Notes on things

QFT and stuff

First Edition

Notes on Things

QFT and stuff

First Edition

Andrea Barontini
Rome, Italy

Self Publishers Worldwide
Seattle San Francisco New York
London Paris Rome Beijing Barcelona

This book was typeset using LaTeX software.

Preface

This is a (non-ordered) collection of notes that I took to learn new things or to avoid forgetting some of the notions I already learned. It is not structured to have a line connecting the different chapters, but in general it will go from classical field theory to more advanced topics that I used in my PhD.

Table of Contents

Introduction

0.1 Units and Scales

This is just to remember something that might be useful later. even though I am going to use natural units ($c = \hbar = 1$).

Nature has three dimensionful constants:

$$[c] = LT^{-1}$$
$$[\hbar] = L^2MT^{-1}$$
$$[G] = L^3M^{-1}T^{-2}.$$

With the use of natural units, we are only left with one scale which we choose to be energy (or equivalently mass), and that is measured in eV. If one want to recover the unit of lenght or time, he has to reintroduce the relevant powers of $\hbar$ and c, as in the Compton wavelenght

$$\lambda = \frac{\hbar}{mc} \tag{1}$$

which gives, for the electron of mass $m_e = 10^6 eV$, $\lambda_e = 10^{-12}m$.

Remember that the highest energy scale which makes sense in a QFT model is the Planck scale ($M_p \approx 10^{19} GeV$), while the highest lenght scale (which is the smallest energy scale) is the cosmological horizon, $10^{27}cm$.

Chapter 1

Classical field theory

1.1 Classical Fields

A *field* is a quantity defined for every point of space and time $(\vec{x}, t)$, which are then considered as labels of the field

$$\phi_a(\vec{x}, t). \tag{1.1}$$

Electromagnetic field. Having defined the field $A^\mu(\vec{x}, t) = (\phi, \vec{A})$, where $\mu = 0, 1, 2, 3$, the electric and magnetic field are[1]

$$\vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t}$$
$$\vec{B} = \nabla \times \vec{A}$$

which ensure that the Maxwell's equations[2]

$$\nabla \cdot \vec{B} = \nabla \cdot \nabla \times \vec{A} = 0 \tag{1.2}$$
$$\frac{d\vec{B}}{dt} = -\nabla \times \vec{E} \tag{1.3}$$

are satisfied.

1.1.1 Lagrangian

The dynamics of the fields is governed by the **Lagrangian**, which is a function $\phi(\vec{x}, t)$, $\dot{\phi}(\vec{x}, t)$ and $\nabla\phi(\vec{x}, t)$[3]. In the system we usually study, it can be written

[1] Remember the definition of $\nabla = \frac{\partial}{\partial x}\hat{\imath} + \frac{\partial}{\partial y}\hat{\jmath} + \frac{\partial}{\partial k}\hat{z}$.

[2] Note that here we used the property that $\nabla \cdot \nabla \times \vec{v} = 0$ and that $\nabla \times \nabla\phi = 0$.

[3] In principle it could depend on further derivatives but that would not be compatible with Lorentz invariance. Also it could depend on the coordinates themselves $\vec{x}$ but that would not be compatible with locality.

in the form

$$L(t) = \int d^3x \mathcal{L}(\phi_a, \partial_\mu \phi_a)\,, \tag{1.4}$$

with $\mathcal{L}$ the *lagrangian density*. The action is

$$S = \int_{t_1}^{t_2} dt \int d^3x \mathcal{L} = \int d^4x \mathcal{L}. \tag{1.5}$$

Note that $\partial_\mu = \partial/\partial x^\mu$ with $\mu = 0, 1, 2, 3$ with $\mu = 0$ corresponding to the time dimension.

The equations of motions are obtained by the principle of *least action*: we fix the path and we require $\delta S = 0$. So

$$\begin{aligned}\delta S &= \int d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi_a} \delta\phi_a + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \delta(\partial_\mu \phi_a) \right] \\ &= \int d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi_a} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \right) \right] \delta\phi_a + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \delta\phi_a \right)\end{aligned}$$

where we used the fact that $\partial_\mu(f(x_\mu)g(x_\mu)) = (\partial_\mu f(x_\mu))g(x_\mu) + f(x_\mu)(\partial_\mu g(x_\mu))$.

Given that the last term is a total derivative, it is zero for every ϕ that decays as spatial infinity and for which $\delta\phi_a(\vec{x}, t_1) = \delta\phi_a(\vec{x}, t_2) = 0$ (which we always require for the fields). The **Euler-Lagrange** equations are then

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \right) - \frac{\partial \mathcal{L}}{\partial \phi_a} = 0\,. \tag{1.6}$$

Klein-Gordon equation. Consider the following Lagrangian for a real scalar field $\phi(\vec{x}, t)$

$$\mathcal{L} = \frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}m^2\phi^2 = \frac{1}{2}\dot{\phi} - \frac{1}{2}(\nabla\phi)^2\frac{1}{2}m^2\phi^2\,, \tag{1.7}$$

with the **Minkowski space metric**

$$\eta^{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} +1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{1.8}$$

Using the Euler-Lagrange equations we need to compute[4]

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \phi} &= -m^2\phi \\ \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} &= \partial^\mu \phi\,,\end{aligned}$$

so that we obtain

$$\partial_\mu\partial^\mu\phi + m^2\phi = \Box\phi + m^2\phi = 0\,, \tag{1.9}$$

that is the *Klein-Gordon Equation*.

[4]Note that as usual we consider $\partial_\mu\phi$ and ϕ as independent variables. Also we used that $\partial_\nu\phi/\partial_\mu\phi = \eta^\mu{}_\nu$.

Locality

One of the important request that we impose for our Lagrangians is that they have to be *local*. This means that in the Lagrangians there are not terms coupling directly $\phi(\vec{x})$ with $\phi(\vec{y})$ with $\vec{x} \neq \vec{y}$. This is not required by the QFT approach but rather by the fact that we do not know equations of *Nature* that are not local.

1.2 Lorentz invariance

The laws of *Nature* are relativistic so we have to construct theories in which space and time are treated on equal footing and that are invariant under Lorentz transformations,

$$x^\mu \to (x')^\mu = \Lambda^\mu{}_\nu x^\nu \tag{1.10}$$

with $\Lambda^\mu{}_\nu$ such that

$$\Lambda^\mu{}_\sigma \eta^{\sigma\tau} \Lambda^\nu{}_\tau = \eta^{\mu\nu} \, . \tag{1.11}$$

Since the Lorentz transformations form a Lie group under matrix multiplication, they have a *representation* on the fields. For a *scalar* field, we have

$$\phi(x) \to \phi'(x) = \phi(\Lambda^{-1}x) \, . \tag{1.12}$$

This is the case of an **active** transformation, in which the field is truly shifted. This is the reason why we have Λ^{-1}. For a **passive** transformation, in which we relabel our choice of coordinates, we would have

$$\phi(x) \to \phi'(x) = \phi(\Lambda x) \, . \tag{1.13}$$

A Lorentz invariant theory is such if, given $\phi(x)$ that solves the equations of motion, also $\phi(\Lambda^{-1}x)$ solves them. We can ensure this looking at the Lorentz invariance of the action.

Example: Klein-Gordon equation. For a real scalar field we have $\phi(x) \to \phi'(x) = \phi(\Lambda^{-1}x)$, while its derivative transforms as a vector, i.e.

$$(\partial_\mu \phi)(x) \to (\Lambda^{-1})^\nu{}_\mu (\partial_\nu \phi)(\Lambda^{-1}x) \, . \tag{1.14}$$

Therefore, the derivatives terms of the Lagrangian transform as

$$\begin{aligned} \partial_\mu \phi(x) \partial_\nu \phi(x) \eta^{\mu\nu} &\to (\Lambda^{-1})^\rho{}_\mu (\partial_\rho \phi)(\Lambda^{-1}x)(\Lambda^{-1})^\sigma{}_\nu (\partial_\sigma \phi)(\Lambda^{-1}x) \eta^{\mu\nu} \\ &= \partial_\rho \phi(\Lambda^{-1}x) \partial_\sigma \phi(\Lambda^{-1}x) \eta^{\rho\sigma} \end{aligned}$$

which shows that it is indeed invariant. The potential term is proportional to $\phi^2(x) \to \phi^2(\Lambda^{-1}x)$, so it is also invariant and, therefore, the full action is invariant,

$$S = \int d^4x \mathcal{L}(x) \to \int d^4x \mathcal{L}(y = \Lambda^{-1}x) = \int d^4y \mathcal{L}(y) = S \, . \tag{1.15}$$

Note that this required $d^4x = d^4y$ which follows from the fact that the Lorentz transformation we are considering must be *connected to the identity*, and as such it has $\det\Lambda = 1$.

Note also that is easy to see whether a Lagrangian is Lorentz invariant: it is if all the Lorentz indices are contracted with Lorentz invariant objects, such as $\eta_{\mu\nu}$ or the *gamma matrices* γ_μ or also the *total antisymmetric tensor* $\epsilon_{\mu\nu\rho\sigma}$.

1.3 Symmetries

The role of *symmetries* in field theory (quantum or not) is extremely important, mostly due to **Noether's theorem**.

1.3.1 Noether's Theorem

Every continuous symmetry of the Lagrangian gives rise to a conserved current $j^\mu(x)$ such that the equations of motion imply

$$\partial_\mu j^\mu = 0 \,. \tag{1.16}$$

Conserved charge. A conserved current always implies a conserved **charge**,

$$Q = \int_{\mathcal{R}^3} d^3x j^0 \,, \tag{1.17}$$

which can be easily shown computing the time derivative

$$\frac{dQ}{dt} = \int_{\mathcal{R}^3} d^3x \frac{\partial j^0}{\partial t} = -\int_{\mathcal{R}^3} d^3x \nabla \cdot \vec{j} = 0 \,. \tag{1.18}$$

Of course in the last step we used the assumption that $\vec{j} \to 0$ sufficiently quickly as $|\vec{x}| \to \infty$.

Note however that the conservation of a current is a much stronger statement than the conservation of a charge, because it implies that the charge is conserved *locally*, i.e. if we compute its time derivative in a finite volume V, we get

$$\frac{dQ_V}{dt} = -\int_V d^3x \nabla \cdot \vec{j} = -\int_A \vec{j} \cdot d\vec{S} \,. \tag{1.19}$$

This equation means that any charge leaving V must be accounted for a flow of the three-vector $\vec{j}$.

Proof of Noether's theorem. We say that the transformation

$$\delta\phi_a(x) = X_a(\phi) \tag{1.20}$$

is a symmetry of the Lagrangian if it changes by a total derivative[5],

$$\delta\mathcal{L} = \partial_\mu F^\mu(\phi)\,. \tag{1.21}$$

Under (1.20), the Lagrangian transforms as

$$\begin{aligned}\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi_a}\delta\phi_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\partial_\mu(\delta\phi_a)\\ &= \left[\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right]\delta\phi_a + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a\right),\end{aligned}$$

which becomes

$$\delta\mathcal{L} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a\right) \tag{1.22}$$

if ϕ_a follow the equation of motions (which make the first term disappear). In the case of a symmetry we know that this has to take the form of (1.21), so we can define a current

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}X_a(\phi) - F^\mu(\phi)\,, \tag{1.23}$$

which by construction would be such that $\partial_\mu j^\mu = 0$.□

Note that this theorem gives us also a constructive method to find the conserved current.

Example: Translations and Energy-Momentum Tensor.

Consider the infinitesimal translation[6]

$$x^\nu \to x^\nu - \epsilon^\nu \Rightarrow \phi_a(x) \to \phi_a(x) + \epsilon^\nu\partial_\nu\phi_a(x)\,, \tag{1.24}$$

which implies that the Lagrangian transforms as

$$\mathcal{L}(x) \to \mathcal{L}(x) + \epsilon^\nu\partial_\nu\mathcal{L}(x)\,. \tag{1.25}$$

Given that we want translation to be a symmetry of our theory, we can use Noether's theorem to get the four conserved currents[7]

$$(j^\mu)_\nu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\partial_\nu\phi_a - \delta^\mu{}_\nu\mathcal{L} \equiv T^\mu{}_\nu\,. \tag{1.26}$$

$T^\mu{}_\nu$ is called *energy-momentum* tensor and, as shown, it satisfies

$$\partial_\mu T^\mu_\nu = 0\,, \tag{1.27}$$

[5] In this way we are sure that the action S would remain the same under the transformation.

[6] Again here we are doing an active transformation so what we did is $\phi'_a(x) = \phi_a(x+\epsilon)$. With a passive transformation we would have got a minus sign instead.

[7] Note that, with respect to the previous definition, $\phi'_a = \phi_a + \epsilon^\nu(\delta\phi_a)_\nu$ and $\mathcal{L}' = \mathcal{L} + \epsilon^\nu(\delta\mathcal{L})_\nu$.

which gives the four conserved charges

$$E = \int d^3x T^{00} \quad P^i = \int d^3x T^{0i}\,, \tag{1.28}$$

where E is the total energy of the field configuration, while P^i is the total momentum of the field configuration.

Example: Lorentz Transformations and Angular Momentum

We plan to do the same of the previous paragraph but this time imposing the symmetry under Lorentz transformations. Given that they implement rotations, we expect that Noether's theorem in this case should gives us the conservation of angular momentum. However, Lorentz transformations also implement boosts, so let's see what they give us.

We start with the infinitesimal version of a Lorentz transfomation

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu} \tag{1.29}$$

where $\omega^{\mu}{}_{\nu}$ is infinitesimal on which we have some constraint due to (1.11), i.e.

$$\begin{aligned}
&(\delta^{\mu}{}_{\sigma} + \omega^{\mu}{}_{\sigma})(\delta^{\nu}{}_{\tau} + \omega^{\nu}{}_{\tau})\eta^{\sigma\tau} = \eta^{\mu\nu} \\
&\Rightarrow \omega^{\mu\nu} + \omega^{\nu\mu} + \mathcal{O}(\omega^2) = 0\,,
\end{aligned}$$

which means that $\omega^{\mu}{}_{\nu}$ must be *anti-symmetric*. This makes sense because the number of degrees of freedom of a four by four anti-symmetric matrix is 6 which indeed is exactly the number of possible independent Lorentz transformations (3 rotations +3 boosts). We need now to compute the variations of the fields and Lagrangian under this infinitesimal transformation. First, the scalar field transforms as

$$\begin{aligned}
\phi(x) \to \phi'(x) &= \phi(\Lambda^{-1}x) \\
&= \phi(x^{\rho} - \omega^{\rho}{}_{\nu}x^{\nu}) \\
&= \phi(x^{\rho}) - \omega^{\rho}{}_{\nu}x^{\nu}\partial_{\rho}\phi(x)\,,
\end{aligned}$$

so that $(\delta\phi)^{\nu}{}_{\rho} = -x^{\nu}\partial_{\rho}\phi$. In the same way, the variation of the Lagrangian density is[8]

$$(\delta\mathcal{L})^{\nu}{}_{\rho} = -x^{\nu}\partial_{\rho}\mathcal{L} = -\partial_{\rho}(x^{\nu}\mathcal{L})\,. \tag{1.30}$$

Applying Noether's theorem we get the following six conserved currents

$$\begin{aligned}
\mathcal{J}^{\mu\nu}{}_{\rho} &= -\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi)}x^{\nu}\partial_{\rho}\phi + \delta^{\mu}{}_{\rho}x^{\nu}\mathcal{L} \\
&= -x^{\nu}T^{\mu}{}_{\rho}\,,
\end{aligned} \tag{1.31}$$

[8]Where we use the fact that $\delta\mathcal{L}$ is contracted with ω and $\omega^{\mu}{}_{\nu} = 0$.

which can also be written as

$$(\mathcal{J}^\mu)^{\rho\sigma} = x^\rho T^{\mu\sigma} - x^\sigma T^{\mu\rho} \,. \tag{1.32}$$

For $\rho, \sigma = 1, 2, 3$, the Lorentz transformation is a spatial rotation and the three conserved charges give the total angolar momentum of the field configuration,

$$Q^{ij} = \int d^3x (x^i T^{0j} - x^j T^{0i}) \,. \tag{1.33}$$

Instead, what about the boosts? We get

$$Q^{0i} = \int d^3x (x^0 T^{0i} - x^i T^{00}) \,, \tag{1.34}$$

from which we get

$$\begin{aligned} 0 = \frac{dQ^{0i}}{dt} &= \int d^3x T^{0i} + t \int d^3x \frac{\partial T^{0i}}{\partial t} - \frac{d}{dt} \int d^3x x^i T^{00} \\ &= P^i + t \frac{dP^i}{dt} - \frac{d}{dt} \int d^3x x^i T^{00} \,. \end{aligned}$$

Now, we know already that P^i is conserved so we are left with

$$\frac{d}{dt} \int d^3x x^i T^{00} = \text{constant} \,, \tag{1.35}$$

which means that the center of energy of the fields travels with a constant velocity (somewhat like Newton's first law).

1.3.2 Internal Symmetries

An *internal symmetry* involves a transformation of the fields and acts the same at every point in spacetime. To illustrate an example, let's take the complex scalar field $\psi(x) = (\phi_1(x) + i\phi_2(x))/\sqrt{2}$. The usual real Lagrangian is then written as

$$\mathcal{L} = \partial_\mu \psi^* \partial^\mu \psi - V(|\psi|^2) \,. \tag{1.36}$$

To find the equations of motion we could expand everything in term of the real fields $\phi_{1,2}$, but it is simpler to keep working with the complex field, considering ψ and ψ^* as two independent variables.

This Lagrangian has the internal continuous symmetry

$$\psi \to e^{i\alpha} \psi \Rightarrow \delta\psi = i\alpha\psi \,, \tag{1.37}$$

where the latter holds with α infinitesimal. The associated conserved current is[9]

$$j^\mu = i(\partial^\mu \psi^*)\psi - i\psi^*(\partial^\mu \psi) \,. \tag{1.38}$$

We will see that these kind of conserved currents are linked to conservation of electric charge or particle number.

[9] Note that in this case $\delta\mathcal{L} = 0$.

Non-Abelian Internal Symmetries Consider a theory involving N scalar fields ϕ_a, all with same mass, described by the Lagrangian

$$\mathcal{L} = \frac{1}{2}\sum_{a=1}^{N} \partial_\mu \phi_a \partial^\mu \phi_a - \frac{1}{2}\sum_{a=1}^{N} m^2 \phi_a^2 - g\left(\sum_{a=1}^{N} \phi_a^2\right)^2 . \tag{1.39}$$

This Lagrangian is invariant under the non-Abelian symmetry group $SO(N)$ (actually $O(N)$). This kind of symmetries are often called *global* symmetries to distinguish them from the *local gauge* symmetries. Isospin is an example of such a symmetry, albeit realized only approximately in Nature.

1.4 The Hamiltonian Formalism

In order to properly link the Lagrangian formalism with the quantum theory, one needs to use the *path integrals methods*. Here we are instead focusing on **canonical quantization**, for which we need the Hamiltonian formalism of field theory.

The *momentum* conjugate to $\phi_a(x)$ is defined as

$$\pi^a(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_a} , \tag{1.40}$$

and it should not be confused with the total momentum P^i which is a single number describing the whole field configuration. The *Hamiltonian density* is

$$\mathcal{H} = \pi^a(x)\dot{\phi}_a(x) - \mathcal{L}(x) , \tag{1.41}$$

where we need to eliminate the dependence on $\dot{\phi}_a(x)$.

Example: Real Scalar Field The usual Lagrangian is

$$\mathcal{L} = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\nabla\phi)^2 - V(\phi) , \tag{1.42}$$

so the momentum is just $\pi = \dot{\phi}$ and so the Hamiltonian is

$$H = \int d^3x \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + V(\phi) . \tag{1.43}$$

Note that it should agree with the definition of the total energy we can get upon application of the Noether's theorem.

Note also that, while the Lagrangian formalism is *manifestly* Lorentz invariant, this is not the case for the Hamiltonian formalism in which we took a preferred time. For example, the equations of motions are

$$\dot{\phi}(\vec{x},t) = \frac{\partial H}{\partial \pi(\vec{x},t)} \quad \text{and} \quad \dot{\pi}(\vec{x},t) = -\frac{\partial H}{\partial \phi(\vec{x},t)} . \tag{1.44}$$

However, of course the physics is unchanged.

Chapter 2

Free fields

2.1 Canonical Quantization

Canonical quantization is a recipe that takes us from the Hamiltonian formalism of classical fields theory to a quantum theory. The recipe states that we must take the fields $\phi_a(\vec{x})$ and their momentum conjugate $\pi^b(\vec{x})$ and promote them to operators. Thus a *quantum field* is an operator valued function of space obeying the commutation relations

$$[\phi_a(\vec{x}), \phi_b(\vec{y})] = [\pi^a(\vec{x}), \pi^b(\vec{y})] = 0$$
$$[\phi_a(\vec{x}), \pi^b(\vec{y})] = i\delta^{(3)}(\vec{x} - \vec{y})\delta^a_b \,.$$

Note that we have lost track of Lorentz invariance since we separated space and time, and this is because we are working in the Schrodinger picture in which the time dependence is in the states, as

$$i\frac{d|\psi\rangle}{dt} = H|\psi\rangle \,. \tag{2.1}$$

Note that in quantum field theory (QFT) the state $|\psi\rangle$ is a *functional*, i.e. a function of every possible configuration of the field ϕ.

The typical information we want to know when solving a quantum system is the spectrum of H. However in QFT this is usually very hard to get, mostly because in QFT we have an infinite amount of degrees of freedom. In the case of *free theories* we can find a way to write the dynamics such that each degree of freedom evolves independently from the others.

The simplest of such theories is the Klein-Gordon equation for a real scalar field,

$$\partial_\mu \partial^\mu \phi + m^2 \phi^2 = 0 \,. \tag{2.2}$$

We can find the coordinates in which the degrees of freedom decouple applying the Fourier transform, as

$$\phi(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}} \phi(\vec{p},t)\,, \tag{2.3}$$

in such a way $\phi(\vec{p},t)$ satisfies

$$\left(\frac{\partial^2}{\partial t^2} + (\vec{p}^2 + m^2)\right)\phi(\vec{p},t) = 0\,. \tag{2.4}$$

Thus, for each value of $\vec{p}$, $\phi(\vec{p},t)$ solves the equation of a *harmonic oscillator* vibrating at frequency

$$\omega_{\vec{p}} = +\sqrt{\vec{p}^2 + m^2}\,. \tag{2.5}$$

This means that the most general solution to KG equation is a superposition of harmonic oscillators, each vibrating at a different frequency and with a different amplitude. To quantize the field is then necessary to quantize all of these harmonic oscillators. Let's recall how to do this.

2.1.1 The simple Harmonic Oscillator.

Let's take the QM Hamiltonian

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2\,, \tag{2.6}$$

with the canonical relations $[q,p] = i$, and let's define the *creation* and *annihilation* operators

$$a = \sqrt{\frac{\omega}{2}}q + \frac{i}{\sqrt{2\omega}}p \quad , a^\dagger = \sqrt{\frac{\omega}{2}}q - \frac{i}{\sqrt{2\omega}}p \tag{2.7}$$

which can be inverted to get

$$q = \frac{1}{\sqrt{2\omega}}(a + a^\dagger) \quad , p = -i\sqrt{\frac{\omega}{2}}(a - a^\dagger)\,. \tag{2.8}$$

The canonical commutator gives us $[a, a^\dagger] = 1$, while the Hamiltonian can be written as

$$H = \omega\big(a^\dagger a + \frac{1}{2}\big)\,. \tag{2.9}$$

The two operators take us between energy eigenstates because

$$[H, a^\dagger] = \omega a^\dagger \quad , [H, a] = -\omega a \tag{2.10}$$

and so, given $|E\rangle$ eigenstate with energy E ($H|E\rangle = E|E\rangle$),

$$Ha^\dagger|E\rangle = (E+\omega)a^\dagger|E\rangle \quad , Ha|E\rangle = (E-\omega)a|E\rangle\,. \tag{2.11}$$

Given that the energy must be *bounded from below*, there must be a **ground state** $|0\rangle$ such that $a|0\rangle = 0$, so that

$$H|0\rangle = \frac{1}{2}\omega|0\rangle\,. \tag{2.12}$$

2.1.2 The Free Scalar Field

Having recalled the simple harmonic oscillator, let's apply its quantization to the free scalar field. We learned before that the most general solution of the KG equation is a infinite superposition of harmonic oscillators. We then write ϕ and π as a linear sum of creation and annihilation operators, indexed by their three-momentum $\vec{p}$,

$$\phi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left[a_{\vec{p}} e^{i\vec{p}\cdot\vec{x}} + a^\dagger_{\vec{p}} e^{-i\vec{p}\cdot\vec{x}}\right]$$
$$\phi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} (-i)\sqrt{\frac{\omega_{\vec{p}}}{2}} \left[a_{\vec{p}} e^{i\vec{p}\cdot\vec{x}} + a^\dagger_{\vec{p}} e^{-i\vec{p}\cdot\vec{x}}\right].$$

It is then easy to show that

$$[\phi(\vec{x}), \phi(\vec{y})] = [\pi(\vec{x}), \pi(\vec{y})] = 0 \quad , \quad [\phi(\vec{x}), \pi(\vec{y})] = i\delta^{(3)}(\vec{x} - \vec{y})\,, \tag{2.13}$$

which is what we have to require, follows immediately from

$$[a_{\vec{p}}, a_{\vec{q}}] = [a^\dagger_{\vec{p}}, a^\dagger_{\vec{q}}] = 0 \quad , \quad [a_{\vec{p}}, a^\dagger_{\vec{q}}] = (2\pi)^3\delta^{(3)}(\vec{p} - \vec{q})\,. \tag{2.14}$$

Note that this is both **a sufficient and a necessary condition**.

Hint for the proof: The proof is rather simple, just remember that

$$\int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot(\vec{x}-\vec{y})} = \delta^{(3)}(\vec{x} - \vec{y})\,. \tag{2.15}$$

The Hamiltonian. We can now write the Hamiltonian in terms of $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$ operators,

$$\begin{aligned} H &= \frac{1}{2}\int d^3x \pi^2 + (\nabla\phi)^2 + m^2\phi^2 \\ &= \text{long but simple calculation...} \\ &= \frac{1}{4}\int \frac{d^3p}{(2\pi)^3} \frac{1}{\omega_{\vec{p}}} [(-\omega^2_{\vec{p}} + \vec{p}^2 + m^2)(a_{\vec{p}} a_{-\vec{p}} + a^\dagger_{\vec{p}} a^\dagger_{-\vec{p}}) + (\omega^2_{\vec{p}} + \vec{p}^2 + m^2)(a_{\vec{p}} a^\dagger_{\vec{p}} + a^\dagger_{\vec{p}} a_{\vec{p}})]\,, \end{aligned}$$

which then gives us

$$H = \int \frac{d^3p}{(2\pi)^3} \omega_{\vec{p}} \left[a^\dagger_{\vec{p}} a_{\vec{p}} + \frac{1}{2}(2\pi)^3\delta^{(3)}(0)\right], \tag{2.16}$$

once the definition of the frequency $w_{\vec{p}} = \vec{p}^2 + m^2$ is used. We found a strange result: the Hamiltonian seems to be proportional to a delta function evaluated at zero. Also, the integral over $\omega_{\vec{p}}$ is divergent as $\vec{p}^2$ goes to infinity. Let's try to understand this issue better looking at the ground state.

2.1.3 The Vacuum

As we did in the case of the simple harmonic oscillator, let's define the *vacuum* $|0\rangle$ by the state that is annihilated by *all* $a_{\vec{p}}$, i.e.

$$a_{\vec{p}}|0\rangle = 0 \quad \forall \vec{p}. \tag{2.17}$$

Using equation (2.16), this means that only its second term contributes to the energy of the ground state because

$$H|0\rangle = E_0|0\rangle = \left[\int d^3p \frac{1}{2}\omega_{\vec{p}}\delta^{(3)}(0)\right]|0\rangle = \infty|0\rangle\,. \tag{2.18}$$

How should we deal with this infinity? The point in this case is that we have an infinity for *two reasons*: the first is that the space is infinitely large (this is a type of **infra-red divergence**) and the second is that we assumed that our theory is valid at arbitrarily high energies, which is clearly wrong (this is a type of **ultra-violet divergence**).

To see the first it is enough to put the theory on a finite box of length L. In this case

$$(2\pi)^3\delta^{(3)}(0) = \lim_{L\to\infty}\int_{-L/2}^{L/2} d^3x e^{i\vec{x}\cdot\vec{p}}\,|_{\vec{p}=0} = \lim_{L\to\infty}\int_{-L/2}^{L/2} d^3 = V\,, \tag{2.19}$$

which tells us that the infinity arises bacause we are computing the total energy instead of the *energy density*

$$\mathcal{E}_0 = \frac{E_0}{V} = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2}\omega_{\vec{p}}\,. \tag{2.20}$$

Note that this is still divergent for the second reason, namely that $\mathcal{E}_0 \to \infty$ when $|\vec{p}| \to \infty$.

The second problem can be cured cutting the integral at high-momentum in order to reflect the fact that our theory is not valid in that region. However, we can adopt a more practical a solution, exploiting the fact that in physics we only measure differences, since absolute quantities cannot be measured (and they do not even make sense). So we can just subtract the infinity which leaves us with the Hamiltonian

$$H = \int \frac{d^3p}{(2\pi)^3}\omega_{\vec{p}} a^\dagger_{\vec{p}} a_{\vec{p}}\,, \tag{2.21}$$

such that $H|0\rangle = 0$. Note that this can be also seen as an ordering ambiguity that arises from the transition between classical and quantum theories. This is the reason why the method we just used is called *normal ordering*.

Normal Ordering: The normal ordering of a string of operators is denoted as

$$:\phi_1(\vec{x}_1)\ldots\phi_n(\vec{x}_n): \tag{2.22}$$

and it consists of putting all the annihilation operators $a_{\vec{p}}$ to the right.

2.2 Particles

We understood the vacuum, so we can now start speaking of the *excitations* of the fields. As in the simple harmonic oscillator case, we can construct energy eigenstates by acting on the vacuum $|0\rangle$ with $a_{\vec{p}^\dagger}$, as

$$|\vec{p}\rangle = a^\dagger_{\vec{p}}|0\rangle\,, \tag{2.23}$$

which has energy $H|\vec{p}\rangle = \omega_{\vec{p}}|\vec{p}\rangle$ with $\omega_{\vec{p}} = \vec{p}^2 + m^2$. Note that we used that

$$[H, a^\dagger_{\vec{p}}] = \omega_{\vec{p}} a^\dagger_{\vec{p}} \quad \text{and} \quad [H, a_{\vec{p}}] = -\omega_{\vec{p}} a_{\vec{p}}\,. \tag{2.24}$$

We recognize this as the *relativistic dispersion* of a particle of mass m and 3-momentum $\vec{p}$, i.e. $E^2_{\vec{p}} = \vec{p}^2 + m^2$. We then interpret the state $|\vec{p}\rangle$ as the momentum eigenstate of a single particle state of mass m (and from now on we will only write $E_{\vec{p}}$ instead of ω_p). However we want to check this particle interpretation by looking at the other quantum numbers of the state. Taking the total momentum $\vec{P}$ and turning it into a quantum operator (and also applying normal ordering), we get

$$\vec{P} = -\int d^3x \pi \vec{\nabla}\phi = \int \frac{d^3p}{(2\pi)^3} \vec{p} a^\dagger_{\vec{p}} a_{\vec{p}}\,. \tag{2.25}$$

Applying $\vec{P}$ on our momentum eigenstate we indeed get that

$$\vec{P}|\vec{p}\rangle = \vec{p}|\vec{p}\rangle\,, \tag{2.26}$$

which means that it is indeed an eigenstate of the total momentum. We can also study its *angular momentum* by once again promoting the classical total angular momentum to its quantum version, as

$$J^i = \epsilon^{ijk} \int d^3x (\mathcal{J}^0)^{jk} \tag{2.27}$$

(remember equation (1.31)), and then applying it to $|\vec{p}\rangle$. It is easy to show that $J^i|\vec{p} = 0\rangle = 0$ (it would not be zero if we chose a state with momentum different from zero), which means that the particle carries no internal angular momentum (indeed the momentum is zero so it can not have external angular momentum). In other words, the scalar field gives rise to a **spin** 0 **particle**.

2.2.1 Multi-Particle States, Boson Statistics and Fock Space

Multi-particle states are obtained acting multiple times with $a^\dagger$ on the vacuum, as

$$|\vec{p}_1, \ldots, \vec{p}_n\rangle = a^\dagger_{\vec{p}_1} \ldots a^\dagger_{\vec{p}_n}|0\rangle\,. \tag{2.28}$$

Since the $a^\dagger$s commute with each other, the state is *symmetric* upon exchange of two particles. In other words, the particles we are describing are *bosons*.

The full Hilbert space is obtained acting on the vacuum with all the possible combinations of $a^\dagger$s. This space is called *Fock space* and it simply is the sum of the $n-$particle Hilbert spaces, for all the $n \geq 0$. In order to count the number of particles of a given state in a Fock space, we have to define the *number operator* N

$$N = \int \frac{d^3p}{(2\pi)^3} a^\dagger_{\vec{p}} a_{\vec{p}} \tag{2.29}$$

which satisfies $N|\vec{p}_1, \ldots, \vec{p}_n\rangle = n|\vec{p}_1, \ldots, \vec{p}_n\rangle$. For a *free theory*, the number operator commutes with the Hamiltonian, $[H, N] = 0$, ensuring the number of particles is conserved. This will no longer be true when we will deal with interacting theories which can create and annihilate particles.

2.2.2 Operator Valued Distributions

So far we have referred to the momentum eigenstates $|\vec{p}\rangle$ as *particles*. However, they are not localized in space in any way. Recall that in quantum mechanics, space and momentum eigenstates are not good states given that they are not properly normalizable (they normalize to delta-functions). Similarly, in QFT neither $\phi(\vec{x})$ nor $a_{\vec{p}}$ are good operators acting on the Fock space. Again, they produce not normalizable states, for example

$$\langle 0|a_{\vec{p}} a^\dagger_{\vec{p}}|0\rangle = \langle \vec{p}|\vec{p}\rangle = (2\pi)^3\delta(0) \quad and \quad \langle 0|\phi(\vec{x})\phi(\vec{x})|0\rangle = \langle \vec{x}|\vec{x}\rangle = \delta(0) \tag{2.30}$$

are *operator valued distributions* rather than functions. This means that, although the vacuum expectation values is well defined as $\langle 0|\phi(\vec{x})|0\rangle = 0$, the fluctuations at a fixed point are infinite $\langle 0|\phi(\vec{x})\phi(\vec{x})|0\rangle = \infty$. A way to construct well defined operators is to smear these distributions, for example creating a wavepacket[1]

$$|\phi\rangle = \int \frac{d^3p}{(2\pi)^3} e^{-i\vec{p}\cdot\vec{x}} \phi(\vec{p})|\vec{p}\rangle\,, \tag{2.31}$$

which is then partially localized in both space and momentum space (note that typically $\phi(\vec{p}) = exp(-\vec{p}^2/2m^2)$).

2.2.3 Relativistic Normalization

According to how we defined the states, we have now

$$\langle \vec{p}|\vec{q}\rangle = (2\pi)^3\delta^{(3)}(\vec{p} - \vec{q})\,. \tag{2.32}$$

Is this Lorentz invariant? Surely the invariance is not evident given that we only have three-vectors. So let's take an object that we know is Lorentz invariant (that is the projection operator on one-particle states)

$$1 = \int \frac{d^3p}{(2\pi)^3} |\vec{p}\rangle\langle\vec{p}|\,, \tag{2.33}$$

[1]Note tha abuse of notation: we are using ϕ now for the field ($\phi(\vec{x})$), for the state ($|\phi\rangle$) and also for the wavefunction ($\phi(\vec{p})$).

This is surely Lorentz invariant but not because the two pieces, the measure d^3p and the projector $|\vec{p}\rangle\langle\vec{p}|$ are individually Lorentz invariant. The Lorentz invariant measure is something different, in particular

$$\frac{d^3p}{2E_{\vec{p}}}\,. \tag{2.34}$$

To show this, let's start by saying that d^4p is obviously Lorentz invariant. The relativistic dispersion of a massive particle,

$$p_\mu p^\mu = m^2 \Rightarrow p_0^2 = E_{\vec{p}}^2 = \vec{p}^2 + m^2\,, \tag{2.35}$$

is also Lorentz invariant. Therefore, putting them together, the following combination must be Lorentz invariant

$$\int d^4p\delta(p_0^2 - \vec{p}^2 - m^2) = \int \frac{d^3p}{2p_0}\,|_{p_0=E_{\vec{p}}}\,, \tag{2.36}$$

which is what we wanted to show.

From this result we can get everything else. For example, the Lorentz invariant $\delta-$function for 3−vectors is

$$2E_{\vec{p}}\delta^{(3)}(\vec{p}-\vec{q}) \tag{2.37}$$

which follows from

$$\int \frac{d^3p}{2E_{\vec{p}}} 2E_{\vec{p}}\delta^{(3)}(\vec{p}-\vec{q}) = 1\,. \tag{2.38}$$

The relativistically normalized momentum states are then

$$|p\rangle = \sqrt{2E_{\vec{p}}}|\vec{p}\rangle = \sqrt{2E_{\vec{p}}}a_{\vec{p}}^\dagger|0\rangle\,, \tag{2.39}$$

from which we get

$$\langle p|q\rangle = (2\pi)^2 2E_{\vec{p}}\delta^{(3)}(\vec{p}-\vec{q})\,. \tag{2.40}$$

2.2.4 Complex Scalar Fields

Let's consider a complex scalar field $\psi(x)$ (as opposed to the real scalar field $\phi(x)$ we considered so far), with the Lagrangian

$$\mathcal{L} = \partial_\mu\psi^*\partial^\mu\psi - M^2\psi\psi^*\,. \tag{2.41}$$

Note that in this case there is no 1/2 in the definition of the Lagrangian, which we would get back if we write the complex field in terms of the real fields, as $\psi = (\phi_1 + i\phi_2)/\sqrt{2}$. The equations of motion are

$$\partial_\mu\partial^\mu\psi + M^2\psi = 0$$
$$\partial_\mu\partial^\mu\psi^* + M^2\psi^* = 0$$

where it is clear that the second is the complex conjugate of the first. Passing to the quantum fields $\psi, \psi^{\dagger}$[2], we expand them as a sum of plane waves

$$\begin{aligned}\psi &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left(b_{\vec{p}}e^{i\vec{p}\cdot\vec{x}} + c^{\dagger}_{\vec{p}}e^{-i\vec{p}\cdot\vec{x}}\right)\\ \psi^{\dagger} &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left(b^{\dagger}_{\vec{p}}e^{-i\vec{p}\cdot\vec{x}} + c_{\vec{p}}e^{i\vec{p}\cdot\vec{x}}\right).\end{aligned}$$

Let's do the same for the quantum version of the classical momentum $\pi = \partial\mathcal{L}/\partial\dot{\psi} = \psi^*$,

$$\begin{aligned}\pi &= \int \frac{d^3p}{(2\pi)^3} i\sqrt{\frac{E_{\vec{p}}}{2}}\left(b^{\dagger}_{\vec{p}}e^{-i\vec{p}\cdot\vec{x}} - c_{\vec{p}}e^{i\vec{p}\cdot\vec{x}}\right)\\ \pi^{\dagger} &= \int \frac{d^3p}{(2\pi)^3}(-i)\sqrt{\frac{E_{\vec{p}}}{2}}\left(b_{\vec{p}}e^{i\vec{p}\cdot\vec{x}} - c^{\dagger}_{\vec{p}}e^{-i\vec{p}\cdot\vec{x}}\right).\end{aligned}$$

The commutation relations are

$$[\psi(\vec{x}), \pi(\vec{y})] = i\delta^{(3)}(\vec{x}-\vec{y}) \quad \text{and} \quad [\psi(\vec{x}), \pi^{\dagger}(\vec{y})] = 0 \tag{2.42}$$

together with the ones that can be obtained by complex conjugation and the usual ones as $[\psi(\vec{x}), \psi^{\dagger}(\vec{y})] = 0$, etc. As in the case of a real scalar field, these commutation relations are equivalent to the usual ones for the creation and annihilation operators.

In summary, quantizing a complex scalar field gives us two creation operators $b^{\dagger}$ and $c^{\dagger}$ that create two types of particles, both of mass M and both with spin zero. They are thus interpreted as particles and anti-particles (as opposed to the case of a real scalar field for which the particle is its own antiparticle).

We have already shown that this theory a classical conserved charge

$$Q = i\int d^3x(\dot{\psi}^*\psi - \psi^*\dot{\psi}) = i\int d^3x(\pi\psi - \psi^*\pi^*), \tag{2.43}$$

which, after normal ordering, becomes the quantum operator

$$Q = \int \frac{d^3p}{(2\pi)^3}(c^{\dagger}_{\vec{p}}c_{\vec{p}} - b^{\dagger}_{\vec{p}}b_{\vec{p}}) = N_c - N_b \tag{2.44}$$

that counts the difference between the number of anti-particles and the number of particles. Since we have that $[H, Q] = 0$, Q is a conserved quantity. For a free theory this is not a big deal because we already know that the number of particles is conserved in a free theory, so that N_c and N_b are individually conserved quantities. However, this difference will still be conserved in an interacting theory, while the absolute number of particles will not.

[2]Since the classical fields are not real, the quantum fields are not *hermitian*.

2.2.5 The Heisenberg Picture

Given that so far we worked in the Schrodinger picture, it is not obvious at all that our construction is Lorentz invariant. For example, our operators $\phi(\vec{x})$ depend on space but not in time and the one-particle states evolve in time as

$$i\frac{d|\vec{p}(t)\rangle}{dt} = H|\vec{p}(t)\rangle \Rightarrow |\vec{p}(t)\rangle = e^{-iE_{\vec{p}}t}|\vec{p}\rangle \,. \tag{2.45}$$

Things look better when one starts to work in the *Heisenberg* picture, where time dependence is given to the operators,

$$\mathcal{O}_H = e^{iHt}\mathcal{O}_S e^{-iHt} \,, \tag{2.46}$$

so that

$$\frac{d\mathcal{O}_H}{dt} = i[H, \mathcal{O}_H] \,. \tag{2.47}$$

In QFT we will not use the subscripts H and S to denote the picture we are using, but we will be using Schrodinger picture when the fields depend only on space ($\phi(\vec{x})$) and Heisenberg picture when the fields depend on spacetime ($\phi(\vec{x}, t)$).

In the Heisenberg picture, the commutations relations become *equal time* commutations relations, as

$$\begin{aligned}
[\phi(\vec{x}, t), phi(\vec{y}, t)] &= [\pi(\vec{x}, t), \pi(\vec{y}, t)] = 0 \\
[\phi(\vec{x}, t), \pi(\vec{y}, t)] &= i\delta^{(3)}(\vec{x} - \vec{y}) \,.
\end{aligned}$$

Now that the operators depend on time we can study how they evolve, as

$$\begin{aligned}
\dot{\phi} = i[H, \phi] &= \frac{i}{2}\left[\int d^3y \pi(y)^2 + \nabla\phi(y)^2 + m^2\phi(y)^2 \quad , \quad \phi(x)\right] \\
&= i\int d^3y \pi(y)^2(-i)\delta^{(3)}(\vec{y} - \vec{x}) = \pi(x) \,,
\end{aligned}$$

and

$$\begin{aligned}
\dot{\pi} = i[H, \pi] &= \frac{i}{2}\left[\int d^3y \pi(y)^2 + \nabla\phi(y)^2 + m^2\phi(y)^2 \quad , \quad \pi(x)\right] \\
&= -\left(d^3y(\nabla_y\delta^{(3)}(\vec{x} - \vec{y}))\nabla_y\phi(y)\right) - m^2\phi(x) \\
&= \nabla^2\phi - m^2\phi \,.
\end{aligned}$$

Putting the last two equation together, we find that the field operator ϕ satisfies the KG equation

$$\partial_\mu\partial^\mu\phi + m^2\phi = 0 \,, \tag{2.48}$$

which starts looking more relativistic.

We can also rewrite the Fourier expansion of the field in the Heisenberg picture as

$$\phi(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}(a_{\vec{p}}e^{-ip\cdot x} + a^\dagger_{\vec{p}}e^{ip\cdot x})\,, \tag{2.49}$$

which looks very similar to the one in Schrodinger picture but now the exponents are written in terms of 4−vectors instead of 3−vectors.

Causality. We almost solved the Lorentz invariance issue with the Heisenberg picture, but there is still a hint of non-Lorentz invariance: ϕ and π satisfy *equal time* commutation relations

$$[\phi(\vec{x},t),\pi(\vec{y},t)] = i\delta^{(3)}(\vec{x}-\vec{y})\,. \tag{2.50}$$

What about arbitrary spacetime separations? We need to require causality in out theory, so that all spacelike separated operators commute,

$$[\mathcal{O}_1(x),\mathcal{O}_2(y)] = 0 \quad \forall (x-y)^2 < 0\,, \tag{2.51}$$

in such a way any measurement happening at x cannot affect a measurement happening at y. In order to check whether our theory satisfies this property, let's define

$$\Delta(x-y) = [\phi(x),\phi(y)]\,, \tag{2.52}$$

which can be written as

$$\Delta(x-y) - \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)})\,, \tag{2.53}$$

showing that $\Delta(x-y)$ is just a complex number function. About it we know

- It is Lorentz invariant, thanks to the measure $d^3p/2E_{\vec{p}}$.
- It does not vanish for timelike separations. Indeed taking $x-y=(t,0,0,0)$, one has $p\cdot(x-y)=mt$ and the integral does not vanish.
- It vanishes for space-like separations. In order to see this, let's start by noting that $\Delta(x-y)=0$ at equal times for all $(x-y)^2 = -(\vec{x}-\vec{y})^2 < 0$, in fact[3]

$$[\phi(\vec{x},t),\phi(\vec{y},t)] = \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{\vec{p}^2+m^2}}(e^{i\vec{p}\cdot(\vec{x}-\vec{y})} - e^{-i\vec{p}\cdot(\vec{x}-\vec{y})})\,, \tag{2.54}$$

  and we can flip the sign of $\vec{p}$ in the second exponent since it is an integration variable. But since $\Delta(x-y)$ is Lorentz invariant, it can depend only on $(x-y)^2$ so if it vanishes at equal time, it has to vanish even when the times are not equal.

So we confirmed that our theory is indeed Lorentz invariant with commutators that vanish outside the lightcone. This property will continue to hold even for interacting theories. However, the fact that $\Delta(x-y)$ is a complex number function is only true for free fields.

[3]Here we compute that integral imposing $x^0 = y^{(0)}$ since we are doing it at equal time.

2.2.6 Propagators

In order to probe causality, we could ask ourselves a different question: what is the amplitude of finding a particle at point x, if it has been prepared at point y? We can compute this amplitude as

$$\begin{aligned}\langle 0|\phi(x)\phi(y)|0\rangle &= \int \frac{d^3p d^3p'}{(2\pi)^6}\frac{1}{\sqrt{4E_{\vec{p}}E_{\vec{p}'}}}\langle 0|a_{\vec{p}}a^\dagger_{\vec{p}'}|0\rangle e^{-ip\cdot x+ip'\cdot y}\\ &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}e^{-ip\cdot(x-y)} = D(x-y)\,.\end{aligned}$$

The function $D(x-y)$ is called *propagator*. For *spacelike* separations, $(x-y)^2 < 0$, it decays like $\sim e^{-m|\vec{x}-\vec{y}|}$. So the quantum field appears to leak outside the lightcone! This is strange because we have just seen that our theory is causal. Why is that then? Well, note that we can rewrite the previous calculation as

$$[\phi(x),\phi(y)] = D(x-y) - D(y-x) = 0 \quad \text{if} \quad (x-y)^2 < 0\,. \tag{2.55}$$

What does it mean? It means that when $(x-y)^2 < 0$ there is no Lorentz invariant ways to order events: if a particle can travel from x to y, it can just as easily travel from y to x. The two amplitudes are then not zero individually, but they cancel with each other.

Even more interesting is the interpretation for a complex scalar field. In this case the amplitude for the particle to propagate from x to y cancels with the amplitude of an *antiparticle* to travel from y to x.

The Feynman Propagator

One of the most important quantities in an interacting theory is the *Feynman propagator*,

$$\Delta_F(x-y) = \langle 0|T[\phi(x)\phi(y)]|0\rangle = \begin{cases} D(x-y) & x^0 > y^0 \\ D(y-x) & y^0 > x^0 \end{cases} \tag{2.56}$$

where $T[]$ stands for *time ordering* and which places all operators evaluated at later times to the left, i.e.

$$T[\phi(x)\phi(y)] = \begin{cases} \phi(x)\phi(y) & x^0 > y^{(0)} \\ \phi(y)\phi(x) & y^0 > x^{(0)}\,. \end{cases} \tag{2.57}$$

Claim: We can write the Feynman propagator as[4]

$$\Delta_F(x-y) = \int \frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2}e^{-ip\cdot(x-y)}\,. \tag{2.58}$$

[4]Note that here we are integrating on the 4−momentum, while so far we only integrated in the 3−momentum, with p^0 fixed by the mass-shell condition.

Note that as it is, this integral is ill defined because for each value of $\vec{p}$, we have a pole on $p^0 = \pm E_{\vec{p}} = \pm\sqrt{\vec{p}^2 + m^2}$. We need a prescription to integrate in p^0 avoiding such singularities. To get the Feynman propagator we must choose a contour that passes below the singularity in $-E_{\vec{p}}$ and passes above the singularity in $E_{\vec{p}}$. **Proof:** The residue at the pole is $\pm 1/2E_{\vec{p}}$, because

$$\frac{1}{p^2 - m^2} = \frac{1}{(p^0 - E_{\vec{p}})(p^0 + E_{\vec{p}})} \,. \tag{2.59}$$

Now we need to treat the two cases, $x^0 > y^0$ and $y^0 > x^0$, separately. In particular, for $x^0 > y^0$ we close the contour in the lower half plane, ensuring $e^{-ip^0(x^0-y^0)} \to 0$, while for $y^0 > x^0$ we close in the upper half plane. In the former case, we get

$$\begin{aligned}\Delta_F(x-y) &= \int \frac{d^3p}{(2\pi)^4} \frac{-2\pi i}{2E_{\vec{p}}} e^{-iE_{\vec{p}}(x^0-y^0) + i\vec{p}\cdot(\vec{x}-\vec{y})} \\ &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} e^{-ip(x-y)} = D(x-y)\,,\end{aligned}$$

and it is easy to show that in the other case we would get $D(y-x)$, showing exactly what we wanted to show.

In the definition of the Feynman propagator we usually do not specify the contour, but rather we write it as

$$\Delta_F(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{ie^{-ip(x-y)}}{p^2 - m^2 \pm i\epsilon}\,, \tag{2.60}$$

with $\epsilon > 0$ and infinitesimal. Of course in this way the poles are shifted off the real axis and this has the same effect of the contour.

2.2.7 Interactions

We will now start to discuss *interacting theories*. In particular, we will see that we usually consider interactions *between particles*, which arise from pieces of the Lagrangian like

$$\Delta\mathcal{L} = \psi^*(\vec{x})\psi^*(\vec{x})\psi(\vec{x})\psi(\vec{x})\,, \tag{2.61}$$

which destroy two particles before creating two new ones.

Chapter 3

Interactions

The free fields we have discussed so far are the only case in which we can determine the spectrum but other than that, nothing happens. We have particles but they do not interact with each other. We start now discussing interacting theories.

The interaction terms are higher order terms in the Lagrangian, as in the real scalar field case

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \sum_{n\geq 3}\frac{\lambda_n}{n!}\phi^n\,, \tag{3.1}$$

where the coefficients λ_n are called *coupling constants*. Also, what we usually ask ourselves is what we need to do in order to make these new terms a *small perturbation* to the original Lagrangian. In order to see what we need to do, let's start doing some dimensional analysis. We know the action S has the same dimensions of $\hbar$, so, given that we took $\hbar = 1$, $[S] = 0$. Given that $S = \int d^4x\mathcal{L}$, this means that $[\mathcal{L}] = 4$. In turn this means that

$$[\phi] = 1 \quad , \quad [m] = 1 \quad , \quad [\lambda_n] = 4 - n\,, \tag{3.2}$$

which tells us that asking λ_n to be small does not make sense in general, because they are, in general, dimensionful quantities. We have then three cases

- $[\lambda_3] = 1$: in this case the dimensionless parameter is λ_3/E, where typically E is the typical energy scale of the process. This means that the term $\lambda_3\phi^3/3!$ is a small perturbation at *high energy*, but they are big at smaller energies.They are called *relevant.*
- $[\lambda_4] = 0$: the perturbation is small simply if $\lambda_4 \ll 1$. They are called *marginal.*
- $[\lambda_n] < 0$: the perturbation is small at low energy and they are called *irrelevant.* Note that this kind is the most problematic because in QFT we cannot always avoid high=energy regions and these perturbations diverge at high energy. Indeed, they give rise to *non-renormalizable* theories.

3.0.1 Examples of Weakly coupled theories

We will only see for the moment theories for which the perturbations are really small at all energies, called *weakly coupled* field theories. In particular, we are going to see

1. ϕ^4 **theory:**
$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 \tag{3.3}$$
with $\lambda \ll 1$. Expanding ϕ^4 in terms of creation and annihilation operators, we get something like
$$a^\dagger_{\vec{p}}a^\dagger_{\vec{p}}a^\dagger_{\vec{p}}a^\dagger_{\vec{p}} \quad \text{and} \quad a^\dagger_{\vec{p}}a^\dagger_{\vec{p}}a^\dagger_{\vec{p}}a_{\vec{p}} \quad \text{etc.} \tag{3.4}$$
These terms will create and destroy particles, making the number operator not conserved, $[H, N] \neq 0$.

2. **Scalar Yukawa Theory:**
$$\mathcal{L} = \partial_\mu\psi^*\partial^\mu\psi + \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - M^2\psi^*\psi - \frac{1}{2}m^2\phi^2 - g\psi^*\psi\phi \tag{3.5}$$
with $g \ll M, m$[1]. This is a theory in which a complex scalar field is coupled to a real scalar field. Again the number operator N is not conserved but the difference between the number of ψ particles and $\bar{\psi}$ anti-particles is still conserved.

Note that also strongly coupled theories can be treated and they give rise to important effects.

3.1 The Interaction Picture

This is a concept that we can introduce even in classical quantum mechanics. We know that in Schrodinger picture states evolve in time, while in Heisenberg picture the operators evolve in time. In the *interaction picture* we split the Hamiltonian in two pieces
$$H = H_0 + H_{\text{int}} \tag{3.6}$$
and the time dependence of the operators is governed by H_0, while the time dependence of states by H_{int}. This is very convenient when H_0 is soluble (for example if it is a free theory).

We will denote states and operators in the interaction picture with a subscript I and they are given by

$$|\psi(t)\rangle_I = e^{iH_{\text{int}}t}|\psi(t)\rangle_S$$
$$\mathcal{O}_I(t) = e^{iH_0 t}\mathcal{O}_S e^{-iH_0 t}\,.$$

[1]Note that this is a case in which $[g] = 1$.

The interaction Hamiltonian in the interaction picture is then

$$H_I = (H_{\text{int}})_I = e^{iH_0 t}(H_{\text{int}})_S e^{-iH_0 t}\,, \tag{3.7}$$

so that one can obtain

$$i\frac{d|\psi\rangle_I}{dt} = H_I(t)|\psi\rangle_I\,. \tag{3.8}$$

3.1.1 Dyson's Formula

In order to solve (3.8), we write the solution as

$$|\psi(t)\rangle_I = U(t,t_0)|\psi(t_0)\rangle_I\,, \tag{3.9}$$

where $U(t,t_0)$ is a *unitary evolution operator* such that $U(t_1,t_2)U(t_2,t_3) = U(t_1,t_3)$ and $U(t,t) = 1$. So we get that

$$i\frac{dU}{dt} = H_I(t)U\,, \tag{3.10}$$

that, if H_I was a function, would have a simple solution

$$U(t,t_0) =^? exp\left(-i\int_{t_0}^{t} H_I(t')dt'\right). \tag{3.11}$$

However, H_I is an operator and we have the usual ordering problems. Let's see why. The exponential of an operator is defined as

$$\exp\left(-i\int_{t_0}^{t} H_I(t')dt'\right) = 1 - i\int_{t_0}^{t} H_I(t')dt' + \frac{(-i)^2}{2}\left(\int_{t_0}^{t} H_I(t')dt'\right)^2 + \dots\,, \tag{3.12}$$

and, when we try to differentiate with respect to t, we find that the quadratic term gives

$$-\frac{1}{2}\left(\int_{t_0}^{t} H_I(t')dt'\right)H_I(t) - \frac{1}{2}H_I(t)\left(\int_{t_0}^{t} H_I(t')dt'\right). \tag{3.13}$$

Now, we need something like the second term but we also have the first term and we cannot commute H_I with itself at different times. What do we do then? **Dyson's Formula:**

$$U(t,t_0) = T\left[\exp\left(-i\int_{t_0}^{t} H_I(t')dt'\right)\right] \tag{3.14}$$

This can also be expanded, to get

$$U(t,t_0) = 1 - i\int_{t_0}^{t} dt' H_I(t') + \frac{(-i)^2}{2}\left[\int_{t_0}^{t} dt' \int_{t'}^{t} dt'' H_I(t'')H_I(t') + \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' H_I(t')H_I(t'')\right] + \dots$$

which can be further simplified noting that the last two terms are exactly the same thing, i.e.

$$U(t,t_0) = 1 - i\int_{t_0}^{t} dt' H_I(t') + (-i)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' H_I(t') H_I(t'') + \dots \quad (3.15)$$

Proof: First note that, under $T[]$ sign, all operators commute (since it is taking care of putting them in the correct order). So

$$\begin{aligned} i\frac{\partial}{\partial t} T\left[\exp\left(-i\int_{t_0}^{t} dt' H_I(t')\right)\right] &= T\left[H_I(t)\exp\left(-i\int_{t_0}^{t} dt' H_I(t')\right)\right] \\ &= H_I(t) T\left[\exp\left(-i\int_{t_0}^{t} dt' H_I(t')\right)\right], \end{aligned}$$

where in the last step we used the fact that t is surely the latest time, so we can put $H_I(t)$ on the left side.

3.2 A First look at scattering

Let's now take as an example the Yukawa theory and let's apply to it the interaction picture,

$$H_{\text{int}} = g\int d^3\psi^\dagger\psi\phi\,. \quad (3.16)$$

Now that the theory is interacting, the particle number is not conserved anymore. In fact, the evolution of a state is given by $|\psi(t)\rangle = U(t,t_0)|\psi(t_0)\rangle$ where $U(t,t_0)$ is given by the Dyson's formula and it thus contain H_{int}. Now, H_{int} contains creation and annihilation operators for each type of particle:

- $\phi \sim a + a^\dagger$: it can create or destroy ϕ particles (we call them *mesons*.
- $\psi \sim b + c^\dagger$: it can destroy ψ particles and create anti-particles $\bar{\psi}$ (we call them *nucleons*[2]).
- $\psi^\dagger \sim b^\dagger + c$: it can create particles and destroy antiparticles.

Note that $Q = N_c - N_b$ is still conserved even if N_c and N_b are not.

We want then to compute some amplitudes with our theory but first we have to make an important assumption: **Initial and final states are eigenstates of the free theory.** In other words, the initial state $|i\rangle$ at $t \to -\infty$ and the final state $|f\rangle$ for $t \to \infty$ are eigenstate of the free Hamiltonian H_0. Therefore, the amplitude to go from $|i\rangle$ to $|f\rangle$ is

$$\lim_{t_\pm \longrightarrow \pm\infty} \langle f|U(t_+,t_-)|i\rangle = \langle f|S|i\rangle\,, \quad (3.17)$$

[2]Note that actually nucleons are spin 1/2 particles and do not arise from quantization of a scalar field. We are using this as a Toy model.

where we implicitly defined the unitary *S-matrix*.

This assumption seems to make sense but there are a couple of points that are at least weird:

- We cannot describe *bound states*. For example, an electron and a proton that collide, bind and then leave as an Hydrogen atom cannot be described by this picture[3].
- In a QFT interacting theory, a single particle is never really alone. We will come back on this point when we will talk about *renormalization*.

Example: Meson Decay. Consider the following initial and final states,

$$\begin{aligned}|i\rangle &= \sqrt{2E_{\vec{p}}}a^\dagger_{\vec{p}}|0\rangle \\ |f\rangle &= \sqrt{4E_{\vec{q}_1}E_{\vec{q}_2}}b^\dagger_{\vec{q}_1}c^\dagger_{\vec{q}_2}|0\rangle\,,\end{aligned}$$

where we have a single meson of momentum p in the initial state and a nucleon anti-nucleon pair with momenta q_1 and q_2. To leading order (LO) in g, the amplitude of passing from $|i\rangle$ to $|f\rangle$ is

$$\langle f|S|i\rangle = -ig\langle f|\int d^4x\psi^\dagger(x)\psi(x)\phi(x)|i\rangle\,. \tag{3.18}$$

Now, of all the contributions proportional to creation and annihilation operators that are inside the definition of the fields, only a few (or even just one) contribute to this amplitude. For example, the part of ϕ proportional to $a^\dagger$, once used on $|i\rangle$, would create a two mesons state which have no way to contribute to this amplitude given that $\psi^{(\dagger)}$ have not way of annihilating mesons. So we are left with

$$\begin{aligned}\langle f|S|i\rangle &= -ig\langle f|\int d^4x\psi^\dagger(x)\psi(x)\int\frac{d^3k}{(2\pi)^3}\frac{\sqrt{2E_{\vec{p}}}}{\sqrt{2E_{\vec{k}}}}a_{\vec{k}}a^\dagger_{\vec{p}}e^{-ik\cdot x}|0\rangle \\ &= -ig\langle f|\int d^4x\psi^\dagger(x)\psi(x)e^{-ip\cdot x}|0\rangle\,,\end{aligned}$$

where we commuted $a_{\vec{k}}$ past $a^\dagger_{\vec{p}}$ and we used the $\delta^{(3)}(\vec{p}-\vec{k})$ to perform the d^3k integral.

We now need to do the same for ψ and $\psi^\dagger$, keeping again only the combination that can give a non-zero overlap to $|f\rangle$. So we get

$$\begin{aligned}\langle f|S|i\rangle &= -ig\langle 0|\int\int\frac{d^4xd^3k_1d^3k_2}{(2\pi)^6}\frac{\sqrt{E_{\vec{q}_1}E_{\vec{q}_2}}}{\sqrt{E_{\vec{k}_1}E_{\vec{k}_2}}}c_{\vec{q}_2}b_{\vec{q}_1}c^\dagger_{\vec{k}_1}b^\dagger_{\vec{k}_2}|0\rangle e^{i(k_1+k_2-p)\cdot x} \\ &= -ig(2\pi)^4\delta^{(4)}(q_1+q_2-p)\,,\end{aligned}$$

[3]However, it is possible to solve this problem as the bound states appear as poles in the S-matrix.

that is our final amplitude.

Notice that the delta function constraints the kinematic of the decay, in particular the decay cannot happen if $m \geq 2M$. In fact, we can move to a reference frame where $p = (m, 0, 0, 0)$ and then the delta function would tell us that $\vec{q}_1 = -\vec{q}_2$ and that $m = 2\sqrt{M^2 + |\vec{q}|^2}$.

How to go from an amplitude to a more physical quantity will be explained later on.

3.3 Wick's Theorem

From Dyson's formula we would like to compute things like $\langle f|T[H_I(x_1)\dots H_I(x_n)]|i\rangle$, where the ordering of the operators is fixed by time-ordering. However, calculation would be much easier if we could move all annihilation operator to the right, where they can kill things in $|i\rangle$. In other words we would like to have those quantity in normal ordering. Wick's theorem tells us how to go from time ordered products to normal ordered products.

Definition: We define the *contraction* of a pair of fields in a string of operators $\dots\phi(x_1)\dots\phi(x_2)\dots$ to mean replacing those operators with the Feynman propagator, leaving all the rest untouched. We use the notation

$$\dots\overbrace{\phi(x_1)\dots\phi(x_2)}\dots \tag{3.19}$$

to denote contraction. The simple case is then

$$\overbrace{\phi(x)\phi(y)} = \Delta_F(x-y)\,. \tag{3.20}$$

The same is then true for complex scalar fields, as

$$\overbrace{\psi(x)\psi^\dagger(y)} = \Delta_F(x-y) \quad \text{and} \quad \overbrace{\psi(x)\psi(y)} = \overbrace{\psi^\dagger(x)\psi^\dagger(y)} = 0\,. \tag{3.21}$$

Wick's theorem: For any collection of fields we have

$$T[\phi_1\dots\phi_n] =: \phi_1\dots\phi_n : + : \text{all possible contractions} : \,. \tag{3.22}$$

Before going to the proof, let's see an example of application

$$T[\phi_1\phi_2\phi_3\phi_4] =: \phi_1\phi_2\phi_3\phi_4 : + \overbrace{\phi_1\phi_2} : \phi_3\phi_4 : + \dots + \overbrace{\phi_1\phi_2}\,\overbrace{\phi_3\phi_4} + \dots \tag{3.23}$$

Proof: It is easy to show for $n = 2$ so let's proceed by induction: suppose it is true for $\phi_2\dots\phi_n$ and let's try to show it is still true once ϕ_1 is added. If we take $x_1^0 > x_k^0$ for all $k = 2, \dots, n$, we can put ϕ_1 to the left of the time ordered product as[4]

$$T[\phi_1\phi_2\dots\phi_n] = (\phi_1^+ + \phi_1^-)(: \phi_2\dots\phi_n : + : \text{contractions} :)\,. \tag{3.24}$$

[4]Note that $\phi^\pm(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}a_{\vec{p}}^{(\dagger)}e^{\mp ip\cdot x}$.

Now, ϕ_1^- is proportional to $a^\dagger$ so it is already normal ordered (it is on the right of ϕ_1^+). However, we need to move ϕ_1^+ on the right of all the ϕ_k^- operators. Each time it moves past one we pick up a factor $\Delta_F(x_1 - x_k)$ from the commutator and it is exactly equal to $\overbrace{\phi_1\phi_k}$.□

Example: Nucleon scattering. Let's look at $\psi\psi \to \psi\psi$ scattering. We have

$$
\begin{aligned}
|i\rangle &= \sqrt{2E_{\vec{p}_1}}\sqrt{2E_{\vec{p}_2}}b^\dagger_{\vec{p}_1}b^\dagger_{\vec{p}_2}|0\rangle \equiv |p_1, p_2\rangle \\
|f\rangle &= \sqrt{2E_{\vec{p}'_1}}\sqrt{2E_{\vec{p}'_2}}b^\dagger_{\vec{p}'_1}b^\dagger_{\vec{p}'_2}|0\rangle \equiv |p'_1, p'_2\rangle \,.
\end{aligned}
$$

We then need to compute $\langle f|S - 1|i\rangle$ (since we are not interested in the case where no scattering happens). At order g^2 we have

$$\frac{(-ig)^2}{2}\int d^4x_1 d^4x_2 T[\psi^\dagger(x_1)\psi(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2)\phi(x_2)]\,, \tag{3.25}$$

on which we can use Wick's theorem to extract the piece

$$:\psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2):\overbrace{\phi(x_1)\phi(x_2)}\,, \tag{3.26}$$

that is the only one contributing to our scattering process. So we have

$$
\begin{aligned}
&\langle p'_1, p'_2|:\psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2):|p_1, p_2\rangle \\
=&\langle p'_1, p'_2|\psi^\dagger(x_1)\psi^\dagger(x_2)|0\rangle\langle 0|\psi(x_1)\psi(x_2)|p_1, p_2\rangle \\
=&\left(e^{ip'_1x_1+ip'_2x_2} + e^{ip'_1x_2+ip'_2x_1}\right)\left(e^{-ip_1x_1-ip_2x_2} + e^{-ip_1x_2-ip_2x_1}\right) \\
=&e^{ix_1(p'_1-p_1)+ix_2(p'_2-p_2)} + e^{ix_1(p'_2-p_1)+ix_2(p'_1-p_2)} + (x_1 \leftrightarrow x_2)\,,
\end{aligned} \tag{3.27}
$$

where we used the fact that

$$\langle 0|\psi(x)|p\rangle = e^{-ipx}\,. \tag{3.28}$$

So, at order g^2 we get

$$\langle f|S|i\rangle = \frac{(-ig)^2}{2}\int d^4x_1 d^4x_2[(3.27)]\int \frac{d^4x}{(2\pi)^4}\frac{ie^{ik(x_1-x_2)}}{k^2 - m^2 + i\epsilon}\,, \tag{3.29}$$

where the integral over k comes from the propagator coming from the ϕ contraction. Given that the x_1, x_2 dependence is symmetric, the term $(x_1 \leftrightarrow x_2)$ double up with the other and cancel the factor 1/2. Also, the integrals over x_1 and x_2 give delta functions, which lead us to

$$i(-ig)^2\left[\frac{1}{(p_1 - p'_1)^2 - m^2 + i\epsilon} + \frac{1}{(p_1 - p'_2)^2 - m^2 + i\epsilon}\right](2\pi)^4\delta^{(4)}(p_1 + p_2 - p'_1 - p'_2)\,. \tag{3.30}$$

In this case the ϵ prescription is not really needed since the denominator is never zero. In fact, in the center of mass frame $\vec{p}_1 = -\vec{p}_2$ we have $|\vec{p}_1| = |\vec{p'}_1$. This means that $k = (0, \vec{p} - \vec{p'})$ so that $k^2 < 0$.

We will see how to recover the same result with the Feynman diagrams.

3.4 Feynman Diagrams

As we saw, computing amplitudes using Wick's theorem is a long process. A much simpler way is to use *Feynman diagrams*, which requires us to write pictures and then associate numbers (or integrals) to them.

Given that the actual term we are interested in is $\langle f|S-1|i\rangle$, we can represent the various terms in the perturbative expansion as[5]

- Draw an external line for each particle in the initial or final states and associate to them a momentum. In this case we choose dotted lines for mesons and solid lines for nucleons. Also, use an arrow to describe its charge. For initial state we choose incoming arrow for ψ and outgoing for $\bar{\psi}$. The opposite is done for final state.

- Join the lines trough vertices, each of them is given by the interaction terms in the Lagrangian.

Each of the diagrams written in this way represents a term in the expansion of $\langle f|S-1|i\rangle$.

Feynman rules. After having drawn the diagrams, one needs to associate a number to them. For that we need the *Feynman rules*:

- Add a momentum k to each *internal line*.

- To each *vertex*, write down a factor[6]

$$(-ig)(2\pi)^4\delta^{(4)}(\sum_i k_i) \tag{3.31}$$

- For each internal dotted (meson) line we write a factor

$$\int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2-m^2+i\epsilon} \tag{3.32}$$

  Same for solid (nucleon) lines in which however we use the nucleon mass M in place of m.

Let's see a few examples:

Nucleon scattering with Feynman Diagrams. Let's look again at the $\psi\psi \to \psi\psi$ scattering. The diagrams contributing to this process are shown in figure 3.4. From them we can get, applying the Feynman rules,

[5] We are not doing this for the Lagrangian describing the mesons and nucleons, but this is general.

[6] This is of course dependent on the particular Lagrangian we choose.

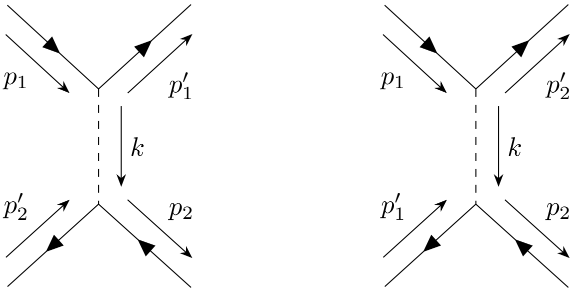


Figure 3.1: Feynman diagrams contributing to nucleon scattering at lowest order.

$$= i(-ig)^2 \left[\frac{1}{(p_1 - p_1')^2 - m^2} + \frac{1}{(p_1 - p_2')^2 - m^2} \right] (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_1' - p_2') \,, \tag{3.33}$$

which is exactly what we got by the Wick's theorem.

Note that in these diagrams there is a meson exchanging between the nucleons with a momentum $k = p_1 - p_1'$. This means that it does not satisfy the usual *energy dispersion relation* $k^2 = m^2$. For this reason, it is called a *virtual meson* and it is said to be *off-shell*. On the contrary, the initial and final state nucleons are *on-shell* (and they better be since they are particles that can be measured).

We only considered the lowest order in perturbation theory, but going up with the orders the number of diagrams grows.

Amplitudes. Having computed $\langle f|S-1|i\rangle$ at order g^2, we want to define the amplitude $\mathcal{A}_{fi}$ as

$$\langle f|S-1|i\rangle = i\mathcal{A}_{fi}\delta^{(4)}(p_F - p_I) \,, \tag{3.34}$$

where p_I and p_F are the sum of the initial and the final 4-momenta. What we usually compute with Feynman diagrams is actually the amplitude so, let's recast the Feynman rules for that

- Draw all possible diagrams with appropriate external legs and impose momentum conservation at each vertex.
- Write a factor $(-ig)$ for each vertex.
- For each internal line write a propagator.
- Integrate over each momentum k flowing trouch each loop, as $\int d^4k/(2\pi)^4$.

We never actually used the last point because our diagrams had no loops, they were *tree level* diagrams.

We can write down several examples but let's concentrate on one which is useful to explain a concept: Nucleon-Anti-Nucleon Scattering. In this case the diagrams are very similar to the previous ones and the final amplitude is

$$i\mathcal{A} = (-ig)^2 \left[\frac{i}{(p_1 - p_1')^2 - m^2} + \frac{i}{(p_1 + p_2)^2 - m^2 + i\epsilon} \right] . \tag{3.35}$$

The second term is quite different from before: in the centre of mass (**CdM**) frame its denominator is $4(M^2 + \vec{p}_1^2) - m^2$. If $m < 2M$ this term never vanishes (and in fact we can remove the $i\epsilon$) and we are back to a situation similar to the one of the other term. However, if $m > 2M$, there exists a value of $\vec{p}_1$ for which the denominator vanishes (and this is the reason we have the $i\epsilon$ term). In this case, however, the meson is unstable (it indeed can decay in two nucleons). This instability can be described by adding an immaginary term like $i\Lambda$ in the denominator (for which of course we can then remove the $i\epsilon$ term). In any case, the amplitude grows a lot when we approach the kinematic region $4(M^2 + \vec{p}_1)^2 = m^2$. This is called a **resonance** and it is what we use to discover particles.

3.4.1 Mandelstam Variables

There exists standard combination of momenta that appears frequently in our amplitudes. They are called *Mandelstam Variables* and they are defined as

$$\begin{aligned} s &= (p_1 + p_2)^2 = (p_1' + p_2')^2 \\ t &= (p_1 - p_1')^2 = (p_2 - p_2')^2 \\ u &= (p_1 - p_2')^2 = (p_2 - p_1')^2 . \end{aligned}$$

To get a feel for what these variables describe, let's put ourselves in the CdM in the case of 4 identical particles so that

$$p_1 = (E, 0, 0, p) \quad \text{and} \quad p_2 = (E, 0, 0, -p) , \tag{3.36}$$

and

$$p_1' = (E, 0, p \sin\theta, p\cos\theta) \quad \text{and} \quad p_2' = (E, 0, -p\sin\theta, -p\cos\theta) . \tag{3.37}$$

Then we get

$$s = 4E^2 \quad \text{and} \quad t = -2p^2(1 - \cos\theta) \quad \text{and} \quad u = -2p^2(1 + \cos\theta) , \tag{3.38}$$

which means that s measures the total center of mass energy while t and u measure the momentum exchanged by the two particles.

In the case of diagrams in which only one particle is exchanged, the amplitudes take simple forms. For example we say that an amplitude has contributions from the t and u channels if it has the form $\mathcal{A} \sim (t - m^2)^-1 + (u - m^2)^-1$.

Note also that there is always a relation between Mandelstam variables, $s + t + u = \sum_i M_i^2$.

ϕ^4 Theory. Let's now look at a different theory with interaction Hamiltonian

$$H_{\text{int}} = \frac{\lambda}{4!}\phi^4 \,. \tag{3.39}$$

The theory has now only one interaction vertex, while the other rules are the same. Note that the vertex rule is $(-i\lambda)$, without the 4!. To see why, let's look at the $\phi\phi \to \phi\phi$ scattering that has the lowest contribution of the form

$$\frac{-i\lambda}{4!}\langle p_1', p_2'| : \phi(x)\phi(x)\phi(x)\phi(x) : |p_1, p_2\rangle \,. \tag{3.40}$$

From Wick's theorem we easily see that anyone of the field can be the one annihilating or creating the particles, leaving us with 4! different contractions which cancel exactly the 1/4! factor.

Using Feynman rules, the amplitude for this process is just $\mathcal{A} = -i\lambda$. Note that it does not depend on the momenta at all: at leading order in this theory the two-particle scattering is completely symmetric.

3.4.2 Connected and Amputated Diagrams.

The rule of writing all possible Feynman diagrams have a couple of caveats. Both of them are related to the assumption that we made for which initial and final states are eigenstates of the free theory. Namely, they are

- We only consider fully connected diagrams, i.e. where every part of the diagram is connected to at least one external line. As we will see, this is related to the fact that the vacuum of the free theory $|0\rangle$ is not the same of the vaccum of the interacting theory Ω.

- We do not consider diagrams with loops on external lines. We will see that they enter our description in some sense but not while computing amplitudes. They are related to the fact that the one-particle states of the free theory are not the same of the ones in an interacting theory. In particular, keeping these diagrams into account will take care of the fact that interacting particles are never alone, but they are always surrounded by a *cloud* of virtual particles. The diagrams in which all loops on the external legs have been removed are called *amputated*.

3.4.3 Cross Sections and Decay Rates.

Now we are going to relate the amplitudes that we have learned to compute to measurable quantities, such as *cross-sections* and *decay rates*.

Fermi's Golden Rule. First of all, let's derive *Fermi's golden rule* from Dyson's formula. For two energy eigenstates $|m\rangle$ and $|n\rangle$, with $E_m \neq E_n$, we have, at leading order,

$$\begin{aligned}\langle m|U(t)|n\rangle &= -i\langle m|\int_0^t dt H_{\rm I}(t)|n\rangle \\ &= -i\langle m|H_{\rm int}|n\rangle\int_0^t dt' e^{i\omega t'} \\ &= -\langle m|H_{\rm int}|n\rangle\frac{e^{i\omega t}-1}{\omega}\,,\end{aligned}$$

where $\omega = E_m - E_n$. The probability of transition between n to m in time t is then[7]

$$\begin{aligned}P_{n\to m}(t) = |\langle m|U(t)|n\rangle|^2 &= 2|\langle m|H_{\rm int}|n\rangle|^2\left(\frac{1-\cos\omega t}{\omega^2}\right) \\ &\to 2\pi|\langle m|H_{\rm int}|n\rangle|^2\rho(E_n)t\,,\end{aligned}$$

which means that for states of similar energy, $E_n \sim E_m = E$, the probability is constant per unit time. This is *Fermi's golden rule.*

Let's suppose now we do not take the limit $t \to 0$, but we choose to compute the probability for the state $|n\rangle$ at $t- \to -\infty$ to transition to $|m\rangle$ at $t \to \infty$. In this case we get, for the amplitude

$$-i\langle m|\int_{-\infty}^{\infty} H_I(t)|n\rangle - -i\langle m|H_{\rm int}|n\rangle 2\pi\delta(\omega)\,, \tag{3.41}$$

which we need to square if we want to get the probability. However, we get now the square of a delta function. The reason for that is we are computing the probability for the transition to happen in an infinity time. Writing the square of the delta function as

$$(2\pi)^2\delta(\omega)^2 = (2\pi)\delta(\omega)T\,, \tag{3.42}$$

where T is a shorthand for $t \to \infty$. So we can get the probability per unit time by dividing by T, getting

$$\dot{P}_{n\to m} = 2\pi|\langle m|H_{\rm int}|n\rangle|^2\delta(\omega)\,. \tag{3.43}$$

The same happen for our QFT calculations, for which we will get square of $\delta^{(4)}-$functions which we will interpret as spacetime volume factors.

Decay Rates. Let's look at the probability for a single particle $|i\rangle$ of momentum p_I to decay into some particles $|f\rangle$ with momentum p_i and total momentum $p_F = \sum_i p_i$. This is

$$P = \frac{|\langle f|S|i\rangle|^2}{\langle f|f\rangle\langle i|i\rangle}\,, \tag{3.44}$$

[7]Note that we are taking the $t \to \infty$ limit where $\left(\frac{1-\cos\omega t}{\omega^2}\right) \to \pi t\delta(\omega)$.

where our states obey the normalization

$$\langle i|i\rangle = (2\pi)^3 2E_{\vec{p}_I}\delta^{(3)}(0) = 2E_{\vec{p}_I}V\,, \tag{3.45}$$

where we substituted $\delta^{(3)}(0)$ with V, and the same for the final states

$$\langle f|f\rangle = \prod_{\text{final states}} 2E_{\vec{p}_i}V\,. \tag{3.46}$$

If we put ourselves in the frame in which the initial particle is at rest, so that $\vec{p}_I = 0$ and $E_{\vec{p}_I} = m$, we get the probability of decay

$$P = \frac{|\mathcal{A}_{fi}|^2}{2mV}(2\pi)^4\delta^{(4)}(p_I - p_F)VT \prod_{\text{final states}} \frac{1}{2E_{\vec{p}_i}V}\,, \tag{3.47}$$

where again we exchanged one of the two delta functions for the spacetime volume VT. Of course now we can divide by T to get the transition function per unit time, but that is not enough because we have to care about summing over all final states. So first we need to integrate over all possible momenta of the final states, $V\int d^3p_i/(2\pi)^3$, which then cancel all the factors $1/V$ and which construct the Lorentz invariant measure for 3−momentum integrals, together with the factors $1/2E_{\vec{p}_i}$. If we now sum over all possible final particles, we finally get the decay probability per unit time

$$\Gamma = \frac{1}{2m}\sum_{\text{final states}}\int |\mathcal{A}_{fi}|^2 d\Pi\,, \tag{3.48}$$

where

$$d\Pi = (2\pi)^4\delta^{(4)}(p_F - p_I)\prod_{\text{final states}} \frac{d^3p_i}{(2\pi)^3}\frac{1}{2E_{\vec{p}_i}}\,. \tag{3.49}$$

Γ is called the width of the particle and it is equal to the reciprocal of the *half-life* $\tau = 1/\Gamma$.

Cross Sections. Let's think now about the case in which we are colliding two beams of particles. The fraction of the particles that actually collide is called *cross-section*. If we define the incoming flux F as the number of incoming particles per area per unit time, then the total number of scattering events per unit time is

$$N = F\sigma\,. \tag{3.50}$$

We would like to compute it. In fact, we can do more: we can compute the *differential cross-section* $d\sigma$ that is the probability for a given scattering process to occur (per unit time) in a solid angle (θ,ϕ). Namely,

$$d\sigma = \frac{\text{Differential probability}}{\text{Unit Time}\times\text{Unit Flux}} = \frac{1}{4E_1E_2V}\frac{1}{F}|\mathcal{A}_{fi}|^2 d\Pi\,, \tag{3.51}$$

where E_1 and E_2 are the energies of the incoming particles. We need now the expression for the unit flux. Let's put ourselves in the CdM frame of the collision. Then $F = |\vec{v}_1 - \vec{v}_2|/V$ so that

$$d\sigma = \frac{1}{4E_1E_2}\frac{1}{|\vec{v}_1 - \vec{v}_2|}|\mathcal{A}_{fi}|^2 d\Pi\,. \tag{3.52}$$

This is our final expression to link the amplitude we can compute in QFT to the differential cross-section, which we can measure.

3.5 Green's Functions

So far we computed scattering amplitudes, which are physical since they are linked to cross-sections and decay rates. For less physical quantities, we need to compute *correlation functions*. Let's see how to compute them.

We denote the true vacuum of the interacting theory as Ω and we normalize H such that

$$H|\Omega\rangle = 0\,, \tag{3.53}$$

and $\langle\Omega|\Omega\rangle = 1$. This is of course different from the vacuum of the free theory, for which $H_0|0\rangle = 0$. Define

$$G^{(n)}(x_1,\ldots,x_n) = \langle\Omega|T[\phi_H(x_1)\ldots\phi_H(x_n)]|\Omega\rangle\,, \tag{3.54}$$

where ϕ_H is ϕ in the Heisenberg picture of the full theory. These functions are called correlation functions, or *Green's functions*. Let's start asking ourselves how to compute them from Feynman diagrams.

Claim: Denote $\phi_{1\mathrm{H}}$ the field in the Heisenberg picture and $\phi_{1\mathrm{I}}$ the field in the interaction picture. Then

$$G^{(n)}(x_1,\ldots,x_n) = \langle\Omega|T[\phi_{1H}\ldots\phi_{nH}]|\Omega\rangle = \frac{\langle 0|T[\phi_{1I}\ldots\phi_{nI}S]|0\rangle}{\langle 0|S|0\rangle}\,. \tag{3.55}$$

Proof: We refer to books for the proof.

So, thanks to this claim, we are close to compute these correlations functions with Feynman diagrams, since we are able to compute both the numerator and denominator with Feynman diagrams. However, what about doing the ratio? What is the interpretation of that?

3.5.1 Connected Diagrams and Vacuum Bubbles

In order to answer these questions, let's take the ϕ^4 theory as an example. We represent the ϕ particles (the only kind of particles in the theory), by solid lines. Then we have the diagrammatic expansion of 3.5.1. The combinatorial factors of these diagrams make sure once can actually write them as an exponential, in such a way the amplitude for the vacuum of the free theory to evolve into itself is $\langle 0|S|0\rangle = \exp(\text{all distinct vacuum bubbles})$. Why is this relevant for

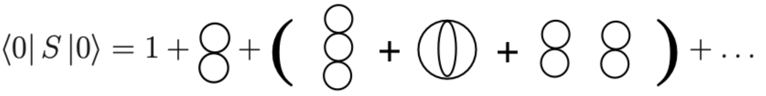


Figure 3.2: Vacuum bubbles diagrams.

the correlation functions? Because we have exactly the same exponential in the correlation functions computation, so that we can write

$$\langle 0|T[\phi_1 \dots \phi_n]S|0\rangle = \left(\sum \text{connected diagrams}\right)\langle 0|S|0\rangle\,. \tag{3.56}$$

This is then the reason of only considering connected diagrams when computing amplitudes: the ratio between $\langle 0|T[\phi_1 \dots \phi_n]S|0\rangle$ and $\langle 0|S|0\rangle$ only depends on connected diagrams. Coming back to the correlation functions, this means that

$$\langle \Omega|T[\phi_H(x_1)\dots\phi_H(x_n)]|\Omega\rangle = \sum \text{Connected diagrams}\,. \tag{3.57}$$

3.5.2 From Green's Functions to S-Matrices

Let's now related these Green's functions to the S-matrices. First let's do the Fourier transform

$$\tilde{G}^{(n)}(p_1,\dots,p_n) = \int \left[\prod_{i=1}^{n} d^4x_i e^{-ip_i x_i}\right] G^{(n)}(x_1,\dots,x_n)\,. \tag{3.58}$$

This is already close to the S-matrix we already computed, the difference being that their Feynman rules include propagators Δ_F for external legs, while the S-matrix ones don't. To solve this, it is enough to cancel these propagators and put their momenta back on shell, as

$$\begin{aligned}\langle p'_1,\dots,p'_{n'}|S-1|p_1,\dots,p_n\rangle =& (-i)^{n+n'}\prod_{i=1}^{n'}(p_i'^2-m^2)\prod_{j=1}^{n}(p_j^2-m^2)\\ &\times \tilde{G}^{(n+n')}(-p'_1,\dots,-p'_{n'},p_1,\dots,p_n)\,.\end{aligned} \tag{3.59}$$

So what did we understand? We understood that ignoring unconnected diagrams is related to shifting to the true vacuum of the interacting theory $|\Omega\rangle$. But then what is the point of introducing the Green's functions? The point is that it provides a framework in which one can deal with the true particle states in the interacting theory trough renormalization. In this way we are correctly treating scattering, considering also the swarm of virtual particles surrounding each asymptotic state. Equation (3.59) is known as the LSZ reductions formula and we will see it better next.

Chapter 4

The Dirac Equation

So far we discussed about fields such that under the Lorentz transformation $x^\mu \to (x')^\mu = \Lambda^\mu_\nu x^\nu$, transforms as

$$\phi(x) \to \phi'(x) = \phi(\Lambda^{-1}x)\,, \tag{4.1}$$

and we have seen that upon quantization, these fields give rise to particles of spin 0. In order to describe particles with an intrinsic angular momentum, or spin, it is necessary to consider fields that transforms non-trivially under Lorentz transformations.

An example of a field that does not transform trivially under Lorentz transformations, is the vector field $A_\mu(x)$, for which

$$A^\mu(x) \to \Lambda^\mu_\nu A^\nu(\Lambda^{-1}x)\,. \tag{4.2}$$

We will deal with *vector fields* later. More generally, a field can transform as

$$\phi^a(x) \to D[\Lambda]^a_b \phi^b(\Lambda^{-1}x)\,, \tag{4.3}$$

where the matrices $D[\Lambda]$ form a *representation* of the Lorentz group, i.e.

$$D[\Lambda_1]D[\Lambda_2] = D[\Lambda_1\Lambda_2] \tag{4.4}$$

and $D[\Lambda^{-1}] = D[\Lambda]^{-1}$ and $D[1] = 1$. How can we find the different representations? One can look at infinitesimal transformations and thus study the *Lie Algebra*. In practice, as we have already seen,

$$\Lambda^\mu_\nu = \delta^\mu_\nu + \omega^\mu_\nu \tag{4.5}$$

for infinitesimal ω, implies that ω is anti-symmetric,

$$\omega^{\mu\nu} + \omega^{\nu\mu} = 0\,, \tag{4.6}$$

which gives a total of 6 independent components. It is then useful to choose a basis for this 6 matrices. It has been proven convenient to write this basis as $(\mathcal{M}^{\rho\sigma})^{\mu\nu}$, where $[\rho\sigma]$ run from 0 to 3 and are anti-symmetric (so that for example $\mathcal{M}^{01} = -\mathcal{M}^{10}$). Then of course $\mathcal{M}$ is also anti-symmetric on the $[\mu\nu]$ indices because they are themselves anti-symmetric matrices. With this notation, we can write our basis as

$$(\mathcal{M}^{\rho\sigma})^{\mu\nu} = \eta^{\rho\mu}\eta^{\sigma\nu} - \eta^{\sigma\mu}\eta^{\rho\nu}\,, \tag{4.7}$$

where μ and ν are the indices of the 4×4 matrix, while ρ and σ denote which basis element we are considering. Note that, for practical reason, we will typically lower one index, getting

$$(\mathcal{M}^{\rho\sigma})^{\mu}_{\ \nu} = \eta^{\rho\mu}\delta^{\sigma}_{\nu} - \eta^{\sigma\mu}\delta^{\rho}_{\nu}\,, \tag{4.8}$$

which of course are no longer necessarily anti-symmetric in the $[\mu\nu]$ indices. Two example of these matrices are

$$(\mathcal{M}^{01})^{\mu}_{\ \nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad (\mathcal{M}^{12})^{\mu}_{\ \nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.9}$$

where the first generates a boost in the x^1 direction and the second generates a rotation in the $(x^1, x^2)-$plane.

Given that we chose a basis, we can then write ω as a linear combination, as

$$\omega^{\mu}_{\ \nu} = \frac{1}{2}\Omega_{\rho\sigma}(\mathcal{M}^{\rho\sigma})^{\mu}_{\ \nu}\,, \tag{4.10}$$

where $\Omega_{\rho\sigma}$ are just 6 numbers. The six basis matrices, $\mathcal{M}$, are called the *generators* of the Lorentz transformations, and they follow the Lie Algebra

$$[\mathcal{M}^{\rho\sigma}, \mathcal{M}^{\tau\nu}] = \eta^{\sigma\tau}\mathcal{M}^{\rho\nu} - \eta^{\rho\tau}\mathcal{M}^{\sigma\nu} + \eta^{\rho\nu}\mathcal{M}^{\sigma\tau} - \eta^{\sigma\nu}\mathcal{M}^{\rho\tau}\,. \tag{4.11}$$

A finite Lorentz transformation can be then expressed by the exponential

$$\Lambda = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma}\right). \tag{4.12}$$

Once we have the Lie Algebra, we can ask ourselves if we can find other matrices that satisfy it.

4.1 The Spinor Representation

One possibility is the *spinor representation.* First, let's define the *Clifford algebra*,

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}1\,, \tag{4.13}$$

where γ^μ with $\mu = 0, 1, 2, 3$ is a set of four matrices and 1 denotes the unit matrix. In order to satisfy the Clifford algebra, we must have

$$\gamma^\mu\gamma^\nu = -\gamma^\nu\gamma^\mu \quad \text{when} \mu \neq \nu \tag{4.14}$$

and

$$(\gamma^0)^2 = 1 \quad , \quad (\gamma^i)^2 = -1 \quad i = 1, 2, 3 \,. \tag{4.15}$$

It is not hard to see that the simplest representation of the Clifford algebra can be found for 4×4 matrices (no 2×2 and 3×3 matrices can satisfy this algebra). One possibility is

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \,, \tag{4.16}$$

where each element is a 2×2 matrix, in particular σ^i are the Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \,, \tag{4.17}$$

for which $\{\sigma^i, \sigma^j\} = 2\delta^{ij}$. One can show that there is only one *irreducible* representation of the Clifford algebra. The representation we have just shown is one example, and it called *Weyl* or *chiral* representation. Now, it is possible to construct other representations simply taking $V\gamma^\mu V^{-1}$ with V invertible matrix, but we will soon restrict to the case in which V is unitary.

So, what does the Clifford algebra have to do with Lorentz transformations? Consider the commutator of two γs,

$$S^{\rho\sigma} = \frac{1}{4}[\gamma^\rho, \gamma^\sigma] = \begin{Bmatrix} 0 & \rho = \sigma \\ \frac{1}{2}\gamma^\rho\gamma^\sigma & \rho \neq \sigma \end{Bmatrix} = \frac{1}{2}\gamma^\rho\gamma^\sigma - \frac{1}{2}\eta^{\rho\sigma} \,. \tag{4.18}$$

The claim now is that $S^{\rho\sigma}$ are a representation of $\mathcal{M}^{\rho\sigma}$. To show this we have to prove equation (4.11), so first let's compute $[S^{\mu\nu}, \gamma^\rho]$, since it will be needed.

When $\mu \neq \nu$ we have

$$\begin{aligned}
[S^{\mu\nu}] &= \frac{1}{2}[\gamma^\mu\gamma^\nu, \gamma^\rho] \\
&= \frac{1}{2}\gamma^\mu\gamma^\nu\gamma^\rho - \frac{1}{2}\gamma^\rho\gamma^\mu\gamma^\nu \\
&= \frac{1}{2}\gamma^\mu\{\gamma^\nu, \gamma^\rho\} - \frac{1}{2}\gamma^\mu\gamma^\rho\gamma^\nu - \frac{1}{2}\{\gamma^\rho, \gamma^\mu\}\gamma^\nu + \frac{1}{2}\gamma^\mu\gamma^\rho\gamma^\nu \\
&= \gamma^\mu\eta^{\nu\rho} - \gamma^\nu\eta^{\rho\mu} \,.
\end{aligned}$$

Now we can compute what we need, namely, for $\rho \neq \sigma$,

$$\begin{aligned}
[S^{\mu\nu}, S^{\rho\sigma}] &= \frac{1}{2}[S^{\mu\nu}, \gamma^\rho\gamma^\sigma] \\
&= \frac{1}{2}[S^{\mu\nu}, \gamma^\rho]\gamma^\sigma + \frac{1}{2}\gamma^\rho[S^{\mu\nu}, \gamma^\sigma] \\
&= \frac{1}{2}\gamma^\mu\gamma^\sigma\eta^{\nu\rho} - \frac{1}{2}\gamma^\nu\gamma^\sigma\eta^{\rho\mu} + \frac{1}{2}\gamma^\rho\gamma^\mu\eta^{\nu\sigma} - \frac{1}{2}\gamma^\rho\gamma^\nu\eta^{\sigma\mu} \,,
\end{aligned}$$

from which we can use $\gamma^\mu\gamma^\sigma = 2S^{\mu\sigma} + \eta^{\mu\sigma}$ to get what we wanted to show,

$$[S^{\mu\nu}, S^{\rho\sigma}] = S^{\mu\sigma}\eta^{\nu\rho} - S^{\nu\sigma}\eta^{\rho\mu} + S^{\rho\mu}\eta^{\nu\sigma} - S^{\rho\nu}\eta^{\sigma\mu} \,. \tag{4.19}$$

4.1.1 Spinors

We have then shown that the $S^{\mu\nu}$ are a representation of the Lorentz group. From now on we will use α and β to denote the matrix entries (they are 4×4 matrices after all). So, if there is a representation, there must a field that transform following such representation. We then introduce the Dirac *spinor* field $\psi^\alpha(x)$, an object with 4 complex components labelled by $alpha = 1, 2, 3, 4$. Under Lorentz transformation we then have

$$\psi^\alpha(x) \to S[\Lambda]^\alpha_\beta \psi^\beta(\Lambda^{-1}x) \,, \tag{4.20}$$

where

$$\begin{aligned} \Lambda &= \exp\left(\frac{1}{2}\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma}\right) \\ S[\Lambda] &= \exp\left(\frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right) . \end{aligned}$$

Note that although the basis of generators is different, we use the same numbers Ω to ensure we are actually doing the same Lorentz transformation on x and on ψ^α.

Now, both Λ and $S[\Lambda]$ are 4×4 matrices so, how can we be sure that we are actually doing something new? Let's then look at specific cases

Rotations. For $i \neq j$ and $i, j = 1, 2, 3$[1],

$$S^{ij} = \frac{1}{2}\begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix} = -\frac{i}{2}\epsilon^{ijk}\begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix} \tag{4.21}$$

So, if we write the rotation parameters as $\Omega_{ij} = -\epsilon_{ijk}\varphi^k$, then we get[2]

$$S[\Lambda] = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right) = \begin{pmatrix} e^{i\vec{\varphi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{i\vec{\varphi}\cdot\vec{\sigma}/2} \end{pmatrix} . \tag{4.22}$$

Consider now a rotation of 2π about, say, the x^3 axis. In this case $\vec{\varphi} = (0, 0, 2\pi)$, and we get a spinor rotation

$$S[\Lambda] = \begin{pmatrix} e^{i\pi\sigma^3} & 0 \\ 0 & e^{i\pi\sigma^3} \end{pmatrix} = -1 \,, \tag{4.23}$$

[1] Here we are using the property of the Pauli matrices $[\sigma^i, \sigma^j] = 2i\epsilon^{ijk}\sigma^k$. Since $\sigma^i\sigma^j = \frac{1}{2}([\sigma^i, \sigma^j] + \{\sigma^i, \sigma^j\})$, one can get $\sigma^i\sigma^j = \frac{1}{2}(2i\epsilon^{ijk}\sigma^k + 2\delta^{ij}) = i\epsilon^{ijk}\sigma^k + \delta^{ij}$.

[2] Here we are using the property $\epsilon^{ijk}\epsilon^{ijn} = 2\delta^{kn}$.

which means that

$$\psi^\alpha \to -\psi^\alpha(x)\,. \tag{4.24}$$

This is definitely not what happens to a vector! A vector, in fact, would transform following Λ that is, in this case,

$$\Lambda = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma}\right) = \exp\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \varphi^3 & 0 \\ 0 & -\varphi^3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.25}$$

so that if $\varphi^3 = 2\pi$, $\Lambda = 1$ and the vector is unchanged, as expected. So, $S[\Lambda]$ is definitely a different representation of the Lorentz group.

Boosts. Doing the same for a boost, we get

$$S^{0i} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma^i \end{pmatrix} \tag{4.26}$$

Writing the boost parameters as $\Omega_{i0} = -\Omega_{0i} = \chi_i$, we have

$$S[\Lambda] = \begin{pmatrix} e^{\vec{\chi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{-\vec{\chi}\cdot\vec{\sigma}/2} \end{pmatrix}\,. \tag{4.27}$$

4.2 Constructing an action

We have now a new field, the Dirac spinor ψ. Let's now construct a Lorentz invariant action. We start with a naive way which will not work but will give us hints on how to proceed. Define

$$\psi^\dagger(x) = (\psi^*)^T(x)\,, \tag{4.28}$$

that is the usual *adjoint* of a multi-component object. A Lorentz scalar could then be $\psi^\dagger\psi$, so let's see how it transforms.

$$\psi(x) \to S[\Lambda]\psi(\Lambda^{-1}x)$$
$$\psi^\dagger(x) \to \psi^\dagger(\Lambda^{-1}x)S[\Lambda]^\dagger$$

So we get $\psi^\dagger(x)\psi(x) \to \psi^\dagger(\Lambda^{-1}x)S[\Lambda]^\dagger S[\Lambda]\psi(\Lambda^{-1}x)$. However, for some Lorentz transformations (as for example the boost we have seen before, $S[\Lambda]$ is not *unitary* and so $S[\Lambda]^\dagger S[\Lambda] \neq 1$. This means that $\psi^\dagger(x)\psi(x)$ was not a good choice. However now we know how to proceed: let's take a representation of the Clifford algebra that, like the spinor representation, satisfies $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^i)^\dagger = -\gamma^i$. Then we have

$$\gamma^0\gamma^\mu\gamma^0 = (\gamma^\mu)^\dagger\,, \tag{4.29}$$

which means that

$$(S^{\mu\nu})^\dagger = \frac{1}{4}[(\gamma^\nu)^\dagger, (\gamma^\mu)^\dagger] = -\gamma^0 S^{\mu\nu}\gamma^0 \,, \tag{4.30}$$

so that we get

$$S[\Lambda]^\dagger = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}(S^{\rho\sigma})^\dagger\right) = \gamma^0 S[\Lambda]^{-1}\gamma^0 \,. \tag{4.31}$$

It is then clear that we can define the *Dirac adjoint*

$$\bar{\psi}(x) = \psi^\dagger(x)\gamma^0 \,, \tag{4.32}$$

with which we can construct a Lorentz invariant object and several Lorentz *covariant* objects. First $\bar{\psi}\psi$ is Lorentz invariant, in fact

$$\begin{aligned}\bar{\psi}(x)\psi(x) &= \psi^\dagger(x)\gamma^0\psi(x) \\ &\to \psi^\dagger(\Lambda^{-1}x)S[\Lambda]^\dagger\gamma^0 S[\Lambda]\psi(\Lambda^{-1}x) \\ &= \psi^\dagger(\Lambda^{-1}x)\gamma^0\psi(\Lambda^{-1}x) \\ &= \bar{\psi}(\Lambda^- 1x)\psi(\Lambda^- 1x) \,.\end{aligned}$$

Also, $\bar{\psi}\gamma^\mu\psi$ is a Lorentz vector, which means that

$$\bar{\psi}\gamma^\mu\psi \to \Lambda^\mu_\nu\bar{\psi}\gamma^\nu\psi \,, \tag{4.33}$$

and $\bar{\psi}\gamma^\mu\gamma^\nu\psi$ behaves as a Lorentz *tensor*.

Having now these new quantities we can construct the Lorentz invariant *Dirac action*

$$S = \int d^4x\bar{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x) \,. \tag{4.34}$$

We will see that this action, after quantization, describes particles and anti-particles of mass m and spin 1/2.

4.3 The Dirac Equation

From the action we can get as usual the equations of motions, varying with respect to ψ and $\bar{\psi}$ independently. We get then (varying with respect to $\bar{\psi}$)

$$(i\gamma^\mu\partial_\mu - m)\psi = 0 \,, \tag{4.35}$$

that is the *Dirac equation*. Varying with respect to ψ, we get

$$i\partial_\mu\bar{\psi}\gamma^\mu + m\bar{\psi} = 0 \,. \tag{4.36}$$

The Dirac equation mixes up the different components of ψ by the matrices γ, but each of the individual components solves the KG equation. To see this, let's write

$$(i\gamma^\nu\partial_\nu + m)(i\gamma^\mu\partial_\mu - m)\psi = -(\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu + m^2)\psi \,. \tag{4.37}$$

But we know $\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu = 1/2\{\gamma^\mu,\gamma^\nu\}\partial_\mu\partial_\nu = \partial_\mu\partial^\mu$, so that we get

$$(\partial_\mu\partial^\mu + m^2)\psi = 0\,, \tag{4.38}$$

where this has no γ matrix, so it is applied to each component of ψ^α, with $\alpha = 1,2,3,4$.

The Slash. An useful notation, that we will use later on, is used when we contract 4−vectors with γ matrices. In particular, we denote

$$A_\mu\gamma^\mu \equiv \not{A}\,, \tag{4.39}$$

so that, for example, the Dirac equation reads

$$(i\not{\partial} - m)\psi = 0\,. \tag{4.40}$$

4.4 Chiral spinors

When we needed an explicit form of the γ matrices we used the chiral representation of equation (4.16). With these, we computed the spinor rotation transformation $S[\Lambda_{\text{rot}}]$ (eq. (4.22)) and boost transformation $S[\Lambda_{\text{boost}}]$ (eq. (4.27)). They both are *block diagonal* and this means that the Dirac spinor representation is *reducible*, i.e it can be decomposed into (two in this case) *irreducible* representations. These two irreducible representations act on two-component spinors $u_\pm$ which, in the chiral representation, are defined as

$$\psi = \begin{pmatrix} u_+ \\ u_- \end{pmatrix}\,. \tag{4.41}$$

They are called *Weyl spinors* or *chiral spinors.* They transforms in the same way under rotations,

$$u_\pm \to e^{i\vec{\varphi}\cdot\vec{\sigma}/2}u_\pm\,, \tag{4.42}$$

but oppositely under boosts,

$$u_\pm \to e^{\pm\vec{\chi}\cdot\vec{\sigma}/2}u_\pm\,. \tag{4.43}$$

In group theory language, u_+ is in the $(1/2,0)$ representation of the Lorentz group, while u_- is in the $(0,1/2)$, so that ψ is in the $(1/2,0)\oplus(0,1/2)$ representation.

4.4.1 The Weyl equation

We can then use the chiral spinors to write our Dirac Lagrangian. We have[3][4]

$$\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi = iu_-^\dagger \sigma^\mu \partial_\mu u_- + iu_+^\dagger \bar{\sigma}^\mu \partial_\mu u_+ - m(u_+^\dagger u_- + u_-^\dagger u_+) = 0\,, \quad (4.44)$$

where we denoted

$$\sigma^\mu = (1, \sigma^i) \quad \text{and} \quad \bar{\sigma}^\mu = (1, -\sigma^i) \quad (4.45)$$

the two 2×2 sectors of the γ matrices. From the Lagrangian now we can see that a massive fermion needs both u_+ and u_- since they couple trough the massive term. However, a massless fermion can be described even by only of the two, with equation of motion

$$i\bar{\sigma}^\mu \partial_\mu u_+ = 0 \quad \text{or} \quad i\sigma^\mu \partial_\mu u_- = 0\,, \quad (4.46)$$

called *Weyl equations*.

Degrees of freedom. The Dirac Fermion has 4 complex componenets = 8 real components. Does this mean that it has 8 degrees of freedom? Keep in mind that a real scalar field has only one degree of freedom, which results in being able to produce only a single type of particle, and a complex scalar field has two degrees of freedom, and in fact it can produce particle and anti-particle. To answer this question, let's look at the momentum conjugate of a Dirac field,

$$\pi_\psi = \partial\mathcal{L}/\partial\dot{\psi} = i\psi^\dagger\,. \quad (4.47)$$

It is not proportional to the time derivative of ψ, which is a direct consequence of the fact that the Dirac equation is linear, as opposed to the KG equation that is quadratic. This means that the phase space for a spinor is parameterized by ψ and $\psi^\dagger$, which means it has 8 real dimensions and correspondingly 4 degrees of freedom. After quantization, this corresponds to two degrees of freedom (spin up and down) for the particle and the other two for the anti-particle. Of course, the Weyl fermion has two degrees of freedom.

4.4.2 The γ^5

The $S[\Lambda]$ came out to be diagonal because of our choice of the chiral representation. Let's see what happens if we choose a different representation γ^μ of the Clifford algebra, so that

$$\gamma^\mu \to U\gamma^\mu U^{-1} \quad \text{and} \quad \psi \to U\psi\,. \quad (4.48)$$

[3] Note that $\bar{\psi}$ is $\begin{pmatrix} u_+^\dagger & u_-^\dagger \end{pmatrix} \gamma^0 = \begin{pmatrix} u_-^\dagger & u_+^\dagger \end{pmatrix}$.

[4] Note also that the γ matrices in the spinor representation are block anti-diagonal, so that they are diagonal on the antidiagonal but they are zero on the diagonal. This means that, when dividing them in two blocks corresponding to the Weyl representation, only the $(1, \sigma_i)$ (corresponding to (γ^0, γ^i)) will contribute to the lower 2×2 part and viceversa. Indeed the fact that they are antidiagonal make sure that their upper 2×2 part will act on the lower 2×2 part of the object they are acting on.

Now $S[\Lambda]$ is not block diagonal anymore. Is there an invariant way to define the chiral spinors? We need to introduce another γ matrix,

$$\gamma^5 = -i\gamma^0\gamma^1\gamma^2\gamma^3 \,, \tag{4.49}$$

which satisfies

$$\{\gamma^5, \gamma^\mu\} = 0 \quad \textit{and} \quad (\gamma^5)^2 = 1 \,. \tag{4.50}$$

The set $\tilde{\gamma}^A = (\gamma^\mu, i\gamma^5)$ satisfies the five-dimensional Clifford algebra $\{\tilde{\gamma}^A, \tilde{\gamma}^B\} = 2\eta^{AB}$, and $[S_{\mu\nu}, \gamma^5] = 0$, which means that γ^5 is a Lorentz scalar. We can then use it to construct Lorentz invariant projector

$$P_\pm = \frac{1}{2}(1 \pm \gamma^5) \,, \tag{4.51}$$

such that $P_+^2 = P_+$, $P_-^2 = P_-$ and $P_+ P_- = 0$ (so that they fulfill the needed projector relations). In the chiral representation, it takes the form

$$\gamma^5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \,, \tag{4.52}$$

which shows that indeed $P_\pm$ project respectively into $u_\pm$. However, we can use $P_\pm$ in an arbitrary representation to define the chiral spinors, as

$$\psi_\pm = P_\pm \psi \,, \tag{4.53}$$

which forms the two irreducible representations of the Lorentz group. Note that ψ_+ is called a *right-handed* spinor and ψ_- a *left-handed* spinor.

4.4.3 Parity

Right and Left handed spinors are related to each other by *parity*. Let's define this concept. Recall that the Lorentz group is defined in such a way

$$\Lambda^\mu_\nu \Lambda^\rho_\sigma \eta^{\nu\sigma} = \eta^{\mu\rho} \,. \tag{4.54}$$

So far we have only considered transformations that are continuously connected to identity, which is the reason why they have an infinitesimal form. However, there are also two *discrete* symmetries that are part of the Lorentz group, namely

$$\begin{aligned} \text{Time Reversal}\, T : x^0 \to -x^0 \,;\, x^i \to x^i \\ \text{Parity}\, P : x^0 \to x^0 \,;\, x^i \to -x^i \,. \end{aligned}$$

While we won't discuss time reversal (because of its limited relevance), parity has an important role in the *Standard Model* (SM).

Under parity, left and right handed spinors are exchanged, namely

$$P : \psi_\pm(\vec{x}, t) \to \psi_\mp(-\vec{x}, t) \,. \tag{4.55}$$

It is easy to show that the action of parity can be described by the γ^0 function, as

$$P : \psi(\vec{x}, t) \to \gamma^0 \psi(-\vec{x}, t) \,. \tag{4.56}$$

Let's then look at how our interactions change under parity and, in general, at how our bilinears change under parity. So we have

$$P : \bar{\psi}\psi(\vec{x}, t) \to \bar{\psi}\psi(-\vec{x}, t) \,, \tag{4.57}$$

which is a scalar, and we have (by components)

$$\begin{aligned} P &: \bar{\psi}\gamma^0\psi(\vec{x}, t) \to \bar{\psi}\gamma^0\psi(-\vec{x}, t) \\ P &: \bar{\psi}\gamma^i\psi(\vec{x}, t) \to \bar{\psi}\gamma^0\gamma^i\gamma^0\psi(-\vec{x}, t) = -\bar{\psi}\gamma^i\psi(-\vec{x}, t) \,, \end{aligned}$$

which is the transformation of a vector, with the spatial part flipping the sign. The same applies to $\bar{\psi}S^{\mu\nu}\psi$ which transforms as a tensor. So everything as expected.

However, now we can form others Lorentz scalar and vectors with our new γ^5,

$$\bar{\psi}\gamma^5\psi \quad \text{and} \quad \bar{\psi}\gamma^5\gamma^\mu\psi \,. \tag{4.58}$$

How do they transform under parity?

$$\begin{aligned} P &: \bar{\psi}\gamma^5\psi(\vec{x}, t) \to -\bar{\psi}\gamma^5\psi(-\vec{x}, t) \\ P &: \bar{\psi}\gamma^5\gamma^\mu\psi(\vec{x}, t) \to \begin{cases} -\bar{\psi}\gamma^5\gamma^0\psi(-\vec{x}, t) & \mu = 0 \\ +\bar{\psi}\gamma^5\gamma^i\psi(-\vec{x}, t) & \mu = i \end{cases} \end{aligned}$$

which means that $\bar{\psi}\gamma^5\psi$ transforms as a *pseudoscalar* (it takes a minus sign) and $\bar{\psi}\gamma^5\gamma^\mu\psi$ transforms as an *axial vector* (the temporal component takes a minus sign). To summarize, we have the following bilinears to construct our theory,

$$\begin{aligned} \bar{\psi}\psi &: \quad \text{scalar} \\ \bar{\psi}\gamma^\mu\psi &: \quad \text{vector} \\ \bar{\psi}S^{\mu\nu}\psi &: \quad \text{tensor} \\ \bar{\psi}\gamma^5\psi &: \quad \text{pseudoscalar} \\ \bar{\psi}\gamma^5\gamma^\mu\psi &: \quad \text{axial vector} \end{aligned}$$

Note that, given that terms involving γ^5 are not really scalar or vector, they will typically break parity, although this will not be always true. A theory that treats $\psi_\pm$ on equal footing is called a *vector-like theory*, while a theory that can distinguish between ψ_+ and ψ_- is called a *chiral theory*.

4.4.4 Majorana Fermions

Our spinors ψ^α are complex objects, and they have to be since $S[\Lambda]$ is also complex, so even a real ψ would become complex in general under a Lorentz transformation. However, there is a way to impose a reality condition. To do that, it is simpler to look at a new basis for the Clifford algebra, called *Majorana basis*,

$$\gamma^0 = \begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix}, \gamma^1 = \begin{pmatrix} i\sigma^3 & 0 \\ 0 & i\sigma^3 \end{pmatrix}, \gamma^2 = \begin{pmatrix} 0 & -\sigma^2 \\ \sigma^2 & 0 \end{pmatrix}, \gamma^3 = \begin{pmatrix} -i\sigma^1 & 0 \\ 0 & -i\sigma^1 \end{pmatrix},$$

which of course satisfy the Clifford algebra. They are special because they are all pure immaginary, $(\gamma^\mu)^* = -\gamma^\mu$, which means that the generators of the Lorentz group, $S^{\mu\nu} = 1/4[\gamma^\mu, \gamma^\nu]$, are real.In this basis we can then work on a real spinor

$$\psi^* = \psi\,, \tag{4.59}$$

which will be preserved real under Lorentz transformations in this basis. Such spinors are called *Majorana spinors*.

Let's now make a step back for a second and define the *charge conjugate* of a Dirac spinor as

$$\psi^{(c)} = C\psi^*\,, \tag{4.60}$$

where C is a 4×4 matrix satisfying

$$C^\dagger C = 1 \quad \text{and} \quad C^\dagger \gamma^\mu C = -(\gamma^\mu)^*\,. \tag{4.61}$$

How does it transform under Lorentz transformations?[5]

$$\psi^{(x)} \to CS[\Lambda]^*\psi^* = CS[\Lambda]C^\dagger C\psi^* = S[\Lambda]C\psi^* = S[\Lambda]\psi^{(x)}\,, \tag{4.62}$$

so that $\psi^{(c)}$ not only transforms nicely under Lorentz transformation, but if ψ satisfies the Dirac equation, $\psi^{(c)}$ does too. In general this is related to the anti-particles.

Coming back to the Majorana spinors, our reality condition implies

$$\psi^{(x)} = \psi\,, \tag{4.63}$$

which means that, after quantization, Majorana spinors give rise to particles that are their own antiparticles. This is exactly the same of the real scalar field case, but for the fact that they were spin 0 particles while the Majorana spinors are spin 1/2 particles.

The matrix C in Majorana basis is just $C_{\text{Maj}} = 1$. In the chiral basis $C_{\text{chiral}} = \begin{pmatrix} 0 & i\sigma^2 \\ -i\sigma^2 & 0 \end{pmatrix}$.

[5] Here we used the fact that $C^\dagger\gamma^\mu C = -(\gamma^\mu)^*$ implies $C(\gamma^\mu)^*C^\dagger = -\gamma^\mu$ and that in $S^{\mu\nu}$ there are two γs so that the minus sign disappear.

Finally, it is interesting to see how the Majorana condition looks like in the chiral representation, so in terms of chiral spinors. What we find is that a Majorana spinor can be written in chiral representation as

$$\psi = \begin{pmatrix} u_+ \\ -i\sigma^2 u_+^* \end{pmatrix} . \tag{4.64}$$

4.5 Symmetries and Conserved currents

The Dirac Lagrangian has a number of symmetries, let's look at them

4.5.1 Spacetime Translations

Under spacetime translations, the spinor transforms as

$$\delta\psi = \epsilon^\mu \partial_\mu \psi \, . \tag{4.65}$$

Applying Noether's theorem we get the energy-momentum tensor

$$T^{\mu\nu} = i\bar{\psi}\gamma^\mu \partial^\nu \psi - \eta^{\mu\nu}\mathcal{L} \, . \tag{4.66}$$

Given that the current is conserved only when the equations of motion are conserved, we may as well apply the equation of motion here without losing anything. In this way we can remove the $\mathcal{L}$ term and get

$$T^{\mu\nu} = i\bar{\psi}\gamma^\mu \partial^\nu \psi \, . \tag{4.67}$$

In particular, we get the total energy

$$E = \int d^3x T^{00} = \int d^3x i\bar{\psi}\gamma^0 \dot{\psi} = \int d^3x \psi^\dagger \gamma^0 (-i\gamma^i \partial_i + m)\psi \, . \tag{4.68}$$

4.5.2 Lorentz transformations

Under an infinitesimal Lorentz transformation we have

$$\delta\psi^\alpha = -\omega^\mu_\nu x^\nu \partial_\mu \psi^\alpha + \frac{1}{2}\Omega_{\rho\sigma}(S^{\rho\sigma})^\alpha_\beta \psi^\beta \, , \tag{4.69}$$

where $\omega^\mu_\nu = \frac{1}{2}\Omega_{\rho\sigma}(\mathcal{M}^{\rho\sigma})^\mu_\nu$. Using the properties of $\mathcal{M}$ we can easily show that $\omega = \Omega$ and so we can write

$$\delta\psi^\alpha = -\omega^{\mu\nu}[x_\nu \partial_\mu \psi^\alpha - \frac{1}{2}(S_{\mu\nu})^\alpha_\beta \psi^\beta] \, . \tag{4.70}$$

Using Noether's theorem now, we can get the conserved currents

$$(\mathcal{J}^\mu)^{\rho\sigma} = x^\rho T^{\mu\sigma} - x^\sigma T^{\mu\rho} - i\bar{\psi}\gamma^\mu S^{\rho\sigma}\psi \, , \tag{4.71}$$

where the first two terms are, as usual, the external angular momentum of the system, while the third term, after quantization, will be responsible for providing the state with internal angular momentum (the spin of the system). In particular, for a one particle state, it will tell us that a Dirac spinor gives rise to a particle of spin 1/2.

4.5.3 Internal Vector Symmetry

The Dirac Lagrangian is invariant under $\psi \to e^{-i\alpha}\psi$, which gives rise to the current

$$j_V^\mu = \bar{\psi}\gamma^\mu\psi\,, \tag{4.72}$$

where V stands for *vector* (it reflects the fact that this symmetry does not distinguish between left and right handed components). The conserved charge arising from this symmetry is

$$Q = \int d^3x\bar{\psi}\gamma^0\psi = \int d^3x\psi^{\dagger\psi} \tag{4.73}$$

which we will see has the interpretation of electric charge of particle number for fermions.

4.5.4 Axial Symmetry

When $m = 0$, the Dirac Lagrangian admits an *axial* symmetry,

$$\psi \to e^{i\alpha\gamma^5}\psi \quad \text{and} \quad \bar{\psi} \to \bar{\psi}e^{i\alpha\gamma^5}\,. \tag{4.74}$$

The conserved current is

$$j_A^\mu = \bar{\psi}\gamma^\mu\gamma^5\psi\,, \tag{4.75}$$

and it is conserved only when $m = 0$. In fact, we can compute

$$\partial_\mu j_A^\mu = (\partial_\mu\bar{\psi})\gamma^\mu\gamma^5\psi + \bar{\psi}\gamma^\mu\gamma^5\partial_\mu\psi = 2im\bar{\psi}\gamma^5\psi \tag{4.76}$$

which vanishes only for $m = 0$. However, this is interesting because, when the theory is coupled with *gauge fields*, the axial symmetry remains a symmetry of the classical Lagrangian, but it does not survive the quantization process! Cases such this are called *anomaly*.

4.6 Plane Wave Solutions

Let's now study the solutions to the Dirac equation

$$(i\gamma^\mu\partial_\mu - m)\psi = 0\,. \tag{4.77}$$

We start by assuming the ansatz

$$\psi = u(\vec{p})e^{-i\vec{p}\cdot x} \tag{4.78}$$

where $u(\vec{p})$ is a $4-$component spinor, independent of spacetime x and possibly dependent on momentum $\vec{p}$. The Dirac equation becomes

$$(\gamma^\mu p_\mu - m)u(\vec{p}) = \begin{pmatrix} -m & p_\mu\sigma^\mu \\ p_\mu\bar{\sigma}^\mu & -m \end{pmatrix} u(\vec{p}) = 0\,. \tag{4.79}$$

The solution to this equation is [6]

$$u(\vec{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\xi \\ \sqrt{p\cdot\bar{\sigma}}\xi \end{pmatrix} \,, \tag{4.80}$$

for any 2−component spinor ξ, normalized as $\xi^\dagger\xi = 1$.

Negative frequency solutions. We can get further solutions using the ansatz

$$\psi = v(\vec{p})e^{ip\cdot x}\,. \tag{4.81}$$

While the solutions we found before are called *positive energy solutions*, the ones we can find with this ansatz have negative energy, thus are called *negative energy solutions*. We can write it as

$$v(\vec{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\eta \\ -\sqrt{p\cdot\bar{\sigma}}\eta \end{pmatrix} \,, \tag{4.82}$$

with again $\eta^\dagger\eta = 1$.

4.6.1 Helicity

The helicity operator is the projection of the angular momentum along the direction of the momentum,

$$h = \frac{i}{2}\epsilon_{ijk}\hat{p}^i S^{jk} = \frac{1}{2}\hat{p}_i \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix} \,. \tag{4.83}$$

The massless field with spin $\xi^T = (1,0)$ has helicity $h = 1/2$ and we say it is *right-handed*, otherwise we say it is *left-handed*.

4.6.2 Some Useful Formulae: Inner and Outer Products

There are a number of identities that involve products of $u(\vec{p})$ and $v(\vec{p})$. It is firstly convenient to introduce a basis ξ^s and η^s with $s = 1,2$ for the two-components spinors such that[7]

$$\xi^{r\dagger}\xi^s = \delta^{rs} \quad \text{and} \quad \eta^{r\dagger}\eta^s = \delta^{rs}\,. \tag{4.84}$$

Starting with the positive energy solutions we have then that the two independent solutions are

$$u^s(\vec{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\xi^s \end{pmatrix} \,. \tag{4.85}$$

[6] We do not provide the proof here.

[7] A simple example is $\xi^1 = \begin{pmatrix}1,0\end{pmatrix}$ and $\xi^2 = \begin{pmatrix}0,1\end{pmatrix}$.

We can take inner products of four-components spinors in two ways, $u^\dagger \cdot u$ or $\bar{u} \cdot u$. The first is not Lorentz invariant but it is still useful for quantization. So we have

$$\begin{aligned} u^{r\dagger}(\vec{p}) \cdot u^s(\vec{p}) &= (\xi^{r\dagger}\sqrt{p\cdot\sigma}, \xi^{r\dagger}\sqrt{p\cdot\bar{\sigma}}) \begin{pmatrix} \sqrt{p\cdot\sigma}\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\xi^s \end{pmatrix} \\ &= \xi^{r\dagger} p\cdot\sigma\xi^s + \xi^{r\dagger} p\cdot\bar{\sigma}\xi^s = 2\xi^{r\dagger} p_0 \xi^s = 2p_0\delta^{rs}\,, \end{aligned}$$

while the Lorentz invariant one is

$$\bar{u}^r(\vec{p}) \cdot u^s(\vec{p}) = \cdots = 2m\delta^{rs}\,. \tag{4.86}$$

Analogous results can be found for the negative energy solutions,

$$\begin{aligned} v^{r\dagger}(\vec{p}) \cdot v^s(\vec{p}) &= 2p_0\delta^{rs} \\ \bar{v}^r(\vec{p}) \cdot v^s(\vec{p}) &= -2m\delta^{rs}\,. \end{aligned}$$

We may also compute the inner product between u and v, obtaining

$$\bar{u}^r(\vec{p}) \cdot v^s(\vec{p}) = \bar{v}^r(\vec{p}) \cdot u^s(\vec{p}) = 0\,, \tag{4.87}$$

and

$$u^{r\dagger}(\vec{p}) \cdot v^s(-\vec{p}) = v^{r\dagger}(\vec{p}) \cdot u^s(-\vec{p}) = 0\,. \tag{4.88}$$

Outer Products. We want to prove[8]

$$\sum_{s=1}^{2} u^s(\vec{p})\bar{u}^s(\vec{p}) = \not{p} + m \tag{4.89}$$

and

$$\sum_{s=1}^{2} v^s(\vec{p})\bar{v}^s(\vec{p}) = \not{p} - m\,. \tag{4.90}$$

Proof: we prove just the first statement.

$$\sum_{s=1}^{2} u^s(\vec{p})\bar{u}^s(\vec{p}) = \sum_{s=1}^{2} \begin{pmatrix} \sqrt{p\cdot\sigma}\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\xi^s \end{pmatrix} (\xi^{s\dagger}\sqrt{p\cdot\bar{\sigma}}, \xi^{s\dagger}\sqrt{p\cdot\sigma}\,, \tag{4.91}$$

but we know $\sum_s \xi^s\xi^{s\dagger} = 1$, the 2×2 unit matrix, so that we get

$$\sum_{s=1}^{2} u^s(\vec{p})\bar{u}^s(\vec{p}) = \begin{pmatrix} m & p\cdot\sigma \\ p\cdot\bar{\sigma} & m \end{pmatrix} \tag{4.92}$$

that is what we wanted to show.

[8] In this case we are not contracting anymore the spinors, but placing them back to back to give a 4×4 matrix. That is the meaning of an *outer product*.

Chapter 5

Quantizing the Dirac Field

We finally want to quantize the Dirac Lagrangian,

$$\mathcal{L} = \bar{\psi}(x)(i\partial\!\!\!/ - m)\psi(x)\,. \tag{5.1}$$

We will start doing the same we did for the scalar field and we will realize things go wrong.

5.1 A Glimpse at the Spin-Statistics Theorem

As usual we have to define the momentum conjugate,

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\bar{\psi}\gamma^0 = i\psi^\dagger\,, \tag{5.2}$$

and then we promote the field ψ and its momentum $\sim \psi^\dagger$ to operators satisfying the canonical relations, as

$$\begin{aligned}
[\psi_\alpha(\vec{x}), \psi_\beta(\vec{y})] &= [\psi^\dagger_\alpha(\vec{x}), \psi^\dagger_\beta(\vec{y})] = 0 \\
[\psi_\alpha(\vec{x}), \psi^\dagger_\beta(\vec{y})] &= \delta_{\alpha\beta}\delta^{(3)}(\vec{x} - \vec{y})\,.
\end{aligned}$$

This step is what we will soon have to reconsider.

Since we are dealing with a free theory, each classical solution is a sum of plane waves, so we can write the quantum operators as

$$\begin{aligned}
\psi(\vec{x}) &= \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\vec{p}}}} [b^s_{\vec{p}} u^s(\vec{p}) e^{ip\cdot x} + c^{s\dagger}_{\vec{p}} v^s(\vec{p}) e^{-i\vec{p}\cdot x}] \\
\psi^\dagger(\vec{x}) &= \sum_{s=1}^{2} \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\vec{p}}}} [b^{s\dagger}_{\vec{p}} u^s(\vec{p})^\dagger e^{-ip\cdot x} + c^s_{\vec{p}} v^s(\vec{p})^\dagger e^{i\vec{p}\cdot x}]\,,
\end{aligned}$$

where $b^{s\dagger}_{\vec{p}}$ create particles associated to spinors $u^s(\vec{p})$, $c^{s\dagger}_{\vec{p}}$ creates particles associated to $v^s(\vec{p})$ and so on. From the canonical commutators, we can get the commutation relations of the operators,

$$[b^r_{\vec{p}}, b^{s\dagger}_{\vec{q}}] = (2\pi)^3\delta^{rs}\delta^{(3)}(\vec{p}-\vec{q})$$
$$[c^r_{\vec{p}}, c^{s\dagger}_{\vec{q}}] = -(2\pi)^3\delta^{rs}\delta^{(3)}(\vec{p}-\vec{q})\,,$$

and all the other commutators vanish.

Note the minus sign on the commutators of the $c^\dagger$ operators. It implies that we cannot define the ground state as $c^r_{\vec{p}}|0\rangle = 0$ because then $c^{s\dagger}_{\vec{p}}|0\rangle$ would get a negative norm. We will then have to flip the interpretation of of c and $c^\dagger$ in such a way the ground state will be defined as $c^{s\dagger}_{\vec{p}}|0\rangle = 0$ and the excited states will be constructed with $c^r_{\vec{p}}$. We will see that this is exactly what we need to change.

5.1.1 The Hamiltonian

Let's proceed constructing the Hamiltonian for the theory. We have

$$\mathcal{H} = \pi\dot{\psi} - \mathcal{L} = \bar{\psi}(-i\gamma^i\partial_i + m)\psi\,, \tag{5.3}$$

with which we can see that $H = \int d^3x\mathcal{H}$ agrees with the conserved energy computed in the previous chapter. We then need to promote the Hamiltonian to an operator. Let's start with

$$(-i\gamma^i\partial_i+m)\psi = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}[b^s_{\vec{p}}(-\gamma^i p_i+m)u^s(\vec{p})e^{ip\cdot x}+c^{s\dagger}_{\vec{p}}(\gamma^i p_i+m)v^s(\vec{p})e^{-i\vec{p}\cdot\vec{x}}]\,,$$

where we left the sum over $s=1,2$ implicit. Now we use the definitions

$$(-\gamma^i p_i + m)u^s(\vec{p}) = \gamma^0 p_0 u^s(\vec{p}) \quad \text{and} \quad (\gamma^i p_i + m)v^s(\vec{p}) = -\gamma^0 p_0 v^s(\vec{p})\,, \tag{5.4}$$

so we can write

$$(-i\gamma^i\partial_i + m)\psi = \int \frac{d^3p}{(2\pi)^3}\sqrt{\frac{E_{\vec{p}}}{2}}\gamma^0[b^s_{\vec{p}}u^s(\vec{p})e^{i\vec{p}\cdot\vec{x}} - c^{s\dagger}_{\vec{p}}v^s(\vec{p})e^{-i\vec{p}\cdot\vec{x}}]\,. \tag{5.5}$$

We can now write the Hamiltonian operator as

$$\begin{aligned}
H &= \int d^3x\psi^\dagger\gamma^0(-i\gamma^i\partial_i + m)\psi \\
&= \int \frac{d^3xd^3pd^3q}{(2\pi)^6}\sqrt{\frac{E_{\vec{p}}}{4E_{\vec{q}}}}[b^{r\dagger}_{\vec{q}}u^r(\vec{q})^\dagger e^{-i\vec{q}\cdot\vec{x}} + c^r_{\vec{q}}v^r(\vec{q})^\dagger e^{i\vec{q}\cdot\vec{x}}]\cdot \\
&\quad [b^s_{\vec{p}}u^s(\vec{p})e^{i\vec{p}\cdot\vec{x}} - c^{s\dagger}_{\vec{p}}v^s(\vec{p})e^{-i\vec{p}\cdot\vec{x}}] \\
&= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2}\big[b^{r\dagger}_{\vec{p}}b^s_{\vec{p}}[u^r(\vec{p})^\dagger\cdot u^s(\vec{p})] - c^r_{\vec{p}}c^{s\dagger}_{\vec{p}}[v^r(\vec{p})^\dagger\cdot v^s(\vec{p})] \\
&\quad - b^{r\dagger}_{\vec{p}}c^{s\dagger}_{-\vec{p}}[u^r(\vec{p})^\dagger\cdot v^s(-\vec{p})] + c^r_{\vec{p}}b^s_{-\vec{p}}[v^r(\vec{p})^\dagger\cdot u^s(-\vec{p})]\big]\,.
\end{aligned}$$

We can now use our inner products equations to get

$$\begin{aligned} H &= \int \frac{d^3p}{(2\pi)^3} E_{\vec{p}} (b^{s\dagger}_{\vec{p}} b^s_{\vec{p}} - c^s_{\vec{p}} c^{s\dagger}_{\vec{p}}) \\ &= \int \frac{d^3p}{(2\pi)^3} E_{\vec{p}} \big(b^{s\dagger}_{\vec{p}} b^s_{\vec{p}} - c^{s\dagger}_{\vec{p}} c^s_{\vec{p}} + (2\pi)^3 \delta^{(3)}(0)\big) \,. \end{aligned}$$

The $\delta^{(3)}$ term is familiar and we know we can deal with it with normal ordering. The $b^\dagger b$ term is also familiar and we can check $b^\dagger$ creates positive energy states as expected,

$$[H, b^{s\dagger}_{\vec{p}}] = E_{\vec{p}} b^{s\dagger}_{\vec{p}} \,. \tag{5.6}$$

The minus sign in front of the $c^\dagger c$ is instead weird. If we interpret $c^\dagger$ as creation operators, then there is no problem because we still find that $c^\dagger$ creates positive energy states. However, we already noted that these states have negative norm so, in order to have a sensible Hilbert space, we are forced to interpret c as creation operators instead. But in this case, the minus sign we got tells us that the Hamiltonian is not bounded from below because

$$[H, c^s_{\vec{p}}] = -E_{\vec{p}} c^s_{\vec{p}} \,. \tag{5.7}$$

This is a real problem as, if this was the case, it would mean that we could get states of lower and lower energies just by producing c particles!

In fact, this minus sign is telling us something important.

5.2 Fermionic Quantization

The key piece we missed is that the particles we are describing are fermions and so they obey Fermi-Dirac statistics, meaning that *the quantum state must pick a minus sign upon the interchange of any two particles.* In fact, the *spin-statistics* theorem tells us that integer spin fields must be quantized as bosons (fields commute) while half-integer spin fields must be quantized as fermions (fields anti-commute). We must be consistent with this.

Remember that the scalar fields obeyed bosonic statistic because creation and annihilation operators obeyed commutation relations, i.e

$$[a^\dagger_{\vec{p}}, a^\dagger_{\vec{q}}] = 0 \Rightarrow a^\dagger_{\vec{p}} a^\dagger_{\vec{q}} |0\rangle \equiv |\vec{p}, \vec{q}\rangle = |\vec{q}, \vec{p}\rangle \,. \tag{5.8}$$

This tells us that, in order to have states obeying fermion statistics, we need to use *anticommutators.* So we must ask

$$\begin{aligned} \{\psi_\alpha(\vec{x}), \psi_\beta(\vec{y})\} &= \{\psi^\dagger_\alpha(\vec{x}), \psi^\dagger_\beta(\vec{y})\} = 0 \\ \{\psi_\alpha(\vec{x}), \psi^\dagger_\beta(\vec{y})\} &= \delta_{\alpha\beta} \delta^{(3)}(\vec{x} - \vec{y}) \,, \end{aligned}$$

then the same expansion of the fields in terms of $b, b^\dagger, c, c^\dagger$ applies but now it tells us

$$\{b^r_{\vec{p}}, b^{s\dagger}_{\vec{q}}\} = (2\pi)^3\delta^{rs}\delta^{(3)}(\vec{p}-\vec{q})$$
$$\{c^r_{\vec{p}}, c^{s\dagger}_{\vec{q}}\} = (2\pi)^3\delta^{rs}\delta^{(3)}(\vec{p}-\vec{q})$$

and

$$\{b^r_{\vec{p}}, b^s_{\vec{q}}\} = \{c^r_{\vec{p}}, c^s_{\vec{q}}\} = \{b^r_{\vec{p}}, c^{s\dagger}_{\vec{q}}\} = \cdots = 0\,. \tag{5.9}$$

Also the computation of the Hamiltonian is the same until

$$\begin{aligned} H &= \int \frac{d^3p}{(2\pi)^3} E_{\vec{p}}[b^{s\dagger}_{\vec{p}} b^s_{\vec{p}} - c^s_{\vec{p}} c^{s\dagger}_{\vec{p}}] \\ &= \int \frac{d^3p}{(2\pi)^3} E_{\vec{p}}[b^{s\dagger}_{\vec{p}} b^s_{\vec{p}} + c^{s\dagger}_{\vec{p}} c^s_{\vec{p}} - (2\pi)^3\delta^{(3)}(0)]\,, \end{aligned}$$

which is the same we got before but for the infamous minus sign. This is what we wanted to get.

5.2.1 Fermi-Dirac Statistics

We define the vacuum $|0\rangle$ as

$$b^s_{\vec{p}}|0\rangle = c^s_{\vec{p}}|0\rangle = 0\,. \tag{5.10}$$

We also note that, although b and c obey anti-commutators relations, the Hamiltonian obey commutation relations with them

$$\begin{aligned} [H, b^r_{\vec{p}}] &= -E_{\vec{p}} b^r_{\vec{p}} \quad \text{and} \quad [H, b^{r\dagger}_{\vec{p}}] = E_{\vec{p}} b^{r\dagger}_{\vec{p}} \\ [H, c^r_{\vec{p}}] &= -E_{\vec{p}} c^r_{\vec{p}} \quad \text{and} \quad [H, c^{r\dagger}_{\vec{p}}] = E_{\vec{p}} c^{r\dagger}_{\vec{p}}\,. \end{aligned}$$

This ensures we can again construct a tower of energy eigenstates by acting on the vacuum with $b^\dagger$ and $c^\dagger$. For example, we have the one-particle state

$$|\vec{p}, r\rangle = b^{r\dagger}_{\vec{p}}|0\rangle\,, \tag{5.11}$$

and the two particle states now satisfy

$$|\vec{p}_1, r_1; \vec{p}_2, r_2\rangle \equiv b^{r_1\dagger}_{\vec{p}_1} b^{r_2\dagger}_{\vec{p}_2}|0\rangle = -|\vec{p}_2, r_2; \vec{p}_1, r_1\rangle\,, \tag{5.12}$$

which confirms that these particles obey Fermi-Dirac statistics. In particular, they satisfy Pauli exclusion principle $|\vec{p}, r; \vec{p}, r\rangle = 0$. We could also check that a stationary particle $|\vec{p} = 0, r\rangle$ carry intrinsic angular momentum $1/2$, as expected.

5.3 Propagators

We define the fermionic propagator to be

$$iS_{\alpha\beta} = \{\psi_\alpha(x), \bar{\psi}_\beta(y)\}\,, \tag{5.13}$$

where in the following we will often drop the indices α, β. Inderting the expansions of the fields we get

$$\begin{aligned} iS(x-y) &= \int \frac{d^3p d^3q}{(2\pi)^6} \frac{1}{\sqrt{4E_{\vec{p}}E_{\vec{q}}}} \big[\{b^s_{\vec{q}}, b^{r\dagger}_{\vec{q}}\} u^s(\vec{p})\bar{u}^r(\vec{q}) e^{-i(p\cdot x - q\cdot y)} \\ &+ \{c^{s\dagger}_{\vec{p}}, c^r_{\vec{q}}\} v^s(\vec{p})\bar{v}^r(\vec{q}) e^{i(p\cdot x - q\cdot y)}\big] \\ &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec{p}}} [u^s(\vec{p})\bar{u}^s(\vec{p}) e^{-ip\cdot(x-y)} + v^s(\vec{p})\bar{v}^s(\vec{p}) e^{ip\cdot(x-y)}] \\ &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec{p}}} [(\not{p}+m) e^{-ip\cdot(x-y)} + (\not{p}-m) e^{ip\cdot(x-y)}]\,. \end{aligned}$$

We can then write

$$iS(x-y) = (i\not{\partial}_x + m)(D(x-y) - D(y-x)) \tag{5.14}$$

where $D(x-y)$ is the scalar propagator that, recall, can be written as

$$D(x-y) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec{p}}} e^{-ip\cdot(x-y)}\,. \tag{5.15}$$

Note that, at least away from singularities, the propagator satisfies

$$(i\not{\partial}_x - m) S(x-y) = 0\,. \tag{5.16}$$

5.4 The Feynman Propagator

Doing similar calculations, we can determine the vacuum expectation value

$$\begin{aligned} \langle 0|\psi_\alpha(x)\bar{\psi}_\beta(y)|0\rangle &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec{p}}} (\not{p}+m)_{\alpha\beta} e^{-ip\cdot(x-y)} \\ \langle 0|\bar{\psi}_\beta(y)\psi_\alpha(x)|0\rangle &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec{p}}} (\not{p}-m)_{\alpha\beta} e^{+ip\cdot(x-y)}\,. \end{aligned}$$

So we can now define the Feynman propagator $S_F(x-y)$ as a time ordered product,

$$S_F(x-y) = \langle 0|T[\psi(x)\bar{\psi}(y)]|0\rangle \equiv \begin{cases} \langle 0|\psi(x)\bar{\psi}(y)|0\rangle & x^0 > y^0 \\ \langle 0| -\bar{\psi}(y)\psi(x)|0\rangle & y^0 > x^0 \end{cases}\,. \tag{5.17}$$

Notice the minus sign: it is necessary for Lorentz invariance. We can also write it in the 4−momentum representation as

$$S_F(x-y) = i\int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} \frac{\gamma\cdot p + m}{p^2 - m^2 + i\epsilon}\,, \tag{5.18}$$

which satisfies $(i\partial\!\!\!/_x - m)S_F(x-y) = i\delta^{(4)}(x-y)$, so that it is a Green function for the Dirac operator.

Also note that fermionic fields also anticommute inside time-ordering products and normal ordered products. This means that Wick theorem work as in the scalar case but with the contraction definition

$$\overbrace{\psi(x)\bar{\psi}(y)} = T[\psi(x)\bar{\psi}(y)] - :\psi(x)\bar{\psi}(y): = S_F(x-y)\,. \tag{5.19}$$

5.5 Yukawa Theory

The interaction between a Dirac fermion of mass m and a real scalar field of mass μ is described by the *Yukawa theory*,

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}\mu^2\phi^2 + \bar{\psi}(i\partial\!\!\!/ - m)\psi - \lambda\phi\bar{\psi}\psi\,, \tag{5.20}$$

which extends the Yukawa theory we already looked at. Couplings of this kind actually appear in the SM for the interaction of fermions with the *Higgs boson*.

Since Yukawa first proposed this theory as an effective theory of interactions between mesons and nucleons, we will use it in this context and still call ϕ the mesons and ψ the nucleons (with the difference, with respect to before, that now nucleons have spin). Note that, since $[\phi] = 1$ and $[\psi] = 3/2$, $[\lambda] = 0$ and it is then a dimensionless coupling. We then proceed as before, computing the amplitude for a particular process and then using it to infer the Feynman rules.

5.5.1 Example: Nucleon scattering with spin

Let's study $\psi\psi \to \psi\psi$ scattering. Same of what we did but now with spin. Our initial and final states are

$$\begin{aligned} |i\rangle &= \sqrt{4E_{\vec{p}}E_{\vec{q}}}\, b^{s\dagger}_{\vec{p}} b^{r\dagger}_{\vec{q}}|0\rangle \equiv |\vec{p}, s; \vec{q}, r\rangle \\ |f\rangle &= \sqrt{4E_{\vec{p}'}E_{\vec{q}'}}\, b^{s'\dagger}_{\vec{p}'} b^{r'\dagger}_{\vec{q}'}|0\rangle \equiv |\vec{p}', s'; \vec{q}', r'\rangle\,, \end{aligned}$$

where now we should be careful about minus signs, in particular when doing the adjoint

$$\langle f| = \sqrt{4E_{\vec{p}'}}E_{\vec{q}'}\langle 0|b^{r'}_{\vec{q}'} b^{s'}_{\vec{p}'}\,. \tag{5.21}$$

We need to compute the order λ^2 terms of $\langle f|S-1|i\rangle$, that is (in the interaction picture)

$$\frac{(-i\lambda)^2}{2}\int d^4x_1 d^4x_2 T[\bar{\psi}(x_1)\psi(x_1)\phi(x_1)\bar{\psi}(x_2)\psi(x_2)\phi(x_2)]\,. \tag{5.22}$$

As before, the only term contributing at this order is

$$: \bar{\psi}(x_1)\psi(x_1)\bar{\psi}(x_2)\psi(x_2) : \overbrace{\phi(x_1)\phi(x_2)} \,. \tag{5.23}$$

Let's start looking at how the fermionic operator acts on $|i\rangle$, for the moment expanding out only the ψ fields and also ignoring the $c^\dagger$ pieces in ψ since they give no contribution in this case. So we have

$$: \bar{\psi}(x_1)\psi(x_1)\bar{\psi}(x_2)\psi(x_2) : b^{s\dagger}_{\vec{p}} b^{r\dagger}_{\vec{q}}|0\rangle = -\int \frac{d^3k_1 d^3k_2}{(2\pi)^6}[\bar{\psi}(x_1)\cdot u^m(\vec{k}_1)][\bar{\psi}(x_2)\cdot u^n(\vec{k}_2)] \frac{e^{-ik_1\cdot x_1 - ik_2\cdot x_2}}{\sqrt{4E_{\vec{k}_1}E_{\vec{k}_2}}} b^m_{\vec{k}_1} b^n_{\vec{k}_2} b^{s\dagger}_{\vec{p}} b^{r\dagger}_{\vec{q}}|0\rangle \,,$$

where the minus sign in front comes from the fact that we moved $\psi(x_1)$ past $\bar{\psi}(x_2)$. Now it is enough to anti-commute b's past $b^\dagger$'s to get

$$= \frac{-1}{2\sqrt{E_{\vec{p}}E_{\vec{q}}}}\Big([\bar{\psi}(x_1)\cdot u^r(\vec{q})][\bar{\psi}(x_2)\cdot u^s(\vec{p})]e^{-ip\cdot x_2 - iq\cdot x_1} - [\bar{\psi}(x_1)\cdot u^s(\vec{p})][\bar{\psi}(x_2)\cdot u^r(\vec{q})]e^{-ip\cdot x_1 - iq\cdot x_2}\Big)|0\rangle \,.$$

Let's now contract this with $\langle f|$. First let's look at

$$\langle 0|b^{r'}_{\vec{q}'} b^{s'}_{\vec{p}'}[\bar{\psi}(x_1)\cdot u^r(\vec{q})][\bar{\psi}(x_2)\cdot u^s(\vec{p})]|0\rangle = \frac{e^{ip'\cdot x_1 + iq'\cdot x_2}}{2\sqrt{E_{\vec{p}'}E_{\vec{q}'}}}[\bar{u}^{s'}(\vec{p}')\cdot u^r(\vec{q})][\bar{u}^{r'}(\vec{q}')\cdot u^s(\vec{p})] - \frac{e^{ip'\cdot x_2 + iq'\cdot x_1}}{2\sqrt{E_{\vec{p}'}E_{\vec{q}'}}}[\bar{u}^{r'}(\vec{q}')\cdot u^r(\vec{q})][\bar{u}^{s'}(\vec{p}')\cdot u^s(\vec{p})] \,,$$

and it is easy to see that the other term, $[\bar{\psi}(x_1)\cdot u^s(\vec{p})][\bar{\psi}(x_2)\cdot u^r(\vec{q})]$, just doubles up with this canceling the 1/2 factor in front. At the end we get

$$\langle f|S-1|i\rangle = (-i\lambda)^2\int\frac{d^4x_1 d^4x_2 d^4k}{(2\pi)^4}\frac{ie^{ik\cdot(x_1-x_2)}}{k^2-\mu^2+i\epsilon}\Big([\bar{u}^{r'}(\vec{q}')\cdot u^r(\vec{q})][\bar{u}^{s'}(\vec{p}')\cdot u^s(\vec{p})]e^{ix_1\cdot(q'-q)+ix_2\cdot(p'-p)} - [\bar{u}^{s'}(\vec{p}')\cdot u^r(\vec{q})][\bar{u}^{r'}(\vec{q}')\cdot u^s(\vec{p})]e^{ix_1\cdot(p'-q)+ix_2\cdot(q'-p)}\Big) \,,$$

and computing the integrals over x_1 and x_2 becomes

$$\int d^4k\frac{(2\pi)^4 i(-i\lambda)^2}{k^2-\mu^2+i\epsilon}\Big([\bar{u}^{r'}(\vec{q}')\cdot u^r(\vec{q})][\bar{u}^{s'}(\vec{p}')\cdot u^s(\vec{p})]\delta^{(4)}(q'-q+k)\delta^{(4)}(p'-p-k) - [\bar{u}^{s'}(\vec{p}')\cdot u^r(\vec{q})][\bar{u}^{r'}(\vec{q}')\cdot u^s(\vec{p})]\delta^{(4)}(p'-q+k)\delta^{(4)}(q'-p-k)\Big) \,.$$

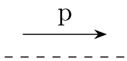
Finally we need to remember that $\langle f|S-1|i\rangle = i\mathcal{A}(2\pi)^4\delta^{(4)}(p+q-p'-q')$, to get the amplitude

$$\mathcal{A} = (-i\lambda)^2\left(\frac{[\bar{u}^{r'}(\vec{q}')\cdot u^r(\vec{q})][\bar{u}^{s'}(\vec{p}')\cdot u^s(\vec{p})]}{(p'-p)^2-\mu^2+i\epsilon} - \frac{[\bar{u}^{s'}(\vec{p}')\cdot u^r(\vec{q})][\bar{u}^{r'}(\vec{q}')\cdot u^s(\vec{p})]}{(q'-p)^2-\mu^2+i\epsilon}\right) \,.$$

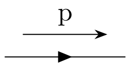
This calculation is already quite long, let's see how Feynman rules can simplify our life

5.6 Feynman Rules for Fermions

- To each incoming fermion with momentum p and spin r, we associate a spinor $u^r(\vec{p})$. To each outgoing fermion we associate $\bar{u}^r(\vec{p})$.
- To each incoming anti-fermion with momentum p and spin r, we associate the spinor $\bar{v}(\vec{p})$. For outgoing anti-fermions we associate $v^r(\vec{p})$.
- Each vertex gets a factor $(-i\lambda)$ (this is of course dependent on the interacting part of the theory).
- Each internal line gets a factor of the relevant propagator



$$\frac{i}{p^2-\mu^2+i\epsilon} \quad \text{for scalars}$$



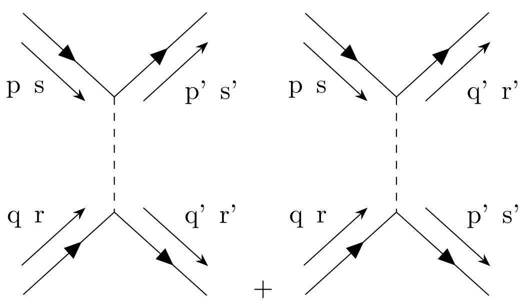
$$\frac{i(\not{p}+m)}{p^2-m^2+i\epsilon} \quad \text{for fermions}\,.$$

  Note that fermion lines should flow consistently through the diagram to ensure fermion number conservation. Note also that the fermion propagator is a 4×4 matrix and it is contracted at each vertex either with spinors or with further propagators.
- Impose momentum conservation at each vertex and integrate over undetermined loop momenta.
- Add extra minus signs for statistics (we will see examples).

5.6.1 Examples

Let's see the same examples we saw for the Yukawa scalar theory.

Again Nucleons scattering with spins. The two diagrams contributing to this process are

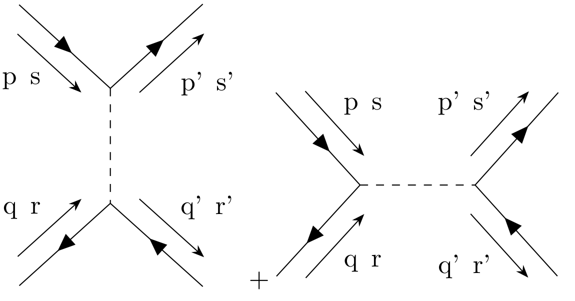


where the second is obtained from the first just switching the two final particles. For this reason, given that we are talking about fermions, we get a minus sign.

Using the Feynman rules we can easily get

$$\mathcal{A} = i(-i\lambda)^2\left(\frac{[u^{s'}(\vec{p'})\cdot u^s(\vec{p})][u^{r'}(\vec{q'})\cdot u^r(\vec{q})]}{(p-p')^2-\mu^2} - \frac{[u^{r'}(\vec{q'})\cdot u^s(\vec{p})][u^{s'}(\vec{p'})\cdot u^r(\vec{q})]}{(p-q')^2-\mu^2}\right).$$

Nucleon-Anti-Nucleon Scattering with spin. For $\psi\bar{\psi} \to \psi\bar{\psi}$ the two lowest order scattering diagrams are



We do not compute the amplitude here, but we just want to note that the first diagram takes a minus sign with respect to the second. In fact it can be obtained from the second, just switching on final particle with an initial particle.

5.6.2 Pseudo-Scalar Coupling

Rather than the standard Yukawa-Coupling we could have considered

$$\mathcal{L}_{\text{Yuk}} = -\lambda\phi\bar{\psi}\gamma^5\psi\,, \tag{5.24}$$

which still preserves parity in case ϕ is pseudoscalar instead of scalar.

Chapter 6

Quantum Electrodynamics

We finally arrive to describe *quantum electrodynamics* (QED) which is the theory of light interacting with charged matter.

6.1 Maxwell's Equations

The Lagrangian from which we can get Maxwell's equations, in absence of external sources, is

$$\mathcal{L} = \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \,, \tag{6.1}$$

where the *field strength tensor* is defined as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,. \tag{6.2}$$

The equation of motion we get from this Lagrangian is

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) = -\partial_\mu F^{\mu\nu} = 0 \,, \tag{6.3}$$

and we can also get the *Bianchi* identity from the field strength

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0 \,. \tag{6.4}$$

If we want to recover the usual Maxwell's equations we need to define some 3−vector notation. In particular, defining $A^\mu = (\phi, \vec{A})$, we get that the electric and magnetic fields are

$$\vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t} \quad \text{and} \quad \vec{B} = \nabla \times \vec{A} \,, \tag{6.5}$$

which can be also written in terms of $F_{\mu\nu}$ as

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix} \,. \tag{6.6}$$

From Bianchi's identity is then easy to get two Maxwell's equations

$$\nabla \cdot \vec{B} = 0 \quad \text{and} \quad \frac{\partial \vec{B}}{\partial t} = -\nabla \times \vec{E} \,. \tag{6.7}$$

These remain true even in presence of electric sources, while the ones obtained from the equations of motion,

$$\nabla \cdot \vec{E} = 0 \quad \text{and} \quad \frac{\partial \vec{E}}{\partial t} = \nabla \times \vec{B} \,, \tag{6.8}$$

get extra terms in presence of charged matter.

6.1.1 Gauge Symmetry

The massless vector field A_μ has 4 components from which one can think that it has 4 degrees of freedom. However, we know that the photon has only 2 degrees of freedom, which we call *polarization states*. There are two points that ensure that, after quantization, we will get 2 degrees of freedom:

- The field A_0 has no kinetic term $\dot{A}_0$ in the Lagrangian, i.e. it is not dynamical. In particular, given the initial conditions A_i and $\dot{A}_i$, we can fully determine A_0 from $\nabla \cdot \vec{E} = 0$, which can be written as

  $$\nabla^2 A_0 + \nabla \cdot \frac{\partial \vec{A}}{\partial t} = 0 \,. \tag{6.9}$$

  The solution is, in fact,

  $$A_0(\vec{x}) = \int d^3x' \, \frac{(\nabla \cdot \partial \vec{A}/\partial t)(\vec{x}'))}{4\pi |\vec{x} - \vec{x}'|} \,. \tag{6.10}$$

  A_0 is then not a free degree of freedom, leaving us with only three (which are however still too many).

- The Lagrangian has a large symmetry group

  $$A_\mu(x) \to A_\mu(x) + \partial_\mu \lambda(x) \tag{6.11}$$

  for any function $\lambda(x)$ that decays suitably quick as $\vec{x} \to \infty$. This is called a *gauge symmetry*. Of course the field strength is invariant under gauge symmetry. This means that we have a theory with an infinite number of symmetries, one for each choice of $\lambda(x)$. We know how to deal with symmetries that affect each point of the spacetime in the same way, but this is not like that. Do we know how to deal with this? The point is that gauge symmetries have a different interpretation than the global symmetries we are used to. The latter take a physical state to another physical state with the same properties, while the former is a *redundancy* of the description (so that two states linked by a gauge symmetry are exactly the same state). This is also very important because otherwise Maxwell's equations would not be able to specify entirely the evolution of A_μ. In fact, the evolution is given by

$$[\eta_{\mu\nu}(\partial^\rho\partial_\rho) - \partial_\mu\partial_\nu]A^\nu = 0\,, \tag{6.12}$$

which is such that any part of A^ν of the form $\partial_\nu\lambda$ is just 0. This means that we cannot distinguish between A^μ and $A^\mu + \partial^\mu\lambda$ and so they better describe exactly the same state!

Different representative configurations of a physical state are called *different gauges*. There are different possibilities that can be easier to use in different contexts. For example we will use:

- **Lorentz Gauge:** $\partial_\mu A^\mu = 0$. It has the advantage of being Lorentz invariant. Also, it does not use completely the freedom of gauge symmetry.
- **Coulomb Gauge:** $\nabla \cdot \vec{A} = 0$. Use the remaining freedom from Lorentz Gauge and, as a result, it is not Lorentz invariant anymore. It is called sometimes the *radiation gauge*.

6.2 The Quantization of the Electromagnetic field

We will perform the quantization twice: one in Coulomb gauge and one in Lorentz gauge. In both case we will have some subtleties.
The first of the subtleties is common to both the gauges and happens when computing the momentum conjugate,

$$\begin{aligned} \pi^0 &= \frac{\partial\mathcal{L}}{\partial\dot{A}_0} = 0 \\ \pi^i &= \frac{\partial\mathcal{L}}{\partial\dot{A}^i} = -F^{0i} \equiv E^i\,, \end{aligned}$$

so that π^0 vanishes (by the way this is expected from what we discussed in the previous section). We can then compute the Hamiltonian

$$\begin{aligned} H &= \int d^3x \pi^i\dot{A}^i - \mathcal{L} \\ &= \int d^3x \frac{1}{2}\vec{E}\cdot\vec{E} + \frac{1}{2}\vec{B}\cdot\vec{B} - A_0(\nabla\cdot\vec{E})\,, \end{aligned}$$

from which we can see that A_0 acts as a Lagrange multiplier which imposes Gauss' law

$$\nabla\cdot\vec{E} = 0\,. \tag{6.13}$$

6.2.1 Coulomb Gauge

In Coulomb gauge the equation of motion for $\vec{A}$ is

$$\partial_\mu \partial^\mu \vec{A} = 0\,, \tag{6.14}$$

that has solution

$$\vec{A} = \int \frac{d^3p}{(2\pi)^3} \vec{\xi}(\vec{p}) e^{ip\cdot x}\,, \tag{6.15}$$

with $p_0^2 = |\vec{p}|^2$. Since in this gauge $\nabla \cdot \vec{A} = 0$, we have

$$\vec{\xi} \cdot \vec{p} = 0\,, \tag{6.16}$$

i.e. $\vec{\xi}$ is perpendicular to the direction of motion. We can then pick $\vec{\xi}(\vec{p})$ to be a linear combination of two orthonormal vectors $\vec{\epsilon}_r$, with $r = 1, 2$, each of them satisfying $\vec{\epsilon}_r(\vec{p}) \cdot \vec{p} = 0$ and

$$\vec{\epsilon}_r(\vec{p}) \cdot \vec{\epsilon}_s(\vec{p}) = \delta_{rs}\,. \tag{6.17}$$

These two vectors correspond to the two polarization states of the photon.
To quantize then, as usual, we need to promote the fields to operators, changing the Poisson brackets into commutators. Naively one would do

$$\begin{aligned}
[A_i(\vec{x}), A_j(\vec{y})] &= [E^i(\vec{x}), E^j(\vec{y})] = 0 \\
[A_i(\vec{x}), E^j(\vec{y})] &= i\delta_i^j \delta^{(3)}(\vec{x} - \vec{y})\,,
\end{aligned}$$

but this is not consistent with $\nabla \cdot \vec{A} = \nabla \cdot \vec{E} = 0$. In fact,

$$[\nabla \cdot \vec{A}(\vec{x}), \nabla \cdot \vec{E}(\vec{y})] = i\nabla^2 \delta^{(3)}(\vec{x} - \vec{y}) \neq 0\,. \tag{6.18}$$

This is something that is already known at the classical level: in order to fulfill the constraints we need to change the second commutation relation in

$$[A_i(\vec{x}), E_j(\vec{y})] = i\left(\delta_{ij} - \frac{\partial_i \partial_j}{\nabla^2}\right)\delta^{(3)}(\vec{x} - \vec{y})\,. \tag{6.19}$$

It is easy to see that with this choice the constraints are fulfilled.

We can now write the operators in the usual expansion,

$$\begin{aligned}
\vec{A}(\vec{x}) &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2|\vec{p}|}} \sum_{r=1}^{2} \vec{\epsilon}_r(\vec{p}) [a_{\vec{p}}^r e^{i\vec{p}\cdot\vec{x}} + a_{\vec{p}}^{r\dagger} e^{-i\vec{p}\cdot\vec{x}}] \\
\vec{E}(\vec{x}) &= \int \frac{d^3p}{(2\pi)^3} (-i) \sqrt{\frac{|\vec{p}|}{2}} \sum_{r=1}^{2} \vec{\epsilon}_r(\vec{p}) [a_{\vec{p}}^r e^{i\vec{p}\cdot\vec{x}} + a_{\vec{p}}^{r\dagger} e^{-i\vec{p}\cdot\vec{x}}]\,.
\end{aligned}$$

As usual the commutation relations between operators imply

$$\begin{aligned}
[a_{\vec{p}}^r, a_{\vec{q}}^s] &= [a_{\vec{p}}^{r\dagger}, a_{\vec{q}}^{s\dagger}] = 0 \\
[a_{\vec{p}}^r, a_{\vec{q}}^{s\dagger}] &= (2\pi)^3 \delta^{rs} \delta^{(3)}(\vec{p} - \vec{q})\,,
\end{aligned}$$

where we used the completeness relation for the polarization vectors,

$$\sum_{r=1}^{2} \epsilon_r^i(\vec{p})\epsilon_r^j(\vec{p}) = \delta^{ij} - \frac{p^i p^j}{|\vec{p}|^2}\,. \tag{6.20}$$

We then get the Hamiltonian

$$H = \int \frac{d^3p}{(2\pi)^3}|\vec{p}|\sum_{r=1}^{2} a_{\vec{p}}^{r\dagger} a_{\vec{p}}^{r}\,, \tag{6.21}$$

where it is clear that the physical degrees of freedom are manifest. However we lost Lorentz invariance. This can be easily seen in the propagator of the transverse part of the photon

$$D_{ij}^{\text{tr}}(x-y) \equiv \langle 0|T[A_i(x)A_j(y)]|0\rangle = \int \frac{d^4p}{(2\pi)^4}\frac{i}{p^2+i\epsilon}\left(\delta_{ij} - \frac{p_i p_j}{|\vec{p}|^2}\right)e^{-ip\cdot(x-y)}\,. \tag{6.22}$$

6.2.2 Lorentz Gauge

We can try to do the same in the Lorentz Gauge, where $\partial_\mu A^\mu = 0$. The equations of motion are then

$$\partial_\mu \partial^\mu A^\nu = 0\,. \tag{6.23}$$

However, we will take a slightly different way. We take as Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\alpha}(\partial_\mu A^\mu)^2\,, \tag{6.24}$$

with arbitrary α. The equations of motion are unchanged and also the quantization is independent of α. We refer to the choice of α again as a *gauge choice* (although formally it is not the same thing). We will use $\alpha = 1$ (*Feynman gauge*), but another common choice is $\alpha = 0$ (*Landau gauge*).
The plan is to quantize the theory and only at the end to impose $\partial_\mu A^\mu = 0$. We will see that the residual gauge symmetry of this theory will be tricky to handle. First, we do the usual definitions,

$$\pi^0 = \frac{\partial \mathcal{L}}{\partial \dot{A}_0} = -\partial_\mu A^\mu$$
$$\pi^i = \frac{\partial \mathcal{L}}{\partial \dot{A}_i} = \partial^i A^0 - \dot{A}^i\,,$$

where now they are all dynamical. We can then easily impose the usual commutation relations

$$[A_\mu(\vec{x}), A_\nu(\vec{y})] = [\pi^\mu(\vec{x}), \pi^\nu(\vec{y})] = 0$$
$$[A_\mu(\vec{x}), \pi_\nu(\vec{y})] = i\eta_{\mu\nu}\delta^{(3)}(\vec{x}-\vec{y})\,,$$

and the usual expansion in terms of creation and annihilation operators,

$$A_\mu(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2|\vec{p}|}} \sum_{\lambda=0}^{3} \epsilon_\mu^\lambda(\vec{p})[a_{\vec{p}}^\lambda e^{i\vec{p}\cdot\vec{x}} + a_{\vec{p}}^{\lambda\dagger} e^{-i\vec{p}\cdot\vec{x}}]$$

$$\pi_\mu(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \sqrt{\frac{|\vec{p}|}{2}} i \sum_{\lambda=0}^{3} \epsilon_\mu^\lambda(\vec{p})[a_{\vec{p}}^\lambda e^{i\vec{p}\cdot\vec{x}} - a_{\vec{p}}^{\lambda\dagger} e^{-i\vec{p}\cdot\vec{x}}] \, .$$

We have now four polarizations 4−vectors instead of two polarization 3−vectors that we had in the Coulomb gauge. We choose the normalization

$$\epsilon^\lambda \cdot \epsilon^{\lambda'} = \eta^{\lambda\lambda'} \, , \tag{6.25}$$

which means

$$\epsilon_\mu^\lambda \epsilon_\nu^{\lambda'} \eta_{\lambda\lambda'} = \eta_{\mu\nu} \, . \tag{6.26}$$

Given the photon 4−momentum $p = (|\vec{p}|, \vec{p})$, we choose ϵ^1 and ϵ^2 to be the transverse:

$$\epsilon^1 \cdot p = \epsilon^2 \cdot p = 0 \, . \tag{6.27}$$

Then ϵ^3 is the longitudinal. For example, if $p \sim (1,0,0,1)$, we have

$$\epsilon^0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} , \epsilon^1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} , \epsilon^2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} , \epsilon^3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} , \tag{6.28}$$

and for other 4−momenta, the polarization vectors will be the appropriate Lorentz transformations.

Now we proceed as usual, getting the commutation relations of the creation and annihilation operators from the fields ones:

$$[a_{\vec{p}}^\lambda, a_{\vec{q}}^{\lambda'}] = [a_{\vec{p}}^{\lambda\dagger}, a_{\vec{q}}^{\lambda'\dagger}] = 0$$

$$[a_{\vec{p}}^\lambda, a_{\vec{q}}^{\lambda'\dagger}] = -\eta^{\lambda\lambda'}(2\pi)^3\delta^{(3)}(\vec{p}-\vec{q}) \, .$$

Here we see something very odd for the timelike part,

$$[a_{\vec{p}}^0, a_{\vec{q}}^{0\dagger}] = -(2\pi)^3\delta^{(3)}(\vec{p}-\vec{q}) \, . \tag{6.29}$$

The fact that we have this minus sign means that if we start with our vacuum

$$a_{\vec{p}}^\lambda|0\rangle = 0 \, , \tag{6.30}$$

and we create a one-particle state as

$$|\vec{p}, \lambda\rangle = a_{\vec{p}}^{\lambda\dagger}|0\rangle \, , \tag{6.31}$$

we get that, if we take the timelike part normalization,

$$\langle\vec{p}, 0|\vec{q}, 0\rangle = \langle 0|a_{\vec{p}}^0 a_{\vec{q}}^{0\dagger}|0\rangle = -(2\pi)^3\delta^{(3)}(\vec{p}-\vec{q}) \, , \tag{6.32}$$

it has negative norm! This makes no sense but we did not use our condition $\partial_\mu A^\mu = 0$ yet. We see now that this condition will cut out the negative norm states and cut the physical polarizations down to two. We work in Heisenberg picture, so that $\partial_\mu A^\mu$ is an operator equation. Then we can impose this condition is several ways, here listed in order of increasingly weak way of doing this:

- We can impose $\partial_\mu A^\mu$ as an equation on operators. However this would not work because the commutation relations would not be obeyed for $\pi^0 = -\partial_\mu A^\mu$.

- We can impose this on the Hilbert space. We can imaging to divide the Hilbert states in bad and good, where the bad ones decouple. Then, in order to define the good states we can try imposing that a good state should be such that
$$\partial_\mu A^\mu |\Psi\rangle = 0 \,. \tag{6.33}$$
Again this would not work because it is too strong. For example, say we decompose $A_\mu = A_\mu^+(x) + A_\mu^-(x)$ with
$$A_\mu^+(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2|\vec{p}|}} \sum_{\lambda=0}^{3} \epsilon_\mu^\lambda a_{\vec{p}}^\lambda e^{-ip\cdot x}$$
$$A_\mu^-(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2|\vec{p}|}} \sum_{\lambda=0}^{3} \epsilon_\mu^\lambda a_{\vec{p}}^{\lambda\dagger} e^{ip\cdot x} \,,$$
then $A_\mu^+|0\rangle = 0$ but $\partial^\mu A_\mu^-|0\rangle \neq 0$. Not even the vacuum would be a good state with this definition.

- This attempt will be the correct one. We can ask that physical state are defined by
$$\partial^\mu A_\mu^+ |\Psi\rangle = 0 \,, \tag{6.34}$$
which ensures
$$\langle \Psi' | \partial_\mu A^\mu |\Psi\rangle = 0 \,, \tag{6.35}$$
or, in other words, that $\partial_\mu A^\mu$ has vanishing matrix elements between physical states. This is called *Gupta-Bleuler* condition.

With this last choice, have we removed these negative norm states? The answer is not yet, but almost.
Let's consider a basis of states for the Fock space. Any element of this basis can be written as $\Psi = |\psi_T\rangle|\phi\rangle$, where ψ_T contains only transverse photons (so created by $a_{\vec{p}}^{1,2\dagger}$). The Gupta-Bleuler condition requires
$$(a_{\vec{p}}^3 - a_{\vec{p}}^0)|\phi\rangle = 0 \,, \tag{6.36}$$

so that physical states should contain combinations of timelike and longitudinal photons. In general $|\phi\rangle$ will be a linear combination of states $|\phi_n\rangle$ containing n pairs of timelike and longitudinal photons,

$$|\phi\rangle = \sum_{n=0}^{\infty} C_n|\phi_n\rangle\,, \tag{6.37}$$

where $|\phi_0\rangle = |0\rangle$ is simply the vacuum. It is rather easy to see that, although the Cupta-Bleuler condition does not remove all negative norm states, all the remaining states involving longitudinal and timelike photons have zero norm, as

$$\langle\phi_m|\phi_n\rangle = \delta_{n0}\delta_{m0}\,. \tag{6.38}$$

So this means that the inner product of the Hilbert space of the physical states $\mathcal{H}_{phys}$ is positive semi-definite. We then have to deal with this states with zero norm.
The idea is to treat these zero norm states as gauge equivalent of the vacuum: two states that differ only in their timelike and longitudinal components are said to be physically equivalent. This however means that we should show that they give the same expectation values for all physical states. It is rather easy to show that. We solved all the problems!

Propagators. Finally let's compute the propagators in Lorentz gauge. They are given by

$$\langle 0|T[A_\mu(x)A_\nu(y)]|0\rangle = \int \frac{d^4p}{(2\pi)^4}\frac{-i\eta_{\mu\nu}}{p^2+i\epsilon}e^{-ip\cdot(x-y)}\,, \tag{6.39}$$

and it is surely a lot nicer than the propagator in Coulomb gauge as it is Lorentz invariant. Even without fixing α in the Lagrangian we still can find

$$\langle 0|T[A_\mu(x)A_\nu(y)]|0\rangle = \int \frac{d^4p}{(2\pi)^4}\frac{-i}{p^2+i\epsilon}\left(\eta_{\mu\nu}+(\alpha-1)\frac{p_\mu p_\nu}{p^2}\right)e^{-ip\cdot(x-y)}\,. \tag{6.40}$$

6.3 Coupling to Matter

We can now build and interacting theory of light and matter. We can write something like

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - j^\mu A_\mu\,, \tag{6.41}$$

with j^μ some function of the matter fields. We then have the equation of motion

$$\partial_\mu F^{\mu\nu} = j^\nu\,, \tag{6.42}$$

so that we require that j^μ should be a conserved current,

$$\partial_\mu j^\mu = 0\,. \tag{6.43}$$

6.3.1 Coupling to Fermions

Remembering that the Dirac Lagrangian has an internal symmetry $\psi \to e^{-i\alpha}\psi$ and $\bar{\psi} \to e^{i\alpha}\bar{\psi}$, which gives rise to the conserved current $j^\mu_V = \bar{\psi}\gamma^\mu\psi$, we can try building our theory coupled to fermions with the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\partial\!\!\!/ - m)\psi - e\bar{\psi}\gamma^\mu A_\mu\psi\,. \tag{6.44}$$

We have seen that for the free theory, it was fundamental that the theory had a gauge symmetry. Is it still the case? The answer is yes. In order to see this, let's rewrite the Lagrangian as

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(iD\!\!\!\!/ - m)\psi\,, \tag{6.45}$$

where $D_\mu\psi = \partial_\mu\psi + ieA_\mu\psi$ is called the *covariant derivative*. This Lagrangian is invariant under the gauge transformation

$$A_\mu \to A_\mu + \partial_\mu\lambda \quad \text{and} \quad \psi \to e^{-ie\lambda}\psi\,, \tag{6.46}$$

for an arbitrary function $\lambda(x)$. The fact that this is a symmetry is not trivial, for example let's look at how the covariant derivative transforms

$$\begin{aligned} D_\mu\psi &= \partial_\mu\psi + ieA_\mu\psi \\ &\to \partial_\mu(e^{-ie\lambda}\psi) + ie(A_\mu + \partial_\mu\lambda)(e^{-ie\lambda}\psi) \\ &= e^{-ie\lambda}D_\mu\psi\,. \end{aligned}$$

Basically it only takes a phase under the gauge transformation. From this we can easily see that the theory has indeed a gauge symmetry.

Electric charge The coupling e has the interpretation of the *electric charge* of the ψ particle. We therefore have the total charge

$$Q = e\int d^3x\bar{\psi}(\vec{x})\gamma^0\psi(\vec{x})\,, \tag{6.47}$$

which is, after quantization,

$$Q = e\int\frac{d^3p}{(2\pi)^3}\sum_{s=1}^{2}(b^{s\dagger}_{\vec{p}}b^s_{\vec{p}} - c^{s\dagger}_{\vec{p}}c^s_{\vec{p}})\,. \tag{6.48}$$

In QED, we usually write the electric charge in terms of the *fine structure constant*

$$\alpha = \frac{e^2}{4\pi\hbar c} \approx \frac{1}{137}\,. \tag{6.49}$$

We need to stop for a second to understand a subtlety. We said that gauge symmetries are different from global symmetries in that they arise from a redundancy of our description and , as such, they do not give conserved currents.

In this case, however, we got the electric charge as a result of a gauge symmetry. Why? The reason is that among all the gauge symmetries that one can get changing the function $\lambda(x)$, there is also the one obtained with $\lambda(x) = \text{constant}$ which is a real symmetry of the theory. More generally, the real symmetries that arise from a gauge symmetry are such that $\lambda(x) \to \alpha = \text{constant}$ as $x \to \infty$.

6.3.2 Coupling to Scalars

For a real scalar field we have no suitable conserved current so we cannot couple a real scalar field with a gauge field. For a complex scalar field φ we have a symmetry $\varphi \to e^{-i\alpha}\varphi$. We can try to use its conserved current to couple φ to the gauge field,

$$\mathcal{L}_{\text{int}} = -i((\partial_\mu \varphi^*)\varphi - \varphi^* \partial_\mu \varphi)A^\mu \,. \tag{6.50}$$

However this does not work because

- The theory is no longer gauge invariant.
- The current that we coupled to A^μ depends on $\partial_\mu \varphi$. This means that the current associated to the symmetry will have a contribution like $j^\mu A_\mu$ that is not consistent.

We can solve both the problems with the covariant derivative

$$\mathcal{D}_\mu \varphi = \partial_\mu \varphi + ieA_\mu \varphi \,, \tag{6.51}$$

which transforms $\mathcal{D}_\mu \varphi \to e^{-ie\lambda}\mathcal{D}_\mu \varphi$ under a gauge transformation. So we can easily get the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \mathcal{D}_\mu \varphi^* \mathcal{D}^\mu \varphi - m^2|\varphi|^2 \,. \tag{6.52}$$

This is very general: if we have a $U(1)$ symmetry that we want to couple to a gauge field, we can do that replacing all derivatives by suitable covariant derivatives. This is called *minimal coupling*.

6.4 QED

The QED Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\not{D} - m)\psi \,, \tag{6.53}$$

where $D_\mu = \partial_\mu + ieA_\mu$.

6.5 Feynman Rules

Let's now work out the Feynman rules for QED. For internal lines and vertices we have

- Vertex: $-ie\gamma^\mu$
- Photon Propagator:
$$-\frac{i\eta_{\mu\nu}}{p^2 + i\epsilon} \tag{6.54}$$
- Fermion Propagator:
$$\frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} \tag{6.55}$$

For external lines we have

- Photons: We add a polarization vector $\epsilon^\mu_{in}/\epsilon^\mu_{out}$ for incoming/outgoing photons.
- Fermions: We add a spinor $u^r(\vec{p})/\bar{u}^r(\vec{p})$ for incoming/outgoing fermions. We add $\bar{v}^r(\vec{p})/v^r(\vec{p})$ for incoming/outgoing anti-fermions.