2 \gg m_\mu^2$. Therefore, in Eq. (6.97) we can neglect the term with m_{mu}^2 ($m_{mu}^2 \sim 0 \Longrightarrow p' = E$) getting

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{16E^2}\left[1+\cos^2\theta\right]. \tag{6.98}$$

Then, we have

$$\sigma = \frac{\alpha^2}{16E^2}\int\left[1+\cos^2\theta\right]d\Omega\,, \tag{6.99}$$

$$= \frac{\alpha^2}{16E^2}2\pi\int_{-1}^{1}\left[1+\cos^2\theta\right]d\cos\theta\,, \tag{6.100}$$

$$= \frac{\alpha^2\pi}{3E^2} = 5.6\cdot10^{-5}\frac{1}{E^2}\,. \tag{6.101}$$

The cross section is now (in natural units) in Energy^{-2}. If $E \sim 100$ GeV, we would obtain

$$\sigma = 5.6\cdot10^{-9}\,\mathrm{GeV}^{-2}\,. \tag{6.102}$$

If we want to express the cross section in barn, we have to remember that

$$1\,\mathrm{GeV}^{-2} = 0.389\,\mathrm{mbarn}\,. \tag{6.103}$$

Therefore:

$$\sigma = 5.6\cdot10^{-9}\,\mathrm{GeV}^{-2} = 2.18\cdot10^{-9}\,\mathrm{mbarn} = 2.18\,\mathrm{pbarn}\,. \tag{6.104}$$