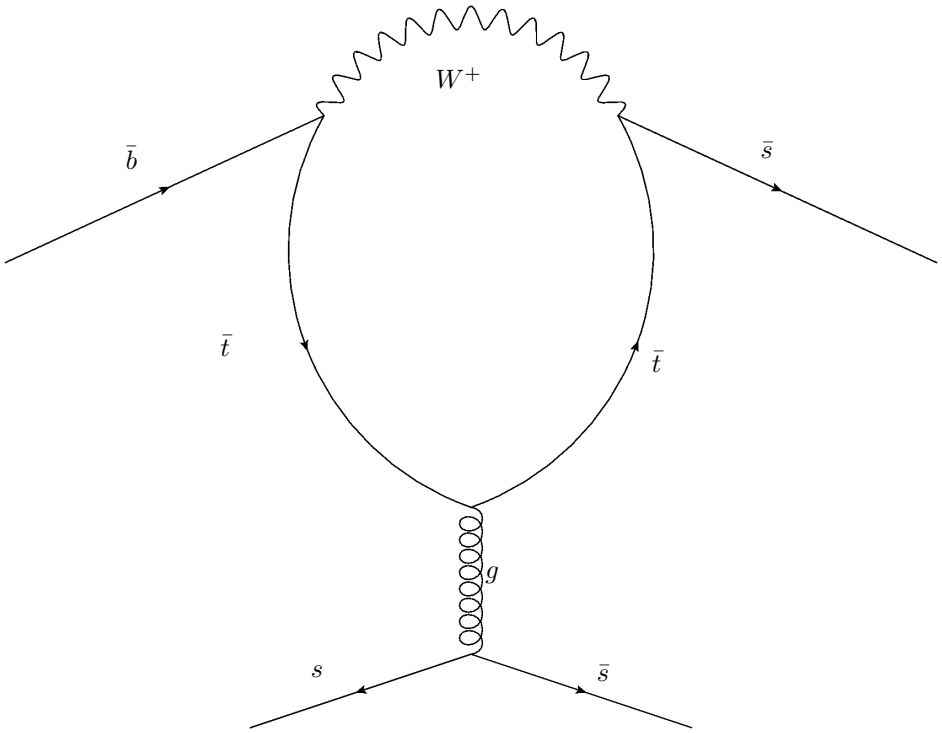
EPFL

Quantum Field Theory



Lecture notes

Based on the course given by Riccardo Rattazzi

First redaction, Johnny Espin, 2012

Latest redaction, Heloise Allaman, Eren Firat, Alessandro Fumagalli, Irwan Le Dorze, Filippo Nardi, Riccardo Rattazzi, 2024

Contents

Part I

Foundations and Matter

Chapter 1

Introduction

Quantum field theory (QFT) is the theoretical "apparatus" that is needed to describe how nature works at its most fundamental level, that is to say at the shortest distances we have explored. In such domain physics is described by elementary particles and by their interactions, and QFT beautifully accounts for that. It also turns out that the notion of *Fundamentality* is here paired with the notion of *Simplicity*. As we shall see towards the end of this course, after having digested the necessary set of mathematical and physical concepts ($\equiv$*formalism*), the laws of particle physics can be written in a few lines with absolute precision and greatest empirical adequacy. In a sense particle physics is orthogonal to the sciences that deal with complexity.

1.1 What is quantum field theory?

Technically, quantum field theory is the application of quantum mechanics to dynamical systems of fields, very much like basic quantum mechanics concerns the quantisation of dynamical systems of particles. Therefore, while quantum mechanics deals with mechanical systems with a finite number of degrees of freedom, quantum field theory describes the quantum systems with infinitely many degrees of freedom. Specifically this course is devoted to relativistic QFT. Relativistic QFT explains the existence of particles and describes their mutual interactions. The fact that nature at its most basic level consists of particles can thus be viewed merely as a consequence of relativistic QFT. The domains of application of the latter in modern physics are quite broad: from the study of collisions among elementary particles in high energy accelerators, to early Universe cosmology. For instance, the primordial density fluctuations that later gave rise to structures like galaxies, the origin of dark matter or black-hole radiation are all described by relativistic QFT. Nevertheless there are also applications of quantum field theory to non-relativistic systems, in particular in condensed-matter physics: superfluidity, superconductivity, quantum Hall effect, ...

1.2 Why quantum field theory?

We just outlined what QFT describes, but why do we need a field theory to describe particles? Is this a necessity?

It basically is: relativistic QFT is the only way to reconcile ordinary quantum mechanics with special relativity. We can understand this necessity with various intuitive arguments.

I. Special relativity implies the existence of a limiting speed. This fact can be seen as a consequence of the request that there exist a relativistically invariant notion of causality. That is, given two events A and B, all observers in relative constant motion should agree on their causal ordering. For instance if A causes B according to one observer, then it should be so for all the others. In Newtonian physics causal ordering is simply, and intuitively, determined by temporal ordering. The extension of the notion to the relativistic domain is however subtle, because the temporal ordering of two events A and B is generally not invariant under Lorentz transformations, unless their separation is timelike, i.e. $c^2(t_A-t_B)^2-(\vec{x}_A-\vec{x}_B)^2 \geq 0$. Then in relativity temporal ordering can provide a sensible

basis for causal ordering only if causation is forbidden among space-like separated events. Stated differently: in relativity, two events can affect one another only if their separation is time-like. This is equivalent to saying that information must propagate within the light-cone, or, more concretely, that it cannot travel faster than light.

In particular a consequence of the above is that instantaneous interactions between two particles separated in space are not possible: one needs a medium (a *field*) to "propagate" interactions. Intuitively the picture is similar to the one offered by two boats separated by a distance on the surface of a lake. The boats affect one another (interact) through the waves they generate in the water. The water surface, or better its deformation, represents in that case the field. The water surface is however only an effective macroscopic description of the complex microscopic dynamics of water molecules. The corresponding field is not fundamental. The simplest example of a more fundamental and also relativistic field is instead given by the electromagnetic field.

The argument we just made purely relies on classical physics. It states the necessity of fields in order to generate interactions, but does not imply that all dynamical variables should be field variables. Indeed, in classical electrodynamics, the fields' sources, the charged particles, are ordinary mechanical point-like objects and are thus not associated to fields. It however turns out that, when taking quantum mechanics into consideration, the only way things can be made compatible (*consistent*) is if matter as well is described by fields at its most fundamental level. Two facts underly this result. The first is that fields, when treated quantum mechanically, do produce a spectrum of excitations that is interpretable in terms of particles: quantum fields can explain the existence of particles. That was realized very early on in the development of quantum theory in a paper of 1926 by Born, Heisenberg and Jordan. The second fact is that the description of a quantum system in terms of a finite number of point-particles leads to severe inconsistencies when married to relativity. These principally have to do with matter stability and/or causality. Let us review them.

II. The Schrödinger equation describing a free particle in ordinary quantum mechanics

$$\left(+i\hbar\partial_t + \frac{\hbar^2\nabla^2}{2m}\right)\psi(t,x) = 0 \tag{1.1}$$

satisfies invariance under Galilei's relativity but not under Einstein's relativity (the equation is first order in time and second order in spacial coordinate, not a good start for Lorentz tranformations which treat time and space on similar footing). Its most obvious generalization satisfying Lorentz invariance is given by the Klein-Gordon equation

$$\left(\hbar^2\partial_t^2 - c^2\hbar^2\nabla^2 + m^2c^4\right)\psi(t,x) \equiv \left(\hbar^2\Box + m^2c^4\right)\psi(t,x) = 0\,. \tag{1.2}$$

The D'Alembertian operator $\Box = \partial_t^2 - c^2\nabla^2$ ensures invariance, but the price to pay is that the equation is of second order in the time derivative: the set of solutions at fixed 3-momentum is therefore doubled. Consequently the energy spectrum is also doubled with respect to the non-relativistic case

$$\text{Schroedinger}: \;\; E = \frac{\hbar^2k^2}{2m} \quad \Rightarrow \quad \text{Klein}-\text{Gordon}: \;\; E = \pm c\sqrt{m^2c^2 + \hbar^2k^2}\,. \tag{1.3}$$

In particular now the spectrum extends form $E = -\infty$ to $E = +\infty$. The existence of unbounded negative energies, or equivalently the absence of a ground state, leads to catastrophic instabilities as soon as interactions between particles are turned on. For instance, energy and momentum conservation do not forbid a system of two interacting particles to evolve towards infinite and opposite energy and momenta, a phenomenon we do not observe in nature! There exist other generalizations of the Schroedinger equation, where spin plays a role. but they all encounter the same problem. The Dirac equation, which we will encounter later, is an example of such relativistic wave equations.[1]

III. Starting from the Klein-Gordon equation one can attempt at disposing of the extra negative energy states. Indeed the equation can formally be written as

$$\left(\hbar^2\Box + m^2c^4\right)\psi(t,x) \equiv (-i\hbar\partial_t - \sqrt{m^2c^4 - \hbar^2c^2\nabla^2})(i\hbar\partial_t - \sqrt{m^2c^4 - \hbar^2c^2\nabla^2})\psi(t,x) = 0 \tag{1.4}$$

where the presence of the two factors, each involving a single power of ∂_t, is also directly associated to the existence of a double set of solution, one of which having negative energy. We can then imagine writing instead an equation involving only one factor in order to eliminate the unwanted solutions

$$(i\hbar\partial_t - \sqrt{m^2c^4 - \hbar^2c^2\nabla^2})\psi(t,x) = 0\,. \tag{1.5}$$

[1] You may wonder why will we bother with such equations if they are associated to a pathology. The answer lies in the interpretation: these equations should not be interpreted as describing the evolution of a wave function, but rather the evolution of a field. With the latter interpretation not only do they make sense but they also play a central role.

Indeed fixing 3-momentum $-i\hbar\vec{\nabla} \to \hbar\vec{k} \equiv \vec{p}$, and thus replacing $-\hbar^2\nabla^2 \to \hbar^2 k^2$, this equation gives

$$E = +c\sqrt{m^2c^2 + \hbar^2 k^2} > 0 \tag{1.6}$$

We must however immediately recognize that eq. (1.5) is not strictly speaking a differential equation, but an integro-differential one. That is because of the presence of ∇^2 under the square root, which makes the corresponding operator act non-locally in position space. This is a "novelty" compared to both Schroedinger equations and its relativistic generalizations. A second comment is that eq. (1.5) is indeed Lorentz covariant, as one can check. That should not come has a surprise though. $H(\vec{p}) = \sqrt{m^2c^4 + c^2p^2}$ is the Hamiltonian for a relativistic particle of momentum p, and the second term in eq. (1.5) is precisely $H(-i\hbar\vec{\nabla})$, consistent with the quantum mechanical relation $\vec{p} = -i\hbar\vec{\nabla}$.

The big problem of eq. (1.5) is that it implies a violation of causality. That is seen by considering the time evolution of the probability amplitude of a particle that is fully localized at position x at time $t = 0$

$$\langle y|e^{-i\hat{H}t/\hbar}|x\rangle = A(x \to y, t)\,. \tag{1.7}$$

Using in the momentum eigenstate basis this can be written as

$$\int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot(\vec{y}-\vec{x})/\hbar} e^{-ic\sqrt{m^2c^4+p^2}\,t/\hbar} \tag{1.8}$$

an expression that can be shown to be non-zero when the space-time event (t, y) sits outside the light cone of $(0, x)$ ($c^2t^2 - (x-y)^2 < 0$). More precisely (see Exercise in Series 1), outside the light-cone one finds

$$A(x \to y, t) = e^{-\frac{mc}{\hbar}\sqrt{(x-y)^2 - c^2t^2}}\, f(x-y, t) \tag{1.9}$$

with $f(x-y,t)$ a rational function. The dependence of the amplitude on the coordinates also marks the appearance of a new length scale, the Compton wavelength $\lambda_C \equiv \hbar/mc$, associated to the particle mass through the fundamental constants of quantum mechanics and relativity. Now, for instance, at $t = 0$ the probability for instantaneous propagation decreases exponentially with $|x-y|$ and for elementary particles, like the electron or the proton, it becomes very tiny already at atomic distances, and even more so at macroscopic distances. Yet it does not vanish. Moreover the probability is sizeable at small separation, that is for $|x-y| \lesssim \lambda_C \equiv \hbar/mc$, which seems unacceptable. This result indicates that the relativistic notion of causality is incompatible with the intepretation of particle coordinates as observables and of the wave function as a probability density.

The same conclusion can be reached by a less formal argument which takes experimental observation into consideration and offers a more physical interpretation of the occurrence of the length scale $\hbar/mc$ below which the point particle description is totally undefendable.

IV. The basic remark is that the famous Einstein relation $E = mc^2$ implies that energy and mass are equivalent. Because of that, in a relativistic process, we have no right to exclude that part of the energy is used to create extra particles, as long other coservation laws such as the conservation of electric charge are respected. Quantum mechanics only makes that possibility more urgent[2] and relates to the fundamental properties of matter. Consider indeed the uncertainty principle with relativity for a particle of mass m:

$$\underline{\text{Relativity}}: \quad \Rightarrow \quad E = c\sqrt{p^2 + m^2c^2} \tag{1.10}$$

$$\underline{\text{Quantum mechanics}}: \quad \Rightarrow \quad \Delta x \Delta p \gtrsim \hbar \tag{1.11}$$

so that if:

$$\Delta x < \frac{\hbar}{mc} = \lambda_{Compton} \tag{1.12}$$

[2]The essence of Quantum Mechanics is nicely captured by the so-called *totalitarian principle*, first stated in writing by Murray Gell-Mann: "Everything not forbidden is compulsory". The Feynman path integral formulation of Quantum Mechanics puts the totalitarian principle up front.

then:

$$\boxed{\Delta E > mc^2} \tag{1.13}$$

In other words, if we try to localise a particle to a distance shorter than its characteristic Compton wavelength, it becomes relativistic and the excess energy can lead to the production of extra particles! The production of these extra particles implies that the one-particle states of ordinary quantum mechanics stop making sense for localisations smaller than $\Delta x \simeq \lambda_{Compton}$. Empirically, it is a fact that particles are indeed created and destroyed in fundamental processes.

The last argument can be made more concrete by considering the process of measurement of the position of an elementary particle, say an electron. In order to measure the position of the electron we need to "see" it, so we must have it interact with a light source. Light consists of the excitations of the electromagnetic field, a field whose quantum description we can assume to know. A result from the quantization of the electromagnetic field is that waves of wavelength λ consist of particles, the photons (normally indicated by γ), carrying momentum $p_\gamma = \hbar/\lambda$ and energy $E_\gamma = cp = \hbar c/\lambda$. Now, the precision in the measurement in the electron position Δx is limited by the finite wavelength of the light we employ to detect it: $\Delta x \gtrsim \lambda$. The energy of the photons will therefore grow as Δx is made smaller: $E_\gamma = \hbar c/\lambda \gtrsim \hbar c/\Delta x$. As a consequence of that, for small enough Δx, the photon will be sufficiently energetic to allow the production of additional particles. For instance one finds that for $E_\gamma > 4m_e c^2$, with m_e the mass of the electron, the energy of the photon is high enough to lead to the production of an additional electron positron pair ($e^+ - e^-$). For such highly energetic photons the single electron picture stops making sense! If we try to make the measurement of the position of the electron more precise by using such short wavelength (and therefore energetic!) photons, we produce instead a different state with 3 particles (2 electrons and 1 positron). We thus conclude there exists a minimal length scale, indeed the Compton wavelength, down to which the single electron picture makes sense:

$$4m_e c^2 > E_\gamma = \frac{\hbar c}{\lambda} \gtrsim \frac{\hbar c}{\Delta x} \quad \Rightarrow \quad \Delta x \gtrsim \frac{\hbar c}{m_e} \equiv \lambda_C^e \,. \tag{1.14}$$

with $\lambda_C^e = 4 \times 10^{-11}$cm the Compton wavelength of the electron. λ_C^e separates the domain of non relativistic quantum mechanics from the domain of relativistic quantum field theory.

1.3 Units of measure

In order to quantitatively describe physical phenomena we must choose units of measure. In principle we will have an independent unit for each different quantity. However physical laws establish relations among physical quantities, which allow to espress the units for one quantity in terms of the units of the others. For instance $F = ma$ allows to express the units of force in terms of the units of mass and acceleration. The laws of classical mechanics and classical electromagnetism allow to reduce the set of independent units to just 3. In the C.G.S. system these are:

$$\underline{\text{Length}}: \quad [L] \quad \rightarrow \quad cm \tag{1.15}$$

$$\underline{\text{Time}}: \quad [t] \quad \rightarrow \quad s \tag{1.16}$$

$$\underline{\text{Mass}}: \quad [m] \quad \rightarrow \quad g \tag{1.17}$$

For instance, force, energy and electric charge are not fundamental quantities and can be expressed in terms of the above three:

$$\underline{\text{Force}}: \qquad F = ma \quad \Rightarrow [F] = \frac{[m][L]}{[t]^2} \quad \rightarrow \; dyne = \frac{g \cdot cm}{s^2} \tag{1.18}$$

$$\underline{\text{Energy}}: \qquad E = \frac{m}{2}v^2 \quad \Rightarrow [E] = \frac{[m][L]^2}{[t]^2} \quad \rightarrow \; erg = \frac{g \cdot cm^2}{s^2} \tag{1.19}$$

$$\underline{\text{Electric charge}}: \qquad F_{el} = \frac{q^2}{r^2} \quad \Rightarrow [q] = \frac{[m]^{1/2}[L]^{3/2}}{[t]} \quad \rightarrow \; franklin = \frac{g^{1/2} \cdot cm^{3/2}}{s} \tag{1.20}$$

But the C.G.S. system is unaware of the further relations dictated by Quantum Mechanics and Relativity. In particular, Relativity implies the speed of light is a fundamental constant. Relativity thus provides one additional relation among the fundamental C.G.S. units (as well as among all the other derived units):

$$v_{light} = c = \frac{[L]}{[t]} \quad \Rightarrow \quad [L] = c[t], \;\; [E] = c[p] = c^2[m] \tag{1.21}$$

We may say that relativity, through the introduction of a fundamental speed c, unifies space and time as well as mass, momentum and energy.

Quantum Mechanics, through the Indetermination Principle $\Delta p \Delta L = \hbar$, introduces instead a relation among length and momentum controlled by the Planck constant

$$[p] = \frac{\hbar}{[L]} \quad \Rightarrow \quad \frac{[m][L]}{[t]} = \frac{\hbar}{[L]} \tag{1.22}$$

$$\Rightarrow \qquad [m] = \hbar \frac{[t]}{[L]^2} = \frac{\hbar}{c^2}\frac{1}{[t]} \tag{1.23}$$

By QM, mass can be expressed in terms of length and time, and with the further use of relativity, mass is made equivalent to inverse length, or inverse time.

In the C.G.S. system we have

$$c = 2.99 \times 10^{10} \; \frac{cm}{s}, \qquad \hbar = 1.05 \times 10^{-27} \; erg \cdot s \,, \tag{1.24}$$

two “huge" numbers whose origin is linked to the contingencies of human life and human history, which made centimeters, seconds and grams very convenient units for everyday affairs. For the purpose of studying the implications of relativity and Quantum Mechanics it is instead clearly more convenient to choose a system of units where c and $\hbar$ are as simple as possible, that is

$$\boxed{c = 1 = \hbar} \tag{1.25}$$

This choice defines what is called a Natural System of Units, which satisfy

$$[t] = [L] = \frac{1}{[m]} = \frac{1}{[p]} = \frac{1}{[E]} \,. \tag{1.26}$$

The natural choice $c = \hbar = 1$ does not fully specify the system: there remains just one unit to be picked, which can be anyone we wish, a length or an energy. In particle physics it is customary to pick a convenient unit of energy. The standard choice is given by the *Gigaelectronvolt*, $GeV = 10^9 \; eV$. Using $1 \; eV = 1.6 \times 10^{-12} \; erg$, we can for instance compute as an exercise the following relations:

Unit	QFT system to CGS
Length	$GeV^{-1} = 1.97 \times 10^{-14}\ cm$
Time	$GeV^{-1} = 6.58 \times 10^{-25}\ s$
Mass	$GeV = 1.78 \times 10^{-24}\ g$
Action	$GeV^0 \quad (\hbar = 1)$
Velocity	$GeV^0 \quad (c = 1)$

(1.27)

This little discussion, in particular eq. 1.26, offers also a persepective as to why it is equivalent to talk either about *high-energy* physics or about *short-distance* physics. High-energy accelerators are truly huge microscopes testing the short distance behaviour of matter and forces!

1.4 Overview of particle physics

The goal of the course is obviously to learn the basics of QFT, nonetheless before that it would be helpful to receive an overview of present understanding of the micro world.

There are four well known fundamental forces:

1. The electromagnetic force which is responsible for the stability of atoms, namely for *binding* nuclei and electrons into atoms. This is a long range force.
2. The weak force: a short-range force ($L \approx 10^{-16}\ cm$) which is responsible for the beta-decay of nuclei, neutrons and for muon-decay:

$$n \to p + e^- + \bar{\nu}_e \tag{1.28}$$

$$\mu^- \to e^- + \bar{\nu}_e + \nu_\mu \tag{1.29}$$

3. The strong force is responsible for the existence of nuclei (for their stability). Also, for instance, more than 99% of the mass of nuclei is due to the strong force alone. The latter explains most of our mass. The strong force is also short-ranged ($L \approx 10^{-13}\ cm$)
4. The gravitational force: it is very weak, however it is long ranged and universally attractive. For these reasons, gravity plays nonetheless an essential role in our life. In order to understand how weak gravity is compared to the other forces, we can calculate the ratio between the gravitational and electromagnetic forces' amplitude between two protons:

$$\frac{F_G}{F_E} \simeq \frac{G_N m_p^2}{e^2} \sim 10^{-40} \tag{1.30}$$

We see immediately that the difference is huge!

As we shall learn, in QFT forces are associated to fields and thus to particles. The electromagnetic force is mediated by the photon γ, the weak force by the $W^\pm$ and Z^0 and the strong force by eight gluons g. These particles are all helicity ± 1 vector bosons. On the other hand, gravity is mediated by the graviton which is a helicity ± 2 boson. This is what makes gravity so different from the other forces. In addition, there is a mysterious "Higgs" force that remains to be studied. It might be mediated by one (or more) scalar particle (spin 0).

We saw which forces are present in Nature as far as we know. They all are mediated by bosonic particles. All other fundamental particles known so far are *fermions* with spin 1/2 and come into two classes:

1. The leptons: they only feel the electromagnetic and weak forces. They come into three distinct generations.

	1^{st}	2^{nd}	3^{rd}
$Q=-1$	electron e^- $m_e = 0.511\ MeV$	muon μ^- $m_\mu = 105\ MeV$	tau τ^- $m_\tau = 1.777\ GeV$
$Q=0$	ν_e	ν_μ	ν_τ

(1.31)

The masses of the neutrinos are non-zero but very small ($m_\nu \lesssim 1\ eV$). In addition, every particle in this table comes in pairs with its antiparticle.

2. The quarks: they also feel the strong force. Because of this, quarks are strongly bound into *hadrons* and are always confined into these ones. We have basically two types of hadrons: the *baryons* (like the proton) made up of three quarks and the *mesons* (like the pions) made up of a quark and an antiquark. They are also organised in three families.

	1^{st}	2^{nd}	3^{rd}
$Q=2/3$	up u $m_u \simeq 1.5-3\ MeV$	charm c $m_c \simeq 1.25 \pm 0.1\ GeV$	top t $m_t \simeq 171.4 \pm 2\ GeV$
$Q=-1/3$	down d $m_d \simeq 3-7\ MeV$	strange s $m_s \simeq 95 \pm 25\ MeV$	bottom b $m_b \simeq 4.20 \pm 0.07\ GeV$

(1.32)

We should understand how these observational facts fit into a coherent theoretical picture. More pragmatically, how well we describe the observed particles and their natural interactions. *Particle physics*, to a beginner, often gives the impression of being all about "finding particles and classifying them" much like botanics. Nothing could be more wrong than that! What is being sought for are indeed the fundamental laws of Nature. Underlying the above empirical fact, there is a well developed theoretical understanding; in certain aspects, our understanding is truly remarkable, while open problems and some mystery remain. What we understand well is the structure of the electromagnetic, weak and strong forces, and quarks and leptons are successfully described by one QFT called the Standard Model (SM). What do we mean by understanding something? Understanding means also being able to describe and predict: for example, some of the existing particles (the weak bosons the charm and the top quarks) have been predicted from logical (or mathematical) consistency of the theory before their experimental discovery. Also, the prediction of some observables is made and verified to a level of accuracy which is unmatched in any other field. One of the best results of our QFT formalism is the derivation of quantum effects in electrodynamics. One example is the study of the magnetic moment of a particle. The latter is proportional to the spin and is given by:

$$\vec{\mu} = g\frac{e}{2m}\vec{s} \tag{1.33}$$

where g is the gyromagnetic ratio. Classically, we expect $g = 1$; for a relativistic electron in Dirac's theory, we have $g = 2$. In QED, however, we can compute the digression from the value of 2:

$$\frac{g-2}{2} = \frac{\alpha}{2\pi} + c_2\left(\frac{\alpha}{2\pi}\right)^2 + c_3\left(\frac{\alpha}{2\pi}\right)^3 + \dots \tag{1.34}$$

The coefficients c_i are predicted by QED and are of order one. Alpha (α) is the fine structure constant of the theory and is given by:

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \sim \frac{1}{137} \tag{1.35}$$

The digression on g is measured experimentally with very high accuracy:

$$\left(\frac{g-2}{2}\right)_{exp} = 1159652180.85(76) \times 10^{-12} \tag{1.36}$$

Once this quantity is measure, we can use it to compute α with precision and then use this value to test the theory with other observable:

$$\left(\frac{g-2}{2}\right)_{exp} \rightarrow \alpha^{-1} = 137.035999705(36) \tag{1.37}$$

and for example, from atomic clocks (Rb and Cs):

$$\alpha^{-1}(Rb) = 137.03599878(91) \tag{1.38}$$

These results are a great confirmation that QFT is the correct description of the subatomic world. The success of QFT and of the SM make the open problems in the theory only more exciting! Here is a list if what we understand less well in decreasing order of understanding:

1. The SM necessarily predicts the existence of a fifth force to explain the origin of the mass of the weak bosons as well as of the quarks and leptons. However, there are various possibilities in nature for this necessary force. The simplest is that it is mediated by a boson of spin zero, a scalar particle called the Higgs boson. However the presence of this boson has not been experimentally confirmed[3] and moreover many physicists question this simple picture proposing alternative theories. This is a very exciting time because before the end of this course, the Large Hadron Collider (LHC) at CERN should have had announced its discovery or it dismissal.
2. Why are there three families of fermions?
3. What is the reason for the big ratios of masses?
4. What is the role of gravity? As far as we understand, gravity makes sense as a quantum field theory only at sufficiently low energy. A full quantum description is however missing. Possibly gravity will require to go beyond QFT (*e.g.* String theory or other alternative Quantum Gravity theories). The problem of gravity is not necessarily urgent. Indeed, the needed energy cannot be reached in any imaginable laboratory. It is however crucial to develop a theory of the very early Universe to answer question such as "How did the Universe start?" or "Was it inevitable?". During our course, we shall neglect gravitational perturbations and work with a static Minkowski metric:

$$g_{\mu\nu} \rightarrow \eta_{\mu\nu} = \text{diag}(+1, -1, -1, -1) \tag{1.39}$$

After this short introduction (for a very broad subject), we shall start learning its technicalities.

[3] As of when these notes were typed.

Chapter 2

Classical Field Theory

2.1 Lagrangian mechanics

In the Newtonian formulation of Mechanics, the time evolution of a system with N degrees of freedom[1] $q_1, \ldots, q_N$ is determined by N second order differential equations[2]

$$\ddot{q}_a = f_a(q, \dot{q}) \qquad a = 1, \ldots, N \tag{2.1}$$

and by $2N$ initial conditions, for q_a and $\dot{q}_a$, at some initial time t_i. In other words, given $q_a(t_i)$, $\dot{q}_a(t_i)$ and eq. (2.1) the time evolution of the system, the trajectory $q_a(t)$, is uniquely fixed. The notion of determinism[3] is explicit in this formulation of the laws of motion.

As it turns out, the laws of Mechanics of *genuinely mechanical systems* (that is, systems without dissipation) admit the alternative Lagrangian formulation. In the Lagrangian formulation the inputs are the coordinates at two different times,

$$q_a(t_i) = q_a^i, \qquad q_a(t_f) = q_a^f \,. \tag{2.2}$$

and the physical trajectories are determined through a variational principle, the Least Action Principle. The way this works, it that, given any trajectory $q_a(t)$ satisfying eq. (2.2), one considers a certain quantity, the action S, given by

$$S\left[q\right] \equiv \int_{t_i}^{t_f} L\left(q, \dot{q}\right) dt \tag{2.3}$$

with the *Lagrangian* $L(q, \dot{q})$ a specific function of coordinates and velocities encapsulating the dynamics of the given system. The action is a *functional*: a real valued function of the trajectory $q_a(t)$ followed by the system between t_i and t_f[4]. The Least Action Principle then consists in the following remarkable realization: the physical trajectories $\bar{q}_a(t)$, i.e. those solving the Newtonian formulation of the dynamics, coincide with the stationary points of $S[q]$

$$\delta S\left[\bar{q}\right] = 0 \tag{2.4}$$

where δS is the linearized variation of S under the change $\bar{q}_a(t) \to \bar{q}_a(t) + \delta q_a(t) \equiv q_a(t)$ with $\delta q_a(t_f) = \delta q_a(t_i) = 0$

[1] By *degrees of freedom* or *dynamical variables* we mean the parameters needed to describe, fully and non-redundantly, the configuration of the system.

[2] Each $f_a(q, \dot{q})$ is in principle a function of all coordinates q and velocities $\dot{q}$

[3] For instance as formulated by Laplace in the *Essai philosophique sur les probabilités* (1814).

[4] Throughout these notes we shall indicate by squared brackets (ex. $S[q]$) the functional dependence and by round brackets the dependence on real (and complex) variables (ex. $L(q, \dot{q})$).

in order to preserve the boundary conditions $q_a(t_i) = q_a^i$, $q_a(t_f) = q_a^f$. That is concretely written as

$$0 = \delta S\left[\bar{q}\right] \equiv \left(S\left[\bar{q} + \delta q\right] - S\left[\bar{q}\right]\right)\Big|_{\text{linearized}} \tag{2.5}$$

$$= \int_{t_i}^{t_f} dt \left(\frac{\partial L}{\partial q_a}\delta q_a + \frac{\partial L}{\partial \dot{q}_a}\delta \dot{q}_a\right) \tag{2.6}$$

$$= \int_{t_i}^{t_f} dt \left(\frac{\partial L}{\partial q_a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_a}\right)\delta q_a + \frac{\partial L}{\partial \dot{q}_a}\delta q_a\Big|_{t_i}^{t_f} \tag{2.7}$$

$$= \int_{t_i}^{t_f} dt \left(\frac{\partial L}{\partial q_a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_a}\right)\delta q_a \tag{2.8}$$

where in the last line we used $\delta q_a = 0$ at the boundary. Notice that here, and throughout the whole course, we employ Einstein's convention for summation over repeated indices. Stationarity of the action is then equivalent to:

$$\boxed{\frac{\partial L}{\partial q_a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_a} = 0} \tag{2.9}$$

This set of N second order differential equations for the generalised coordinates is called the *Lagrange equations*. The Least Action Principle then amounts to this simple fact: for mechanical systems there always exists a Lagrangian function $L\left(q, \dot{q}\right)$ such that eq. (2.1) and eq. (2.9) coincide.

Equivalently, the dynamics of the system can be described in the Hamiltonian formulation where one ends up with $2N$ first order differential equations for the generalised coordinates and momenta. The Hamiltonian is defined as the Legendre transform of the Lagrangian with respect to the generalised momenta:

$$p_a \equiv \frac{\partial L}{\partial \dot{q}_a} \tag{2.10}$$

This relation is then inverted and the velocities are expressed in terms of the coordinates and momenta:

$$\dot{q}_i \to \dot{q}_a(q_a, p_a) \tag{2.11}$$

The Hamiltonian is finally given by:

$$H \equiv H(q, p, t) = p_a \dot{q}_a - L(q, \dot{q}) \tag{2.12}$$

In the Hamiltonian formulation the Lagrange equations then read

$$\boxed{\begin{aligned} \dot{q}_a &= \frac{\partial H}{\partial p_a} \\ \dot{p}_a &= -\frac{\partial H}{\partial q_a} \end{aligned}} \tag{2.13}$$

An important structure in Hamiltonian mechanics is offered by the *Poisson bracket* of pairs of real functions on the phase space. Given two such functions $A \equiv A(q, p)$ and $B \equiv B(q, p)$, their Posson bracket is defined by

$$\{A, B\} \equiv \frac{\partial A}{\partial p_a}\frac{\partial B}{\partial q_a} - \frac{\partial A}{\partial q_a}\frac{\partial B}{\partial p_a}\,. \tag{2.14}$$

By its definition the Poisson bracket satisfies the following two properties

$$\{A,B\}+\{B,A\}=0 \qquad \text{anti}-\text{symmetry} \tag{2.15}$$

$$\{A,\{B,C\}\}+\{B,\{C,A\}\}+\{C,\{A,B\}\}=0 \qquad \text{Jacobi identity} \tag{2.16}$$

The space $\mathcal{F}$ of real functions on the phase space can be viewed as a linear vector space. That is because any linear combinations of functions is again a function. The Poisson bracket can then be viewed as a product operation mapping pairs of functions (pairs of vectors) back to $\mathcal{F}$. Together with the properties 2.15,2.16 the Poisson bracket then endows $\mathcal{F}$ with the structure of a *Lie Algebra*. We will encounter and study Lie Algebras in the next section when considering continuous symmetry groups. Here we neither want nor can further explore this fact. We will limit ourselves to noticing that by eqs. (2.15,2.16) the Poisson bracket of any function $A(p,q)$ with q_a and p_a generates an infinitesimal canonical tranformation. That is, defining

$$q'_a \equiv q_a + \epsilon\,\{A,q_a\} \qquad\qquad p'_a \equiv p_a + \epsilon\,\{A,p_a\} \tag{2.17}$$

one can check that $\{p'_a,q'_b\}=\delta_{ab}+O(\epsilon^2)$. In other words, the Poisson bracket allows to write canonical transformations in an algebraic form. Notice that symmetry transformations are a particular example of canonical transformations, they too are then generated by the Poisson brackets of certain functions acting as symmetry generators.

The simplest example of all the above is given by time evolution itself. This can be viewed as a canonical transformation associated to the symmetry of time translation invariance. The generator is nothing but the Hamiltonian function and indeed we have

$$\dot{q}_a = \{H,q_a\} \tag{2.18}$$

$$\dot{p}_a = \{H,p_a\} \tag{2.19}$$

while for any observable $O\,(q,p)$, the time dependence is:

$$\begin{aligned}\frac{d}{dt}O &= \frac{\partial O}{\partial p_i}\dot{p}_a + \frac{\partial O}{\partial q_a}\dot{q}_a \\ &= -\frac{\partial O}{\partial p_a}\frac{\partial H}{\partial q_a} + \frac{\partial O}{\partial q_a}\frac{\partial H}{\partial p_a} = \{H,O\}\end{aligned} \tag{2.20}$$

Thus the Poisson bracket allows to formulate time evolution in algebraic form. But as we mentioned this property applies equally well to more involved transformations, like for instance rotations. In that case the infinitesimal generators are nothing but the three angular momentum components L_i, $i=1,2,3$, which famously satisfy the algebra

$$\{L_i,L_j\}=\epsilon_{ijk}L_k\,. \tag{2.21}$$

The algebraic structure induced by the Poisson bracket in Hamiltonian mechanics is *remarkably* unchanged when going to the quantum description of the system. The "quantization" of the system *simply, yet profoundly,* amounts to the replacement of the dynamical variables q_a, p_a with operators $\hat{q}_a$, $\hat{p}_a$ acting on a Hilbert space and to the replacement of the Poisson brackets with commutators, according to the scheme

Classical	Quantum
$q_a\,,\,p_a\,,\,A(q_a,p_a)$	$\hat{q}_a\,,\,\hat{p}_a\,,\,A(\hat{q}_a,\hat{p}_a)$
$H(q_a,\,p_a)$	$\hat{H}(\hat{q}_a,\,\hat{p}_a)$
$\{\cdot,\cdot\}$	$i\,[\cdot,\cdot]$

(2.22)

The above correspondence based on Hamiltonian mechanics defines what is known as *canonical quantization*. It is the approach that is followed in non-relativistic quantum mechanics in its Schrödinger's equation formulation. This approach, however, while offering a straighforward path to quantization, makes space-time symmetries not fully manifest. For instance, the Hamiltonian is directly connected to a choice of the time parameter t, but in relativity different inertial observers employ different time parameters. Because of that, in relativistic quantum field theory one makes a greater use of the Lagrangian formulation, where relativistic invariance is instead manifest, and which corresponds to the *path integral* formulation of quantum mechanics. In this course (QFT I and II) we will follow a hybrid path, in which we use the Lagrangian approach to characterize the system, in particular its symmetries, and the Hamiltonian approach to canonically quantize it. The path integral will be developed in QMIV and its methodology fully employed in the advanced courses of the third master semester.

2.2 From mechanics to field theory

The notion of a *field* arises when we consider a coordinate space $\mathcal{X}$ (*e.g.* $\mathbb{R}^n$, S^n,...) and we associate a set of *dynamical* variables $\phi_\alpha(\mathbf{x}, t)$ ($\alpha = 1, \dots, N$) to each spacial point $\mathbf{x} \in \mathcal{X}$. The dependence on the *time* coordinate t encodes the dynamics of the system. From a mechanical perspective, the *total* number of degrees of freedom is given by $N \times \dim\mathcal{X}$, where by $\dim\mathcal{X}$ we indicate the number of points in $\mathcal{X}$. Of course if $\mathcal{X}$ is a continuum, like $\mathbb{R}^n$ or S^n, then $\dim\mathcal{X}$ is infinite. One example of dynamical fields is given by the electric and magnetic fields:

$$\vec{E}(\mathbf{x}, t) \tag{2.23}$$

$$\vec{B}(\mathbf{x}, t) \tag{2.24}$$

Another example is offered by fluid mechanics, where the dynamical field variables are respectively the mass density, the pressure and the flow velocity

$$\rho(\mathbf{x}, t) \tag{2.25}$$

$$P(\mathbf{x}, t) \tag{2.26}$$

$$\vec{\mathbf{v}}(\mathbf{x}, t) \tag{2.27}$$

In both of the above examples the coordinate space is $\mathbb{R}^3 \ni \mathbf{x}$. This is infinite dimensional, both because the points form a continuum and because $\mathbf{x}$ extends to infinity. Thus the fields of both examples have an infinite number of degrees of freedom! Ordinary mechanics concerns systems with a finite number of dynamical variables $q_a(t)$, $a = 1, \dots, M$. Field theory, instead, can be viewed as a limit of mechanics where the discrete label a is replaced by $(\alpha, \mathbf{x})$, which, as long as $\mathbf{x}$ is a continuous, is continuous and thus infinite. As in the relevant quantum field theories $\mathbf{x}$ takes values in a continuum we will be forced us to deal with infinitely many degrees of freedom throughout this course. In particular, most of the time the coordinate space will be given by $\mathbb{R}^3$ unless specifically indicated.

In field theory, at any given time t, the state of the system is specified by a set of functions, the fields $\phi_\alpha(\mathbf{x})$ and their time derivatives $\dot{\phi}_\alpha(\mathbf{x})$, rather than by a finite set of real variables $(q_a, \dot{q}_a)$ as in mechanics. Consequently the Lagrangian, rather than a simple function, will now be a *functional*

$$L \equiv L[\phi, \dot{\phi}]\,. \tag{2.28}$$

In principle there is a lot of freedom for the form of the $L[\phi, \dot{\phi}]$. However, in the field theories that describe physical systems $L[\phi, \dot{\phi}]$ is itself the integral of a function of the fields and their derivatives at any given point $\mathbf{x}$

$$L[\phi, \dot{\phi}] = \int d^3\mathbf{x}\, \mathcal{L}\left(\phi(\mathbf{x}, t), \dot{\phi}(\mathbf{x}, t), \vec{\nabla}\phi(\mathbf{x}, t)\right)\,. \tag{2.29}$$

Notice that the functional dependence on $\vec{\nabla}\phi(\mathbf{x}, t)$ does not need to be indicated in $L[\phi, \dot{\phi}]$ as $\vec{\nabla}\phi$ is fully fixed once $\phi(\mathbf{x}, t)$ is specified as a function of $\mathbf{x}$ at some fixed t. A lagrangian of the above form is said to have the property of *locality*: L is the integral of a density $\mathcal{L}$ which purely depends of the local properties (the fields and their derivatives) at each point. For instance no product of fields at far away points appear if $L[\phi, \dot{\phi}]$. It is intuitively clear that, if a field theory description is to help us to get rid of instantaneous interactions between

spacelike separated objects, then its Lagrangian should be *local*, i.e. given by the space integral of a *Lagrangian density* $\mathcal{L}$. Notice also that we assumed $\mathcal{L}$ depends only on the fields and their first derivative with respect to the space and time coordinates. This is analogous to what is normally assumed in ordinary mechanics where L depends on $q, \dot{q}$. The systems we will consider have this property but it is straightforward to generalise to actions involving higher derivatives. Indeed one can show (see one of the exercises) that the dynamics implied by any lagrangian involving higher derivatives is equivalent to that of a lagrangian with only up to first derivatives but with a larger set of dynamical variables, in our case a larger set of fields. Notice finally, that very much like in mechanics the lagrangian does not depend explictly on t, the lagrangian density in field theory does not depend explicitly on either t or $\mathbf{x}$. As we will better discuss later, but as it is intuitively clear from the study of mechanics, this property corresponds to the invariance of the laws of motion under space-time translations.

Indicating by x^μ, $\mu = 0, 1, 2, 3$, the space-time coordinates with $x^0 = t$, $x^i = \mathrm{x}^i$ we can collectively indicate space-time derivatives by a four-vector $\partial_\mu \equiv \partial/\partial x^\mu$. We can then write the action as

$$S[\phi] = \int dt L[\phi, \dot{\phi}] = \int d^4x \, \mathcal{L}(\phi(x), \partial_\mu \phi(x)) \tag{2.30}$$

where we have joined the integrals of over space and over time into a single integral over space-time with measure $d^4x \equiv dt d^3\mathbf{x}$. For concreteness we have considered 4-dimensional spacetime, sometimes indicated as $(1+3)$-dimensional spacetime. It is however evident that our construction can be used to define the action in field theories over a spacetime of arbitrary dimension D. We would simply have to consider spacetime indices taking values $\mu = 0, \ldots, D-1$ and integration measure $d^D x$. In this respect, ordinary mechanics can be viewed as field theory in $(1+0)$-dimensional spacetime. In view of this analogy the lagrangian density $\mathcal{L}$ is normally simply referred to as "the Lagrangian".

2.3 Least action principle in field theory

We are now ready to generalize the formulation of the Least Action Principle to field theory.

Given any compact spacetime region Ω bounded by a surface $\partial\Omega$, we consider the action 2.30 limited to the region Ω, which is the integral of $\mathcal{L}$ over Ω

$$S_\Omega[\phi] = \int_\Omega d^4x \, \mathcal{L} \,. \tag{2.31}$$

Like in mechanics given an infinitesimal variation of a field configuration, $\phi_\alpha \to \phi_\alpha + \delta\phi_\alpha$, we indicate by $\delta S_\Omega[\phi]$ the linearized variation of the action, with $\delta\phi_\alpha$ vanishing on $\partial\Omega$. The Least Action Principle is then formulated as follows:

The solutions of the dynamics are given by the field configurations $\bar{\phi}_\alpha$, defined over the full $\mathbb{R}^4$ spacetime, and satisfying $\delta S_\Omega[\bar{\phi}] = 0$ for any Ω and for any $\delta\phi_\alpha$ satisfying $\delta\phi_\alpha|_{\partial\Omega} = 0$.

To see what that concretely implies consider a putative solution $\bar{\phi}_\alpha(x)$ and a general small variation $\bar{\phi}_\alpha \to \bar{\phi}_\alpha + \delta\phi_\alpha$ satisfying $\delta\phi_\alpha(x_{\partial\Omega}) = 0$ for any $x_{\partial\Omega} \in \partial\Omega$. The linearized variation of the action is then

$$\begin{aligned}
\delta S_\Omega[\phi] &\equiv \int_\Omega d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi_\alpha} \delta\phi_\alpha + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_\alpha)} \delta(\partial_\mu \phi_\alpha) \right) \\
&= \int_\Omega d^4x \left[\left(\frac{\partial \mathcal{L}}{\partial \phi_\alpha} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_\alpha)} \right) \delta\phi_\alpha + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_\alpha)} \delta\phi_\alpha \right) \right] \\
&= \int_\Omega d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi_\alpha} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_\alpha)} \right) \delta\phi_\alpha + \int_{\partial\Omega} d\sigma_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_\alpha)} \delta\phi_\alpha \\
&= \int_\Omega d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi_\alpha} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_\alpha)} \right) \delta\phi_\alpha
\end{aligned} \tag{2.32}$$

where we first integrated by parts, then used Gauss theorem to express the integral of a total derivative as a boundary integral [5] and finally used that $\delta\phi_\alpha = 0$ on the boundary $\partial\Omega$. The vanishing of the final expression for

[5] The boundary measure is given by $d\sigma_\mu = \epsilon_{\mu\nu\rho\sigma} dx^\nu dx^\rho dx^\sigma / 3!$ with $\epsilon_{\mu\nu\rho\sigma}$ the totally anti-symmetric Levi-Civita tensor and dx^μ tangent to $\partial\Omega$. Notice that $d\sigma_\mu$ is a 3-volume element.

arbitrary infinitesimal $\delta\phi_\alpha$ corresponds to the set of equations

$$\boxed{\frac{\partial\mathcal{L}}{\partial\phi_\alpha} - \partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_\alpha)} = 0} \tag{2.33}$$

that define the dynamics of a system of classical fields according to the Least Action Principle. These are the *Euler-Lagrange* equations.

In our formulation of the least action principle we used the notion of an arbitrary compact region Ω. But the steps in eq. (2.32) make it clear that we can directly formulate it by taking $\Omega = \mathbb{R}^4$, i.e. the whole spacetime, and by constraining $\delta\phi_\alpha$ to vanish fast enough at infinity so as to allow the dropping of the boundary terms.

Like in mechanics, we can consider a Newtonian formulation where the solution is univocally determined, besides by the Euler-Lagrange equations, by fixing $\phi_\alpha(\mathbf{x}, t_i)$ and $\dot\phi_\alpha(\mathbf{x}, t_i)$, for any $\mathbf{x}$ at some time t_i. The $\mathbb{R}^3$ spacetime slice $t = t_i$, where these initial conditions are given, is called the Cauchy surface.

It should be appreciated that the least action principle in field theory is the natural generalization of the case of mechanics, where the region Ω reduces to the compact segment $[t_i, t_f]$. In our discussion we have left Ω arbitrary, also given, as we argued, the principle applies to any choice. However, in order to get a more concrete picture, and to make closer contact with mechanics, we could choose a solid cylinder $\Omega = [t_i, t_f] \times B_R$ with B_R the spacial ball of radius R: $|\mathbf{x}| \leq R$. A small variation $\bar\phi_\alpha \to \bar\phi_\alpha + \delta\phi_\alpha$ will then satisfy boundary conditions

$$\delta\phi_\alpha(\mathbf{x}, t_i) = \delta\phi_\alpha(\mathbf{x}, t_f) = 0 \quad \text{and} \quad \delta\phi_\alpha(\mathbf{x}, t)|_{|\mathbf{x}|=R} = 0\,. \tag{2.34}$$

and the linear variation of the action be

$$\begin{aligned}
\delta S_\Omega[\phi] &\equiv \int_\Omega d^4x \left(\frac{\partial\mathcal{L}}{\partial\phi_\alpha}\delta\phi_\alpha + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_\alpha)}\delta(\partial_\mu\phi_\alpha)\right) \\
&= \int_\Omega d^4x \left(\frac{\partial\mathcal{L}}{\partial\phi_\alpha} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_\alpha)}\right)\delta\phi_\alpha + \int_{B_R} d^3x\, \frac{\partial\mathcal{L}}{\partial\dot\phi_\alpha}\delta\phi_\alpha\Big|_{t_i}^{t_f} + \int_{t_i}^{t_f} dt \int_{\partial B_R} d^2\Sigma\; \vec n \cdot \frac{\partial\mathcal{L}}{\partial(\vec\nabla\phi_\alpha)}\delta\phi_\alpha \\
&= \int_\Omega d^4x \left(\frac{\partial\mathcal{L}}{\partial\phi_\alpha} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_\alpha)}\right)\delta\phi_\alpha
\end{aligned} \tag{2.35}$$

from which we draw the same conclusions as in our previous general discussion.

The expression appearing in brackets in the last line defines the *functional derivative* of the functional S

$$\boxed{\frac{\delta S_\Omega}{\delta\phi(x)} \equiv \left(\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)} \tag{2.36}$$

so that we can write (now taking $\Omega = \mathbb{R}^4$ and assuming $\delta\phi_\alpha$ vanishes fast enough to drop boundary terms)

$$S[\phi + \delta\phi] - S[\phi] = \int d^4x \frac{\delta S}{\delta\phi(x)}\delta\phi(x) \;+\; \mathcal{O}(\delta\phi(x))^2\,. \tag{2.37}$$

This result is in full analogy in analogy with ordinary derivatives of functions of many variables, where, given a function $f(q_i)$, we have

$$f(q + \delta q) - f(q) = \sum_i \frac{\partial f}{\partial q_i}\delta q_i + \mathcal{O}(\delta q_i)^2\,. \tag{2.38}$$

The concept of functional derivatives is obviously straightforwardly generalised to arbitrary space dimensions and to densities that depend on coordinates as well. For example, define a functional F:

$$F[\phi] = \int d^n x f(\phi, \partial_i\phi, x) \tag{2.39}$$

Then:

$$\frac{\delta F}{\delta \phi(x)} \equiv \frac{\partial f}{\partial \phi} - \partial_i \frac{\partial f}{\partial(\partial_i \phi)} \tag{2.40}$$

In particular, we can choose as density:

$$f(\phi, \partial_i \phi, x) = \phi(x)\delta^{(n)}(x - x_0) \tag{2.41}$$

So that $F[\phi] = \phi(x_0)$. This yields on the one hand:

$$\frac{\delta F}{\delta \phi(x)} = \frac{\delta \phi(x_0)}{\delta \phi(x)} \tag{2.42}$$

while on the other hand:

$$\frac{\delta F}{\delta \phi(x)} = \frac{\partial f}{\delta \phi(x)} = \delta^{(n)}(x - x_0) \tag{2.43}$$

Yielding finally an important generalisation of $\frac{\partial q_i}{\partial q_j} = \delta^i_j$:

$$\frac{\delta \phi(x_0)}{\delta \phi(x)} \equiv \delta^{(n)}(x - x_0) \tag{2.44}$$

2.4 Hamiltonian formalism

In the Lagrangian description of field theory, see eq. (2.30), the time and the space coordinates are treated on equal grounds. This is an advantage of the Lagrangian formalism when considering relativistic field theory, because, as we shall see, it makes spacetime symmetries manifest.

The Hamiltonian approach, on the other hand, singles out the time coordinate t among the spacetime coordinates. Because of that, spacetime symmetries are less manifest in the Hamiltonian approach. There are however other advantages in this approach, in particular its algebraic structure, which directly carries on to the quantum theory.

The Hamiltonian formalism can be developed along the same lines as in ordinary mechanics. This formalism is based on a specific choice of time slicing and on the corresponding spacially integrated lagrangian density L in eq. (2.29). The canonically conjugated momentum $\pi(x)$ is naturally defined by

$$\pi(x) \equiv \frac{\delta L}{\delta \dot{\phi}(x)} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \tag{2.45}$$

which simply extends the ordinary derivative of eq. (2.10) to a functional derivative. The above local relation is then inverted at each point to express $\dot{\phi}$ in terms of $\pi(x)$, $\phi(x)$ and $\nabla\phi(x)$

$$\dot{\phi}(x) \to \dot{\phi}(\pi(x), \phi(x), \nabla\phi(x)) \tag{2.46}$$

The Hamiltonian is defined in analogy with mechanics by a Legendre transform:

$$H(\pi, \phi) = \int d^3\mathbf{x} \left(\pi(x)\dot{\phi}(x) - \mathcal{L}(\phi, \pi)\right) \tag{2.47}$$

where:

$$\mathcal{H} \equiv \pi(x)\dot{\phi}(x) - \mathcal{L}(\phi, \pi) \tag{2.48}$$

is the Hamiltonian density. The Hamiltonian is a *local* functional of the fields and of their conjugated momenta over three dimensional space! As expected, in the Hamiltonian formalism the time coordinate has been singled out so that in this formalism relativistic invariance is no longer manifest. Also, as we work with functionals, the functional derivative formalism we developed in the previous section becomes handy. "Equal time" Poisson brackets (we make the time dependence implicit from here on) can then be straightforwardly generalised by employing functional derivatives. Given two functionals:

$$A[\pi, \phi] = \int d^3\mathbf{x}\ a(\pi, \phi) \tag{2.49}$$

$$B[\pi, \phi] = \int d^3\mathbf{x}\ b(\pi, \phi) \tag{2.50}$$

we define:

$$\{A, B\} \equiv \int d^3\mathbf{x} \left(\frac{\delta A}{\delta \pi} \frac{\delta B}{\delta \phi} - \frac{\delta A}{\delta \phi} \frac{\delta B}{\delta \pi} \right) \tag{2.51}$$

In particular, using Eq. 2.44, our definition implies:

$$\{\pi(\mathbf{x}, t), \phi(\mathbf{y}, t)\} = \delta^{(3)}(\mathbf{x} - \mathbf{y}) \tag{2.52}$$

$$\{\pi(\mathbf{x}, t), \pi(\mathbf{y}, t)\} = 0 = \{\phi(\mathbf{x}, t), \phi(\mathbf{y}, t)\} \tag{2.53}$$

Similarly to ordinary mechanics, the equations of motion can be written as

$$\boxed{\begin{aligned} \dot{\pi} &= \{H, \pi\} = -\frac{\delta H}{\delta \phi} = \vec{\nabla} \cdot \frac{\partial \mathcal{H}}{\partial(\vec{\nabla}\phi)} - \frac{\partial \mathcal{H}}{\partial \phi} \\ \dot{\phi} &= \{H, \phi\} \ = \frac{\delta H}{\delta \pi} = \frac{\partial \mathcal{H}}{\partial \pi} \end{aligned}} \tag{2.54}$$

Chapter 3

Symmetries

Symmetries play a central role in physics, both from a fundamental perspective and from a practical one. Relativity, which is one of the conceptual bases of Quantum Field Theory, provides an example of the fundamental role played by symmetries. The practical relevance of symmetries is instead elucidated by the solution of innumerable physics problems. Example range, for instance, from the use of spherical or cylindrical symmetry in electrostatics, to the use of isospin symmetry in the study of the strong interactions. Indeed symmetries are also useful, sometimes even more so, when they are only approximately exact and provide deep insight into the dynamics of physical systems through the so-called selection rules.

Mathematically symmetries are associated with the notion of *group*. The corresponding area of Mathematics, *Group Theory*, provides then both the language to formulate many physics problems and the technique to solve them. There is no escape: a physicist, even more so a theoretical one, has better learn group theory!

3.1 Group theory

3.1.1 Groups

Let us start with a short summary of basic group theory. It is highly suggested to any reader, familiar or not with the topic, to have a look at one among any reference book on group theory (*e.g. Lie Algebras in Particle Physics* by H. Georgi Chap. 1-3).

Definition. *A group G is a set of elements $\{g_i\}$, $i = 1, \ldots, |G|$, where $|G|$ is called the order of the group, endowed with a binary operation that assigns to each ordered pair of elements a third element of the group:*

$$\begin{aligned} \circ \ &: G \times G \mapsto G \\ &(g_1, g_2) \mapsto g_1 \circ g_2 = g_3 \end{aligned} \tag{3.1}$$

This binary operation is usually written with a multiplicative notation:

$$g_1 \circ g_2 \ \rightarrow g_1 g_2 \tag{3.2}$$

For G to be a group, the binary operation must satisfy the following axioms:

1. *Associativity:*

$$g_1(g_2 g_3) = (g_1 g_2) g_3, \quad \forall \ g_1, g_2, g_3 \in G \tag{3.3}$$

2. *Existence of a (left-)identity:*

$$\exists\ e \in G \ \ s.t. \ \ \forall\ g \in G \ \ eg = g \tag{3.4}$$

3. *Existence of a (left-)inverse:*

$$\forall\ g \in G, \ \ \exists\ g^{-1} \in G \ \ s.t. \ \ g^{-1}g = e \tag{3.5}$$

In many books, it is required both the identity and the inverse to act also from the right. It is actually enough to have a left-identity/inverse. Indeed, from these three axioms it is easy to prove the following corollaries:

Corollary. *The group axioms have the following consequences:*

1. *The left-inverse is also a right-inverse:*

$$gg^{-1} = e \tag{3.6}$$

2. *The left-identity acts in the same way from the right:*

$$ge = g \tag{3.7}$$

3. *Uniqueness of the identity:*

$$\exists!\ e \in G \ \ s.t. \ \ \forall\ g \in G \ \ eg = g = ge \tag{3.8}$$

4. *Uniqueness of the inverse :*

$$\forall\ g \in G, \ \ \exists!\ g^{-1} \in G \ \ s.t. \ \ gg^{-1} = e = g^{-1}g \tag{3.9}$$

The proofs are left as an exercise.

Two elements g_1 and g_2 are said to *commute* if $g_1 g_2 = g_2 g_1$. Moreover if $g_1 g_2 = g_2 g_1\ \forall\ g_1, g_2 \in G$, then G is called an *abelian group.* A subset $H \subset G$ such that

1. $e \in H$
2. $\forall\, h_1\, h_2 \in H\,,\ h_1 h_2 \in H$
3. $\forall h \in H\,,\ h^{-1} \in H$

is itself a group and is said to be a subgroup of G.

Let us enumerate a few examples of groups.

1. $(\mathbb{R}^+, \cdot)$ the set of real, strictly positive numbers with the operation of multiplication forms a group. This is swiftly verified. Associativity is indeed satisfied by the arithmetic product. Each positive real number has an inverse which is also a real positive number. The identity element is 1.
2. $(\mathbb{Z}, +)$ the set of integers with addition is another example of group. Again, it is associative. The inverse of an integer a is given by $-a$, and the identity element is 0.
3. The group of permutations of n objects, denoted by $\mathcal{S}_n$, is a third example. It is composed of the $n!$ bijections π between the numbers $1, ..., n$ and themselves:

$$(1, .., n) \mapsto (\pi(1), ..., \pi(n))$$

The operation in this group is the usual function composition, which satisfies associativity. $\mathcal{S}_n$, unlike $(\mathbb{R}^+, \cdot)$ and $(\mathbb{Z}, +)$, is a non-abelian group.

4. The group of rotations in $\mathbb{R}^3$, denoted by $SO(3)$, is another important example of a non-abelian group. To show it is non-abelian, consider the unit vector in the x direction. Apply two rotations of $\pi/2$, the first about the z direction and the second about the x direction. You will get the unit vector in the z direction. Reversing the order of those two operations yields the unit vector in the y direction:

$$\vec{e_z} = R_x(\pi/2)R_z(\pi/2)\vec{e_x} \neq R_z(\pi/2)R_x(\pi/2)\vec{e_x} = \vec{e_y}$$

The above examples also illustrate how groups can be distinguished in some basic classes. One distinction is between *finite groups*, like $\mathcal{S}_n$, which contain a finite number of elements, and *infinite groups*, like $(\mathbb{Z}, +)$, which do not. Another distinction is between *discrete* groups, like $\mathcal{S}_n$ and $(\mathbb{Z}, +)$, and *continuous* groups, like $(\mathbb{R}^+, \cdot)$ and $SO(3)$. This intuitively correspond to the distinction between discrete and continuous spaces. Concretely and more mathematically, continuous groups are locally isomorphic to $\mathbb{R}^n$ for some $n \in \mathbb{N}$ with the groups operations satisfying continuity with respect to the standard topology in $\mathbb{R}^n$. In other words, given an element g in a continuous group, we can find a succession $\{g_i\}$ of elements not including g whose properties get arbitrarily close to those of g. In particular, given another group element h, the succession of products $\{g_i h\}$ will approximate gh arbitrarily well. Discrete groups, instead do not satisfy this property or, as a mathematician would say, they more trivially respect continuity with respect to the finest possible topology, the discrete one. As a side note, notice a continuous group always has an infinite number of elements whereas a discrete group it may not be the case, as exemplified by $\mathcal{S}_n$ and $(\mathbb{Z}, +)$.

We can also consider another example, which is also central to the occurrence of group theory in physics.

5. Consider a configuration space given by a manifold parametrized by N coordinates $\{q_i\}$, $i = 1, \ldots, N$. For example for $N = 1$ and $\{q_i\} \equiv \theta \in [0, 2\pi)$, the manifold is a circle. Now on any space there exists *freedom* in the choice of coordinates (*e.g.* cartesian, spherical,...). A change of coordinates corresponds to a change in the "point of view" on the system. Mathematically:

$$q_i \mapsto q_i' = f_i\left(\{q_i\}\right) \tag{3.10}$$

such that for any configuration there is one and only one $\{q_i'\}$. Then f is a *bijection* and the set of all possible coordinate changes is a group with respect to the binary operation which is the composition of functions! Such bijections are also called *reparametrisations*. Obviously if f and g are two reparametrisations, then $f \circ g$ is also a reparametrisation Therefore, to prove that we deal with a group, we only need to check the three axioms a group is required to obey.

   (a) Associativity. Let f, g, h be three changes of coordinates and let us work, without loss of generality, with space described by a single coordinate ($\{q_i\} \equiv q$). Then by the associativity of function composition (like for S_n):

$$(f \circ g) \circ h(q) = f(g(h(q))) = f \circ (g \circ h(q))$$

   (b) The identity is simply:

$$q \mapsto q$$

   (c) Inverse. As we consider the set of bijective functions, by definition, for any reparametrisation f there exists f^{-1}, its inverse, such that:

$$f^{-1} \circ f(q) = q = f \circ f^{-1}(q)$$

Hence, the set of changes of coordinates endowed with the operation of functions' composition is indeed a group.

This example is interesting for its own sake. However, what is even more interesting is that all the previous examples can themselves be viewed as reparametrisations. S_n was already defined as a reparametrisations of a discrete set. Moreover $(\mathbb{R}^+, \cdot)$ can be viewed as describing rescalings of a real parameter, i.e. a changes of units of measure, $(\mathbb{Z}, +)$ as describing translations on a discrete lattice, $SO(3)$ as describing rigid rotations to another Cartesian frame.

3.1.2 Groups in Physics: Symmetry

To study a physical system, it is necessary to choose a set of coordinates parametrizing its configurations. Very generally, we choose coordinates (q_i) belonging to a space X. For instance, for a system of N point particles in mechanics $X = \mathbb{R}^{3N}$ while for the rigid body X is a 3-sphere with antipodal points identified, which can be parametrized by the three Euler angles. However, nothing prevents another observer to use a different set of coordinates (q'_i) belonging to X' while describing the same physical system. The change of coordinates remains valid as long as there exists a bijection f allowing to go from one set of coordinates to the other. In the general case, X and X' can have very different structures. Well-known examples are the cartesian coordinates $\mathbb{R}^2$ and the polar coordinates $[0, 2\pi[\times[0, \infty[$ [1]. For our discussion, we will focus on the case where $X = X'$. The fact we may have two different faithful descriptions of the same system tells us this mapping precisely belongs to the group of reparametrisations of X. And this is where groups enter the Physics arena.

Now, the specific interest arises when considering the way a change of coordinates affects the description of the dynamics of the system. An arbitrary change of coordinates may completely change the form of the equations of motion. For instance, it could make them look simpler or more complicated. While the first case has clearly great practical interest, the case which has instead great fundamental interest is that in which the change of variables does not change the form in which the dynamics is described. This is the subset of transformations that leave the equations of motion invariant. As it should be evident, this subset itself forms a subgroup of the general group of bijections from X to itself. We call this subset the *Symmetries* of the system.

Definition. *The set of symmetries of a physical system is the set of reparametrizations leaving the form of the equations of motion unchanged.*

Each different parametrization of a physical system can be pictured as a different observer. The elements of the symmetry group of a system therefore parametrize a family of different observers that describe physics in exactly the same way. Just by considering the form of their equations of motion we cannot tell one of these equivalent observers from another.

Consider for instance the equation of motion of a single free non-relativistic particle:

$$m\frac{d^2\vec{x}}{dt^2} = 0 \tag{3.11}$$

This equation of motion is invariant under the *Galileo* group, which is the symmetry group of Newtonian mechanics (inlcuding in particular the case of an arbitrary number of interacting particles). The Galileo group consists of three-dimensional rotations, uniform translations of spacetime, as well as linear uniform boosts:

$$\vec{x} \;\mapsto\; \vec{x}'(t') = R(\alpha,\beta,\gamma)\vec{x}(t) + \vec{v}_0 t + \vec{x}_0\,, \qquad t' = t + t_0 \tag{3.12}$$

One can immediately check, by standard calculus, that eq. (3.11) translates into an equation of the same precise form in the new coordinates

$$m\frac{d^2\vec{x}'}{dt'^2} = 0 \tag{3.13}$$

The Galileo group transformation is described by 10 real parameters $(\alpha, \beta, \gamma, \vec{v}_0, \vec{x}_0, t_0)$ taking values in a continuum. For example, $\vec{x}_0 \in \mathbb{R}^3$. Therefore, the Galileo group is another example of a continuous group. As we will see shortly, it is indeed an example of a Lie group.

3.1.3 Lie groups

Let us first introduce or recall some basic notions in differential geometry. A differentiable manifold is a space that is at all its points locally homeomorphic to some patch of $\mathbb{R}^N$ and such that the transition maps between

[1] In this case, the change of coordinates is a bijection provided we identify the points $[0, 2\pi[\times 0$ in polar coordinates, which all correspond to the origin.

overlapping patches (or charts) are differentiable. A differentiable manifold, in the mind of a physicist, is normally a smooth differentiable N-dimensional surface embedded in a higher dimensional $\mathbb{R}^M$. An analytic manifold is the same as a differentiable manifold with the added stronger request that the transition maps be not only differentiable but also analytic, i.e. their Taylor expansion is a convergent series. Again, we physicists can think of such a manifold as a surface in $\mathbb{R}^M$ determined by some analytic constraints on the coordinates. An example of that is the n-sphere, S^n, corresponding to the points $\mathbf{x} \in \mathbb{R}^{n+1}$ satisfying the constraint $\mathbf{x} \cdot \mathbf{x} = 1$. Other types of quadratic constraints determine other analytic manifolds, like paraboloids or hyperboloids, and similarly higher order constraints, like cubics etc. In a sense analytic manifolds are "smoother than smooth", in that it suffices to know a part of it to reconstruct the rest.

Definition. *G is an N-dimensional Lie group if*

1. *G is an analytic manifold of dimension N. G is then obviously continuous: its elements are labelled by N continuous coordinates.*
2. *the group operations of product and inversion are also described by analytic functions of the coordinates. This means that, given any two group elements labelled respectively by coordinates $\alpha \equiv (\alpha_1, \ldots, \alpha_N)$ and $\beta \equiv (\beta_1, \ldots, \beta_N)$ and the products and inversions*

$$g(\alpha)g(\beta) = g(p(\alpha, \beta))\,, \tag{3.14}$$

$$g^{-1}(\alpha) = g(r(\alpha))\,, \tag{3.15}$$

the functions $p(\alpha, \beta)$ and $r(\alpha)$ are analytic.

We have defined Lie groups, as analytic manifolds where the group operations are also analytic functions. It would thus seem that Lie groups are very special continuous groups. It however turns out that any continuous groups is in reality a Lie group! This fact was only proven in the 1950's, decades after the use of Lie groups had already become widespread in Mathematics and Physics. The question of the equivalence between continuous groups and Lie Groups, was indeed the 5th problem of the list Hilbert presented to mathematicians in the summer of 1900: *Is every locally euclidean topological group a Lie group?*

In what follows we will indicate compactly indicate the Lie group coordinates as a vector

$$\alpha \equiv \vec{\alpha} = (\alpha_1, \ldots, \alpha_N)\,. \tag{3.16}$$

The α's will also be referred to as *Lie parameters*. We will also adhere to the convenient convention to choose coordinates where the identity group element corresponds to $\vec{\alpha} = \vec{0}$

$$g(\vec{0}) = e \tag{3.17}$$

Let us consider some examples Lie groups emphasizing their being (analytic) manifold

1. $(\mathbb{R}^+, \cdot)$: $\mathbb{R}^+$ is a manifold isomorphic to $\mathbb{R}$ as any element $x \in \mathbb{R}^+$ can be uniquely written as $x = e^\alpha$ for $\alpha \in \mathbb{R}$. Notice that in this exponential parametrization of $(\mathbb{R}^+, \cdot)$, the identity corresponds to $\alpha = 0$ in compliance with the convention stated above.
2. $U(1)$, the group of rotations in the plane. Parametrizing any point in the plane by a complex number $z \equiv x + iy$, rotations are given by $z \to e^{i\alpha} z$, with $\alpha \in \mathbb{R}$. Of course α and $\alpha + 2\pi$ give rise to the same group element, so that the $U(1)$ group manifold is the circle S^1. Again we have chosen a parametrization where $\alpha = 0$ corresponds to the identity element.
3. $SO(3)$ is a Lie group corresponding to the manifold $S^3/\mathbb{Z}_2$, i.e. the 3-dimensional sphere in $\mathbb{R}^4$ with antipodal points identified. We will later discuss its parametrization in terms of the Lie parameters.
4. $SU(2)$ is truly the whole manifold S^3 with no identification of the antipodal points. We can then informally say that $SU(2)$ is "twice as big as $SO(3)$".

Lie groups are crucial objects in Physics, and even more so in Particle Physics. The Galileo group was an early example of a Lie group in Physics. But the relevance of groups and Lie groups was boosted by the advent of Quantum Mechanics. Indeed, the structure of the Hilbert space of states of physical systems is often controlled by groups theory to a significant extent.

3.1.4 Realizations and Representations

Groups, and Lie groups in particular, can be thought as abstract mathematical objects, defined by their properties and by the rules they satisfy. However, and that it is especially the case in physics, we often work with an explicit *realization* of the group. A realization is a concrete way of writing the group elements in terms of transformations over some space X. Let us clarify this notion by considering some examples.

1. The group of affine transformation on the real line $\mathrm{Aff}(\mathbb{R})$ can be defined through its realization as a group of bijective mappings of $\mathbb{R} \mapsto \mathbb{R}$ (i.e. the space X over which we realize our group is $\mathbb{R}$)

$$x \longmapsto g(\alpha_1, \alpha_2)x = e^{\alpha_1}x + \alpha_2, \qquad x \in \mathbb{R}, \quad (\alpha_1, \alpha_2) \in \mathbb{R}^2 \tag{3.18}$$

   $\mathrm{Aff}(\mathbb{R})$ is a two-dimensional Lie group whose manifold is isomorphic to $\mathbb{R}^2$.

2. The euclidean group in 2 dimensions $ISO(2)$ (i.e. the set of transformations that leave the euclidean metric unchanged) is another example. It is simply given by the combination of translations and rotations on the plane (i.e. the space X over which we realize our group is $\mathbb{R}^2$)

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\alpha_3 & \sin\alpha_3 \\ -\sin\alpha_3 & \cos\alpha_3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \tag{3.19}$$

   Its Lie parameters $(\alpha_1, \alpha_2, \alpha_3)$ clearly span the manifold $\mathbb{R}^2 \times S^1$, corresponding to respectively to translations and rotations.

An important class of realizations of Lie groups is given by *group representations*.

Definition. *Given a group G and given a linear vector space V with the set (actually the group) $GL(V)$ of invertible linear operators in V, a representation D is an application from G to $GL(V)$*

$$\begin{array}{rcl} D : G & \to & GL(V) \\ g & \mapsto & D(g) \end{array} \tag{3.20}$$

satisfying

$$1. \quad D(g_1)D(g_2) = D(g_1 g_2), \qquad \forall\ g_1, g_2 \in G \tag{3.21}$$

$$2. \quad D(e) = \mathbb{1} \tag{3.22}$$

By the above properties, a representation D is a *homomorphism* between G and $GL(V)$. The vector space V is said to be the *basis of the representation*. If V has dimension N, then the representation D is said to have dimension N.

Concretely, choosing a basis in V, we can parametrize any $v \in V$ as a vector $v \equiv (v_1, \ldots, v_N)$ and any element of the representation $D(g)$ as an $N \times N$ matrix $D(g) \equiv D(g)_i^{\ j}$. We can then concretely express the abstract operation $v \mapsto D(g)v$ as

$$v_i \longmapsto D(g)_i^{\ j} v_j \tag{3.23}$$

A representation can then be concretely viewed as a mapping between a group and a set of matrices.

Then, the realisation of the Lie group on the vector space V is as follows:

$$\forall\ g \in G, \quad \vec{v} \mapsto D(g)\vec{v} \in V, \quad \forall\ \vec{v} \in V \tag{3.24}$$

The "fame" of group representations is largely due to Quantum Mechanics. That is because according to Quantum Mechanics the states of a physical system (physical reality itself!) correspond to vectors in a Hilbert space.

Symmetry operations and changes of point of view on the physical system must then correspond to operations on the states of the Hilbert space. By the superposition principle in Quantum Mechanics this must be a correspondence with linear operators, i.e. it must be a *representation* of the group of transformations. It thus so happens that on our way towards the study of the role of symmetries in Quantum Mechanics, it is necessary to develop the subject of group representations. In what follows we will provide some basic notions and some basic results in the theory of groups representations.

Definition. *Consider a representation D of a group G on an N-dimensional vector space V.*

1. *D is **reducible** if there exists a non-trivial invariant subspace $U \neq V, \{0\}$ of V. Formally:*
$$\exists\ U \subset V, \quad s.t.\ \ \forall\ \vec{u} \in U, \quad D(g)\vec{u} \in U,\ \forall g \in G \tag{3.25}$$
2. *D is **irreducible** if it is not reducible. That is, the only invariant subspaces are V itself and $\{0\}$.*
3. *D is **completely reducible** if and only if we can decompose V into a direct sum of invariant subspaces:*
$$V = \oplus_i U_i = U_1 \oplus U_2 \oplus \dots \tag{3.26}$$
such that D acts irreducibly on each U_i. In other words, there exists a basis of V such that:
$$\forall\ g \in G, \quad D(g) = \left(\begin{array}{c|c|c} D_1(g) & 0 & 0 \\ \hline 0 & D_2(g) & 0 \\ \hline 0 & 0 & \ddots \end{array}\right) \tag{3.27}$$
*where D_i is a $n_i \times n_i$ matrix, $n_1 + n_2 + \dots = N$ and all D_i are irreducible. D is said to be a **direct sum** of irreducible representations:*
$$D = \oplus_i D_i = D_1 \oplus D_2 \oplus \dots \tag{3.28}$$
Notice that an irreducible representation is by definition completely reducible, only that the decompositions consist of just a single invariant subspace coinciding with the whole V.
4. *Two representations D_1 and D_2 are **equivalent** if they are related by a similarity transformation:*
$$\exists\ S \in GL(n, \mathbb{C}),\ s.t.\ \ S^{-1} D_1(g) S = D_2(g), \quad \forall\ g \in G \tag{3.29}$$
In other words, D_1 and D_2 are related by a change of basis described by S.
5. *D is **unitary** if it is equivalent to a representation in which :*
$$D^{-1}(g) = D^{\dagger}(g), \quad \forall\ g \in G \tag{3.30}$$
Once can easily prove that a unitary representation is always completely reducible.

Notice that two representations of the same group can be very different. In particular representations can range from the so-called *trivial* representation $D^{(0)}$ for which $D^{(0)}(g) = \mathbb{1}$ for any $g \in G$ of the group, to the so-called *faithful* representations for which
$$g_1 \neq g_2 \implies D(g_1) \neq D(g_2)\,. \tag{3.31}$$
The set of matrices in a faithful representation offer therefore a concrete realization of G as a subgroup of $GL(V)$. While the trivial representation is maximally unfaithful, in the sense that it does not contain any information about the group structure, the faithful ones encode all the structure. Between these two extremes, there normally are non-trivial unfaithful representations that retain only a part of the structure of G.

Let us illustrate with an example the notion of invariant subspace.

1. Let V be a N-dimensional $\mathbb{K}$-vector space. Let $M < N$ and V' be a M-dimensional invariant subspace of V. Up to changing the basis, a vector in $v \in V'$ can be written:

$$v \in \begin{pmatrix} w_M \\ 0_{N-M} \end{pmatrix} \tag{3.32}$$

for some $w_M \in \mathbb{K}^M$ and from which it can be seen V' is invariant under all transformations of the form:

$$\begin{pmatrix} M \times M & M \times (N-M) \\ 0_{(N-M)\times M} & (N-M)\times(N-M) \end{pmatrix} \tag{3.33}$$

Irreducible representations play a crucial role in the application of group theory to physics. That is particularly the case in Quantum Mechanics where the representations have to be unitary as the norm of vectors is associated to a probability. Based on our previous results, group representations over the Hilbert space of a quantum mechanical system must therefore be fully reducible, i.e. built as a direct sum of irreducible representations. Such representations are like the "atomic components" of all possible representations and their classification is of paramount importance. The results that follow play a crucial role in characterization of the irreducible components of a representation.

3.1.5 Schur's lemma

A crucial result on the reducibility of representations is given by Schur's lemmas.

Schur's lemma. *Let D_1 and D_2 be two irreducible representations of G acting on V_1 and V_2 respectively and an intertwining operator $\Lambda : V_1 \to V_2$ such that:*

$$\Lambda D_1 = D_2 \Lambda \tag{3.34}$$

Then:

1. *Either $\Lambda \equiv 0$ or Λ is invertible. In the second case, V_1 and V_2 have therefore the same dimension and $D_1 = \Lambda^{-1} D_2 \Lambda$, that is to say D_1 and D_2 are equivalent representations.*

   Corollary: If, under the hypotheses of Schur's Lemma, D_1 and D_2 are further assumed to be inequivalent, then it must be $\Lambda = 0$.

2. *If $V_1 = V_2 = V$ and $D_1 = D_2 = D$ then $\Lambda = \lambda \mathbb{1}$.*

Now let D be a completely reducible representation and Λ an hermitian operator such that:

$$\Lambda D = D \Lambda \tag{3.35}$$

3. *Then Λ is block-diagonal in the same basis in which D is block-diagonal:*

$$D = \left(\begin{array}{c|c|c} D_1 & 0 & 0 \\ \hline 0 & D_2 & 0 \\ \hline 0 & 0 & \ddots \end{array}\right), \quad \Lambda = \left(\begin{array}{c|c|c} \lambda_1 \mathbb{1} & 0 & 0 \\ \hline 0 & \lambda_2 \mathbb{1} & 0 \\ \hline 0 & 0 & \ddots \end{array}\right) \tag{3.36}$$

We give a short proof of the first and second Schur's lemmas:

Proof. *To prove the first lemma, consider the two following objects:*

$$[Ker(\Lambda) \equiv \textit{subspace of } V_1 \textit{ annihilated by } \Lambda] \quad \textit{and} \quad [Im(\Lambda) \equiv \Lambda V_1] \tag{3.37}$$

Eq. 3.34 shows that $Ker(\Lambda) \subseteq V_1$ *and* $Im(\Lambda) \subseteq V_2$ *are invariant subspaces for the representations* D_1 *and* D_2 *respectively:*

$$\forall\ \vec{v}_1 \in Ker(\Lambda)\ :\quad \Lambda(D_1\vec{v}_1) = D_2(\Lambda\vec{v}_1) = 0\ \Rightarrow\ \forall\ \vec{v}_1 \in Ker(\Lambda),\ D_1\vec{v}_1 \in Ker(\Lambda) \tag{3.38}$$

and

$$\forall\ \vec{v}_2 \in Im(\Lambda),\ \exists\ \vec{v}_1 \in V_1,\ s.t.\ \vec{v}_2 = \Lambda\vec{v}_1 :\ \begin{matrix} D_2\vec{v}_2 = D_2\Lambda\vec{v}_1 \\ D_2\vec{v}_2 = \Lambda(D_1\vec{v}_1) \end{matrix} \quad \Rightarrow\ \forall\ \vec{v}_2 \in Im(\Lambda),\ D_2\vec{v}_2 \in Im(\Lambda) \tag{3.39}$$

This proves that the kernel and the image of the intertwining operator are invariant subspaces under D_1 *and* D_2 *respectively. Now, since we assume these two representations to be irreducible (meaning that there is no non-trivial invariant subspace for any of them), either* $\Lambda \equiv 0$ *or these subspaces are trivial:*

$$Ker(\Lambda) = \{0\}, \qquad Im(\Lambda) = V_2 \tag{3.40}$$

In the latter case, Λ *is injective and surjective, i.e. bijective and invertible.*

The second lemma can be seen as a corollary since it uses the first one to be proven. Eq. 3.34 becomes in this case:

$$\Lambda D = D\Lambda \tag{3.41}$$

This equation remains obviously true if Λ *is replaced by* $\Lambda_\lambda = \Lambda - \lambda\mathbb{1}$ *for any real (or complex)* λ*:*

$$\Lambda_\lambda D = D\Lambda_\lambda \tag{3.42}$$

Now, by Schur's first lemma we have that either $\Lambda_\lambda \equiv 0$ *or* Λ_λ *is invertible. However, by the fundamental theorem of algebra,*

$$det(\Lambda_\lambda) = det(\Lambda - \lambda\mathbb{1}) \tag{3.43}$$

is a polynomial of order n *and has at least one root. If we choose* λ *to be this root, the polynomial is zero and* $det(\Lambda_\lambda) = 0$*, meaning it is not invertible. Therefore, by Schur's first lemma:*

$$\Lambda_\lambda \equiv 0\ \Rightarrow \Lambda = \lambda\mathbb{1} \tag{3.44}$$

□

The third lemma is proven in the same spirit using the first two lemmas. The proof is left as an exercise.

The proof relies strongly on the fact that the representations we consider are irreducible. Therefore, Schur's lemma can be seen as a tool to prove representations are reducible by finding a non-zero, non-invertible intertwiner satisfying Eq. 3.34.

Notice that Schur's third lemma means that, if we find an operator which commutes with all the elements $D(g)$ of the representation, the eigenvalues of this operator can be used to label the different irreducible representations in the direct sum $D = \oplus_i D_i$. For example, in quantum mechanics, the representations of the rotation group commutes with the squared angular momentum operator $J^2 \equiv J_i J_i$. The eigenvalues of this operator therefore label the irreducible representations of the rotation group. The possible eigenvalues of J^2 have the form $j(j+1)$ for half-integer j, corresponding to the total angular momentum quantum number.

Another instance concerns the Hamiltonian operator, H, which commutes by definition with the representation of the symmetry group. By Schur's lemma, the Hamiltonian is then block-diagonal diagonal in the basis where the representation decomposes into irreducible blocks.

The results on group representations we presented apply to any group, whether discrete or continuous. In particular the main result, Schur's lemma 3.1.5, applies to both cases. Whereas in some areas of solid state physics, like crystallography, discrete groups are the most crucial, continuous groups have instead the lion's share in Particle Physics. We are thus obliged to specialise our discussion and focus on Lie groups. Our first encounter in this study will be the notion of Lie algebra.

3.1.6 Lie algebras

Definition. *A **Lie algebra** $\mathcal{A}$ is an n-dimensional vector space endowed with a second multiplication, the Lie bracket or Lie product:*

$$\begin{array}{rccc} [\cdot,\cdot] & : & \mathcal{A}\times\mathcal{A} & \to & \mathcal{A} \\ & & (X,Y) & \mapsto & [X,Y] \end{array} \tag{3.45}$$

satisfying:

Antisymmetry

$$[X,Y] = -[Y,X] \tag{3.46}$$

Linearity

$$[aX+bY,Z] = a[X,Z] + b[Y,Z] \tag{3.47}$$

Jacobi identity

$$[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0 \tag{3.48}$$

Notice that, even though the Lie bracket $[\cdot,\cdot]$ is represented with the same symbol as the usual commutator of matrices, one should keep in mind that the elements of a Lie Algebra do not necessarily have to be matrices and therefore the Lie product does not necessarily have to be a matrix commutator. It is true that, for any n, the set of $n\times n$ matrices is a vector space which, with the operation of matrix commutation, does form a Lie algebra. The properties of Antisymmetry and Linearity are immediately obvious, while the satisfaction of the Jacobi identity is easily verified by explicitly writing the commutators. This fact justifies the notation used to represent the Lie product. However the abstract notion of Lie algebra remains much more general. For instance we have previously listed the properties of Poisson brackets, and we can now see that the space of smooth functions on the phase space, with the Poisson bracket, forms indeed a Lie Algebra. Interestingly, this fucntional space is infinite-dimensional, and therefore the corresponding Lie algebra is also infinite-dimensional.

Since $\mathcal{A}$ is a vector space, we can find a basis $\{X_i\}$ for $i = 1,\dots,q$ with $q =$ the dimension of $\mathcal{A}$, such that any of its element A can be written as

$$A = \alpha^1 X_1 + \cdots + \alpha^q X_q \equiv \alpha^i X_i \tag{3.49}$$

The coefficients α^i can belong to either $\mathbb{R}$ or $\mathbb{C}$, corresponding to either a real or a complex Lie Algebra. By the property of linearity, the Lie product of any two elements A and B is then fully defined by the product among the basis vectors

$$[X_i, X_j] = i f_{ij}^k X_k \tag{3.50}$$

according to

$$[A, B] = [\alpha^i X_i, \beta^j X_j] = i\alpha^i \beta^j f_{ij}^k X_k \,. \tag{3.51}$$

The f_{ij}^k, a set of numbers called the *structure constants*, then completely determine the Lie algebra, very much like the metric $g_{ij} = e_i^T \cdot e_j$ completely determines the geometry in standard Riemannian spaces. At this stage, one may wonder why there is an i factor in eq. (3.50). The reason is that in physical applications we will be interested in the case where the X_i are hermitian matrices or hermitian operators acting on the Hilbert space. The factor of i then guaramteed that the structure constants are real numbers. Notice also that the propery of antisymmetry and the Jacobi identity imply that the structure constants f_{ij}^k cannot be any set of numbers, and must instead satisfy the following constraints:

$$\begin{gathered} f_{ij}^k = -f_{ji}^k \\ f_{ij}^\ell f_{\ell k}^m + f_{jk}^\ell f_{\ell i}^m + f_{ki}^\ell f_{\ell j}^m = 0 \end{gathered} \tag{3.52}$$

Conversely, given a set of numbers satisfying this constraint, it is possible to construct a Lie algebra.

We can think of very general Lie algebras. However, as the following theorem precisely states, there is always a way to reduce the problem to the study of matrices.

Ado's theorem. *Any finite-dimensional Lie algebra can be faithfully represented by $n \times n$ matrices, for n large enough.*

Thus, we have a convenient way to work with finite-dimensional Lie algebras. What about the infinite-dimensional cases? The way we understand this case in Physics is as a limit of finite-dimensional cases.

3.1.7 Lie Groups and Topology

We have already recalled (even if superficially) some basic notions in differential geometry. We would now like to mention/recall a few more, especially in connection to topology. These will be necessary to properly qualify the connection between Lie algebras and Lie groups.

As we have already mentioned a smooth manifold can be thought of as a surface embedded in some $\mathbb{R}^n$ space, for n large enough. The following examples may help clarifying:

1. The one-dimensional sphere S_1, the circle, can be embedded in $\mathbb{R}^2$. More generally, the n-dimensional sphere can be embedded in $\mathbb{R}^{n+1}$.
2. We are not forced to consider spheres. Paraboloids and hyperboloids are other surfaces that can be embedded in higher dimensional real spaces.
3. The de Sitter space that appears in General Relativity, and which may describe the future of our Universe, can be viewed as a 4-dimensional surface embedded in $\mathbb{R}^5$. More precisely given the coordinates in 5-dimensional Minkowsky space $x^M = x^0, \ldots, x^4$ and the metric $\eta_{MN} = \text{diag}(-1, 1, 1, 1, 1)$, de Sitter space is the set of points satisfying $x^M x^N \eta_{MN} = R^2$. The parameter R is the Hubble radius of the 4D geometry on the embedded surface.

The are a few facts about the topological structure of manifolds that are relevant for our discussion. They are linked to the notion of connectedness.

Definition. *A manifold is said to be connected if any two of its points can be continuously connected by a line inside the manifold.*

If a manifold is not connected, we say it is disconnected.

A mathematician may phrase the definition differently. They would interpret the manifold as a topology, roughly speaking the set of all open sets. A disconnected manifold is a manifold that can be written as the union of two disjoint open sets. A connected manifold cannot. However, the two pictures are equivalent.

Definition. *A manifold is said to be simply connected if any closed curve in it can be continuously contracted to a point.*

Some simple examples

1. The 2-sphere S^2 is simply connected.
2. A torus, of any dimensionality, is not simply connected. In particular the 1-dimensional torus, which coincides with the circle S^1, is not simply connected.

The simplest example relevant to physics is the group of 3×3 real orthogonal matrices: $O(3)$. This group, we recall, is given by the set of 3×3 real matrices M satisfying $MM^T = 1$. From its definition, it can immediately be seen to be disconnected. This is because

$$\det(MM^T) = 1 \implies \det(M)\det(M^T) = 1 \implies \det(M)^2 = 1 \implies \det M = \pm 1 \tag{3.53}$$

so that the group consists of, at least, two components, one with determinant $+1$ and the other with determinant -1. As the determinant is a continuous function of the matrix elements, there is no way to connect these two components: $O(3)$ is a disconnected manifold. Notice that any of these two components could in principle further decompose into disconnected components, but in reality they don't: $O(3)$ consists of just two disconnected components. How can we view $O(3)$ as a manifold? It can be viewed as a subset of $\mathbb{R}^9$ to which we apply 6 constraints. We thus get a 3-dimensional manifold. The component where the determinant is $+1$ contains the identity and is a subgroup: $SO(3)$. It is a subgroup simply because given any two matrices g_1 and g_2 the determinant of their product satisfies $\det(g_1 g_2) = \det(g_1)\det(g_2)$, and therefore it is positive if the determinants of g_1 and g_2 are. As we already mentioned, $SO(3)$ is connected and contains the identity. It thus contains all the points that can be continuously connected to the identity: it is the *connected component* of the identity in $O(3)$.

In the case of $O(3)$, as we have just argued, the connected component of the identity is itself a group, $SO(3)$. This property is indeed true for any Lie Group as one can prove that :

The connected componentof the identity of a Lie group, i.e. the set of points that can be continuously connected to the identity, is itself a Lie Group and is connected.

The proof of this result relies on the continuity of the group product. Any two elements g_1 and g_2 are connected to the identity, if there exists two continuous mappings f_1 and f_2 of $t \in [0,1] \to G$ such that $f_i(0) = e$, $f_i(1) = g_i$ for $i = 1, 2$. Consider then the product $f_1(t)f_2(t)$: as t varies from 0 to 1, by the continuity of the group product, this product provides a curve that continuously connects e to $g_1 g_2$. The product of two elements connected to the identity is then also connected to the identity. A similar argument can be made to prove that if g is connected to the identity also g^{-1} is. Thus proving that the connected component of the identity is a group.

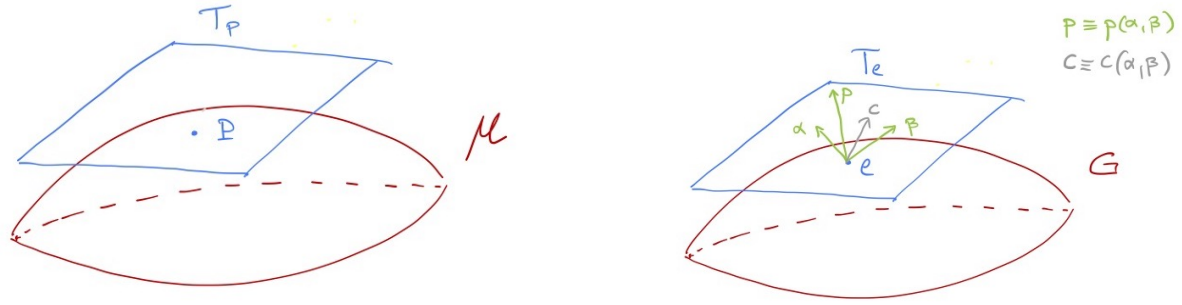
3.1.8 Lie Theorems

Central in the study of Lie groups and Lie algebras is a set of three theorems due to Sophus Lie. We are now ready to state, and partly prove, them.

Lie theorem I. *The tangent space T_g at any $g \in G$ defines one and the same Lie algebra $\mathcal{A}_G$.*

Lie theorem II. *Given an abstract Lie algebra $\mathcal{A}$, one can always construct an associated Lie group. (Indeed there generally exists more than one group associated to the same $\mathcal{A}$)*

Lie theorem III. *Any Lie algebra $\mathcal{A}$ is associated to one and only one connected and simply connected Lie group $G_{\mathcal{A}}$.*



(a) (b)

Figure 3.1: (a): Schematic picture of the tangent plane T_P at P. (b): Group operations projected on T_e, the tangent space at the identity of a Lie group G

We will prove Theorem I, sort of prove Theorem II, and not prove Theorem III. The interested student is referred to the excellent book "Lie Algebras, Lie Groups and some of their Applications" by Robert Gilmore.

Theorem I.

Before starting we should refresh/illustrate the notion of tangent space. We will proceed like physicists, hoping no mathematician will ever read this, but secretely desiring some of them will stumble on it one day. Consider a manifold $\mathcal{M}$ and picture it as a D-dimensional surface embedded in $\mathbb{R}^n$, with $n > D$ (see Fig.3.1). Consider then any point $P \in \mathcal{M}$ and construct the D-dimensional plane T_P tangent at $\mathcal{M}$ in P. Locally, in a sufficiently small neighbourhood of P, we can parametrise the points on the manifold $\mathcal{M}$ by projecting them onto the tangent plane T_P. Mathematically, this defines a map $\mathcal{M} \to T_P$. The plane T_P, with the choice of the origin at P, is a linear, real, vector space of dimension D. The manifold $\mathcal{M}$ is therefore locally isomorphic to a linear vector space.

We can now specialize to the case where $\mathcal{M}$ is a Lie group G and $P = e$, the identity element. T_e will be $\mathbb{R}^n$ where n the dimension of G. In a sufficiently small neighbourhood of e we can parametrize the group elements by the $\mathbb{R}^n$ coordinate vector $\alpha \equiv (\alpha_1, \dots \alpha_n)$ of their projection on T_e. Choosing e as the origin in T_e, its coordinate vector is $\alpha = 0$. [2] Close to the origin, group product and inversion can be viewed as operations on elements of the tangent space $T_e \equiv \mathbb{R}^n$. In particular the product corresponds to a mapping $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$

$$g(\alpha)g(\beta) = g(p(\alpha,\beta)) \qquad (\alpha,\beta) \mapsto p(\alpha,\beta) \tag{3.54}$$

while the inversion to a mapping $\mathbb{R}^n \to \mathbb{R}^n$

$$g(\alpha)^{-1} = g(r(\alpha)) \qquad \alpha \mapsto r(\alpha)\,. \tag{3.55}$$

As $g(0) = e$ these mappings satisfy

$$p(\alpha,0) = p(0,\alpha) = \alpha \qquad p(\alpha, r(\alpha)) = 0 \qquad r(0) = 0\,. \tag{3.56}$$

Now, as we have already explained, p and r are not just smooth function but analytic functions admitting a convergent Taylor expansion. For α, β small enough, we can thus reliably expand p and r to the lowest powers of their arguments. Taking eqs. (3.56) into account one easily concludes that the expansion takes the form

$$p^i(\alpha,\beta) = \alpha^i + \beta^i + T^i_{jk}\alpha^j\beta^k + B^i_{jkl}\alpha^j\alpha^k\beta^\ell + \dots \tag{3.57}$$

$$r^i(\alpha) = -\alpha^i + S^i_{jk}\alpha^j\alpha^k + \dots \tag{3.58}$$

with $S^i_{jk} = (T^i_{jk} + T^i_{kj})/2$. Notice indeed that S^i_{jk} is by construction symmetric in the jk indices, while T^i_{jk} is a generic tensor which decomposes in general as the sum of symmetric and antisymmetric components $T^i_{jk} = S^i_{jk} + A^i_{jk}$. In fact the tensor S^i_{jk} fully depends on the choice of coordinates. In particular, by choosing coordinates

[2] We deal with the tangent space, in the physicist's way: we work in a small neighbourhood of e where the tangent plane approximate well the group manifold. Practically that means we assume α is small enough. In mathematics instead, the tangent space is defined is such a way that α remains infinitesimally small at each stage. We will stick to the most pragmatic way and keep α small.

$\alpha'^i = \alpha^i - S^i_{jk}\alpha^j\alpha^k + O(\alpha^3)$, the inversion function takes the form $r'(\alpha') = -\alpha' + O(\alpha'^3)$. In fact one is easily convinced that coordinates can be chose such that $r'(\alpha') = -\alpha'$ exactly. In other words, we can choose coordinates where $S^i_{jk} = 0$ and $T^i_{jk} = A^i_{jk}$ is antisymmetric in jk. In what follows we shall adopt such simpler coordinates and assume antisymmetric T^i_{jk}.

Interestingly, in the limiting case of arbitrary small α and β, eqs. (3.57,3.58) imply that the group product and inversion linearize to respectively vector sum and inversion on the tangent space

$$p(\alpha,\beta) \to \alpha + \beta \qquad\qquad r(\alpha) \to -\alpha \tag{3.59}$$

We thus conclude that the group structure of our manifold G naturally endows the tangent space T_e with the standard vector space operations of sum and inversion.

Considering the group commutator

$$g(\alpha)^{-1}g(\beta)^{-1}g(\alpha)g(\beta) = g(c(\alpha,\beta)) \tag{3.60}$$

one indeed finds that the tangent space T_e ineherits one additional operation. The group commutator defines a mapping $c : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$. Notice that interchanging α and β eq. (3.60) becomes $g(c(\alpha,\beta))^{-1}$, which at lowest order in the Taylor expansion corresponds to $g(-c(\alpha,\beta))$. This implies that, at the lowest order, $c(\alpha,\beta)$ must be antisymmetric under the exchange of α and β. Indeed by eqs. (3.57,3.58) one finds

$$c^i(\alpha,\beta) = (T^i_{jk} - T^i_{kj})\alpha^j\beta^k + O(\alpha^2\beta,\alpha\beta^2) \equiv 2T^i_{jk}\alpha^j\beta^k + O(\alpha^2\beta,\alpha\beta^2)\,. \tag{3.61}$$

where we made use of $T^i_{jk} = -T^i_{kj}$. At lowest order in the Taylor expansion the above equation defines a bilinear antisymmetric product on the tangent space $(\alpha,\beta) \mapsto [\alpha,\beta]$ with

$$[\alpha,\beta]^i = 2T^i_{jk}\alpha^j\beta^k\,, \tag{3.62}$$

This results provides part of the conditions for having a Lie algebra, where the T^i_{jk} play the role of structure constants. In order to prove that we indeed have a Lie algebra, we now only need to check that the product satisfies the Jacobi identity. As it will be proven explicitly in one of the exercises, the Jacobi identity follows from a property which we had not yet used, the associativity of the group product

$$[g(\alpha)g(\beta)]\,g(\gamma) = g(\alpha)\,[g(\beta)g(\gamma)] \implies p(p(\alpha,\beta),\gamma) = p(\alpha,p(\beta,\gamma))\,. \tag{3.63}$$

Inserting in the above equation the Taylor series expansion of p one indeed finds T must satisfy the Jacobi identity

$$T^\ell_{ij}T^m_{\ell k} + T^\ell_{jk}T^m_{\ell i} + T^\ell_{ki}T^m_{\ell j} = 0 \tag{3.64}$$

establishing T_e inherits from the group product the structure of Lie algebra. As we shall see below, the structure constants T^i_{jk}, which we derived from the group product in a neighbourhood of the identity, determine to a significant extent the whole group G. They are sort of like the group's DNA.

The above discussion considered a small neighbourhood of the identity, its projection on T_e and the operations 3.59 and 3.62 induced by the group at lowest order in the Taylor expansion of the functions p, r and c. Using the group product, one can however "export" the above discussion to any other element $g_* \in G$ and to its tangent space T_{g_*}. Consider indeed a neighbourhood U_{g_*} of g_*, mathematically an open set containing g_*. Like per the previous discussion, for small enough U_{g_*} we can describe its elements via their projections on the tangent plane T_{g_*}. Now, left multiplication by g_*^{-1} defines a continuous and bijective mapping between U_{g_*} and a neighbourhood $g_*^{-1}U_{g_*}$ of the identity. [3] The mapping $U_{g_*} \mapsto g_*^{-1}U_{g_*}$ and its inverse $g_*^{-1}U_{g_*} \mapsto g_*g_*^{-1}U_{g_*} = U_{g_*}$ establish therefore a correspondence between T_e and T_{g^*}.[4] That allows to export to T_{g^*} the notions of sum and Lie product on T_e.

[3] $g_*^{-1}U_{g_*}$, consisting of the points obtained by acting with g_*^{-1} on any element of U_{g_*}, is a neighbourhood of e because the group product $g_*^{-1}g$ is a continuous fuction of g and because $e \in g_*^{-1}U_{g_*}$.

[4] Notice that working as usual at the lowest order, i.e. for points infinitesimally close to g_*, this is a linear mapping between vector spaces. In particular the origins e and g_* of respectively T_e and T_{g^*} are mapped into one another.

Indeed any two elements $g_1, g_2 \in U_{g_*}$ correspond to elements $g_*^{-1}g_1$ and $g_*^{-1}g_2$ in $g_*^{-1}U_{g_*}$. Inversion, product and commutation of the latter, as we have shown, define at the lowest order Lie algebra operations in T_e. Exporting the results of these operation back to U_{g_*} through left multiplication by g_*

$$\begin{aligned}
&\text{inversion} && g_1 && \to && g_*(g_*^{-1}g_1)^{-1} = g_* g_1^{-1} g_* && (3.65)\\
&\text{product} && (g_1, g_2) && \to && g_*(g_*^{-1}g_1)(g_*^{-1}g_2) = g_1 g_*^{-1} g_2 && (3.66)\\
&\text{commutator} && (g_1, g_2) && \to && g_*(g_*^{-1}g_1)^{-1}(g_*^{-1}g_2)^{-1}(g_*^{-1}g_1)(g_*^{-1}g_2) = g_* g_1^{-1} g_* g_2^{-1} g_1 g_*^{-1} g_2 && (3.67)
\end{aligned}$$

we obtain a set of operations on T_{g^*} that are isomorphic to eqs. (3.59, 3.62) and thus define the same Lie algebra structure on T_{g^*}.

Theorem II.

We must prove that, given a Lie algebra $\mathcal{A}$ we can construct a Lie group $G_{\mathcal{A}}$ whose Lie algebra, constructed on the basis of Theorem I, is precisely $\mathcal{A}$. The proof of this result crucially relies on Ado's Theorem, which states that, given the abstract Lie algebra 3.50, we can represent it by a set of $n \times n$ matrices $\tilde{X}_i$

$$[\tilde{X}_i, \tilde{X}_j] = i f_{ij}^k \tilde{X}_k \tag{3.68}$$

The claim now is that the set of matrices

$$D(\alpha) = e^{i\alpha^i \tilde{X}_i} \equiv \sum_k \frac{(i\alpha \cdot \tilde{X})^k}{k!} \tag{3.69}$$

obtained by *exponentiating* the Lie algebra corresponds to precisely such group $G_{\mathcal{A}}$. In order to prove this result we must check that the set $\{D(\alpha)\}$ is mapped onto itself under inversion and multiplication. For inversion this is obvious as $D(\alpha)^{-1} = D(-\alpha)$. Concerning the product, we must prove

$$D(\alpha)D(\beta) = e^{i\alpha_i X_i} e^{i\beta_j X_j} = e^{i\bar{p}^k(\alpha,\beta)X_k} = D(\bar{p}(\alpha,\beta))\,. \tag{3.70}$$

for some function $\bar{p}^k$ which describes the group product for this choice of coordinates. This equation is proven using the *Campbell-Baker-Haussdorf* (CBH) formula for the product of exponentials of matrices

$$\exp(A)\exp(B) = \exp\left(A + B + \frac{1}{2}[A,B] + \frac{1}{12}([A,[A,B]] + [B,[B,A]]) + \dots\right) \tag{3.71}$$

where the dots represent an infinite series that involves only multiple commutators of A and B. Applying CBH to the case where $A = i\alpha^k \tilde{X}_k$ and $B = i\beta^j \tilde{X}_j$ and using eq (3.68), we conclude that the exponential on the right-hand side in the above equation is also given by a linear combination of X_i. This proves eq. (3.70). Notice moreover that the whole CBH series in the exponent, and therefore the product function $\bar{p}^k(\alpha,\beta)$ is fully determined by the structure constants f_{ij}^k. The latter are indeed the DNA of a Lie group!

Having established that the set 3.69 forms a Lie group it remains to show that its Lie algebra is precisely $\mathcal{A}$. Around the origin the group operations and the Lie algebra operations precisely correspond as dictated by Lie Theorem 1

$$D(\alpha)D(\beta) = \mathbb{1} + i(\alpha_i + \beta_i)\tilde{X}_i + \dots \tag{3.72}$$

$$D(\alpha)^{-1} = \mathbb{1} - i\alpha_i \tilde{X}_i + \dots \tag{3.73}$$

$$D(\alpha)^{-1}D(\beta)^{-1}D(\alpha)D(\beta) = \mathbb{1} + i[\tilde{X}_j, \tilde{X}_k]\alpha^j\beta^k + \dots \tag{3.74}$$

so $\mathcal{A}$ is indeed the Lie algebra of the group.

Two comments are now in order. First notice that the set of elements $D(\alpha)$, by the continuity of the exponential function, can be continuously connected to the identity. This is for instance done by mapping $\alpha_i \mapsto t\alpha_i$ and then letting t go from 1 to 0. We are thus let to conclude that the group $G_{\mathcal{A}}$ generated by the exponential map should be connected. The second comment is that we have been a bit superficial in our application of CBH: this formula gives an infinite series in the exponent, which raises issues of convergence. In other words, for some products the function $\bar{p}$ in the exponent could diverge, While our reasoning should work (in fact it does!) for sufficently small α's and $\beta's$, i.e. close to the origin, it is not obvious what happens globally. Otherwise stated: how much of $G_{\mathcal{A}}$ is covered by the set $\{\exp(i\alpha \cdot X)\}$?

The answer to the above question is given by the following two theorems, which we state without proof.

Theorem. *If G is a compact and connected Lie group*[5]*, then the exponential map covers the whole group.*

Intuitively this means that for compact groups the structure constants are such that the series defining the function $\bar{p}$ through the CBH formula is always convergent. Our naive use of CBH is then justified as well as our proof of the connectedness of $G_{\mathcal{A}}$.

It is sometimes the case in physics that the groups we have to deal with are connected but not compact. In that case another theorem applies.

Theorem. *If G is a connected but non-compact Lie group, then every element of G can be written as*

$$\prod_{n}^{N} e^{i\alpha_i^{(n)} X_i} \tag{3.75}$$

where, generally, the number N of elements in the product needed to span the whole group is finite.

For instance $SL(2,\mathbb{R})$ falls into the above category. The whole group is covered by considering the product of at most 2 exponential elements, i.e. $N = 2$.

Theorem III.

We will not prove this results, but only make a few comments. As we have seen, the basis of Lie Theorem II is Ado's theorem, which allows to represent the algebra by matrices. Exponentiating the matrix representation of the algebra we obtained a group. But which group? Ado's theorem is indeed an existence theorem, and there generally exist different, inequivalent, matrix representations of the same algebra $\mathcal{A}$. For example two representations of different dimensionality are certainly inequivalent. There is then no guarantee that by exponentiating two of of such inequivalent representations we will obtain the same group. We may, but we also may not. The resulting groups are all connected[6], but they can differ as concerns simple-connectedness. Theorem III, states that among the set of groups that we can obtain by exponentiating the representations of a certain Lie algebra, the exists a unique group that is both connected and simply connected. This group is called the *universal covering* of the Lie algebra.

A physically relevant example of the above is given by the angular momentum algebra

$$[L_i, L_j] = i\epsilon_{ijk} L_k \,. \tag{3.76}$$

This algebra, among its many representations, possesses a representation by 3×3 matrices $(L_i)_{ab} = i\epsilon_{iab}$ and one by the 2×2 Pauli matrices $(L_i)_{ab} = (\sigma_i)_{ab}/2$. The exponentiation of these two inequivalent represention produces respectively $SO(3)$ and $SU(2)$. While for $SU(2)$, the manifold S_3, is connected and simply connected, the $SO(3)$ group manifold is $S_3/\mathbb{Z}_2$, which is connected but not simply connected.

To conclude this section we should emphasize some practical aspects concerbning the use of the exponential map. First, notice that, given any matrix representation of the Lie algebra of a certain group, we can construct a matrix representation of the group simply by exponentiating the Lie algebra. Secondly, given instead a matrix representation $\tilde{D}(\alpha)$ that is not in exponential form, we can always change variables by bringing it in exponential form. All that is needed for that purpose is to consider the expansion near the origin $\tilde{D}(\alpha) = \mathbb{1} + i\alpha^i \tilde{X}_i + \dots$. One can easily conclude by following through the same steps of the proof of Lie Theorem I, that the $\tilde{X}$ form the Lie algebra of the group. An alternative parametrization of the group elelemnts can then be obtained by exponentiating $\tilde{X}_i$ or, equivalently, by the limit procedure

$$\boxed{D(\alpha) \equiv \lim_{n\to\infty} \tilde{D}(\alpha/n)^n = \lim_{n\to\infty} \left(1 + i\frac{\alpha_i}{n}\tilde{X}_i + \mathcal{O}(1/n^2)\right)^n \equiv e^{i\alpha_i \tilde{X}_i}}\,. \tag{3.77}$$

[5]Meaning the G manifold is both compact and connected

[6]The very existence of disconnected Lie groups tells us that some groups cannot be obtained by exponentiating a set of matrices. In that case the whole group can however be obtained by combining exponentiation with a discrete set of group elements. That will be seen explicitly for the Poincarè group.

We conclude that finding the representations of a Lie group is as hard as finding the representations of its Lie algebra. The latter is a rather manageable task, as you are already familiar from the study of angular momentum, and as well we see better when studying the Ploincarè and Lorentz groups.

3.1.9 The Adjoint Representation

Consider *any* matrix representation with basis (X_i) of the Lie algebra of a group G. By exponentiating it we obtain a representation $D(\alpha) \equiv D(g(\alpha))$:

$$D(\alpha) = e^{i\alpha^i X_i} \tag{3.78}$$

of G . Consider now an element $v = v^i X_i$ in the Lie algebra and consider the mapping

$$v \mapsto e^{i\alpha^j X_j}\, v\, e^{-i\alpha^j X_j} \tag{3.79}$$

We claim the matrix on the right-hand side also belongs to the Lie algebra. At first order, this is seen by linearizing $D(\alpha) \simeq \mathbb{1} + i\alpha^i X_i$

$$\begin{aligned}(\mathbb{1} + i\alpha^j X_j)v(\mathbb{1} - i\alpha^j X_j) &= v + i\alpha^j [X_j, v] \\ &= v + i\alpha^j v^k [X_j, X_k] \\ &= v^i X_i + i\alpha^j v^k i f^l_{jk} X_l \\ &= (v^i - \alpha^j v^k f^i_{jk}) X_i \\ &\equiv v'^i X_i\end{aligned} \tag{3.80}$$

which belongs the Lie algebra. As shown in one of the exercises, this remains true to all orders. We can make now a further claim:$D_{Adj}(\alpha)$ defined by

$$D_{Adj}(g(\alpha)) :\ v \mapsto D(g(\alpha))\, v\, D(g(\alpha))^{-1} \tag{3.81}$$

is a representation of the group. Clearly, it satisfies the representation axioms as it acts linearly on v and respects the group product

$$D_{Adj}(g(\beta))D_{Adj}(g(\alpha)) : v \mapsto D(g(\beta)) \left[D(g(\alpha))\, v\, D(g(\alpha))^{-1}\right] D(g(\beta))^{-1} = [D(g(\beta))D(g(\alpha))]\ v\ [D(g(\beta))D(g(\alpha))]^{-1}\ . \tag{3.82}$$

Moreover, according to eq. (3.80), on the v vector coordinates the mapping takes the form

$$v^i \mapsto v^i - \alpha^j v^k f^i_{jk} = (\delta^i_k - \alpha^j f^i_{jk}) v^k \equiv \left(\delta^i_k + i\alpha^j (\tilde{X}_j)^i_k\right) v^k \tag{3.83}$$

with $(\tilde{X}_j)^i_k = i f^i_{jk}$, so that the infinitesimal generators for the adjoint representation are simply given by the structure constants. The structure constants f^i_{jk} are a set of numbers with three indices. If we single one index out (say j), we are left with a matrix, with rows and colums labelled by i and k. The resulting matrix, as we have shown offers a particular representation of the Lie algebra, known as the *adjoint representation*. Indeed, satisfaction of the Jacobi identity by the f^i_{jk} implies the $(\tilde{X}_j)^i_k$ precisely satisfy the Lie algebra. Finally, by exponentiating eq. (3.83), we can represent the arbitrary finite transformation of eq. (3.79) as

$$v^i \mapsto D_{Adj}(g(\alpha))^i_j v^j \equiv (e^{i\alpha^k \tilde{X}_k})^i_j v^j\ . \tag{3.84}$$

The matrices $D_{Adj}(g(\alpha))^i_j$ concrete express the group elements in the adjoint representations. Notice that in order to construct them we just need the structure constants!

As it is the case for any representation, we can equivalently view the group action (3.80) on the vector v either as a transformation of its coordinates v^i or as a tranformation of the basis vectors X_i: $v \to X_i v'^i \equiv X'_i v^i$. According to eq. (3.79) and eq. (3.80), the action of $g(\alpha)$ on X_i can then be written as

$$g(\alpha) : X_i \mapsto e^{i\alpha^j X_j} X_i e^{-i\alpha^j X_j} = X_i D_{Adj}(g(\alpha))^i_j \tag{3.85}$$

At the infinitesimal level this amounts to $X_i \mapsto X_i + i\alpha^j [X_j, X_i] + \dots$. Thus the commutator

$$[X_j, X_i] = i f^k_{ji} X_k \tag{3.86}$$

can be viewed as the transformation operated on X_i by the infinitesimal Lie group transformation controlled by X_j. In jargon, we say the commutator expresses how X_i transforms under the action of X_j. In other words, the infinitesimal action of the group on any element of the Lie algebra is given by the commutator.

In what follows we will now put all this machinery to work in the crucial symmetry of fundamental physics.

3.2 Lorentz and Poincaré groups

3.2.1 Construction

Let us start with a bit of history. Before the 20^{th} century, the physics of mechanical systems was governed by Newton's laws of motion. These laws are left invariant by the action of a symmetry group: the Galileo group. As we saw before, its realization on space and time is given by

$$\vec{x} \mapsto \vec{x}' = R(\alpha, \beta, \gamma)\vec{x} + \vec{v}_0 t + \vec{x}_0 \tag{3.87}$$
$$t \mapsto t + t_0 \tag{3.88}$$

which summarise space and time homogeneity ($\vec{x}_0,\ t_0$), space isotropy ($R \in SO(3)$) and the equivalence of all inertial observers, i.e. invariance under constant boosts ($\vec{v}_0$). Galileo's group is a 10 dimensional Lie group. This described pretty much all of physics before Maxwell came up with his set of equations: these were not invariant under the Galileo group! At that time, three options seemed to be given:

i. Maxwell equations are wrong.

ii. Galilean relativity does not apply to electromagnetism: Maxwell equations only apply in a preferential reference frame that is at rest with respect to a fluid filling the universe, the aether.

iii. Galileo's group is in reality not describing the symmetries of Nature: Maxwell is right and at the same time all inertial observers describe physics by the same laws. Therefore, the group must be different (and as we shall see it is simply a deformation of Galileo's group).

The first option seemed crazy already back then, given the great success the Maxwell equations were having in describing electromagnetic phenomena. The second option looked at odds with a series of experiments trying to detect the aether, most prominently the Michelson-Morley experiment, and it is the third option that emerged as the correct one. Indeed work on this possiblity started around the end of the 19^{th} century. First Voigt (1887) and then Lorentz (1895) led the way, though they only had an incomplete solution in their hands. It was Larmor (1897, 1900) and Lorentz (1899, 1904) who found the "Lorentz transformations", and it was Poincaré (1905) who

understood their group structure and gave them a name. But it was Einstein who understood how to properly formulate the laws of physics according to the third option. In that formulation, which appeared in a famous 1905 paper, and took the name of *theory of special relativity* the symmetries of physics were derived using solely two principles

I. Spacetime is homogeneous, isotropic and the laws of Nature are the same for all inertial observers.

II. The speed of light c is the same in all frames.

The same principles can be encapsulated in the request that *the symmetries of physics be given by the group of transformations that preserves the form of the D'Alembert operator*, which is the differential operator describing the propagation of light waves. We will now study the consequences of this request, but for that purpose we must first introduce some notation. This simply amounts to defining the so called Minkowsky metric[7]

$$\eta^{\mu\nu} = \text{diag}(+1,-1,-1,-1) \qquad \eta_{\mu\nu} = \text{diag}(+1,-1,-1,-1) \qquad \eta^{\mu\rho}\eta_{\rho\nu} = \delta^\mu_\nu = \text{diag}(+1,+1,+1,+1) \tag{3.89}$$

and to use it to "raise" and "lower" indices:

$$x_\mu \equiv \eta_{\mu\nu}x^\nu \qquad \partial^\mu \equiv \eta^{\mu\nu}\partial_\nu = \eta^{\mu\nu}\frac{\partial}{\partial x^\nu}\,. \tag{3.90}$$

Consider now the D'Alembert operator

$$\Box \equiv \partial_t^2 - \vec{\nabla}^2 \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu \tag{3.91}$$

and perform a change of observer described by the coordinate change $x'^\mu = f^\mu(x)$. According to the above postulate this is a symmetry if

$$\Box \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu = \eta^{\mu\nu}\partial'_\mu\partial'_\nu \equiv \Box'\,. \tag{3.92}$$

By applying the chain rule this identity reads

$$\begin{aligned}\Box &= \eta^{\mu\nu}\frac{\partial x'^\rho}{\partial x^\mu}\frac{\partial}{\partial x'^\rho}\frac{\partial x'^\sigma}{\partial x^\nu}\frac{\partial}{\partial x'^\sigma} \\ &= \eta^{\mu\nu}\left(\frac{\partial x'^\rho}{\partial x^\mu}\frac{\partial x'^\sigma}{\partial x^\nu}\partial'_\rho\partial'_\sigma + \frac{\partial x'^\rho}{\partial x^\mu}\left(\partial'_\rho\frac{\partial x'^\sigma}{\partial x^\nu}\right)\partial'_\sigma\right) = \eta^{\rho\sigma}\partial'_\rho\partial'_\sigma\end{aligned} \tag{3.93}$$

which implies (using compact notation for partial derivatives)

$$A)\ \eta^{\mu\nu}\partial_\mu x'^\rho\partial_\nu x'^\sigma = \eta^{\rho\sigma} \qquad\qquad B)\ \Box x'^\sigma = 0\,. \tag{3.94}$$

The identity B) corresponds to the request that the coefficient of the term of first order in ∂' derivatives vanish

$$\eta^{\mu\nu}\partial_\mu x'^\rho\partial'_\rho\partial_\nu x'^\sigma = \eta^{\mu\nu}\partial_\mu\partial_\nu x'^\sigma = \Box x'^\sigma = 0 \tag{3.95}$$

It turns out that condition A implies that x'^μ must be at most linear in x^ν: its second derivatives must vanish and B is automatically satisfied. Let us then study condition A. Multiplying at both sides by $\partial'_\sigma x^\lambda$ it becomes

$$\partial^\lambda x'^\rho = \partial'^\rho x^\lambda \tag{3.96}$$

Taking the derivative of this equation, repeatedly using the same very equatiom and the commutativity of partial derivatives we obtain a chain of identities

$$\partial^\nu\partial^\lambda x'^\rho = \partial^\nu\partial'^\rho x^\lambda = \partial^\nu x'^\mu\partial'_\mu\partial'^\rho x^\lambda = \partial^\nu x'^\mu\partial'^\rho\partial^\lambda x'_\mu = \partial'^\mu x^\nu\partial'^\rho\partial^\lambda x'_\mu \tag{3.97}$$

$$= -\partial'^\rho\partial'^\mu x^\nu\partial^\lambda x'_\mu = -\partial'^\mu\partial^\nu x'^\rho\partial^\lambda x'_\mu = -\partial^\lambda\partial^\nu x'^\rho \tag{3.98}$$

where when going from the first to the second line we used $\partial'^\rho(\partial'^\mu x^\nu\partial^\lambda x'_\mu) = \partial'^\rho\eta^{\nu\lambda} = 0$. The first and the last steps in the chain imply $\partial^\nu\partial^\lambda x'^\rho = 0$. We thus conclude that the change of variables must have the general form $x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu$ with $\Lambda^\mu{}_\nu$ and a^μ constants. The matrix $\Lambda^\mu{}_\nu$ is however further subjected by constraint A to satisfy $\eta^{\mu\nu}\Lambda^\rho{}_\mu\Lambda^\sigma{}_\nu = \eta^{\rho\sigma}$. All these results combine to give the symmetry transformations of special relativity

[7] At this stage we are calling $\eta_{\mu\nu}$ a metric without any particular motivation. Nonetheless we follow the notational convention of differential geometry, where given a metric $g_{\mu\nu}$ one indicates by $g^{\mu\nu}$ its inverse. Then, since our metric is diagonal with only ±1 entries, the metric and its inverse are given by the same matrix $\text{diag}(+1,-1,-1,-1)$.

$$\boxed{x^\mu \;\mapsto\; x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu, \quad \eta^{\mu\nu}\Lambda^\rho{}_\mu\Lambda^\sigma{}_\nu = \eta^{\rho\sigma}} \tag{3.99}$$

These transformations form a group, the *Poincaré group*, sometimes also called the *inhomogeneous Lorentz group*.

Before studying the properties of the Poincarè group, let us complete our notation and add some concepts. Defining $\Lambda_\rho{}^\sigma \equiv \eta_{\rho\mu}\eta^{\sigma\nu}\Lambda^\mu{}_\nu$ we can rewrite the above constraint as

$$\eta^{\mu\nu}\Lambda^\rho{}_\mu\Lambda^\sigma{}_\nu = \eta^{\rho\sigma} \quad \Longrightarrow \quad \Lambda_\mu{}^\rho\Lambda^\nu{}_\rho = \delta^\nu_\mu \tag{3.100}$$

which, in standard matrix notation, means that $\Lambda_\mu{}^\rho = (\Lambda^{-1T})_\mu{}^\rho$. [8] Given that $AB = \mathbb{1}$ implies $BA = \mathbb{1}$, we then also have $\Lambda_\rho{}^\mu\Lambda^\rho{}_\nu = \delta^\mu_\nu$, from which we can deduce

$$\Lambda_\rho{}^\mu\Lambda^\rho{}_\nu = \delta^\mu_\nu \quad \Longrightarrow \quad \eta_{\rho\sigma}\eta^{\mu\lambda}\Lambda^\sigma{}_\lambda\Lambda^\rho{}_\nu = \delta^\mu_\nu \quad \Longrightarrow \quad \boxed{\eta_{\rho\sigma}\Lambda^\sigma{}_\mu\Lambda^\rho{}_\nu = \eta_{\mu\nu}} \tag{3.101}$$

Notice that $\Lambda_\mu{}^\nu$ controls the transformation of x_μ

$$x'_\mu = \Lambda_\mu{}^\nu x_\nu + a_\nu \,. \tag{3.102}$$

The 4-vectors, like x^μ, whose transformation is controlled by $\Lambda^\mu{}_\nu$ are called *contravariant*, while those like x_μ, whose transformation is controlled by $\Lambda_\mu{}^\nu$ are called *covariant*. These different properties is what motivates the upper and lower index notation. For partial derivatives we have

$$\partial_\mu = \partial_\mu x'^\nu \partial'_\nu = \Lambda^\nu{}_\mu \partial'_\nu \quad \Longrightarrow \quad \boxed{\partial'_\mu = \Lambda_\mu{}^\nu \partial_\nu} \tag{3.103}$$

consistent with the assignment of a low index to $\partial/\partial x^\mu$ as it suits a covariant vector.

Finally, eq. (3.101) relates to the alternative, but equivalent, definition of the Poincarè group as the set of transformations that preserve the form of the Minkowsky space time distance

$$ds^2 \equiv \eta_{\mu\nu}dx^\mu dx^\nu \,. \tag{3.104}$$

Indeed, given a change of coordinates $x^\mu \mapsto x'^\mu(x)$ the request $ds^2 = ds'^2 \equiv \eta_{\mu\nu}dx'^\mu dx'^\nu$ immediately implies $x'^\mu(x)$ should be at most linear in x^μ. Considering then a general transformation of the form in eq. (3.99) one easily sees that invariance is achieved if and only if $\Lambda^\mu{}_\nu$ satisfies eq. (3.101)

$$ds'^2 \equiv \eta_{\mu\nu}dx'^\mu dx'^\nu = \eta_{\mu\nu}\partial_\rho x'^\mu \partial_\sigma x'^\nu dx^\rho dx^\sigma = \eta_{\mu\nu}\Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma dx^\rho dx^\sigma = \eta_{\rho\sigma}dx^\rho dx^\sigma \equiv ds^2 \,. \tag{3.105}$$

The Poincaré group $\mathbb{P}$ is a deformation of the Galileo group under whose action Maxwell's equations are invariant. It is remarkable that before special relativity was developed, the equation of classical electrodynamics already "incorporated" it. Though, it took a few decades, till Einstein's revolutionary 1905 paper, to fully appreciate that.

Let $g_1 = (\Lambda_1, a_1), g_2 = (\Lambda_2, a_2) \in \mathbb{P}$. The group product structure is as follows:

$$g_1 g_2 = (\Lambda_1\Lambda_2, \Lambda_1 a_2 + a_1) \neq g_2 g_1 = (\Lambda_1\Lambda_2, \Lambda_2 a_1 + a_2) \tag{3.106}$$

and the inverse is given by:

$$g_i^{-1} = (\Lambda_i^{-1}, -\Lambda_i^{-1}a_i) \tag{3.107}$$

[8] Eq. (3.96) precisely corresponds to $\Lambda_\mu{}^\rho = (\Lambda^{-1T})_\mu{}^\rho$ once the form $x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu$ is taken.

Consistent with being a deformation of the Galileo group, it is also described by 10 parameters. Let us count them:

$$a^\mu \rightarrow 4 \tag{3.108}$$
$$\Lambda^\mu{}_\nu \rightarrow 16 \tag{3.109}$$

It seems to have 20 parameters, however there are 10 constraints coming from

$$\eta_{\mu\nu}\Lambda^\mu{}_\alpha\Lambda^\nu{}_\beta = \eta_{\alpha\beta}, \tag{3.110}$$

leaving us with exactly 10 parameters. The set of homogeneous Poincaré transformations, that is those with $a^\mu = 0$, is easily seen to from a subset closed with respect to multiplication and inversion. It is thus a subgroup known as the *Lorentz group*. It is denoted by $O(1,3)$ much like we denote by $O(3)$ the group of homogeneous isometries of the 3-dimensional euclidean space, i.e. $\mathbb{R}^3$ endowed with the euclidean metric. More generally $O(p,q)$ is the group of homogeneous transformations that leave invariant the metric

$$g_{\mu\nu} = \mathrm{diag}(\underbrace{+1,\ldots,+1}_{p},\underbrace{-1,\ldots,-1}_{q}) \tag{3.111}$$

In other words, it is the group of *orthogonal* matrices (hence the "O"(p,q)) with respect to the metric 3.111.

3.2.2 Global structure of the Lorentz Group

This section is devoted to a detailed study of the Lorentz group. We shall first see what Eq. 3.110 implies on the latter. Taking the determinant of the equation, we obtain:

$$\det(\Lambda) = \pm 1 \tag{3.112}$$

Thus the group $O(1,3)$ is disconnected. The elements belonging to the class with $\det(\Lambda) = 1$ form a subgroup called the *proper Lorentz group* denoted $SO(1,3)$. The so-called *improper* elements do not form a group as they do not contain the identity. The other constraint concerns $\Lambda^0{}_0$, which tells us how time is measured in different frames: observers related by $\Lambda^0{}_0 > 0$ are termed *orthochronous*, while those related by $\Lambda^0{}_0 < 0$, have mutually inverted notions of past and future, and are termed *non-orthochronous*. Now, the the "00" component of Eq. 3.110 gives

$$1 = \left(\Lambda^0{}_0\right)^2 - \sum_i \left(\Lambda^i{}_0\right)^2 \;\Rightarrow\; \left(\Lambda^0{}_0\right)^2 \geq 1 \tag{3.113}$$

which implies the existence of two options, $\Lambda^0{}_0 \geq 1$ and $\Lambda^0{}_0 \leq -1$. These are clearly associated to the existence of two disconnected subsets of elements in the Lorentz group. The set with $\Lambda^0{}_0 \geq +1$, contains the identity and is closed under multiplication and inversion. It thus represents a subgroup termed the *orthochronous Lorentz group* $O^\uparrow(1,3)$. Similarly to the improper elements, the *non-orthochronous* elements do not form a group. To summarise, the Lorentz group has 4 disconnected components:

	$\Lambda^0{}_0$	$\det(\Lambda)$	Group?
$\mathscr{L}_+^\uparrow \equiv SO^\uparrow(1,3)$	$\geq +1$	$+1$	✓
$\mathscr{L}_+^\downarrow$	≤ -1	$+1$	
$\mathscr{L}_-^\uparrow$	≥ 1	-1	
$\mathscr{L}_-^\downarrow$	≤ -1	-1	

(3.114)

Notice that $\mathscr{L}_+^\uparrow \oplus \mathscr{L}_+^\downarrow = SO(1,3)$ and $\mathscr{L}_+^\uparrow \oplus \mathscr{L}_-^\uparrow = O^\uparrow(1,3)$ are two subgroups. However, we fill focus on the smallest subgroup of the Lorentz group, the proper-orthochronous Lorentz group. The three other components are connected to the latter by parity and time reversal transformations:

$$P \ : \ (t,\vec{x}) \ \mapsto \ (t,-\vec{x}) \ \in \mathscr{L}_-^\uparrow \tag{3.115}$$

$$T \ : \ (t,\vec{x}) \ \mapsto \ (-t,\vec{x}) \ \in \mathscr{L}_-^\downarrow \tag{3.116}$$

$$PT \ : \ (t,\vec{x}) \ \mapsto \ (-t,-\vec{x}) \ \in \mathscr{L}_+^\downarrow \tag{3.117}$$

$$\tag{3.118}$$

So that the full Lorentz group can be seen as the composition of the proper-orthochronous Lorentz subgroup and these spacetime transformations, schematically:

$$O(1,3) = \mathscr{L}_+^\uparrow \circ \{\mathbb{1}, P, T, PT\} \tag{3.119}$$

This van more explicitly be expressed as follows. Consider first the parity transformation:

$$x^\mu \mapsto x'^\mu = P^\mu{}_\nu x^\nu \tag{3.120}$$

where $P^\mu{}_\nu \equiv \text{diag}(1,\text{-}1,\text{-}1,\text{-}1)$.[9] Notice that $P \in \mathscr{L}_-^\uparrow$. Take then any $\Lambda_-^\uparrow \in \mathscr{L}_-^\uparrow$. One has

$$\Lambda_-^\uparrow P \equiv \Lambda_+^\uparrow \in \mathscr{L}_+^\uparrow \tag{3.121}$$

which, upon using $P^2 = \mathbb{1}$, implies:

$$\Lambda_-^\uparrow = \Lambda_+^\uparrow P \tag{3.122}$$

In other words, an arbitrary $\Lambda_-^\uparrow$ can be written as the product of an element in the orthochronous proper Lorentz group and parity P. Interestingly, since $\mathscr{L}_+^\uparrow$ can be obtained by exponentiating its Lie algebra, this property defines a simple manner to obtain an element in $\mathscr{L}_-^\uparrow$. Proceeding in exactly the same way for respectively the time-reversal operator $T^\mu{}_\nu = \text{diag}(-1,1,1,1)$ and $(PT)^\mu{}_\nu = -\mathbb{1}$ allows to conclude $\mathscr{L}_-^\downarrow = \mathscr{L}_+^\uparrow T$ and $\mathscr{L}_+^\downarrow = \mathscr{L}_+^\uparrow PT$.

[9] Be aware that $P^\mu{}_\nu$ and $\eta_{\mu\nu}$ even if they have the same entries, cannot be compared because of the different (covariant vs. contravariant) nature of their indices. A proper comparison could be made by raising one of the indices in $\eta_{\mu\nu}$ which would give $\delta^\mu{}_\nu = \text{diag}(1,1,1,1) \neq P^\mu{}_\nu$.

From now on, when mentioning the Lorentz group we shall mean $\mathscr{L}_+^\uparrow \equiv SO^\uparrow(1,3)$. Also, some authors refer to $SO(1,3)$ or $O(1,3)$ when they actually mean $SO^\uparrow(1,3)$, we shall make the same misuse of notation and denote the (proper-orthochronous) Lorentz group by $SO(1,3)$ as it is now clear which subgroup we are working with.

Now, what is the smallest possible group we can have and with which we can talk about relativity? Can we have just a little part of the proper orthochronous group $\mathscr{L}_+^\uparrow$? Or is the full subgroup required? In fact, the full $\mathscr{L}_+^\uparrow$ is required. Indeed, as soon as we have the identity and one of its neighbourhoods, we can combine the transformations and produce (for example with the exponential map) each element in the group. But what about the other subgroups? This time, there is no mathematical constrain on whether they should be used. It is only up to Nature to decide whether they are symmetries or not. As it turns out, P and T, very famous transformations in Physics, are not considered by Nature to be valid symmetries. Subtly, PT combined with charge conjugation (as we will see later) happens, in a profound way, to be a symmetry of Nature.

3.2.3 Raising and lowering

This section offers both an alternative discussion and a reiteration of the notions of contravariant and covariant, raising and lowering, which we already encountered. Notice first of all that the mapping $\Lambda \to (\Lambda^T)^{-1}$ furnishes an alternative *representation* of the Lorentz group. One can indeed easily check that the above mapping satisfies all the properties of representations, in particular $\Lambda_1\Lambda_2 \to [(\Lambda_1\Lambda_2)^T]^{-1} = (\Lambda_1^T)^{-1}(\Lambda_2^T)^{-1}$. Moreover the defining property of Lorentz tranformations, written in matrix form as

$$\Lambda^T \eta \Lambda = \eta \tag{3.123}$$

gives

$$(\Lambda^T)^{-1} = \eta \Lambda \eta^{-1}\,, \tag{3.124}$$

implying that Λ and $(\Lambda^T)^{-1}$ are equivalent representations that are obtained one from the other by the simple action of the Minkowsky metric. Defining η^{-1} as an objet with upper indices $\equiv \eta^{\mu\nu}$, the relation between the two equivalent representations conveniently reduces to raising and lowering the indices with the aid of $\eta_{\mu\nu}$ and $\eta^{\mu\nu}$. The fact that $\eta_{\mu\nu} = \eta^{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$, is just an additional incidental fact.

Putting indices back, eq. (3.124) thus reads

$$[(\Lambda^T)^{-1}]_\mu{}^\nu = \eta_{\mu\rho}\Lambda^\rho{}_\sigma \eta^{\sigma\nu} \equiv \Lambda_\mu{}^\nu \tag{3.125}$$

which further implies

$$\Lambda^\mu{}_\tau \Lambda_\mu{}^\nu = \delta^\nu_\tau \qquad\qquad \Lambda_\mu{}^\nu \Lambda^\tau{}_\nu = \delta^\tau_\mu\,. \tag{3.126}$$

A 4-vector V^μ transforming according to $V'^\mu = \Lambda^\mu{}_\nu V^\nu$ is called *contravariant*, while a vector W_μ transforming according to $W'{}_\mu = \Lambda_\mu{}^\nu W_\nu$ is called *covariant*. By eq. (3.125) we have that $V_\mu \equiv \eta_{\mu\nu}V^\nu$ is covariant, while $W^\mu = \eta^{\mu\nu}W_\nu$ is contravariant. It is easy to see that the 4-gradient $\partial/\partial x^\mu$ forms a covariant vector $\partial_\mu \equiv \partial/\partial x^\mu$. Indeed, using $x'^\mu = \Lambda^\mu{}_\nu x^\nu$ one has

$$\partial_\nu = \frac{\partial x'^\rho}{\partial x^\nu}\partial'_\rho = \Lambda^\rho{}_\nu \partial'_\rho\,. \tag{3.127}$$

which upon multiplication by $\Lambda_\mu{}^\nu$ and use of eq. (3.126) becomes

$$\partial'_\mu = \Lambda_\mu{}^\nu \partial_\nu\,. \tag{3.128}$$

3.2.4 The Lie algebra of $SO(1,3)$

Let us now derive the Lie algebra of the Lorentz group. We will later extend it to the full algebra of the Poincaré group. The Lorentz group is realized in a standard way on four vectors:

$$g(\Lambda)x^\mu = \Lambda^\mu{}_\nu x^\nu \tag{3.129}$$

Recall that to study the Lie algebra, we should consider group elements around the origin (represented by the identiy on spacetime):

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu + \mathcal{O}(\omega^2), \quad |\omega| \ll 1 \tag{3.130}$$

We shall exploit the constraints imposed by the group to derive constraints on the Lie algebra. Eq. 3.110 reads at an infinitesimal level:

$$\eta_{\mu\nu} = \eta_{\alpha\beta}(\delta^\alpha{}_\mu + \omega^\alpha{}_\mu + \mathcal{O}(\omega^2))(\delta^\beta{}_\nu + \omega^\beta{}_\nu + \mathcal{O}(\omega^2)) = \eta_{\mu\nu} + \omega_{\mu\nu} + \omega_{\nu\mu} + \mathcal{O}(\omega^2) \tag{3.131}$$

Since this equation must be true order by order in ω, we obtain:

$$\boxed{\omega_{\mu\nu} = -\omega_{\nu\mu}} \tag{3.132}$$

Therefore the matrix of generators is equivalent to a four-by-four antisymmetric matrix and thus has 6 components. In other words, the tangent space has dimension 6. We now would like to write transformations near the identity as $\Lambda = \mathbb{1} + i\alpha^A X_A + O((\alpha^A)^2)$ for some set α^A of Lie parameters and for some matrices X_A which will represent the basis of the Lorentz Lie algebra. For that purpose we define:

$$(\mathcal{J}^{\rho\sigma})^\mu{}_\nu \equiv i(\eta^{\rho\mu}\delta^\sigma{}_\nu - \eta^{\sigma\mu}\delta^\rho{}_\nu) \tag{3.133}$$

for each pair $\rho\sigma$ the above equation defines a matrix with rows and colums labelled by μ and ν. Since $(\mathcal{J}^{\rho\sigma})^\mu{}_\nu = -(\mathcal{J}^{\sigma\rho})^\mu{}_\nu$ this is a set of 6 matrices labelled by the antisymmetric pair $\rho\sigma$. Now, through some simple algebra, we can write eq. (3.130) as

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu - \frac{i}{2}\omega_{\rho\sigma}\,(\mathcal{J}^{\rho\sigma})^\mu{}_\nu + \dots \tag{3.134}$$

According to the above equation the $(\mathcal{J}^{\rho\sigma})$ form the basis of the Lorentz Lie algebra and the $\omega_{\rho\sigma}$ are the Lie parameters, i.e. the coordinates on the group manifold. At this stage, we could avoid putting the i factor. Indeed, we are still working with real spacetime. The reason it appears is that Λ will at some point be replaced by a representation acting on the Hilbert space. In this case, $\omega_{\alpha\beta}\left(\mathcal{J}^{\alpha\beta}\right)^\mu{}_\nu$ will become a hermitian operator on the Hilbert space. Hence, it is more convenient to have the factor i instead of working with anti-hermitian operators.

It is instructive to write down explicitly the $\mathcal{J}^{\rho\sigma}$ matrices. One finds

$$\mathcal{J}^{12} = i\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{J}^{23} = i\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \mathcal{J}^{13} = i\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{3.135}$$

$$\mathcal{J}^{10} = -i\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{J}^{20} = -i\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{J}^{30} = -i\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{3.136}$$

The $\mathcal{J}^{ij}$ purely act on the spacial coordinates generating rotations, while the $\mathcal{J}^{i0}$ combine the time and space coordinates are represent the genuine Lorentz boost generators. There is a crucial difference between these two sets of generators: the generators of rotations $\mathcal{J}^{ij}$ are hermitean, while the generators of boosts $\mathcal{J}^{i0}$ are anti-hermitean. The latter fact reflects the non-compactedness of the boosts as we will exemplify below.

The commutation relations of the Lie algebra of the Lorentz group, denoted $\mathfrak{so}(1,3)$, can be derived from Eq. 3.133 (exercise):

$$\boxed{[\mathcal{J}^{\mu\nu}, \mathcal{J}^{\rho\sigma}] = i\left(\eta^{\mu\sigma}\mathcal{J}^{\nu\rho} + \eta^{\nu\rho}\mathcal{J}^{\mu\sigma} - \eta^{\mu\rho}\mathcal{J}^{\nu\sigma} - \eta^{\nu\sigma}\mathcal{J}^{\mu\rho}\right)} \tag{3.137}$$

With the Lie algebra at hands, we can then construct finite Lorentz transformations through exponentiation

$$\boxed{\Lambda(\omega) \equiv e^{-\frac{i}{2}\omega_{\alpha\beta}\mathcal{J}^{\alpha\beta}}} \tag{3.138}$$

As we have already stated $\mathcal{J}^{ij}$ generate rotations while $\mathcal{J}^{i0}$ a generate boosts. The last property can for instance be seen by considering

$$\Lambda = \exp(-i\omega_{10}\mathcal{J}^{10}) \equiv \exp(-i\eta\mathcal{J}^{10}) = \exp\left(-\eta\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\right) \tag{3.139}$$

where η is called the *rapidity*. Writing the series expansion of the exponential and using simple matrix algebra one finds:

$$\Lambda = \begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0 \\ -\sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.140}$$

which is nothing but a boost in the x direction. Now, with respect to the standard form of the boosts, we have the following relations:

$$\cosh\eta = \gamma, \quad \tanh\eta = \beta \tag{3.141}$$

where, in units for which $c = 1$, $\beta = v$. Since $|v| \leq 1$, we have from Eq. 3.141:

$$\eta \in (-\infty, +\infty) \tag{3.142}$$

which reflects the non-compactness of the Lorentz group.

3.2.5 Representations

Now that we have the Lie algebra (Eq. 3.137), we can find its representations. We already know one: the *defining* representation Eq. 3.133. It is called the defining representation as the Lorentz group is defined through its action on four-dimensional spacetime (Lorentz transformations). In the same way, $SO(3)$ is defined as the group of rotations in three dimensions and as such its defining representation will be the usual rotation matrices in three dimensions. We shall however classify *all* representations of the Lorentz group.

On notation: from here on we shall indicate by $\mathcal{J}^{\mu\nu}$ the Lorentz group generators in the defining representation, that is the 4×4 matrices defined in eq. (3.133), while we shall indicate by $J^{\mu\nu}$ the generators in a generic representation of the Lorentz group. It then goes without saying that the $J^{\mu\nu}$ satisfy the same algebra as the $\mathcal{J}^{\mu\nu}$.

In representation theory, one often tries to reduce the problem to another one that is already solved. To make the Lie algebra more explicit, let us separate the $J^{\mu\nu}$ into the generators of rotations (which we already know well!) and the boost generators (of which we know less)

$$\underline{\text{Rotations}}: \quad J_i \equiv \frac{1}{2}\epsilon_{ijk}J^{jk}, \quad \theta_i \equiv \frac{1}{2}\epsilon_{ijk}\omega_{jk} \tag{3.143}$$

$$\underline{\text{Boosts}}: \quad K_i \equiv J^{i0}, \quad \eta_i \equiv \omega_{i0} \tag{3.144}$$

so that we can represent the elements of the Lorentz group as

$$\exp\left(-i\frac{\omega_{\mu\nu}}{2}J^{\mu\nu}\right) = \exp(-i\theta_i J_i + i\eta_i K_i)\,. \tag{3.145}$$

We stress that only when choosing $J^{\mu\nu} = \mathcal{J}^{\mu\nu}$ do the above group elements correspond to the 4×4 defining representation of the Lorentz group. The Lie algebra Eq. 3.137 can then be written as

$$[J_i, J_j] = i\epsilon_{ijk}J_k \tag{3.146}$$

$$[J_i, K_j] = i\epsilon_{ijk}K_k \tag{3.147}$$

$$[K_i, K_j] = -i\epsilon_{ijk}J_k \tag{3.148}$$

We recognise in Eq. 3.146 the $SU(2)$ algebra, that is the algebra of 3D rotations. Being embedded in a bigger algebra, it is here a *subalgebra*. Notice also that from Eq. 3.147, the boost generators transform as vectors under rotations. So far, there is nothing new compared to Galileo's group. The novelty of relativity comes with the non-commutativity of boosts! In a sense, the Galileo group can be obtained by a *contraction* of the Poincaré group [10]. The non-relativistic limit intuitivelly corresponds to infinitesimally small boosts, or equivalently to the limit in which the speed of light is sent to ∞. This limit can then be formally attained by rescaling the boost generators $K_i \mapsto \tilde{K}_i \equiv K_i\epsilon$ and treating ϵ as infinitesimally small. Equations 3.146, 3.147 and 3.148 now read

$$[J_i, J_j] = i\epsilon_{ijk}J_k \tag{3.149}$$

$$[J_i, \tilde{K}_j] = i\epsilon_{ijk}\tilde{K}_k \tag{3.150}$$

$$[\tilde{K}_i, \tilde{K}_j] = -i\epsilon_{ijk}J_k \times \epsilon^2 \tag{3.151}$$

and, in the $\epsilon \to 0$ limit, we recover the Lie algebra of the Galileo group.

We found that we have an $SU(2)$ subgroup spanned by the rotation generators, however the full Lie algebra still mixes these generators with the boost generators. It is therefore natural to look for combinations of generators that simplifies thecommutation relations. Consider indeed:

[10] We refer to the previously mentioned book by Robert Gilmore for a detailed discussion of this mathematical procedure, known as the Inonu-Wigner contraction.

$$\boxed{J_i^{\pm} = \frac{1}{2}(J_i \pm iK_i)} \tag{3.152}$$

in the terms of which the Lie algebra becomes:

$$[J_i^{\pm}, J_j^{\pm}] = i\epsilon_{ijk}J_k^{\pm} \tag{3.153}$$
$$[J_i^{\pm}, J_j^{\mp}] = 0 \tag{3.154}$$

In this case, the analog to equation 3.145 is:

$$\exp\left(-i\frac{\omega_{\mu\nu}}{2}J^{\mu\nu}\right) = \exp\left(-i(\theta^a - i\eta^a)J_a^- + (\theta^a + i\eta^a)J_a^+\right) \tag{3.155}$$

We managed to decompose the Lorentz algebra into two independent $SU(2)$ subalgebras:

$$\mathfrak{so}(1,3) \simeq \mathfrak{su}(2)_- \oplus_{\mathbb{C}} \mathfrak{su}(2)_+ \tag{3.156}$$

where $\oplus_{\mathbb{C}}$ denotes a complexified direct sum. This identity holds at the level of the algebra, much like the $SO(3)$ Lie algebra is isomorphic to the $SU(2)$ algebra. However it is not true that at the group level the Lorentz group decomposes into two $SU(2)$ subgroups.

Let us expand on the notion of complexified Lie algebra. In our case, it is made of vectors of the form:

$$v = \alpha_i J_i^- + \beta_i J_i^+ \quad (\alpha_i, \beta_i) \in \mathbb{C}^6 \tag{3.157}$$

We now have a 6-dimensional complex vector space. But, there exists $\mathbb{R}^6$ embeddings in $\mathbb{C}^6$. If we take $\alpha_i = \beta_i^*$, we get the Lie algebra of the Lorentz group $SO(3,1) \cong SL(2,\mathbb{C})$. If we take $(\alpha_i, \beta_i) \in \mathbb{R}^6$, we get the Lie algebra of the group $SO(3) \times SO(3) \cong SO(4)$. We can see these two algebras as living in different sections of the complexified Lie algebra.

This decomposition of the algebra nonetheless implies that we can construct all the representations of the Lorentz group if we know how to build $SU(2)$ representations. This is well known from spin quantum mechanics and we will recall the main results here.

- The irreducible representations are labelled by half integer j and of dimension $2j+1$
- The basis of each irreducible representation is a set of vectors $|m\rangle$ for $m = -j, ..., j$ such that:
$$J_3|m\rangle = m|m\rangle \tag{3.158}$$
- The combinations
$$J^{\pm} = \frac{J_1 \pm iJ_2}{\sqrt{2}} \tag{3.159}$$
acting as raising and lowering operators of J_3
- The generators J_3 and $J^{\pm}$ are explicitly found to be
$$\langle j, m'|J_3|j, m\rangle \equiv L^3_{m',m} = \delta_{m,m'}m \tag{3.160}$$
$$\langle j, m'|J^+|j, m\rangle \equiv L^+_{m',m} = \sqrt{(j+m+1)(j-m)/2}\delta_{m',m+1} \tag{3.161}$$
$$\langle j, m'|J^-|j, m\rangle \equiv L^-_{m',m} = \left(L^+_{m',m}\right)^{\dagger} \tag{3.162}$$
$$\tag{3.163}$$

- A general state can be written $|\psi\rangle = \psi_m |m\rangle$ (summation intented). The action of the generators on this state is then:

$$J_i|\psi\rangle = L^i_{m',m}\psi_m|m'\rangle \tag{3.164}$$

so that the entries of the vector changed the following way:

$$\psi_m \mapsto L^i_{m,m'}\psi_{m'} \tag{3.165}$$

As the Lorentz Lie algebra consists of two independent copies of the $SU(2)$ algebra, its irreducible representations (irreps) will simply consists of two independent irreducible representations of $SU(2)$ *glued together*, and thus be labelled by a pair of half integers

$$(j_-, j_+)\,. \tag{3.166}$$

More concretely, a vector in the irrep will be labelled by a pair of indices

$$\psi_{m_+,m_-}\,, \qquad m_- = -j_-, ..., j_- \qquad m_+ = -j_+, ..., j_+ \tag{3.167}$$

upon each of which the $SU(2)_-$ and $SU(2)_+$ algebras are independently represented by respectively the j_- and j_+ representations of $SU(2)$ we already constructed

$$J^-_i \;:\; \psi_{m_-,m_+} \to \left(L^-_{m_-,m_+}\right)^i \psi_{m_-,m_+} \tag{3.168}$$

$$J^+_i \;:\; \psi_{m_-,m_+} \to \left(L^+_{m_-,m_+}\right)^i \psi_{m_-,m_+} \tag{3.169}$$

$$\tag{3.170}$$

As $(L^-)^i$ and $(L^+)^i$ are respectively $(2j_-+1)\times(2j_-+1)$ and $(2j_++1)\times(2j_++1)$ matrices, the dimensionality of the resulting representation is

$$\dim(j_-, j_+) = (2j_++1)(2j_-+1) \tag{3.171}$$

The pair (j_-, j_+) labels the casimirs of the two components

$$J^i_- J^i_- = j_-(j_-+1) \qquad J^i_+ J^i_+ = j_+(j_++1) \tag{3.172}$$

Here is a short table describing the main representations of the Lorentz group:

(j_-, j_+)	dim	Type	Example
$(0,0)$	1	Scalar	$\pi^0, \pi^\pm$, Higgs
$(1/2,0)$	2	Left-handed spinor	Neutrinos
$(0,1/2)$	2	Right-handed spinor	Anti-neutrinos
$(1/2,0)\oplus(0,1/2)$	4	Dirac spinor	$e^\pm, p, n$
$(1/2,1/2)$	4	Vector	$\gamma, W^\pm, Z^0, g$
$(1,1)$	9	Traceless metric tensor	"Gravity"

(3.173)

As our goal isto study quantum field theory, we must now construct the representations act on fields. We will see that each representation will act on a given type of particles as indicated in 3.173.

3.2.6 Representating Lie groups on fields

The notion of group representation is general and flexible. To represent a group all we need is a vector space of the suitable dimensionality. An example of particular relevance for physics is given by the representation of groups on the linear vector space of functions, i.e. fields, on a space X. As we will see, this also allows to represent the Lie algebra in terms of differential operators and hence to straightforwardly derive the commutation relations. We will then concretely illustrate our results in the case of the Poincaré group.

Consider now a space X and a field φ [11]:

$$\begin{aligned} \varphi \ &: X \to \mathbb{R} \\ &x \mapsto \varphi(x) \end{aligned} \tag{3.174}$$

The space of such fields is obviously a linear vector space V:

- the null field $\varphi(x) = 0$ is a possible field
- if $\varphi \in V$ then $\lambda\varphi \in V$ for $\lambda \in \mathbb{R}$
- if $\varphi_1 \, , \, \varphi_2 \, \in V$ then $\varphi_1 + \varphi_2 \in V$

and as such, we can try to construct a representation of a Lie group on it. Assume indeed there exists a Lie group G that can be realized in terms of coordinate changes on X:

$$x^\mu \mapsto x'^\mu = f^\mu_g(x) \qquad\qquad g \equiv g(\alpha^1, ..., \alpha^N) \in G \tag{3.175}$$

The functions f^μ_g constitue a concrete realization of G on X. As such, the set f^μ_g satisfies the group product: $f_{g_2}(f_{g_1}(x)) = f_{g_2 g_1}(x)$. Eq. (3.175) can be viewed as a change of observers: a point x^μ in the referential of observer O is parametrized by $x'^\mu = f^\mu_g(x)$ by observer O'. In the same manner, the observers parametrize physics through respectively the field $\varphi(x)$ and $\varphi'('x)$. Now, given x and x' represent the same physical point, the two observers "measure" the same field value at these corresponding coordinates:

$$\forall x' : \varphi'(x') \equiv \varphi(x) \tag{3.176}$$

which implies the functions $\varphi'(x)$ and $\varphi(x)$ are related by :

$$\varphi'(x) = \varphi(f_{g^{-1}}(x)) \tag{3.177}$$

We now claim this gives us a representation of the group on the vector space of fields. As it will be useful, let us introduce the following notation:

$$\mathcal{D}_g[\varphi] \equiv \varphi \circ f_{g^{-1}} \tag{3.178}$$

First of all it is evident that $\mathcal{D}_g[\varphi]$ is a linear operator since

$$\mathcal{D}_g[a\varphi + b\psi](x) = a\mathcal{D}_g[\varphi](x) + b\mathcal{D}_g[\psi](x) \tag{3.179}$$

[11] Here we focus on the simplest example, corresponding to a single real field, but a similar discussion holds for a general $\varphi \ : \ X \to \mathbb{R}^n$.

Moreover $\mathcal{D}_e[\varphi] = \varphi$ as expected for the representation of the identity. It only remains to check that the product rule of the group is satisfied:

$$\mathcal{D}_{g_2}\left[\mathcal{D}_{g_1}[\varphi]\right](x) = \mathcal{D}_{g_1}[\varphi](g_2^{-1}x) = \varphi(g_1^{-1}g_2^{-1}x) = \varphi((g_2g_1)^{-1}x) = \mathcal{D}_{g_2g_1}[\varphi](x)\,. \tag{3.180}$$

proving we have a group representation over the vector field of scalar functions on X. By considering group elements that are close to the identity, we can now derive the representation of the Lie algebra. For that purpose, let us focus on infinitesimally small α and expand x'^μ in α, keeping only first-order terms

$$x'^\mu = f^\mu_{g(\alpha)}(x) = x^\mu - \alpha^j \epsilon^\mu_j(x) + \mathcal{O}(\alpha^2) \tag{3.181}$$

This infinitesimal coordinate transformation is reflected on the field as

$$\begin{aligned}\mathcal{D}_{g(\alpha)}[\varphi](x) &= \varphi(f_{g^{-1}}(x)) \\ &= \varphi(x^\mu + \alpha^j\epsilon^\mu_j + \mathcal{O}(\alpha^2)) \\ &= \varphi(x) + \alpha^j\epsilon^\mu_j\partial_\mu\varphi(x) + \mathcal{O}(\alpha^2) \\ &= (\mathbb{1} + i\alpha^j X_j)\varphi(x) + \mathcal{O}(\alpha^2)\end{aligned} \tag{3.182}$$

where we have made the identification

$$X_j = -i\epsilon^\mu_j(x)\partial_\mu \tag{3.183}$$

For infinitesimally small Lie parameters, i.e. close to the identity, we can thus formally write

$$D_g \simeq \mathbb{1} + i\alpha^j X_j \tag{3.184}$$

from which we identify $\alpha^j X_j$ as a representation of $\mathcal{A}_G$, the Lie algebra of G. The Lie algebra is thus here represented by differential operators acting on the linear vector space of fields. In a sense this is akin to a matrix representation, but of infinite dimensional ones. The X_j represent the basis and as such satisfy the commutation relations of $\mathcal{A}_G$

$$[X_i, X_j] = if^k_{ij}X_k \tag{3.185}$$

This is a non-trivial relation that descends from the fact that the $\epsilon^\mu_j(x)$, which define the X_j, precisely encode the infinitesimal action of G on the coordinates. Eq. (3.183) concretely gives

$$[X_i, X_j] = (-i)^2[\epsilon^\mu_i\partial_\mu, \epsilon^\nu_i\partial_\nu] = \left(\epsilon^\mu_j\partial_\mu\epsilon^\nu_i - \epsilon^\mu_i\partial_\mu\epsilon^\nu_j\right)\partial_\nu \tag{3.186}$$

so that eq. (3.185) implies

$$if_{ij}{}^k\epsilon^\nu_k\partial_\nu = \left(\epsilon^\mu_j\partial_\mu\epsilon^\nu_i - \epsilon^\mu_i\partial_\mu\epsilon^\nu_j\right)\partial_\nu \tag{3.187}$$

which is the infinitesimal reflection of the fact that the coordinate transformations $x^\mu \mapsto x'^\mu = f^\mu_g(x)$ form a group.

3.2.7 The Poincarè Lie Algebra

We can now apply the results of the previous section to the Poincaré group. Considering infinitesimal transformations

$$\begin{aligned} X'^{\mu} &= \Lambda^{\mu}{}_{\nu} x^{\nu} + a^{\mu} \\ &= (\delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu} + \mathcal{O}(\omega^2)) x^{\nu} + a^{\mu} \\ &= x^{\mu} + (\omega^{\mu}{}_{\nu} x^{\nu} + a^{\mu}) \\ &= x^{\mu} - \epsilon^{\mu} \end{aligned} \tag{3.188}$$

we identify $\epsilon^{\mu} \equiv -(\omega^{\mu}{}_{\nu} x^{\nu} + a^{\mu})$, while $\omega^{\mu\nu}$ and a^{ν} are identified with the Lie parameters. Considering a scalar field φ and following the lines of the previous section we have

$$\begin{aligned} \mathcal{D}_{(\Lambda,a)}[\varphi] &= \varphi(x^{\mu} - \omega^{\mu}{}_{\nu} x^{\nu} - a^{\mu} + \dots) \\ &= \varphi(x) - (\omega^{\mu}{}_{\nu} x^{\nu} + a^{\mu}) \partial_{\mu} \varphi + \dots \\ &\equiv \left(1 - i \frac{\omega^{\mu\nu} J_{\mu\nu}}{2} + i a^{\mu} P_{\mu}\right) \varphi + \dots \end{aligned} \tag{3.189}$$

with

$$J_{\mu\nu} = i(x_{\mu}\partial_{\nu} - x_{\nu}\partial_{\mu}) \tag{3.190}$$

$$P_{\mu} = i\partial_{\mu} \tag{3.191}$$

where for the Lorentz generators we chose a sign convention coinciding with our previous discussion (see eq. 3.138), while for translations our convention coincides with the familiar one in Quantum Mechanics. In particular $P^0 = i\partial^0 = i\partial_t$ corresponds to P^0 representing the energy.

Using the above generators we can now derive the full set of commutation relations of the Poincarè group

$$[J_{\mu\nu}, J_{\rho\sigma}] = i\left(\eta_{\nu\rho} J_{\mu\sigma} + \eta_{\mu\sigma} J_{\nu\sigma} - \eta_{\mu\rho} J_{\nu\sigma} - \eta_{\nu\sigma} J_{\mu\rho}\right) \tag{3.192}$$

$$[J_{\mu\nu}, P_{\rho}] = i\left(\eta_{\nu\rho} P_{\mu} - \eta_{\mu\rho} P_{\nu}\right) \tag{3.193}$$

$$[P_{\mu}, P_{\nu}] = 0 \tag{3.194}$$

Eq. (3.192) is not new and coincides with the commutation relations of the Lorentz generators derived in eq. (3.137) by considering the defining representation of the Lorentz group. Notice also that, by raising the μ and ν indices, eq. (3.193) can be written as

$$[J^{\mu\nu}, P_{\rho}] = P_{\sigma}\left\{i\left(\eta^{\mu\sigma}\delta^{\nu}{}_{\rho} - \eta^{\nu\sigma}\delta^{\mu}{}_{\rho}\right)\right\} = P_{\sigma}(\mathcal{J}^{\mu\nu})^{\sigma}{}_{\rho} \tag{3.195}$$

where (up to index relabelling) the expression in curly brackets precisely matches the expression for the generators in the defining representation given in eq. 3.133. We can formally express this result by saying that P_{ρ} transforms as a vector under Lorentz transformations. Indeed, by considering a finite Lorentz transformation and by iterating eq. (3.195) via the exponentiation trick (see also sect. (3.1.9)), we have ($\omega \cdot J \equiv \omega_{\mu\nu} J^{\mu\nu}$)

$$\begin{aligned} e^{-\frac{i}{2}\omega\cdot J} P_{\rho} e^{\frac{i}{2}\omega\cdot J} &= \lim_{N\to\infty} (\mathbb{1} - \frac{i}{2N}\omega\cdot J)^N P_{\rho} (\mathbb{1} + \frac{i}{2N}\omega\cdot J)^N = \lim_{N\to\infty} \left\{ P_{\sigma}\left[(\mathbb{1} - \frac{i}{2N}\omega\cdot\mathcal{J})^N\right]^{\sigma}{}_{\rho} + O(1/N)\right\} \\ &= P_{\sigma}\left(e^{-\frac{i}{2}\omega\cdot\mathcal{J}}\right)^{\sigma}{}_{\rho} = P_{\sigma}\Lambda^{\sigma}{}_{\rho}\,. \end{aligned} \tag{3.196}$$

Notice that the Lorentz transformation Λ "acts on the right" on the Lie algebra basis elements P_{ρ}, as already seen in sect. (3.1.9). On a vector $a \equiv a^{\mu} P_{\mu}$, eq. (3.196) is equivalent to the normal tranformation $a^{\mu} \to \Lambda^{\mu}{}_{\nu} a^{\nu}$ (that is the action is "on the left").

We can repeat a similar discussion for eq. (3.192). We can first rewrite it as

$$[J^{\mu\nu}, J_{\rho\sigma}] = J_{\lambda\sigma}(\mathcal{J}^{\mu\nu})^{\lambda}{}_{\rho} + J_{\rho\lambda}(\mathcal{J}^{\mu\nu})^{\lambda}{}_{\sigma} \tag{3.197}$$

and, secondly, by considering a finite transformation, we find

$$e^{-\frac{i}{2}\omega\cdot J} J_{\rho\sigma} e^{\frac{i}{2}\omega\cdot J} = J_{\mu\nu}\Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}\,. \tag{3.198}$$

These results express the fact that $J_{\rho\sigma}$ transforms as a two index tensor under Lorentz transformations.

A generic vector in the Poincarè Lie algebra

$$V = \omega^{\rho\sigma} J_{\rho\sigma} + a^{\rho} P_{\rho} \tag{3.199}$$

is parametrized by 10 coefficients $\omega^{\rho\sigma}$ (6) and a^{ρ} (4). As discussed in section 3.1.9, these coefficients constitue the adjoint representation of the Poincarè group. In particular, according to eqs. (3.196,3.198), under the Lorentz subgroup the transformation takes the form $a^{\rho} \to \Lambda^{\rho}{}_{\mu} a^{\mu}$ and $\omega^{\rho\sigma} \to \Lambda^{\rho}{}_{\mu}\Lambda^{\sigma}{}_{\nu}\omega^{\mu\nu}$. Notice that, as a^{ρ} and $\omega^{\rho\sigma}$ do not mix under Lorentz transformations, the adjoint representation of Poincaré is reducible under the Lorentz subgroup.

3.2.8 Representations of the Poincaré group on general fields

We have seen how to represent the Poincarè group on a scalar field. This construction can be generalized to a field associated with a general (j_-, j_+) representation of the Lorentz group. Such general fields will have not just one component like the scalar, but a number $(2j_- + 1)(2j_+ + 1)$ corresponding to the dimension of (j_-, j_+). Let us see how this works.

Consider the matrices representing Lorentz transformations in its (j_-, j_+) representation

$$D^{B}{}_{A}(\Lambda)\ ,\ \ A = 1, \dots (2j_- + 1)(2j_+ + 1) \tag{3.200}$$

and assume we have a set of fields $\phi^A(x)$ with $A = 1, \dots (2j_- + 1)(2j_+ + 1)$. We postulate the following transformation for the fields under Poincarè tranformations

$$\begin{cases} x'^{\mu} & = \Lambda^{\mu}{}_{\nu} x^{\nu} + a^{\mu} \equiv f^{\mu}_{(\Lambda,a)}(x) \\ \phi'^{A}(x') & \equiv D^{A}{}_{B}(\Lambda)\phi^{B}(x) \end{cases} \tag{3.201}$$

In order to verify that this is a valid definition we must consider how it behaves under composition. Using the same logic and a similar notation as in the previous section, we indicate the field of the O' observer as

$$\mathcal{D}_{(\Lambda,a)}[\phi]^{A}(x) \equiv D^{A}{}_{B}(\Lambda)\phi^{B}(f^{-1}_{\Lambda,a}(x')) \tag{3.202}$$

Focussing now on pure Lorentz transformations $a^{\mu} = 0$, we can check that the group composition law is respected

$$\mathcal{D}_{\Lambda_2}[\mathcal{D}_{\Lambda_1}[\phi]]^{A}(x) = D(\Lambda_2)^{A}{}_{B} D(\Lambda_1)^{B}{}_{C}\, \phi^{C}\left(\Lambda_1^{-1}\Lambda_2^{-1}x\right) = D(\Lambda_2\Lambda_1)^{A}{}_{C}\, \phi^{C}\left((\Lambda_2\Lambda_1)^{-1}x\right) = \mathcal{D}_{\Lambda_2\Lambda_1}[\phi]^{A}(x) \tag{3.203}$$

where we used that D is a group representation. The more general case where $a^{\mu} \neq 0$ is left as an exercise.

Let us consider a concrete field $\phi^A = A^{\mu}(x)$, like the vector potential in electrodynamics. Under the change of reference frame, A^{μ} transforms like x^{μ} (i.e. as a $(1/2, 1/2)$ field)

$$A'^{\mu}(x') = \Lambda^{\mu}{}_{\nu} A^{\nu}(x) \tag{3.204}$$

so that the functional form of this field changes as:

$$A'^{\mu}(x) = \Lambda^{\mu}{}_{\nu} A^{\nu}(\Lambda^{-1}x) \tag{3.205}$$

We can also consider the transformation 3.201 at the infinitesimal level. For that purpose we write

$$D^{A}{}_{B}(\Lambda) = \mathbb{1}^{A}{}_{B} - \frac{i}{2}\omega_{\mu\nu}(\Sigma^{\mu\nu})^{A}{}_{B} + \dots \tag{3.206}$$

where the $(\Sigma^{\mu\nu})^{A}{}_{B}$ represent $\mathcal{J}^{\mu\nu}$ in (j_-, j_+). For instance, we will later write these matrices explicitly in the case of spinors. At the infinitesimal level, Eq. 3.201 becomes:

$$\phi'^{A}(x) = \left(\mathbb{1}^{A}{}_{B} - \frac{i}{2}\omega_{\mu\nu}(\Sigma^{\mu\nu})^{A}{}_{B}\right)(1 - \frac{i}{2}\omega_{\mu\nu} i(x^{\mu}\partial^{\nu} - x^{\nu}\partial^{\mu}))\phi^{B}(x) \tag{3.207}$$

$$= (\mathbb{1}^{A}{}_{B} - \frac{i}{2}\omega_{\mu\nu}(\mathbb{1}^{A}{}_{B}(ix^{\mu}\partial^{\nu} - ix^{\nu}\partial^{\mu}) + (\Sigma^{\mu\nu})^{A}{}_{B}))\phi^{B}(x) \tag{3.208}$$

The first-order term $\mathbb{1}^{A}{}_{B}(ix^{\mu}\partial^{\nu} - ix^{\nu}\partial^{\mu}) + (\Sigma^{\mu\nu})^{A}{}_{B}$ is the sum of a differential operator and a matrix, both of which satisfy the Lorentz Lie algebra independently. Summing them gives another representation of this Lie algebra. We recognize the structure of the total angular momentum in Quantum Mechanics: the sum of an orbital part and a spin part.

3.3 Dynamical Consequences of Symmetry: Noether's theorem

3.3.1 Statement

Noether's theorem is a statement about the *dynamical* consequences of symmetry. But first, let us define properly what a symmetry is. Let us consider a change of coordinates corresponding to a Lie group with parameters $\{\alpha_i\}$:

$$\begin{cases} x'^{\mu} &= f^{\mu}(x, \alpha) \\ \phi'_a(x') &= F_a(\phi(x), \alpha) \end{cases} \tag{3.209}$$

In general we will have

$$S = \int_{\Omega} d^4x \mathcal{L}(\phi(x), \partial\phi(x)) = \int_{f(\Omega,\alpha)} d^4x' \mathcal{L}'(\phi'(x'), \partial'\phi'(x')) \tag{3.210}$$

with $\mathcal{L} \neq \mathcal{L}'$. In general, the form of the lagrangian, and consequently, the form of the equations of motion will change. The transformation 3.209 is said to be a *symmetry* if

$$\boxed{\mathcal{L}'(\phi', \partial\phi', x') = \mathcal{L}(\phi'(x'), \partial'\phi'(x')) + \partial'_{\mu}K^{\mu}(\phi')} \tag{3.211}$$

or, equivalenty, choosing an infinitesimal Ω in Eq.(3.210),

$$\boxed{d^4x\left[\mathcal{L}(\phi(x), \partial\phi(x))\right] = d^4x'\left[\mathcal{L}(\phi'(x'), \partial'\phi'(x')) + \partial'_{\mu}K^{\mu}(\phi')\right]} \tag{3.212}$$

When the above condition holds true, the form of the equations of motion of the two observers is the same, given their lagrangians differ only by a total derivative $\partial'_\mu K^\mu$, which does not affect the equations of motion. Equivalently, consider the action as computed by the two observers. By the above equations we have

$$S = \int_\Omega d^4x \mathcal{L}(\phi(x), \partial\phi(x)) = \int_{f(\Omega,\alpha)} d^4x' \mathcal{L}(\phi'(x'), \partial'\phi'(x')) + \int_{\partial f(\Omega,\alpha)} d\sigma'^\mu K^\mu(\phi') \,. \tag{3.213}$$

Indicating by $\Gamma(\phi,\partial)^a = 0$ the equations of motion, an infinitesimal shift $\delta\phi_a$ with $\delta\phi_a = 0$ on the boundary $\partial\Omega$ will satisfy

$$\delta S = \int_\Omega d^4x \, \Gamma(\phi,\partial)^a \delta\phi_a = \int_{f(\Omega,\alpha)} d^4x' \, \Gamma(\phi',\partial')^a \delta\phi'_a \; + \; 0 \tag{3.214}$$

where $\delta\phi'_a = (\partial F_a/\partial\phi_b)\delta\phi_b$, and where we doropped the variation of the boundary term as $\delta\phi'_a = 0$ on $\partial f(\Omega,\alpha)$. By the above equation if $\phi_a(x)$ is a solution, i.e. if $\delta S = 0$, then also $\phi'_a(x')$ is a solution, and of the same equations of motion. Symmetries then imply a degeneracy of solutions: if $\phi_a(x)$ is a solution, so will be $\phi'_a(x') = F_a(\phi(f^{-1}(x',\alpha)),\alpha)$ for all choices of the Lie parameters α. Eq. 3.211.

We can synthesize the above discussion as follows. A symmetry is associated to the existence of a class of equivalent observers describing physics in the same way. These observers are in one to one correspondence with the elements fo the symmetry group. Given a solution of the dynamics in the description of one observer, the descriptions of the same physical situation by the other equivalent observers correspondingly offers a class of solutions of the dynamics.

Examples. *1. Free scalar field:*

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 \tag{3.215}$$

Let us check that spacetime translations

$$\begin{cases} x'^\mu &= x^\mu - a^\mu \\ \phi'(x') &= \phi(x) \end{cases} \tag{3.216}$$

are indeed a symmetry. The volume element is not affected by translations:

$$d^4x = d^4x' \tag{3.217}$$

nor is the mass term of the Lagrangian:

$$\frac{1}{2}m^2\phi^2(x) = \frac{1}{2}m^2\phi'^2(x') \tag{3.218}$$

It just remains to check the derivative term:

$$\partial'_\mu\phi'(x') = \partial'_\mu\phi(x) = \frac{\partial x^\nu}{\partial x'^\mu}\partial_\nu\phi(x) = \delta^\nu{}_\mu\partial_\nu\phi(x) = \partial_\mu\phi(x) \tag{3.219}$$

Eq. (3.212) is therefore satisfied with $K^\mu = 0$, and spacetime translations are indeed a symmetry. By the previous discussion we also conclude that given a solution $\phi(x)$, for any a^μ, $\phi(x+a)$ is also a solution. The resulting multiplicity of solutions obviously corresponds to the existence of a multiplicity of equivalent choices of the origin of space-time coordinate space.

The occurrence of translation invariance in eq. (3.215), which seems cheap and automatic, actually originates from the fact that $\mathcal{L}$ does not explicitly depend on the coordinates. A coordinate dependent Lagrangian is for instance obtained by taking m^2 to be some spacetime dependent function $m^2 = g(x)$, rather than a constant. Such coordinate dependence would break the translation symmetry as

$$\mathcal{L}'(\phi,\partial\phi,g(x)) = \mathcal{L}(\phi'(x'),\partial'\phi'(x'),g(x'+a)) \neq \mathcal{L}(\phi'(x'),\partial'\phi'(x'),g(x')) \,. \tag{3.220}$$

2. *Complex scalar field:*

$$\mathcal{L} = \partial_\mu \phi \partial^\mu \phi^* - m^2 \phi \phi^* \tag{3.221}$$

We will check that this Lagrangian respects a global $U(1)$ symmetry:

$$\begin{cases} x'^\mu &= x^\mu \\ \phi'(x') &= e^{i\alpha}\phi(x) \end{cases} \tag{3.222}$$

The symmetry is said global, because the coordinates are not affected. We obviously have

$$d^4x' = d^4x \tag{3.223}$$
$$\phi'\phi'^* = \phi e^{i\alpha} e^{-i\alpha}\phi^* = \phi\phi^* \tag{3.224}$$
$$\partial_\mu \phi' \partial^\mu \phi'^* = \partial_\mu \phi e^{i\alpha} e^{-i\alpha} \partial^\mu \phi^* = \partial_\mu \phi \partial^\mu \phi^* \tag{3.225}$$

where in the third line we used the constancy of α. Eq.(3.212) is therefore satisfied. Again, in order to break this symmetry it would for instance suffice to add to the lagrangian the term $\phi^2 + (\phi^*)^2$.

We can know proceed at studying in more detail the dynamical consequences of symmetry. As we saw previously, basically all the structure of Lie groups (at least of the connected component of the identity) is encoded in the infinitesimal transformations: given all infinitesimal trasformations, or equivalently given the Lie algebra, all finite elements can be obtained by iterating several infinitesimal transformations. For the purpose of studying the dynamical consequences of symmetry it is then sufficient to focus on infinitesimal transformations. Choosing the usual convention where $\vec{\alpha} = 0$ corresponds to the identity

$$\begin{cases} x^\mu &= f^\mu(x, 0) \\ \phi_a(x) &= F_a(\phi(x), 0) \end{cases} \tag{3.226}$$

we expand the coordinate and field transformations to linear order in α neglecting higher orders

$$\begin{cases} x'^\mu &= x^\mu - \epsilon_i^\mu(x)\alpha^i &\equiv x^\mu - \epsilon^\mu(x) \\ \phi'_a(x') &= \phi_a(x) + \mathcal{E}_{ai}(\phi(x))\alpha^i &\equiv \phi_a(x) + \mathcal{E}_a(x) \end{cases} \tag{3.227}$$

We will also need to consider the infinitesimal change in the functional form of the field. Consider:

$$\phi'_a(x') = \phi'_a(x) - \epsilon^\mu(x)\partial_\mu \phi'_a(x) \tag{3.228}$$

Then:

$$\phi'_a(x) = \phi_a(x) + \left(\mathcal{E}_{ai}(\phi(x)) + \epsilon_i^\mu(x)\partial_\mu \phi_a(x)\right)\alpha^i \equiv \phi_a(x) + \Delta_{ai}(\phi(x))\alpha^i \equiv \phi_a(x) + \Delta_a(x) \tag{3.229}$$

In our translation symmetry example 3.216, we may explicitly identify:

$$\begin{cases} \epsilon^\mu &= a^\mu \\ \mathcal{E}_a &= 0 \\ \Delta_a &= a^\mu \partial_\mu \phi_a \end{cases} \tag{3.230}$$

In our second $U(1)$ symmetry example:

$$\begin{cases} x'^\mu = x^\mu & \Longrightarrow \ \epsilon^\mu = 0 \\ \phi' = e^{i\alpha}\phi \approx (1+i\alpha)\phi & \Longrightarrow \ \mathcal{E} = i\alpha\phi \\ \phi'^* = e^{-i\alpha}\phi^* \approx (1-i\alpha)\phi' & \Longrightarrow \ \mathcal{E}^* = -i\alpha\phi^* \\ \Delta = \mathcal{E} \end{cases} \tag{3.231}$$

Before stating Noether's theorem, let us make a final remark concerning the behaviour of $\partial'_\mu K^\mu$ as $\alpha \to 0$. Given that, for $\alpha = 0$, coordinates and fields do not transform, we have

$$\lim_{\alpha\to 0} \mathcal{L}(\phi'(x'), \partial'\phi'(x'))d^4x' = \mathcal{L}(\phi(x), \partial\phi(x))d^4x \tag{3.232}$$

Eq. (3.212), which holds for any α, then implies $\lim_{\alpha\to 0} \partial'_\mu K^\mu = 0$, or

$$K^\mu = \alpha^i K^\mu_i + \mathcal{O}(\alpha^2) \equiv \tilde{K}^\mu + \mathcal{O}(\alpha^2) \tag{3.233}$$

Noether's theorem. *Let us consider a Lie group G with coordinates $\{\alpha_i\}$, $i = 1, \dots, N$ which induces the transformation 3.209. If G is a symmetry of the system, then there is one conservation law for each $i = 1, \dots, N$ taking the form*

$$\boxed{\partial_\mu J^\mu_i = 0} \tag{3.234}$$

with

$$\boxed{J^\mu_i \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\Delta_{ai} - \epsilon^\mu_i \mathcal{L} + K^\mu_i} \tag{3.235}$$

Proof. *The fact that G is a symmetry is encapsulated in eq. (3.212) The strategy now is to write the right-hand-side of eq. (3.212) in terms of the x coordinates and ϕ_a fields and to expand it in powers of α. As the left hand side is α independent all the terms involving non-zero powers of α on the right hand side will have to vanish identically. This will give an infinite set of identities, the first of which, corresponding to terms linear in α, will coincide with Noether's theorem. The vanishing of the higher order terms is dictated by the group structure and does not add additional information.*

For an infinitesimal transformation (we will work from here on at linear order in the parameters α^i and thus we omit in all equation a $\mathcal{O}(\alpha^2)$), the Jacobian will be given by:

$$\left|\frac{\partial x'^\mu}{\partial x^\nu}\right| = \det\left(\delta^\mu_\nu - \partial_\nu \epsilon^\mu\right) = 1 - \mathrm{Tr}\left(\partial_\nu \epsilon^\mu\right) = 1 - \partial\epsilon \tag{3.236}$$

The Lagrangian term becomes:

$$\mathcal{L}(\phi'(x'), \partial'\phi'(x')) = \mathcal{L}(\phi'(x), \partial\phi'(x)) - \epsilon^\mu \partial_\mu \mathcal{L}(\phi', \partial\phi') \tag{3.237}$$

$$= \mathcal{L}(\phi'(x), \partial\phi'(x)) - \epsilon^\mu \partial_\mu \mathcal{L}(\phi, \partial\phi) \tag{3.238}$$

where when going from the first to the second line we have neglected terms of order α^2 and higher (notice that $\epsilon^\mu = O(\alpha)$). Introducing the field variation Δ_a as defined in Eq. 3.229, we get:

$$\mathcal{L}(\phi'(x'), \partial'\phi'(x')) = \mathcal{L}(\phi(x) + \Delta(x), \partial(\phi(x) + \Delta(x)) - \epsilon^\mu \partial_\mu \mathcal{L}(\phi, \partial\phi) \tag{3.239}$$

$$= \mathcal{L}(\phi(x), \partial\phi(x)) + \Delta_a \frac{\partial \mathcal{L}}{\partial \phi_a} + \partial_\mu \Delta_a \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} - \mathcal{L}(\phi'(x'), \partial'\phi'(x')) \tag{3.240}$$

Now, let us go back to our initial equation Eq. 3.212 and summarize.

$$\mathcal{L}(\phi,\partial\phi)d^4x = \left[\mathcal{L}(\phi,\partial\phi) + \Delta_a\frac{\partial\mathcal{L}}{\partial\phi_a} + \partial_\mu\Delta_a\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} - \epsilon^\mu\partial_\mu\mathcal{L} + \partial_\mu K^\mu\right](1-\partial_\nu\epsilon^\nu)d^4x \tag{3.241}$$

Simplifying and keeping only terms up to linear order in α we obtain

$$0 = -\partial_\mu(\epsilon^\mu\mathcal{L}) + \Delta_a\frac{\partial\mathcal{L}}{\partial\phi_a} + \partial_\mu\Delta_a\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} + \partial_\mu\tilde{K}^\mu \tag{3.242}$$

The first and last term have the form of the divergence of vectors, but the second and third don't. However by simple manipulations this can be rewitten as

$$0 = -\partial_\mu(\epsilon^\mu\mathcal{L}) + \partial_\mu\left(\Delta_a\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right) - \Delta_a\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} + \Delta_a\frac{\partial\mathcal{L}}{\partial\phi_a} + \partial_\mu\tilde{K}^\mu \tag{3.243}$$

$$= \partial_\mu\left(-\epsilon^\mu\mathcal{L} + \Delta_a\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} + \tilde{K}^\mu\right) + \Delta_a\left(\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right) \tag{3.244}$$

which, besides the divergence of a current, features a term proportional to the equations of motion. Assuming, as we will do in what follows, that the fields satisfy the equations of motion, the above equation establishes a certain current must be conserved. Making the dependence on α^i explicit this reads

$$0 = \alpha_i\partial_\mu\left[-\epsilon_i^\mu\mathcal{L} + \Delta_{ai}\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} + K_\mu^i\right] \tag{3.245}$$

and using that this identity holds true for arbitrary α^i we have N identities

$$\forall i \;:\; 0 = \partial_\mu\left[-\epsilon_i^\mu\mathcal{L} + \Delta_{ai}\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} + K_\mu^i\right] = \partial_\mu J_i^\mu \qquad i = 1,\dots,N \tag{3.246}$$

which coincide with the thesis of Noether's theorem.

Which means that, when the Lie group G is a symmetry of the system and the field configuration considered satisfies the equations of motion, for each Lie parameter α^i, that is for each element of the Lie algebra, there exists an associated corserved current J_i^μ. □

Let us make some remarks about this theorem. First, this result implies the conservation of the *total* charge associated to each current J_i^μ:

$$Q_i \equiv \int d^3\mathbf{x}J_i^0 \tag{3.247}$$

Consider the spacetime region $\mathbb{R}^3\times\Delta T$, where $\Delta T = [0,t]$. Then, from Noether's theorem:

$$0 = \int_0^t dt\int d^3\mathbf{x}\,\partial_\mu J_i^\mu = \int_0^t dt\int d^3\mathbf{x}\,(\partial_t J_i^0 + \vec{\nabla}\cdot\vec{J}_i) = \int d^3\mathbf{x}\,(J_i^0(t,\mathbf{x}) - J_i^0(0,\mathbf{x})) + \int_0^t dt\int d\vec{\Sigma}\cdot\vec{J}_i \tag{3.248}$$

If the fields vanish fast enough at infinity, then the space boundary integral vanishes, so that:

$$Q_i(t) = \int d^3\mathbf{x}\, J_i^0(t,\mathbf{x}) = \int d^3\mathbf{x}\, J_i^0(0,\mathbf{x}) = Q_i(0), \quad \forall\, t \tag{3.249}$$

In other words, Noether's theorem implies that the total charge is time independent

$$\boxed{\frac{d}{dt}Q_i(t) = 0} \tag{3.250}$$

Finally, Noether's theorem implies a *local* conservations law as a result of a *global* symmetry. The locality of the conservation law suggests that a local derivation should exist. Indeed, consider a *local* transformation where the Lie coordinates $\alpha^i \equiv \alpha^i(x)$ now depend on spacetime. Define:

$$\delta\phi_a(x) \equiv \phi'_a(x) - \phi_a(x) = \Delta_{ai}\alpha^i(x) \tag{3.251}$$

$$\delta\partial_\mu\phi_a(x) \equiv \partial_\mu(\Delta_{ai}\alpha^i(x)) = (\partial_\mu\Delta_{ai})\alpha^i(x) + \Delta_{ai}\partial_\mu\alpha^i(x) \tag{3.252}$$

Now:

$$\delta\mathcal{L} = \mathcal{L}(\phi',\partial\phi') - \mathcal{L}(\phi,\partial\phi) = \frac{\partial\mathcal{L}}{\partial\phi_a}\Delta_{ai}\alpha^i + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}((\partial_\mu\Delta_{ai})\alpha^i(x) + \Delta_{ai}\partial_\mu\alpha^i(x)) \tag{3.253}$$

Using first the invariance under the symmetry as expressed by eq. (3.242) and then the equations of motion, this becomes

$$\begin{aligned}\delta\mathcal{L} &= \partial_\mu(\epsilon_i^\mu\mathcal{L} - K_i^\mu)\alpha^i + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\Delta_{ai}\partial_\mu\alpha^i(x) \\ &= J_i^\mu\partial_\mu\alpha^i + \partial_\mu(\epsilon_i^\mu\mathcal{L}\alpha^i - K_i^\mu\alpha^i)\end{aligned} \tag{3.254}$$

So that for a local transformation we have:

$$\Delta S \equiv \int d^4x\,[\mathcal{L}(\phi',\partial\phi') - \mathcal{L}(\phi,\partial\phi)] = \int_\Omega J_i^\mu\partial_\mu\alpha^i d^4x \;+\; \text{boundary term} \tag{3.255}$$

which shows that the current associated with a global symmetry can be derived by *formally* promoting the symmetry to a local one.

We will now give an important example and application of Noether's theorem.

3.3.2 The Noether currents of Poincarè invariance

The Poincarè group consists of the combination of spacetime translations and Lorentz transformations. Let us consider translations first:

$$\left\{\begin{array}{lcll} x'^\mu & = & x^\mu - a^\mu & \equiv\; x^\mu - \epsilon_i^\mu(x)\alpha^i \\ \phi'_a(x') & = & \phi_a(x) & \equiv\; \phi_a(x) + \mathcal{E}_{ai}(\phi(x))\alpha^i \end{array}\right. \tag{3.256}$$

As we have already seen, invariance under translation is guaranteed by the absence of an explicit dependence of the Lagrangian on the coordinates:

$$\mathcal{L}(\phi(x), \partial\phi(x)) = \mathcal{L}(\phi'(x'), \partial'\phi'(x')) \tag{3.257}$$

Applying the definition in eqs. (3.227,3.229) and identifying a^ν with the Lie parameters α^i

$$\epsilon^\mu = a^\mu = a^\nu \delta^\mu_\nu \quad \Rightarrow \quad \epsilon^\mu_\nu \equiv \delta^\mu_\nu, \tag{3.258}$$
$$\mathcal{E}_a = 0 \tag{3.259}$$
$$\Delta_a = a^\nu \partial_\nu \phi_a \quad \Rightarrow \quad \Delta_{a\nu} \equiv \partial_\nu \phi_a \tag{3.260}$$
$$\tag{3.261}$$

We see that the Lie coordinates are the components of a spacetime four-vector. Therefore the currents J^μ_i actually form in this case a rank 2 tensor called the *stress-energy tensor*:

$$T^\mu{}_\nu \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \partial_\nu \phi_a - \delta^\mu_\nu \mathcal{L} \tag{3.262}$$

Following the well known result in mechanics the four corresponding conserved charges correspond to the total energy and momentum of the system, which can be fit into the four-momentum vector

$$P_\mu \equiv \int T^0{}_\mu d^3\mathbf{x} \tag{3.263}$$

Spacetime translations invariance of the action implies thus the conservation of energy and momentum:

$$\frac{d}{dt} P_\mu = 0 \tag{3.264}$$

Consider now Lorentz invariance. A finite Lorentz transformation reads:

$$\left\{ \begin{array}{lcl} x'^\mu & = & \Lambda^\mu{}_\nu x^\nu \\ \phi'_a(x') & = & D(\Lambda)_a{}^b \phi_b(x) \end{array} \right. \tag{3.265}$$

where

$$D(\Lambda)_a{}^b \equiv \left(\exp \ -\frac{i}{2} \omega_{\mu\nu} \Sigma^{\mu\nu} \right)^b_a \tag{3.266}$$

and where the $\Sigma^{\mu\nu}$ provide a representation of the Lorentz algebra. At the infinitesimal level these become

$$\left\{ \begin{array}{lcl} x'^\mu & = & x^\mu + \omega^\mu{}_\nu x^\nu \\ \phi'_a(x') & = & \phi_a(x) - \frac{i}{2} (\omega_{\mu\nu} \Sigma^{\mu\nu})_a{}^b \phi_b(x) \end{array} \right. \tag{3.267}$$

and thus

$$\epsilon^\mu = -\omega^{\mu\nu} x_\nu \tag{3.268}$$

and

$$\begin{aligned}\Delta_a &= -\frac{i}{2}(\omega_{\mu\nu}\Sigma^{\mu\nu})_a{}^b \phi_b(x) - \omega^{\mu\nu} x_\nu \partial_\mu \phi_a(x) \\ &= -\frac{i}{2}\omega^{\mu\nu}\left(\Sigma_{\mu\nu} + i(x_\mu\partial_\nu - x_\nu\partial_\mu)\right)_a{}^b \phi_b(x) \\ &\equiv -\frac{i}{2}\omega^{\mu\nu}\left(\Sigma_{\mu\nu} + \mathrm{J}_{\mu\nu}\right)_a{}^b \phi_b(x)\,.\end{aligned} \tag{3.269}$$

This last expression shows that the infinitesimal Lorentz transformation of a field decomposes into the sum of a spin part and of an orbital angular momentum part.

Now, for the sake of simplicity, let us consider the case of a single scalar field $\phi(x)$. Using the above results for the infinitesimal coordinates and field variations in eq. (3.235) with $K^\mu = 0$[12] we then find for any $\omega^{\mu\nu}$ the conserved current

$$J^\rho = \frac{\partial \mathcal{L}}{\partial(\partial_\rho \phi)}(-\omega^{\mu\nu} x_\nu \partial_\mu \phi) + \omega^{\rho\nu} x_\nu \mathcal{L}\,. \tag{3.270}$$

Factoring out the antisymmetric $\omega^{\mu\nu}$ we can write

$$J^\rho = \frac{1}{2}\omega^{\mu\nu}\left\{\left(\frac{\partial \mathcal{L}}{\partial(\partial_\rho\phi)} x_\nu \partial_\mu \phi - \delta^\rho_\mu x_\nu \mathcal{L}\right) \;-\; (\mu \leftrightarrow \nu)\right\} = \frac{1}{2}\omega^{\mu\nu}\left(x_\mu T^\rho{}_\nu - x_\nu T^\rho{}_\mu\right) \equiv \frac{1}{2}\omega^{\mu\nu} J^\rho_{\mu\nu} \tag{3.271}$$

Hence, we notice that we can write the currents associate to Lorentz invariance in terms of the stress-energy tensor and of the coordinates. Their conservation

$$\partial_\rho J^\rho_{\mu\nu} = 0\,, \tag{3.272}$$

by using $\partial_\rho T^\rho{}_\mu = 0$, is easily seen to imply

$$T_{\mu\nu} \equiv \eta_{\mu\rho} T^\rho{}_\nu = \eta_{\nu\rho} T^\rho{}_\mu \equiv T_{\nu\mu} \tag{3.273}$$

Therefore, if a field theory is both translationally and Lorentz invariant, i.e. if it is Poincarè invariant, its stress-energy tensor is symmetric. Our discussion was based on the case of a scalar field. In the case of a field with spin things are more complicated, and Lorentz invariance does not immediately imply the symmetry of $T^{\mu\nu}$. However one can prove that even in the more general case of a spinning field one can *improve* the energy momentum tensor so as to make it symmetric. This amounts to the following. One defines a new energy momentum tensor

$$\Theta^{\mu\nu} \equiv T^{\mu\nu} + \partial_\sigma A^{\sigma\mu\nu} \tag{3.274}$$

[12] In all interesting cases Lorentz invariance is realized *exactly*, that is with $K^\mu = 0$ in eq. (3.211)

with $A^{\rho\mu\nu}$ a suitable tensor expression in fields and derivatives, antisymmetric in the indices σ and μ. The resulting $\Theta^{\mu\nu}$ is conserved when $T^{\mu\nu}$ is conserved given $\partial_\mu\partial_\sigma A^{\sigma\mu\nu} = 0$ identically, because of $A^{\sigma\mu\nu} = -A^{\mu\sigma\nu}$. Moreover the conserved charges P_μ are the same for $\Theta^{\mu\nu}$ and $T^{\mu\nu}$, because the corresponding integrands only differ by a total space derivative, which gives a vanishing contributions when fields vanish sufficiently fast at infinity. $\Theta^{\mu\nu}$ is thus physically equivalent to $T^{\mu\nu}$. The non-trivial result is that one can prove that in the case of a Lorentz invariant theory one can always choose $A^{\sigma\mu\nu}$ so as to make $\Theta^{\mu\nu}$ symmetric. The resulting symmetric energy momentum tensor is also known as the Belinfante tensor.

It is useful to consider the constraint placed by Lorentz invariance on $T^{\mu\nu}$, purely in terms of quantum numbers. As each index of $T^{\mu\nu}$ transforms as a four-vector, that is as $(1/2, 1/2)$, the tensor transforms as:

$$(1/2,1/2)\otimes(1/2,1/2) = (1/2\otimes 1/2, 1/2\otimes 1/2) \tag{3.275}$$
$$= (0\oplus 1, 0\oplus 1) \tag{3.276}$$
$$= (0,0)\oplus(0,1)\oplus(1,0)\oplus(1,1) \tag{3.277}$$

where the last line offers the decomposition of a generic two index tensor into irreducible representations of $SO(3,1)$: $(0,0)$ is a singlet, corresponding to the trace, while $(0,1)\oplus(1,0)$ and $(1,1)$ are respectively the antisymmetric and the symmetric traceless components. In a theory without other spacetime symmetries other than translations, these components do not satisfy any other constraint beside current conservation $\partial_\mu T^{\mu\nu} = 0$. However, Lorentz invariance implies that we have 6 additional conserved currents associated to its 6 generators. These generators sit in the $(0,1)\oplus(1,0)$ representation, and constrain the corresponding component in $T^{\mu\nu}$ to vanish. This results in a symmetric tensor decomposing as

$$T^\mu{}_\nu \sim (0,0)\oplus(1,1) \tag{3.278}$$

Having seen how things work for Lorentz transformations we can immediately grasp what will happen by adding a further spacetime symmetry, dilatations: $x'^\mu = \lambda x^\mu$. Now the generator is a Lorentz scalar and the natural guess is that the corresponding conservation law will constrain the $(0,0)$ component of $T^{\mu\nu}$ to vanish, that is $T^\mu{}_\mu = 0$. This guess is correct, as one can check in the simplest examples.

Let us see what the conservation of the Lorentz currents implies physically. The conserved charges are:

$$J_{\mu\nu} \equiv \int d^3\mathbf{x}\, J^0_{\mu\nu} = \int d^3\mathbf{x}\,(x_\mu T^0{}_\nu - x_\nu T^0{}_\mu) \tag{3.279}$$

Consider first purely spacelike indices, $\mu\nu = ij$, which are associated to rotations. We have

$$J_{ij} = \int d^3\mathbf{x}\,(x_i T^0{}_j - x_j T^0{}_i) \equiv \int d^3\mathbf{x}\,(x_i p_j - x_j p_i) \tag{3.280}$$

where p_i is the density of three-momentum. Define now:

$$J_i = \frac{1}{2}\epsilon_{ijk}J_{jk} = \int d^3\mathbf{x}(\vec{x}\wedge\vec{p})_i \tag{3.281}$$

Therefore, the spatial components of the conserved charges correspond to the total angular momentum of the system! In other words, invariance under space rotations implies the conservation of angular momentum.

Consider then the mixed time-space indices $\mu\nu = i0$, which are associated to boosts. The conserved charges are

$$K_i \equiv J_{i0} = \int d^3\mathbf{x}(x_i\rho - x_0 p_i) \tag{3.282}$$

where $\rho = T^0{}_0$ is the energy density. Defining the center of mass coordinates

$$X_i^{CM} \equiv \frac{\int d^3\mathbf{x} x_i \rho}{\int d^3\mathbf{x}\rho} \tag{3.283}$$

we can then write

$$K_i \equiv P_0 X_i^{CM} - tP_i \tag{3.284}$$

Notice that the K_i depend explictly on time. So, what does its conservation imply?

$$0 = \frac{d}{dt}K_i = P_0\dot{X}_i^{CM} - P_i \;\Rightarrow \dot{X}_i^{CM} = \frac{P_i}{P_0} = \text{ const.} \tag{3.285}$$

In other words, the center of mass moves at a constant velocity determined by the ratio of the total 3-momentum and the total energy. This is similar to what implied by Galilean invariance. In that case, however, the role of P_0 is played by the total mass.

Chapter 4

Scalar fields

4.1 The Klein-Gordon field

The goal of this chapter is to consider the simplest field theory and study its quantization. The simplest field content clearly corresponds to a single scalar field $\varphi(x)$. The simplest meaningful dynamics, like in any other mechanical system, is described by the most general quadratic Lagrangian, i.e. involving up to quadratic terms in $\varphi(x)$. Let us recall this general remark by considering the mechanics of a single variable q. The most general $L(q,\dot{q})$ can be written as an expansion in powers of $\dot{q}$

$$\mathcal{L} = F_0(q) + F_1(q)\dot{q} + F_2(q)\dot{q}^2 + F_3(q)\dot{q}^3 + \dots \tag{4.1}$$

Notice that the second term can be written as a total time derivative $F_1(q)\dot{q} = dI(q)/dt$ with $I(q) = \int F_1(q)$. Thus this term does not affect the equations of motion and can be dropped. Now, making the generic assumption that $-F_0(q) \equiv V(q)$ possesses a minimum at $q = q_0$ and that $F_2(q_0) \equiv m/2 > 0$, we can meaningfully expand around this point according to $q(t) = q_0 + \delta(t)$ and write, up to an irrelevant constant, the Lagrangian as

$$L = -\frac{1}{2}m\omega^2\delta^2 + \frac{m}{2}\dot{\delta}^2 + O(\delta^3, \delta\dot{\delta}^2, \dot{\delta}^3)\,. \tag{4.2}$$

For small enough δ and $\dot{\delta}$ the dynamic is then well approximated by a simple harmonic oscillator, while the terms of higher order in δ and $\dot{\delta}$ can be treated as small perturbations.

Applying the same logic to a Poincarè invariant theory of a scalar field φ, in particular using integration by parts, we can write

$$\mathcal{L}(\varphi, \partial\varphi) = -V(\varphi) + \frac{1}{2}F_2(\varphi)\partial_\mu\varphi\partial^\mu\varphi + \dots\,, \tag{4.3}$$

where the dots represent terms involving more than two powers of $\partial\varphi$. Expanding around the minimum of $V(\varphi)$ at $\varphi = \varphi_0$ according to $\varphi(x) = \varphi_0 + \delta\varphi(x)$, at quadratic order in $\delta\varphi$, we can write

$$\mathcal{L} = -\frac{1}{2}V''(\varphi_0)\delta\varphi^2 + \frac{1}{2}F(\varphi_0)(\partial_\mu\delta\varphi)(\partial^\mu\delta\varphi) + \dots\,. \tag{4.4}$$

Assuming $F(\varphi_0) > 0$ and further redefining $\delta\varphi(x) \equiv \phi(x)/\sqrt{F(\varphi_0)}$ we can write

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 + \dots\,. \tag{4.5}$$

where $m^2 = V''(\varphi_0)/F_2(\varphi_0) \geq 0$. The resulting quadratic terms define the Klein-Gordon Lagrangian, which is the field-theoretic analog of the harmonic oscillator. The Klein-Gordon Lagrangian offers a meaningful approximation of the dynamics of a scalar field ϕ in those situations where the terms of cubic and higher order can be treated as small perturbations. As it turns out, the latter situation occurs quite frequently in nature. An example is given by the scalar field describing the Higgs boson. It, therefore, makes sense to study, as a first approximation, the quantum theory associated with the Klein-Gordon Lagrangian.

Alternatively one could arrive at the Klein-Gordon Lagrangian by considering the most general Lorentz invariant Lagrangian for $\phi(x)$ involving at most two derivatives and at most two powers of $\phi(x)$

$$\mathcal{L} = \frac{c_1}{2}\partial_\mu\phi\partial^\mu\phi + \frac{c_2}{2}\phi^2 + c_3\phi + c_4\partial_\mu\partial^\mu\phi + c_5\phi\partial_\mu\partial^\mu\phi\,. \tag{4.6}$$

Using integration by parts, dropping total derivative terms, and shifting the origin of field space, the above Lagrangian is easily seen to be equivalent to the quadratic term in eq. (4.5)

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m^2}{2}\phi^2\,. \tag{4.7}$$

Let us study the basic properties of this Lagrangian. Applying dimensional analysis, we have

$$\left[\int d^4x\mathcal{L}\right] = E^0\,, \tag{4.8}$$

and thus

$$\begin{array}{rcl}\left[d^4x\right] &=& E^{-4}\,,\\ \left[\mathcal{L}\right] &=& E^4\,,\\ \left[\phi\right] &=& E\,,\\ \left[m^2\right] &=& E^2\,.\end{array} \tag{4.9}$$

The equations of motion are

$$\partial_\mu\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi\right)} - \frac{\partial\mathcal{L}}{\partial\phi} = 0 \quad\rightarrow\quad \left(\Box + m^2\right)\phi = 0. \tag{4.10}$$

where $\Box \equiv \partial_\mu\partial^\mu$. This is the Klein-Gordon equation.

Defining the conjugate momentum $\pi \equiv \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \dot{\phi}$ we find the Hamiltonian density

$$\mathcal{H} = \pi\dot{\phi} - \mathcal{L} = \frac{1}{2}\pi^2 + \frac{1}{2}\left(\nabla\phi\right)^2 + m^2\phi^2\,. \tag{4.11}$$

The Hamiltonian density is positive definite for $m^2 \geq 0$. This result is a consequence of our original hypotheses, $V''(\varphi_0) \geq 0$, $F_2(\varphi_0) > 0$.

It is important to appreciate that, with the identification $\dot{\phi} = \pi$, $\mathcal{H}$ coincides with the T^0_0 component of the energy momentum tensor derived by applying eq. (3.262) to the scalar field Lagrangian of eq. (4.7)

$$T^\mu{}_\nu = \partial^\mu\phi\partial_\nu\phi - \frac{1}{2}\delta^\mu_\nu\left[\left(\partial\phi\right)^2 - m^2\phi^2\right]\,. \tag{4.12}$$

The component T^0_i instead gived the 3-momentum density, for that for the total 3-momentum the above equation gives

$$P_i = \int d^3\mathbf{x}\,T^0{}_i = \int d^3\mathbf{x}\,\dot{\phi}\partial_i\phi = \int d^3\mathbf{x}\,\pi\partial_i\phi\,. \tag{4.13}$$

4.2 Quantized Klein-Gordon field

To construct the quantized Klein-Gordon field theory we proceed via the canonical formalism, like in ordinary quantum mechanics. As outlined in eq. (2.1) this amounts to

At equal times this reads

· Replace	$\phi(\mathbf{x},t), \pi(\mathbf{x},t)$	with hermitian operators acting on a Hilbert space	$\hat{\phi}, \hat{\pi}$
· Replace	Poisson brackets {-,-}	with commutators	$i[-,-]$

Classical	Quantum
$\{\pi(\mathbf{x},t), \phi(\mathbf{y},t)\} = \delta^3(\mathbf{x}-\mathbf{y})$	$i\left[\hat{\pi}(\mathbf{x},t), \hat{\phi}(\mathbf{y},t)\right] = \delta^3(\mathbf{x}-\mathbf{y})$
$\{\pi(\mathbf{x},t), \pi(\mathbf{y},t)\} = \{\phi(\mathbf{x},t), \phi(\mathbf{y},t)\} = 0$	$[\hat{\pi}(\mathbf{x},t), \hat{\pi}(\mathbf{y},t)] = \left[\hat{\phi}(\mathbf{x},t), \hat{\phi}(\mathbf{y},t)\right] = 0$

(4.14)

We now have to find the Hilbert space that realises this algebra. Just like for the harmonic oscillator in quantum mechanics, a suitable basis of the Hilbert space is given by the eigenvectors of the Hamiltonian [1]

$$H = \int \mathcal{H} d^3\mathbf{x} = \int d^3\mathbf{x} \left[\frac{1}{2}\pi^2 + \frac{1}{2}\left(\vec{\nabla}\phi\right)^2 + \frac{1}{2}m^2\phi^2\right] . \tag{4.15}$$

Since $\frac{1}{2}\left(\vec{\nabla}\ \phi\right)^2$ couples the field at neighboring points, the Hamiltonian is not diagonal in the position space variables. However, the Hamiltonian is invariant under spatial translations. As such, it will be diagonal when expressed in terms of eigenstates of translations. These are nothing else but the spatial Fourier modes of $\pi(\mathbf{x},t)$ and $\phi(\mathbf{x},t)$.

To study the system in Fourier space, let us first focus on the case where space is compactified on a 3-torus of length L: $0 \le \mathrm{x}^1 < L,\ 0 \le \mathrm{x}^2 < L,\ 0 \le \mathrm{x}^3 < L$. This can be technically achieved by imposing periodic boundary conditions $\phi(\mathrm{x}^1 + L, \mathrm{x}^2, \mathrm{x}^3) = \phi(\mathrm{x}^1, \mathrm{x}^2, \mathrm{x}^3)$. For finite values of L we will have to deal with discrete sums rather than integrals making the manipulations more familiar. The complete set of orthonormalized periodic Fourier modes is given by

$$\psi_{\mathbf{n}}(\mathbf{x}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k_n}\cdot\mathbf{x}}, \tag{4.16}$$

with $V = L^3$ and $\mathbf{k_n} = \frac{2\pi}{L}\mathbf{n}$, $\mathbf{n} \in \mathbb{Z}^3$. They are clearly eigenstates of translations as

$$\psi_{\mathbf{n}}(\mathbf{x}+\mathbf{a}) = e^{i\mathbf{k_n}\cdot\mathbf{a}}\psi_{\mathbf{n}}(\mathbf{x}) . \tag{4.17}$$

Orthonormality and completeness respectively correspond to the relations

$$(\psi_{\mathbf{n}}, \psi_{\mathbf{m}}) \equiv \int \psi_{\mathbf{n}}(\mathbf{x})^* \psi_{\mathbf{m}}(\mathbf{x}) d^3\mathbf{x} = \frac{1}{V}\int e^{i(\mathbf{k_m}-\mathbf{k_n})\cdot\mathbf{x}} d^3\mathbf{x} = \delta_{\mathbf{n},\mathbf{m}}, \tag{4.18}$$

$$\sum_{\mathbf{n}} \psi_{\mathbf{n}}(\mathbf{x})^* \psi_{\mathbf{n}}(\mathbf{y}) = \delta^3(\mathbf{x}-\mathbf{y}) . \tag{4.19}$$

The Fourier modes of the fields are then

$$\phi_{\mathbf{n}}(t) = \int \psi^*_{\mathbf{n}}(\mathbf{x})\phi(\mathbf{x},t) d^3\mathbf{x}, \qquad \pi_{\mathbf{n}}(t) = \int \psi^*_{\mathbf{n}}(\mathbf{x})\phi(\mathbf{x},t) d^3\mathbf{x}, \tag{4.20}$$

and inverting these relations using orthonormality and completeness we arrive at the mode expansion

$$\phi(\mathbf{x},t) = \sum_{\mathbf{n}} \psi_{\mathbf{n}}(\mathbf{x})\phi_{\mathbf{n}}(t), \qquad \pi(\mathbf{x},t) = \sum_{\mathbf{n}} \psi_{\mathbf{n}}(\mathbf{x})\pi_{\mathbf{n}}(t). \tag{4.21}$$

Notice that by the hermiticity of the field $\phi(\mathbf{x},t) = \phi(\mathbf{x},t)^\dagger$ we have

$$\phi^\dagger_{\mathbf{n}}(t) = \int d^3\mathbf{x}\ (\psi^*_{\mathbf{n}}(\mathbf{x}))^*\ \phi(\mathbf{x},t)^\dagger = \int d^3\mathbf{x}\, \psi_{\mathbf{n}}(\mathbf{x})\phi(\mathbf{x},t) = \int d^3\mathbf{x}\, \psi_{-\mathbf{n}}(\mathbf{x})^*\phi(\mathbf{x},t) = \phi_{-\mathbf{n}}(t), \tag{4.22}$$

and similarly, $\pi^\dagger_{\mathbf{n}} = \pi_{-\mathbf{n}}$ comes from $\pi(\mathbf{x}) = \pi^\dagger(\mathbf{x})$.

From eq. (4.20) we can derive the equal time commutation relations of the Fourier modes $\phi_{\mathbf{n}}$ and $\pi_{\mathbf{n}}$. Without surprise, we get

$$[\phi_{\mathbf{n}}(t), \phi_{\mathbf{m}}(t)] = [\pi_{\mathbf{n}}(t), \pi_{\mathbf{m}}(t)] = 0, \tag{4.23}$$

[1] As from now on we will always work with quantum fields, we will drop the hat on the field operators.

while for the less trivial commutators, we have

$$[\phi_{\mathbf{n}}(t), \pi_{\mathbf{m}}(t)] = \int d^3\mathbf{x}\, d^3\mathbf{y}\, \psi^*_{\mathbf{n}}(\mathbf{x})\psi^*_{\mathbf{m}}(\mathbf{y})\, [\phi(\mathbf{x},t), \pi(\mathbf{y},t)] = i \int d^3\mathbf{x}\, \psi^*_{\mathbf{n}}(\mathbf{x})\psi^*_{\mathbf{m}}(\mathbf{x})$$
$$= \frac{i}{V}\int d^3\mathbf{x} \exp\left\{-\frac{i2\pi\,(\mathbf{n}+\mathbf{m})\cdot\mathbf{x}}{L}\right\} = i\delta_{\mathbf{n}+\mathbf{m},0} \equiv i\delta_{\mathbf{n},-\mathbf{m}}, \tag{4.24}$$

which, by the hermiticity condition $\pi^\dagger_{\mathbf{n}} = \pi_{-\mathbf{n}}$, can be written as

$$\left[\phi_{\mathbf{n}}(t), \pi^\dagger_{\mathbf{m}}(t)\right] = i\delta_{\mathbf{n},\mathbf{m}} \tag{4.25}$$

Using eq. (4.21), we can write the Hamiltonian in terms of the Fourier modes of the fields. For the various terms, we have

$$\begin{aligned}
\textstyle\int_V \pi^2(\mathbf{x}) &= \sum_{\mathbf{n},\mathbf{m}} \int d^3\mathbf{x}\, \psi^*_{\mathbf{n}}(\mathbf{x})\psi^*_{\mathbf{m}}(\mathbf{x})\pi_{\mathbf{n}}\pi_{\mathbf{m}} \\
&= \sum_{\mathbf{n},\mathbf{m}} \delta_{\mathbf{n},-\mathbf{m}}\pi_{\mathbf{n}}\pi_{\mathbf{m}} \\
&= \sum_{\mathbf{n}} \pi_{\mathbf{n}}(t)\pi^\dagger_{\mathbf{n}}, \\
\textstyle\int_V \phi^2(\mathbf{x}) &= \sum_{\mathbf{n}} \phi_{\mathbf{n}}\phi^\dagger_{\mathbf{n}}, \\
\textstyle\int_V (\nabla\phi(\mathbf{x}))^2 &= \sum_{\mathbf{n},\mathbf{m}} \int d^3\mathbf{x}\, (i\mathbf{k}_{\mathbf{n}}\cdot i\mathbf{k}_{\mathbf{m}})\, \psi^*_{\mathbf{n}}(\mathbf{x})\psi^*_{\mathbf{m}}(\mathbf{x})\phi_{\mathbf{n}}\phi_{\mathbf{m}} \\
&= \sum_{\mathbf{n},\mathbf{m}} (-\mathbf{k}_{\mathbf{n}}\cdot\mathbf{k}_{\mathbf{m}})\, \delta_{\mathbf{n},-\mathbf{m}}\phi_{\mathbf{n}}\phi_{\mathbf{m}} \\
&= \sum_{\mathbf{n}} \mathbf{k}^2_{\mathbf{n}}\, \phi_{\mathbf{n}}\phi^\dagger_{\mathbf{n}}.
\end{aligned} \tag{4.26}$$

Putting everything together we arrive at

$$\begin{aligned}
H = \int \mathcal{H}\, d^3\mathbf{x} &= \frac{1}{2}\int d^3\mathbf{x}\left[\pi^2 + (\nabla^i\phi)^2 + m^2\phi^2\right] \\
&= \frac{1}{2}\sum_{\mathbf{n}}\left[\pi_{\mathbf{n}}\pi^\dagger_{\mathbf{n}} + (m^2 + \mathbf{k}^2_{\mathbf{n}})\,\phi_{\mathbf{n}}\phi^\dagger_{\mathbf{n}}\right], \qquad \mathbf{k}_{\mathbf{n}} \equiv \frac{2\pi}{L}\mathbf{n}.
\end{aligned} \tag{4.27}$$

This Hamiltonian corresponds to an infinite set of harmonic oscillators with a frequency $\omega_{\mathbf{n}} = \sqrt{\mathbf{k}^2_{\mathbf{n}} + m^2}$. Notice that, because of the commutation relations, since $\phi_{-\mathbf{n}} = \phi^\dagger_{\mathbf{n}}$ and $\pi_{-\mathbf{n}} = \pi^\dagger_{\mathbf{n}}$, the contribution from $\mathbf{n}$ and $-\mathbf{n}$ in the summation are identical.

4.2.1 The continuum limit

In the infinite volume limit, $L \to \infty$, the discrete eigenvalues $\mathbf{k}_{\mathbf{n}} = \frac{2\pi}{L}\mathbf{n}$ will form a continuum. Mathematically, this corresponds to the fact that, for any given $\mathbf{k}$, there exists an integer vector function $\mathbf{n}(\mathbf{k}, L)$ such that

$$\lim_{L\to\infty} \frac{2\pi}{L}\mathbf{n}(\mathbf{k}, L) = \mathbf{k}\,. \tag{4.28}$$

Clearly $\mathbf{n}(\mathbf{k}, L) \propto L \to \infty$ in the infinite volume limit.

In the sums over $\mathbf{n}$, as long as the summand is a smooth function of $\mathbf{k}_{\mathbf{n}}$, we can thus replace the sums by integrals:

$$\sum_{\mathbf{n}}(\ldots) \to \int d^3\mathbf{n}\,(\ldots) = \left(\frac{L}{2\pi}\right)^3 \int d^3\mathbf{k}\,(\ldots). \tag{4.29}$$

Now, when $\phi(\mathbf{x},t)$ is a localized ($\equiv$ integrable) field configuration, the discrete Fourier transform $\phi_{\mathbf{n}}$ vanishes in the limit $L\to\infty$. The finite and well-behaved quantity in this limit is

$$\sqrt{V}\phi_{\mathbf{n}}(t)=\int_V d^3\mathbf{x}\, e^{-i\mathbf{k_n}\cdot\mathbf{x}}\phi(\mathbf{x},t)\xrightarrow{V\to\infty}\tilde\phi_{\mathbf{k}}(t)\equiv\int d^3\mathbf{x}\, e^{-i\mathbf{k}\cdot\mathbf{x}}\phi(\mathbf{x},t)\,. \tag{4.30}$$

where, in taking the $L\to\infty$ limit, we have assumed $\mathbf{n}=\mathbf{n}(\mathbf{k},L)$ as defined in eq. (4.28). Applying eq. (4.29) to the inverse Fourier transform and using eq. (4.30) we have

$$\phi(\mathbf{x},t)=\frac{1}{\sqrt{V}}\sum_{\mathbf{n}}e^{i\mathbf{k_n}\cdot\mathbf{x}}\phi_{\mathbf{n}}(t)\sim\frac{\sqrt{V}}{(2\pi)^3}\int d^3\mathbf{k_n}e^{i\mathbf{k_n}\cdot\mathbf{x}}\phi_{\mathbf{n}}(t)\to\int\frac{d^3\mathbf{k}}{(2\pi)^3}e^{i\mathbf{k}\cdot\mathbf{x}}\tilde\phi_{\mathbf{k}}(t). \tag{4.31}$$

Performing the infinite volume Fourier transform to the above equation and using

$$\int d^3\mathbf{x}\, e^{i(\mathbf{k}-\mathbf{p})\cdot\mathbf{x}}=(2\pi)^3\delta^3(\mathbf{k}-\mathbf{p})\,, \tag{4.32}$$

we indeed get back eq. (4.30)

$$\int d^3\mathbf{x}\, e^{-i\mathbf{p}\cdot\mathbf{x}}\phi(\mathbf{x},t)=\int d^3\mathbf{x}\,\frac{d^3\mathbf{k}}{(2\pi)^3}e^{i(\mathbf{k}-\mathbf{p})\cdot\mathbf{x}}\tilde\phi_{\mathbf{k}}(t)=\tilde\phi_{\mathbf{p}}(t)\,. \tag{4.33}$$

The same result obviously applies to the conjugated momentum, for which we define

$$\int d^3\mathbf{x}\, e^{-i\mathbf{p}\cdot\mathbf{x}}\pi(\mathbf{x},t)=\tilde\pi_{\mathbf{p}}(t)\,. \tag{4.34}$$

Using again eq. (4.32) and the commutation relations of the coordinate space fields we derive the commutation relation for their continuous Fourier modes

$$\big[\tilde\phi_{\mathbf{k}}(t),\tilde\phi_{\mathbf{p}}(t)\big]=[\tilde\pi_{\mathbf{k}}(t),\tilde\pi_{\mathbf{p}}(t)]=0\quad\big[\tilde\phi_{\mathbf{k}}(t),\tilde\pi_{\mathbf{p}}(t)\big]=i\int d^3\mathbf{x}e^{i(\mathbf{k}+\mathbf{p})x}=i(2\pi)^3\delta^3(\mathbf{p}+\mathbf{k}) \tag{4.35}$$

Similarly, we have

$$\begin{aligned}
\textstyle\int d^3\mathbf{x}\,\pi(\mathbf{x},t)^2 &= \int\frac{d^3\mathbf{p}}{(2\pi)^3}\,\frac{d^3\mathbf{k}}{(2\pi)^3}\,d^3\mathbf{x}\,e^{i(\mathbf{k}+\mathbf{p})\cdot\mathbf{x}}\tilde\pi_{\mathbf{k}}(t)\tilde\pi_{\mathbf{p}}(t)\\
&= \int\frac{d^3\mathbf{k}}{(2\pi)^3}\,\tilde\pi_{\mathbf{k}}(t)\,\tilde\pi^\dagger_{\mathbf{k}}(t),\\
\textstyle\int d^3\mathbf{x}\,\phi(\mathbf{x},t)^2 &= \int\frac{d^3\mathbf{k}}{(2\pi)^3}\,\tilde\phi_{\mathbf{k}}(t)\,\tilde\phi^\dagger_{\mathbf{k}}(t),\\
\textstyle\int d^3\mathbf{x}\,(\nabla\phi(x,t))^2 &= \int\frac{d^3\mathbf{k}}{(2\pi)^3}\,\mathbf{k}^2\tilde\phi_{\mathbf{k}}(t)\,\tilde\phi^\dagger_{\mathbf{k}}(t).
\end{aligned} \tag{4.36}$$

It follows that the Hamiltonian in the infinite volume limit is

$$H=\frac12\int\frac{d^3\mathbf{k}}{(2\pi)^3}\left(\tilde\pi_{\mathbf{k}}\tilde\pi^\dagger_{\mathbf{k}}+\omega^2_{\mathbf{k}}\tilde\phi_{\mathbf{k}}\tilde\phi^\dagger_{\mathbf{k}}\right),\qquad\qquad\omega^2_{\mathbf{k}}\equiv\mathbf{k}^2+m^2. \tag{4.37}$$

This result could have been equivalently derived by applying eq. (4.29), and definition (4.30), to the finite volume Hamiltonian (4.27).

Physically, we are interested in the infinite volume case rather than in the finite volume one. However, as it will become clear momentarily, some quantities can grow like the volume and become infinite when $V\to\infty$. In those cases, in order to properly interpret the result, it is useful to quickly deduce what would happen at finite volume. The relevant relation for that purpose is given by the one connecting the discrete Kronecker δ of the finite volume case to the Dirac δ of the infinite volume limit

$$V\delta_{\mathbf{n},\mathbf{m}}=\int_V d^3\mathbf{x}\, e^{i(\mathbf{k_n}-\mathbf{k_m})\cdot\mathbf{x}}\to(2\pi)^3\delta^3(\mathbf{k_n}-\mathbf{k_m}), \tag{4.38}$$

which for $\mathbf{n}=\mathbf{m}$ gives the formal relation

$$V=(2\pi)^3\delta^3(0). \tag{4.39}$$

We will soon be using this equation.

Our system corresponds to a collection (an infinite one!) of harmonic oscillators. In order to construct the Hilbert space and find the eigenvalues of the Hamiltonian, it is useful to recall the procedure followed in the case of the simple harmonic oscillator in quantum mechanics, where the Hamiltonian is

$$H_{SHO} = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2 \,. \tag{4.40}$$

To proceed one first defines the ladder operators a and $a^\dagger$

$$\begin{array}{ll} q = \frac{1}{\sqrt{2\omega}}(a^\dagger + a) & \qquad\qquad a = \frac{1}{\sqrt{2\omega}}\left(\omega q + ip\right) \\ & \Longrightarrow \\ p = i\sqrt{\frac{\omega}{2}}(a^\dagger - a) & \qquad\qquad a^\dagger = \frac{1}{\sqrt{2\omega}}\left(\omega q - ip\right) \end{array} \tag{4.41}$$

which, by the canonical commutation relation $[q,p] = i$, satisfy $\left[a, a^\dagger\right] = 1$. The Hamiltonian is then written as

$$H_{SHO} = \omega\left(a^\dagger a + \frac{1}{2}\right) . \tag{4.42}$$

$a^\dagger$ and a act as raising and lowering operators for the energy. In order for the Hamiltonian spectrum to be bounded from below one must then postulate the existence of a ground state $|0\rangle$ annihilated by a: $a|0\rangle = 0$. The basis of the Hilbert space is constructed by acting on the ground state with the raising operator multiple times. The resulting states $|n\rangle \equiv (a^\dagger)^n|0\rangle$, characterized by one integer quantum number n form an eigenbasis of H with eigenvalues $E_n = \omega(n + \frac{1}{2})$.

Making the parallel between eq. (4.37) and eq. (4.40) we then define

$$a_{\mathbf{k}} = \frac{1}{\sqrt{2\omega_{\mathbf{k}}}}\left(\omega_{\mathbf{k}}\tilde{\phi}_{\mathbf{k}} + i\tilde{\pi}_{\mathbf{k}}\right) , \qquad a^\dagger_{\mathbf{k}} = \frac{1}{\sqrt{2\omega_{\mathbf{k}}}}\left(\omega_{\mathbf{k}}\tilde{\phi}_{-\mathbf{k}} - i\tilde{\pi}_{-\mathbf{k}}\right) , \tag{4.43}$$

so that

$$\tilde{\phi}_{\mathbf{k}} = \frac{1}{\sqrt{2\omega_{\mathbf{k}}}}\left(a_{\mathbf{k}} + a^\dagger_{-\mathbf{k}}\right) , \qquad \tilde{\pi}_{\mathbf{k}} = -i\sqrt{\frac{\omega_{\mathbf{k}}}{2}}\left(a_{\mathbf{k}} - a^\dagger_{-\mathbf{k}}\right) . \tag{4.44}$$

The commutation relations become

$$\begin{aligned} [a_{\mathbf{k}}, a_{\mathbf{p}}] &= \frac{1}{2}\sqrt{\frac{\omega_{\mathbf{k}}}{\omega_{\mathbf{p}}}}\left[\tilde{\phi}_{\mathbf{k}}, i\tilde{\pi}_{\mathbf{p}}\right] + \frac{1}{2}\sqrt{\frac{\omega_{\mathbf{p}}}{\omega_{\mathbf{k}}}}\left[i\tilde{\pi}_{\mathbf{k}}, \tilde{\phi}_{\mathbf{p}}\right] = \frac{(2\pi)^3}{2}\left(-\sqrt{\frac{\omega_{\mathbf{k}}}{\omega_{\mathbf{p}}}}\delta^3(\mathbf{k}+\mathbf{p}) + \sqrt{\frac{\omega_{\mathbf{p}}}{\omega_{\mathbf{k}}}}\delta^3(\mathbf{k}+\mathbf{p})\right) \\ &= \frac{(2\pi)^3}{2}\left(-\delta^3(\mathbf{k}+\mathbf{p}) + \delta^3(\mathbf{k}+\mathbf{p})\right) \\ &= 0, \end{aligned} \tag{4.45}$$

$$\left[a^\dagger_{\mathbf{k}}, a^\dagger_{\mathbf{p}}\right] = -\left[a_{\mathbf{k}}, a_{\mathbf{p}}\right]^\dagger = 0, \tag{4.46}$$

$$\begin{aligned} \left[a_{\mathbf{k}}, a^\dagger_{\mathbf{p}}\right] &= \frac{1}{2}\sqrt{\frac{\omega_{\mathbf{k}}}{\omega_{\mathbf{p}}}}\left[\tilde{\phi}_{\mathbf{k}}, i\tilde{\pi}_{-\mathbf{p}}\right] + \frac{1}{2}\sqrt{\frac{\omega_{\mathbf{p}}}{\omega_{\mathbf{k}}}}\left[i\tilde{\pi}_{\mathbf{k}}, \tilde{\phi}_{-\mathbf{p}}\right] = \frac{(2\pi)^3}{2}\left(\delta^3(\mathbf{k}-\mathbf{p}) + \delta^3(\mathbf{k}-\mathbf{p})\right) \\ &= (2\pi)^3\delta^3(\mathbf{k}-\mathbf{p}). \end{aligned} \tag{4.47}$$

Therefore, by substituting the above expressions in the Hamiltonian, we get

$$H = \frac{1}{2}\int \frac{d^3\mathbf{k}}{(2\pi)^3}\left[\frac{\omega_{\mathbf{k}}}{2}\left(a_{\mathbf{k}} - a^\dagger_{-\mathbf{k}}\right)\left(a^\dagger_{\mathbf{k}} - a_{-\mathbf{k}}\right) + \frac{\omega_{\mathbf{k}}}{2}\left(a_{\mathbf{k}} + a^\dagger_{-\mathbf{k}}\right)\left(a^\dagger_{\mathbf{k}} + a_{-\mathbf{k}}\right)\right] . \tag{4.48}$$

Finally by simplifying the cross terms we have

$$\begin{aligned} H &= \frac{1}{2}\int \frac{d^3\mathbf{k}}{(2\pi)^3}\omega_{\mathbf{k}}\left[a_{\mathbf{k}}a^\dagger_{\mathbf{k}} + a^\dagger_{-\mathbf{k}}a_{-\mathbf{k}}\right] = \frac{1}{2}\int \frac{d^3\mathbf{k}}{(2\pi)^3}\omega_{\mathbf{k}}\left[a_{\mathbf{k}}a^\dagger_{\mathbf{k}} + a^\dagger_{\mathbf{k}}a_{\mathbf{k}}\right] = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\omega_{\mathbf{k}}\left[a^\dagger_{\mathbf{k}}a_{\mathbf{k}} + \frac{1}{2}\left[a_{\mathbf{k}}, a^\dagger_{\mathbf{k}}\right]\right] \\ &= \int \frac{d^3\mathbf{k}}{(2\pi)^3}\omega_{\mathbf{k}}\left[a^\dagger_{\mathbf{k}}a_{\mathbf{k}} + \frac{1}{2}(2\pi)^3\delta(0)\right]. \end{aligned} \tag{4.49}$$

The second term in the final expression represents the zero-point energy. It consists of the contribution from the continuum (an infinity) of harmonic oscillators, and it is therefore not surprising that it diverges as signified by the $\delta^3(0)$. Indeed, by eq. (4.39), we can associate the singular term $\delta^3(0)$ to the infinity of the volume and write the zero point contribution to the energy as

$$E_0 = V\int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{\omega_{\mathbf{k}}}{2}, \tag{4.50}$$

and interpret

$$\rho_0 \equiv \int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{\omega_{\mathbf{k}}}{2}, \tag{4.51}$$

as the zero point energy density. The divergence of E_0 with V is termed an *infrared divergence*, as it is associated with long-distance physics, i.e. the existence of an infinite volume. The divergence is linear in V, in accordance with the expectation that the energy of a spatially homogeneous configuration is *extensive*. The zero point energy density ρ_0 is not infrared divergent, but just a glance at its expression in eq. (4.51) suffices to realize that it suffers from another pathology: the integral over $\mathbf{k}$ diverges in the region $|\mathbf{k}| \to \infty$. This is the region of infinitely short Fourier wavelengths. Divergences from this region of wavelengths are therefore termed *ultraviolet*. Both infrared and ultraviolet divergences arise from the existence of an infinity of degrees of freedom associated with the infinity of the points $\mathbf{x}$ of physical space. This infinity is of a double nature: it is infrared because $\mathbf{x}$ can go to infinity, and it is ultraviolet because $\mathbf{x}$ is continuous, i.e. there exist points that are infinitely close. The infrared divergence of E_0 is regulated when the volume is finite and the integrals over $\mathbf{k}$ are replaced by sums over $\mathbf{n}$. However to regulate the ultraviolet divergence one would need a more dramatic change. One option, for instance, would be to make space a discrete lattice with spacing $|\Delta\mathbf{x}| = a$ between neighboring points. In that case, the integral in eq. (4.51) would be effectively cut off at $|\mathbf{k}| \sim 2\pi/a \equiv \Lambda$ and we would have $\rho_0 \sim \Lambda^4/16\pi^2$ (assuming $\Lambda \gg m$). Other options would give similar results: they would always entail a fundamental microscopic energy scale Λ where the integral in eq. (4.51) is cut-off.

We have gone through the above discussion, in good part, to illustrate the notions of infrared and ultraviolet divergences, which play a big role in more advanced treatments of quantum field theory. It should however be clear that the constant term E_0 in the Hamiltonian does not have any dynamical consequence in the system we are considering: it simply amounts to a universal unobservable phase in the quantum mechanical time evolution of the states and it does not affect the Heisenberg picture evolution of the operators (see discussion below). So we can drop this term, or, equivalently, we can add a suitable constant to $\mathcal{L}$ so as to exactly cancel it in the Hamiltonian. This is what we shall do in what follows. Notice, however, that in a more ambitious treatment of fundamental physics, where gravity is also taken into account, we would have to reckon with the consequences of this constant contribution to the Hamiltonian density. That is because it corresponds to the T^0_0 entry of the energy-momentum tensor, and the latter tensor plays the role of the source of the gravitational field in Einstein's equation. More precisely, eq. (4.51) is associated to a contribution $\rho_0\delta^\mu_\nu$ to T^μ_ν and the coefficient ρ_0 is known as the *cosmological constant*. Cosmological observations imply that ρ_0 is a very tiny energy density, in fact as tiny as if there existed some magical cancellation among the various contributions (classical and quantum) to it. The observed smallness of the cosmological constant is a gigantic open question of modern physics. It is perhaps amusing that in our first real computation in QFT we met with a quantity where QFT itself seems to fail. But, what is truly remarkable is that, putting the cosmological constant and gravity aside, the description of particle physics and fundamental interactions offered by QFT, as we shall discover[2], is unfailingly precise and successful!

Before closing this section, let us go back for a moment to the use of eq. (4.39), to interpret the $\delta^3(0)$ in eq. (4.49). We just want to stress that we could have equivalently gotten back to the finite volume case by making the reverse of eqs. (4.29,4.30) in eq. (4.48)

$$\int \frac{d^3\mathbf{k}}{(2\pi)^3} \quad\to\quad \frac{1}{V}\sum_{\mathbf{n}}, \tag{4.52}$$

$$a_{\mathbf{k}} \quad\to\quad \sqrt{V}a_{\mathbf{n}}, \tag{4.53}$$

[2]Well, at least in part: full appreciation would require the advanced course on the Standard Model.

after which by simplifying and using the commutation relations $[a_{\mathbf{n}}, a^\dagger_{\mathbf{m}}] = \delta_{\mathbf{n},\mathbf{m}}$ the Hamiltonian would have read (see also eq. (4.27))

$$H = \sum_{\mathbf{n}} \omega_n a^\dagger_{\mathbf{n}} a_{\mathbf{n}} \,. \tag{4.54}$$

The constant piece is a sum over a smooth (constant!) function of $\mathbf{n}$, thus, at large V, it is well approximated by eq. (4.50).

4.2.2 Fock space and Bose-Einstein statistics

As we just argued the dynamics of the Klein-Gordon field is not affected by the overall constant term in H, so we will eliminate it and work with the Hamiltonian

$$H = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\, \omega_{\mathbf{k}}\, a^\dagger_{\mathbf{k}} a_{\mathbf{k}} \,. \tag{4.55}$$

To construct the Hilbert space of states, notice first of all that

$$\left[H, a^\dagger_{\mathbf{k}}\right] = \omega_{\mathbf{k}} a^\dagger_{\mathbf{k}} \,, \qquad [H, a_{\mathbf{k}}] = -\omega_{\mathbf{k}} a_{\mathbf{k}}, \tag{4.56}$$

corresponding to $a^\dagger_{\mathbf{k}}$ and $a_{\mathbf{k}}$ being respectively energy-raising and energy-lowering operators. Indeed, given $H|\psi\rangle = E|\psi\rangle$ these equations imply

$$\begin{aligned} H a^\dagger_{\mathbf{k}}|\psi\rangle &= a^\dagger_{\mathbf{k}} H|\psi\rangle + \omega_{\mathbf{k}} a^\dagger_{\mathbf{k}}|\psi\rangle = (E+\omega_{\mathbf{k}}) a^\dagger_{\mathbf{k}}|\psi\rangle & (4.57)\\ H a_{\mathbf{k}}|\psi\rangle &= a_{\mathbf{k}} H|\psi\rangle - \omega_{\mathbf{k}} a_{\mathbf{k}}|\psi\rangle = (E-\omega_{\mathbf{k}}) a_{\mathbf{k}}|\psi\rangle. & (4.58)\end{aligned}$$

Secondly, one notices that the Hamiltonian in equation (4.55) is positive definite. Indeed, considering a generic normalizable state $|\psi\rangle$ one has

$$\langle\psi|H|\psi\rangle = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\, \omega_{\mathbf{k}}\, \langle\psi|a^\dagger_{\mathbf{k}} a_{\mathbf{k}}|\psi\rangle = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\, \omega_{\mathbf{k}}\, |a_{\mathbf{k}}|\psi\rangle|^2 \geq 0. \tag{4.59}$$

Then, precisely as for the harmonic oscillator, the positivity of the energy implies that the energy lowering procedure in equation (4.58) cannot be carried out indefinitely, so that there must exist a ground state $|0\rangle$ which is annihilated by all the lowering operators

$$a_{\mathbf{k}}|0\rangle = 0 \qquad \forall \mathbf{k} \tag{4.60}$$

and which, by eq. (4.55), also satisfies

$$H|0\rangle = 0. \tag{4.61}$$

(Notice that we have dropped the additive infinite constant in H). All the states can now be built by repeatedly acting with $a^\dagger_{\mathbf{k}}$ on $|0\rangle$

$$|\mathbf{k}_1, \ldots \mathbf{k}_n\rangle = a^\dagger_{\mathbf{k}_1} \ldots a^\dagger_{\mathbf{k}_n}|0\rangle, \tag{4.62}$$

which satisfies

$$H|\mathbf{k}_1, \ldots \mathbf{k}_n\rangle = (\omega_{\mathbf{k}_1} + \cdots + \omega_{\mathbf{k}_n})\, |\mathbf{k}_1, \ldots \mathbf{k}_n\rangle. \tag{4.63}$$

By a generalization of the harmonic oscillator, the basis of the Hilbert space is then given by the set

$$\left\{ \prod_j a^\dagger_{\mathbf{k}_j}|0\rangle, j = 0, 1, 2.... \right\} \equiv \{|0\rangle, |\mathbf{k}_1\rangle, |\mathbf{k}_1, \mathbf{k}_2\rangle, ...\} \,. \tag{4.64}$$

These states have respectively energies $0, \omega_{\mathbf{k}_1}, \omega_{\mathbf{k}_1}+\omega_{\mathbf{k}_2}, \ldots$, and correspond to the ground state and excited states with $E>0$.

H is the 0 is component of the 4-momentum P^μ, while for the 3-momentum we have (see eq. (4.13))

$$P^i = -P_i = -\int d^3\mathbf{x}\dot{\phi}\partial_i\phi = -\int d^3\mathbf{x}\,\pi\partial_i\phi \equiv \int \frac{d^3\mathbf{k}}{(2\pi)^3}\,\mathbf{k}^i\, a^\dagger_{\mathbf{k}} a_{\mathbf{k}}. \tag{4.65}$$

Not surprisingly the structure is the same as for the Hamiltonian, and $a^\dagger_{\mathbf{k}}$ and $a_{\mathbf{k}}$ are also raising and lowering operators for the 3-momentum:

$$\left[P^i, a^\dagger_{\mathbf{k}}\right] = \mathbf{k}^i a^\dagger_{\mathbf{k}}\ , \qquad \left[P^i, a_{\mathbf{k}}\right] = -\mathbf{k}^i a_{\mathbf{k}}. \tag{4.66}$$

Eqs. (4.56,4.66) can then be synthesised as

$$\begin{array}{llll} a^\dagger_{\mathbf{k}} & \text{creates} & \text{4-momentum} & k^\mu = (\omega_{\mathbf{k}}, \mathbf{k}) = (\sqrt{\mathbf{k}^2+m^2}, \mathbf{k}) \\ a_{\mathbf{k}} & \text{destroys} & \text{4-momentum} & k^\mu = (\omega_{\mathbf{k}}, \mathbf{k}) = (\sqrt{\mathbf{k}^2+m^2}, \mathbf{k}) \end{array} \tag{4.67}$$

where

$$k^\mu k_\mu = m^2. \tag{4.68}$$

Similarly $a^\dagger_{\mathbf{k}_1} a^\dagger_{\mathbf{k}_2}$ creates 4-momentum $k_1^\mu + k_1^\mu$ where $k_1^\mu(k_1)_\mu = k_2^\mu(k_2)_\mu = m^2$. This generalises to arbitrary products of creation operators. The $a^\dagger_{\mathbf{k}}$ create 4-momentum quanta satisfying the relation $k^\mu k_\mu = m^2$. It is natural to interpret these discrete quanta as **particles** of mass m. According to this view, the $a^\dagger_{\mathbf{k}}$ and $a_{\mathbf{k}}$ are also respectively called creation and destruction operators.

The basis of the Hilbert space 4.64 consists then of vectors $|\mathbf{k}_1, \ldots, \mathbf{k}_n\rangle \equiv a^\dagger_{\mathbf{k}_1} \ldots a^\dagger_{\mathbf{k}_2}|0\rangle$ that correspond to states with n-particles with 4-momenta $k_1^\mu = (\omega_{\mathbf{k}_1}, \mathbf{k}_1)\ \ldots\ , k_n^\mu = (\omega_{\mathbf{k}_n}, \mathbf{k}_n)$. A Hilbert space endowed with such a structure is called a **Fock Space**.

A general state is given by a general superposition of the Hilbert space basis vectors in eq. (4.64). Among these general superposition we can consider superpositions with a fixed number of particles

- vacuum - 0 particles: $|0\rangle$,
- 1 particle: $\int d^3\mathbf{k}\, f(\mathbf{k}) a^\dagger_{\mathbf{k}}|0\rangle$,
- 2 particles: $\int d^3\mathbf{k}_1 d^3\mathbf{k}_2\, f(\mathbf{k}_1, \mathbf{k}_2) a^\dagger_{\mathbf{k}_1} a^\dagger_{\mathbf{k}_2}|0\rangle$,
- $\ldots$,
- n particles: $\int \prod_{i=1}^n d^3\mathbf{k}_i\, f(\mathbf{k}_1, \ldots \mathbf{k}_n) a^\dagger_{\mathbf{k}_1} \ldots a^\dagger_{\mathbf{k}_n}|0\rangle$.

The function $f(\mathbf{k}_1, \ldots, \mathbf{k}_n)$ describes the quantum mechanical superposition of states with quanta of different momenta and plays the role of the wave function of the n-particle state. Notice that, as the creation operators commute, the basis states $|\mathbf{k}_1, \ldots, \mathbf{k}_n\rangle \equiv a^\dagger_{\mathbf{k}_1} \ldots a^\dagger_{\mathbf{k}_2}|0\rangle$ are totally symmetric under permutation of any pair of $\mathbf{k}_i$'s. As a consequence of this fact, the wave function $f(\mathbf{k}_1, \ldots, \mathbf{k}_n)$ can also be taken to be fully symmetric under the permutation of the $\mathbf{k}_i$. Indeed the contribution to the state of any antisymmetric component of the wave function trivially vanishes upon integration over $\prod_{i=1}^n d^3\mathbf{k}_i$, precisely because of the permutational symmetry of the basis states $|\mathbf{k}_1, \ldots, \mathbf{k}_n\rangle$. One can easily see this in the case of a two-particle state. Indeed by simple manipulations, one has

$$|\Psi_f\rangle \equiv \int d^3\mathbf{k}_1 d^3\mathbf{k}_2\, f(\mathbf{k}_1, \mathbf{k}_2)|\mathbf{k}_1, \mathbf{k}_2\rangle = \int d^3\mathbf{k}_1 d^3\mathbf{k}_2\, f(\mathbf{k}_1, \mathbf{k}_2)|\mathbf{k}_2, \mathbf{k}_1\rangle = \int d^3\mathbf{k}_1 d^3\mathbf{k}_2\, f(\mathbf{k}_2, \mathbf{k}_1)|\mathbf{k}_1, \mathbf{k}_2\rangle. \tag{4.69}$$

where, in the last step, we simply renamed the integration variables. From the above, we conclude that the wave functions $f(\mathbf{k}_1, \mathbf{k}_2)$ and $f(\mathbf{k}_2, \mathbf{k}_1)$ correspond to the same physical state. Correspondingly the symmetric component $f_S(\mathbf{k}_1, \mathbf{k}_2) \equiv [f(\mathbf{k}_1, k_2) + f(\mathbf{k}_2, \mathbf{k}_1)]/2$ also corresponds to the same state $|\Psi_f\rangle$, while the antisymmetric component $f_A(\mathbf{k}_1, \mathbf{k}_2) \equiv [f(\mathbf{k}_1, \mathbf{k}_2) - f(\mathbf{k}_2, \mathbf{k}_1)]/2$ corresponds to the null vector, i.e. not a state.

By this last results we have thus established that the particles associated to a relativistic quantum scalar field obey Bose-Einstein statistics. In particular, by the symmetry of the wave function $f(\mathbf{k}_1, \ldots, \mathbf{k}_n)$, a single momentum $\mathbf{k}$ can be occupied by many quanta. As we shall see later the statistics obeyed by particles depend on their spin: integer spin particles obey Bose-Einstein statistics, while half-integer spin particles obey Fermi-Dirac statistics.

4.3 Properties of the Klein-Gordon field

4.3.1 Relativistic normalization of states and measure

Let us consider the momentum operator

$$P^\mu = \int \frac{d^3\mathbf{k}}{(2\pi)^3} k^\mu a^\dagger_{\mathbf{k}} a_{\mathbf{k}}. \tag{4.70}$$

This should be a Lorentz vector, however from this way of writing it, this property is not manifest. For example, the measure of integration, $\frac{d^3k}{(2\pi)^3}$, is rotation invariant but not Lorentz invariant (not invariant under Lorentz boosts). It is better to use a Lorentz invariant measure which is simply given by

$$d\Omega_k \equiv \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{k}}}, \tag{4.71}$$

where we recall $\omega_{\mathbf{k}} = \sqrt{\mathbf{k}^2 + m^2}$. To show this consider the measure

$$\frac{dk^0 d^3\mathbf{k}}{(2\pi)^3} \delta(k^2 - m^2)\theta(k^0) = \frac{d^4k}{(2\pi)^3} \delta(k^2 - m^2)\theta(k^0), \tag{4.72}$$

which is manifestly invariant under the orthochronous Lorentz group: the first two parts are obviously invariant, while the $\theta(k^0)$ is not. The δ-function forces the 4-momentum to be inside either the forward or backward light cone. Under orthochronous Lorentz transformations, the forward light cone goes into itself, i.e. k^0 does not change sign and $\theta(k^0)$ is invariant. By integrating the measure in eq. 4.72 times a generic function $f(\mathbf{k})$ we get

$$\begin{aligned} \int \frac{d^4k}{(2\pi)^4} (2\pi)\delta(k^2 - m^2)\theta(k^0) f(\mathbf{k}) &= \int \frac{d^3\mathbf{k}}{(2\pi)^3} dk^0 \theta(k^0) \left(\frac{\delta(k^0 - \sqrt{\mathbf{k}^2 + m^2})}{2\sqrt{\mathbf{k}^2 + m^2}} + \frac{\delta(k^0 + \sqrt{\mathbf{k}^2 + m^2})}{2\sqrt{\mathbf{k}^2 + m^2}} \right) f(k) \\ &= \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{k}}} f(\mathbf{k}) \\ &= \int d\Omega_{\mathbf{k}} f(\mathbf{k}). \end{aligned} \tag{4.73}$$

With the new normalization we rewrite

$$P^\mu = \int \frac{d^3\mathbf{k}}{(2\pi)^3 2\omega_{\mathbf{k}}} k^\mu \bar{a}^\dagger_{\mathbf{k}} \bar{a}_{\mathbf{k}}, \qquad \bar{a}_{\mathbf{k}} = \sqrt{2\omega_{\mathbf{k}}} a_{\mathbf{k}}, \tag{4.74}$$

and

$$\left[\bar{a}_{\mathbf{k}}, \bar{a}^\dagger_{\mathbf{p}}\right] = (2\pi)^3 2\omega_{\mathbf{k}} \delta^3(\mathbf{k} - \mathbf{p}), \tag{4.75}$$

which is also Lorentz invariant. It follows that the normalization of the 1-particle state is

$$|\mathbf{k}\rangle \equiv \bar{a}^\dagger_{\mathbf{k}}|0\rangle, \qquad \langle \mathbf{p}|\mathbf{k}\rangle = \langle 0|\bar{a}_{\mathbf{p}} \bar{a}^\dagger_{\mathbf{k}}|0\rangle = (2\pi)^3 2\omega_{\mathbf{p}} \delta^3(\mathbf{p} - \mathbf{k}). \tag{4.76}$$

This normalization differs from the one we have used in non-relativistic quantum mechanics. The above normalization is Lorentz invariant. Using $p^i_\Lambda \equiv \Lambda^i_\mu p^\mu$ and $\omega_{\mathbf{p}_\Lambda} \equiv \Lambda^0_\mu p^\mu$ we can perform a Lorentz boost

$$\langle \mathbf{p}_\Lambda|\mathbf{k}_\Lambda\rangle = (2\pi)^3 2\omega_{\mathbf{p}_\Lambda} \delta^3(\mathbf{p}_\Lambda - \mathbf{k}_\Lambda) \equiv (2\pi)^3 2\omega_{\mathbf{p}} \delta^3(\mathbf{p} - \mathbf{k}), \tag{4.77}$$

This means that $\langle \mathbf{p}|\mathbf{k}\rangle = \langle \mathbf{p}_\Lambda|\mathbf{k}_\Lambda\rangle$. The advantage of this normalization is that Lorentz transformations are realized without additional factors. All observers have the same normalization convention.

How do we get $|\mathbf{k}_\Lambda\rangle$? Simply by acting with a representation of the Lorentz group U_Λ. So we would have

$$U_\Lambda|\mathbf{k}\rangle = c_{\mathbf{k},\Lambda}|\mathbf{k}_\Lambda\rangle \text{ and } U_\Lambda|\mathbf{p}\rangle = c_{\mathbf{p},\Lambda}|\mathbf{p}_\Lambda\rangle. \tag{4.78}$$

So,

$$(2\pi)^3 2\omega_{\mathbf{p}}\delta^3(\mathbf{p}-\mathbf{k}) = \langle \mathbf{p}|\mathbf{k}\rangle = \langle \mathbf{p}|U_\Lambda^+ U_\Lambda|\mathbf{k}\rangle = |c_{\mathbf{p},\Lambda}|^2\langle \mathbf{p}_\Lambda|\mathbf{k}_\Lambda\rangle = |c_{\mathbf{p},\Lambda}|^2(2\pi)^3 2\omega_{\mathbf{k}_\Lambda}\delta^3(\mathbf{k}_\Lambda-\mathbf{p}_\Lambda). \tag{4.79}$$

This shows that $|c_{\mathbf{p},\Lambda}|^2 = 1$.

4.3.2 States localized in space

Let us now consider the state

$$\phi(x)|0\rangle = \int \frac{d^3\mathbf{k}}{(2\pi)^3 2\omega_{\mathbf{k}}}\left(\bar{a}_{\mathbf{k}} + \bar{a}^\dagger_{\mathbf{k}}\right)e^{i\mathbf{k}\cdot\mathbf{x}}|0\rangle = \int \frac{d^3\mathbf{k}}{(2\pi)^3 2\omega_{\mathbf{k}}}e^{i\mathbf{k}\cdot\mathbf{x}}|\vec{\mathbf{k}}\rangle. \tag{4.80}$$

When $|\mathbf{k}| \ll m$ then $\omega_{\mathbf{k}} \simeq m$ we recover the usual non-relativistic expression for the position eigenstate $|\mathbf{x}\rangle$. We can thus propose the same interpretation and claim that $\phi(\mathbf{x})$ acting on the vacuum creates a particle at position $\mathbf{x}$. Consistent with this interpretation we have

$$\langle \mathbf{p}|\mathbf{x}\rangle \equiv \langle \mathbf{p}|\phi(\mathbf{x})|0\rangle = \int \frac{d^3\mathbf{k}}{(2\pi)^3 2\omega_{\mathbf{k}}}e^{i\mathbf{k}\cdot\mathbf{x}}\langle 0|\bar{a}_{\mathbf{p}}\bar{a}^\dagger_{-\mathbf{k}}|0\rangle = \int \frac{d^3\mathbf{k}}{(2\pi)^3 2\omega_{\mathbf{k}}}e^{i\mathbf{k}\cdot\mathbf{x}}(2\pi)^3 2\omega_{\mathbf{k}}\delta^3(\mathbf{k}+\mathbf{p}) = e^{-i\mathbf{p}\cdot\mathbf{x}}. \tag{4.81}$$

This is the same result as in the non-relativistic theory. $\langle \mathbf{x}|\mathbf{p}\rangle = \langle \mathbf{p}|\mathbf{x}\rangle^* = e^{i\mathbf{p}\cdot\mathbf{x}}$ is the position-space representation of the wave function of a single particle momentum eigenstate $|\mathbf{p}\rangle$.

4.3.3 Time evolution

We have so far considered the theory at a fixed time t. We could now, given our Hamiltonian, consider the Schrödinger picture evolution of our Hilbert space states

$$|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle. \tag{4.82}$$

In field theory, it is often more convenient to use the Heisenberg picture of time evolution, for which the operators rather than the states evolve with time. We can go from the Schrödinger to the Heisenberg picture (considering the operator $\hat{A}$:

$$A(t) = \langle\psi(t)|\hat{A}|\psi(t)\rangle = \langle\psi(0)|e^{iHt}\hat{A}e^{-iHt}|\psi(0)\rangle, \tag{4.83}$$

where now $e^{iHt}\hat{A}e^{-iHt} \equiv \hat{A}(t)$ is an operator in the Heisenberg picture. Let us find

$$\begin{array}{rcl}\phi(t,\mathbf{x}) & \equiv & \phi(x) \equiv e^{iHt}\phi(0,\mathbf{x})e^{-iHt}, \\ \pi(t,\mathbf{x}) & \equiv & \pi(x) \equiv e^{iHt}\pi(0,\mathbf{x})e^{-iHt}.\end{array} \tag{4.84}$$

Since ϕ and π are linear combination of a and $a^\dagger$ we must find

$$a_{\mathbf{p}}(t) \equiv e^{iHt}a_{\mathbf{p}}e^{-iHt}. \tag{4.85}$$

This can be computed by solving an ordinary differential equation

$$\begin{aligned}\frac{d}{dt}a_{\mathbf{p}}(t) &= e^{iHt}(iHa_{\mathbf{p}} - ia_{\mathbf{p}}H)e^{-iHt} \\ &= ie^{iHt}\left[H, a_{\mathbf{p}}\right]e^{-iHt} \\ &= ie^{iHt}(-\omega_{\mathbf{p}}a_{\mathbf{p}})e^{-iHt} \\ &= -i\omega_{\mathbf{p}}a_{\mathbf{p}}(t).\end{aligned} \tag{4.86}$$

The solution satisfying $a_{\mathbf{p}}(0) = a_{\mathbf{p}}$ is just

$$a_{\mathbf{p}}(t) = e^{-i\omega_{\mathbf{p}}t}a_{\mathbf{p}}. \tag{4.87}$$

We could play the same game to compute $a^\dagger_{\mathbf{p}}(t)$. This is most trivially done by noting that

$$a^\dagger_{\mathbf{p}}(t) = e^{iHt}a^\dagger_{\mathbf{p}}e^{-iHt} = \left[e^{iHt}a_{\mathbf{p}}e^{-iHt}\right]^\dagger = \left[e^{-i\omega_{\mathbf{p}}t}a_{\mathbf{p}}\right]^\dagger = e^{i\omega_{\mathbf{p}}t}a^\dagger_{\mathbf{p}}. \tag{4.88}$$

Our final result is thus

$$\begin{aligned}\phi(x) &= \int d\Omega_{\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{x}}\left(e^{-i\omega_{\mathbf{p}}t}a_{\mathbf{p}} + e^{i\omega_{\mathbf{p}}t}a^\dagger_{-\mathbf{p}}\right) \\ &= \int d\Omega_{\mathbf{p}}\left(e^{-i(\omega_{\mathbf{p}}t-\mathbf{p}\cdot\mathbf{x})}a_{\mathbf{p}} + e^{i(\omega_{\mathbf{p}}t-\mathbf{p}\cdot\mathbf{x})}a^\dagger_{\mathbf{p}}\right) \\ &= \int d\Omega_{\mathbf{p}}\left(e^{-ip_\mu x^\mu}a_{\mathbf{p}} + e^{ip_\mu x^\mu}a^\dagger_{\mathbf{p}}\right).\end{aligned} \tag{4.89}$$

The first term corresponds to the positive frequency modes (creation), while the second corresponds to the negative frequency modes (destruction). To conclude, we notice that $\phi(x)$ trivially solves the Klein-Gordon equation

$$\begin{aligned}(\Box + m^2)\phi(x) &= \int d\Omega_{\mathbf{p}}\left((\Box + m^2)e^{-ip_\mu x^\mu}a_{\mathbf{p}} + (\Box + m^2)e^{ip_\mu x^\mu}a^\dagger_{\mathbf{p}}\right) \\ &= \int d\Omega_{\mathbf{p}}(-p^2 + m^2)\left(e^{-ip_\mu x^\mu}a_{\mathbf{p}} + e^{ip_\mu x^\mu}a^\dagger_{\mathbf{p}}\right) \\ &= 0.\end{aligned} \tag{4.90}$$

This is a generalization of the well-known Ehrenfest theorem of quantum mechanics. The quantum operators in the Heisenberg picture solve the same equations of the classical theory.

Similarly to e^{iHt}, we can consider the action of $e^{-i\mathbf{P}\cdot\mathbf{x}}$. Notice that $[P^\mu, a_{\mathbf{k}}] = -k^\mu a_{\mathbf{k}}$. Consider

$$e^{-i\mathbf{P}\cdot\mathbf{b}}a_k e^{i\mathbf{P}\cdot\mathbf{b}} = e^{iP^jb_j}a_{\mathbf{k}}e^{-iP^jb_j} \equiv a_{\mathbf{k}}(\mathbf{b}), \tag{4.91}$$

and hence

$$\frac{d}{db_j}a_{\mathbf{k}}(\mathbf{b}) = e^{iP^jb_j}i\left[P^j, a_{\mathbf{k}}\right]e^{-iP^jb_j} = -ik^j a_{\mathbf{k}}(\mathbf{b}), \tag{4.92}$$

The solution to this differential equation is

$$a_{\mathbf{k}}(\mathbf{b}) = e^{-ik^jb_j}a_{\mathbf{k}} = e^{i\mathbf{k}\cdot\mathbf{b}}a_{\mathbf{k}}. \tag{4.93}$$

Combining eqs. (4.87, 4.93) and defining $P_\mu b^\mu \equiv Hb^0 - \mathbf{P}\cdot\mathbf{b}$ we can then write

$$e^{iP_\mu b^\mu}a_{\mathbf{p}}e^{iP_\mu b^\mu} = e^{-ip_\mu b^\mu}a_{\mathbf{p}} \tag{4.94}$$

which by eq. (4.89) implies

$$e^{iPb}\phi(x)e^{-iPb} = \phi(x+b) \tag{4.95}$$

and in particular

$$\phi(x) = e^{iPx}\phi(0)e^{-iPx}. \tag{4.96}$$

These results directly show that P^μ acts as the generator of space-time translations on the Heisenberg picture field $\phi(x)$. That makes us ready to consider the fate of Poincaré invariance in the quantum theory.

4.3.4 Poincarè Invariance in the quantum theory

We have already found the expression for P^μ. We now need to focus on $J^{\mu\nu}$. Using equation (4.89) for the Heisenberg picture field and the defining eqs. (3.279) and (3.262) we find (notice that $\partial/\partial_{k_i} = -\partial/\partial_{k^i}$)

$$\begin{aligned} J^{ij} &= \int d^3\mathbf{x}\,(x^i T^{0j} - x^j T^{0i}) \\ &= -i\int d\Omega_{\mathbf{k}}\, a^\dagger_{\mathbf{k}}(t)\left(k^i\frac{\partial}{\partial k^j} - k^j\frac{\partial}{\partial k^i}\right) a_{\mathbf{k}}(t) \\ &= -i\int d\Omega_{\mathbf{k}}\, a^\dagger_{\mathbf{k}}\left(k^i\frac{\partial}{\partial k^j} - k^j\frac{\partial}{\partial k^i}\right) a_{\mathbf{k}}, \end{aligned} \tag{4.97}$$

where in the last step we explicitly used eq. (4.87) to show that J^{ij} does not depend on time. Similarly we can work out the Lorentz boost charge, and find

$$\begin{aligned} J^{j0} &= \int d^3\mathbf{x}\,(x^j T^{00} - tT^{0j}) \\ &= \int d\Omega_{\mathbf{k}}\, a^\dagger_{\mathbf{k}}(t)\left(i\omega_{\mathbf{k}}\frac{\partial}{\partial k^j} - tk^j\right) a_{\mathbf{k}}(t) \\ &= \int d\Omega_{\mathbf{k}}\, a^\dagger_{\mathbf{k}} i\omega_{\mathbf{k}}\frac{\partial}{\partial k^j} a_{\mathbf{k}}. \end{aligned} \tag{4.98}$$

where again in the last step we used the explicit form of $a_{\mathbf{k}}(t)$ to show that J^{j0} is time independent. To draw physical consequences we need the commutators

$$\begin{aligned} \left[J^{ij}, a_{\mathbf{p}}\right] &= i\left(p^i\frac{\partial}{\partial p^j} - p^j\frac{\partial}{\partial p^i}\right) a_{\mathbf{p}}, \\ \left[J^{i0}, a_{\mathbf{p}}\right] &= -i\omega_{\mathbf{p}}\frac{\partial}{\partial p^i} a_{\mathbf{p}}. \end{aligned} \tag{4.99}$$

The action of $J^{\mu\nu}$ on the states determines their Lorentz transformation properties, as we shall now discuss. Note that $J^{\mu\nu}|0\rangle = 0$, corresponding to the vacuum being invariant under both rotations and boosts.

Let us focus first on angular momentum $J^k = \frac{1}{2}\epsilon^{kij}J^{ij}$ (see eq. (3.143)), which by eq. (4.97) reads

$$J^k = \int d\Omega_{\mathbf{p}}\, a^\dagger_{\mathbf{p}}(-i\mathbf{p}\wedge\nabla_{\mathbf{p}})^k a_{\mathbf{p}} \equiv \int d\Omega_{\mathbf{p}}\, a^\dagger_{\mathbf{p}}\hat{L}^k a_{\mathbf{p}}\,. \tag{4.100}$$

Where in the last step, by recalling that $i\nabla_{\mathbf{p}}$ is the coordinate operator $\mathbf{x}$ in the momentum representation, we have identified the differential operator in brackets with the ordinary orbital angular momentum $\vec{L} = \mathbf{x}\wedge\mathbf{p} = -\mathbf{p}\wedge\mathbf{x}$.

Let us now consider the most general single particle state $|f\rangle = \int d\Omega_{\mathbf{p}} f(\mathbf{p}) a^\dagger_{\mathbf{p}}|0\rangle$, where $f(\mathbf{p})$ is a a generic square integrable function, and compute the action of J^k on it. By the above equations we have

$$\left[J^k, a^\dagger_{\mathbf{p}}\right] = i(\mathbf{p}\wedge\nabla_{\mathbf{p}})^k a^\dagger_{\mathbf{p}} \tag{4.101}$$

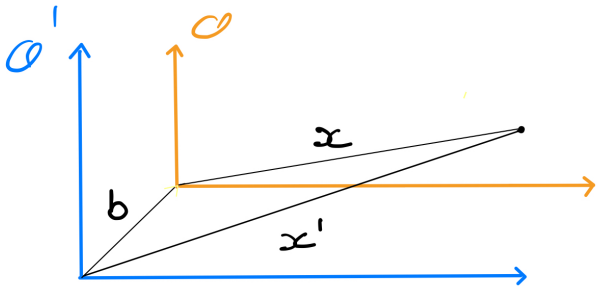


Figure 4.1: The reference frames of observers $\mathcal{O}$ and $\mathcal{O}'$ related by a translation b^μ.

so that

$$\begin{aligned} J^k|f\rangle &= \int d\Omega_{\mathbf{p}} f(\mathbf{p}) \left[J^k, a^\dagger_{\mathbf{p}}\right]|0\rangle \\ &= \int d\Omega_{\mathbf{p}} f(\mathbf{p}) i(\mathbf{p}\wedge\nabla_{\mathbf{p}})^k a^\dagger_{\mathbf{p}}|0\rangle \\ &= \int d\Omega_{\mathbf{p}} \left[-i(\mathbf{p}\wedge\nabla_{\mathbf{p}})^k f(\mathbf{p})\right] a^\dagger_{\mathbf{p}}|0\rangle \\ &= \int d\Omega_{\mathbf{p}} \left[\hat{L}^k f(\mathbf{p})\right] a^\dagger_{\mathbf{p}}|0\rangle. \end{aligned} \tag{4.102}$$

which shows that the action of J^k on the state $|f\rangle$ is expressed as the pure action of the orbital angular momentum differential operator $\hat{L}^k$ of the wave function $f(k)$. This matches the action of the quantum mechanical angular momentum operator J^k for a particle without spin. We conclude that the quanta associated with ϕ have zero spin.

We can know study the Poincaré transformations of $\phi(x)$. Applying eqs (4.99) to the Heisenberg picture field we find [3]

$$[P^\mu, \phi(x)] = -i\partial^\mu\phi(x), \qquad [J^{\mu\nu}, \phi(x)] = -i\left(x^\mu\partial^\nu - x^\nu\partial^\mu\right)\phi(x). \tag{4.103}$$

These equation explicitly show that P^μ and $J^{\mu\nu}$ do act as the infinitesimal generators of the Poincaré group. To find a finite Poincaré transformation we just need to exponentiate the infinitesimal transformation. For translations that was already accomplished in the previous section, while we shall take care of boosts later. What we would like to do first is to discuss the action of space-time translations as implementing a "Change of Observer".

Consider two observers $\mathcal{O}$ and $\mathcal{O}'$ whose coordinates are related by $x'^\mu = x^\mu + b^\mu$, as shown in Fig. (4.1). The fields used by the two observer are related by $\phi(x) = \phi'(x')$ or, equivalently, $\phi(x-b) = \phi'(x)$ which, by eq. (4.95) can be written as

$$\boxed{e^{-iPb}\phi(x)e^{iPb} = \phi(x-b) = \phi'(x)\,.} \tag{4.104}$$

The interpretation of this last equation is then that the change of frame from $\mathcal{O}$ to $\mathcal{O}'$ is operated by a unitary operator e^{-iPb}. The same interpretation holds for the creation and destruction operators and for the corresponding states. Indeed, decomposing ϕ and ϕ' in Fourier modes the above equation implies (remember $kb \equiv k_\mu b^\mu$)

$$a'_{\mathbf{k}} = e^{-iPb}a_{\mathbf{k}}e^{iPb} = e^{ikb}a_{\mathbf{k}} \tag{4.105}$$

and hence

$$|\mathbf{k}\rangle' \equiv a'^\dagger_{\mathbf{k}}|0\rangle = e^{-iPb}a^\dagger_{\mathbf{k}}e^{iPb}|0\rangle = e^{-iPb}|\mathbf{k}\rangle = e^{-ikb}|\mathbf{k}\rangle \tag{4.106}$$

[3] These results can equivalently and more directly be derived by expressing the generators in eqs. (3.263,3.279) in terms of the quantum field operators $\phi(x)$ and $\pi(x)$ and by then using the canonical commutation relation (4.14).

We would next like to apply the same "Change of Observer" perspective to Lorentz transformations. Consider then two observers $\mathcal{O}$ and $\mathcal{O}'$ whose coordinates are related by a Lorentz transformation $x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu}$. The corresponding fields are related by $\phi'(x') = \phi(x)$, or, equivalently, $\phi'(x) = \phi(\Lambda^{-1}x)$. Taking $\omega_{\mu\nu}$ to be the Lie parameters that define Λ according to eq. (3.138), the representation of Λ on the Hilbert space is given by

$$U(\Lambda) = e^{-\frac{i}{2}\omega_{\mu\nu}J^{\mu\nu}} . \tag{4.107}$$

Now, by using the commutation relation in eq. (4.103) and applying the standard exponentiation trick as we will momentarily show, one finds

$$\boxed{U^{\dagger}(\Lambda)\phi(x)U(\Lambda) = \phi(\Lambda^{-1}x) \equiv \phi'(x)\,.} \tag{4.108}$$

Precisely like for translation the intepretation of this equation is that the change of frame from $\mathcal{O}$ to $\mathcal{O}'$ is operated by the unitary operator that represents Λ on the Hilbert space. We can similarly expend the discussion to ladder operators and to states. However before moving to that we howe to the reader a proof of eq. (4.108).

The idea is to express the representative $U(\Lambda)$ of the finite Lorentz transformation Λ as the product of "large" number N of "small" Lorentz transformation

$$U(\Lambda) = e^{-\frac{i}{2}\omega_{\mu\nu}J^{\mu\nu}} = \left(e^{-\frac{i}{2}\frac{\omega_{\mu\nu}}{N}J^{\mu\nu}}\right)^N = \left(\mathbb{1} - \frac{i}{2}\frac{\omega_{\mu\nu}}{N}J^{\mu\nu} + O(\frac{1}{N^2})\right)^N \tag{4.109}$$

and then use that eq. (4.103) fixes the action of an infinitesimal transformation. Considering such an infinitesimal transformation with Lie parameter $\epsilon_{\mu\nu} \ll 1$, by eq. (4.103) we can then write

$$\begin{aligned} e^{\frac{i}{2}\epsilon_{\mu\nu}J^{\mu\nu}}\phi(x)e^{-\frac{i}{2}\epsilon_{\mu\nu}J^{\mu\nu}} &= \phi(x) + \frac{i}{2}\epsilon_{\mu\nu}\left[J^{\mu\nu},\phi(x)\right] + O(\epsilon^2) \\ &= \phi(x) + \frac{\epsilon_{\mu\nu}}{2}\left(x^{\mu}\partial^{\nu} - x^{\nu}\partial^{\mu}\right)\phi(x) + O(\epsilon^2) \\ &= \phi\left(x^{\mu} - \epsilon^{\mu}{}_{\rho}x^{\rho}\right) + O(\epsilon^2) \\ &= \phi\left((\delta^{\mu}{}_{\rho} - \epsilon^{\mu}{}_{\rho})x^{\rho}\right) + O(\epsilon^2) \\ &\equiv \phi\left((\mathbb{1} - \epsilon)x\right) + O(\epsilon^2), \end{aligned} \tag{4.110}$$

where one should appreciate that up to $O(\epsilon^2)$ terms the matrix acting on x in the last step is just the infinitesimal Lorentz transformation $\Lambda^{\mu}{}_{\rho} \simeq \delta^{\mu}{}_{\rho} - \epsilon^{\mu}{}_{\rho}$. By identifying $\epsilon_{\mu\nu} = \omega_{\mu\nu}/N$ we can finally write

$$U(\Lambda)^{\dagger}\phi(x)U(\Lambda) = \left(e^{\frac{i}{2}\frac{\omega_{\mu\nu}}{N}J^{\mu\nu}}\right)^N \phi(x) \left(e^{-\frac{i}{2}\frac{\omega_{\mu\nu}}{N}J^{\mu\nu}}\right)^N \tag{4.111}$$

$$= \phi\left((\mathbb{1} - \frac{\omega}{N})^N x\right) + O(\frac{1}{N}) \tag{4.112}$$

$$\overset{N\to\infty}{=} \phi(\Lambda^{-1}x) \tag{4.113}$$

where one should appreciate that going from the first to the second line we simply iterated N times eq. (4.110), also noticing that, by adding up the N operations, the $O(\epsilon^2) = O(1/N^2)$ remainders sum up to an $O(\epsilon) = O(1/N)$ contribution. In any case this residual term goes to zero as the limit $N \to \infty$ is taken, in the last step.

We are now ready to discuss the implication of eq. (4.108) on ladder operators and states. Using the decomposition of ϕ and ϕ' in ladder operators for the first equation we find

$$\begin{aligned} &\int d\Omega_{\mathbf{p}}\left(e^{-ip_{\mu}x^{\mu}}U^{\dagger}(\Lambda)a_{\mathbf{p}}U(\Lambda) + e^{ip_{\mu}x^{\mu}}U^{\dagger}(\Lambda)a^{\dagger}_{\mathbf{p}}U(\Lambda)\right) = \int d\Omega_{\mathbf{p}}\left(e^{-ip_{\mu}(\Lambda^{-1}x)^{\mu}}a_{\mathbf{p}} + e^{ip_{\mu}(\Lambda^{-1}x)^{\mu}}a^{\dagger}_{\mathbf{p}}\right) \\ &= \int d\Omega_{\Lambda^{-1}\mathbf{p}'}\left(e^{-i(\Lambda^{-1}p')_{\mu}(\Lambda^{-1}x)^{\mu}}a_{\Lambda^{-1}\mathbf{p}'} + e^{i(\Lambda^{-1}p')_{\mu}(\Lambda^{-1}x)^{\mu}}a^{\dagger}_{\Lambda^{-1}\mathbf{p}'}\right) \\ &= \int d\Omega_{\mathbf{p}'}\left(e^{-ip'_{\mu}x^{\mu}}a_{\Lambda^{-1}\mathbf{p}'} + e^{ip'_{\mu}x^{\mu}}a^{\dagger}_{\Lambda^{-1}\mathbf{p}'}\right) = \int d\Omega_{\mathbf{p}}\left(e^{-ip_{\mu}x^{\mu}}a_{\Lambda^{-1}\mathbf{p}} + e^{ip_{\mu}x^{\mu}}a^{\dagger}_{\Lambda^{-1}\mathbf{p}}\right). \end{aligned} \tag{4.114}$$

where in the second line we change integration variable according to $p = \Lambda^{-1}p'$, while in the third we used the Lorentz invariance of the measure $d\Omega_{\mathbf{p}'} = d\Omega_{\Lambda^{-1}\mathbf{p}'}$ and of the product $(\Lambda^{-1}p')_\mu(\Lambda^{-1}x)^\mu = p'_\mu x^\mu$. In the final identity we simply relabeled $p' \to p$. Identifying the coefficients of the Fourier modes in the first and in the last expression we then conclude that $U^\dagger(\Lambda)a_{\mathbf{p}}U(\Lambda) = a_{\Lambda^{-1}\mathbf{p}}$. On the other hand equality of the fiurst and third term in eq. (4.108) simply reads $U^\dagger(\Lambda)a_{\mathbf{p}}U(\Lambda) = a'_{\mathbf{p}}$. Putting all this together, we have that eq. (4.108) is expressed in momentum space as

$$\boxed{U^\dagger(\Lambda)a_{\mathbf{p}}U(\Lambda) = a_{\Lambda^{-1}\mathbf{p}} = a'_{\mathbf{p}}.} \tag{4.115}$$

and similarly for the creation operators. Using the Lorentz invariance of the vacuum, $U(\Lambda|0\rangle = |0\rangle$, for single particle states these relations imply

$$U^\dagger(\Lambda)|\mathbf{p}\rangle = |\Lambda^{-1}\mathbf{p}\rangle = |\mathbf{p}\rangle' \tag{4.116}$$

As it should have been expected, a particle of momentum $\mathbf{p}$ in the $\mathcal{O}'$ frame corresponds to a particle of momentum $\Lambda^{-1}\mathbf{p}$ in the $\mathcal{O}$ frame. Moreover the states $|\mathbf{p}\rangle$ and $|\mathbf{p}\rangle'$, carrying the same momentum but in respectively $\mathcal{O}$ and $\mathcal{O}'$ are related to one another by a unitary operator $U(\Lambda)$ which represents the Lorentz transformation Λ relating the two frames. Adding this result to what we previously found for translations, we conclude that the quantum fields and the states of different inertial observers are related by unitary transformations that realize a representation of the Poincaré group.

What we have just touched upon is just one incarnation of Wigner's theorem, which states that the elements g of a classical symmetry transformation group G are realized at the quantum level by unitary operators $U(g)$ offering a representation of G on the Hilbert space.

4.4 The charged scalar field

4.4.1 Construction

A charged scalar field is a scalar field under Poincaré transformations that does not satisfy the hermiticity condition. As such, it can be represented in term of two scalar fields

$$\phi = \frac{\phi_1 + i\phi_2}{,}\sqrt{2} \tag{4.117}$$

where both ϕ_1, ϕ_2 are real. We can write the Lagrangian as a sum of two Klein-Gordon Lagrangians

$$\mathcal{L} = \mathcal{L}_{KG}(\phi_1) + \mathcal{L}_{KG}(\phi_2) = \frac{1}{2}\partial_\mu\phi_1\partial^\mu\phi_1 - \frac{1}{2}m^2\phi_1^2 + \frac{1}{2}\partial_\mu\phi_2\partial^\mu\phi_2 - \frac{1}{2}m^2\phi_2^2 = \partial_\mu\phi^*\partial^\mu\phi - m^2\phi^*\phi. \tag{4.118}$$

This lagrangian is not only Poincarè invariant. It has a $U(1)$ internal symmetry that amounts to the invariance with respect to the transformation

$$\begin{aligned} &x^\mu \to x'^\mu = x^\mu, \\ &\phi(x) \to \phi'(x) = e^{-i\alpha}\phi(x) \qquad \alpha \in \mathbb{R}. \end{aligned} \tag{4.119}$$

Note that this is realized as an $SO(2)$ rotation on the real fields ϕ_1 and ϕ_2.

4.4.2 Internal current and charge

Using the field's transformation law

$$\begin{cases} \Delta_\phi = \dfrac{\phi'(x) - \phi(x)}{\alpha} = -\dfrac{i\alpha\phi(x)}{\alpha} = -i\phi(x) \\ \Delta_{\phi^*} = \dfrac{\phi'^*(x) - \phi^*(x)}{\alpha} = \dfrac{i\alpha\phi^*(x)}{\alpha} = i\phi^*(x) \end{cases} \tag{4.120}$$

the Noether current is found to be,

$$J^\mu = -i\partial_\mu\phi^*\phi + i\partial_\mu\phi\phi^* = i\left(\phi^*\partial_\mu\phi - (\partial_\mu\phi^*)\,\phi\right) \equiv i\phi^*\overleftrightarrow{\partial_\mu}\phi = \phi_2\partial_\mu\phi_1 - \phi_1\partial_\mu\phi_2. \tag{4.121}$$

The equations of motion are

$$\begin{array}{l}\partial_\mu\dfrac{\partial\mathcal{L}}{\partial\partial_\mu\phi^*} - \dfrac{\partial\mathcal{L}}{\partial\phi^*} = 0,\\ \partial_\mu\partial^\mu\phi - m^2\phi = 0,\end{array} \tag{4.122}$$

which are just the Klein-Gordon equation. It is easy to check that $\partial_\mu J^\mu = 0$ on the equations of motion.

4.4.3 Quantisation

The canonical field conjugate are

$$\pi(\mathbf{x},t) = \frac{\partial\mathcal{L}}{\partial\dot\phi(\mathbf{x},t)} = \dot\phi^\dagger(\mathbf{x},t), \quad \pi^\dagger(\mathbf{x},t) = \frac{\partial\mathcal{L}}{\partial\dot\phi^*(\mathbf{x},t)} = \dot\phi(\mathbf{x},t), \tag{4.123}$$

and hence the Hamiltonian is

$$H = \int d^3\mathbf{x}\left[\pi^\dagger(\mathbf{x},t)\pi(\mathbf{x},t) + |\nabla\phi(\mathbf{x},t)|^2 + m^2\phi^\dagger(\mathbf{x},t)\phi(\mathbf{x},t)\right]. \tag{4.124}$$

The commutation relations can be written in complex notation as

$$\begin{array}{rcl}[\phi(\mathbf{x},t),\phi(\mathbf{y},t)] &=& \left[\phi^\dagger(\mathbf{x},t),\phi^\dagger(\mathbf{y},t)\right] = \left[\phi(\mathbf{x},t),\phi^\dagger(\mathbf{y},t)\right] = 0,\\ [\pi(\mathbf{x},t),\pi(\mathbf{y},t)] &=& \left[\pi^\dagger(\mathbf{x},t),\pi^\dagger(\mathbf{y},t)\right] = \left[\pi(\mathbf{x},t),\pi^\dagger(\mathbf{y},t)\right] = 0,\\ [\phi(\mathbf{x},t),\pi(\mathbf{y},t)] &=& \left[\phi^\dagger(\mathbf{x},t),\pi^\dagger(\mathbf{y},t)\right] = i\delta^3(\mathbf{x}-\mathbf{y}).\end{array} \tag{4.125}$$

Or in terms of the real fields

$$\begin{array}{rcl}[\phi_i(\mathbf{x},t),\phi_j(\mathbf{y},t)] &=& 0,\\ [\pi_i(\mathbf{x},t),\pi_j(\mathbf{y},t)] &=& 0,\\ [\phi_i(\mathbf{x},t),\pi_j(\mathbf{y},t)] &=& i\delta_{ij}\delta^3(\mathbf{x}-\mathbf{y}),\end{array} \tag{4.126}$$

for $i,j = 1,2$. Since $\mathcal{L}$ is the sum of the Lagrangian for ϕ_1 and ϕ_2, we can proceed with the canonical quantisation of ϕ_1 and ϕ_2 independently. Let us do so. We shall go back to complex notation at the end. Applying the results of the previous chapter we have in the Heisenberg picture

$$\begin{array}{rcl}\phi_1(\mathbf{x}) &=& \displaystyle\int d\Omega_\mathbf{k}\, a_{\mathbf{k}\,1}e^{-i\mathbf{k}\cdot\mathbf{x}} + a^\dagger_{\mathbf{k}\,1}e^{+i\mathbf{k}\cdot\mathbf{x}}\\ \phi_2(\mathbf{x}) &=& \displaystyle\int d\Omega_\mathbf{k}\, a_{\mathbf{k}\,2}e^{-i\mathbf{k}\cdot\mathbf{x}} + a^\dagger_{\mathbf{k}\,2}e^{+i\mathbf{k}\cdot\mathbf{x}},\end{array} \tag{4.127}$$

with $\omega_k = \sqrt{m^2+\mathbf{k}^2}$ as usual. Now

$$P^\mu = \int d\Omega_\mathbf{p}\, p^\mu(a^\dagger_{\mathbf{p}\,1}a_{\mathbf{p}\,1} + a^\dagger_{\mathbf{p}\,2}a_{\mathbf{p}\,2}) \tag{4.128}$$

The complex scalar field consists of two types of particles, generated by $a_1^\dagger$ and $a_2^\dagger$, with the same mass m. In complex notation

$$\begin{array}{rcl}\phi(x) &=& \int d\Omega_\mathbf{k}\, a_\mathbf{k}e^{-ik\cdot x} + b^\dagger_\mathbf{k}e^{+ik\cdot x},\\ \phi^\dagger(x) &=& \int d\Omega_\mathbf{k}\, b_\mathbf{k}e^{-ik\cdot x} + a^\dagger_\mathbf{k}e^{+ik\cdot x}\end{array} \tag{4.129}$$

with the new ladder operators

$$\begin{aligned} a_{\mathbf{k}} &= \frac{a_{\mathbf{k}\,1} + i a_{\mathbf{k}\,2}}{\sqrt{2}} \\ b_{\mathbf{k}} &= \frac{a_{\mathbf{k}\,1} - i a_{\mathbf{k}\,2}}{\sqrt{2}} \end{aligned} \tag{4.130}$$

The notation with a, b is more convenient to keep track of the charge of the states. The commutation relations are

$$\begin{aligned} &[a_{\mathbf{k}}, a_{\mathbf{p}}] = [a_{\mathbf{k}}, b_{\mathbf{p}}] = [b_{\mathbf{k}}, b_{\mathbf{p}}] = 0, \\ &[a_{\mathbf{k}}, a^\dagger_{\mathbf{p}}] = [b_{\mathbf{k}}, b^\dagger_{\mathbf{p}}] = (2\pi)^3 2p^0 \delta^{(3)}(\mathbf{p} - \mathbf{k}). \end{aligned} \tag{4.131}$$

So we can rewrite P^μ in terms of these operators

$$P^\mu = \int d\Omega_{\mathbf{p}}\, p^\mu \left(a^\dagger_{\mathbf{p}} a_{\mathbf{p}} + b^\dagger_{\mathbf{p}} b_{\mathbf{p}} \right), \tag{4.132}$$

and

$$Q = \int J^0(\mathbf{x}, 0)\, d^3\mathbf{x} = \int d\Omega_{\mathbf{p}} \left(a^\dagger_{\mathbf{p}} a_{\mathbf{p}} + b^\dagger_{\mathbf{p}} b_{\mathbf{p}} \right), \tag{4.133}$$

and

$$\left[Q, a^\dagger_{\mathbf{k}} \right] = a^\dagger_{\mathbf{k}}, \qquad \left[Q, b^\dagger_{\mathbf{k}} \right] = -b^\dagger_{\mathbf{k}}. \tag{4.134}$$

The commutators are

$$\begin{aligned} \left[Q, a^\dagger_{\mathbf{k}} \right] &= a^\dagger_{\mathbf{k}}, \qquad [Q, a_{\mathbf{k}}] = -a_{\mathbf{k}}, \\ \left[Q, b^\dagger_{\mathbf{k}} \right] &= -b^\dagger_{\mathbf{k}}, \qquad [Q, b_{\mathbf{k}}] = b_{\mathbf{k}}. \end{aligned} \tag{4.135}$$

Thus $a^\dagger_{\mathbf{k}}$ raises the charge by one unit (i.e. creates particles of charge $+1$). While $b^\dagger_{\mathbf{k}}$ decreases the charge by one unit (i.e. creates particle of charge -1). Notice that $e^{-i\alpha Q}$ generates charge transformation

$$\begin{aligned} e^{i\alpha Q} a_{\mathbf{k}} e^{-i\alpha Q} &= (1 + i\alpha Q) a_{\mathbf{k}} (1 - i\alpha Q) = a_{\mathbf{k}} - i\alpha + .. = e^{-i\alpha} a_{\mathbf{k}}, \\ e^{i\alpha Q} b_{\mathbf{k}} e^{-i\alpha Q} &= e^{i\alpha} b_{\mathbf{k}}. \end{aligned} \tag{4.136}$$

From this follows $e^{i\alpha Q} \phi e^{-i\alpha Q} = e^{-i\alpha} \phi$. The particles generated by $a^\dagger, b^\dagger$ have exactly the same 4-momentum $k^\mu = \left(\sqrt{k^2 + m^2}, k^i \right)$, mass, spin (here 0), but equal and opposite charge. It is conventional to say that $a^\dagger_{\mathbf{k}}$ generates a particle, while $b^\dagger_{\mathbf{k}}$ generates an anti-particle. A deep symmetry of the theory (CPT) requires the masses of particle and antiparticle to be the same, basically, a and b are associated with the same complex field ϕ.

Examples of particle-antiparticle pairs with spin-0 are π^+, π^- and K^+, K^-. Both pairs have opposite electromagnetic charges. In the case of K^0 and $\bar{K}^0$ the particles have opposite strangeness. Notice that the original operators $a^\dagger_1$ and $a^\dagger_2$ create particles that are coherent quantum superpositions of particles and antiparticles. This state is not an eigenstate of the charge Q.

$$\begin{aligned} |\psi\rangle &= (b^\dagger + a^\dagger)|0\rangle, \\ Q|\psi\rangle &= (b^\dagger - a^\dagger)|0\rangle \neq |\psi\rangle. \end{aligned} \tag{4.137}$$

Chapter 5

Spinor fields

5.1 Spinor representations of the Lorentz group

Let us recall that the irreducible representations of the Lorentz group are labelled by two (half)-integer numbers (j_-, j_+) so that the most general irreducible representation of Lorentz transformations on fields is:

$$\Phi_a(x) \;\mapsto\; \underbrace{D_a{}^b(\Lambda)}_{\text{spin part}}\,\Phi_b(\;\underbrace{\Lambda^{-1}x}_{\text{orbital part}}\;) \tag{5.1}$$

where $D_a{}^b(\Lambda)$ are matrices realizing the (j_-, j_+) representation on a vector space of dimension $(2j_-+1)(2j_++1)$. As we saw, the trivial representation $(0,0)$ corresponds to a scalar field with one index: $\phi(x)$. Indeed, $j_\pm = 0 \Rightarrow J^i_\pm \equiv 0$ so that $D_a{}^b(\Lambda) \equiv \mathbb{1}$. The first non trivial cases correspond to the $(1/2, 0)$ and $(0, 1/2)$ representations. Recalling that

$$J^i_\pm \equiv \frac{J^i \pm iK^i}{2}, \qquad [J^i_\pm, J^j_\pm] = i\epsilon^{ijk}J^k_\pm \tag{5.2}$$

this leads to the so-called *left-handed* spinorial representation of the Lorentz group algebra:

$$(1/2,0) \;\Rightarrow\; \left\{\begin{array}{ll} J^i_- & = \frac{\sigma^i}{2} \\ J^i_+ & = 0 \end{array}\right. \;\Rightarrow\; \begin{array}{ll} J^i & = \frac{\sigma^i}{2} \\ K^i & = +i\frac{\sigma^i}{2} \end{array} \tag{5.3}$$

and to the *right-handed* one:

$$(0,1/2) \;\Rightarrow\; \left\{\begin{array}{ll} J^i_- & = 0 \\ J^i_+ & = \frac{\sigma^i}{2} \end{array}\right. \;\Rightarrow\; \begin{array}{ll} J^i & = \frac{\sigma^i}{2} \\ K^i & = -i\frac{\sigma^i}{2} \end{array} \tag{5.4}$$

where σ^i are the 2×2 Pauli matrices satisfying the algebra

$$\sigma^i\sigma^j = \delta^{ij} + i\epsilon^{ijk}\sigma^k \tag{5.5}$$

and offering the spin-$\frac{1}{2}$ representation of $SO(3)$: $[\frac{\sigma^i}{2}, \frac{\sigma^j}{2}] = i\epsilon^{ijk}\frac{\sigma^k}{2}$.

Finally, recalling[1]:

[1] Throughout the discussion in this chapter we use both $\omega_{\mu\nu}$ and $(\vec{\theta}, \vec{\eta})$ related by eqs. (3.143), (3.144)

$$D(\Lambda(\omega)) = e^{-\frac{i}{2}\omega_{\mu\nu}J^{\mu\nu}} = e^{-i\vec{\theta}\cdot\vec{J}+i\vec{\eta}\cdot\vec{K}} \tag{5.6}$$

We obtain the left-handed and right-handed spinorial representations of the Lorentz transformations:

$$\Lambda_L(\omega) \equiv D_{(1/2,0)} = e^{-\frac{1}{2}(\vec{\eta}+i\theta)\cdot\sigma} \tag{5.7}$$
$$\Lambda_R(\omega) \equiv D_{(0,1/2)} = e^{-\frac{1}{2}(-\vec{\eta}+i\theta)\cdot\sigma} \tag{5.8}$$

which are 2×2 complex matrices that, according to the general construction in sect. 3.2.5, act on 2-dimensional complex vectors, called respectively the *left-handed spinor* and the *right-handed spinor*:

$$(1/2,0) \;\Rightarrow\; \psi_L(x), \qquad\qquad (0,1/2) \;\Rightarrow\; \psi_R(x) \tag{5.9}$$

For both representations the base space is then the two-dimensional complex space $\mathbb{C}^2$. Given a Lorentz transformation $\Lambda(\omega) = e^{-i\omega_{\mu\nu}\mathcal{J}^{\mu\nu}/2}$, the general transformation of eq. (5.1) is realized on ψ_L and ψ_R according to

$$\psi'_L(x) = \Lambda_L(\omega)\psi_L(\Lambda(\omega)^{-1}x)\,, \qquad\qquad \psi'_R(x) = \Lambda_R(\omega)\psi_R(\Lambda(\omega)^{-1}x)\,. \tag{5.10}$$

One important fact to notice is that the two representations $\Lambda_L(\omega)$ and $\Lambda_R(\omega)$ are non-unitary:

$$\Lambda^\dagger_{L/R} \neq \Lambda^{-1}_{L/R} \tag{5.11}$$

That is because finite dimensional representations of non-compact groups are non-unitary. Among the finite dimensional representations, only the trivial one is unitary. On the other hand, one readily verifies that

$$\Lambda_L = \left(\Lambda^\dagger_R\right)^{-1} \tag{5.12}$$

Moreover, since Λ_L and Λ_R are exponentials of traceless matrices, they have unit determinant. Thus they actually span the group $SL(2,\mathbb{C})$:

$$\det\Lambda_L = \det\Lambda_R = 1 \quad\Rightarrow\quad \Lambda_L, \Lambda_R \in SL(2,\mathbb{C}) \tag{5.13}$$

We will actually see later that:

$$SO(1,3) \simeq SL(2,\mathbb{C})/\mathbb{Z}_2 \tag{5.14}$$

In both representations, spatial rotations act in the same way:

$$e^{-i\vec{\theta}\cdot\vec{\sigma}/2}\psi_{L/R}(R_\theta^{-1}x) \tag{5.15}$$

The above transformation is precisely the same undergone by the wave function of a particle of spin-$\frac{1}{2}$ in non-relativitic quantum mechanics. The action on the coordinates $x \to R_\theta^{-1}x$ corresponds to orbital angular momentum, while that operated by $\sigma^i/2$ corresponds to intrinsic angular momentum, i.e. spin. As will become clear later on, this does not happen by chance: the quanta of the spinor fields are indeed particles carrying spin 1/2.

5.2 Spinor miscellanea

5.2.1 Parity

One question that arises when dealing with spinors is "why do we have two inequivalent 2-dimensional representations?". The doubling is indeed related to the existence of *parity*, $x^\mu \to P^\mu{}_\nu x^\nu$, see eq. (3.120). Parity is just one very specific Lorentz transformation. Its action on the generators is then simply

$$P \ : \ \mathcal{J}^{\mu\nu} \to P^\mu{}_\rho P^\nu{}_\sigma \mathcal{J}^{\rho\sigma} \tag{5.16}$$

which on respectively rotations J^i and boosts K^i reads

$$P \ : \ \begin{cases} J^i & \mapsto \ J^i, \quad \text{axial vector} \\ K^i & \mapsto -K^i, \quad \text{polar vector} \end{cases} \tag{5.17}$$

That the generators of spatial rotations and boosts form respectively an axial and a polar vector can also be explicitly seen by their realization on a scalar field

$$[J^i, \phi] = i(\vec{x} \wedge \vec{\nabla})^i \phi, \qquad [K^i, \phi] = -i(x^i \partial_t + t\nabla^i)\phi \, . \tag{5.18}$$

taking into account that under parity $t \to t$ and $\vec{x} \to -\vec{x}$.

In terms of $J^i_\pm$, eq. (5.17) reads

$$P \ : \ J^i_\pm \mapsto J^i_\mp \tag{5.19}$$

which implies that, under a parity transformation, Lorentz representations undergo a swap:

$$P \ : \ (j_-, j_+) \mapsto (j_+, j_-) \tag{5.20}$$

This means that if we want to represent parity, we need to double the number of fields! Let us make this argument more explicit. Choose for example the left-handed representation of the Lorentz group:

$$\Lambda_L(\vec{\theta}, \vec{\eta}) \tag{5.21}$$

Since a representation of parity changes the sign of the boost generators, we have:

$$P\Lambda_L(\vec{\theta}, \vec{\eta})P = \Lambda_L(\vec{\theta}, -\vec{\eta}) \tag{5.22}$$

But the last quantity is nothing but the right-handed representation, so that:

$$\boxed{P\Lambda_{L/R}(\vec{\theta}, \vec{\eta})P = \Lambda_{R/L}(\vec{\theta}, \vec{\eta})} \tag{5.23}$$

Therefore, in order to represent parity, we must consider the direct sum of representation

$$(1/2,0)\oplus(0,1/2) \quad : \quad \Lambda_D = \Lambda_L \oplus \Lambda_R \tag{5.24}$$

which defines a *Dirac spinor* field

$$\Psi_D(x) \equiv \begin{pmatrix} \psi_L(x) \\ \psi_R(x) \end{pmatrix} . \tag{5.25}$$

Notice that $\Psi_D(x)$ is a 4-dimensional complex spinor, consisting of two 2-dimensional complex spinors, ψ_L and ψ_R. As parity exchanges $(1/2,0)$ and $(0,1/2)$ it is realized on $\Psi_D(x)$ as

$$P \ : \ \Psi_D(x) = \begin{pmatrix} \psi_L(x) \\ \psi_R(x) \end{pmatrix} \ \mapsto \ \Psi_D^P(x) \equiv \begin{pmatrix} \psi_R(Px) \\ \psi_L(Px) \end{pmatrix} = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \begin{pmatrix} \psi_L(Px) \\ \psi_R(Px) \end{pmatrix} \equiv \gamma_0 \Psi_D(Px) \tag{5.26}$$

The matrix γ_0 swaps ψ_L and ψ_R and represents parity on the spinorial indices. Indeed one as $\gamma_0^2 = \mathbb{1}_{4\times 4}$ as befits parity (remember $P^2 = 1$). Moreover, as Lorentz transformations are represented on Ψ_D by the block diagonal matrices

$$\Lambda_D \equiv \begin{pmatrix} \Lambda_L(\theta,\eta) & 0 \\ 0 & \Lambda_R(\theta,\eta) \end{pmatrix} \tag{5.27}$$

according to eqs. (5.7,5.8) we have

$$\gamma_0 \Lambda_D(\vec{\theta},\vec{\eta}) = \Lambda_D(\vec{\theta},-\vec{\eta})\gamma_0 \, . \tag{5.28}$$

5.2.2 Complex Conjugation

Let us first define the 2×2 matrix

$$\epsilon \equiv i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} , \tag{5.29}$$

Notice that ϵ is just the two-dimensional *Levi-Civita* tensor. It satisfies the properties

$$\epsilon^T = \epsilon^{-1} = -\epsilon \qquad \epsilon^2 = -\mathbb{1}. \tag{5.30}$$

and acts on the Pauli matrices as

$$\epsilon^{-1}\sigma_i\epsilon = \epsilon\sigma_i\epsilon^{-1} = -\sigma_i^* = -\sigma_i^T \qquad\qquad \epsilon^{-1}\sigma_i^*\epsilon = \epsilon\sigma_i^*\epsilon^{-1} = -\sigma_i \, . \tag{5.31}$$

The above relations imply that $\sigma_i/2$ and $-\sigma_i^*/2$ are equivalent representations of the $SU(2)$ algebra. To see what this implies consider H a two dimensional representation of $SU(2)$ and its complex conjugate H^*. They transform respectively like

$$H \to e^{-i\frac{i}{2}\vec{\sigma}\cdot\vec{\theta}}H \qquad\qquad H^* \to e^{\frac{i}{2}\vec{\sigma}^*\cdot\vec{\theta}}H^* \tag{5.32}$$

which are at first sight inequivalent. However, by eq. (5.31) the above two tranformations are actually equivalent as made explicit by the fact that ϵH^* transforms like H

$$\epsilon H^* \to \epsilon e^{\frac{i}{2}\vec{\sigma}^*\cdot\vec{\theta}}H^* = e^{i\frac{i}{2}\epsilon\vec{\sigma}^*\cdot\epsilon^{-1}\vec{\theta}}\epsilon H^* = e^{-\frac{i}{2}\vec{\sigma}\cdot\vec{\theta}}\epsilon H^* \, . \tag{5.33}$$

Consider now instead the implications of 5.31 on the spinorial representation of $SO(3,1)$. We have

$$\Lambda_R^T = e^{\frac{1}{2}(\vec{\eta}-i\vec{\theta})\vec{\sigma}^T} = e^{-\frac{1}{2}(\vec{\eta}-i\vec{\theta})\epsilon^{-1}\vec{\sigma}\epsilon} = \epsilon^{-1}\Lambda_R^{-1}\epsilon = \epsilon^{-1}\Lambda_L^\dagger\epsilon \tag{5.34}$$

so that:

$$\boxed{\Lambda_R = \epsilon^{-1}\Lambda_L^*\epsilon} \tag{5.35}$$

Because of this we have:

$$\epsilon\psi_L^* \overset{\text{Lorentz}}{\mapsto} \epsilon\Lambda_L^*\psi_L^* = \epsilon\Lambda_L^*\epsilon^{-1}\epsilon\psi_L^* = \Lambda_R(\epsilon\psi_L^*) \tag{5.36}$$

which means that:

$$\boxed{\epsilon\psi_{L/R}^* \sim \psi_{R/L}} \tag{5.37}$$

In other words, from Eq. 5.35, we see that the complex conjugate of $(1/2, 0)$ is unitarily equivalent to $(0, 1/2)$ and similarly the complex conjugate of $(0, 1/2)$ is unitarily equivalent to $(1/2, 0)$

$$(1/2, 0) \sim (0, 1/2)^* \tag{5.38}$$

This result will be useful to derive Lorentz covariant equations of motions for the spinorial fields.

5.2.3 The $(\frac{1}{2}, \frac{1}{2})$ irrep, or the relation between $SL(2, \mathbb{C})$ and $\mathscr{L}_+^\uparrow$

In this section we shall study the $(\frac{1}{2}, \frac{1}{2})$ representation of $SO(3, 1) \equiv \mathscr{L}_+^\uparrow$. Our results will elucidate the relation between $SL(2, \mathbb{C})$ and $\mathscr{L}_+^\uparrow$, as well as set the basis for the construction of covariant "wave equations" in the next section.

An aside on notation[2]
The spinors in eq. (5.9) carry a spinor index that we have so far been omitting. As with Lorentz indices, when performing computations it is sometimes convenient to omit them and some other times it is convenient to display them explicitly, especially to construct covariant expressions. A standard notation [3] is to use *lower undotted* indices for the left-handed spinors and *upper dotted* indices for the right-handed ones

$$(\psi_L)_\alpha\,, \qquad (\psi_R)^{\dot\alpha} \tag{5.39}$$

with $\alpha = 1, 2$ and $\dot\alpha = \dot 1, \dot 2$. The Lorentz transformation matrices are then indicated by

$$(\Lambda_L)_\alpha{}^\beta\,, \qquad (\Lambda_R)^{\dot\alpha}{}_{\dot\beta} \tag{5.40}$$

so that the action of the Lorentz group reads

$$(\psi_L)_\alpha \to (\Lambda_L)_\alpha{}^\beta(\psi_L)_\beta, \qquad (\psi_R)_{\dot\alpha} \to (\Lambda_L)^{\dot\alpha}{}_{\dot\beta}(\psi_L)^{\dot\beta} \tag{5.41}$$

As a convention, we contract undotted indices from top-left to bottom-right $\searrow$ and dotted indices from bottom-left to top-right $\nearrow$.

[2] The content of this "aside" is not strictly necessary. Overall the very basic notion is that we use different indices (dotted and undotted) to indicate the components of respectively the right and the left spinors.

[3] See for instance the appendix of Phys. Rept. 128 (1985) 39-204. Other conventions are however in use.

The 2-dimensional Levi-Civita tensor is a 2×2 matrix, see eq. (5.29), acting in spinor space. We define

$$\epsilon_{\alpha\beta} = -\epsilon_{\beta\alpha} = \epsilon^{\alpha\beta}\,, \qquad \epsilon_{\dot\alpha\dot\beta} = -\epsilon_{\dot\beta\dot\alpha} = \epsilon^{\dot\alpha\dot\beta}\,, \qquad \epsilon_{12} = -\epsilon_{\dot1\dot2} = 1\,. \tag{5.42}$$

Notice that $\epsilon_{\alpha\beta}$ corresponds to the matrix ϵ in eq. (5.29), while $\epsilon_{\dot\alpha\dot\beta}$ corresponds to $\epsilon^T = \epsilon^{-1}$. Now, it is convenient to define the operation of raising an lowering indices by contracting with the Levi-Civita tensor in the the $\searrow$ and the $\nearrow$ convention for respectively left and right spinors

$$(\psi_L)^\alpha \equiv \epsilon^{\alpha\beta}\,(\psi_L)_\beta\,, \qquad (\psi_R)_{\dot\alpha} \equiv \epsilon_{\dot\alpha\dot\beta}\,(\psi_R)^{\dot\beta}\,. \tag{5.43}$$

Consistently, by eq. (5.42) and by $\epsilon^T = \epsilon^{-1}$, we have

$$(\psi_L)^\beta\,\epsilon_{\beta\alpha} = (\psi_L)_\alpha\,, \qquad (\psi_R)_{\dot\beta}\,\epsilon^{\dot\beta\dot\alpha} = (\psi_R)^{\dot\alpha} \tag{5.44}$$

and morever index raising applies consistently to the Levi-Civita tensor itself

$$\epsilon^{\gamma\alpha}\epsilon^{\delta\beta}\epsilon_{\alpha\beta} = \epsilon^{\gamma\delta} \qquad \epsilon_{\dot\alpha\dot\beta}\epsilon^{\dot\alpha\dot\gamma}\epsilon^{\dot\beta\dot\delta} = \epsilon^{\dot\gamma\dot\delta}\,. \tag{5.45}$$

The matrix relation $\epsilon^T\epsilon = \mathbb{1}$ implies for the mixed tensors $\epsilon_\alpha{}^\beta \equiv \epsilon^{\gamma\beta}\epsilon_{\gamma\alpha}$ and $\epsilon^\beta{}_\alpha \equiv \epsilon^{\beta\gamma}\epsilon_{\gamma\alpha}$

$$\epsilon_\alpha{}^\beta = -\epsilon^\beta{}_\alpha = \delta^\beta_\alpha \tag{5.46}$$

with δ^β_α the Kronecker delta. For the dotted tensor we instead have $\epsilon_{\dot\alpha}{}^{\dot\beta} \equiv \epsilon_{\dot\alpha\dot\gamma}\epsilon^{\dot\gamma\dot\beta}$ and $\epsilon^{\dot\beta}{}_{\dot\alpha} \equiv \epsilon_{\dot\gamma\dot\alpha}\epsilon^{\dot\gamma\dot\beta}$

$$\epsilon^{\dot\beta}{}_{\dot\alpha} = -\epsilon_{\dot\alpha}{}^{\dot\beta} = \delta^{\dot\beta}_{\dot\alpha} \tag{5.47}$$

And finally, by eqs. (5.35,5.36), the operation of complex conjugation turns dotted into undotted indices and viceversa. Using also that $(\epsilon_{\alpha\beta})^* = \epsilon_{\dot\beta\dot\alpha} = -\epsilon_{\dot\alpha\dot\beta}$ by eq. (5.42), we have that eq. (5.36) consistently reads

$$[(\psi_L)_\alpha]^* = (\psi_L^*)_{\dot\alpha}\,, \quad [(\psi_L)^\alpha]^* = (\psi_L^*)^{\dot\alpha}\,, \qquad [(\psi_R)^{\dot\alpha}]^* = (\psi_R^*)^\alpha\,, \quad [(\psi_R)_{\dot\alpha}]^* = (\psi_R^*)_\alpha \tag{5.48}$$

while eq. (5.35) reads instead:

$$(\Lambda_R)_{\dot\alpha\dot\beta} \equiv \epsilon_{\dot\alpha\dot\gamma}\,(\Lambda_R)^{\dot\gamma}{}_{\dot\beta} = \Big[(\Lambda_L)_\alpha{}^\delta\Big]^*\,\epsilon_{\dot\delta\dot\beta} = \Big[-\,(\Lambda_L)_\alpha{}^\delta\,\epsilon_{\delta\beta}\Big]^* \equiv -\Big[(\Lambda_L)_{\alpha\beta}\Big]^* \tag{5.49}$$

Notice that, when displaying the indices explicitly, it become unecessary to add the L and R labels: spinors with undotted indices are by definition left, while those with dotted ones are by definition right. A convenient and often used notation is then offered by the following identifications

$$(\psi_L)_\alpha \equiv \psi_\alpha\,, \quad (\psi_R)^{\dot\alpha} \equiv \bar\psi^{\dot\alpha}\,, \qquad (\Lambda_L)_\alpha{}^\beta \equiv \Lambda_\alpha{}^\beta\,, \quad (\Lambda_R)^{\dot\alpha}{}_{\dot\beta} \equiv \bar\Lambda^{\dot\alpha}{}_{\dot\beta} \tag{5.50}$$

by which complex conjugation simply reads

$$(\psi_\alpha)^* = \bar\psi_{\dot\alpha}\,, \quad (\bar\psi^{\dot\alpha})^* = \psi^\alpha\,, \qquad \bar\Lambda_{\dot\alpha\dot\beta} = -(\Lambda_{\alpha\beta})^* \tag{5.51}$$

$\star\quad\star\quad\star$

Consider now the $(1/2,1/2)$ representation. As $(1/2,1/2) = (1/2,0)\otimes(0,1/2)$ we can describe its generic element by a tensor with two spinorial indices $\hat v_\alpha{}^{\dot\beta}$ transforming under Lorentz as

$$\hat v_\alpha{}^{\dot\beta} \;\to\; (\Lambda_L)_\alpha{}^\gamma\,(\Lambda_R)^{\dot\beta}{}_{\dot\delta}\,\hat v_\gamma{}^{\dot\delta}\,. \tag{5.52}$$

precisely as the product of a left and a right spinor

$$\hat{v}_\alpha{}^{\dot\beta} \sim (\psi_L)_\alpha (\psi_R)^{\dot\beta} \,. \tag{5.53}$$

In matrix notation (i.e. without indicating the spinor indices) eq. (5.52) simply reads

$$\hat{v} \mapsto \Lambda_L \hat{v} \Lambda_R^T \tag{5.54}$$

It is possible, and useful, to rewrite the spinorial tensor $\hat{V}$ so as to make its Lorentz transformation involve only Λ_L or only Λ_R. For that purpose, considering the comment just below eq. (5.42) we can write

$$v_{\alpha\dot\beta} \equiv \epsilon_{\dot\beta\dot\gamma} \hat{v}_\alpha{}^{\dot\gamma} \;\to v = \hat{v}\epsilon \tag{5.55}$$

so that eq. (5.35) implies

$$v \;\mapsto\; \Lambda_L \hat{v} \Lambda_R^T \epsilon = \Lambda_L \hat{v} \epsilon \Lambda_L^\dagger = \Lambda_L v \Lambda_L^\dagger \tag{5.56}$$

Similarly, defining

$$\tilde{v}^{\alpha\dot\beta} \equiv \epsilon^{\alpha\gamma} \hat{v}_\gamma{}^{\dot\beta} \;\to \tilde{v} \equiv \epsilon \hat{v} \tag{5.57}$$

eq. (5.35) implies

$$\tilde{v} \;\mapsto\; \epsilon \Lambda_L \hat{v} \Lambda_R^T = \Lambda_R^* \epsilon \hat{v} \Lambda_R^T = \Lambda_R^* \tilde{v} \Lambda_R^T \tag{5.58}$$

This is good, but slightly better is to instead further define

$$\bar{v}^{\dot\beta\alpha} \equiv \tilde{v}^{\alpha\dot\beta} \;\to \bar{v} = \hat{v}^T \epsilon^T = \epsilon v^T \epsilon^T \tag{5.59}$$

so that

$$\bar{v} \;\mapsto\; \Lambda_R \bar{v} \Lambda_R^\dagger \tag{5.60}$$

Eqs. (5.56) and (5.60) offer different but equivalent ways to represent $SO(3,1)$ in the $(1/2, 1/2)$. Indeed V and $\bar{v}$ are simply related by a change of basis. Now, independent of which parametrization we use, $\hat{v}$, v or $\bar{v}$ belong to the vector space of complex 2×2 matrices $M(2,\mathbb{C}) \sim \mathbb{C}^4$. However eq. (5.56), and analogously (5.60), show that the subspaces of hermitian and anti-hermitean matrices are both invariant under the action of the Lorentz group. A generic complex V ($\bar{V}$) thus offers a reducible representation. As we are interested in the irreducible representations, because they constitue the smallest building blocks, we can assume v and $\bar{v}$ span the subspace of hermitean matrices [4]

$$v,\; \bar{v} \in MH(2,\mathbb{C}) \,. \tag{5.61}$$

Notice that $MH(2,\mathbb{C})$ is a real vector space equivalent to $\mathbb{R}^4$. The $(1/2, 1/2)$, is thus a *real representation*, while the $(0, 1/2)$ and $(1/2, 0)$ are genuinely *complex representations*. A basis for $MH(2,\mathbb{C})$ is given by the 4 hermitean matrices

$$\sigma^\mu = (\sigma^0, \sigma^i), \quad \sigma^0 = \mathbb{1}, \quad \sigma^i \;=\; \text{Pauli matrices} \tag{5.62}$$

[4] Anti-hermitean matrices offer an equivalent space as they are obtained from the hermitean ones by multiplying by i.

We can then expand V on this basis:

$$v = \eta_{\mu\nu} v^\mu \sigma^\nu = \begin{pmatrix} v^0 - v^3 & -v^1 + iv^2 \\ -v^1 - iv^2 & v^0 + v^3 \end{pmatrix} \tag{5.63}$$

On the other hand, for $\bar{v}$ eqs. 5.31, 5.59 give

$$\bar{v} = \eta_{\mu\nu} v^\mu \bar{\sigma}^\nu \tag{5.64}$$

where

$$\bar{\sigma}^\nu = (\sigma^0, -\sigma^i) \tag{5.65}$$

Notice that, for reasons that will become momentarily clear, we have been treating the basis matrices σ^μ and $\bar{\sigma}^\mu$ as a 4-vector. An important property of these matrices is (spinorial indices not displayed)

$$\boxed{\sigma_\mu \bar{\sigma}_\nu + \sigma_\nu \bar{\sigma}_\mu = 2\eta_{\mu\nu} = \bar{\sigma}_\mu \sigma_\nu + \bar{\sigma}_\nu \sigma_\mu} \tag{5.66}$$

When displaying spinorial indices explicitly, the above discussion implies

$$\sigma^\mu \ \to (\sigma^\mu)_{\alpha\dot{\beta}}\,, \qquad\qquad \bar{\sigma}^\mu \ \to (\bar{\sigma}^\mu)^{\dot{\beta}\alpha}\,, \tag{5.67}$$

so that σ^μ and $\bar{\sigma}^\mu$ act according to our contraction conventions on respectively the right and the left spinors

$$(\sigma^\mu \psi_R)_\alpha = (\sigma^\mu)_{\alpha\dot{\beta}} \psi_R^{\dot{\beta}}\,, \qquad\qquad (\bar{\sigma}^\mu \psi_L)^{\dot{\alpha}} = (\bar{\sigma}^\mu)^{\dot{\alpha}\beta} (\psi_L)_\beta \tag{5.68}$$

The next thing to work out is the explicit transformation properties of the v^μ. A clear hint of what is going on is gained by noticing that, by eq. (5.56) and by $\det \Lambda_L = 1$, the determinant of V is Lorentz invariant[5]

$$v' = \Lambda_L V \Lambda_L^\dagger \qquad \Rightarrow \quad \det v' = \det v |\det \Lambda_L|^2 = \det v \tag{5.69}$$

which concretely reads

$$v'^\mu v'_\mu = \det V' = \det V = (v^0)^2 - (v^1)^2 - (v^2)^2 - (v^3)^2 \equiv v^\mu v_\mu \tag{5.70}$$

compatibly with v^μ transforming as a 4-vector. One of the exercises confirms that is indeed the case. More precisely one finds that, given a generic Lorentz transformation $\Lambda(\omega) = \exp(-i\omega_{\mu\nu}\mathcal{J}^{\mu\nu}/2)$, one has

$$\Lambda_L(\omega) v^\mu \sigma_\mu \Lambda_L^\dagger(\omega) = \Lambda(\omega)^\mu{}_\nu v^\nu \sigma_\mu\,, \tag{5.71}$$

which establishes v^μ transforms like a 4-vector. Similarly one has

$$\Lambda_R(\omega) v^\mu \bar{\sigma}_\mu \Lambda_R^\dagger(\omega) = \Lambda(\omega)^\mu{}_\nu v^\nu \bar{\sigma}_\mu\,. \tag{5.72}$$

These results show an important property of the Lorentz group:

$$\forall\ U \in SL(2,\mathbb{C}) \quad \text{and} \quad \mathrm{X} = \mathrm{x}^\mu \sigma_\mu \tag{5.73}$$

The matrix X will transform as:

[5] Again, the same considerations can equivalently be made working in the $\bar{V}$ basis.

$$X \;\mapsto\; UXU^\dagger = X' = (\Lambda^\mu{}_\nu x^\nu)\sigma_\mu \tag{5.74}$$

This defines a mapping:

$$\Lambda \;:\; SL(2,\mathbb{C}) \;\to\; \mathscr{L}_+^\uparrow \tag{5.75}$$

which is a homomorphism:

$$\Lambda(U_1U_2) = \Lambda(U_1)\Lambda(U_2) \tag{5.76}$$

It is surjective and two-to-one:

$$\Lambda(U) = \Lambda(-U), \quad \forall\, U \in SL(2,\mathbb{C}) \tag{5.77}$$

In particular:

$$\Lambda(\mathbb{1}) = \Lambda(-\mathbb{1}) = \delta^\mu_\nu \tag{5.78}$$

$SL(2,\mathbb{C})$ is locally isomorphic to the Lorentz group, implying that they share the same Lie algebra, but the two groups are globally different. $SL(2,\mathbb{C})$ is connected and simply connected whereas $\mathscr{L}_+^\uparrow$ is connected but not simply connected. The above mapping describes the isomorphism:

$$\mathscr{L}_+^\uparrow \simeq SL(2,\mathbb{C})/\mathbb{Z}_2 \tag{5.79}$$

$SL(2,\mathbb{C})$ is the universal covering group of the proper orthochronous Lorentz group much like $SU(2)$ is the universal covering group of $SO(3)$. Spinors are those representations that transform under the bigger group $SL(2,\mathbb{C})$. As we already are familiar from ordinary quantum mechanics, spinors flip sign under a 2π rotation. This is why, in physical quantities, spinors enter twice: this phase is not observable. It is Nature (not us!) who chose to make full use of group theory to realise Lorentz.

5.3 Covariant wave equations

We are now ready to derive the relativistic dynamics involving spinorial fields. Consistently with our approach in this first part of the course, we are interested in the case of *free fields*, i.e. satisfying linear equations of motion. One obvious way to proceed is to write down the most general Poincarè invariant quadratic action. We will however find it more convenient to write down directly the covariant equations of motion. The general solutions of linear field equations consist of a superposition of plane waves, hence the terminology *wave equations*.

5.3.1 Weyl spinors

We have now all the ingredients to try and write down relativistic "wave equations" for spinors. Given $\psi_L(x)$, we want to generalise the scalar field equation (Klein-Gordon) and find a differential equation that transforms covariantly under Lorentz transformations. To do so, it is convenient to work in Fourier space:

$$\psi_L(p) = \int d^4x e^{ipx} \psi_L(x) \tag{5.80}$$

Under Lorentz, it transforms as:

$$\psi_L(p) \;\mapsto\; \Lambda_L \psi_L(\Lambda^{-1} p) \tag{5.81}$$

Consider the two following quantities in Fourier space:

$$P \equiv p_\mu \sigma^\mu = p \cdot \sigma, \qquad \bar{P} \equiv p_\mu \bar{\sigma}^\mu = p \cdot \bar{\sigma} \tag{5.82}$$

Their transformation properties under Lorentz are:

$$P' = p' \cdot \sigma = (\Lambda p) \cdot \sigma = \Lambda_L (p \cdot \sigma) \Lambda_L^\dagger = \Lambda_L (p \cdot \sigma) \Lambda_R^{-1} \tag{5.83}$$

$$\bar{P}' = p' \cdot \bar{\sigma} = (\Lambda p) \cdot \bar{\sigma} = \Lambda_R (p \cdot \bar{\sigma}) \Lambda_R^\dagger = \Lambda_R (p \cdot \bar{\sigma}) \Lambda_L^{-1} \tag{5.84}$$

We have therefore a good candidate for the wave equation in Fourier space. Indeed, the quantity:

$$\bar{P} \psi_L(p) \tag{5.85}$$

Transforms covariantly:

$$\bar{P}' \psi'_L(p') = \Lambda_R \left(\bar{P} \psi_L(p) \right) \tag{5.86}$$

Which means that if:

$$\bar{P} \psi_L(p) = 0 \tag{5.87}$$

in one frame, then the equality will hold in any frame! Notice that the wave equation transforms as a right-handed spinor. Indeed, the left-handed field belongs to $(1/2, 0)$ and $\bar{P}$ to $(1/2, 1/2)$. We then contract the unprimed indeces to form a Lorentz scalar and it only remains a free primed index:

$$\bar{P} \psi_L(p) \equiv \left(\bar{P} \right)^{A' A} (\psi_L)_A(p) = \left(\bar{P} \psi_L \right)^{A'} (p) \tag{5.88}$$

The equation of motion for the spinors is therefore:

$$\boxed{i \bar{\sigma}^\mu \partial_\mu \psi_L(x) = 0} \tag{5.89}$$

This is the so-called *Weyl equation* to which we associate *Weyl spinors*. Notice further that:

$$P \bar{P} \psi_L(p) \equiv p^2 \psi_L(p) = 0 \to -\Box \psi_L(x) = 0 \tag{5.90}$$

so that we expect Weyl spinors, upon quantization, to be associated to massless particles. In Fourier space:

$$\bar{P}\psi_L(p) = 0 \tag{5.91}$$

has a non trivial solution iff:

$$\det(\bar{P}) = 0 \Leftrightarrow p^2 = 0 \tag{5.92}$$

which leads to the same conclusion. The right-handed Weyl spinors satisfy a similar wave equation:

$$\boxed{i\sigma^\mu \partial_\mu \psi_R(x) = 0} \tag{5.93}$$

which, this time, transforms as a left-handed spinor.

Notice that the Weyl spinor wave equations, being homoneneous in either ψ_L or ψ_R are invariant under constant phase rotations, $\psi_L \to e^{i\alpha_L}\psi_L$ or $\psi_R \to e^{i\alpha_R}\psi_R$. The existence of this symmetry implies then, according to Noether's theorem, that the corresponding particles would carry a conserved charge. We conclude that the particles associated to Weyl spinors have vanishing mass but carry a conserved charge.

5.3.2 Majorana spinors

We would like to describe something with mass and to do so we need to "add" something to the RHS of Eq. 5.89. Since the LHS transforms as a right-handed spinor, so must do the RHS. We saw earlier that we can get a right-handed spinor from a left-handed one by a parity transformation:

$$\epsilon\psi_L^*(x) \in (0, 1/2) \tag{5.94}$$

we therefore write:

$$\boxed{i\bar{\sigma}^\mu \partial_\mu \psi_L(x) = c\epsilon\psi_L^*(x) \overset{\text{Fourier}}{\to} \bar{P}\psi_L(p) = c\epsilon\psi_L^*(-p)} \tag{5.95}$$

From Eq. 5.31, we have:

$$\epsilon^{-1}\bar{\sigma}_\mu^*\epsilon = \sigma_\mu \tag{5.96}$$

This leads to the conjugate equation:

$$i\sigma^\mu p_\mu \epsilon\psi_L^*(x) = c^*\psi_L(x) \tag{5.97}$$

Finally, this implies:

$$P\bar{P}\psi_L(p) = p^2\psi_L(p) = |c|^2\psi_L(p) \tag{5.98}$$

Or in position space:

$$(\Box + |c|^2)\psi_L(x) = 0 \tag{5.99}$$

This second wave equation corresponds to a particles of mass $|c|^2 = m^2$ and is the so-called *Majorana equation* describing *Majorana spinors*.

Neutrinos likely are, as far as we understand, Majorana particles. Notice that the Majorana spinor equation is not invariant under phase rotations:

$$\psi_L(x) \;\mapsto\; e^{i\alpha}\psi_L(x) \tag{5.100}$$

Thus, there is no conserved charge (electric or of any other type) associated to Majorana particles. Majorana particles have mass but do not carry a conserved charge.

5.3.3 Dirac spinors

As we have seen, Weyl spinors are expected to describe massless charged particles, while Majorana spinors describe massive neutral particles. It thus remains to be seen if we can arrange things in such a way that spinors describe particles that carry both nonzero mass and nonzero charge. With just a single ψ_L or a single ψ_R, as we have seen, there are only two options, Weyl and Majorana. In order to go beyond we must then consider a situation where we have both one ψ_L and one independent ψ_R. With these two independent fields at hand we can imagine a set of two covariant paired wave equations of the form

$$\begin{cases} i\bar{\sigma}\cdot\partial\psi_L(x) &= c\psi_R(x) \\ i\sigma\cdot\partial\psi_R(x) &= \tilde{c}\psi_L(x) \end{cases} \tag{5.101}$$

By rescaling the field:

$$\psi_R \;\to\; \lambda\psi_R, \quad \lambda \in \mathbb{C} \tag{5.102}$$

such that:

$$c\lambda = \frac{\tilde{c}}{\lambda} = m, \quad m \in \mathbb{R} \tag{5.103}$$

The equations become:

$$\begin{cases} i\bar{\sigma}\cdot\partial\psi_L(x) &= m\psi_R(x) \\ i\sigma\cdot\partial\psi_R(x) &= m\psi_L(x) \end{cases} \tag{5.104}$$

This set of equations is clearly endowed with a global $U(1)$ symmetry:

$$\begin{cases} \psi_L(x) &\mapsto e^{i\alpha}\psi_L(x) \\ \psi_R(x) &\mapsto e^{i\alpha}\psi_R(x) \end{cases} \tag{5.105}$$

This tells us that there is a conserved charge associated with the quanta of these fields. We will see later that when we couple this set of equations to electromagnetism, this leads to the conservation of the electric charge.

The paired wave equations in eq. (5.104), when iterated, are quickly seen to imply satisfaction of the massive Klein-Gordon equation

$$-\Box\psi_L = (i\sigma\cdot\partial)(i\bar{\sigma}\cdot\partial)\psi_L = m(i\sigma\cdot\partial)\psi_R = m^2\psi_L \tag{5.106}$$

and similarly for ψ_R.

By the above remarks, we thus expect a pair ψ_L, ψ_R satisfying eq. (5.104) to give rise to massive charged particles upon quantization. A pair of fermions satisfying eq. (5.104) constitutes a *Dirac spinor* and the system of two coupled wave equations is known as the *Dirac equation*. As we shall see, the quanta not only carry mass and charge but also an intrinsic 1/2 spin. The relevance of the Dirac spinor stems from the fact that it describes the fundamental constituents of matter, electrons and protons.

As a side remark we also notice that eq. (5.104) is invariant under parity transformations:

$$\begin{cases} \psi_L(t,\vec{x}) & \mapsto \psi_R(t,-\vec{x}) \\ \psi_R(t,\vec{x}) & \mapsto \psi_L(t,-\vec{x}) \end{cases} \tag{5.107}$$

The systems described by this set of equations is suitable to describe representations of parity. We saw earlier that in terms of spinors, it is the *Dirac spinor* which carries a representation of that symmetry and indeed this set of equations can be rewritten in a more compact way. Recall:

$$\Psi_D \equiv \psi_L \oplus \psi_R \tag{5.108}$$

Now define a set of matrices, called the *Dirac matrices* or *gamma matrices*:

$$\gamma^\mu \equiv \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix} \tag{5.109}$$

They satisfy the *Dirac algebra*:

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \tag{5.110}$$

The set of equations can then be rewritten in a compact form:

$$\boxed{(i\partial\!\!\!/ - m)\Psi_D(x) = 0} \tag{5.111}$$

where we used the Feynman slash notation $\partial\!\!\!/ \equiv \gamma^\mu\partial_\mu$. This is the *Dirac equation* for four-components spinors. It describes charged particles with spin 1/2 (*e.g.* e^-, μ^-, p, n, ...).

5.3.4 Lagrangians

We would like to have a Lagrangian formulation of the dynamics of the spinor fields. In order to do so, we must learn how to construct local expressions of the spinors and, in principle, of their derivatives that transforms like Lorentz scalars. Let us then consider a general change of Lorentz frame $x'^\mu = \Lambda^\mu{}_\nu x^\nu$ under which spinors transform according to (see eq. (5.10))

$$\psi'_L(x') = \Lambda_L(\omega)\psi_L(x)\,, \qquad \psi'_R(x') = \Lambda_R(\omega)\psi_R(x)\,. \tag{5.112}$$

where $\Lambda_{L,R}$ represent Λ. As we already did in the section on wave equations, putting together the results of section 3.2.3 and eqs. (5.71, 5.72), we can express the transformation of spacetime derivatives as

$$\partial' \cdot \sigma \equiv \partial'_\mu \sigma^\mu = \Lambda_\mu{}^\nu \partial_\nu \sigma^\mu = \Lambda_L (\partial_\mu \sigma^\mu) \Lambda_L^\dagger \equiv \Lambda_L \partial \cdot \sigma \Lambda_L^\dagger \tag{5.113}$$

$$\partial' \cdot \bar{\sigma} \equiv \partial'_\mu \bar{\sigma}^\mu = \Lambda_\mu{}^\nu \partial_\nu \bar{\sigma}^\mu = \Lambda_R (\partial_\mu \bar{\sigma}^\mu) \Lambda_R^\dagger \equiv \Lambda_R \partial \cdot \bar{\sigma} \Lambda_R^\dagger \,. \tag{5.114}$$

Taking into account that $\Lambda_L^\dagger \Lambda_R = \Lambda_R^\dagger \Lambda_L = \mathbb{1}$, by eq. (5.112), and the above two equations, one can then write the following set of Lorentz scalar bilinears in the spinors

$$S_1(x) \equiv \psi_L(x)^\dagger \psi_R(x), \qquad S_1^\dagger(x) \equiv \psi_R(x)^\dagger \psi_L(x), \tag{5.115}$$

$$S_2(x) \equiv \psi_L(x) \partial \cdot \bar{\sigma} \psi_L(x), \qquad S_3(x) \equiv \psi_R(x) \partial \cdot \sigma \psi_R(x). \tag{5.116}$$

In other words, and in an obvious notation, one has $S'_i(x') = S_i(x)$ for $i = 1, 2, 3$. One is easily convinced, given two independent spinors ψ_L and ψ_R, that there are no other quadratic invariant other than the four above. Notice, in particular, that $\psi_L^\dagger \psi_L$ is <u>not</u> an invariant, and that simply because $\Lambda_L^\dagger \Lambda_L \neq \mathbb{1}$.

We have now gathered all the necessary ingredients to build the Lagrangians. Let us start with the Weyl Lagrangian. Simply:

$$S_W = \int d^4x \mathcal{L}_W = \int d^4x \; i\psi_L^\dagger(x) \bar{\sigma} \cdot \partial \psi_L(x) \tag{5.117}$$

Notice that the action is real up to boundary terms. The equations of motion lead as expected to the Weyl equation:

$$0 = \frac{\delta S_W}{\delta \psi_L^\dagger} = i\bar{\sigma} \cdot \partial \psi_L(x) \tag{5.118}$$

The next to simplest case is the Dirac Lagrangian. For the mass terms we simply add the first scalars we built:

$$S_D = \int d^4x \mathcal{L}_D = \int d^4x \left(i\psi_L^\dagger(x) \bar{\sigma} \cdot \partial \psi_L(x) + i\psi_R^\dagger(x) \sigma \cdot \partial \psi_R(x) - m\psi_L^\dagger(x) \psi_R(x) - m^* \psi_R^\dagger(x) \psi_L(x) \right) \tag{5.119}$$

Do we have a complex mass? No, we can simply redefine:

$$\psi_R \;\mapsto\; e^{-i\theta} \psi_R, \quad m = |m| e^{i\theta} \tag{5.120}$$

This is a so-called *chiral rotation*. To build the Majorana Lagrangian we use:

$$\psi_R \;\to\; \epsilon \psi_L^* \tag{5.121}$$

in the Dirac Lagrangian. We then write:

$$S_M = \int d^4x \mathcal{L}_M = \int d^4x \left(i\psi_L^\dagger(x) \bar{\sigma} \cdot \partial \psi_L(x) - \frac{m}{2} \psi_L^T \epsilon \psi_L - \frac{m^*}{2} \psi_L^\dagger \epsilon \psi_L^* \right) \tag{5.122}$$

Here comes the tricky part:

$$\psi_L^T \epsilon \psi_L = (\psi_L)_A \epsilon^{AB} (\psi_L)_B = -(\psi_L)_A \epsilon^{BA} (\psi_L)_B = -(\psi_L)_B \epsilon^{BA} (\psi_L)_A = -\psi_L^T \epsilon \psi_L \tag{5.123}$$

Where we assumed in the last equality that spinors *commute*, just as classical variables. However, this implies that $\psi_L^T \epsilon \psi_L = 0$ so that the Majorana mass would not exist! The solution of this paradox relies on the fact that spinors do not have a classical limit! Spinors are Grassman variables that *anti-commute*:

$$\{(\psi_L)_A, (\psi_L)_B\} = 0 \tag{5.124}$$

Notice that we did not need to use this fact in the construction of the Weyl and Dirac Lagrangians.

Finally, the Dirac Lagrangian can be rewritten in terms of a Dirac spinor. Define the *Dirac conjugate* of a Dirac field Ψ_D:

$$\bar{\Psi}_D \equiv \Psi_D^\dagger \gamma^0 = \psi_R^\dagger \oplus \psi_L^\dagger \tag{5.125}$$

Then, the Dirac Lagrangian can be rewritten:

$$S_D = \int d^4x \mathcal{L}_D = \int d^4x \ \bar{\Psi}_D (i\partial\!\!\!/ - m)\Psi_D, \qquad 0 = \frac{\delta S_D}{\delta \bar{\Psi}_D} = (i\partial\!\!\!/ - m)\Psi_D \tag{5.126}$$

Very much like we did with the charged scalar field, one can prove that the Dirac Lagrangian is the most general Lagrangian that is Poincaré invariant, at most quadratic in ψ_L and ψ_R, with up to one-derivative terms and invariant under a global "charge" $U(1)$ symmetry.

5.4 Towards quantization

5.4.1 Classical solutions

The dirac equation is $(-i\not{\partial} - m)\psi = 0$. We're going to look for the most general plane wave solution $\psi_p(x) = e^{-ipx}u(p)$ with $p = (p^0, \vec{p})$

$$(i\not{\partial} - m)\psi_p = e^{-ipx}\left(\gamma^\mu p_\mu - m\right)u(p) = 0 \qquad \text{from now on } :\not{p} = \gamma^\mu p_\mu \tag{5.127}$$

In order to neglect trivial solutions, we require a null determinant. Note that $\not{p}\not{p} = \frac{\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu}{2}p_\mu p_\nu = p^2$

$$\left(\not{p} + m\right)\left(\not{p} - m\right)u(p) = \left(p^2 - m^2\right)u(p) = 0 \quad \rightarrow \quad p^2 \stackrel{!}{=} m^2 \tag{5.128}$$

So we get positive and negative energy solutions where $p_0 = \sqrt{m^2 + p^2}$ with $\psi_p = e^{-ipx}u(p)$ and $\psi_p e^{ipx}v(p)$. I.e. the usual red hyperboloid 8.2 and the symmetric one. In particular one can write the ψ_p as

$$\psi(p) = \delta(p^2 - m^2)\tilde{\psi}(p) = \tag{5.129}$$
$$= \theta(p_0)\delta(p^2 - m^2)u(p) + \theta(-p_0)\delta(p^2 - m^2)v(-p) = \tag{5.130}$$
$$= \delta(p_0 - \sqrt{m^2 + \vec{p}^2})u(\vec{p}) + \delta(p_0 - \sqrt{m^2 + \vec{p}^2})v(-\vec{p}). \tag{5.131}$$

Where v and u are the solutions above. Then adding back time evolution $e^{-ip_0 t}$ we get

$$\psi(x) = \int d\omega_p e^{-ip\cdot x}u(p) + e^{ip\cdot x}v(p), \tag{5.132}$$

where the momenta are 4-momenta.

To find explicitly u and v, we first study the reference momentum $\bar{p}$: defined as $\vec{p} = 0$ which implies $p^0 = m$. which with little algebra implies

$$m\left(\gamma^0 - 1\right)u(p) = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} \xi_L \\ \xi_R \end{pmatrix} = 0 \quad \rightarrow \quad u = \begin{pmatrix} \xi \\ \xi \end{pmatrix} \text{ with } \xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \tag{5.133}$$

Where the 1's in the matrix are two by two identities. One can look at Peskin for more details. Let's write ξ_1 and ξ_2 as basis vectors (of which they were the indices in the above definition of ξ). We define $\xi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\xi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ which will correspond to eigenvalues of σ_3: $\sigma_3\xi_1 = \frac{1}{2}$ and $\sigma_3\xi_2 = -\frac{1}{2}$.

To get the most general solution, we can notice that the fields at a given momenta can be obtained by boosting the ones at the reference momentum. Notice in particular that the Λ_D acting on Dirac spinors are a representation of the Lorentz group, therefore

$$\Lambda_D p^\mu \gamma_m u \lambda_D^{-1} = (\Lambda^\mu_\nu)p^\nu. \tag{5.134}$$

Consider the wave equation for $u(\bar{p})$ and multiply times Λ_D:

$$\Lambda_D(\bar{\not{p}} - m)\Lambda_D^{-1}\Lambda_D u(\bar{p}) = \tag{5.135}$$
$$(\not{p} - m)u(p) = 0. \tag{5.136}$$

where we defined

$$u(p) = \Lambda_D(\bar{p}). \tag{5.137}$$

We know that to get any momentum starting from $\bar{p}$ we can use

$$\Lambda = e^{i\vec{\eta}_p \cdot \vec{\mathbb{K}}}, \tag{5.138}$$

where $\vec{\mathbb{K}}$ is the vector of the generators of boosts and

$$\vec{\eta}_p = \left(\tanh^{-1}\left(\frac{|\vec{p}|}{p_0}\right)\right)\frac{\vec{p}}{|\vec{p}|}. \tag{5.139}$$

It can be shown (exercises) that most general solution is given by $u(p) = \begin{pmatrix}\sqrt{\sigma p}\xi \\ \sqrt{\bar{\sigma} p}\xi\end{pmatrix}$ where $\sqrt{\sigma p}$ is defined as $\sqrt{\sigma^\mu p_\mu} = \frac{1}{2}\sqrt{\sqrt{E+m} - \frac{\vec{p}\vec{\sigma}}{\sqrt{E+m}}}$

So let's check it is at least a solution

$$(\not{p} - m)\, u_p = \begin{pmatrix}-m & p\sigma \\ p\bar{\sigma} & -m\end{pmatrix}\begin{pmatrix}\sqrt{\sigma p}\xi \\ \sqrt{\bar{\sigma} p}\xi\end{pmatrix} = 0 \tag{5.140}$$

Generically the $\psi(x)$ will be a sum of the different solutions u_i and v_i

$$u_i(p) = \sum_{i=1}^{2} a_i(\vec{p}) u_i(\vec{p}), \tag{5.141}$$

we have

$$\psi(x) = \int d^3p \sum_s \left(a_s(p)u_s(p)e^{-ipx} + b_s(p)v_s(p)e^{ipx}\right) \tag{5.142}$$

And the coefficients a and b will turn out to be the analogous of creation and annihilation operator for the Klein-Gordon field.

5.4.2 Chirality and helicity

We saw that the Dirac field is putting together a left and right field which are two irreducible representations

$$\psi_{\mathrm{D}} = \begin{pmatrix}\psi_L \\ \psi_R\end{pmatrix} \tag{5.143}$$

. We define $\gamma^5 = \begin{pmatrix}-1 & 0 \\ 0 & 1\end{pmatrix}$ so that $\gamma_5\psi_D = \begin{pmatrix}-\psi_L \\ \psi_R\end{pmatrix}$. We define the left and right projectors

$$P_L = \frac{1-\gamma_5}{2} = \begin{pmatrix}1 & 0 \\ 0 & 0\end{pmatrix} \qquad P_R = \frac{1+\gamma_5}{2} = \begin{pmatrix}0 & 0 \\ 0 & 1\end{pmatrix} \tag{5.144}$$

This will be useful in the concept of chirality, where eigenstates of the projector will be 'left' or 'right'.

Now let's define the concept of helicity. Helicity is the projection of the spin on the momentum $\vec{s}\cdot\frac{\vec{p}}{|\vec{p}|}$.

$$\hat{h} = \frac{\vec{p}\cdot\vec{\sigma}}{2|\vec{p}|} \tag{5.145}$$

with the helicity operator $h = \begin{pmatrix}\hat{h} & 0 \\ 0 & \hat{h}\end{pmatrix}$

One can show that in the case $m \to 0$, $u(p) = \sqrt{\frac{E}{2}} \begin{pmatrix} (1-2\hat{h})\xi \\ (1+2\hat{h})\xi \end{pmatrix}$. What does the helicity do when acting on the solution ?

$$\begin{pmatrix} \hat{h} & 0 \\ 0 & \hat{h} \end{pmatrix} u(p) = \sqrt{\frac{E}{2}} \begin{pmatrix} (\hat{h}-\frac{1}{2})\xi \\ (\hat{h}+\frac{1}{2})\xi \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} u_p \tag{5.146}$$

So if we have a left handed spinor (upper component) it will be reversed under helicity, whereas the right handed spinor will not. Do solutions with definite chirality or helicity exist in the massless limit? If you choose ξ such that $\xi(1+2\hat{h}) = 0$, then $u(p) = \sqrt{\frac{E}{2}} \begin{pmatrix} \xi \\ 0 \end{pmatrix}$.

For example, $\vec{p} = (0,0,p)$, $\hat{h} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$, then we have, defining $\xi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\xi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$$u(p) = \sqrt{\frac{E}{2}} \begin{pmatrix} \xi_- \\ 0 \end{pmatrix} \text{ Chirality = -1, helicity = - } \frac{1}{2} \qquad u(p) = \sqrt{\frac{E}{2}} \begin{pmatrix} 0 \\ \xi_+ \end{pmatrix} \text{ Chirality = 1, helicity = } \frac{1}{2} \tag{5.147}$$

In the massless limit, chirality and helicity are synonyms and are lorentz invariant. In the massive case, helicity is not lorentz invariant since if we overtake a particle in a boosted frame, it will look like it is going backwards with respect to us, but its spin will still be in the same direction, so helicity is not lorentz invariant.

The most general positive energy solution is

$$\psi(x) = e^{-ipx} \begin{pmatrix} \sqrt{\sigma p}\xi_s \\ \sqrt{\bar{\sigma} p}\xi_s \end{pmatrix} \quad \text{s = 1,2} \tag{5.148}$$

They have some orthogonality relations

$$\bar{u}^s(p)u^r(p) = u^s(p)^\dagger \gamma_0 u^r(p) = (\xi_s^\dagger \sqrt{\sigma p}, \xi_s^\dagger \sqrt{\bar{\sigma} p}) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{\sigma p}\xi_s \\ \sqrt{\bar{\sigma} p}\xi_s \end{pmatrix} = 2m\xi_s^\dagger \xi_r = 2m\delta_{rs} \tag{5.149}$$

Moreover,

$$\bar{u}^s(p)\gamma^\mu u^r(p) = 2p^\mu \delta_{rs} \tag{5.150}$$

Negative energy solutions are

$$\psi(x) = e^{ipx} v_s(p) = e^{ipx} \begin{pmatrix} \sqrt{\sigma p}\xi_s \\ -\sqrt{\bar{\sigma} p}\xi_s \end{pmatrix} \tag{5.151}$$

Where

$$\bar{v}^s(p)v^r(p) = -2m\delta_{rs} \tag{5.152}$$

$$\bar{v}^s(p)\gamma^\mu v^r(p) = 2p^\mu \delta_{rs} \tag{5.153}$$

Orthogonality

$$\bar{u}^r(p)v^s(p) = \left(\xi^{r\dagger}\sqrt{p\sigma}, \xi^{r\dagger}\sqrt{p\bar{\sigma}}\right) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{p\sigma}\xi_s \\ -\sqrt{p\bar{\sigma}}\xi_s \end{pmatrix} = 0 \tag{5.154}$$

In other words, $u_1(p), u_2(p), v_1(p), v_2(p)$ form a basis of $\mathbb{C}^4$

spin sums : Completeness relation

$$A^{\alpha\beta} = \sum_s u_s^\alpha(p)\bar{u}_s^\beta(p) = \sum_s \begin{pmatrix} \sqrt{p\sigma}\xi_s \\ \sqrt{p\bar{\sigma}}\xi_s \end{pmatrix} \begin{pmatrix} \xi_s^\dagger\sqrt{p\sigma} & \xi_s^\dagger\sqrt{p\bar{\sigma}} \end{pmatrix} \tag{5.155}$$

So

$$A_{\alpha\beta} = \begin{pmatrix} m & p\sigma \\ p\bar{\sigma} & m \end{pmatrix}_{\alpha\beta} = (\not{p} + m)_{\alpha\beta} \tag{5.156}$$

So all in all, the most general ψ is

$$\psi_{\text{general}}(x) = \int d^3p \sum_s \left(a_s(p)u_s(p)e^{-ipx} + b_s(p)v_s(p)e^{ipx}\right) \tag{5.157}$$

Reminiscent of what we did in KG Fields, we'll quantize the fourrier coefficients to warp the theory to the quantum world.

5.5 Quantization of the Dirac field

5.5.1 How not to quantize the Dirac field

We'll do exactly what we did with the Klein-Gordon field.

$$\mathcal{L} = \frac{i}{2}\left(\bar{\psi}\not{\partial}\psi - \partial_\mu\bar{\psi}\gamma^\mu\psi\right) - m\bar{\psi}\psi \tag{5.158}$$

$$T^\mu_\nu = \frac{\partial\mathcal{L}}{\partial\partial_\mu\psi}\partial_\nu\psi + \frac{\partial\mathcal{L}}{\partial\partial_\mu\bar{\psi}}\partial_\mu\bar{\psi} - \delta^\mu_\nu\mathcal{L} \tag{5.159}$$

$$H = \int\mathcal{H} = \int T^0_0 = \int \frac{1}{2}\bar{\psi}\left(-i\vec{\gamma}\vec{\nabla} + m\right)\psi + \frac{1}{2}\bar{\psi}\left(i\overleftarrow{\nabla}\gamma + m\right)\psi d^3x \tag{5.160}$$

$$= \int \bar{\psi}\left(-i\vec{\gamma}\vec{\nabla} + m\right)\psi d^3x \underbrace{\equiv}_{h_D\equiv\gamma_0\left(-i\vec{\gamma}\vec{\nabla}+m\right)} \int \psi^\dagger h_D\psi d^3x \tag{5.161}$$

$$P_i = \int T^0_i d^3x = \frac{1}{2}i2\int \bar{\psi}\gamma^0\partial_i\psi - \partial_i\bar{\psi}\gamma_0\psi = i\int\psi^\dagger\partial_i\psi d^3x \tag{5.162}$$

We recognize the 3-momentum from quantum mechanics, however with a sign flip. Our ordinary 3-Momentum is P^i not P_i.

Focussing on the Hamiltonian H, notice that on configurations that satisfy the equations motion (i.e. the Dirac equation) we have

$$H = \mathcal{H} = \int\psi^\dagger h_D\psi d^3x \underbrace{=}_{\text{using eq.o.m}} \int\psi^T i\partial_0\psi d^3x \tag{5.163}$$

So both expressions look quite similar and we're happy. One has the 3 components (3-momentum) and the other one has the one that is left, the 0-th component. Now let's quantize the whole thing. There's a problem with the fact that the complex conjugated field is troublesome for writing the canonical conjugate variable. In particular

$$\Pi_\alpha \equiv \frac{\partial \mathcal{L}}{\partial \dot{\psi}_\alpha} = \frac{i}{2}\psi^\dagger_\alpha \quad \Pi^\dagger_\alpha = \frac{i}{2}\psi^\dagger. \tag{5.164}$$

I.e. ψ and $\psi^\dagger$ are conjugate variables. The procedure to correctly work with systems that exhibits such features is the Dirac bracket procedure. However, we can avoid the problem by taking another Lagrangian which should be equivalent.

$$\mathcal{L} = i\bar{\psi}\partial\psi - m\bar{\psi}\psi \tag{5.165}$$

The canonical variable conjugated to ψ

$$(\pi_\psi)^\alpha = \frac{\partial \mathcal{L}}{\partial \dot{\psi}^\alpha} = i\left(\bar{\psi}\gamma^0\right)^\alpha = i\psi^{\dagger\alpha} \tag{5.166}$$

So $\left(\psi_\alpha, \pi_\alpha \equiv i\psi^\dagger_\alpha\right)$ are the canonically conjugated momenta. We try to impose commutation relations as for the Klein-Gordon field:

$$[\psi_\alpha(\vec{x}), \psi_\beta(\vec{y})] = 0 \qquad [\psi_\alpha(\vec{x}), \pi_\beta(\vec{x})] = i\delta_{\alpha\beta}\delta^3(x-y) \qquad \left[\psi_\alpha(\vec{x}), \psi^\dagger_\beta(\vec{y})\right] = \delta_{\alpha\beta}\delta^3(\vec{x}-\vec{y}) \tag{5.167}$$

Exercise: Show $\dot{\psi}(x) \equiv i\,[H, \psi(x)]$.

Moving on :

$$H = \int d^3x \left(\psi^\dagger\left(-i\vec{\gamma}\vec{\nabla} + m\right)\psi\right) \tag{5.168}$$

Let's now expand ψ into the basis of eigenfunctions to the hamiltonian so that it will trivially be diagonal. What is the eigenbasis of h_D ? We had already found some stuff :

$$\text{Positive energy :} \quad u_s(p)e^{-ipx} \qquad \text{Negative energy} \quad v_s(p)e^{ipx} \tag{5.169}$$

In particular for the positive energy solution

$$\left(i\gamma^0\partial_0 + i\vec{\gamma}\vec{\nabla} - m\right)u_s(p)e^{-ipx} = 0 \rightarrow -\left(i\vec{\gamma}\vec{\nabla} - m\right)u_s(p)e^{ipx} = E(p)u_s(p)e^{ipx} \qquad \text{t=0} \tag{5.170}$$

For negative ones

$$\left(i\gamma^0\partial_0 + i\vec{\gamma}\vec{\nabla} - m\right)v_s(p)e^{ipx} = 0 \rightarrow -\left(i\vec{\gamma}\vec{\nabla} - m\right)v_s(-p)e^{ipx} = -E(p)u_s(-p)e^{ipx} \qquad \text{t=0} \tag{5.171}$$

So we have 4 linearly independent eigenvectors $(v_1(-p), v_2(-p), u_1(p), u_2(p))$ when fixing p (so the hamiltonian is a 4 by 4 matrix). By varying p, one spans the whole basis. The most general $\psi(x)$ is

$$\psi(x) = \int \frac{d^3p}{(2\pi)^3 2E(p)} e^{ipx}\left(\sum_{s=1}^{2} Q^s_p u_s(p) + b^s_{-p}v_s(-p)\right) \tag{5.172}$$

Remember for the scalar field we had

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3 2E(k)} e^{ipx} \underbrace{\left(a_p + a_p^\dagger\right)}_{\phi(p)} \tag{5.173}$$

Our choice of v_s and u_s is very useful since as we will check, it diagonalizes the hamiltonian.

$$\bar{u}^r(p)\gamma^0 \int \psi(x) e^{-ipx} d^3x = \frac{1}{2E(p)} \bar{u}_r(p)\gamma_0 \sum_s \left(a_s u_s(p) + b_s(-p)v_s(-p)\right) = \sum_s \delta_{rs} a_s(p) = a_r(p) \tag{5.174}$$

$$\bar{v}^r(p)\gamma^0 \int \psi(x) e^{-ipx} d^3x = b_r(-p) \tag{5.175}$$

So the ψ's is a linear combination of a's and b's. So the commutation relations on ψ imply commutation relations on a and b.

$$\left[a_r(p), a_s(k)^\dagger\right] = (2\pi)^3 2E(p)\delta_{rs}\delta^3(k-p) \qquad \left[b_r(p), b_s^\dagger(k)\right] = (2\pi)^3 2E(p)\delta_{rs}\delta^3(k-p) \tag{5.176}$$

Coming back to our sheep

$$H = \int d^3\psi^\dagger h_D \psi = \int d\Omega_k d\Omega_p \sum_s \sum_r \left(\bar{u}_s(p)a_s^\dagger(p) + \bar{v}_s(-p)b_s^\dagger(-p)\right)\left(E(k)u_r(k)a_r(k) - E(k)v_r(k)b_r(-k)\right) e^{i(k-p)x} d^3x$$
$$\underbrace{=}_{\text{Trivial manipulation}} \int d\Omega_p E_P \sum_s a_s^\dagger(p)a_s(p) - b_s^\dagger(p)b_s(p) \tag{5.177}$$

One can check $\left[H, a_s^\dagger(p)\right] = E(p)a_s^\dagger(p)$ and $[H, a_s(p)] = -E(p)a_s(p)$ while for the b's it's the other way around: $\left[H, b_s^\dagger(p)\right] = -E(p)b_s^\dagger(p)$ and $[H, b_s(p)] = E(p)b_s(p)$.

The daggered a is an energy raising operator but the daggered b is an energy lowering operator ! So the specturm of the Hamiltonian is not bound (if one defines the vacuum as $a_s(p)|0\rangle = 0$ and $b_s(p)|0\rangle = 0$). One can not fix this by redefining $b^\dagger$ to b and conversely. This is because of their commutation relations. One would get $bb^\dagger - b^\dagger b \underbrace{=}_{\text{sort of}} 1 > 0$. If we would switch both definitions, one would get negative probability norms.

$$bb^\dagger - b^\dagger b = 1 \tag{5.178}$$

Attempt to define $b^\dagger|0\rangle = 0$ and $b|0\rangle \neq 0$

$$\langle 0|bb^\dagger - b^\dagger b|0\rangle = \langle 0|0\rangle = -\langle 0|b^\dagger b|0\rangle = -||b|0\rangle||^2 \tag{5.179}$$

So either the vacuum or the state where we created a particle has negative norm. We can't do it. The question is : is the way to quantize the field unique w.r.t to the commutation relations one imposes ? No ! There's also the possibility of imposing anti-commutation relations !

5.5.2 The correct approach to quantization

So we impose the following commutation relations

$$\{\psi_\alpha(\vec{x}), \psi_\beta(\vec{y})\} = 0 \qquad \{\psi_\alpha(\vec{x}), \pi_\beta(\vec{x})\} = i\delta_{\alpha\beta}\delta^3(x-y) \qquad \left\{\psi_\alpha(\vec{x}), \psi_\beta^\dagger(\vec{y})\right\} = \delta_{\alpha\beta}\delta^3(\vec{x}-\vec{y}) \tag{5.180}$$

As an exercise, one can show that by using:

$$[A, BC] = B\,[A, C] + [A, B]\,C \text{ or } = -B\,\{A, C\} + \{A, B\}\,C \tag{5.181}$$

The relation for the time evolution still holds:

$$\dot{\psi} = i\,[H, \psi] \iff \text{Dirac equation} \tag{5.182}$$

So we have a quantum theory of the Dirac field that reproduces classical mechanics. Let us go on and see what are the properties of the theory. First, the anti-commutation relations for ψ give anti-commutation relations for a and b:

$$\left\{a_r(p), a_s(k)^\dagger\right\} = (2\pi)^3 2E(p)\delta_{rs}\delta^3(k-p) \qquad \left\{b_r(p), b_s^\dagger(k)\right\} = (2\pi)^3 2E(p)\delta_{rs}\delta^3(k-p) \tag{5.183}$$

The Hamiltonian is now the same as before, provided that we switch the roles of b and $b^\dagger$, by defining $b_s^\dagger \equiv \tilde{b}_s$.

5.5.3 Hilbert space

$$H = \int d\Omega_p E_P \sum_s a_s^\dagger(p)a_s(p) - b_s^\dagger(p)b_s(p) \equiv \int d\Omega_p E_P \sum_s a_s^\dagger(p)a_s(p) + \tilde{b}_s^\dagger(p)\tilde{b}_s(p) \tag{5.184}$$

Then, we define the vacuum as $|0\rangle$

$$\tilde{b}_s(p)|0\rangle = a_s(p)|0\rangle = 0 \tag{5.185}$$

The commutation relations between a , $\tilde{b}$ and the Hamiltonian now are

$$\left[H, a_r^\dagger(P)\right] = \omega_p a_r^\dagger(P); \tag{5.186}$$
$$\left[H, \tilde{b}_r^\dagger(P)\right] = \omega_p \tilde{b}_r^\dagger(P); \tag{5.187}$$
$$\tag{5.188}$$

So both the operators now create positive energy quanta, this time without negative norm problems.

Let us see how this works in a toy model where one has only one set of operators (i.e. only one r and one momentum). It is the equivalent of the harmonic oscillator with anti-commutation relations.

$$\left\{b, b^\dagger\right\} = 1 \quad \to \quad bb^\dagger + b^\dagger b = 1 \qquad \text{and} \qquad \{b, b\} = \left\{b^\dagger, b^\dagger\right\} = 0 \tag{5.189}$$

Call $|0\rangle$ such that $b|0\rangle = 0$. Then, let's call $|1\rangle = b^\dagger|0\rangle$. Is there $|2\rangle \equiv \left(b^\dagger\right)^2|0\rangle$? No, because from the anti-commutation $|b^\dagger\rangle^2 = 0$. So the Hilbert space is 2 dimensional.

$$|0\rangle = \begin{pmatrix}0\\1\end{pmatrix} \qquad b^\dagger = \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix} \qquad b = \begin{pmatrix}0 & 0\\1 & 0\end{pmatrix} \qquad |1\rangle = \begin{pmatrix}1\\0\end{pmatrix} \tag{5.190}$$

But so far there's no notion of what is the excited state and what is the ground state. However, one notices the exclusion principle at work here.

We can now turn back to the original system and consider how the Fock space looks like. We have that $a_r^\dagger(p)$ and $\tilde{b}_r^\dagger(p)$create a quantum of energy E(p) and 3-momentum $\vec{p}$, polarization r, and we know from the example above that we cannot have two excited states with the same r and p (exclusion principle). We will have thus one particle states, two-particle, three-particle etc. states with the restriction above. Moreover, consider the following two-particle state:

$$a_s^\dagger(k)a_r^\dagger(p)|0\rangle = |(\vec{p},s);(\vec{k},r)\rangle = -|(\vec{k},r);(\vec{p},s)\rangle = |\psi\rangle \tag{5.191}$$

1. Notice that if you get identical quanta, then $|\psi\rangle = -|\psi\rangle$ so that $|\psi\rangle = 0$. Therefrom follows exclusion principle. You cannot create two particles of the same quantum numbers.

2. Ex: a two particle state

$$|\Psi\rangle = \int d\Omega_k d\Omega_p \sum_{r,s} f((\vec{k},r);(\vec{p},s))|(\vec{k},r);(\vec{p},s)\rangle \tag{5.192}$$

Since the ket is totally anti-symmetric, the only relevant piece in f is the anti-symmetric part, the symmetric part will drop out.

One should conclude that the Dirac field quanta obeys Fermi-Dirac statistics.

The spin = 0 is quantized according to commutation relations so the resulting quanta satisfies Bose-Einstein statistics.
The spin = $\frac{1}{2}$ is quantized according to anti-commutation relations so the resulting quanta satisfies Fermi-Dirac statistics.

These are the two examples of a theorem :

Spinor-statistics Theorem:

Given a local-relativistic invariant unitary bounded Hamiltonian implies a connection between the spin of the quanta and their statistics : integer spin quanta follow bose-Einstein and half integer satisfy Fermi-Dirac.

5.6 Properties of the Dirac field

5.6.1 Four momentum

Let us first rename $\tilde{b} \to b$ for convenience. We have that

$$\psi = \int d\Omega_p e^{ipx} \left[\sum_r a_r(p)u_r(p) + b_r^\dagger(-p)v_r(-p)\right]. \tag{5.193}$$

Then we can derive the momentum

$$\vec{P} = -i\int d^3\vec{x}\psi^\dagger\vec{\nabla}\psi = \tag{5.194}$$

$$= \int d\Omega_p \vec{p}\left[\sum_r a_r(p)u_r(p) + b_r^\dagger(p)v_r(p)\right]. \tag{5.195}$$

so that

$$P^\mu = \int d\Omega_p p^\mu \left[\sum_r a_r(p)u_r(p) + b_r^\dagger(p)v_r(p)\right]. \tag{5.196}$$

We have four types of quanta with same mass m:

$$a_1^\dagger,\ a_2^\dagger,\ b_1^\dagger,\ b_2^\dagger. \tag{5.197}$$

Corresponding to the fact that we have particle and anti-particles with two polarizations (spins).

5.6.2 Internal charge

We can see that the Dirac Lagrangian is invariant under a global $U(1)$ transformation:

$$\psi \to e^{-i\alpha}\psi \tag{5.198}$$
$$\bar{\psi} \to e^{i\alpha}\bar{\psi}. \tag{5.199}$$
$$\tag{5.200}$$

The Noether current associated to this symmetry is

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)}\Delta\psi = i\bar{\psi}\gamma^\mu(-i\psi) = \bar{\psi}\gamma^\mu\psi, \tag{5.201}$$

and the charge reads (after straightforward manipulations and after dropping the constant term coming from the delta functions of the commutators)

$$Q = \int d^3\vec{x} = \int d\Omega_p \left[\sum_r a_r(p)u_r(p) - b_r^\dagger(p)v_r(p)\right]. \tag{5.202}$$

We can now see that $a^\dagger$ creates quanta with $Q = +1$ (fermions) while $b^\dagger$ creates quanta with $Q = -1$ (anti-fermions), corresponding to the fact that particles and anti-particles have opposite electric ($U(1)$) charge.

5.6.3 Time evolution

Like for the Klein-Gordon field we can consider the Heisenberg picture operators. For the time evolution of the ladder operators one finds

$$e^{iHt}a_r(p)e^{-iHt} = a_r(p)e^{-i\omega_p t}, \tag{5.203}$$
$$e^{iHt}b_r(p)e^{-iHt} = b_r(p)e^{-i\omega_p t}. \tag{5.204}$$

which inserted in the mode expansion of the field at $t = 0$ gives

$$e^{iHt}\psi(\mathbf{x})e^{-iHt} = \psi(t,\mathbf{x}) \equiv \psi(x) = \tag{5.205}$$
$$= \int d\Omega_p \left[\sum_r e^{-ipx}a_r(p)u_r(p) + e^{ipx}b_r^\dagger(p)v_r(p)\right]. \tag{5.206}$$

5.6.4 Angular momentum and boosts

For a scalar field the action of Lorentz transformations was

$$\phi(x) \to \phi(\Lambda^{-1}x). \tag{5.207}$$

In the case of a spinor field we have:

$$\psi_\alpha(x) \to D_{\alpha\beta}(\Lambda)\psi_\beta(\Lambda^{-1}x). \tag{5.208}$$

Where $D_{\alpha\beta}(\Lambda)$ is the rep of Λ on (Dirac) spinors:

$$D_{\alpha\beta}(\Lambda) = \Lambda_D = \begin{pmatrix} \Lambda_L & 0 \\ 0 & \Lambda_R \end{pmatrix}. \tag{5.209}$$

One can show that this corresponds to:

$$\Lambda_D = e^{-i\frac{1}{2}\omega_{\mu\nu}S^{\mu\nu}}, \tag{5.210}$$

where $S^{\mu\nu} = \frac{i}{4}[\gamma^m u, \gamma^n u]$ is the generator of the Lorentz group in this representation. In particular one has

$$S^{0i} = -\frac{i}{2}\begin{pmatrix}\sigma^i & 0\\ 0 & -\sigma^i\end{pmatrix}, \tag{5.211}$$

$$S^{ij} = \frac{1}{2}\varepsilon^{ijk}\begin{pmatrix}\sigma^k & 0\\ 0 & \sigma^k\end{pmatrix}. \tag{5.212}$$

The Noether current associated to Lorentz boosts can be computed noticing that the variation of the field $\Delta\psi$ is

$$\Delta\psi = \omega_{\mu\nu}\left[-i\frac{S^{\mu\nu}}{2} + \frac{1}{2}\left(x^\mu\gamma^\nu - x^\nu\gamma^\mu\right)\right]. \tag{5.213}$$

and then one can compute the charges

$$J^\rho_{\mu\nu} \to J_{\mu\nu} \equiv \int d^3\vec{x} J^0_{\mu\nu}. \tag{5.214}$$

In particular we have that

$$J^{ij} = \varepsilon^{ijk}J_k, \tag{5.215}$$

where J_k will be associated with the angular momentum, we will see in the exercises that

$$J_k = \int d^3\vec{x}\psi^\dagger\left[\vec{x}\wedge(-\vec{\nabla}) + \frac{\vec{\Sigma}}{2}\right]\psi. \tag{5.216}$$

where

$$\vec{\Sigma} = \begin{pmatrix}\vec{\sigma} & 0\\ 0 & \vec{\sigma}\end{pmatrix}. \tag{5.217}$$

and the two components represent the "orbital" angular momentum and the spin. We will see in the exercise that the angular momentum operator acts on single particle states

$$|\Psi\rangle = \int d^3\vec{x} f_\alpha(x)\psi^\dagger_\alpha(x)|0\rangle = \tag{5.218}$$

$$= \int d^3\vec{x}\sum_r u^\dagger_\alpha(r,p)e^{ipx}f_\alpha(x)|p,r\rangle = \tag{5.219}$$

$$= \int d\Omega_p F(r,p)|p,r\rangle, \tag{5.220}$$

as

$$\vec{J}|\Psi\rangle = \sum_{r,r'}\int d\Omega_p\left[(-ip\wedge\partial_p)\delta_{rr'} + (\vec{\sigma})_{rr'}/2\right]F(r',p). \tag{5.221}$$

Chapter 6

Representations of the Poincaré group

6.1 Poincaré transformations and Irreducible Representations

From the previous discussion on the Poincaré group, we recall that a Poincaré transformation $g(\Lambda, a) \in ISO(3,1)$ acts on space-time coordinates as

$$x'^{\mu} = \Lambda^{\mu}_{\ \nu} x^{\nu} + a^{\mu}, \qquad \Lambda \in SO(3,1), \qquad a \in \mathbb{R}^4. \tag{6.1}$$

It follows the group composition rule

$$g(\Lambda_1, a_1) g(\Lambda_2, a_2) = g(\Lambda_1 \Lambda_2, \Lambda_1 a_2 + a_1), \tag{6.2}$$

which implies

$$\begin{aligned} &g(\Lambda, a) = g(\mathbb{1}, a) g(\Lambda, 0), \\ &g^{-1}(\Lambda, a) = g(\Lambda^{-1}, -\Lambda^{-1} a). \end{aligned} \tag{6.3}$$

In this section, we will study the representations of the Poincaré group on the Hilbert space of states $\mathcal{H}$. To any $g(\Lambda, a) \in ISO(3,1)$, there corresponds a unitary operator that we call

$$\begin{aligned} &U(g(\Lambda, a)) \equiv U(\Lambda, a), \\ &U^{\dagger}(g(\Lambda, a)) = U(g^{-1}(\Lambda, a)). \end{aligned} \tag{6.4}$$

They are generated by the hermitian operators, P^{μ} and $\mathcal{J}^{\mu\nu}$, via the exponential map

$$\begin{aligned} &U(g(\Lambda, 0)) \equiv U(\Lambda) = \exp(-\frac{i}{2} \omega_{\mu\nu} \mathcal{J}^{\mu\nu}), \\ &U(g(0, a)) \equiv U(a) = \exp(i a_{\mu} P^{\mu}). \end{aligned} \tag{6.5}$$

For a generic $g(\Lambda, a) \in ISO(3,1)$ we thus have

$$U(\Lambda, a) = U(a) U(\Lambda). \tag{6.6}$$

The rotations and boosts generators can be extracted as usual as

$$J^i = \frac{1}{2} \varepsilon^{ijk} \mathcal{J}^{jk}, \qquad K^i = \mathcal{J}^{i0}. \tag{6.7}$$

The generators must satisfy the algebra

$$\begin{aligned} &[P^{\mu}, P^{\nu}] = 0, \\ &[P^{\mu}, \mathcal{J}^{\rho\sigma}] = i(\eta^{\mu\rho} P^{\sigma} - \eta^{\mu\sigma} P^{\rho}), \\ &[\mathcal{J}^{\mu\nu}, \mathcal{J}^{\rho\sigma}] = i(\eta^{\nu\rho} \mathcal{J}^{\mu\sigma} + \eta^{\mu\sigma} \mathcal{J}^{\nu\rho} - \eta^{\mu\rho} \mathcal{J}^{\nu\sigma} - \eta^{\nu\sigma} \mathcal{J}^{\mu\rho}), \end{aligned} \tag{6.8}$$

to ensure the validity of the group product rule (6.2). To prove so, let us consider two Poincaré transformations, the second of which is infinitesimal: (Λ, a) and $(1+\omega, \varepsilon)$. From the group product rule (6.2) we have

$$U^\dagger(\Lambda, a)U(1+\omega,\varepsilon)U(\Lambda,a) = U(1+\Lambda^{-1}\omega\Lambda, \Lambda^{-1}(\omega a+\epsilon)). \tag{6.9}$$

From the exponential form in (6.5), it follows that

$$\begin{aligned} &U^\dagger(a)P^\mu U(a) = P^\mu, \\ &U^\dagger(\Lambda)P^\mu U(\Lambda) = \Lambda^\mu_{\ \nu}P^\nu, \\ &U^\dagger(\Lambda)\mathcal{J}^{\mu\nu}U(\Lambda) = \Lambda^\mu_{\ \rho}\Lambda^\nu_{\ \sigma}\mathcal{J}^{\rho\sigma}. \end{aligned} \tag{6.10}$$

As we should have expected, the Hermitian operators P^μ and $\mathcal{J}^{\mu\nu}$ transform respectively as a vector and as a two-rank tensor (more precisely as the $(1,0)\oplus(0,1)$ representation) under Lorentz transformations. Taking also Λ and a to be infinitesimal transformations we can find the Poincaré algebra written above.

In a Poincaré invariant theory, the Hilbert space can be decomposed into unitary irreducible representations of the $ISO(3,1)$ group. It is therefore of fundamental importance to classify all the possible representations. To begin with, let us recall that these unitary representations must be infinite-dimensional because the group is non-compact. As for the angular momentum in quantum mechanics, the first step amounts to identify the Casimir operators of the group. In fact, their eigenvalues label the invariant sub-spaces of $\mathcal{H}$, the sets of states that are left invariant under Poincaré transformations. After that, we can find a basis of states for each of these sub-spaces.

By definition, the Casimir operators are Poincaré invariant operators

$$U^\dagger(\Lambda,a)\mathcal{C}\,U(\Lambda,a) = \mathcal{C}, \qquad \forall\ g(\Lambda,a)\in ISO(3,1), \tag{6.11}$$

built from the generators P^μ and $\mathcal{J}^{\mu\nu}$. This condition implies the commutation relations

$$\begin{aligned} [\mathcal{C}, \mathcal{J}_{\mu\nu}] &= 0, \\ [\mathcal{C}, P_\mu] &= 0. \end{aligned} \tag{6.12}$$

The first equation can be satisfied by constructing the Casimir as a product of generators where all the Lorentz indices are contracted (either using $\eta^{\mu\nu}$ or $\varepsilon^{\mu\nu\rho\sigma}$). The constraints coming from translation invariance are less obvious and we are now going to exploit them. You can show that among all the possible Lorentz invariant contractions there are only four independent operators (none of them can be built as a polynomial of the other ones)

$$P^\mu P_\mu,\ \ \mathcal{J}^{\mu\nu}\mathcal{J}_{\mu\nu},\ \ \varepsilon^{\mu\nu\rho\sigma}\mathcal{J}_{\mu\nu}\mathcal{J}_{\rho\sigma},\ \ W^\mu W_\mu, \tag{6.13}$$

where

$$W^\mu = \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\mathcal{J}_{\nu\rho}P_\sigma, \tag{6.14}$$

is the Pauli-Lubanski pseudo-vector. Note that there is no ambiguity in the operator ordering because by commuting P_σ with $\mathcal{J}_{\nu\rho}$ we generate a symmetric object that vanishes in the contraction with the Levi-Civita tensor. This operator satisfies

$$\begin{aligned} &W_\mu P^\mu = 0, \\ &[P^\mu, W^\nu] = 0, \\ &[\mathcal{J}^{\mu\nu}, W^\rho] = i(\eta^{\nu\rho}W^\mu - \eta^{\mu\rho}W^\nu) \\ &[W^\mu, W^\nu] = i\varepsilon^{\mu\nu\rho\sigma}W_\rho P_\sigma. \end{aligned} \tag{6.15}$$

By an explicit computation, you can prove that only

$$M^2 = P^\mu P_\mu \tag{6.16}$$

and

$$W^2 = W^\mu W_\mu \tag{6.17}$$

are translational invariant. These are the two Casimir operators of the Poincaré group. Since they commute with every transformation, their eigenspaces are Poincaré invariant. It is immediate to recognize P^2 as the invariant

mass while the physical meaning of W^2 is going to be identified shortly. Calling m^2 and w^2 the eigenvalues of M^2 and W^2 respectively, we have that the Hilbert space decomposes as

$$\mathcal{H} = \bigoplus_{m^2,w^2} \mathcal{H}(m^2, w^2). \tag{6.18}$$

We want to identify a basis of states for each $\mathcal{H}(m^2, w^2)$. They will provide different representations of the Poincaré group. As in quantum mechanics, we can choose them to be labeled by the eigenvalues of a maximal set of commuting operators. Part of them can be chosen to be the operators P^μ so that the state has a definite energy and momentum. Note that because translations commute, we can simultaneously diagonalize all the momentum components. Usually, this is not enough to uniquely identify a state: there can be several states with the same quantum numbers we have defined up to now, m^2, w^2, and p^μ. These states are generically shuffled among each other when you perform a Lorentz transformation that leaves the momentum p^μ invariant. Let us call σ the additional quantum numbers that remove the degeneracy of the states with fixed momentum (we will see that there is just one of them). Then, under $\mathbf{p}$, a transformation that leaves p^μ invariant, we can write

$$U(\Lambda)|\mathbf{p},\sigma\rangle = \sum_{\sigma'} \mathcal{D}_{\sigma'\sigma}(\Lambda,\mathbf{p})|\mathbf{p},\sigma'\rangle, \qquad \text{for } \Lambda^\mu{}_\nu p^\nu = p^\mu. \tag{6.19}$$

Indeed, the state $U(\Lambda)|\mathbf{p},\sigma\rangle$ has the same momentum p^μ as $|\mathbf{p},\sigma\rangle$

$$\begin{aligned} P^\mu U(\Lambda)|\mathbf{p},\sigma\rangle &= U(\Lambda)U^\dagger(\Lambda)P^\mu U(\Lambda)|\mathbf{p},\sigma\rangle = U(\Lambda)\Lambda^\mu{}_\nu P^\nu|\mathbf{p},\sigma\rangle \\ &= (\Lambda p)^\mu U(\Lambda)|\mathbf{p},\sigma\rangle = p^\mu\, U(\Lambda)|\mathbf{p},\sigma\rangle. \end{aligned} \tag{6.20}$$

As such, it can be written as a linear combination of the states with the same momentum. The numbers $\mathcal{D}_{\sigma'\,\sigma}(\Lambda,p)$ are just the coefficients of the linear decomposition. The set of transformations that leave p^μ invariant is a subgroup of the Lorentz group called the little group G_L. We now recognize that equation (6.19) is the statement that states with fixed momentum are a basis for a representation of the related little group G_L. This is a very important conclusion that allows us to build representations of the Poincaré group starting from representations of the Little group. We will get into the details shortly but let us explain the general idea first. Imagine you have found a basis for the states with a reference momentum $\bar{p}^\mu$: $\{|\bar{\mathbf{p}},\,\sigma\rangle\}$. This furnishes a representation of G_L. Then, we can build a basis for the states with any arbitrary momentum p^μ by choosing a reference Lorentz transformation Λ_p such that[1]

$$\Lambda_p: \quad p^\mu = \Lambda_p{}^\mu{}_\nu \bar{p}^\nu, \tag{6.21}$$

and defining the states with momentum p^μ as

$$|\mathbf{p},\sigma\rangle \equiv U(\Lambda_p)|\bar{p}\,\sigma\rangle = H(p)|\bar{p}\,\sigma\rangle. \tag{6.22}$$

For notational simplicity, we have defined the unitary operator implementing the reference Lorentz transformation as $U(\Lambda_p) = H(p)$. The previous definition completely fixes the transformation low of any state under a generic Poincaré transformation. For example, under a translation,

$$\begin{aligned} U(a)|\mathbf{p},\,\sigma\rangle =& U(a)H(p)|\bar{\mathbf{p}},\,\sigma\rangle \\ =& H(p)H^\dagger(p)e^{i\hat{P}\cdot a}H(p)|\bar{\mathbf{p}},\,\sigma\rangle \\ =& H(p)e^{i\Lambda^\mu_{p\,\nu}\bar{p}^\nu a_\mu}|\bar{\mathbf{p}},\,\sigma\rangle \\ =& H(p)e^{ip^\mu a_\mu}|\bar{\mathbf{p}},\,\sigma\rangle \\ =& e^{ip^\mu a_\mu}|\mathbf{p},\,\sigma\rangle. \end{aligned} \tag{6.23}$$

Under a Lorentz transformation, we have

$$\begin{aligned} U(\Lambda)|\mathbf{p},\,\sigma\rangle =& U(\Lambda)H(p)|\bar{\mathbf{p}},\,\sigma\rangle \\ =& H(\Lambda p)H^\dagger(\Lambda p)U(\Lambda)H(p)|\bar{\mathbf{p}},\,\sigma\rangle \\ =& H(\Lambda p)U(\mathcal{W}(\Lambda,\mathbf{p}))|\bar{\mathbf{p}},\,\sigma\rangle, \end{aligned} \tag{6.24}$$

with $\mathcal{W}(\Lambda,\mathbf{p}) = \Lambda^{-1}_{\Lambda p}\Lambda\Lambda_p$. Note that

$$\bar{p} \xrightarrow{\Lambda_p} p \xrightarrow{\Lambda} \Lambda p \xrightarrow{\Lambda^{-1}_{\Lambda_p}} \bar{p}. \tag{6.25}$$

[1] There exist an infinite set of transformations that map $\bar{p}$ to p. They are related by a right multiplication of an element of G_L. Different choices give rise to different, but equivalent, bases of the Hilbert space.

Therefore, $\mathcal{W}(\Lambda, \mathbf{p})$ belongs to the little group G_L and we know how it acts on states. At the same time, $H(\Lambda p)$ realizes our reference transformation that defines[2] $|\mathbf{p}_\Lambda, \sigma\rangle = H(\Lambda\, p)|\bar{\mathbf{p}}, \sigma\rangle$. We thus have

$$U(\Lambda)|\mathbf{p}, \sigma\rangle = \sum_{\sigma'} \mathcal{D}_{\sigma'\sigma}(\mathcal{W}(\Lambda, \mathbf{p}))H(\Lambda p)|\bar{\mathbf{p}}, \sigma'\rangle = \sum_{\sigma'} \mathcal{D}_{\sigma'\sigma}(\mathcal{W}(\Lambda, \mathbf{p}))|\mathbf{p}_\Lambda, \sigma'\rangle. \tag{6.26}$$

The little group transformation $\mathcal{W}(\Lambda, \mathbf{p})$ that depends on the starting momentum $\mathbf{p}$ and on the Lorentz transformation Λ is usually called a Wigner transformation. From the previous equation, it appears clear how representations of the little group G_L induce representations of the full Poincaré group. This is the general strategy we will follow to build representations of $ISO(3,1)$. Let us summarize the steps:

- for fixed values of the Casimir operators P^2 and W^2 we choose a reference frame where the state has a particular momentum $\bar{p}^\mu$ and we identify its little group G_L (as well as the meaning of the eigenvalue of the Casimir W^2),
- we build irreducible representations of G_L,
- we choose a reference Lorentz transformation to build a basis for the states with any momentum p^μ.

The set of states built in this way is a basis for a representation of the full Poincaré group. We can now follow these steps in more details.
We start noting that there exist different representations depending on the eigenvalue of P^2 and the sign of P^0:

- **Massive representations**: $m^2 > 0$ and $p^0 > 0$,
- **Massless representations**: $m^2 = 0$ and $p^0 > 0$,
- **Zero momentum states**: $m^2 = 0$ and $p^0 = 0$.

One could have also considered the situation where $m^2 < 0$. In this case, we would have to build representations of the little group $SO(2,1)$. However, the sign of the energy of the states in these representations wouldn't be defined. As such, they cannot correspond to stable configurations of the system, they are not physical. Regarding the states with zero momentum, there is just one state of this type, building a trivial representation. This is the unique Poincaré invariant vacuum $|0\rangle$

$$P^\mu|0\rangle = 0, \quad J^{\mu\nu}|0\rangle = 0. \tag{6.27}$$

Its little group is thus the full Lorentz group $SO(3,1)$. We can now discuss massive and massless representations. Table ?? summarizes the little groups for the different representations, some of which are going to be introduced in the next paragraphs.

P^2	P^0	Reference vector	G_L
$m^2 = 0$	$P^0 = 0$	$\bar{p}^\mu = (0,0,0,0)$	$SO(3,1)$
$m^2 > 0$	$P^0 > 0$	$\bar{p}^\mu = (m,0,0,0)$	$SO(3)$
$m^2 = 0$	$P^0 > 0$	$\bar{p}^\mu = (\omega,0,0,\omega)$	$ISO(2)$
$m^2 < 0$	–	$\bar{p}^\mu = (0,0,0,m)$	$SO(2,1)$

6.2 Massive representations

In this section, we will be interested in the construction of massive representation for which $m^2 > 0$ and $p^0 > 0$. In this situation, we can choose the reference momentum to be $\bar{p} = (m,0,0,0)$. The little group is thus $G_L = SO(3)$, the group of spatial rotation. We can compute the action of the Pauli-Lubanski pseudovector on states of this form. We find

$$W^\mu = \frac{1}{2}\epsilon^{\mu\nu\rho 0} J_{\nu\rho} m \qquad \Leftrightarrow \qquad W^0 = 0 \quad \text{and} \quad W^i = -mJ^i. \tag{6.28}$$

[2]For ease of notation we indicate by $\mathbf{p}_\Lambda$ the spacial component of the 4-vector Λp.

Thus, the components of the Pauli-Lubanski pseudovector are the generators of the Little group. The Casimir W^2 takes the form

$$W^2 = -m^2 J^2. \tag{6.29}$$

The operator J^2 is nothing else but the square of the angular momentum. As such, for massive representations, $w^2 = -m^2 s(s+1)$ and we can trade w^2 for the total spin s of the particle[3]. The basis $\{|\bar{\mathbf{p}}, \sigma\rangle\}$ furnishes a representation of the $SO(3)$ little group with dimension $2s+1$ and we can identify σ as the eigenvalue of the operator J^3

$$J^3|\bar{\mathbf{p}}, \sigma\rangle = \sigma|\bar{\mathbf{p}}, \sigma\rangle \qquad \sigma = -s, ..., +s. \tag{6.30}$$

To summarize, the Casimir of the representation acts on these states as

$$P^2 = m^2, \quad W^2 = -m^2 s(s+1). \tag{6.31}$$

and we simultaneously diagonalize the momentum operators P^μ and one of the components of the Pauli-Lubanski pseudo-vector. Now that the representation of G_L is constructed, it remains to Lorentz transform these states to get a representation of the full Lorentz group. We thus chose a particular Lorentz transformation Λ_p and define

$$|\mathbf{p}, \sigma\rangle \equiv H(p)|\bar{\mathbf{p}}, \sigma\rangle, \quad \text{where } H(p) \equiv U(\Lambda_p). \tag{6.32}$$

As we already noticed, the reference transformation is defined up to the right multiplication of an element of the little group, $SO(3)$ in this case. Note that σ, for the states with generic momentum p, is the eigenvalue of the operator $H(p)J^3H^\dagger(p)$ whose interpretation depends explicitly on the choice of the reference Lorentz transformation. The normalization of the states is

$$\langle p, \sigma|\bar{p}, \sigma'\rangle = (2\pi)^3 2E_{\bar{\mathbf{p}}}\delta^{(3)}(\mathbf{p} - \bar{\mathbf{p}})\delta_{\sigma\sigma'}. \tag{6.33}$$

Finally, we can consider a generic Lorentz transformation acting on these states (there is nothing more to add to the case of translations than what has been said before)

$$\begin{aligned} U(\Lambda)|\mathbf{p}, \sigma\rangle =& H(\Lambda p)U(W(\Lambda, p))|\bar{\mathbf{p}}, \sigma\rangle \\ =& H(\Lambda p)\sum_{\sigma'} \mathcal{D}^{(s)}_{\sigma'\sigma}(W(\Lambda, p))|\bar{\mathbf{p}}, \sigma'\rangle \\ =& \sum_{\sigma'} \mathcal{D}^{(s)}_{\sigma'\sigma}(W(\Lambda, p))|\mathbf{p}_\Lambda, \sigma'\rangle, \end{aligned} \tag{6.34}$$

where in this case $\mathcal{D}^{(s)}_{\sigma'\sigma}$ is a $(2s+1)\times(2s+1)$ $SU(2)$ matrix. It is important to notice that the Wigner rotation depends on the choice made for Λ_p. Different choices will give different expressions for these matrices. In the following, one particular case will be studied, namely the spin basis.

The spin basis

In the spin basis, the reference Lorentz transformation consists in taking a pure boost in the direction of $\mathbf{p}$

$$\Lambda_p^{Spin} \equiv e^{i\eta_{\mathbf{p}}\cdot\mathbf{K}}, \tag{6.35}$$

where

$$\eta_{\mathbf{p}} \equiv \tanh^{-1}\left|\frac{\mathbf{p}}{p_0}\right|\mathbf{n}, \quad \text{with} \quad \mathbf{n} \equiv \frac{\mathbf{p}}{|\mathbf{p}|}, \tag{6.36}$$

This choice can be interpreted as taking a state at rest, with spin σ along the third direction, and boosting it to get a state with non-vanishing momentum and the same spin. The goal now is to find the expression for the Wigner rotations. We consider at first the case of pure rotations. The correspondent unitary operator is $U(R_\theta) = e^{-i\theta\cdot\mathbf{J}}$. It acts on the state $|\mathbf{p}, \sigma\rangle$ as

[3] We want to include half integers spin particles. As such, we actually consider $SU(2)$ representations.

$$\begin{aligned} U(R_\theta)|\mathbf{p},\,\sigma\rangle =& U(R_\theta)e^{i\eta_{\mathbf{p}}\cdot\mathbf{K}}U^\dagger(R_\theta)U(R_\theta)|\bar{\mathbf{p}},\,\sigma\rangle \\ =& e^{i(R_\theta\eta_{\mathbf{p}})\cdot\mathbf{K}}U(R_\theta)|\bar{\mathbf{p}},\,\sigma\rangle \\ =& e^{i(R_\theta\eta_{\mathbf{p}})\cdot\mathbf{K}}\sum_{\sigma'}\mathcal{D}^{(s)}_{\sigma'\sigma}(R_\theta)|\bar{\mathbf{p}},\sigma'\rangle \\ =& \sum_{\sigma'}\mathcal{D}^{(s)}_{\sigma'\sigma}(R_\theta)|R_\theta\mathbf{p},\sigma'\rangle. \end{aligned} \tag{6.37}$$

We see that rotations act in two ways. First, they rotate the momentum, and second, they act on σ. The first action can be thought of as generated by the orbital angular momentum while the second one by the spin. In this basis, we thus recognize the decomposition $\mathbf{J} = \mathbf{L} + \mathbf{S}$ we are used to from quantum mechanics. It is a good exercise to consider an infinitesimal rotation and explicitly verify the form of the generators. This shows that the spin basis is particularly suitable when considering the non-relativistic limit, as it maps to the known properties of quantum mechanics. We do not see this decomposition when using other bases. Another important basis is the helicity basis, where the states will be labeled by their helicity, that is the spin projected along their momentum. This basis is more suitable when dealing with the relativistic limit. It will not be described in these notes for the massive case. However, we will choose this basis when dealing with massless states, where only the helicity is defined.

Last, we consider the states' transformation under pure boosts. For simplicity we specialize to the infinitesimal case $\Lambda_\eta = e^{i\eta\cdot\mathbf{K}}$, with $\eta \ll 1$. The states transform as:

$$U(\Lambda_\eta)|\mathbf{p},\,\sigma\rangle = \sum_{\sigma'}\mathcal{D}^{(s)}_{\sigma'\sigma}(\mathcal{W}(\Lambda_\eta,p))|\mathbf{p}_{\Lambda_\eta},\sigma'\rangle, \tag{6.38}$$

where the Wigner rotation for boosts can be proven to be

$$\mathcal{W}(\Lambda_\eta,p) = \mathbb{1} + i\frac{(\eta\times\mathbf{p})\cdot\mathbf{J}}{p^0+m} + \mathcal{O}(\eta^2). \tag{6.39}$$

This can be interpreted as a rotation in the direction given by $(\eta\times\mathbf{p})$.

6.3 Massless Representations

We now come to massless representations. In this case, $P^2 = 0$ and we can always find a frame where $\bar{p}^\mu = (\omega, 0, 0, \omega)$. The components of the Pauli-Lubanski pseudo-vector are

$$W^0 = -\omega J^3, \quad W^1 = \omega(K^2 - J^1), \quad W^2 = \omega(-K^1 - J^2), \quad W^3 = W^0. \tag{6.40}$$

Trading W^3 for J^3, we get the following commutation relations:

$$[J^3, W^1] = iW^2, \quad [J^3, W^2] = -iW^1, \quad [W^1, W^2] = 0. \tag{6.41}$$

This is the algebra of $ISO(2)$, the isometry group of the plane[4]. Rotations in the plane are generated by J^3 while translations are generated by W^1 and W^2. This is the little group G_L of $\bar{p}^\mu$. We proceed to build irreducible representations for $ISO(2)$. For that purpose, the commutators can be rewritten conveniently in terms of $W^\pm = W^1 \pm iW^2$.

$$[J^3, W^\pm] = \pm W^\pm, \quad [W^\pm, W^\mp] = 0. \tag{6.42}$$

We notice that W^+ and W^- act respectively as raising and lowering operators for J_3. This means that for any eigenstate $|\bar{\mathbf{p}},\,\lambda\rangle$ of J_3

$$\begin{aligned} J_3|\bar{\mathbf{p}},\,\lambda\rangle = \lambda|\bar{\mathbf{p}},\,\lambda\rangle \quad \Rightarrow \quad & J_3W^+|\bar{\mathbf{p}},\,\lambda\rangle = (\lambda+1)W^+|\bar{\mathbf{p}},\,\lambda\rangle \\ & J_3W^-|\bar{\mathbf{p}},\,\lambda\rangle = (\lambda-1)W^-|\bar{\mathbf{p}},\,\lambda\rangle. \end{aligned} \tag{6.43}$$

The Casimir operator W^2 is proportional to the identity for any irreducible representation (can you tell why?)

$$(W)^2 = W^\mu W_\mu = -W^+W^- = -c\,\mathbb{1}. \tag{6.44}$$

[4]Pay attention to the fact that this is not a spatial plane.

with c a positive number. Indeed

$$c\,\mathbb{1} = W^- W^+ = (W^+)^\dagger W^+ \geq 0, \tag{6.45}$$

and for any operator A, $A^\dagger A$ is an hermitian positive semi-definite operator. There are then two possibilities.

$c > 0$. In this case $W^+/\sqrt{c}$ is a unitary operator and translations of the $ISO(2)$ little group are non-trivially realized on the states. This group is non-compact and its (non-trivial) unitary representations are thus infinite-dimensional. To see so, it is enough to recall equation (6.43) to realize that given any real value of λ, the eigenvalue of J^3, we can build a ladder of states by the repeated application of $W^\pm$

$$\begin{aligned}(W^+)^n|\bar{\mathbf{p}},\,\lambda\rangle &\propto |\bar{\mathbf{p}},\lambda+n\rangle \neq 0,\\ (W^-)^n|\bar{\mathbf{p}},\,\lambda\rangle &\propto |\bar{\mathbf{p}},\lambda-n\rangle \neq 0.\end{aligned} \tag{6.46}$$

A particle of this sort has infinite degrees of freedom and it has never been observed in nature[5]. For this reason, we will not consider this type of representations.

$c = 0$. In this case, translations are trivially realized. In fact

$$(W^-)^\dagger W^- = 0 \quad \Rightarrow \quad |W^-|\bar{\mathbf{p}},\,\lambda\rangle| = 0, \tag{6.47}$$

for any state in the representation. As such, we can write $W^- = W^+ = 0$. This leads to the conclusion that only the compact $SO(2)$ subgroup of the little group is non-trivially realized. This does admit finite dimensional representations. Indeed, they are one-dimensional. In this situation, the Pauli-Lubanski pseudo-vector is proportional to the momentum. In fact

$$W^\mu|\bar{p},\lambda\rangle = -\lambda\bar{p}^\mu|\bar{p},\lambda\rangle. \tag{6.48}$$

Considering now the reference transformation $H(p)$ we get

$$H(p)W^\mu H^\dagger(p)H(p)|\bar{p},\lambda\rangle = (\Lambda_{\mathbf{p}}^{-1}W)^\mu|\mathbf{p},\,\lambda\rangle = -\lambda\bar{p}^\mu|\mathbf{p},\,\lambda\rangle, \tag{6.49}$$

from which it follows the more general relation

$$W^\mu = -\lambda P^\mu. \tag{6.50}$$

This result, being covariant, also proves that λ is Lorentz invariant. To give it a more transparent physical meaning recall that the states $|\mathbf{p},\,\lambda\rangle$ are eigenvectors of the operator $H(p)J^3H^\dagger(p)$. You can realize that in the case of massless representations, this operator does not depend on the choice of the reference transformation $\Lambda_{\mathbf{p}}$ (J^3 commutes with the Little group that leaves $\bar{p}^\mu$ invariant). As such, we can consider the most convenient map from $\bar{p}^\mu$ to p^μ to compute it. Performing first a boost along the third direction that maps $\bar{p}^0$ to p^0 and then a rotation that aligns the momentum to $\hat{\mathbf{p}}$ we find

$$\lambda = \frac{\mathbf{J}\cdot\mathbf{p}}{|\mathbf{p}|}. \tag{6.51}$$

This means that λ is the projection of the angular momentum along the direction of motion, the helicity.

It is now time to construct the full representation with arbitrary momentum as we did for massive representations. We first need to define the Lorentz transformation that sends the reference momentum $\bar{p} = (\omega, 0, 0, \omega)$ to a generic null vector $p^\mu = (|\mathbf{p}|, \mathbf{p})$. Contrary to the massive case, we choose the helicity basis. This is done by first performing a boost in the third direction to get the correct energy $|\mathbf{p}|$, followed by a spatial rotation. The boost is given by

$$\begin{pmatrix}\omega\\0\\0\\\omega\end{pmatrix} \xrightarrow{boost} \begin{pmatrix}|\mathbf{p}|\\0\\0\\|\mathbf{p}|\end{pmatrix} = \begin{pmatrix}\cosh(\eta_p) & 0 & 0 & \sinh(\eta_p)\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ \sinh(\eta_p) & 0 & 0 & \cosh(\eta_p)\end{pmatrix}\begin{pmatrix}\omega\\0\\0\\\omega\end{pmatrix} = e^{i\eta_p\mathbf{K}_3}\begin{pmatrix}\omega\\0\\0\\\omega\end{pmatrix}, \tag{6.52}$$

with $\eta_p \equiv \ln\left(\frac{|\mathbf{p}|}{\omega}\right)$. We then consider the rotation $\mathcal{R}(\hat{\mathbf{p}})$ such that

$$\begin{pmatrix}|\mathbf{p}|\\\mathbf{p}^1\\\mathbf{p}^2\\\mathbf{p}^3\end{pmatrix} = \mathcal{R}(\hat{\mathbf{p}})\begin{pmatrix}|\mathbf{p}|\\0\\0\\|\mathbf{p}|\end{pmatrix}. \tag{6.53}$$

[5]For example, a multi-particle system made up of constituents of this sort has infinite heat capacity.

Overall we define

$$\Lambda_p \equiv \mathcal{R}(\mathbf{n}) e^{i\eta_p \mathbf{K}_3}. \tag{6.54}$$

States with arbitrary p^μ are then defined as

$$|\mathbf{p}, \lambda\rangle \equiv U(\Lambda_p)|\bar{\mathbf{p}}, \lambda\rangle \tag{6.55}$$

Let's study how they transform under a generic Lorentz transformation. By doing similar manipulations as in (6.24)

$$U(\Lambda)|\mathbf{p}, \lambda\rangle = U(\Lambda_{\Lambda p})U(\mathcal{W}(\Lambda, p))|\mathbf{p}, \lambda\rangle. \tag{6.56}$$

In this case, the Wigner transformation is just

$$U(\mathcal{W}(\Lambda, p)) = e^{-i\theta(\Lambda,p)J^3}, \tag{6.57}$$

and it boils down to just a phase

$$U(\Lambda)|\mathbf{p}, \lambda\rangle = e^{-i\theta(\Lambda,p)\lambda}|\mathbf{p}_\Lambda, \lambda\rangle. \tag{6.58}$$

To conclude the discussion, let us understand what are the possible values of the helicity λ. Indeed, as you know very well, the spin of a massive particle can assume either integer or half integer values. This can be proven by using the $SO(3)$ algebra. On the contrary, massless particle states are a representation of the $ISO(2)$ little group and the same argument does not apply. However, from topological considerations one can still derive the quantization of the helicity in integers or half integers[6]. The reason has to be found in the fact that the Lorentz group is not simply but doubly connected. Its topology is the one of $\mathbb{R}^3 \times S^3/\mathbb{Z}_2$. Therefore, any 4π rotation must give back the identity. This means that

$$e^{-i4\pi J^3}|\mathbf{p},\, \sigma\rangle = e^{-i4\pi\lambda}|\mathbf{p},\, \sigma\rangle = |\mathbf{p},\, \sigma\rangle, \tag{6.59}$$

from which

$$2\lambda \in \mathbb{Z}, \tag{6.60}$$

follows.

Up to know, the basis of massless particles we have constructed are representations of $ISO(3,1)$. A fundamental property of any local, Lorentz invariant quantum field theory with an hermitian Hamiltonian bounded from below is the symmetry with respect to a CPT transformation. Let's call θ the operator realizing it. Then, as it will be clear later, the generators transform as

$$\theta \mathbf{J} \theta = -\mathbf{J} \quad \text{and} \quad \theta \mathbf{P} \theta = \mathbf{P}. \tag{6.61}$$

It follows that

$$\theta|\mathbf{p},\, \lambda\rangle = |\mathbf{p},\, -\lambda\rangle. \tag{6.62}$$

This means that massless particle states, differently from the massive ones, are not a basis for a representation of CPT by themselves, they must come in pairs of opposite helicities. This is way there are two photon states $\lambda = \pm 1$ as well as two graviton states $\lambda = \pm 2$.

[6] See Weinberg Volume 1, section 2.7 for more details.

Part II

Forces and Interactions

Chapter 7

Vector Fields

7.1 Classical Electrodynamics

Recall the classification of the smallest representations of the Lorentz group:

(0,0)	(1/2,0)	(0,1/2)	(1/2, 1/2)	...
ϕ	χ_L^α	$\chi_R^{\dot\beta}$	$A^{\alpha\dot\beta}$	...

We had previously seen that the field $A^{\alpha\dot\beta}$, a matrix with one left and one right spinorial index, can be packaged into a 4-vector according to $A^{\alpha\dot\beta} \equiv \sigma_\mu^{\alpha\dot\beta} A^\mu$. We had also verified that A^μ does transform as a 4-vector under the Lorentz group:

$$A^{\alpha\dot\beta} \longrightarrow \Lambda_L^{\alpha\gamma}\Lambda_R^{\dot\beta\dot\delta}\sigma_\mu^{\gamma\dot\delta}A^\mu = \sigma_\nu^{\alpha\dot\beta}\Lambda_\mu^\nu A^\mu. \tag{7.1}$$

We would now like to construct the most general quadratic ($\equiv$ free) Poincaré invariant Lagrangian involving the field A^μ and at most two derivatives. Up to a total derivative it has the form:

$$\mathcal{L} = a_1 \partial_\mu A_\nu \partial^\mu A^\nu + a_2 (\partial_\mu A^\mu)^2 + a_3 A_\mu A^\mu \,. \tag{7.2}$$

Notice that by assigning A^μ scaling dimension 1, the above terms have at most dimension 4. At this point it seems we have an *embarassment of riches*. On one hand the above Lagrangian appears to have more free parameters than one would expect to describe a particle of given spin and arbitrary mass: we could always rescale A_μ to make $a_1 = 1$, but we would still be left with two free real parameters a_2 and a_3, which is one too many. On the other A_μ corresponds to 4 independent degrees of freedom (d.o.f.) describing in principle 4 independent polarizations. If we wanted to use A_μ to describe electromagnetic waves, which come in just two polarizations (left and right polarized light), we would need to somehow get rid of the extra polarizations.

The solution to both these problems comes from just one principle: symmetry. We can restrict the form of the Lagrangian (7.2) and at the same time lessen the number of d.o.f. *demanding* the action to be invariant under the following transformation:

$$A_\mu \longrightarrow A'_\mu \equiv A_\mu - \partial_\mu \alpha \,, \qquad S[A] = S[A - \partial\alpha] = S[A']. \tag{7.3}$$

One can easily check that this request translates in a constraint on the coefficients of (7.2): $a_1 = -a_2$, $a_3 = 0$. Finally we can determine the sign of a_1 requiring the Hamiltonian of the theory to be positive and choose the overall normalization asking for a canonical kinetic term. This fixes $a_1 = -1/2$ so that the final Lagrangian assumes the form

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \,, \qquad F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{7.4}$$

The equations of motion derived from the above Lagrangian reproduce the Maxwell equations (see Exercise Set 15). It is not immediately obvious to deduce from Maxwell equations that they imply the propagation of just two type of waves, the two transverse polarizations of light, but we already know from the study of classical electrodynamics that this is the case. In the following lectures we shall investigate this result in even more detail,

including its quantum realization. It is remarkable how the Maxwell equations, and their physical implications regarding the property of light waves (and light quanta), can be derived by a symmetry principle.

Notice that the symmetry transformations (7.3) do not involve a finite set of parameters but an infinite one, namely a function $\alpha(x)$ over spacetime. This kind of symmetry is called a *gauge symmetry*. A gauge symmetry, more than a symmetry, is a *redundancy* in the description of the physical d.o.f of a system. Indeed if we declare that the only physical observables are gauge invariant quantities, then $F_{\mu\nu}$ is and observable while A_μ, *per se* is not. Now $F_{\mu\nu}$ is fully determined by A_μ, but only 3 out of the 4 components of A_μ contribute: the configuration $A_\mu = \partial_\mu \alpha$ gives $F_{\mu\nu} = 0$ for any α. This can also be seen by making use of the Bianchi identity $\epsilon^{\mu\nu\rho\sigma}\partial_\mu F_{\nu\rho} = 0$. One finds that the longitudinal component of the magnetic field $\vec{B}_L$ is identically vanishing while the transverse part $\vec{B}_\perp$ is completely determined in terms of the curl of $\vec{E}_\perp$ by a first order (in time) differential equation (see Exercise Set 15). Hence only the electric field is unconstrained: this corresponds to 3 unconstrained variables, as compared to the 4 carried by A_μ. Here we haven't yet used the equations of motion $\partial_\mu F^{\mu\nu} = 0$: it is according to these other equations that 2 of these 3 variables turn out to be genuinely dynamical (corresponding to propagating waves), while the third variable, corresponding to a static Coulomb field, is not dynamical. (All this is clarified in the exercise on Coulomb gauge).

Gauge invariance ensures that the lagrangian and the physical quantities depend on less local variables than we have fields in our description. It thus allows to construct and describe systems that, is a certain sense, have a minimal number of d.o.f.. In order for this to make full sense, gauge symmetry should be extendable to interacting theories. Now, not only that is the case but, in addition, invariance under (7.3) turns out to fix in a very stringent way the form of the interaction between the vector field and the matter fields (bosons and fermions).

The matter Lagrangians are made gauge invariant by *postulating* the general transformation

$$\Psi_a(x) \longrightarrow \Psi'_a(x) \equiv e^{iq_a\alpha(x)}\Psi_a(x), \tag{7.5}$$

where Ψ_a is a general field and q_a is real number called the *charge* of the field. It is clear that with the above definition the simple kinetic term that one is used to write is not invariant since the transformation is spacetime dependent and doesn't commute with derivatives. On the other hand one can show that the combination

$$D_\mu\Psi_a \equiv (\partial_\mu + iq_aA_\mu)\,\Psi_a \tag{7.6}$$

transforms in the same way as the field Ψ_a under gauge transformation:

$$D_\mu\Psi_a \longrightarrow (\partial_\mu + iq_a(A_\mu - \partial_\mu\alpha))\,e^{iq_a\alpha(x)}\Psi_a = e^{iq_a\alpha}\,(\partial_\mu + iq_aA_\mu - iq_a\partial_\mu\alpha + iq_a\partial_\mu\alpha)\,\Psi_a = e^{iq_a\alpha}D_\mu\Psi_a. \tag{7.7}$$

$D_\mu\Psi_a$ is called *covariant derivative* of the field Ψ_a and has the important property that transforms as the field itself. At this point is straightforward to construct invariant Lagrangians. Let's see two examples.

Fermions

$$\begin{aligned}\mathcal{L} &= i\underbrace{\overline{\Psi}_a}_{e^{-iq_a\alpha}}\gamma^\mu\underbrace{D_\mu\Psi_a}_{e^{iq_a\alpha}} - m\overline{\Psi}_a\Psi_a = \overline{\Psi}_a(i\partial\!\!\!/ - m)\Psi_a + J_\mu A^\mu,\\ J^\mu &\equiv -q_a\overline{\Psi}_a\gamma^\mu\Psi_a \equiv q_aJ^\mu_{Noether}.\end{aligned} \tag{7.8}$$

Scalars

$$\begin{aligned}\mathcal{L} &= \underbrace{(D_\mu\Phi_a)^\dagger}_{e^{-iq_a\alpha}}\underbrace{(D^\mu\Phi_a)}_{e^{iq_a\alpha}} - m^2\Phi_a^\dagger\Phi_a = \partial_\mu\Phi_a^\dagger\partial^\mu\Phi_a + J_\mu A^\mu + q_a^2\Phi_a^\dagger\Phi_aA_\mu A^\mu - m^2\Phi_a^\dagger\Phi_a,\\ J_\mu &\equiv -iq_a\left(\Phi_a^\dagger\partial_\mu\Phi_a - \Phi_a\partial_\mu\Phi_a^\dagger\right) \equiv q_aJ^\mu_{Noether}.\end{aligned} \tag{7.9}$$

Notice that in the special case $\alpha =$ constant, the transformations (7.5) reduces to a global phase rotation. Therefore a necessary condition for gauge invariance is invariance under global phase rotations, which correspond to a $U(1)$ symmetry. This implies, via Noether theorem, the existence of a conserved current J_μ. This current is exactly the

EM current that appears in Maxwell equations and is the one called J_μ in the previous examples. Indeed from a Lagrangian of the form

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \mathcal{L}_{\text{matter}}, \tag{7.10}$$

one can derive the equations of motion $\partial_\mu F^{\mu\nu} = J^\nu$, where $J^\nu \equiv \partial\mathcal{L}_{\text{matter}}/\partial A_\nu$. The antisymmetric nature of $F^{\mu\nu}$ indeed also implies the conservation of the current:

$$\partial_\nu\partial_\mu F^{\mu\nu} = 0 = \partial_\nu J^\nu. \tag{7.11}$$

Current conservation, in this case, is both a consequence of Maxwell equations and of Noether theorem. In order to understand why the EM currents coincides with the Noether current associated to the global phase rotation symmetry, one could also recall the discussion in the Exercise Set 9 (Exercise 2). Given a matter Lagrangian invariant under a global symmetry one can formally perform a transformation where the parameters $q_a\alpha^a$ of the transformation are spacetime dependent. In this case the Lagrangian is, in general, no longer invariant and we have:

$$\delta\mathcal{L}_{\text{matter}} = J^\mu_a\partial_\mu\alpha_a\,, \tag{7.12}$$

where J^μ_a is exactly the Noether current that one would obtain with the standard procedure. In order to construct a Lagrangian $\mathcal{L}'_{\text{matter}}$ invariant under gauge symmetry, one can start adding to $\mathcal{L}_{\text{matter}}$ the term $A_\mu J^\mu$ so that under gauge transformations we have

$$\delta\mathcal{L}'_{\text{matter}} \equiv \delta(\mathcal{L}_{\text{matter}} + A_\mu J^\mu) = \delta\mathcal{L}_{\text{matter}} + J^\mu\delta A_\mu + A_\mu\delta J^\mu. \tag{7.13}$$

For simplicity let us assume that $\delta J^\mu = 0$. Then using (7.12) and (7.3) one gets $\delta\mathcal{L}'_{\text{matter}} = 0$. The assumption $\delta J^\mu = 0$ is true for the Noether current of fermions but not for the one for scalars. In the second case a further step is needed but this will not affect the discussion. We see than that the current $J^\nu = \partial\mathcal{L}'_{\text{matter}}/\partial A_\nu$ appearing in the Maxwell equation coincides with the Noether current extracted from the gauge invariant Lagrangian $\mathcal{L}'_{\text{matter}}$.

7.2 Exercise

Consider the following two Lagrangians:

$$\mathcal{L}_{QED} = i\bar\psi\not{D}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{7.14}$$

$$\mathcal{L}_{sQED} = (D_\mu\phi)^\dagger D^\mu\phi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{7.15}$$

The above Lagrangians are invariant under the transformations

$$\psi(x) \longrightarrow \psi'(x) = e^{i\alpha(x)}\psi(x) \implies \delta\psi(x) = ie\alpha(x)\psi(x), \tag{7.16}$$

$$\phi(x) \longrightarrow \phi'(x) = e^{i\alpha(x)}\phi(x) \implies \delta\phi(x) = ie\alpha(x)\phi(x), \tag{7.17}$$

$$A_\mu(x) \longrightarrow A'_\mu(x) = A_\mu(x) - \partial_\mu\alpha(x). \tag{7.18}$$

With these conventios the covariant derivative is defined as $D_\mu = \partial_\mu + ieA_\mu$. One can expand these Lagrangians:

$$\mathcal{L}_{QED} = i\bar\psi\not{\partial}\psi - e\bar\psi\gamma^\mu\psi\,A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{7.19}$$

$$\mathcal{L}_{sQED} = \partial_\mu\phi^\dagger\partial^\mu\phi + ie\left(\partial^\mu\phi^\dagger\phi - \phi^\dagger\partial^\mu\phi\right)A_\mu + e^2\phi^\dagger\phi A_\mu A^\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{7.20}$$

The first term of both the Lagrangians coincides with the original "ungauged" Lagrangian, that is to say invariant under the *global* $U(1)$ transformations (those with α = constant). The additional terms are indeed required to achieve the invariance under *local* transformations (those with $\alpha = \alpha(x)$). We recall the Noether currents associated to the global $U(1)$ obtained from the old Lagrangians:

$$\mathcal{L}^{old}_{QED} = i\bar\psi\not{\partial}\psi \implies J^\mu_{Noether} = \frac{\partial\mathcal{L}^{old}_{QED}}{\partial(\partial_\mu\psi)}\Delta_\psi = -\bar\psi\gamma^\mu\psi, \tag{7.21}$$

$$\mathcal{L}^{old}_{sQED} = \partial_\mu\phi^\dagger\partial^\mu\phi \implies J^\mu_{Noether} = \frac{\partial\mathcal{L}^{old}_{sQED}}{\partial(\partial_\mu\phi)}\Delta_\phi + \frac{\partial\mathcal{L}^{old}_{sQED}}{\partial(\partial_\mu\phi^\dagger)}\Delta_{\phi^\dagger} = i\left(\partial^\mu\phi^\dagger\phi - \phi^\dagger\partial^\mu\phi\right). \tag{7.22}$$

If we consider the linear term in A_μ inside $\mathcal{L}_{QED}$, $\mathcal{L}_{sQED}$ we notice that the combination of fields appearing are exactly the obtained currents. The motivation for this is related to how the gauge invariant Lagrangian has been constructed. Let us repeat briefly the steps:

- We start with a Lagrangian invariant under a global symmetry: $\delta\mathcal{L} = 0$.
- We consider a transformation of the fields with parameters depending on the coordinates $e\alpha(x)$; in this case the variation of the Lagrangian induced by the transformation will not be zero. We have indeed shown that $\delta\mathcal{L} = eJ^\mu_{Noether}\partial_\mu\alpha(x)$.
- In order to make the Lagrangian invariant we add a term of the form $eJ^\mu_{Noether}A_\mu$. Then the variation of the new Lagrangian:

$$\delta(\mathcal{L} + eJ^\mu_{Noether}A_\mu) = eJ^\mu_{Noether}\partial_\mu\alpha(x) - eJ^\mu_{Noether}\partial_\mu\alpha(x) + e\,A_\mu\delta J^\mu_{Noether} \tag{7.23}$$

- If the Noether current is gauge invariant then the procedure ends here. This is the case for fermions. The obtained Lagrangian is the covariantized Dirac Lagrangian. In the case of scalars the Noether current is not gauge invariant (since it contains derivatives) therefore one should iterate the procedure and add further terms to the Lagrangian. In the particular case of scalars indeed

$$eA_\mu\delta J^\mu_{Noether} = ieA^\mu\left(-i\phi^\dagger\phi\partial_\mu\alpha(x) - i\phi^\dagger\phi\partial_\mu\alpha(x)\right) = 2eA^\mu\phi^\dagger\phi\,\partial_\mu\alpha(x) \tag{7.24}$$

  Hence adding to the Lagrangian a term like $e^2A_\mu A^\mu\phi^\dagger\phi$ we can cancel this variation. Finally one observe that no other source of non-invariance are introduced. Therefore the sum $\mathcal{L}^{old}_{sQED} + eA_\mu\delta J^\mu_{Noether} + e^2A_\mu A^\mu\phi^\dagger\phi$ is gauge invariant.

This explains why the gauge vector couples to the Noether current of the global symmetry: we have engineered the construction in this way. Finally let us compute the Noether current associated to the global symmetry starting from the gauge invariant Lagrangian:

$$J^\mu_{QED} = \frac{\partial\mathcal{L}_{QED}}{\partial(\partial_\mu\psi)}\Delta_\psi = -\bar\psi\gamma^\mu\psi, \tag{7.25}$$

$$J^\mu_{sQED} = \frac{\partial\mathcal{L}_{sQED}}{\partial(\partial_\mu\phi)}\Delta_\phi + \frac{\partial\mathcal{L}_{sQED}}{\partial(\partial_\mu\phi^\dagger)}\Delta_{\phi^\dagger} = i\left(\partial^\mu\phi^\dagger\phi - \phi^\dagger\partial^\mu\phi\right) + 2eA^\mu\phi^\dagger\phi = i\left((D^\mu\phi)^\dagger\phi - \phi^\dagger D^\mu\phi\right). \tag{7.26}$$

The new currents are the equal to the old ones where we substitute normal derivatives with covariant derivatives and coincide exactly with the currents appearing in the Maxwell equations of motion:

$$J^\mu_{QED} = \frac{\partial\mathcal{L}_{QED}}{\partial A_\mu}\,, \qquad J^\mu_{sQED} = \frac{\partial\mathcal{L}_{sQED}}{\partial A_\mu}\,. \tag{7.27}$$

7.3 A fresh look at Maxwell equations: counting degrees of freedom

Before approaching the Hamiltonian description of the EM field and its quantization, it is worth having a fresh look at the structure of its (classical) dynamics, and in particular at the counting of its dynamical degrees of freedom. For that purpose it is useful to recall that, given any vector field $V^i(\vec{x})$ on an euclidean space, we can decompose it into longitudinal and transverse components $V^i = V^i_L + V^i_\perp$ according to

$$V^i_L \equiv \Pi^{ij}_L V^j \equiv \frac{\nabla^i\nabla^j}{\nabla^2}V^j\,, \quad V^i_\perp \equiv \Pi^{ij}_\perp V^j \equiv \left(\delta^{ij} - \frac{\nabla^i\nabla^j}{\nabla^2}\right)V^j\,, \qquad \Longrightarrow \qquad \nabla^iV^i_L = \nabla^iV^i\,, \;\; \nabla^iV^i_\perp = 0. \tag{7.28}$$

In R^D, with $i = 1, \ldots, D$, the transverse and longitudinal fields have respectively $D-1$ and 1 components. Therefore in our R^3 case, $A_\perp$ is a vector determined by two independent functions.

In our case we can package the four components in A^μ as $(A^0, A^i_L, A^i_\perp)$. Notice also that $A^i_\perp$ (fully determining F^{ij}) and $\nabla^2A^0 + \nabla^i\dot{A}^i$ $(= \partial_iF^{0i})$ are gauge invariant, while any other linear combination of A^0 and A^i_L (for instance just A^0 or A^i_L) is gauge variant.

Consider now the equations of motion for A_μ that are determined by the general lagrangian in eq. (7.10)

$$\frac{\delta S}{\delta A_\mu} = 0 \;\Rightarrow\; \partial_\mu F^{\mu\nu} = J^\mu \,. \tag{7.29}$$

The current J^μ depends on the matter fields and their covariant derivatives (the latter only if there are scalars), but not on derivatives of A_μ. The derivatives of A_μ only appear in the lagrangian via the $F^2_{\mu\nu}$ term. Now, the structure of $F_{\mu\nu}$ is such that the lagrangian does not depend on the time derivative of A_0, indicating that this variable is not dynamical. Indeed its equation of motion reads [1]

$$0 = \frac{\delta S}{\delta A_0} \;\Longrightarrow\; \partial_i F^{i0} = -\nabla^2 A^0 - \nabla^i \dot{A}^i = J^0 \,. \tag{7.30}$$

This equation, which is just the familiar *Gauss Law* ($F^{i0} \equiv E^i$), involves only terms with at most one time derivative acting on the fields. [2] It therefore represents a *constraint*, rather than a standard dynamical equation with second order time derivatives. In particular, initial conditions must satisfy this constraint! This equation can be solved to express, at any time, one variable in terms of the others. In particular it is convenient to solve for A^0

$$A^0 = -\nabla^{-2}\left(J^0 + \nabla^i \dot{A}^i\right) \,. \tag{7.31}$$

This shows that A^0 is not an independent dynamical variable [3]. One out of the four field variables in A_μ is thus "eliminated" by *Gauss law*.

Consider next the equation for A^i, that reads

$$\partial_0 F^{0i} + \partial_i F^{ij} = \nabla^i \dot{A}^0 + \ddot{A}^i - \nabla^2 A^i - \nabla^i \nabla^j A^j = J^i \,. \tag{7.32}$$

Using current conservation and eq. (7.31) this equation can be written purely in terms of transverse fields

$$(\partial_t^2 - \nabla^2) A^i_\perp = J^i_\perp \,. \tag{7.33}$$

This is a consequence of the gauge invariance of eq. (7.32): $A^i_\perp$ is gauge invariant while A^i_L is not, so that once eq. (7.31) is used to eliminate A^0, the final equation can only depend on $A^i_\perp$. Eq. (7.33) thus describes the dynamics of two d.o.f.: the two polarizations of light waves. At this stage we have not yet made any use of our freedom to fix the gauge, and the single remaining field to play with is A^i_L. Various options are given.

- Coulomb gauge: $\nabla_i A^i = 0 \iff A^i_L = 0$. For the purpose of quantization this is the best choice as it allows for the simplest Hamiltonian in terms of the 2 physical d.o.f. The problem with this choice is the lack of manifest Lorentz invariant. Of course the final result of any computation respects Lorentz invariance, but that is not manifest in the intermediate stages.
- Temporal gauge: $A^0 = 0 \iff J^0 + \nabla^i \dot{A}^i_L = 0$, slighly less simple from the Hamiltonian perspective and also not manifestly Lorentz invariant.
- Lorentz gauge $\partial_\mu A^\mu = 0 \iff \dot{A}^0 + \nabla^i A^i_L = 0$; less straightforward Hamiltonian description but manifestly Lorentz covariant.
- Axial gauge: $A_3 = 0 \iff A^3_L = -A^i_\perp$; not manifestly rotationally invariant.

In the end the counting of degrees of freedom is already done: one linear combination of A^0 and A^i_L is made non dynamical by Gauss law, another can be eliminated by gauge fixing, so that only two dynamical degrees of freedom, which we can parameterize with $A^i_\perp$, are left.

[1] For ∇^i we use the euclidean metric while for ∂_i we use the Minkowski one: $\nabla^i \equiv (\partial/\partial x^i) \equiv \partial_i = -\partial^i$.

[2] From a mechanical perspective we ould say it only depends on the q_i (coordinates), the $\dot{q}_i$ (velocities), but not on $\ddot{q}_i$ (accelerations).

[3] Notice that ∇^2 only involves space derivative, so it inverse ∇^{-2} relates quantities at the same time, though not at the same space position. The inverse $\Box^{-1}$ of the D'alembertian would be a different story, as it relates quantities at different times (cfr. for instance the retarded or advanced Green's functions in classical electrodynamics).

7.4 A first step towards quantization

The simplicity (and beauty) of gauge theories implies however some technical difficulty when deriving the Hamiltonian description of their dynamics. In turn, since canonical quantization is performed in the Hamiltonian formulation, this procedure is more involved than for scalar or fermionic fields. The origin of these complications is precisely the variable redundance discussed in the previous section: the vector field A_μ is determined by four local variables, but only two of them are dynamical. Let us see how this issues show up when trying to derive the Hamiltonian description for the general interacting case in eq. (7.10).

Consider now the conjugated momenta as defined by eq. (7.10):

$$\Pi^0 \equiv \frac{\partial \mathcal{L}}{\partial(\partial_0 A_0)} = 0\,, \tag{7.34}$$

$$\Pi^i \equiv \frac{\partial \mathcal{L}}{\partial(\partial_0 A_i)} = -F^{0i} = E^i. \tag{7.35}$$

We have already discussed in the previous section the non-dynamical status of A^0. The first equation expresses that fact in Hamiltonian languange: the conjugate momentum Π^0 vanishes identically! This confirms that (Π^0, A_0) does not form an ordinary pair of canonical Hamiltonian variables. But the peculiarity of A_0 also determines the peculiarity of its equation of motion, Gauss law, that, as we have already seen, is a constraint rather than a dynamical equation. Indeed, using eq. (7.35), eq. (7.30) can be rewritten as

$$-\nabla^i \Pi^i = J^0\,, \tag{7.36}$$

which represents a constraint on the Hamiltonian variables. In particular it relates the space divergence of the conjugated momentum Π^i to other Hamiltonian variables, associated with matter fields, and appearing in J^0. By this equation, the quantity $\nabla^i \Pi^i \equiv \nabla^i \Pi^i_L$ is not an independent dynamical variable. This indicates that a second Hamiltonian pair (A^i_L, Π^i_L) should be eliminated in addition to (A^0, Π^0), and thus reduces the number of genuine local degrees of freeedom down to two. The elimination of A^i_L is achieved via a choice of gauge, rather then by an equation of motion. Because of the above constraint, we cannot proceed to a straightforward quantization since a canonical commutation relation

$$[A^i(\vec{x}, t), \Pi^j(\vec{y}, t)] = i\delta^{ij}\delta^3(\vec{x} - \vec{y}) \tag{7.37}$$

would not be in agreement with Gauss Law. In order to proceed we must eliminate the gauge redundancy of the A_μ field by "choosing a gauge". That corresponds to making a suitable transformation (7.3) in such a way that some linear combination of A^μ and its first derivatives vanishes. Some examples are:

- Coulomb gauge: $\partial_i A^i = 0$; it allows for the simplest Hamiltonian in terms of the 2 physical d.o.f but is not explicitly Lorentz invariant.
- Axial gauge: $A_3 = 0$; not rotational invariant.
- Lorentz gauge $\partial_\mu A^\mu = 0$; less straightforward Hamiltonian description but manifestly Lorentz covariant.

7.5 Quantization à la Gupta-Bleuler

In order to proceed to quantization in a way that is Lorentz invariant and allows the use of a simple Lagrangian, we will quantize the EM field using a different procedure: we will start with a Lagrangian that is *inequivalent* to the Maxwell Lagrangian but for which the the mentioned conditions are satisfied. The equivalence to the Maxwell theory will be obtained at the end selecting only the states $|\psi\rangle$ for which the following gauge condition holds:

$$\langle\psi|\partial_\mu A^\mu|\psi\rangle = 0. \tag{7.38}$$

We will discuss extensively the meaning and the origin of the above constraint. The method we are going to discuss was first proposed by Fermi and later formalized by Gupta and Bleuler.

We have seen in the past lecture that gauge invariance has the effect of reducing the number of dynamical degrees of freedom in the theory and this introduces complication in the definition of the commutation relations. The idea

is to make all the variables in A_μ *formally* dynamical by adding an explicit breaking of the gauge invariance to the theory so that our starting Lagrangian is

$$\mathcal{L}_{GB} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{\xi}{2}(\partial_\mu A^\mu)^2 = -\frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu + \frac{1-\xi}{2}(\partial_\mu A^\mu)^2, \tag{7.39}$$

where in the second equality we have integrated by part a term. The parameter ξ is a *free* parameter. The choice $\xi = 1$, that simplifies a lot the form of the Lagrangian, is called Fermi's or Feynman's choice.

Now the action depends on all 4 d.o.f A_μ. In practice we have added a quadratic term for the component $A_\mu \sim \partial_\mu \alpha$ which, being a pure gauge, in the Maxwell equation has an exactly vanishing action. For all the values $\xi \neq 0$ we can straightforwardly compute the conjugate momenta:

$$\frac{\partial \mathcal{L}}{\partial \dot{A}_\mu} \neq 0 \quad \text{for any } \mu. \tag{7.40}$$

The equations of motion are analogously modified with respect to the Maxwell equations:

$$\partial_\mu F^{\mu\nu} + \xi \partial^\nu (\partial_\mu A^\mu) = J^\nu. \tag{7.41}$$

Let us define the scalar field $\chi = \partial_\mu A^\mu$. As we can observe from the above equation, the gradient of χ acts as a source for the EM fields, therefore the system is not in general equivalent to the Maxwell theory. On the other hand if we take the divergence of eq. (7.41) we get

$$\xi \Box \chi = \underbrace{\partial_\nu J^\nu}_{=0 \text{ by conservation}} - \underbrace{\partial_\nu \partial_\mu F^{\mu\nu}}_{=0 \text{ by antisymmetry}} = 0. \tag{7.42}$$

The above relation implies that the field χ is free, since nothing acts as source in its equation: the χ waves come and go without being affected by anything. This doesn't mean that this field is completely decoupled because we have seen that the equation of motion for $\vec{E}$ and $\vec{B}$ are influenced by its presence. The important consequence of (7.42) is that it is consistent to choose $\chi \equiv 0$ at all times since if it was vanishing at $t = -\infty$ it will not be generated later on by the dynamics. When this is the case (7.41) goes back to Maxwell equations at all times. This is a intuitive justification of the condition (7.38).

Let us then proceed with the canonical quantization and let us work in the Feynman gauge $\xi = 1$. The Lagrangian

$$\mathcal{L}_{GB} = -\frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu \tag{7.43}$$

is equivalent to 4 scalars $(A^0,\, A^i)$, three of which have the proper kinetic term while one has a negative kinetic term. As consequence, the energy of the system will not be positive definite. We can introduce canonical commutation relations

$$\begin{aligned} &[A_\mu(\vec{x},t),\, \Pi^\nu(\vec{y},t)] = i\delta^\nu_\mu \delta^3(\vec{x}-\vec{y}), \qquad \Pi^\mu = -\dot{A}^\mu \\ &\big[A_\mu(\vec{x},t),\, \dot{A}_\nu(\vec{y},t)\big] = -i\eta_{\mu\nu}\delta^3(\vec{x}-\vec{y}), \end{aligned} \tag{7.44}$$

while all the other commutators are zero. Notice the sign flip between the commutator of the 0-component and the space components: it is a consequence of covariance.
The Hamiltonian is thus:

$$\mathcal{H} = \Pi^\mu \dot{A}_\mu - \mathcal{L} = -\frac{1}{2}\left(\Pi_0^2 + (\partial_i A^0)^2\right) + \frac{1}{2}\left(\Pi^j \Pi^j + (\partial_i A^j)(\partial_i A^j)\right) = -\mathcal{H}_0 + \sum_{i=1}^{3} \mathcal{H}_i, \tag{7.45}$$

The above Hamiltonian contains an oscillator with negative energy. Working in the Fourier space we can define, as we did for the Klein-Gordon fields, the expansion at a fixed time ($t = 0$):

$$\begin{aligned} &\widetilde{A}_\mu(\vec{k}) = \int d^3x\; A_\mu(\vec{x})\, e^{-i\vec{k}\cdot\vec{x}}, \qquad \widetilde{\Pi}(\vec{k})^\mu = \int d^3x\; \Pi^\mu(\vec{x})\, e^{-i\vec{k}\cdot\vec{x}}, \qquad w_k \equiv |k|\,, \\ &a_\mu(\vec{k}) = w_k \widetilde{A}_\mu(\vec{k}) - i\widetilde{\Pi}_\mu(\vec{k}), \\ &a_\mu(\vec{k})^\dagger = w_k \widetilde{A}_\mu(-\vec{k}) + i\widetilde{\Pi}_\mu(-\vec{k}). \end{aligned} \tag{7.46}$$

Notice the $(-)$ sign in parenthesis in the second equation to reproduce the right combination as for the Klein-Gordon fields in the case of spatial components. Indeed the right conjugate momentum of A_i is $-\Pi_i$. Following the same computation as in the Klein-Gordon case we can show that

$$\begin{aligned}\left[a_\mu(\vec{k}), a_\nu(\vec{p})\right] &= \left[a^\dagger_\mu(\vec{k}), a^\dagger_\nu(\vec{p})\right] = 0\\ \left[a_\mu(\vec{k}), a^\dagger_\nu(\vec{p})\right] &= -\eta_{\mu\nu}\delta^3(\vec{k}-\vec{p})(2\pi)^3 2w_k\end{aligned} \tag{7.47}$$

An important outcome is that the time photon (a_0) has "wrong" sign. The inverse relation of (7.46) has the form

$$A_\mu(\vec{x},0) = \int d\Omega_{\vec{k}} \left(a_\mu(\vec{k})\, e^{+i\vec{k}\cdot\vec{x}} + a^\dagger_\mu(\vec{k})\, e^{-i\vec{k}\cdot\vec{x}}\right). \tag{7.48}$$

At this point one can plug the above expansion in the expression of the energy-momentum tensor and derive the associated conserved charge P_μ (see Exercise Set 17):

$$P^\mu = \int d\Omega_{\vec{k}}\, k^\mu \left(a^\dagger_i(\vec{k})a_i(\vec{k}) - a^\dagger_0(\vec{k})a_0(\vec{k})\right) = \int d\Omega_{\vec{k}}\, k^\mu \left(-a^\dagger_\nu(\vec{k})a^\nu(\vec{k})\right). \tag{7.49}$$

Notice that we can use the expression of A_μ at $t=0$ since the conserved charges like P^μ are independent of time. At this point we can compute the vector potential in the Heisenberg picture

$$A_\mu(\vec{x},t) = e^{iHt}\, A_\mu(\vec{x},0)\, e^{-iHt}. \tag{7.50}$$

For the spatial components the computation is the standard one. For the time component one should pay attention to the $(-)$ signs in (7.45) and (7.47). These two factors $(-)$ compensate and the result is simply

$$A_\mu(\vec{x},t) = \int d\Omega_{\vec{k}} \left(a_\mu(\vec{k})\, e^{-ikx} + a^\dagger_\mu(\vec{k})\, e^{ikx}\right). \tag{7.51}$$

The result is consistent with Lorentz invariance: since A_μ is a four-vector, the integral measure and the exponential in the integral are scalars, hence the operator $a_\mu(\vec{k})$ must be a four-vector. This is also in accord with the commutation relations (7.47).

We are now ready to construct the Fock space of the theory. Insisting on Lorentz covariance we define the vacuum as the state annihilated by all destroying operators

$$a_\mu(\vec{k})|0\rangle = 0\,, \qquad \text{for any } \vec{k}\,. \tag{7.52}$$

The above requirement is surely consistent with the Lorentz invariance of the vacuum, $U(\Lambda)|0\rangle = |0\rangle$, but, as we will see in the following, it's a stronger requirement.
At this point the Fock space is constructed applying repeatedly the operators $a^\dagger_\mu(\vec{k})$. However we immediately encounter a serious issue: if we compute the norm of a single particle state $a^\dagger_\mu(\vec{k})|0\rangle$ we get

$$\langle 0|a_\mu(\vec{k})a^\dagger_\mu(\vec{k})|0\rangle = -2k_0(2\pi)^3\eta_{\mu\mu}, \qquad \text{no summation over } \mu. \tag{7.53}$$

Therefore for $\mu = i$ spatial the above norm is positive, while $a^\dagger_0(\vec{k})|0\rangle$ is a state with negative norm. This pathology, if not solved, would compromise the probabilistic interpretation of Quantum Mechanics. On the other hand we should recall that the system we are considering doesn't coincide with Maxwell's. As we saw at the very beginning, the equivalence can be achieved imposing the condition $\partial_\mu A^\mu = 0$. This constraint has been derived at the classical level in order to reproduce Maxwell equations. At the quantum level the field A_μ is promoted to an operator and we cannot impose this condition naively at the operator level since it would be in conflict with the commutation relations (7.44). What we can do instead is to impose the vanishing of the expectation value of $\partial_\mu A^\mu$. More precisely, given the Fock space $\mathcal{F}$ we define the subset $\mathcal{F}'$ of *physical states* as the states $|\Psi\rangle$ satisfying

$$\langle\Psi|\,\partial_\mu A^\mu|\Psi\rangle = 0. \tag{7.54}$$

We stress that the above condition is not a constraint on the field A_μ, which is unconstrained, but a restriction of the states of $\mathcal{F}$: only a subset of them is selected.
Defining

$$\begin{aligned}&\partial^\mu A_\mu = \partial^\mu A^+_\mu + \partial^\mu A^-_\mu\,,\\ &\partial^\mu A^+_\mu = \int d\Omega_{\vec{k}}(ik^\mu)a^\dagger_\mu(\vec{k})e^{ikx}\,, \qquad \partial^\mu A^-_\mu = \int d\Omega_{\vec{k}}(-ik^\mu)a_\mu(\vec{k})e^{-ikx},\end{aligned} \tag{7.55}$$

a sufficient condition for (7.54) is that $\partial^\mu A^-_\mu|\Psi\rangle = 0$. We can also define $L(\vec{k}) \equiv k^\rho a_\rho(\vec{k})$ as the Fourier transform of the operator $\partial^\mu A^-_\mu$. Then we will postulate that a physical state is a state that satisfies the condition

$$L(\vec{k})|\Psi\rangle = 0 \quad \text{or equivalently} \quad \langle\Psi|L^\dagger(\vec{k}), \qquad \text{for any } \vec{k}\,. \tag{7.56}$$

Let us see what this implies for a single particle state $|\Psi\rangle \sim \varepsilon^\mu a^\dagger_\mu(\vec{q})|0\rangle$. Imposing the above condition and using the commutation relations (7.47) we get

$$L(\vec{q})|\Psi\rangle = q^\rho a_\rho(\vec{q})\varepsilon^\mu a^\dagger_\mu(\vec{k})|0\rangle = -2k_0(2\pi)^3\delta^3(\vec{k}-\vec{q})k^\mu\varepsilon_\mu|0\rangle = 0. \tag{7.57}$$

Here ε_μ is the polarization of the photon and for a physical state it has to satisfy the transversality condition $k^\mu\varepsilon_\mu = 0$. This means that the number of physical polarization is reduced by one. However we will now show that the vector ε_μ contains in fact only 2 physical polarizations. Indeed the general expression for the polarization vector satisfying $k^\mu\varepsilon_\mu = 0$ is

$$\varepsilon_\mu = \varepsilon^\perp_\mu + a_L k_\mu \qquad \varepsilon^\perp_\mu = c_1\varepsilon^1_\mu + c_2\varepsilon^2_\mu, \tag{7.58}$$

where 1 and 2 are the directions transverse w.r.t $\vec{k}$. For example if $k_\mu = (k,0,0,k)$ and $\varepsilon_\mu = (a,b,c,d)$ we have:

$$k^\mu\varepsilon_\mu = 0 \;\Rightarrow\; a = d, \qquad \varepsilon^\perp_\mu = (0,b,c,0)\,,\; a_L = \frac{a}{k}. \tag{7.59}$$

The term proportional to k_μ is called *longitudinal polarization* and its presence doesn't affect any physical quantity. Instead $\varepsilon^\perp_\mu$ is called transverse polarization and satisfies $\vec{\varepsilon}^\perp \cdot \vec{k} = 0$. The first check of the irrelevance of the longitudinal polarization is given by the computation of the norm of a physical state:

$$\langle\Psi|\Psi\rangle = \int d\Omega_{\vec{p}}d\Omega_{\vec{q}}\;\varepsilon^{\mu*}(\vec{p})\varepsilon^\nu(\vec{q})\langle 0|a_\mu(\vec{p})a^\dagger_\nu(\vec{q})|0\rangle = -\int d\Omega_{\vec{p}}\;\varepsilon^*_\mu(\vec{p})\varepsilon^\mu(\vec{p}) = \int d\Omega_{\vec{p}}|\vec{\varepsilon}^\perp(\vec{p})|^2 \geq 0. \tag{7.60}$$

In the last equality we have used $k_\mu k^\mu = 0$, $\epsilon^\perp_\mu k_\mu = 0$ and the fact that $\epsilon^\perp_\mu$ has zero time component. We conclude that the longitudinal polarization doesn't affect the norm of physical states: its only a redundancy of the theory. One can check that the same holds for any other physical quantity, for example energy and momentum:

$$\begin{aligned}\langle\Psi|P^\mu|\Psi\rangle &= -\int d\Omega_{\vec{p}}d\Omega_{\vec{q}}d\Omega_{\vec{k}}\;\varepsilon^{\rho*}(\vec{p})\varepsilon^\nu(\vec{q})k^\mu\langle 0|a_\rho(\vec{p})a^\dagger_\sigma(\vec{k})a^\sigma(\vec{k})a^\dagger_\nu(\vec{q})|0\rangle \\ &= \int d\Omega_{\vec{p}}d\Omega_{\vec{k}}\varepsilon^{\rho*}(\vec{p})\varepsilon^\nu(\vec{k})k^\mu\langle 0|a_\rho(\vec{p})a^\dagger_\nu(\vec{k})|0\rangle = \int d\Omega_{\vec{p}}\;p^\mu|\vec{\varepsilon}^\perp(\vec{p})|.\end{aligned} \tag{7.61}$$

Again the longitudinal polarizations doesn't give any contribution. This property is a consequence of the cancellation between the time photon $a_0(\vec{k})$ and the 3rd photon $a_3(\vec{k})$ (in general the photon aligned along $\vec{k}$). This cancellation is enforced by the condition of being physical, as can be seen by considering the square of the vanishing state $L(k)|\Psi\rangle$; taking again $k_\rho = (k,0,0,k)$ we have

$$0 = L(k)|\Psi\rangle \implies ka_0(\vec{k})|\Psi\rangle + ka_3(\vec{k})|\Psi\rangle \tag{7.62}$$

$$\langle\Psi|a^\dagger_0(\vec{k})a_0(\vec{k})|\Psi\rangle = \langle\Psi|a^\dagger_3(\vec{k})a_3(\vec{k})|\Psi\rangle. \tag{7.63}$$

The above relation expresses the fact that, for a physical state, the occupation number of the time photon is equal to the occupation number of 3rd component photons. Hence they are always equal in number and their contribution to observables quantities always cancel. Only transverse polarizations contribute to physical quantities.

Let us discuss more in details the form of the space of physical states. Given a physical state $|\Psi\rangle$ and a second (even infinite) set of other physical states $\{|\varphi_1\rangle, |\varphi_2\rangle, ..., |\varphi_n\rangle\}$ we construct the state

$$|\Psi'\rangle = |\Psi\rangle + \int d\Omega_{k_1}f_1(k_1)L^\dagger(k_1)|\varphi_1\rangle + \int d\Omega_{k_1}d\Omega_{k_2}f_2(k_1,k_2)L^\dagger(k_1)L^\dagger(k_2)|\varphi_2\rangle + \dots\,. \tag{7.64}$$

It's easy to check that the above combination is a physical state: $L(\vec{k})|\Psi'\rangle = 0$; indeed, by $k^\mu k_\mu = 0$, one has

$$[L(\vec{k}_1), L^\dagger(\vec{k}_2)] = k^\mu_1 k^\nu_2[a_\mu(\vec{k}_1), a^\dagger_\nu(\vec{k}_2)] = -(2\pi)^3 2k^0_1\delta^3(\vec{k}_1-\vec{k}_2)k^\mu_1 k_{2\mu} = 0 \tag{7.65}$$

and using that the $|\varphi_i\rangle$ are physical states one has

$$L(\vec{q})\int d\Omega_{\vec{k}_1}f_1(\vec{k}_1)L^\dagger(\vec{k}_1)|\varphi_1\rangle = \int d\Omega_{\vec{k}_1}f_1(\vec{k}_1)L^\dagger(\vec{k}_1)L(\vec{q})|\varphi_1\rangle = 0\,, \tag{7.66}$$

and similarly for the other terms. However, not only $|\Psi'\rangle$ is physical but it contributes to all the physical quantities in the same way as $|\Psi\rangle$. Indeed the states $L(k_1)^\dagger ... L(k_i)^\dagger|\varphi_i\rangle$ are all states with null norm and are orthogonal to any other physical state. This means that the norm of $|\Psi\rangle$ and $|\Psi'\rangle$ are identical and the scalar product of the two states with any other physical state coincide:

$$\langle\Psi|\Psi\rangle = \langle\Psi'|\Psi'\rangle \qquad \langle\Phi|\Psi\rangle = \langle\Phi|\Psi'\rangle\,, \quad \text{for any } |\Phi\rangle \text{ physical.} \tag{7.67}$$

The above relations imply that the physical content of the tho states $|\Psi\rangle$ and $|\Psi'\rangle$ is identical and therefore they describe the same state. This means that we can define an *equivalence relation* on the Hilbert space of physical states and consider equivalent two states if they coincide modulus a combination of states of the form $L(k_1)^\dagger ... L(k_i)^\dagger|\varphi_i\rangle$:

$$|\Psi\rangle \sim |\Psi\rangle + L^\dagger(k_1)|\varphi_1\rangle + L^\dagger(k_1)L^\dagger(k_2)|\varphi_2\rangle + ... + L^\dagger(k_1)...L^\dagger(k_n)|\varphi_n\rangle + ...\,. \tag{7.68}$$

In the end any physical state is defined by a tower of equivalent states; in particular also the physical vacuum is described by all the set of combinations

$$|0\rangle + a_1 L^\dagger(k_1)|\varphi_1\rangle + a_2 L^\dagger(k_1)L^\dagger(k_2)|\varphi_2\rangle + ... + a_n L^\dagger(k_1)...L^\dagger(k_n)|\varphi_n\rangle + ...\,. \tag{7.69}$$

Let us now define in a more concrete way what is a physical observable. The basic requirement that must be satisfied is that the action of a physical observable on a physical state must produce a physical state. Then if O is an operator associated to a physical observable and $|\Psi\rangle$ is a physical state we must have

$$L(k)|\Psi\rangle = 0, \qquad L(k)O|\Psi\rangle = 0. \tag{7.70}$$

A sufficient condition to satisfy the above requirement is that the commutator $[L(k), O]$ be proportional to some operator that has an L on the right:

$$[L(k), O] \sim A\,L(k)\,, \qquad [O, L(k)^\dagger] \sim L(k)^\dagger A^\dagger. \tag{7.71}$$

It is clear that the action of O on a physical state produces again a physical state:

$$L(k)O|\Psi\rangle = [L(k), O]|\Psi\rangle = AL|\Psi\rangle = 0. \tag{7.72}$$

In addition this requirement ensures that the expectation value of an observable O on two physical states is independent of the vector chosen to represent the states:

$$\langle\Phi'|O|\Psi'\rangle \sim (\langle\Phi| + \langle\varphi_1|L(k_1) + ...)\,O\,\big(|\Psi\rangle + L(k_2)^\dagger|\varphi_2\rangle + ...\big) = \langle\Phi|O|\Psi\rangle. \tag{7.73}$$

where we have repeatedly made use of the commutation relations (7.71).
Here we have defined observables all the operator satisfying equation (7.70). This definition is a stronger request than simply requiring equation (7.73), that instead is verified whenever $[L(k), O] \sim AL(k) + L^\dagger(k)B$. The problem with this second definition is that is not preserved by the composition of observables. For instance the observable O^2 is not independent of the vector that represents a physical state. In practice however we don't have any simple example where B is different from 0.

7.6 The Massive Vector Field

Let us consider the Maxwell Lagrangian plus a real scalar φ:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}\partial_\mu\varphi\,\partial^\mu\varphi\,. \tag{7.74}$$

The above Lagrangian is invariant under the shift symmetry $\varphi \to \varphi + c$, as long as c is a constant. The field A_μ doesn't transform under this symmetry. The Noether current associated to this global symmetry is simply

$$J_\mu = \partial_\mu\varphi\,, \tag{7.75}$$

and the conservation of the current is guaranteed by the equation of motion of the scalar field: $\partial_\mu J^\mu = \Box\varphi = 0$. The above Lagrangian describes therefore 3 physical degrees of freedom: one is the real scalar field and the other two are described by the massless vector A_μ.

We can now try to make the shift symmetry of φ local, exactly as we did for the phase symmetry in the case of spinor fields and complex scalar fields. In that case, imposing the invariance under the new local symmetry, we got a Lagrangian describing the interaction between electromagnetism and matter. What shall we obtain in this case? The local transformation must have the form:

$$\varphi \to \varphi + M\alpha(x)\,. \tag{7.76}$$

Notice that we were forced to add a parameter M with energy (mass) dimension $[M] = E$, because the function $\alpha(x)$, describing the transformation of A_μ in eq. (7.3), is by consruction dimensionless and φ has mass dimension 1. The Lagrangian (7.6), as it stands, is not invariant under the tranformation 7.76, since derivatives act non-trivially on the function $\alpha(x)$. Again we can define a *new covariant derivative* involving the vector field A_μ to compensate for the additional terms. One can easily check that the covariant derivative

$$D_\mu\varphi \equiv \partial_\mu\varphi + MA_\mu \tag{7.77}$$

is invariant under the combined set of transformations

$$\varphi \to \varphi + M\alpha(x), \qquad A_\mu \to A_\mu - \partial_\mu\alpha(x)\,. \tag{7.78}$$

Indeed one has

$$D_\mu\varphi \equiv \partial_\mu\varphi + MA_\mu \longrightarrow \partial_\mu(\varphi + M\alpha) + M(A_\mu - \partial_\mu\alpha) = D_\mu\varphi\,. \tag{7.79}$$

so that the Lagrangian invariant under the local transformation 7.78 is

$$\mathcal{L}^{\rm inv} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}(\partial_\mu\varphi + MA_\mu)(\partial^\mu\varphi + MA^\mu)\,. \tag{7.80}$$

In order to more directly extract the physical meaning of this Lagrangian we must perform a gauge fixing. For this system there exists an obvious choice:

$$\alpha(x) = -\frac{\varphi(x)}{M}\,, \tag{7.81}$$

upon which φ simply vanishes! The resulting gauge fixed lagrangian

$$\mathcal{L}^{\rm fixed} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}M^2A_\mu A^\mu\,. \tag{7.82}$$

describes a free vector field of mass M, as we shall now discuss. In particle physics, this phenomenon, according to which a massless scalar shifting under gauge transformations implies that the photon obtains a mass, goes under the broad name of *Higgs mechanism*[4]. In condensed matter an analogous phenomenon leads to superconductivity. In that case the role of φ is played by the "phase" of the complex fermion bilinear field describing Cooper pairs. As already observed the original Lagrangian (7.6) describes 3 physical degrees of freedom, and it is a fair question to ask how many degrees of freedom the new Lagrangian (7.82) describes. One would argue that given one can go from to (7.82) by turning on a continuous parameter M, the number of degrees of freedom, an integer, should not change. It is instructive to check that is indeed the case. Let us then consider the equations of motion

$$\partial_\mu F^{\mu\nu} + M^2A^\nu = 0\,. \tag{7.83}$$

Taking the diverge we obtain

$$\partial_\nu\partial_\mu F^{\mu\nu} + M^2\partial_\nu A^\nu = M^2\partial_\nu A^\nu \implies M^2\partial_\nu A^\nu = 0\,. \tag{7.84}$$

The dynamics then forces the field A_μ to be transverse. Expanding the field strength $F_{\mu\nu}$ and dropping the vanishing terms, the equations of motion can then be cast as the combined set

$$(\Box + M^2)A^\nu = 0, \quad \partial_\nu A^\nu = 0 \tag{7.85}$$

The first corresponds to four independent Klein-Gordon equations describing fields with mass M, while the second constrains one linear combination of these four Klein-Gordon fields to vanish. The plane wave solutions have the form

$$A^\mu(\vec{x},t) = \varepsilon^\mu(\vec{k})e^{ik_0t - \vec{k}\cdot\vec{x}}, \qquad k_0 = \sqrt{|\vec{k}|^2 + M^2}. \tag{7.86}$$

[4]The mechanism was fully understood in relativistic QFT in the 60's. Notice that the famous Higgs boson does not correspond to the "unphysical" φ field here introduced but to another degree of freedom that accompanies the φ in theories endowed with additional properties, in particular renormalizability. The association of a photon mass to a scalar shifting under a local symmetry goes back to the 30's in the work of Stueckelberg.

with polarization vector $\varepsilon^\nu(k)$ constrained to be transverse by the second equation in (7.85)

$$k^\mu \varepsilon_\mu = 0\,. \tag{7.87}$$

This constraint implies that only 3 out of the 4 components of ε_μ are independent. For example in the special case of $\vec{k}=0$, $k_0=M$, the polarization is of the form $\varepsilon_\mu=(0,\vec{\varepsilon})$. We thus conclude that the massive vector field describes the propagation of 3 independent polarizations, precisely like the system we started with, for which the 3 polarizations where distributed among a massless vector (2 polarizations) and a massless scalar (1 polarization). The spin of the resulting particle is evident by considering the polarization 4-vector in the rest frame ($\vec{k}=0$). There it reduces to a 3-vector transforming as the $J=1$ representation of the group of rotations in the rest frame. Hence we conclude that the spin of the massive vector field A_μ is 1.

As usual the plane wave solution corresponds to an eigenstate of the 4-momentum, and the general solution is obtained by writing the most general superposition of plane waves.

7.6.1 Canonical quantization

We now proceed with the standard technique to quantize the theory: we first compute the conjugate momentum of the fields A_μ and the Hamiltonian, then we impose the canonical commutation relations. In the following we will work in the Schroedinger picture: the quantum states evolve in time while the operators do not. Hence all the fields, unless explicitly stated, will be evaluated at $t=0$.

From the Lagrangian (7.82) we can extract the conjugate momentum:

$$\begin{aligned}
&\Pi^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_\mu)} = -F^{0\mu},\\
&\Pi^0 = 0\,, \qquad\qquad \Pi^i = E^i = \partial^i A_0 - \partial_0 A^i.
\end{aligned} \tag{7.88}$$

As in the massless case the field A_0 is not dynamical since its conjugate momentum vanishes. On the other hand, in the present case its equation of motion does not represent a constraint on the spatial components A^i but rather it allows to solve for A_0 and express it in terms of the other degrees of freedom:

$$\begin{aligned}
&\frac{\delta S}{\delta A_0(x)} = 0 \quad \Longrightarrow \quad \partial^\nu F_{\nu 0} + M^2 A_0 = 0,\\
&\partial^i(F^{i0}) - M^2 A_0 = 0 \quad \Longrightarrow \quad A_0 = -\frac{\vec{\nabla}\cdot\vec{\Pi}}{M^2}.
\end{aligned} \tag{7.89}$$

The time component of the vector field A_μ is completely determined in term of the others. Keeping the constraint (7.89) in mind we can start imposing the usual canonical commutation relations for the three pairs of unconstrained variables (A_i,Π^i):

$$\begin{aligned}
&[A_i(0,\vec{x}),\, A_j(0,\vec{y})] = [\Pi^i(0,\vec{x}),\, \Pi^j(0,\vec{y})] = 0,\\
&[A_k(0,\vec{x}),\, \Pi^j(0,\vec{y})] = i\delta^j_k \delta^3(\vec{x}-\vec{y}).
\end{aligned} \tag{7.90}$$

We cannot adopt the standard prescription for the field A_0. At classical level this field is indeed a function of the conjugate momenta, hence we cannot impose the vanishing of the commutator $[A_0, A_i]$. Rather we would like to take into account the classical constraint at the quantum level. We impose therefore

$$[A_0(0,\vec{x}),\, A_j(0,\vec{y})] = [-\frac{\vec{\nabla}\cdot\vec{\Pi}(0,\vec{x})}{M^2},\, A_j(0,\vec{y})] = \frac{i}{M^2}\partial^{(x)}_j \delta^3(\vec{x}-\vec{y}). \tag{7.91}$$

Finally we can compute the Hamiltonian density associated to the Lagrangian density (7.82):

$$\begin{aligned}
\mathcal{H} = \Pi^i \dot{A}_i - \mathcal{L} \;&=\; \Pi^i\left(\partial_i\left(-\frac{\vec{\nabla}\cdot\vec{\Pi}}{M^2}\right) - \Pi_i\right) + \frac{1}{2}\Pi_i\Pi^i + \frac{1}{4}F_{ij}F^{ij} - \frac{1}{2}\frac{(\vec{\nabla}\cdot\vec{\Pi})^2}{M^2} - \frac{1}{2}M^2 A_i A^i\\
&=\; \frac{1}{2}\Pi^i\Pi^i + \frac{1}{2}\frac{(\vec{\nabla}\cdot\vec{\Pi})^2}{M^2} + \frac{1}{4}F_{ij}F_{ij} + \frac{1}{2}M^2 A_i A_i + (\text{total derivative}).
\end{aligned} \tag{7.92}$$

When considering the Hamiltonian $H = \int \mathcal{H} d^3x$ the last terms vanishes. Notice that we have substituted everywhere the constraint (7.89) and we have expressed $\dot{A}_i$ in terms of the other variables using the relation (7.88). As consistency check we can compute the Hamilton equations of motion and show that they coincide with those derived with the Lagrangian formalism:

$$\begin{aligned}
\left[H,\, \Pi^j(0,\vec{x})\right] \equiv -i\dot{\Pi}^j &= \int d^3y \left(\frac{1}{2}F^{mn}\left[F_{mn}(0,\vec{y}),\, \Pi^j(0,\vec{x})\right] - M^2 A^m \left[A_m(0,\vec{y}),\, \Pi^j(0,\vec{x})\right]\right) \\
&= \int d^3y \left(F^{ij}\left[\partial_m^{(y)} A_n(0,\vec{y}),\, \Pi^j(0,\vec{x})\right] - M^2 A^m \left[A_m(0,\vec{y}),\, \Pi^j(0,\vec{x})\right]\right) \\
&= i\int d^3y \left(F^{mj}\partial_m^{(y)}\delta^3(\vec{x}-\vec{y}) - M^2 A^j \delta^3(\vec{x}-\vec{y})\right) = -i\partial_m\, F^{mj}(0,\vec{x}) - iM^2 A^j(0,\vec{x}), \qquad (7.93) \\
\left[H,\, A_j(0,\vec{x})\right] \equiv -i\dot{A}_j &= \int d^3y \left(-\Pi_i\left[\Pi^i(0,\vec{y}),\, A_j(0,\vec{x})\right] + \frac{1}{M^2}(\partial_m\Pi^m)\left[(\partial_n\Pi^n)(0,\vec{y}),\, A_j(0,\vec{x})\right]\right) \\
&= \int d^3y \left(i\Pi_j(0,\vec{y})\delta^3(\vec{x}-\vec{y}) - i\frac{1}{M^2}(\partial_m\Pi^m)\partial_j\delta^3(\vec{x}-\vec{y})\right) = i\Pi_j(0,\vec{x}) + \frac{i}{M^2}\partial_j(\partial_m\Pi^m)(0,\vec{x}).
\end{aligned} \qquad (7.94)$$

Finally taking the time derivative of the second equation and using the first one we get:

$$\begin{aligned}
&\ddot{A}^j = -\dot{\Pi}^j - \frac{1}{M^2}\partial_0\partial^j(\partial_m\Pi^m) = -\dot{\Pi}^j + \partial_0\partial^j A_0 = -\partial_m\, F^{mj} - M^2 A^j + \partial_0\partial^j A_0 \\
&\Longrightarrow \partial_0\partial^0 A^j - \partial_0\partial^j A^0 + \partial_m\, F^{mj} + M^2 A^j = 0.
\end{aligned} \qquad (7.95)$$

In order to construct the Fock space, we express all the fields and the Hamiltonian in Fourier space:

$$A_i(\vec{k}) = \int d^3x\, A_i(0,\vec{x}) e^{-i\vec{x}\cdot\vec{k}}\,, \qquad \Pi^i(\vec{k}) = \int d^3x\, \Pi^i(0,\vec{x}) e^{-i\vec{x}\cdot\vec{k}}\,. \qquad (7.96)$$

The reality of the fields in the coordinate space translates in the condition

$$A_i^\dagger(\vec{k}) = A_i(-\vec{k})\,, \qquad \Pi^{i\dagger}(\vec{k}) = \Pi^i(-\vec{k}). \qquad (7.97)$$

Thus we can straightforwardly express the Hamiltonian in terms of the new fields:

$$\begin{aligned}
H &= \int \frac{d^3k}{(2\pi)^3}\frac{1}{2}\left(\Pi^i(\vec{k})\Pi^i(-\vec{k}) + \frac{(k^i\Pi^i(\vec{k}))(k^j\Pi^j(-\vec{k}))}{M^2} + \left[k_iA_j(\vec{k}) - k_jA_i(\vec{k})\right]k_iA_j(-\vec{k}) + M^2A_i(\vec{k})A_i(-\vec{k})\right) \\
&= \int \frac{d^3k}{(2\pi)^3}\frac{1}{2}\left(|\vec{\Pi}(\vec{k})|^2 + \frac{|\vec{k}\cdot\vec{\Pi}(\vec{k})|^2}{M^2} + \left(|\vec{k}|^2|\vec{A}(\vec{k})|^2 - |\vec{k}\cdot\vec{A}(\vec{k})|^2\right) + M^2|\vec{A}(\vec{k})|^2\right).
\end{aligned} \qquad (7.98)$$

In order to make manifest the difference between the massive case and the massless case let us decompose the fields in their longitudinal and transverse part with respect to the direction of the momentum $\vec{k}$. Introduce the projectors

$$\begin{aligned}
&(P_L)_{ij} = \frac{k_ik_j}{k^2}, \qquad (P_\perp)_{ij} = \delta_{ij} - (P_L)_{ij}\,, \qquad k^2 \equiv |\vec{k}|^2\,, \\
&(P_L)^2 = P_L\,, \qquad (P_\perp)^2 = P_\perp\,, \qquad P_\perp + P_L = \mathbb{1}\,, \qquad P_LP_\perp = P_\perp P_L = 0,
\end{aligned} \qquad (7.99)$$

and define the projected fields:

$$A_i(\vec{k}) = A_{Li}(\vec{k}) + A_{\perp i}(\vec{k})\,, \qquad \Pi_i(\vec{k}) = \Pi_{Li}(\vec{k}) + \Pi_{\perp i}(\vec{k}), \qquad (7.100)$$

where:

$$\begin{aligned}
&A_{Li}(\vec{k}) = (P_L)_{ij}A_j(\vec{k})\,, && A_{\perp i}(\vec{k}) = (P_\perp)_{ij}A_j(\vec{k}), \\
&\Pi_{Li}(\vec{k}) = (P_L)_{ij}\Pi_j(\vec{k})\,, && \Pi_{\perp i}(\vec{k}) = (P_\perp)_{ij}\Pi_j(\vec{k}), \\
&\vec{k}\cdot\vec{A}_\perp(\vec{k}) = \vec{k}\cdot\vec{\Pi}_\perp(\vec{k}) = 0. &&
\end{aligned} \qquad (7.101)$$

Pugging the decomposition (7.100) in the Hamiltonian we obtain:

$$\begin{aligned} H &= H_L + H_\perp \,, \\ H_L &= \int \frac{d^3k}{(2\pi)^3}\frac{1}{2}\left(\frac{w_k^2}{M^2}|\vec{\Pi}_L(\vec{k})|^2 + M^2|\vec{A}_L(\vec{k})|^2\right), \\ H_\perp &= \int \frac{d^3k}{(2\pi)^3}\frac{1}{2}\left(|\vec{\Pi}_\perp(\vec{k})|^2 + w_k^2|\vec{A}_\perp(\vec{k})|^2\right), \end{aligned}$$

where $w_k = k^2 + M^2$. Notice that the transverse modes and the longitudinal ones decouple. Moreover the former have a standard Hamiltonian, schematically of the form $H \sim p^2 + w^2q^2$. The latter instead don't seem to have a standard Hamiltonian, however it is sufficient a field redefinition to bring it in a more familiar form. Let us review the procedure first for the simple case of quantum mechanics. Suppose to start from a unidimensional system where

$$H = \frac{w^2}{M^2}p^2 + M^2q^2\,, \qquad [q,p] = i. \tag{7.102}$$

Define the canonical transformation

$$p' = \frac{w}{M}p\,, \qquad q' = \frac{M}{w}q\,, \qquad [q',p'] = [q,p] = i\,, \qquad H = p'^2 + w^2q'^2. \tag{7.103}$$

Hence the starting Hamiltonian is equivalent to the above one and define the same spectrum. The same procedure can be followed for the quantum field theory of a massive vector. Hence we define the creation and annihilation operators for the transverse modes in the usual way:

$$\begin{aligned} \vec{A}_\perp(\vec{k}) &= \frac{1}{2w_k}\left(\vec{a}_\perp(\vec{k}) + \vec{a}^\dagger_\perp(-\vec{k})\right), \\ \vec{\Pi}_\perp(\vec{k}) &= \frac{i}{2}\left(\vec{a}_\perp(\vec{k}) - \vec{a}^\dagger_\perp(-\vec{k})\right). \end{aligned} \tag{7.104}$$

Instead, as suggested by the above discussion, we rescale the longitudinal mode:

$$\begin{aligned} \vec{A}_L(\vec{k}) &= \frac{1}{2M}\left(\vec{a}_L(\vec{k}) + \vec{a}^\dagger_L(-\vec{k})\right), \\ \vec{\Pi}_L(\vec{k}) &= \frac{iM}{2w_k}\left(\vec{a}_L(\vec{k}) - \vec{a}^\dagger_L(-\vec{k})\right). \end{aligned} \tag{7.105}$$

In order to express the Hamiltonian in terms of the lowering and raising operator we need first to translate the canonical commutation relation for A_i, Π^j to commutation relations between a_i, $a_j^\dagger$. The first straightforward step is to pass to Fourier space:

$$\begin{aligned} &[A_i(\vec{k}), \Pi_j(\vec{k}')] = \int d^3x\, d^3x' e^{-i\vec{x}\cdot\vec{k} - i\vec{x}'\cdot\vec{k}'}[A_i(\vec{x}), \Pi_j(\vec{x}')] = i\delta_{ij}(2\pi)^3\delta^3(\vec{k}+\vec{k}'), \\ &[A_i(\vec{k}), A_j(\vec{k}')] = [\Pi_i(\vec{k}), \Pi_j(\vec{k}')] = 0. \end{aligned} \tag{7.106}$$

Then we can invert the decomposition

$$\begin{aligned} \vec{A}(\vec{k}) &= \frac{1}{2w_k}\left[\left(\vec{a}_\perp(\vec{k}) + \vec{a}^\dagger_\perp(-\vec{k})\right) + \frac{w_k}{M}\left(\vec{a}_L(\vec{k}) + \vec{a}^\dagger_L(-\vec{k})\right)\right], \\ \vec{\Pi}(\vec{k}) &= \frac{i}{2}\left[\left(\vec{a}_\perp(\vec{k}) - \vec{a}^\dagger_\perp(-\vec{k})\right) + \frac{M}{w_k}\left(\vec{a}_L(\vec{k}) - \vec{a}^\dagger_L(-\vec{k})\right)\right], \end{aligned} \tag{7.107}$$

to obtain $\vec{a}(\vec{k})$ in terms of $\vec{A}(\vec{k})$ and $\vec{\Pi}(\vec{k})$

$$\begin{aligned} \vec{a}(\vec{k}) &= \left(w_k\vec{A}_\perp(\vec{k}) - i\vec{\Pi}_\perp(\vec{k})\right) + \left(M\vec{A}_L(\vec{k}) - i\frac{w_k}{M}\vec{\Pi}_L(\vec{k})\right), \\ \vec{a}^\dagger(\vec{k}) &= \left(w_k\vec{A}_\perp(-\vec{k}) + i\vec{\Pi}_\perp(-\vec{k})\right) + \left(M\vec{A}_L(-\vec{k}) + i\frac{w_k}{M}\vec{\Pi}_L(-\vec{k})\right), \end{aligned} \tag{7.108}$$

and finally extract:

$$\begin{aligned} &[\vec{a}(\vec{k}),\, \vec{a}(\vec{k}')] = -iw_k[\vec{A}_\perp(\vec{k}),\, \vec{\Pi}_\perp(\vec{k}')] - iw_k[\vec{\Pi}_\perp(\vec{k}),\, \vec{A}_\perp(\vec{k}')] - iw_k[\vec{A}_L(\vec{k}),\, \vec{\Pi}_L(\vec{k}')] - iw_k[\vec{\Pi}_L(\vec{k}),\, \vec{A}_L(\vec{k}')] = 0, \\ &[\vec{a}^\dagger(\vec{k}),\, \vec{a}^\dagger(\vec{k}')] = 0, \\ &[a_i(\vec{k}),\, a_j^\dagger(\vec{k}')] = iw_k[A_{\perp i}(\vec{k}),\, \Pi_{\perp j}(-\vec{k}')] - iw_k[\Pi_{\perp i}(\vec{k}),\, A_{\perp j}(-\vec{k}')] \\ &\qquad\qquad + iw_k[A_{Li}(\vec{k}),\, \Pi_{Lj}(-\vec{k}')] - iw_k[\Pi_{Li}(\vec{k}),\, A_{Lj}(-\vec{k}')] \\ &\qquad\qquad = 2w_k\delta_{ij}(2\pi)^3\delta^3(\vec{k}-\vec{k}'). \end{aligned} \tag{7.109}$$

In the end the Hamiltonian expressed in terms of the new quantities reads:

$$
\begin{aligned}
H_L &= \int \frac{d^3k}{(2\pi)^3}\frac{1}{4}\left(\vec{a}_L(\vec{k})\cdot\vec{a}_L^\dagger(\vec{k})+\vec{a}_L^\dagger(-\vec{k})\cdot\vec{a}_L(-\vec{k})\right),\\
H_\perp &= \int \frac{d^3k}{(2\pi)^3}\frac{1}{4}\left(\vec{a}_\perp(\vec{k})\cdot\vec{a}_\perp^\dagger(\vec{k})+\vec{a}_\perp^\dagger(-\vec{k})\cdot\vec{a}_\perp(-\vec{k})\right),\\
H = H_L + H_\perp &= \int \frac{d^3k}{(2\pi)^3}\frac{1}{4}\left(\vec{a}(\vec{k})\cdot\vec{a}^\dagger(\vec{k})+\vec{a}^\dagger(\vec{k})\cdot\vec{a}(\vec{k})\right) = \int \frac{d^3k}{2w_k(2\pi)^3} w_k \sum_i \vec{a}_i(\vec{k})^\dagger\vec{a}_i(\vec{k})\\
&= \int d\Omega_{\vec{k}}\; w_k \sum_i \vec{a}_i(\vec{k})^\dagger\vec{a}_i(\vec{k}).
\end{aligned}
\tag{7.110}
$$

7.6.2 Heisenberg Picture Field

We can now consider the time dependent fields in the Heisenberg picture

$$
A^0(x)\equiv A^0(t,\vec{x})\equiv e^{iHt}A^0(0,\vec{x})e^{-iHt} = -e^{iHt}\frac{\vec{\nabla}\cdot\vec{\Pi}(0,\vec{x})}{M^2}e^{-iHt} \qquad A^i(x)\equiv A^i(t,\vec{x})\equiv e^{iHt}A^i(0,\vec{x})e^{-iHt}
\tag{7.111}
$$

The Hamiltonian straightfowardly dicatates the time dependende of creation and destruction operators: $a^i(t,\vec{k}) = e^{-i\omega_k t}a^i(\vec{k})$. By using eq. (7.107) we can then write

$$
\begin{aligned}
A^0(x) &= \int d\Omega_k \left(\frac{\vec{k}\cdot\vec{a}(\vec{k})}{M}e^{-ikx}+\text{h.c.}\right)\\
\vec{A}(x) &= \int d\Omega_k \left[\left(\vec{a}_\perp(\vec{k})+\frac{\omega_k}{M}\vec{a}_L(\vec{k})\right)e^{-ikx}+\text{h.c.}\right]
\end{aligned}
\tag{7.112}
$$

where we used $\vec{k}\cdot\vec{a}_\perp = 0$ and $\vec{k}\cdot\vec{a}_L = \vec{k}\cdot\vec{a}$. We can check that the above quantum fields satisfy the equations of motion 7.85. The first equation is obviously satisfied given $k^\mu = (\omega_k,\vec{k})$ satisfies by construction the on-shell condition $k^\mu k_\mu = M^2$. Consider then the diverge of $\partial_\mu A^\mu$. We find

$$
\dot{A}^0(x) = -i\int d\Omega_k \left(\omega_k\frac{\vec{k}\cdot\vec{a}(\vec{k})}{M}e^{-ikx}-\text{h.c.}\right)
\tag{7.113}
$$

$$
\vec{\nabla}\vec{A}(x) = i\int d\Omega_k \left[\vec{k}\cdot\left(\vec{a}_\perp(\vec{k})+\frac{\omega_k}{M}\vec{a}_L(\vec{k})\right)e^{-ikx}-\text{h.c.}\right]
\tag{7.114}
$$

which, upon using $\vec{k}\cdot\vec{a}_\perp = 0$, are easily seen to give $\dot{A}^0+\vec{\nabla}\cdot\vec{A} = \partial_\mu A^\mu = 0$.

7.6.3 Polarization vectors

Eq. (7.112) can be more compactly written

$$
A^\mu(x) = \int d\Omega_k \left(\epsilon_i^\mu(\vec{k})a^i(\vec{k})e^{-ikx}+\text{h.c.}\right)
\tag{7.115}
$$

with $\epsilon_i^\mu(\vec{k})$ a 4×3 matrix with entries given by

$$
\mu = 0 \quad\to\quad \epsilon_i^0(\vec{k}) = \frac{k^i}{M} = -\frac{k_i}{M}
\tag{7.116}
$$

$$
\mu = j \quad\to\quad \epsilon_i^j(\vec{k}) = P_\perp(\vec{k})_i^j + \frac{\omega_k}{M}P_L(\vec{k})_i^j = \delta_i^j - \frac{k^jk^i}{k^2}+\frac{\omega_k}{M}\frac{k^jk^i}{k^2}\,.
\tag{7.117}
$$

For $i = 1,2,3$ the $\epsilon_i^\mu(\vec{k})$ form a basis of the 3-dimensional subspace of 4-vectors satisfying $k_\mu\epsilon^\mu = 0$. Recalling eq. (7.86), this shows that (like all quantum fields encountered so far) A^μ is written as the most general linear superposition of plane wave solutions with quantized coefficients $a^i(\vec{k})$, $a^{i\dagger}(\vec{k})$. The $\epsilon_i^\mu(\vec{k})$ are thus the analogue

for the vector field of the $u_r^\alpha(\vec{k})$ and $v_r^\alpha(\vec{k})$ of the Dirac field (the 4-spinor index α is here made explicit). Like in the case of the Dirac field, one can show that the polarization $\epsilon_i^\mu(\vec{k})$ for arbitrary $\vec{k}$ is simply obtained by Lorentz transforming the result in the center of mass $\epsilon_i^\mu(\vec{0})$. The latter by eq. (7.116) reads

$$\epsilon_i^0(\vec{0}) = 0\,, \qquad \epsilon_i^j(\vec{0}) = \delta_i^j\,, \tag{7.118}$$

and can be compactly represented in block form as the 4×3 matrix (4 rows and 3 columns)

$$\epsilon_i^\mu(\vec{0}) = \begin{pmatrix} 0 \\ \mathbb{1} \end{pmatrix}. \tag{7.119}$$

For arbitrary $\vec{k}$ eq. (7.116) can similarly be written as

$$\epsilon_i^\mu(\vec{k}) = \begin{pmatrix} \gamma_k \vec{\beta}_{\vec{k}} \\ P_\perp(\vec{k}) + \gamma_k P_L(\vec{k}) \end{pmatrix} \tag{7.120}$$

where, to make contact with the standard notation for Lorentz boosts, we defined $\gamma_k = \omega_k/M$ and $\vec{\beta}_{\vec{k}} = \vec{k}/\omega_k$. By the above two equations we can thus write

$$\epsilon_i^\mu(\vec{k}) = H(k)^\mu{}_i = H(k)^\mu{}_\nu \epsilon_i^\nu(\vec{0}) = \begin{pmatrix} \gamma_k & \gamma_k \vec{\beta}_{\vec{k}} \\ \gamma_k \vec{\beta}_{\vec{k}} & P_\perp(\vec{k}) + \gamma_k P_L(\vec{k}) \end{pmatrix} \begin{pmatrix} 0 \\ \mathbb{1} \end{pmatrix}. \tag{7.121}$$

where, in the same notation of Chapter 6,

$$H(k)^\mu{}_\nu \equiv (e^{i\vec{\eta}(k)\cdot\vec{\mathbb{K}}})^\mu{}_\nu \qquad \vec{\eta}(k) = \frac{\vec{k}}{|k|}\tanh^{-1}(|k|/\omega_k) \tag{7.122}$$

is the pure boost that relates the generic 4-momentum $k^\mu = (\omega_k, \vec{k})$ to the center of mass momentum $\bar{k}^\mu = (M, \vec{0})$

$$k^\mu = H(k)^\mu{}_\nu \bar{k}^\nu\,. \tag{7.123}$$

For instance, in the case of $\vec{k} = (k, 0, 0)$, we have

$$H(k)^\mu{}_\nu = \begin{pmatrix} \frac{\omega_k}{M} & \frac{k}{M} & 0 & 0 \\ \frac{k}{M} & \frac{\omega_k}{M} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{7.124}$$

The polarization vectors $\epsilon_i^\mu(\vec{k})$ for arbitary k^μ are thus obtained from those in the center of mass via a Lorentz transformation, precisely like for the Dirac field. Indeed, in our choice of basis the polarization vectors $\epsilon_i^\mu(\vec{k})$ are for $i = 1, 2, 3$ given by respectively the second, third and fourth column of the boost matrix $H(k)^\mu{}_\nu$. As we will see in the next subsection when discussing the action of Lorentz transformation of the physical states, the appearance of the boost matrix $H(k)^\mu{}_\nu$ in the construction of the polarization vectors is not by chance.

The relation $\epsilon_i^\mu(\vec{k}) = H(k)^\mu{}_i$, through the basic property of Lorentz transformation in eq. (3.101) and through eq. (7.123), also makes the transversality of the polarization vectors manifest

$$k^\mu H(h)_\mu{}^i = H(h)_\mu{}^i H(k)^\mu{}_\nu \bar{k}^\nu = \delta_\nu^i \bar{k}^\nu = \bar{k}^i = 0\,. \tag{7.125}$$

The polarization vectors $\epsilon_i^\mu(\vec{k})$ satisfy a set of *orthogonality* and *completeness* relations that are fully analogous to those satisfied by the Dirac polarization wave functions $u_r^\alpha(k)$ and $v_r^\alpha(k)$. These relations are readily derived using the explicit form $\epsilon_i^\mu(\vec{k}) = H(k)^\mu{}_i$:

- Orthogonality

$$\epsilon_i^\mu(\vec{k})\epsilon_{\mu j}(\vec{k}) = \eta_{\mu\nu} H(k)^\mu{}_i H(k)^\nu{}_j = \eta_{ij} = -\delta_{ij} \tag{7.126}$$

- Completeness

$$\epsilon_i^\mu(\vec{k})\epsilon_i^\nu(\vec{k}) = H(k)^\mu{}_{\mu'}H(k)^\nu{}_{\nu'}(-\eta^{\mu'\nu'}+\delta_0^{\mu'}\delta_0^{\nu'}) = H(k)^\mu{}_{\mu'}H(k)^\nu{}_{\nu'}(-\eta^{\mu'\nu'}+\frac{\bar{k}^{\mu'}\bar{k}^{\nu'}}{M^2}) = -\eta^{\mu\nu}+\frac{k^\mu k^\nu}{M^2}\,, \quad (7.127)$$

where we used $\bar{k}^\mu = \delta_0^\mu M$ for the center of mass momentum. In particular in the center of mass frame we have

$$\epsilon_i^\mu(\vec{0})\epsilon_i^\nu(\vec{0}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (7.128)$$

7.6.4 Lorentz transformations

Has shown from the study of the irreducible representations of the Lorentz group, the single particle states of a massive particle with spin s can be constructed by boosting the states in the center of mass according to

$$|p,\sigma\rangle \equiv H(p)|\bar{p},\sigma\rangle \qquad \bar{p} = (M,0)\,. \quad (7.129)$$

Under a general Lorentz transformation we then found

$$U(\Lambda)|p,\sigma\rangle = |\Lambda p,\sigma'\rangle\mathcal{D}^{(s)}_{\sigma'\sigma}(\mathcal{W}(\Lambda,p)) \quad (7.130)$$

with $\mathcal{D}^{(s)}_{\sigma'\sigma}(\mathcal{W}(\Lambda,p))$ the spin s representation of the Wigner rotation $\mathcal{W}(\Lambda,p) = H^\dagger(\Lambda p)U(\Lambda)H(\Lambda p)$.

Let us then consider the realization of this general result in the special case of the vector field, where the single particle states are

$$|p,i\rangle \equiv a_i^\dagger(\mathbf{p})|0\rangle\,. \quad (7.131)$$

As $i = 1,2,3$ this corresponds (as we shall now check) to $s = 1$ where $\mathcal{D}^{(s)}$ reduces to an ordinary 3×3 orthogonal matrix.

Let us consider pure rotation. The action on the ladder operators and hence on the states, can be derived by focussing on the transformation properties of the spacial vector component $\vec{A}(t,\mathbf{x})$

$$U(R)^\dagger A_i(t,\mathbf{x})U(R) = R_i{}^j A_j(t,R^{-1}\mathbf{x})\,. \quad (7.132)$$

Using eq. (7.112) and the commutation properties of the projectors (which are easily checked)

$$RP_\perp(\mathbf{p}) = P_\perp(R\mathbf{p})R\,, \qquad RP_L(\mathbf{p}) = P_L(R\mathbf{p})R\,, \quad (7.133)$$

eq. (7.132) on the ladder operators reads

$$U(R)^\dagger a_i^\dagger(\mathbf{p})U(R) = R_i{}^j a_j^\dagger(R^{-1}\mathbf{p}) \quad (7.134)$$

which on single particle states implies (using $U(R)|0\rangle = |0\rangle$)

$$U(R)|\mathbf{p},i\rangle = U(R)a_i^\dagger(\mathbf{p})|0\rangle = a_j(R\mathbf{p})|0\rangle R^{-1}{}_i{}^j = a_j(R\mathbf{p})|0\rangle R^j{}_i \quad (7.135)$$

where in the last step we used that $R^{-1} = R^T$. This result coincides with eq. (6.37) for the special case where $s = 1$ and $\mathcal{D}_{\sigma\sigma'}$ coincides with the R matrix itself. This proves that the quanta of the massive vector field carry spin 1. For instance in the case of a particle at rest, we have three basis states $|\mathbf{0},i\rangle = a_i^\dagger(\mathbf{0})|0\rangle$ and the general state is thus written as

$$|\psi\rangle = \alpha_i|\mathbf{0},i\rangle\,. \quad (7.136)$$

By the transformation properties of $|\mathbf{0},i\rangle$ the canonically normalized eigenstates of J_3 correspond then to

$$J_3 = 0 \;\to\; \alpha = (0,0,1) \qquad J_3 = \pm 1 \;\to\; \alpha = (1,\pm i,0)/\sqrt{2}\,. \quad (7.137)$$

In the center of mass, the $a^{i\dagger}(k_{CM})$ and $a^i(k_{CM})$ are thus associated to quanta polarized in the direction i. These three polarization states transform as a $J = 1$ representation (a vector) under space rotations, corresponding to the particles having spin =1.

Consider now instead a general Lorentz transformation

$$U(\Lambda)^\dagger A^\mu(x) U(\Lambda) = \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x)\,. \tag{7.138}$$

Using the expansion in ladder operators, the left hand side simply reads

$$U(\Lambda)^\dagger A^\mu(x) U(\Lambda) = \int d\Omega_k \left(\epsilon^{\mu i}(\mathbf{k}) U(\Lambda)^\dagger a_i(\mathbf{k}) U(\Lambda) e^{-ikx} + \text{h.c.}\right)\,, \tag{7.139}$$

while, by a standard change of integration variables, the right hand side can be written as

$$\begin{aligned} \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x) &= \int d\Omega_k \left[\Lambda^\mu{}_\nu \epsilon^{\nu i}(\mathbf{k}) a_i(\mathbf{k}) e^{-ik(\Lambda^{-1}x)} + \text{h.c.}\right] \\ &= \int d\Omega_k \left[\Lambda^\mu{}_\nu \epsilon^\nu_i(\Lambda^{-1}\mathbf{k}) a^i(\Lambda^{-1}\mathbf{k}) e^{-ikx} + \text{h.c.}\right] \end{aligned} \tag{7.140}$$

(as elsewhere by $\Lambda^{-1}\mathbf{k}$ we obviously mean the action of Λ^{-1} on the 4-vector $k^\mu = (\omega_k, \mathbf{k})$). To deduce the transformation property of $a^i(\mathbf{k})$ by comparing eqs. (7.139) and (7.140), we must properly rewrite $\Lambda^\mu{}_\nu \epsilon^\nu_i(\Lambda^{-1}k)$. Using eq. (7.121) we have

$$\Lambda^\mu{}_\nu \epsilon^\nu_i(\Lambda^{-1}\mathbf{k}) = \Lambda^\mu{}_\nu H(\Lambda^{-1}k)^\nu{}_i = H(k)^\mu{}_\rho H^{-1}(k)^\rho{}_\sigma \Lambda^\sigma{}_\nu H(\Lambda^{-1}k)^\nu{}_i \equiv H(k)^\mu{}_\rho \widetilde{\mathcal{W}}(\Lambda,k)^\rho{}_i \tag{7.141}$$

and moreover, by the same reasoning used in eq. (6.24), we have that $\widetilde{\mathcal{W}}(\Lambda,k)$ is an element of the little group $SO(3)$, i.e. a pure rotation[5]

$$\widetilde{\mathcal{W}}(\Lambda,k) = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \tilde{\mathcal{R}}(\Lambda,k) \end{pmatrix}\,. \tag{7.142}$$

In fact by comparing to the definition of the Wigner rotation in eq. (6.24) to eq. (7.141), we find the relation $\widetilde{\mathcal{W}}(\Lambda,k) = \mathcal{W}^{-1}(\Lambda^{-1},k)$. Eqs. (7.141, 7.142) toghether then imply

$$\Lambda^\mu{}_\nu \epsilon^\nu_i(\Lambda^{-1}\mathbf{k}) = H(k)^\mu{}_\rho \widetilde{\mathcal{W}}(\Lambda,k)^\rho{}_i = H(k)^\mu{}_j \widetilde{\mathcal{W}}(\Lambda,k)^j{}_i = \epsilon^\mu_j(\mathbf{k}) \widetilde{\mathcal{W}}(\Lambda,k)^j{}_i \tag{7.143}$$

so that by comparing eq. (7.139) and eq. (7.140) we obtain

$$U^\dagger(\Lambda) a^j(\mathbf{k}) U(\Lambda) = \widetilde{\mathcal{W}}(\Lambda,k)^j{}_i a^i(\Lambda^{-1}\mathbf{k})\,, \tag{7.144}$$

which shows how the unitary Wigner rotation controls the Lorentz tranformation of the ladder operators. To make contact with the notation of chapter 6, it is convenient to consider the inverse of the above

$$U(\Lambda) a_j(\mathbf{k}) U^\dagger(\Lambda) = U^\dagger(\Lambda^{-1}) a^j(\mathbf{k}) U(\Lambda^{-1}) = \widetilde{\mathcal{W}}(\Lambda^{-1},k)^j{}_i a^i(\Lambda\mathbf{k}) = \left[\mathcal{W}(\Lambda,k)^{-1}\right]_j{}^i a_i(\Lambda\mathbf{k})\,, \tag{7.145}$$

so that for the creation operator we have

$$U(\Lambda) a^\dagger_j(\mathbf{k}) U^\dagger(\Lambda) = \left[\mathcal{W}(\Lambda,k)^{-1*}\right]^i{}_j a^\dagger_i(\Lambda\mathbf{k}) = \mathcal{W}(\Lambda,k)^i{}_j a^\dagger_i(\Lambda\mathbf{k})\,, \tag{7.146}$$

where in the last step we used unitarity of the Wigner matrix: $\mathcal{W}^{-1*} = \mathcal{W}^T$. The last equation dictates the Lorentz transformation property of single particle states

$$U(\Lambda)|\mathbf{k},j\rangle = U(\Lambda) a^\dagger_j(\mathbf{k}) U^\dagger(\Lambda)|0\rangle = \mathcal{W}(\Lambda,k)^i{}_j a^\dagger_i(\Lambda\mathbf{k})|0\rangle = |\Lambda\mathbf{k},i\rangle \mathcal{W}(\Lambda,k)^i{}_j\,, \tag{7.147}$$

which nicely matches, for the specific case $s=1$, the general result in eq. (6.34).

[5] More explicitly, one can easily check that $\widetilde{\mathcal{W}}^\mu{}_\nu(\Lambda,k)\bar{k}^\nu = \bar{k}^\mu$.

Chapter 8

Causality in quantum field theory

8.1 Introduction

Causality is a very important notion in physics. Nevertheless, it is a very simple and intuitive one: it has to do with the intrinsic difference between future and past and it is needed to distinguish whether an event can influence another event in the future or be influenced by an event in the past.

Non relativistic theory. In a non-relativistic theory causality is a fact, Newton's equations are local in time, in solving a problem in this framework one sets initial conditions and then finds the time evolution of the observables using the equations of motion. The reason why causality is intrinsically implemented in a non-relativistic theory is that the space-time symmetries are given by the Galileo group that does not change time ordering for different observers.

Relativistic theory. We know that in a relativistic theory information cannot be exchanged at a speed larger than the speed of light, if an event happens at $x^\mu = (t, \mathbf{x})$ we know that the region in causal contact with this event is represented by the future and past light-cone. What we want to do is to make this statement quantitative, at the level of the operators that one considers in any theory.

8.2 Causality in classical field theory

Let us start by considering the implementation of this notion in classical field theory. We consider a free scalar field for simplicity, but the discussion applies more generally. In classical field theory one starts by defining the initial conditions at a fixed time slice $t = 0$ (this hypersurface is usually called Cauchy surface), meaning that we specify the value of the fields $\phi(t = 0, \mathbf{x}) \equiv \phi_0(\mathbf{x})$ and of its time derivative $\dot{\phi}(t = 0, \mathbf{x}) = \dot{\phi}_0(\mathbf{x}) \equiv \pi_0(\mathbf{x})$. Notice that $\phi_0(\mathbf{x})$ and $\dot{\phi}_0(\mathbf{x})$ are functions defined on $\mathbb{R}^3$. The solution to the equations of motion then gives the time evolution of the fields in the form

$$\phi(t, \mathbf{y}) = F[t, \mathbf{y}; \phi_0, \pi_0], \tag{8.1}$$

where F is clearly a functional of ϕ_0, π_0. To define causality we proceed as follows. Imagine we modify the initial conditions in a region of the Cauchy surface V as indicated in figure 8.1. Causality is then naturally defined as the independence of $\phi(y)$ on such modification when $y = (t, \mathbf{y})$ lies outside the union of the lightcones of the points $x \in V$. Denoting spacelike separated points as $x \bigtimes y$, the above statement is mathematically expressed in terms of functional derivatives as

$$x \bigtimes y \implies \frac{\delta\phi(y)}{\delta\phi_0(x)} = \frac{\delta\phi(y)}{\delta\pi_0(x)} = 0. \tag{8.2}$$

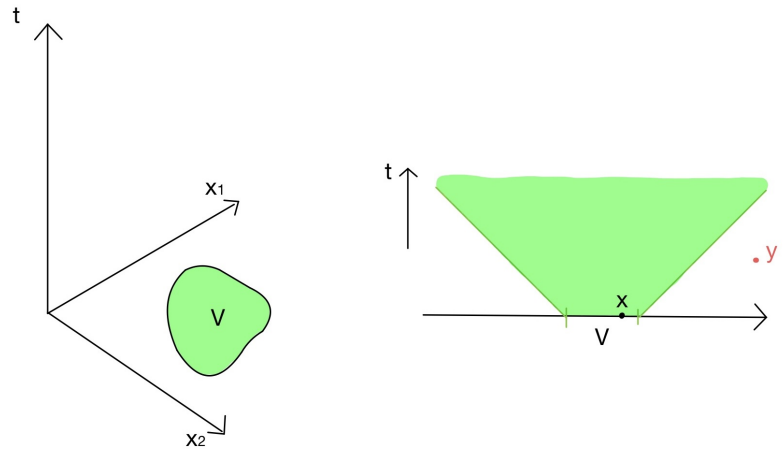


Figure 8.1: Left: region of the $t = 0$ cauchy surface where we modify the initial conditions. Right: the union of the future lightcones of the points $x \in V$. If y lies outside this region (red point case), the solution $\phi(y)$ should remain the same even if one modifies the initial conditions in V.

These equations can also be written using our definition of Poisson brackets as

$$x \bigtimes y \implies \frac{\delta\phi(y)}{\delta\phi_0(x)} = \{\phi(y), \phi_0(x)\} = 0, \tag{8.3}$$

$$x \bigtimes y \implies \frac{\delta\phi(y)}{\delta\pi_0(x)} = \{\phi(y), \pi_0(x)\} = 0, \tag{8.4}$$

As an exercise, one can check that this holds for Klein-Gordon fields.

8.3 Causality in quantum field theory

When canonically quantizing the fields, one promotes the fields to operators and the Poisson brackets to commutators, thus the statement of causality for two observables $A(x)$ and $B(x)$ (fields or product of fields) becomes:

$$x \bigtimes y \implies [A(x), B(y)] = 0. \tag{8.5}$$

8.3.1 Example: the Klein-Gordon field

For the Klein Gordon field the commutator looks like

$$[\phi(x), \phi(y)] = \int d\Omega_p d\Omega_k \left[a_p e^{-ip\cdot x} + a_p^\dagger e^{ip\cdot x}, a_k e^{-ip\cdot y} + a_k^\dagger e^{ip\cdot y}\right] = \tag{8.6}$$

$$= \int d\Omega_p \left(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}\right) \equiv D(x-y) - D(y-x). \tag{8.7}$$

Notice that $D(x)$ is Lorentz invariant since

$$D(x) \to D(\Lambda x) = \int d\Omega_p e^{-ip\cdot(\Lambda x)} = \int d\Omega_p e^{-i(\Lambda^{-1}p)\cdot x} = \int d\Omega_{p'} e^{-ip'\cdot x} = D(x), \tag{8.8}$$

where we changed integration variable to $p' = \Lambda^{-1}p$ and used the Lorentz invariance of the integration measure. We want to show that the commutator expressed as in (8.7) vanishes for space-like separated points, i.e. for $(x-y)^2 < 0$. Using translation invariance, we can pick a frame where $x = (0, 0)$. Furthermore, by applying a Lorentz transformation, we move, since the points are space-like, along the blue hyperboloid of figure 8.2 to a frame where and $x - y \to (0, \mathbf{y})$. In this frame (8.7) becomes

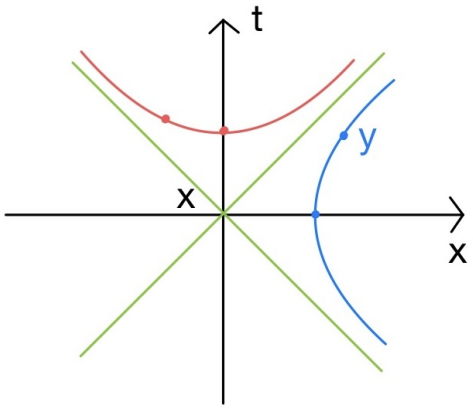


Figure 8.2: Hyperbolae where we can move using Lorentz transformation (the point x is in placed at the origin): for space-like separated points (blue) and time-like separated points (red).

$$D(x-y) - D(y-x) = \int d\Omega_p \left(e^{-i\mathbf{p}\cdot\mathbf{y}} - e^{i\mathbf{p}\cdot\mathbf{y}}\right) = 0, \tag{8.9}$$

since we can change $\mathbf{p} \to -\mathbf{p}$ and the measure of integration is parity invariant.
For time-like separated points, i.e. $(x-y)^2 > 0$ we will instead move along the red hyperboloid in figure 8.2 to a frame where $y - x = (y^0, 0)$, getting

$$[\phi(x), \phi(y)] = \int d\Omega_p \left(e^{i\omega_p y^0} - e^{-i\omega_p y^0}\right). \tag{8.10}$$

Now, as $\omega_p = \sqrt{\mathbf{p}^2 + m^2} > 0$, we cannot perform the analogue of the $\mathbf{p} \to -\mathbf{p}$ change of variables we did before. Consequently the commutator does not vanish. In the same way one can show that the commutator with $\pi(x)$ vanishes for space-like separated points, in fact [1]

$$x \bigtimes y \implies [\phi(x), \pi(y)] = \partial_{y^0} [\phi(x), \phi(y)] = 0 \tag{8.11}$$

For fermions, one can show in a similar way that what vanishes outside the light-cone is the anti-commutator of fields:

$$x \bigtimes y \implies \{\psi(x), \overline{\psi}(y)\} = 0. \tag{8.12}$$

8.3.2 Relation to "naive" Relativistic Quantum Mechanics

In section 1.2 we have seen that when trying to formulate a relativistic theory of quantum mechanics using a wave-function, and a relativistic (non-local) Hamiltonian $H = \sqrt{p^2 + m^2}$ the probability amplitude

$$\langle y | e^{-iHt} | x \rangle = \int \frac{d^3p}{(2\pi)^3} e^{-i\sqrt{p^2+m^2}} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} \neq 0, \tag{8.13}$$

also for space-like separated points, implying that causality is violated. How does this get solved in the correct formulation of Quantum Field Theory? The reason is that we have seen that information about causality is embedded in the commutator of observables, for example consider the following

$$[\phi(x), \pi(y)] = \partial_{y^0} [\phi(x), \phi(y)] = \partial_{y^0} \int \frac{d^3p}{(2\pi)^3 2\omega_p} \left(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}\right) = \tag{8.14}$$

$$= \int \frac{d^3p}{(2\pi)^3} \frac{-i}{2} \left(e^{-ip\cdot(x-y)} + e^{ip\cdot(x-y)}\right) = 0 \text{ when } (x-y)^2 < 0. \tag{8.15}$$

[1] Notice that at fixed times we have $[\phi(0, \mathbf{x}), \pi(0, \mathbf{y})] = -i\delta^{(3)}(\mathbf{x} - \mathbf{y})$, compatibly with causality.

Looking at equation (8.13) we see that the commutator vanishing implies that what vanishes not the amplitude above but the sum of the following two amplitudes:

$$\langle y|e^{-iHt}|x\rangle + \langle y|e^{iHt}|x\rangle. \tag{8.16}$$

Thus, the single contribution to the amplitude is nonzero and violates causality, but in the "naive" version of relativistic QM we missed an extra contribution that compensates this and makes the commutator vanishes. This has an interesting particle interpretation: the two terms in the last equation can be interpreted as the amplitude for a particle travelling from x to y and another particle going in the opposite direction, i.e.

$$\langle y|e^{-iHt}|x\rangle + \langle y|e^{iHt}|x\rangle = \tag{8.17}$$

$$= \quad x \xrightarrow{\quad} y \quad + \quad y \xrightarrow[\quad]{} x \ , \tag{8.18}$$

where the arrows denote the direction of propagation. If fields had charge we could interpret the particle moving from y to x as a particle with opposite charge from x to y: an antiparticle

$$\langle y|e^{iHt}|x\rangle + \langle y|e^{-iHt}|x\rangle = \tag{8.19}$$

$$= \quad x \xrightarrow{\quad} y \quad + \quad y \xrightarrow[\quad]{} x \ , \tag{8.20}$$

where the bold arrow denotes the flow of charge. This means that to ensure causality one has to consider both the propagation of particles and antiparticles, and that causality implies the existence of antiparticles.

8.3.3 Causality from Lorentz invariance of the commutator

There is another way to prove the vanishing of the commutator of fields outside the light-cone just using canonical quantization and Lorentz invariance of the commutators. Consider two points at $t = 0$ in a different position, from canonical quantization (at fixed times) we have

$$[\phi(0, \mathbf{x}), \phi(0, \mathbf{y})] = 0, \tag{8.21}$$

for any $\mathbf{y} \neq \mathbf{x}$. Now we boost this relation recalling that $U(\Lambda)\phi(x)U^\dagger(\Lambda) = \phi(\Lambda x)$, where $U(\Lambda)$ is the representation of a Lorentz boost in the Hilbert space. We take $\mathbf{x} = 0$ and multiply the above equation times U from the left and $U^\dagger$ on the right and inserting $1 = U^\dagger U$ inside product of fields in the commutator:

$$U(\Lambda)\,[\phi(0), \phi(y)]\,U^\dagger(\Lambda) = [\phi(0), \phi(\Lambda y)] = 0. \tag{8.22}$$

Since by choosing any $\mathbf{y}$ and by moving with any Λ on the blue hyperboloid in figure 8.2 we can reach any point $y' = \Lambda y$ with space-like separation from $x = 0$, we proved that the commutator vanishes for any point such that $(x - y)^2 < 0$.

8.4 Measurements, commuting observables and causality

In this section, we want to explain how the notion of causality is related to the measurement process in quantum mechanics. In particular, we will show that with the implementation of causality in quantum field theory, two measurements are not correlated if performed at space-like separated points in space-time.

Consider a quantum field theory in which we specify an Hamiltonian H_0. Any state $|\psi\rangle$ is evolved in time using the Hamiltonian $|\psi_0(t)\rangle = e^{-iH_0 t}|\psi\rangle$. In quantum mechanics, a measurement works in the following way: one makes a number of copies of a system and then add a perturbation to the Hamiltonian, for example a term that represents the interaction of the system with a detector. Say the detector is placed around $\mathbf{x} = 0$ in a certain frame, we can model this interaction in the following way:

$$H_0 \to H_0 + f(t)\Delta H \equiv H_0 + H_I(t), \tag{8.23}$$

where $f(t)$ is a function peaked around $t = 0$ and ΔH is the integral of an operator localized around $\mathbf{x} = 0$

$$\Delta H = \int d^3\mathbf{x}\, \mathcal{O}(\mathbf{x}). \tag{8.24}$$

Let us write the evolution operator $U(t)$ of the full system as $U(t) \equiv U_0(t)\Omega(t)$, where $U_0(t) = e^{-iH_0 t}$ then the time evolution equation for $U(t)$

$$i\frac{d}{dt}U(t) = (H_0 + H_I)U(t), \tag{8.25}$$

reads

$$\frac{d}{dt}\Omega = i\frac{d}{dt}U_0 + iU_0\frac{d}{dt}\Omega = H_0U_0\Omega + H_IU_0\Omega \tag{8.26}$$

$$\frac{d}{dt}\Omega = -i(U_0^\dagger H_I U_o)\Omega, \tag{8.27}$$

where we used that $\frac{d}{dt}U_0 = -iH_0U_0$. Now we define $H_I(t) \equiv U_0^\dagger(t)H_IU_0(t)$, that gives

$$H_I(t) = f(t)\int d^3\mathbf{x}U_0(t)\mathcal{O}(\mathbf{x})U_0(t) = \int d^3\mathbf{x}\mathcal{O}(x), \tag{8.28}$$

where $\mathcal{O}(x) = \mathcal{O}(t, \mathbf{x})$ is the operator evolved in the Heisemberg picture with H_0. The final equation for Ω is

$$\frac{d}{dt}\Omega = -iH_I(t)\Omega. \tag{8.29}$$

The solution can be expressed through the Dyson series and at first order is

$$\Omega = 1 - i\int_{-T}^{t} H_I(t')dt' + \ldots = 1 - i\int_{-\infty}^{t} H_I(t')dt' + \ldots, \tag{8.30}$$

where we extended the time integral since the interaction acts only around $t = 0$. Now we can compute the evolution of the initial wave function $|\psi\rangle$ with the combined Hamiltonian of the system plus the detector, we have

$$|\psi(t)\rangle = U(t)|\psi\rangle = U_0(t)\Omega(t)|\psi\rangle. \tag{8.31}$$

With this new evolved state at time t, we do another measurement, namely we measure the expectation value of an observable $A(\mathbf{x})$, now at time t, i.e. we compute

$$\langle\psi(t)|A(\mathbf{x})|\psi(t)\rangle, \tag{8.32}$$

and we ask when is this different with respect to the situation where we do the second measurement without having performed any measurement at $t = 0$, $\mathbf{x} = 0$:

$$\langle\psi_0(t)|A(\mathbf{x})|\psi_0(t)\rangle = \langle\psi|A(t, \mathbf{x})|\psi\rangle, \tag{8.33}$$

where $A(t, \mathbf{x}) = A(x) = U_0^\dagger(t)A(\mathbf{x})U_0(t)$ is again the operator evolved with H_0 and $|\psi_0(t)\rangle = U_0(t)|\psi\rangle$. We get

$$\langle\psi(t)|A(\mathbf{x})|\psi(t)\rangle = \langle\psi|\Omega^\dagger U_0^\dagger(t)A(\mathbf{x})U_0(t)\Omega|\psi\rangle = \langle\psi|\Omega^\dagger A(t, \mathbf{x})\Omega|\psi\rangle = /\text{expand at 1st order } \Omega/ = \tag{8.34}$$

$$= \langle\psi|A(t, \mathbf{x})|\psi\rangle + i\int_{-\infty}^{t} dt'\langle\psi|\left[H_I(t'), A(x)\right]|\psi\rangle. \tag{8.35}$$

Then taking the difference with equation (8.33) we get

$$i\int_{-\infty}^{t} dt\langle\psi|\left[H_I(t'), A(x)\right]|\psi\rangle. \tag{8.36}$$

So when one does a measurement of $A(\mathbf{x})$ at time t, the difference between having previously performed or not a measurement at $t = 0$ is controlled by the commutator of fields, and will vanish if the points are space-like separated, meaning that the causal relation between measurements is respected in the way we stated above.

As a concrete case, consider $f(t) = \delta(t)$ and $\Delta H = \phi(t, \mathbf{x} = 0)$ being a Klein Gordon field, we get a difference between the two expectation values that is

$$i\langle\psi|\left[\phi(0), A(x)\right]|\psi\rangle, \tag{8.37}$$

which vanishes if $(x - 0)^2 < 0$, i.e. if the two points do not lie one in the light-cone of the other a measurement done in the first point does not affect the second one.

Chapter 9

Discrete symmetries

9.1 Introduction

As we have seen, symmetries play an essential role in Quantum Field Theory (QFT). So far, we have mostly concentrated on *continuous* symmetries (*i.e.* Lie groups) for which the main dynamical consequences are Noether theorem and of course the presence of a degeneracy of solutions related by symmetry transformations. We have already briefly encountered examples of *discrete* (as opposed continuous) symmetries when discussing the Lorentz group:

- Parity: $P: \ (t, \vec{x}) \mapsto (t, -\vec{x})$ (9.1)
- Time Reversal: $T: \ (t, \vec{x}) \mapsto (-t, \vec{x})$ (9.2)

These symmetries are familiar already from non-relativistic physics. In relativistic QFT there is a third important discrete symmetry which does not act on spacetime but which exchanges particles and antiparticles:

- Charge Conjugation: $C: \ \left\{ \begin{smallmatrix} \text{Particles} \\ x^\mu \\ \phi_a(x) \end{smallmatrix} \right\} \mapsto \left\{ \begin{smallmatrix} \text{Antiparticles} \\ x^\mu \\ \phi_a^*(x) \end{smallmatrix} \right\}$ (9.3)

The free field Lagrangians we have been considering so far are all invariant, not only under the proper orthochronous Poincaré group (proper orthochronous Lorentz group plus space-time translations), but also under the three discrete symmetries P, T and C. Basically, for the free lagrangians, the latter symmetries arise *gratis* from the invariance under the proper orthochronous Poincaré group. More specifically, the lagrangians for free scalar fields, free vector fields and free Dirac fields are all invariant under P, T and C, separately, while the lagrangian for the free massless Weyl spinor is invariant under T and under the combined action of parity and charge conjugation CP. The case of the Weyl spinor is special, in that C and P cannot even be independently defined. Indeed parity must act on it by exchanging handedness:

$$\psi_L(x) \in (1/2, 0) \xrightarrow{P} \varepsilon \psi_L^*(x_P) \in (0, 1/2) \tag{9.4}$$

where $x_P = (t, -\vec{x})$. The original field and the parity reflected one transform under $U(1)$ (phase rotations) as:

$$\psi_L \to e^{iq} \psi_L \qquad \varepsilon \psi_L^* \to e^{-iq} \varepsilon \psi_L^* \tag{9.5}$$

We therefore see that the parity transformed field has opposite charge ($q \to -q$): P and C cannot act independently, but only together as a CP transformation. One can check that the free lagrangian for the Weyl field is invariant under this CP transformation. The system is then only invariant under CP but not under P nor C separately!

The presence of the above discrete symmetries in free field theories can however be considered an accidental fact. More generally, while preserving invariance under the proper orthochronous Poincaré group, all the three discrete symmetries can in principle be broken in an interacting theory. There is however a theorem stating that, in a local quantum field theory invariant under the proper orthochronous Poincarè group, the combined transformation:

$$\Theta = CPT \tag{9.6}$$

is always a symmetry. This is the *CPT theorem* (Schwinger '51, Lüders and Pauli '54). These are theoretical results obtained in the QFT framework, but how do they compare to experimental observations? The proper orthochronous Poincaré group as well as Θ are symmetries of Nature as far as we can see. Quantum Electrodynamics (QED) respects P, T and C at leading order as "accidental symmetries" very much like the free Lagrangians respect them. Quantum Chromodynamics (QCD, theory of strong interactions) respects both T and P. This is a surprising puzzle (Strong CP problem) as one could in principle add in QCD an interaction term that breaks T and P, but apparently Nature chose not to do so. However, neither T nor P are symmetries of the Standard Model (SM) of particle physics, which encompasses all the known interactions. Indeed, Electroweak (EW) interactions violate parity in a maximal fashion (as the building blocks of the SM are Weyl fermions for which we can define CP but not P) and violate time reversal through small effects associated to quark flavor violation (for instance in Kaon oscillations).

9.2 Parity

9.2.1 Foreword

In classical mechanics a parity transformation reverses the direction of the canonical variables:

$$P \ : \begin{pmatrix} \vec{q} \\ \vec{p} \end{pmatrix} \mapsto \begin{pmatrix} -\vec{q} \\ -\vec{p} \end{pmatrix} \tag{9.7}$$

Quantities transforming in the same way are called *polar vectors* or simply *vectors*. On the other hand, quantities like the angular momentum have a different transformation rule:

$$P \ : \vec{L} = \vec{q} \wedge \vec{p} \mapsto \vec{L} \tag{9.8}$$

These are called *axial vectors* or *pseudovectors*. From here we can proceed to a canonical quantisation of the system passing from Poisson brackets to canonical commutation relations (CCRs):

$$\{q_i, p_j\} = \delta_{ij} \to [\hat{q}_i, \hat{p}_j] = i\delta_{ij} \tag{9.9}$$

It is obvious that parity transformations will preserve the CCRs and therefore the operation P (which acts on spacetime) must be realised by an unitary operator U_P on the Hilbert space (Wigner's theorem):

$$U_P^\dagger \begin{pmatrix} q_i \\ p_i \end{pmatrix} U_P = \begin{pmatrix} -q_i \\ -p_i \end{pmatrix} \tag{9.10}$$

So in quantum mechanics, parity is realized by unitary transformations, we need now to go to a relativistic theory, and parity acting on the 4-vector x^μ is:

$$x^\mu \to x^\mu_P = P^\mu_\nu x^\nu \quad P^\mu_\nu = \mathrm{Diag}(1, -1, -1, -1) \tag{9.11}$$

So again $P^\mu_\rho P^\rho_\nu = \delta^\mu_\nu$, this implies that if we are considering **faithful** representation of parity, $U_P^2 = \mathbb{1}$. Recall that faithful representation must preserve the group multiplication group $U_P U_P = U_{P^2} = U_{\mathbb{1}} = \mathbb{1}$.

Let see now the representations of parity are indeed faithful, let first look at the action of parity on states and fields:

$$\begin{aligned} &U_P|\{\vec{p}_i\},\{s_i\}\rangle = \prod_i \eta_i |\{-\vec{p}_i\},\{s_i\}\rangle \Rightarrow U_P^2|\{\vec{p}_i\},\{s_i\}\rangle = \prod_i \eta_i^2|\{\vec{p}_i\},\{s_i\}\rangle = \eta^2|\{\vec{p}_i\},\{s_i\}\rangle \\ &U_P^\dagger \phi_a(x) U_P = \hat{P}_{ab}\phi_b(x_P) \implies (U_P^2)^\dagger \phi_a(x) U_P^2 = \hat{P}^2_{ab}\phi(x) \end{aligned} \tag{9.12}$$

The states considered here is an element of the Fock space obtained by applying creation operators on the vacuum, the parity operator will act independently on each of those quantas by flipping the momenta, the spin is an axial vector and remains unchanged under parity and all other quantum number that are not written down are not concerned by parity. The η_i are phase up to which parity is defined and are not known yet. The representation of P will be faithful only if $\eta^2 = 1 \implies \eta_i^2 = 1$ and $\hat{P}^2_{ab} = \delta_{ab}$.
Notice that U_P^2 and $\hat{P}^2_{ab}$ are internal transformations, they do not change the coordinates, then only two choices are possible, either they belong to a Lie group (element of a continuous symmetry) or they are discrete symmetries.
Suppose U_P^2 belongs to a Lie group. Using the exponential map, an element of the connected to the identity part of the group can be written down as $U_P^2 = e^{i\alpha_I Q^I}$, where Q_I are the generators of the Lie group. Now we can define a new operator $\tilde{U}_P = U_P e^{-i\frac{\alpha_I}{2}Q^I}$. Then,

$$\tilde{U}_P^2 = U_P e^{-i\frac{\alpha_I}{2}Q^I} U_P e^{i\frac{-\alpha_I}{2}Q^I} = U_P^2 e^{-i\alpha_I Q^I} = \mathbb{1} \tag{9.13}$$

The second equality holds as parity acts only on spatial components, and the Q_I generators an internal symmetry that will thus only act on the quantum numbers of the state. With this redefined parity $\eta_i^2 = 1$, U_P^2 can always be redefined in such a way that U_P is a faithful representation of parity as long as U_P^2 is an element of a Lie group. However, if U_P^2 is a discrete symmetry, there is no redefinition leading to $U_P^2 = \mathbb{1}$. Previously, we used that the Lie parameter is a continuous real variable, so we could take half the Lie parameter $\alpha/2$ to define the square root of U_P^2 that remains an internal symmetry. In the case of a discrete symmetry, this is no longer possible. To illustrate this, consider the situation where $U_P^2 = (-1)^F$, where F is the fermion number. Then, $U_P^2 = e^{i2\pi J_3}$. U_P^2 can be split into two parts $U_P^2 = e^{i\pi J_3}e^{i\pi J_3}$, but now $e^{i\pi J_3}$ is not an internal symmetry, it will act on the coordinates to by rotation of π along the 3$^{\text{rd}}$ axis! This is precisely what happens to Majorana fermions. Invariance of the mass term of the Majorana fermion Lagrangian $\mathcal{L}^{\text{Majorana}} \supset m\Psi_L^T\epsilon\Psi_L + h.c$ under parity requires that $\Psi_L \to \eta\epsilon\Psi_L^*$ with $\eta = \pm i$. Therefore acting twice with parity leads to:

$$(P^\dagger)^2\Psi_L P^2 = \eta\epsilon P^\dagger(\Psi_L^*)P = \eta\eta^*\epsilon\epsilon\Psi_L = -\Psi_L \tag{9.14}$$

Where we used that $\epsilon\epsilon = -\mathbb{1}$. Beside the Majorana spinor, all the fields transform under faithful representation of parity.
Let consider then the ordinary case where $\hat{P}^2_{ab} = \delta_{ab}$ this implies that $\hat{P}_{ab}$ can be diagonalized. Indeed trough a similarity transformation, the matrix $\hat{P}$ can be brought to the normal Jordan form:

$$\hat{P} \to S\hat{P}S^{-1} = \begin{pmatrix} \hat{P}_1 & 0 & 0 & 0 \\ 0 & \hat{P}_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \hat{P}_n \end{pmatrix} \hat{P}_i = \begin{pmatrix} \eta_1^i & a_1^i & 0 & 0 \\ 0 & \eta_2^i & a_2^i & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \eta_k^i \end{pmatrix} \tag{9.15}$$

Where a_k^i are either 0 for all k or 1. But it is easy to prove that the condition $\hat{P}_i^2 = \mathbb{1}$ requires $a_k^i = 0$ and $\eta_k^i = \pm 1$ for all k. Thus $U_P^\dagger\phi(x)U_P = \eta_a\phi_a(x_P)$ and η_a is called the intrinsic parity of field ϕ_a.

9.2.2 Spin 0

We start by considering the case of a real scalar:

$$\boxed{U_P^\dagger\phi(x)U_P = \eta_P\phi(x_P)} \tag{9.16}$$

where $\eta_P = 1$ corresponds to a *scalar* and $\eta_P = -1$ corresponds to a *pseudoscalar*. Notice that that the Klein-Gordon lagrangian is invariant under the above parity transformation, regardless of the sign of η_P. Consequently

the Klein-Gordon field equation is also invariant, as one can easily check. More generally parity invariance of a lagrangian of the form $(\partial\phi)^2 + V(\phi)$ requires

$$V(\phi) = V(\eta_P \phi) \tag{9.17}$$

which is always true for $\eta_P = 1$, while for $\eta_P = -1$ it requires $V(\phi)$ to be an even function of ϕ.

Focussing on the Klein Gordon field, we now want to figure out how parity acts on the Fock space. In order to do so, consider the plane waves expansion of the fields:

$$U_P^\dagger \phi(x) U_P = \int d\Omega_k \left(U_P^\dagger a_k U_P e^{-ikx} \quad + \quad h.c. \right) = \eta_P \phi(x_P) \tag{9.18}$$

Now it suffices to notice that the scalar product in the plane waves of the RHS can be rewritten simply:

$$k \cdot x_P = k_P \cdot x \tag{9.19}$$

so that by a change of variables in the integral of the LHS $\vec{k} \leftrightarrow -\vec{k}$ $(k \leftrightarrow k_P)$ we obtain :

$$\boxed{\begin{aligned} U_P^\dagger a_k U_P &= \eta_P a_{k_P} \\ U_P^\dagger a_k^\dagger U_P &= \eta_P a_{k_P}^\dagger \end{aligned}} \tag{9.20}$$

The second equation can be obtained either by looking at the hermitian conjugate part of the plane waves expansion or simply by "taking the dagger" of the first equation and recalling that η_P is real.
The case of many real scalar fields is a trivial generalisation of the latter but it is sensible to look at the charged scalar case. We *define* the parity transformation so that it does not flip the charge:

$$U_P^\dagger \phi(x) U_P = \eta_P \phi(x_P) \tag{9.21}$$

$$U_P^\dagger \phi^*(x) U_P = \eta_P \phi^*(x_P) \tag{9.22}$$

instead of:

$$U_P^\dagger \phi(x) U_P = \eta_P \phi^*(x_P) \tag{9.23}$$

Here particles and antiparticles have the same intrinsic parity.
In free field theory the vacuum is invariant under parity transformations:

$$U_P |0\rangle = |0\rangle \tag{9.24}$$

and the quantum parity operator U_P can be explicitly constructed out of the creation and annihilition operators. We can also check that the quantum operators of the Hilbert space transform as expected. For instance the four-momentum:

$$U_P^\dagger P^\mu U_P = \int d\Omega_k k^\mu U_P^\dagger a_k^\dagger a_k U_P = \int d\Omega_k \eta_P^2 k^\mu a_{k_P}^\dagger a_{k_P} = \int d\Omega_k k_P^\mu a_k^\dagger a_k = P_P^\mu \equiv P_\mu \tag{9.25}$$

Similarly, one can check that:

$$\underline{\text{Boosts}}: \qquad U_P^\dagger K^i U_P = -K^i \tag{9.26}$$

$$\underline{\text{Rotations}}: \quad U_P^\dagger J^i U_P = J^i \tag{9.27}$$

So that we have:

$$\boxed{U_P^\dagger J_\pm^i U_P = J_\mp^i} \tag{9.28}$$

9.2.3 Spin 1/2

Let us now focus on just one Dirac spinor. We want to implement a parity transformation on it:

$$U_P^\dagger \psi_\alpha(x) U_P = \eta_P P_{\alpha\beta} \psi_\beta(x_P) \tag{9.29}$$

where η_P is an overall intrisic parity factor. We need to find the matrix $P_{\alpha\beta}$ and the action of parity on b_p^s and d_p^s. We already know from the discussion on the Dirac spinor how parity acts on the latter. Let us recall some facts. Since parity acts on the Lorentz group generators as:

$$P \ : \ J_\pm^i \mapsto J_\mp^i \tag{9.30}$$

we have that:

$$P \ : \ \psi_\alpha \in (j_1, j_2) \mapsto \psi_\alpha^P \in (j_2, j_1) \tag{9.31}$$

This implies that we cannot represent a parity transformed object from (j_1, j_2) into the same basis (unless $j_1 = j_2$). Therefore ψ_L and ψ_R separately do not form a basis for a representation of parity. This is why parity is broken in the Standard Model (SM) as these two representations appear differently in it. However there exists a spinor representation which is also a representation of parity, the Dirac spinor:

$$\Psi_D \equiv \psi_L \oplus \psi_R = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \qquad \Lambda_D = \begin{pmatrix} \Lambda_L & 0 \\ 0 & \Lambda_R \end{pmatrix} \tag{9.32}$$

Any parity transformation will act as:

$$P \ : \ \begin{pmatrix} \psi_L(x) \\ \psi_R(x) \end{pmatrix} \mapsto \eta_P \begin{pmatrix} \psi_R(x_P) \\ \psi_L(x_P) \end{pmatrix} \tag{9.33}$$

So that:

$$\boxed{U_P^\dagger \Psi_D(x) U_P \equiv \eta_P \gamma_0 \Psi_D(x_P)} \tag{9.34}$$

Thus we found our matrix $P_{\alpha\beta}$ and it only remains to find its action on the creation and annihilation operators. We proceed in the exact same way as before. We write (9.34) in terms of plane waves and do a change of variables in the LHS to obtain:

$$U_P^\dagger b_k^r U_P u^r(k) = \eta_p b_{k_P}^r \gamma_0 u^r(k_P) \tag{9.35}$$

$$U_P^\dagger d_k^{\dagger r} U_P v^r(k) = \eta_p d_{k_P}^{\dagger r} \gamma_0 v^r(k_P) \tag{9.36}$$

Now recall the form of the polarisation spinors we obtained:

$$u^s(k) = \begin{pmatrix} \sqrt{k\cdot\sigma}\xi^s \\ \sqrt{k\cdot\bar{\sigma}}\xi^s \end{pmatrix} \tag{9.37}$$

$$v^s(k) = \begin{pmatrix} \sqrt{k\cdot\sigma}\xi^s \\ -\sqrt{k\cdot\bar{\sigma}}\xi^s \end{pmatrix} \tag{9.38}$$

and since $k_P \cdot \sigma \equiv k \cdot \bar{\sigma}$ it is straightforward to see that:

$$\gamma_0 u^s(k_P) = u^s(k) \tag{9.39}$$
$$\gamma_0 v^s(k_P) = -v^s(k) \tag{9.40}$$

Hence we finally obtain:

$$\boxed{\begin{array}{c} U_P^\dagger b_k^s U_P = \eta_p b_{k_P}^s \\ U_P^\dagger d_k^{\dagger s} U_P = -\eta_p d_{k_P}^{\dagger s} \end{array}} \tag{9.41}$$

These equations imply that particles and antiparticles have opposite intrinsic parities, fact that has big phenomenological consequences. Moreover the transformations we just derived imply that:

$$\begin{pmatrix} \vec{S} \\ \vec{P} \end{pmatrix} \xrightarrow{P} \begin{pmatrix} \vec{S} \\ -\vec{P} \end{pmatrix} \tag{9.42}$$

Let us now give an example of parity violation in Weak interactions. Consider the following decay of a pion into a muon and a muon-antineutrino:

$$\pi^- \ \rightarrow \ \mu^- \ + \ \bar{\nu}^\mu \tag{9.43}$$

The pion is a scalar particle and hence has spin 0. Both the muon and its antineutrino are fermions of spin 1/2. Since the pion has no orbital angular momentum in its center-of-mass frame we have:

$$0 = \vec{J}_{init} = \vec{J}_{final} = \sum_i \vec{L}(i) + \vec{S}(i), \quad i = \mu^-, \ \bar{\nu}^\mu \tag{9.44}$$

This equation can be projected on $\vec{p}(\mu^-) = -\vec{p}(\bar{\nu}^\mu) \equiv \vec{p}$. Recalling that $\vec{L}(i) \cdot \vec{p}(i) = 0$ we thus obtain

$$0 = \vec{p} \cdot \left[\sum_i \vec{L}(i) + \vec{S}(i) \right] = \vec{p}(\mu^-) \cdot \vec{S}(\mu^-) - \vec{p}(\bar{\nu}^\mu) \cdot \vec{S}(\bar{\nu}^\mu) \tag{9.45}$$

and conclude that the muon and its antineutrino should have the same helicity

$$h(\mu^-) \equiv \frac{\vec{p}(\mu^-) \cdot \vec{S}(\mu^-)}{|\vec{p}(\mu^-)|} = \frac{\vec{p}(\bar{\nu}^\mu) \cdot \vec{S}(\bar{\nu}^\mu)}{|\vec{p}(\bar{\nu}^\mu)|} \equiv h(\bar{\nu}^\mu) \tag{9.46}$$

Now, as one is easily convinced, the composition of parity with a 180° rotation around any axis orthogonal to $\vec{p}$, leaves the final state momenta unaffected while flipping the helicities. As the original state, a pion at rest, is invariant under such tranformation, if parity were a symmetry we would observe with the same probablity (50%) muons with the two opposite helicities: $P(+1/2) = P(-1/2)$. Instead we only observe muons with helicity $+1/2$, indicating that parity is maximally violated in this process!

This experimental result has a simple theoretical explanation in the SM. There, neglecting the absolutely miniscule effects of the neutrino mass, the neutrino is described by a massless Weyl spinor, for which the particle (the neutrino) has helicity $-1/2$ and the antiparticle (the antineutrino) has helicity $+1/2$: in this case, as we already mentioned, only CP is defined, and neither P nor C make sense independently. The practical manifestation of

the impossibility to define parity is that antineutrinos come only in one helicity, so that by helicity conservation (eq. (9.46)) the accompanying muon must always carry the same helicity $+1/2$. We should add here that experimentally we mostly observe the muon as the neutrino interacts very weakly and is thus very hard to detect.

Four fermions bilinears are mostly important when dealing with the phenomenology of fermions in the Standard Model. Their transformation properties under parity are as follows:

Bilinear	$P\ :\ \cdot \mapsto$	
$\bar{\Psi}\Psi$	$\bar{\Psi}\Psi$	scalar
$\bar{\Psi}\gamma_5\Psi$	$-\bar{\Psi}\gamma_5\Psi$	pseudoscalar
$\bar{\Psi}\gamma_\mu\Psi$	$\bar{\Psi}\gamma^\mu\Psi$	vector
$\bar{\Psi}\gamma_5\gamma_\mu\Psi$	$-\bar{\Psi}\gamma_5\gamma^\mu\Psi$	axial vector

(9.47)

9.2.4 Spin 1

We finally consider spin one particles. Parity acts on vectors as:

$$\boxed{U_P^\dagger V_\mu U_P = \eta_P P_\mu^\nu V_\nu} \tag{9.48}$$

where $\eta_P = 1$ corresponds to a *polar vector* and $\eta_P = -1$ corresponds to an *axial vector*. One more time, we use the plane waves expansion in terms of polarisation vectors to obtain:

$$U_P^\dagger a_k^r U_P \epsilon_\mu(k,r) = \eta_p a_{k_P}^r \epsilon_\mu(k_P,r) \tag{9.49}$$

Recall that the polarisation vectors are given as a pure boost acting on the rest frame vectors (the standard momentum is denoted q here):

$$\epsilon_\mu(k,s) = \lambda(k)_\mu{}^\nu \epsilon_\nu(q,s) \tag{9.50}$$

So that:

$$\epsilon_\mu(k_P,s) = \lambda(k_P)_\mu{}^\nu \epsilon_\nu(q,s) = \lambda^{-1}(k)_\mu{}^\nu \epsilon_\nu(q,s) \tag{9.51}$$

In order to to relate $\epsilon_\mu(k,s)$ to $\epsilon_\mu(k_P,s)$, we need to use some identities of the Lorentz boosts:

$$\lambda^{-1}(k)_\mu{}^\nu = \lambda(k)^\nu{}_\mu \equiv \lambda(p)^\mu{}_\nu \tag{9.52}$$

The second equality comes from the fact that for a boost the matrix is symmetric. Finally recalling that the rest frame polarisation vectors are purely spatial, we have:

$$\epsilon_\nu(q,s) \equiv -\epsilon^\nu(q,s) \tag{9.53}$$

Therefore:

$$\epsilon_\mu(k_P,s) = \lambda(k_P)_\mu{}^\nu \epsilon_\nu(q,s) \equiv -\lambda(k)^\mu{}_\nu \epsilon^\nu(q,s) \equiv -\epsilon^\mu(k,s) \tag{9.54}$$

Using this equation together with (9.48) and (9.49), we finally obtain:

$$\boxed{U_P^\dagger a_k^r U_P = -\eta_p a_{k_P}^r} \tag{9.55}$$

We give now another derivation of the transformation properties of the ladder operators. Consider a gauge in which the vector is purely spatial. Then:

$$\vec{A}(x) = \int d\Omega_k \left(e^{-ikx} \left(\vec{a}_\perp(k) + \frac{\omega}{M} \vec{a}_L(k) \right) \quad + \quad h.c. \right) \tag{9.56}$$

Parity transformations (9.48) become here:

$$\boxed{U_P^\dagger \vec{A}(x) U_P = -\eta_P \vec{A}(x_P)} \tag{9.57}$$

Using the same change of variables as usual, it gives us:

$$\boxed{U_P^\dagger \vec{a}_k U_P = -\eta_P \vec{a}_{k_P}} \tag{9.58}$$

9.3 Time reversal

9.3.1 Foreword

In classical physics time reversal changes the direction of time while leaving space invariant as its name indicates. This implies that the transformation properties of the quantities of interest of the theory are:

$$T \ : \begin{pmatrix} \vec{q} \\ \vec{p} \end{pmatrix} \mapsto \begin{pmatrix} \vec{q} \\ -\vec{p} \end{pmatrix} \tag{9.59}$$

The transformation property of the conjugate momenta can be easily understood as roughly they describe the change in time of their canonical variable. This implies that, for example:

$$T \ : \vec{L} = \vec{q} \wedge \vec{p} \mapsto -\vec{L} \tag{9.60}$$

So that the spin follows the same transformation law:

$$T \ : \vec{S} \mapsto -\vec{S} \tag{9.61}$$

Finally, the reader should have noticed that time reversal does not preserve the Poisson brackets between canonical variables:

$$T \ : \ \{q_i, p_j\} = \delta_{ij} \mapsto \{q_i, -p_j\} = -\delta_{ij} \tag{9.62}$$

Hence, when we quantise the theory the operation T cannot be represented by an unitary operator. Nevertheless, by Wigner's theorem it can be represented by an *antiunitary* operator A_T. Let us recall some definitions:

Definition. *An operator $A : \mathscr{H} \mapsto \mathscr{H}$ is **antilinear** if:*

$$A\left(\alpha|\psi_1\rangle + \beta|\psi_2\rangle\right) = \alpha^* A|\psi_1\rangle + \beta^* A|\psi_2\rangle, \quad \forall \alpha, \beta \in \mathbb{C},\ |\psi_i\rangle \in \mathscr{H} \tag{9.63}$$

Let us list some properties antilinear operators sarisfy:

Properties. *1. If two operators $A, B : \mathscr{H} \mapsto \mathscr{H}$ are antilinear, then their product $AB : \mathscr{H} \mapsto \mathscr{H}$ is a linear operator.*

2. If an operator $A : \mathscr{H} \mapsto \mathscr{H}$ is antilinear, then:

$$\langle\psi_1|A\psi_2\rangle \equiv \langle\psi_2|A^\dagger\psi_1\rangle \tag{9.64}$$

The reader should prove these two properties as an exercise. Finally, we are interested in antiunitary operators:

Definition. *An operator $A : \mathscr{H} \mapsto \mathscr{H}$ is **antiunitary** if it is antilinear and it satisfies:*

$$A^\dagger A = AA^\dagger = 1 \tag{9.65}$$

Now that we have these definitions in hand, we can study certain properties of time reversal transformations. As before $T^2 = \mathbb{1}$ and previous discussion for parity can be extended for time reversal. For faithful representation of time reversal the following holds

$$A_T \cdot A_T = A_{T^2} = A_{\mathbb{1}} = \mathbb{1} \tag{9.66}$$

$$A_T^\dagger = A_{T^{-1}} = A_T \tag{9.67}$$

Operators on the Hilbert space will transform under time reversal as:

$$T : O \mapsto A_T^\dagger O A_T, \quad \forall O \in \mathscr{H} \otimes \mathscr{H}^* \tag{9.68}$$

In particular, covariance of the time evolution of the system is in a one-to-one correspondance with the fact that its Hamiltonian is invariant under time reversal:

$$\begin{aligned}\langle f|e^{-iH\Delta t}|i\rangle =& \langle f|e^{-iA_T H A_T \Delta t}|i\rangle = \langle f|A_T e^{+iH\Delta t} A_T|i\rangle \\ =& \langle e^{+iH\Delta t} A_T i|A_T f\rangle = \langle A_T i|e^{-iH\Delta t}|A_T f\rangle \\ =& \langle i_T|e^{iH\Delta t}|f_T\rangle\end{aligned} \tag{9.69}$$

Where $\Delta t = t_f - t_i$. If T is a symmetry of the system, the amplitude of $i \to f$ is the same as the amplitude associated to the process $f_T \to i_T$ where the sub-index T indicate a time reverse state $|f_T\rangle = A_T|f_T\rangle$

9.3.2 Spin 0

We start again considering the case of a real Klein-Gordon scalar:

$$\boxed{A_T^\dagger \phi(x) A_T = \eta_T \phi(x_T)} \tag{9.70}$$

where $\eta_T = \pm 1$ and $x_T = (-t, \vec{x}) = -x_P$. We want to figure out how parity acts on the Fock space. In order to do so, consider the plane waves expansion of the fields:

$$A_T^\dagger \phi(x) A_T = \int d\Omega_k \left(A_T^\dagger a_k e^{-ikx} A_T \quad + \quad h.c. \right) = \int d\Omega_k \left(A_T^\dagger a_k A_T e^{+ikx} \quad + \quad h.c. \right) = \eta_P \phi(x_T) \tag{9.71}$$

Now it suffices to notice that the scalar product in the plane waves of the RHS can be rewritten simply:

$$k \cdot x_T = -k \cdot x_P = -k_P \cdot x \tag{9.72}$$

so that by a change of variables in the integral of the LHS $\vec{k} \leftrightarrow -\vec{k}$ ($k \leftrightarrow k_P$) we obtain :

$$\boxed{\begin{array}{c} A_T^\dagger a_k A_T = \eta_T a_{k_P} \\ A_T^\dagger a_k^\dagger A_T = \eta_T a_{k_P}^\dagger \end{array}} \tag{9.73}$$

And states transform as:

$$A_T |k\rangle = A_T a_k^\dagger A_T A_T |0\rangle = \eta_T a_{k_P}^\dagger |0\rangle = \eta_T |k_P\rangle \tag{9.74}$$

9.3.3 Spin 1/2

In the case of fermions, time reversal should reverse both spin and momentum. Let us take a guess working with a Weyl fermion:

$$\chi^\alpha = \begin{pmatrix} \chi^1 \\ \chi^2 \end{pmatrix} \tag{9.75}$$

where each component χ^α carries either spin up or down depending on the nature of the field (particle or antiparticle). We would like the transformed field to have opposite spin and hopefully we have a naturally well suited object at our disposal in order to invert the field components, the spinor metric $\epsilon^{\alpha\beta}$. As we previously saw, time reversal of x together with the antilinearity of A_T implied the right transformation on p, thus an educated guess is:

$$A_T^\dagger \chi^\alpha(x) A_T = \eta_T \epsilon^{\alpha\beta} \chi^\beta(x_T) \equiv \chi_T^\alpha \tag{9.76}$$

Let us see why this is a good guess:

- It exchanges the spinor components, which implies:

$$\vec{S} \to -\vec{S} \tag{9.77}$$

- It does not change the charge of the field, under the QED $U(1)$ gauge group:

$$\chi \ \mapsto \ e^{iq}\chi \quad \Leftrightarrow \quad \chi_T \ \mapsto \ e^{iq}\chi_T \tag{9.78}$$

The reader will check as an exercise that this transformation indeed leaves the Weyl Lagragian invariant and will deduce the transformation of the Dirac spinor. In the end, the transformation law of the Dirac field is:

$$\boxed{A_T^\dagger \Psi_D(x) A_T = \eta_T T \Psi_D(x_T) \equiv \eta_T i\gamma_5\gamma^0\gamma^2\Psi_D(x_T)} \tag{9.79}$$

and for the ladder operators:

$$\boxed{\begin{array}{l} A_T^\dagger b_k^s A_T = \eta_T \epsilon^{sr} b_{k_P}^r \\ A_T^\dagger d_k^s A_T = \eta_T \epsilon^{sr} d_{k_P}^r \end{array}} \tag{9.80}$$

9.3.4 Spin 1

The transformation rules of the spin one field will be derived in a simple way below. For further details the reader should refer to the exercises.
As in the case of parity transformations, when the field is expressed only in terms of its spatial components, we have:

$$\boxed{A_T^\dagger \vec{A}(x) A_T = \eta_T \vec{A}(x_T)} \tag{9.81}$$

Using the same change of variables as usual, it gives us:

$$\boxed{A_T^\dagger \vec{a}_k A_T = \eta_T \vec{a}_{k_P}} \tag{9.82}$$

9.4 Charge conjugation

9.4.1 Foreword

Charge is an internal quantity associated with any gauge symmetry of the theory. In particular, in the case of $U(1)$ symmetry (*e.g.* QED), any general field which is charged under the group will be (gauge-)rotated as:

$$\phi_a(x) \overset{U(1)}{\longrightarrow} e^{iq}\phi_a(x) \tag{9.83}$$

Charge conjugation is defined as:

$$C \ : \ \phi_a(x) \mapsto \eta_C C_{ab} \phi_b^*(x) \tag{9.84}$$

$$q \mapsto -q \tag{9.85}$$

with $C^2 = 1$, so that $|\eta_C^2| = 1$. More generally, charge conjugation maps all internal quantum numbers (*e.g.* electric charge, baryon number, lepton number, etc.) into their opposite. Notice that charge conjugation transformations commute with Lorentz transformations as the former do not affect spacetime. It is represented by a unitary operator which as usual has the following properties:

$$U_C \cdot U_C = U_{C^2} = U_1 = 1 \tag{9.86}$$

$$U_C^\dagger = U_{C^{-1}} = U_C \tag{9.87}$$

9.4.2 Spin 0

Firstly we consider the case of scalar fields. Real scalar fields are of no interest as they do not carry any charge. However their complex extension does and charge conjugation is defined as:

$$\boxed{U_C^\dagger \phi(x) U_C = \eta_C \phi^*(x)} \tag{9.88}$$

As it is expected the transformation of the ladder operators is (exercise):

$$\boxed{U_C^\dagger a_k U_C = \eta_C b_k} \tag{9.89}$$

Notice that in general we could have $\eta_C \in \mathbb{C}$, but without loss of generality we can have $\eta_C \in \mathbb{R}$ or even $\eta_C = 1$, indeed we can make a field redefinition:

$$\phi \longrightarrow e^{i\theta}\phi \equiv \phi_\theta \tag{9.90}$$

so that:

$$U_C^\dagger \phi_\theta U_C = e^{i\theta} U_C^\dagger \phi U_C = e^{i\theta}\eta_C \phi^* = e^{2i\theta}\eta_C \phi_\theta^* \tag{9.91}$$

Therefore we can define $e^{2i\theta}\eta_C \equiv \eta_C' \in \mathbb{R}$ or $e^{2i\theta} = \eta_C^*$ so that $\eta_C' = 1$.
Lastly, recall that the complex field can be rewritten as:

$$\phi \equiv \phi_1 + i\phi_2, \quad \phi_i \in \mathbb{R} \tag{9.92}$$

In terms of the two real fields, $U(1)$ phase rotations become $SO(2)$ rotations and charge conjugation acts as:

$$C \; : \; \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \mapsto \begin{pmatrix} \phi_1 \\ -\phi_2 \end{pmatrix} \tag{9.93}$$

Hence charge conjugation acts as a mirror symmetry on the field space about their second component.

9.4.3 Spin 1/2

We start with two Weyl fermions since, as we saw in the introduction, charge conjugation cannot be represented on a single Weyl fermion. Charge keeps handedness but exchanges particles and antiparticles, therefore Weyl spinors will transform as:

$$\boxed{C \; : \; \begin{cases} \psi_L \\ \psi_R \end{cases} \mapsto \begin{matrix} \eta_L \epsilon \psi_R^* \\ \eta_R \epsilon \psi_L^* \end{matrix}} \tag{9.94}$$

That the Weyl Lagrangian is invariant is left as an exercise. Its invariance, considering real ηs implies:

$$\eta_L \eta_R = -1 \tag{9.95}$$

so that without loss of generality $\eta_L = -\eta_R = 1$. The extension of these transformations to Dirac spinors is straightforward:

$$\boxed{C \ : \ \Psi_D \ \mapsto \eta_C i\gamma^2 \Psi_D^* = \eta_C i\gamma^2\gamma^0 \bar{\Psi}_D^T \equiv \eta_C \mathcal{C} \bar{\Psi}_D^T} \tag{9.96}$$

where we define the matrix $\mathcal{C} \equiv i\gamma^2\gamma^0$. The reader should be aware that the matrix $\mathcal{C}$ so defined does not correspond to a proper charge conjugation matrix (as it can be seen when we compute its hermitian conjugate or its square) but is a mathematical trick so that the charge conjugation of the field yields its "bar" conjugate which enters naturally the Dirac Lagrangian. A defining property of the matrix which allows us to compute it in any basis is:

$$\mathcal{C}^T \gamma^{\mu T} \mathcal{C} = -\gamma^\mu \tag{9.97}$$

Equation (9.96) translates itself into the transformation rule for the ladder operators (with $\eta_C = 1$):

$$U_C^\dagger b_k^s U_C = -\epsilon^{sr} d_k^r \tag{9.98}$$

As it is known, if b_k^s destroys a particle of spin s, then d_k^s destroys an antiparticle of spin $-s$, therefore $-\epsilon^{sr} d_k^r$ destroys an antiparticle of spin s as expected. Let us then define new creation and annihilation operators for the antiparticles:

$$-\epsilon^{sr} d_k^r \to \tilde{d}_k^s \tag{9.99}$$

so that $\tilde{d}_k^s$ effectively destroys a antiparticle of spin s. The new transformation rules become:

$$\boxed{U_C^\dagger b_k^s U_C = \tilde{d}_k^s} \tag{9.100}$$

9.4.4 Spin 1

The transformation properties of the vector field can be extracted very easily from the invariance of the QED Lagrangian under charge conjugation. The interacting term connects the four-vector potential to the current:

$$\mathcal{L}_{int} \propto J^\mu A_\mu \tag{9.101}$$

Since by definition the current changes sign under charge conjugation:

$$C \ : \ J^\mu \ \mapsto \ -J^\mu \tag{9.102}$$

we must have:

$$\boxed{U_C^\dagger A^\mu U_C = -A^\mu} \tag{9.103}$$

which straightforwardly leads to:

$$\boxed{U_C^\dagger a_k^r U_C = -a_k^r} \tag{9.104}$$

This shows how the photon is its own antiparticle.

Chapter 10

Interacting Fields

10.1 Asymptotic States

One basic result of our study of free field theories is the structure of their Hilbert space $\mathscr{H}$ of states, the Fock space:

$$\mathscr{H} = \text{span}\,(|0\rangle,\ |k\rangle,\ |k_1, k_2\rangle, \ldots,\ |k_1, \ldots k_n\rangle, \ldots) \tag{10.1}$$

In reality, particles are associated, in position space, with suitable wave packets. The ith particle can be associated with a blob around position x_0^i. For example, for a one particle state we have:

$$|f\rangle = \int d\Omega_k \hat{f}(k)|k\rangle \tag{10.2}$$

where $\hat{f}(k)$ represents the Fourier transform of its wave function. In position space, the distribution probability will be given by $\langle f|\phi^2(x)|f\rangle - \langle 0|\phi^2(x)|0\rangle \sim |f(x - x_0)|^2$ as depicted in Fig. 10.1.

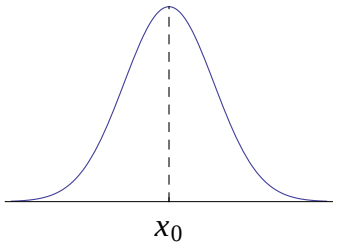


Figure 10.1: Blob associated to one particle in position space

One crucial feature of free field theory is that the presence of one particle does not affect the others, *i.e.* the superposition principle applies exactly. Indeed, the equations of motion of a free field theory are linear in the fields, so that if ϕ_1 and ϕ_2 are two solutions, so is $\phi = \phi_1 + \phi_2$, see the equations for a free real scalar field and Dirac fermions for instance. This implies that even if two wave packets travel towards each other, Fig. 10.2, after overlapping their shape and direction is not affected and they will pursue their journey as if they never crossed, Fig. 10.3

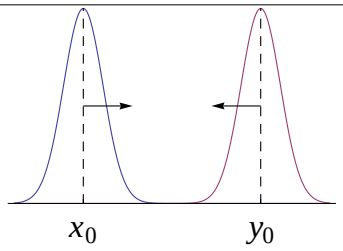


Figure 10.2: Two wave packets travelling towards each other that will overlap in future.

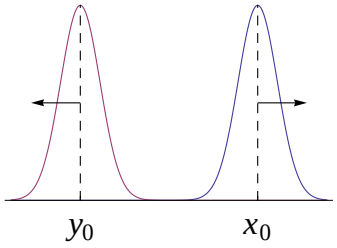


Figure 10.3: After crossing each other, the wave packets follow their original path without being disturbed.

In physical reality, however, particles do interact and the wave packets of the previous example are affected by each other. On the other hand, if the range of the interaction is finite or decreases fast enough with distance, it is intuitive and reasonable to expect that the states are well described by individual particle wave packets of a free field theory, provided the mutual separation of these wave packets is sufficiently large. In other words, well separated wave packets can be expand into a Fock space of states characterised by a set of particles $\{A\}$ with quantum numbers $\{m_A,\ s_A,\ Q_A, \ldots\}$ and the Lagrangian describing it evolution can be written as

$$\mathcal{L}_0 = \sum_A \mathcal{L}_0^A \quad \Rightarrow \quad \mathcal{H}_0 = \sum_A \mathcal{H}_0^A \tag{10.3}$$

Where by $\mathcal{L}_0^A$ ($\mathcal{H}_0^A$) we indicate the free Lagrangian (Hamiltonian) density describing particle A. Supported by empirical observation and by direct computation whenever applicable (basically in quantum field theories that can be studied with the use of perturbation theory), what we stated just above will be our basic working assumption in the following.

With this idea in mind, we can pick a convenient basis for the Hilbert space $\mathscr{H}_0 = \{|\phi_\alpha\rangle\}$ where states $|\phi_\alpha\rangle$ are defined as eigenstates of the momentum operator P_0^μ:

$$P_0^\mu |\phi_\alpha\rangle = p_\alpha^\mu |\phi_\alpha\rangle \tag{10.4}$$

The Label 0 is here to recall that we consider here the free Hamiltonian $P_0^0 = H_0$ and α denotes all discreet quantum numbers such as mass or quantum charges but also continuous one as relative angles between momenta in a multi particle state. Well separated wave packets are described by superpositions of these states:

$$|\phi\rangle = \int d\alpha\, g(\alpha) |\phi_\alpha\rangle \tag{10.5}$$

It is essential to think of wave packets; momentum eigenstates that we picked as a basis are fully delocalised and two states will overlap everywhere, our construction become useless as interactions cannot be neglected anymore. The typical situation which is realised in particle physics experiments is the following: at $t \to -\infty$ all separated wave packets (well described by the above state) move towards the collision point. At finite time, interactions

(perhaps very complicated) take place and our description using a Fock space breaks down. Finally, at $t \to +\infty$, the wave packets separate away from the collision point and interactions can be neglected again. In general the outgoing state is different with respect to the ingoing one, it is therefore a different superposition of $|\phi_\alpha\rangle$.

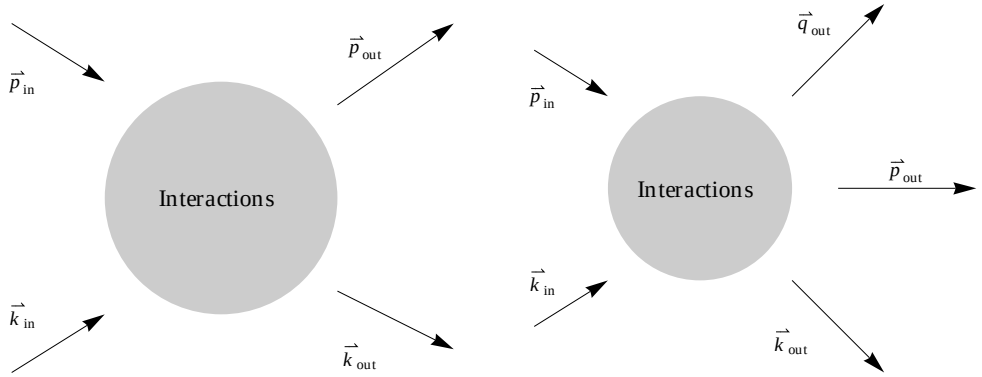


Figure 10.4: Different scattering processes.

We could for example consider the processes for which:

$$|\phi_\alpha\rangle_{in} = \int d\Omega_{p_{in}} d\Omega_{k_{in}} \hat{f}_1(p_{in}) \hat{f}_2(k_{in}) |p_{in}, k_{in}\rangle \tag{10.6}$$

We can then have as a final state:

$$|\phi_\alpha\rangle_{out} = \int d\Omega_{p_{out}} d\Omega_{k_{out}} \hat{f}'_1(p_{out}) \hat{f}'_2(k_{out}) |p_{out}, k_{out}\rangle \tag{10.7}$$

as in the LHS of Fig. 10.4, but we could also have a final state with a different number of particles: whereas its final state is:

$$|\phi_\alpha\rangle_{out} = \int d\Omega_{p_{out}} d\Omega_{k_{out}} d\Omega_{q_{out}} \hat{f}'_1(p_{out}) \hat{f}'_2(k_{out}) \hat{f}'_3(q_{out}) |p_{out}, k_{out}, q_{out}\rangle \tag{10.8}$$

We can formalise this physically interesting situation by defining two complete sets of states of **the full interacting theory**, the *in* states $|\psi^+_\alpha\rangle$ and the *out* states $|\psi^-_\alpha\rangle$ in the following way:

- The *in* states $|\psi^+_\alpha\rangle$ are such that for $t \to -\infty$, they become well described by $|\phi_\alpha\rangle$.
- The *out* states $|\psi^-_\alpha\rangle$ are such that for $t \to +\infty$, they become well described by $|\phi_\alpha\rangle$.

These are the so-called *asymptotic states*. In interacting theories, we have in general $|\psi^+_\alpha\rangle \neq |\psi^-_\alpha\rangle$. For example, when we start before the interactions with two field configurations ϕ_1, ϕ_2, we have in general $\phi_i \not\to \phi_i$ after the interaction took place. The above statement can be made more formal:

$$e^{-iHt} \int d\alpha \; g(\alpha) |\psi^\pm_\alpha\rangle \overset{t \to \mp\infty}{\longrightarrow} e^{-iH_0 t} \int d\alpha \; g(\alpha) |\phi_\alpha\rangle \tag{10.9}$$

where H and H_0 are the fully interacting and free Hamiltonian respectively. Let us make a few comments. We have kept the wave packet smearing $d\alpha \; g(\alpha)$ in order to stress that our states must be localised in time. Indeed, if they were localised in momentum (*i.e.* eigenstates of the Hamiltonian), they would not evolve in time (stationarity

of the Hamiltonian eigenstates). However, the equality holds for any $g(\alpha)$ so that we could as well more sloppily write $e^{-iHt}|\psi_\alpha^\pm\rangle \overset{t\to\mp\infty}{\longrightarrow} e^{-iH_0t}|\phi_\alpha\rangle$, so that:

$$\boxed{|\psi_\alpha^\pm\rangle = \lim_{t\to\mp\infty} e^{iHt}e^{-iH_0t}|\phi_\alpha\rangle \equiv \Omega(\mp\infty)|\phi_\alpha\rangle} \tag{10.10}$$

where we defined $\Omega(t) = e^{iHt}e^{-iH_0t}$ and $\Omega(\pm\infty)$ are called the *Möller wave operators*. Let us derive an important property of the latter. If the limit in Eq. 10.10 exists, it should not be affect by a shift in the origion of time (after all we chose $t_0 = 0$ by convention, but $t_0 \neq 0$ is an equally good origin). Then:

$$\Omega(\mp\infty) = \lim_{t\to\mp\infty} e^{iH(t-t_0)}e^{-iH_0(t-t_0)} \tag{10.11}$$

Since the limit is independent of the origin of time, we have:

$$0 = \frac{\partial}{\partial t_0}\Omega(\mp\infty) = -iH\Omega(\mp\infty) + i\Omega(\mp\infty)H_0 \tag{10.12}$$

$$\Rightarrow H\Omega(\mp\infty) = \Omega(\mp\infty)H_0 \tag{10.13}$$

This has the following important implication:

$$\boxed{H|\psi_\alpha^\pm\rangle = H\Omega(\mp\infty)|\phi_\alpha\rangle = \Omega(\mp\infty)H_0|\phi_\alpha\rangle = E_\alpha|\psi_\alpha^\pm\rangle} \tag{10.14}$$

So that the asympotic states are eigenstates of the full interacting Hamiltonian when their corresponding free states are eigenstates of the free Hamiltonian, and both have the same eigenvalue. This fact is reassuring as it is natural to expect energy conservation when comparing both sets of states. The Möller wave operators transform a free state into its interacting analogue without changing the energy. Physically:

$$\underbrace{H_0|\phi_\alpha\rangle = E_\alpha|\phi_\alpha\rangle}_{\text{Purely kinetic, no interaction}} \qquad \underbrace{H|\psi_\alpha^\pm\rangle = E_\alpha|\psi_\alpha^\pm\rangle}_{\text{Kinetic and potential}} \tag{10.15}$$

What physically happens is that the free states give up some kinetic energy for the interaction. One last property is that since the Möller operators are built up from the limit of an unitary operator, they are themselves unitary operators, so that the asymptotic states are normalised in the same way as the free states:

$$\langle\psi_\alpha^\pm|\psi_\beta^\pm\rangle = \langle\phi_\alpha|\Omega^\dagger(\mp\infty)\Omega(\mp\infty)|\phi_\beta\rangle = \langle\phi_\alpha|\phi_\beta\rangle = \delta(\alpha-\beta) \tag{10.16}$$

For sake of clarity we can redefine[1]:

$$\Omega_\pm \equiv \Omega(\mp\infty) \tag{10.17}$$

10.1.1 Lippmann-Schwinger Equation

We will here derive an explicit solution for the Möller wave operators. Start defining:

[1]Note that $\Omega_\pm$ so defined are the Möller wave operators one encounters in the litterature

$$H_I \equiv H - H_0 \tag{10.18}$$

From Eq. 10.14, we have:

$$(E_\alpha - H)\,|\psi_\alpha^\pm\rangle = 0 \quad \Rightarrow \quad (E_\alpha - H_0)\,|\psi_\alpha^\pm\rangle = H_I|\psi_\alpha^\pm\rangle \tag{10.19}$$

We see from this equation that in the limit where $H_I \to 0$ we have $|\psi_\alpha^\pm\rangle \to |\phi_\alpha\rangle$. This leads us to guess a solution of the form $|\psi_\alpha^\pm\rangle = a^\pm|\phi_\alpha\rangle + \Delta|\psi_\alpha^\pm\rangle$. Plugging it into Eq. 10.19 we get:

$$(E_\alpha - H_0)\left(\cancel{a^\pm|\phi_\alpha\rangle} + \Delta|\psi_\alpha^\pm\rangle\right) = H_I|\psi_\alpha^\pm\rangle \tag{10.20}$$

$$\Rightarrow \Delta|\psi_\alpha^\pm\rangle = \frac{H_I}{E_\alpha - H_0}|\psi_\alpha^\pm\rangle \tag{10.21}$$

Added to this the limit where $H_I \to 0$ which implies $a^\pm = 1$, we obtain:

$$\boxed{|\psi_\alpha^\pm\rangle = |\phi_\alpha\rangle + \frac{H_I}{E_\alpha - H_0 \pm i\epsilon}|\psi_\alpha^\pm\rangle} \tag{10.22}$$

where the $\pm i\epsilon$ factor is there to regulate the pole. This formal solution to the asymptotic states is the *Lippmann-Schwinger* equation. Recall that the kets entering the equation are actually wavepackets. Then the prescription $\pm i\epsilon$ is important and chosen so that (Fig. 10.5):

$$e^{-iHt}\int d\alpha\; g(\alpha)|\psi_\alpha^\pm\rangle \overset{t\to\mp\infty}{\longrightarrow} e^{-iH_0t}\int d\alpha\; g(\alpha)|\phi_\alpha\rangle \tag{10.23}$$

The different possible choices follow from Cauchy's theorem. Explicitly:

$$\begin{aligned} e^{-iHt}\int d\alpha\; g(\alpha)|\psi_\alpha^\pm\rangle &= \int d\alpha\; g(\alpha)e^{-iE_\alpha t}|\psi_\alpha^\pm\rangle = \int d\alpha\; g(\alpha)e^{-iE_\alpha t}\left(|\phi_\alpha\rangle + \frac{H_I}{E_\alpha - H_0 \pm i\epsilon}|\psi_\alpha^\pm\rangle\right) \\ &= \int d\alpha\; g(\alpha)e^{-iE_\alpha t}|\phi_\alpha\rangle + \int d\alpha d\beta\; g(\alpha)\frac{e^{-iE_\alpha t}}{E_\alpha - E_\beta \pm i\epsilon}|\phi_\beta\rangle \underbrace{\langle\phi_\beta|H_I|\psi_\alpha^\pm\rangle}_{T^\pm_{\beta\alpha}} \end{aligned} \tag{10.24}$$

where $T^\pm_{\beta\alpha}$ is called the transition matrix and its meaning will become clear shortly. Now for the *in* states we must have:

$$e^{-iHt}\int d\alpha\; g(\alpha)|\psi_\alpha^+\rangle \overset{t\to-\infty}{\longrightarrow} e^{-iH_0t}\int d\alpha\; g(\alpha)|\phi_\alpha\rangle \tag{10.25}$$

and therefore the second term in Eq. 10.24 must vanish. Consider a Cauchy integral in the complex plane. In our equation, we are integrating over real variables, but we can continue them analytically on the complex plane. If we want to determine the value of the integral on the real line, we choose a contour which contains the real line and, for example, a semi-circle which we choose to close either in the upper or the lower-half plane. This choice depends on where the poles of the integral lay. Indeed, if they lay in the upper-half plane, we will choose to close our contour in the lower-half plane so that we can argue, using Cauchy's theorem, that the value of the integral on the real line is equal to the value of the integral on the infinite semi-circle (no poles inside the contour). In our case we have the freedom to choose where the poles lay (through the $\pm i\epsilon$ prescription) and thefore we will first focus on the fact that we want the integral to vanish.

Figure 10.5: Asymptotic states and free states relation.

As $t \to -\infty$, the exponential will behave as $\sim e^{iE_\alpha|t|}$ so that we need a positive imaginary part to make the Cauchy integral converge and vanish. This implies that we will choose to close the contour in the upper-half plane. Then, to make the integral vanish, the poles must lay in the lower-half plane. Assuming smooth $g(\alpha)$ and $T^{\pm}_{\beta\alpha}$ the poles of the $d\alpha$ integral become relevant and are exponetially cancelled at finite E_α as $t \to -\infty$ if we choose the prescription $+i\epsilon$. Indeed, this is achieved by selecting $E_\alpha = E_\beta - i\epsilon$. Similarly, we choose the $-i\epsilon$ prescription for the *out* states.

On the other hand, as $t \to +\infty$ the $+i\epsilon$ contribution does *not* vanish. The Cauchy integral will be evaluated as the sum of the residues and we obtain:

$$e^{-iHt}\int d\alpha\ g(\alpha)|\psi^+_\alpha\rangle \overset{t\to+\infty}{\longrightarrow} \int d\alpha\ g(\alpha)e^{-iE_\alpha t}\left(|\phi_\alpha\rangle - 2\pi i\int d\beta\ \delta(E_\alpha - E_\beta)T^+_{\beta\alpha}|\phi_\beta\rangle\right). \tag{10.26}$$

This expression will be useful in the next section.

10.2 The S-Matrix

In a typical scattering experiment we are interested in measuring the transition probability from a given initial state to a final one. An initial state is prepared such that long before the interaction process happens it has a particle content labelled by a complete set of quantum numbers α. Thus *in*-states $|\psi^+_\alpha\rangle$ constitute a natural basis for the initial state of the process we want to measure. Similarly long after the interaction took place we measure a system with a particle content β, thus we use *out*-states $|\psi^-_\beta\rangle$ as basis for final states.
We define the *S-Matrix* to be the amplitude for finding the final result $|\psi^-_\beta\rangle$, starting with $|\psi^+_\alpha\rangle$. Thus the *S-Matrix* is just the following scalar product:

$$\boxed{S_{\beta\alpha} = \langle\psi^-_\beta|\psi^+_\alpha\rangle}. \tag{10.27}$$

Few remarks are in order. In the absence of interactions particles propagate freely. In this case *in* and *out* states coincide, $|\psi^+_\alpha\rangle = |\psi^-_\alpha\rangle$ and we would just have

$$S_{\beta\alpha}\big|_{H_I=0} = \delta(\alpha - \beta). \tag{10.28}$$

We stress that, in general, $|\psi^+_\alpha\rangle$ and $|\psi^-_\beta\rangle$ are just two different complete bases of the Hilbert space. The S-matrix thus just expresses the expansion of the elements of one basis in terms of the other one. Thefore it is *unitary* in the sense that

$$\int d\beta\ S^*_{\beta\gamma}S_{\beta\alpha} = \int d\beta\ \langle\psi^+_\gamma|\psi^-_\beta\rangle\langle\psi^-_\beta|\psi^+_\alpha\rangle = \langle\psi^+_\gamma|\psi^+_\alpha\rangle = \delta(\gamma - \alpha). \tag{10.29}$$

Similary $S\cdot S^\dagger = 1$.
We can associate $S_{\beta\gamma}$ to the matrix elements of an operator S if we work with the states $|\phi_\alpha\rangle$. Since the latter are related to the *in* and *out states* through the Moller operators (10.10), we easily obtain

$$S_{\beta\alpha} \equiv \langle\phi_\beta|S|\phi_\alpha\rangle \tag{10.30}$$

where we defined

$$\boxed{S = \Omega^\dagger(+\infty)\Omega(-\infty) \equiv U(+\infty, -\infty)} \tag{10.31}$$

This expression will be useful in the next section to examine Lorentz invariance but also later to find a formula in perturbation theory.

Using the Lippmann-Schwinger equation we can also write S in a more explicit way. First of all notice the trivial set of equalities

$$S_{\beta\alpha} = \langle \psi^-_\beta | \psi^+_\alpha \rangle = \langle \psi^-_\beta | \underbrace{e^{iHt} e^{-iHt}}_{=\mathbb{1}} | \psi^+_\alpha \rangle \equiv \langle \psi^-_\beta(t) | \psi^+_\alpha(t) \rangle \qquad \forall t. \tag{10.32}$$

We can take the limit $t \to \infty$ in the last expression:

$$S_{\beta\alpha} = \lim_{t\to+\infty} \langle \psi^-_\beta(t) | \psi^+_\alpha(t) \rangle. \tag{10.33}$$

In this limit we can use the definition of *out* states for $\langle \psi^-_\beta(t)|$ and the Lippmann-Schwinger equation in the form (10.26) for $|\psi^+_\alpha(t)\rangle$:

$$\langle \psi^-_\beta(t)| \xrightarrow{t\to+\infty} \langle \phi_\beta | e^{iE_\beta t}, \tag{10.34}$$

$$|\psi^+_\alpha(t)\rangle \xrightarrow{t\to+\infty} e^{-iE_\alpha t} \left(|\phi_\alpha\rangle - 2\pi i \int d\beta\ \delta(E_\alpha - E_\beta) T^+_{\beta\alpha} |\phi_\beta\rangle \right). \tag{10.35}$$

Then, plugging in (10.33) and using $\langle \phi_\beta | \phi_\alpha \rangle = \delta(\alpha - \beta)$, we obtain

$$S_{\beta\alpha} = \langle \phi_\beta | e^{iE_\beta t} \left[e^{-iE_\alpha t} \left(|\phi_\alpha\rangle - 2\pi i \int d\beta\ \delta(E_\alpha - E_\beta) T^+_{\beta\alpha} |\phi_\beta\rangle \right) \right] \tag{10.36}$$

$$\Longrightarrow \qquad \boxed{S_{\beta\alpha} = \delta(\alpha - \beta) - 2\pi i \delta(E_\alpha - E_\beta) T^+_{\beta\alpha}} \tag{10.37}$$

Notice that, in light of this discussion, equation (10.26) can be rewritten as

$$|\psi^+_\alpha\rangle \xrightarrow{t\to+\infty} \int d\beta\ S_{\beta\alpha} |\phi_\beta\rangle. \tag{10.38}$$

Indeed, what happens is that the *in* states, when evolved up until infinite positive times will have passed through the interactions and thus we do not simply recover their corresponding free states but rather a superposition of them. The coefficients mixing free states components correspond to the S-Matrix (matrix-)elements, Fig. 10.6.

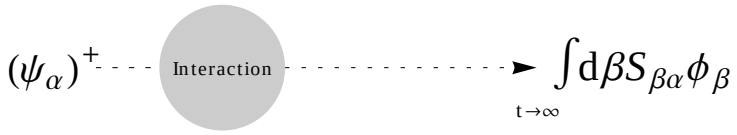


Figure 10.6: Asymptotic states and their relation to the S-Matrix.

We already commented that, if there is no interaction, then $S_{\beta\alpha} = \delta(\alpha - \beta)$. The *transition rate* is thus defined to be proportinal to:

$$|S_{\beta\alpha} - \delta(\alpha - \beta)|^2 \sim \left| T^+_{\beta\alpha} \right|^2 \tag{10.39}$$

Finally, as in non-relativistic quantum mechanics, we can derive the *Born approximation*. Indeed, for weak interactions we have:

$$T^+_{\beta\alpha} = \langle \phi_\beta | H_I | \psi^+_\alpha \rangle \simeq \langle \phi_\beta | H_I | \phi_\alpha \rangle \equiv V_{\beta\alpha} \tag{10.40}$$

so that :

$$S^{(1)}_{\beta\alpha} = \delta(\alpha - \beta) - 2\pi i \delta(E_\alpha - E_\beta) V_{\beta\alpha} \tag{10.41}$$

where the upperscript (1) denotes the first order Born approximation.

10.2.1 Symmetries of the S-Matrix

What is meant by the invariance of the S-Matrix under different symmetries and what are the conditions on the Hamiltonian that will ensure such invariance properties?

Poincaré Invariance:
Poincaré transformations act on spacetime as:

$$\mathbb{P} \ : \ x \ \rightarrow \ \Lambda x + a \tag{10.42}$$

We define an unitary operator $U(\Lambda, a)$ for the interacting theory by how it acts on *either in* or *out* states. It should act in the same way as $U_0(\Lambda, a)$ (unitary operator of Poincaré transformations of the free theory) does on the free states. Our interacting theory will be called Poincaré invariant if the *same* $U(\Lambda, a)$ acts on *both* sets of asymptotic states as $U_0(\Lambda, a)$ acts on the free states. In other words, the interactions do not affect the action of the Poincaré generators. We consider here for simplicity a theory of one interacting scalar field. We define $U_0(\Lambda, a)$ such that its action on the free states $|\phi_\alpha\rangle \equiv |\{p_i\}\rangle$ is given by:

$$U_0(\Lambda, a)|\{p_i\}\rangle = e^{ia_\mu \Lambda^\mu{}_\nu \sum_i p_i^\nu} |\{\Lambda p_i\}\rangle \tag{10.43}$$

and therefore, from the unitarity of $U(\Lambda, a)$ we obtain the Poincaré *covariance* of the S-Matrix:

$$S(\{k_i\};\{p_j\}) = \langle\{k_i\}, out|\{p_j\}, in\rangle = \langle\{k_i\}, out|U^\dagger \cdot U|\{p_j\}, in\rangle = e^{ia_\mu \Lambda^\mu{}_\nu \left(\sum_j p_j^\nu - \sum_i k_i^\nu\right)} S(\{\Lambda k_i\};\{\Lambda p_j\}) \tag{10.44}$$

where $|\{p_j\}, in\rangle \equiv |\psi_\alpha^+\rangle$ and similarly for the *out* state and we required them both to transform in the same way as the free states. Since the LHS is independent of a, so must be the RHS. In other words:

$$\boxed{S \neq 0 \ \Leftrightarrow \ P_i \equiv \sum_j p_j^\nu = \sum_i k_i^\nu \equiv P_f} \tag{10.45}$$

This equation allows us to factorise a four-dimensional delta-function in the definition of the S-Matrix and rewrite:

$$\boxed{S_{\beta\alpha} = \delta_{\alpha\beta} - (2\pi)^4 \delta^{(4)}(P_f - P_i) i\mathcal{M}_{\beta\alpha}} \tag{10.46}$$

We finally need to extract the condition on the Hamiltonian. In order to do so, recall:

$$\langle\psi_\beta^-|\psi_\alpha^+\rangle = S_{\beta\alpha} = \langle\phi_\beta|S|\phi_\alpha\rangle \tag{10.47}$$

Since the *in* and *out* states transform under "interacting" Poincaré transformations in the same way as the free states do under the "free" transformations, we have:

$$\langle\psi_\beta^-|\underbrace{U^\dagger \cdot U}_{1}|\psi_\alpha^+\rangle = \langle\phi_\beta|\underbrace{U_0^\dagger \cdot U_0 \cdot S \cdot U_0^\dagger \cdot U_0}_{S}|\phi_\alpha\rangle = \langle\phi_\beta|U_0^\dagger \cdot S \cdot U_0|\phi_\alpha\rangle \tag{10.48}$$

where the last equality is imposed. This means that for the fields of the full theory to transform in the same way with $U(\Lambda, a)$ as the free fields with $U_0(\Lambda, a)$, we must have:

$$[S, U_0(\Lambda, a)] = 0 \tag{10.49}$$

Therefore the set of free generators $G_0 = \{H_0, \vec{P}_0, \vec{J}_0, \vec{K}_0\}$ all commute with the scattering matrix. We can then define the *exact* generators $G = \{H, \vec{P}, \vec{J}, \vec{K}\}$ that generate the transformations of the asymptotic states. Their group structure will be the one of the Poincaré group. For interacting theories we have in general:

$$H_0 \to H_0 + H_I = H \tag{10.50}$$
$$\vec{P} = \vec{P}_0 \tag{10.51}$$
$$\vec{J} = \vec{J}_0 \tag{10.52}$$
$$\vec{K} \neq \vec{K}_0 \tag{10.53}$$

The first two consequences are:

$$\left[\vec{J}, H\right] = 0 = \left[\vec{P}, H\right] \quad \to \quad \left[\vec{J}_0, H_I\right] = 0 = \left[\vec{P}_0, H_I\right] \tag{10.54}$$

so that the interactions should be rotational and translational invariant. Also from $\left[\vec{K}, \vec{P}\right] \sim H$, since we already have $\vec{P} = \vec{P}_0$, we must have $\vec{K} = \vec{K}_0 + \Delta\vec{K} \neq \vec{K}_0$ and finally, from $\left[\vec{K}, H\right] \sim \vec{P} = \vec{P}_0$, we have:

$$\left[\vec{K}, H_I\right] + \left[\Delta\vec{K}, H_0\right] = 0 \tag{10.55}$$

This equation for the exact boost generators allows us to build them. Finally, since we require the asymptotic states to transform in the same way as the free states we obtain, denoting $U|\psi^{\pm}_{\alpha}\rangle = u_\alpha|\psi^{\pm}_{\alpha}\rangle$ and $U_0|\phi_\alpha\rangle = u_\alpha|\phi_\alpha\rangle$:

$$\langle\psi^{\pm}_{\beta}|U|\psi^{\pm}_{\alpha}\rangle = \begin{cases} u_\alpha\delta(\alpha-\beta) = \langle\phi_\beta|U_0|\phi_\alpha\rangle \\ \langle\psi^{\pm}_{\beta}|\Omega_{\pm}\cdot\Omega^{\dagger}_{\pm}\cdot U\cdot\Omega_{\pm}\cdot\Omega^{\dagger}_{\pm}|\psi^{\pm}_{\alpha}\rangle = \langle\phi_\beta|\Omega^{\dagger}_{\pm}\cdot U\cdot\Omega_{\pm}|\phi_\alpha\rangle \end{cases} \tag{10.56}$$

so that:

$$U\Omega_{\pm} = \Omega_{\pm}U_0 \tag{10.57}$$

This is the generalisation of Eq. 10.14 which implies that for any generator of the Poincaré group, we have:

$$\boxed{G\Omega_{\pm} = \Omega_{\pm}G_0} \tag{10.58}$$

Internal symmetry:
Suppose that our theory respect some additional symmetry, it can be discrete as parity or can be associated to some conserved quantum number, one can say $U(1)_{EM}$ associated to the electric charge for definiteness. When acting on the states $|\phi_\alpha\rangle$, the symmetry will rotate the quantum numbers gather under the global index α:

$$U(T)|\phi_\alpha\rangle = \mathcal{D}(T)_{\alpha\alpha'}|\phi_{\alpha'}\rangle \tag{10.59}$$

Where $U(T)$ is the unitary operator associated to the symmetry transformation T and $\mathcal{D}(T)_{\alpha\alpha'}$ is it realisation on the state. Consider then the S-matrix element associated to transform states:

$$\mathrm{S}_{\beta\alpha} = \langle\psi^{-}_{\beta'}|\mathcal{D}^{\dagger}(T)_{\beta\beta'}\mathcal{D}(T)_{\alpha\alpha'}|\psi^{+}_{\alpha'}\rangle = \mathcal{D}^{*}(T)_{\beta'\beta}\mathcal{D}(T)_{\alpha\alpha'}\mathrm{S}_{\alpha'\beta'} \tag{10.60}$$

Here we have implicitly assume that the transformation is realized in exactly the same way on both *in* and *out* states. For this to hold we must assume that there is $U_0(T)$ that induces these transformation into the free states just as in equation 10.59. Then using that the definition of the asymptotic states:

$$|\psi^{\pm}_{\alpha}\rangle = \Omega(\mp\infty)|\phi_\alpha\rangle \to U^{\pm}(T)|\psi^{\pm}_{\alpha}\rangle = \Omega(\mp\infty)U^0(T)|\phi_\alpha\rangle \tag{10.61}$$

If $[\Omega(\mp\infty), U^0(T)] = 0$, then we have that $U^0(T) = U^{\pm}(T)$, so the unperturbed operator U^0 must commute with both H_0 and H_I, the symmetry must leave invariant the free theory and the interactions independently.

10.3 Phenomenology

The square of an S-matrix element represents the probability to go from an asymptotic initial state to an asymptotic final state. Experimental observations at colliders, such as the LHC, are stated in terms of suitable observables that provide a measure of that probability. The purpose of the two next subsections is to describe the two main observables, the cross-section and the decay rate. For further discussions one can refer to Chapter 4 of Peskin & Schroeder's textbook, *An Introduction to Quantum field theory* .

10.3.1 Cross-sections

Let us consider a simplified particle physics experiment where a beam of B-particles interacts with a target of T-particles. To have a picture in mind, the beam and the target are represented in Fig. 10.7 by two bunches of particles (shaped like rectangular bricks) with respectively lengths l_B and l_T, and common surface S:

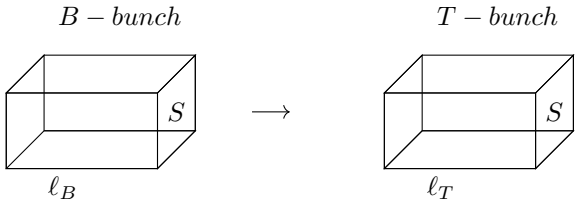


Figure 10.7: Pictorial view of the two interacting bunches of particles, we choose to be in the rest frame of the target beam

Let us call $\rho_{B,T}$ the particle density of respectively the beam and the target bunch. Assuming the beam and the target are sufficiently diluted that only interactions of pairs (one B and one T) contribute, we expect the total number of collisions N after the bunches have crossed to satisfy the proportionality[2]

$$N \propto \rho_B \ell_B \rho_T \ell_T S = \frac{N_B N_T}{S} = L_i \tag{10.62}$$

The quantity L_i, called the **integrated luminosity**, has the dimension of an inverse area and gives a measure of the combined intensity of beam and target. As the number of events N is dimensionless, the above relation should be completed by a proportionality constant σ with the dimension of an area

$$N = \sigma \frac{N_B N_T}{S} = \sigma L_i \tag{10.63}$$

The quantity σ, called the **cross-section**, measures the interaction probability between a beam particle B and a target particle T. Notice that the cross-section σ measures an intrinsic physical properties of the particles in question, while the integrated luminosity L_i measures the experimental conditions. In simple words σ is set by Nature while L_i is set by the Experimenter.
For a single event, many possible outcomes are possible. The differences can concern both the kinematical variables and the very particle content of the final state. In the parlance of particle physics each particular particle content in the final state is termed a *channel*. Consider for instance electron-electron scattering, where both B and T, consist of electrons e^-. At low energy, the most likeley outcome is the reaction of Møller scattering: $e^-e^- \to e^-e^-$. Indeed, as we repeat the experiment, the outgoing electrons will follow some distribution in the angular variables. Møller scattering is termed an *elastic channel*, because the initial and final state content is identical. But it can also happen that the final state involves different particles with respect to the initial state. One terms those cases as *inelastic channels*. At low enough energy the most common inelastic channel, beyond e^-e^-, involves the emission of an additional photon, i.e. the reaction $e^-e^- \to e^-e^-\gamma$. However at sufficiently high energies one can even have channels that do not involve electrons for $e^-e^- \to \nu_e\nu_e\pi^-\pi^-$ which features two neutrinos and two pions in the final state. In order to characterize physics in a more faithful way, rather than simply counting the total

[2]This formula is easily understood by noticing for instance that by doubling any of its factors corresponds to performing two identical experiments tghuse obtaining twice as many events. Notice also that in a more general situation where the beam and target buches have surfaces $\mathcal{B}$ and $\mathcal{T}$ of different area, what would matter is just the value S of the overlap surface $\mathcal{B} \cap \mathcal{T}$.

number of events, it is then better to sort out the final states according to their particle content and according to the value of the kinematic variables (momenta and, when measurable, spins). Indicating by N_a the number of events in channel a, we can then write $N = \sum_a N_a$. Moreover we can decompose each N_a in the contribution of the different kinematical regions

$$N = \sum_a N_a = \sum_a \int \mathrm{d}N_a, \quad \mathrm{d}N_a = \frac{\mathrm{d}N_a(p_1, ..., p_{n_a})}{\mathrm{d}\Omega_1 ... \mathrm{d}\Omega_{n_a}} \mathrm{d}\Omega_1 ... \mathrm{d}\Omega_{n_a}, \quad \mathrm{d}\Omega_i = \frac{\mathrm{d}^3 k_i}{(2\pi)^3 2E_i}\,. \tag{10.64}$$

The number of events dN_a, in each phase space cell of each channel, is also an observable and obviously one that contains more information than the total number of events N. Like for eq. (10.63) one can then associate to the differential number of events dN_a a differential cross section $d\sigma_a$ according to

$$\mathrm{d}\sigma_a = \frac{\mathrm{d}N_a}{L_i}\,. \tag{10.65}$$

The output of observations in particle physics experiments are mainly expresses in terms of differential cross sections.

The discussion above focussed on the total number of event that happen during the whole duration of the experiment. One might also be interested in the number of event happening per unit time. That simply corresponds to the time derivative of N in eq. (10.63)

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \sigma \frac{\mathrm{d}L_i}{\mathrm{d}t}, \quad \frac{\mathrm{d}L_i}{\mathrm{d}t} = \rho_B \frac{\mathrm{d}\ell_B}{\mathrm{d}t} \rho_T \ell_T S = \rho_B v_B \rho_T \ell_T S \equiv L \tag{10.66}$$

where v_B is the beam velocity (the target is assumed at rest). The quantity L is called the (instantaneous) luminosity. In the first equation above, we also used that σ is time independent, while in the second we used that the only time dependence is given by the motion of the B-bunch across the T-bunch, which is effectively captured by $d\ell_B/dt = v_B$.

having defined the physical observable, the cross-section, we must now derive its relation to the more formal object, the scattering amplitude.

According to the above discussion, we are interested in the interaction of two incoming particles, one from the beam and one in the target. In a real experiment several of these interactions take place, as both beam and target involve many particles. However as the interaction of pairs are independent of one another, we can picture the outcome by considering many repetitions of the interaction of a single beam particle and a single target particle. We must also allow the relative separation, or impact parameter, of the two incoming wave packets to vary each time according to a certain distribution, which is determined by the experimental conditions. For simplicity we can consider the state of the target particle T to be fixed, while the B-particle wave packets follow some distribution in impact parameter. Fig. 10.8 offers a picture of the set-up we have in mind. We characterize the B particle initial trajectories in terms of their crossing point on some chosen 2D plane. Any choice for the plane is good as long as the initial B trajectory does not lie in it. To make things simpler we could chose the plane orthogonal to the trajectories, but that is not strictly necessary. We parametrize the crossing points of the incoming B particles by a two dimensional impact parameter vector $\vec{b}$. The beam will thus be described by an integrated flux $F(\vec{b})$ of B-particles across each point on the plane (in other words, the total number of B-particles that, at the end of the many repetitions of the two particle scattering, will have crossed an infinitesimal area $\mathrm{d}^2 b$ at $\vec{b}$ is simply $F(\vec{b})\mathrm{d}^2 b$). It is convenient, but not necessary, to choose coordinates on the plane such that $\vec{b} = 0$ corresponds to a B-trajectory that maximally overlaps with the wavepacket of the target particle. In that case $|\vec{b}|$ does indeed measure the impact parameter.

The state of each particle is described by some wave packet function. We will assume that all B-particles have the same wavepacket, modulo a translation. The initial state of one T-particle and one B-particle crossing at $\vec{b}$ can then be written as

$$|\phi_{in}(\vec{b})\rangle = \int \mathrm{d}\Omega_B \mathrm{d}\Omega_T \phi_B(k_B) e^{-i\vec{k}_B \cdot \vec{b}} \phi_T(k_T) |k_B, k_T\rangle \equiv |\phi_B, \phi_T\rangle_{\vec{b}} \tag{10.67}$$

where $e^{-i\vec{k}_B \cdot \vec{b}}$ realizes the translation by $\vec{b}$ of the B-particle position (and thus of its initial trajectory)[3] . One has

$$_{\vec{b}}\langle \phi_B, \phi_T | \phi_B, \phi_T \rangle_{\vec{b}} = \int \mathrm{d}\Omega_B |\phi_B(k_B)|^2 \times \int \mathrm{d}\Omega_T |\phi_T(k_T)|^2 \tag{10.68}$$

[3] As $\vec{b}$ is a vector on our chosen reference plane, by $\vec{k}_B \cdot \vec{b}$ we mean $\vec{k}_B^{\perp} \cdot \vec{b}$ with $\vec{k}_B^{\perp}$ the projection of $\vec{k}_B$ on the plane.

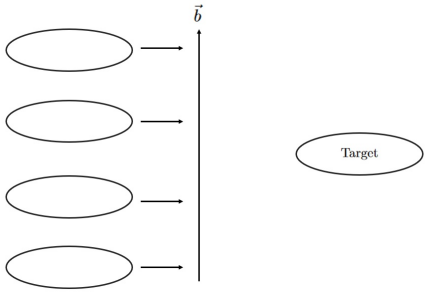


Figure 10.8: Schematic representation of the experiment

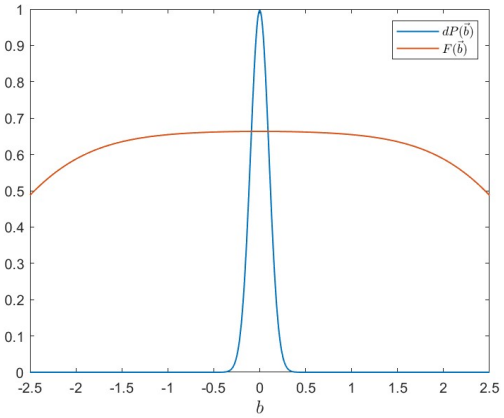


Figure 10.9: Situation of interest for the beam distribution $F(\vec{b})$ and for the differential probability controlled by the region of interaction

so that the proper normalization is obtained by taking both factors in the right-hand side of the above equation equal to 1.

We want to consider the differential probability for the initial state $|\phi_B, \phi_T\rangle_{\vec{b}}$ to evolve into some general final state $|\{p_i\}_{i=1,\dots,n}, \{s_i\}_{i=1,\dots,n}\rangle$. Here $\{p_i\}_{i=1,\dots,n}$ denotes the momenta of the outgoing particles and $\{s_i\}_{i=1,\dots,n}$ is a global label to indicate the type of particle, it spin and so on. In what follows, to simplify the notation we will only characterize the final states by the particles' momentum: $\equiv |\{p_i\}_{i=1,\dots,n}\rangle$. The properly normalized differential probability per unit cell of the phase space is found to be

$$\mathrm{d}P(T, B \to \{p_i\}_{i=1,\dots,n})_{\vec{b}} = |\langle\{p_i\}_{i=1,\dots,n}|S|\phi_B, \phi_T\rangle_{\vec{b}}|^2 \mathrm{d}\Omega_1 \dots \mathrm{d}\Omega_n \,. \tag{10.69}$$

Indeed one can easily show that, given any region Ω of the $n-$particle phase space, the operator

$$\mathbb{P}_\Omega \equiv \int_\Omega |\{p_i\}_{i=1,\dots,n}\rangle\langle\{p_i\}_{i=1,\dots,n}| \mathrm{d}\Omega_1 \dots \mathrm{d}\Omega_n \tag{10.70}$$

is a projector on the set of states $\in \Omega$. That amounts to checking that

$$\mathbb{P}_\Omega^2 = \mathbb{P}_\Omega \,, \qquad \mathbb{P}_\Omega \mathbb{P}_{\Omega'} = 0 \tag{10.71}$$

where Ω' is any set such that $\Omega \cap \Omega' = \emptyset$. By the axioms of quantum mechanics, the probability to measure the final state $S|\phi_B, \phi_T\rangle_{\vec{b}}$ as an $n-$particle state in the phase space cell $\mathrm{d}\Omega_1 \dots \mathrm{d}\Omega_n$ around $\{p_i\}_{i=1,\dots,n}$ is then given by eq. (10.69).

Given the differential probability 10.69 and the integrated flux $F(\vec{b})$, the differential number of events is then

$$\mathrm{d}N(T, B \to \{p_i\}_{i=1,\dots,n}) = \int \mathrm{d}^2\vec{b} F(\vec{b}) \, \mathrm{d}P(T, B \to \{p_i\}_{i=1,\dots,n})_{\vec{b}} \to F(0) \int \mathrm{d}^2\vec{b} \, \mathrm{d}P(T, B \to \{p_i\}_{i=1,\dots,n})_{\vec{b}} \,, \tag{10.72}$$

where in the last step we have assumed, which is normally the case, that $F(\vec{b})$ is much more spread than $\mathrm{d}P_{\vec{b}}$ (see Fig.10.9) so that it can be approximated as constant in the region where $\mathrm{d}P_{\vec{b}} \neq 0$, i.e. $F(\vec{b}) \simeq F(0)$. This corresponds to the situation where the interaction range between B-particles and T-particles is (to a good extend) finite and much smaller than the size of the beam. Notice now that in our present situation $F(0)$ is nothing but the integrated luminosity L_i. This can be checked by considering eq. (10.62) and noticing that in the present case we only have one target particle, $N_T = 1$, while $F(0)$ corresponds to the integrated flux of beam particles per unit area, that is N_B/S. Applying eq. (10.65) we then have

$$\mathrm{d}\sigma(T, B \to \{p_i\}_{i=1,\dots,n}) = \int \mathrm{d}^2\vec{b}\, \mathrm{d}P(T, B \to \{p_i\}_{i=1,\dots,n})_{\vec{b}} \tag{10.73}$$

We must now develop the above equation using the general form of $\mathrm{d}P(T, B \to \{p_i\}_{i=1,\dots,n})_{\vec{b}}$ and performing the integral over the impact parameter. Let us focus first on the matrix element in eq. (10.69) and write it in terms of the reduce S-matrix element as:

$$\begin{aligned}\langle\{p_i\}_{i=1,\dots,n}|S|\phi_T\phi_B\rangle_{\vec{b}} &= \int \langle p_1, \cdots p_n|S|k_T k_B\rangle\, \phi_T\left(k_T\right)\phi_B\left(k_B\right) e^{-ik_B\cdot b} d\Omega_T d\Omega_B \\ &= \int (2\pi)^4\delta^4\left(P_f - k_T - k_B\right)\mathcal{M}(TB \to f)\phi_T\left(k_T\right)\phi_B\left(k_B\right) e^{-ik_B\cdot b} d\Omega_T d\Omega_B\end{aligned} \tag{10.74}$$

where $P_f = \sum_i^n p_i$ is the total 4-momentum of the final state. Taking its squared absolute value according to eq. (10.69) we thus get:

$$\begin{aligned}d\sigma = \int d^2b \prod_i^n d\Omega_i \int d\Omega_T d\Omega_B d\bar{\Omega}_T d\bar{\Omega}_B e^{-i\left(k_B - \bar{k}_B\right)\cdot b}\phi_T\left(k_T\right)\phi_T^*\left(\bar{k}_T\right)\phi_B\left(k_B\right)\phi_B^*\left(\bar{k}_B\right) \\ (2\pi)^4\delta^4\left(\bar{k}_T + \bar{k}_B - P_f\right)(2\pi)^4\delta^4\left(k_T + k_B - P_f\right)\mathcal{M}(T, B \to f)\mathcal{M}(\bar{T}, \bar{B} \to f)^*\end{aligned} \tag{10.75}$$

The integral over b leads to ad addtional delta-function, $(2\pi)^2\delta^2(k_B^\perp - \bar{k}_B^\perp)$ (we recall that the upper script $\perp$ indicates the momentum component lying on the $\vec{b}$-plane). The 6 constraints from $\delta^2(k_B^\perp - \bar{k}_B^\perp)\delta^4\left(\bar{k}_T + \bar{k}_B - P_f\right)$ can be used to easily integrate over $\bar{k}_{B,T}$, as these indeed consist of 6 variables. Leaving aside the other factors, to perform this intergration it suffices to focus on the integration measure for the initial particles and on the delta-functions

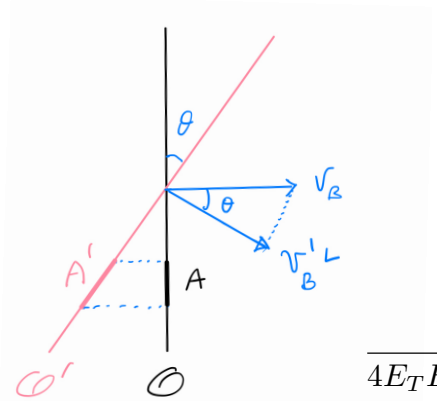
$$\begin{aligned}&\int d\Omega_T d\Omega_B d\bar{\Omega}_T d\bar{\Omega}_B (2\pi)^{10}\delta^2(k_B^\perp - \bar{k}_B^\perp)\delta^4\left(\bar{k}_T + \bar{k}_B - p_f\right)\delta^4\left(k_T + k_B - p_f\right) = \\ &\int \left[(2\pi)^6\delta^2(k_B^\perp - \bar{k}_B^\perp)\delta^4\left(\bar{k}_T + \bar{k}_B - k_T + k_B\right)\right]\frac{d^3\bar{k}_B d^3\bar{k}_T}{(2\pi)^6 4\bar{E}_B\bar{E}_T}\left((2\pi)^4\delta^4\left(k_T + k_B - P_f\right)d\Omega_T d\Omega_B\right)\end{aligned} \tag{10.76}$$

Now, one obvious solution of the constraints imposed by the delta functions in square brackets is $\bar{k}_B = k_B$, $\bar{k}_T = k_T$. As one can check there is also another one where $\bar{k}_B \neq k_B$ and $\bar{k}_T \neq k_T$. However the contribution of this second solution can be made arbitrarily small by taking the wave functions ϕ_B and ϕ_T to be arbitrarily peaked around respectively some momenta $\vec{p}_B$ and $\vec{p}_T$. Basically in that case the product $\phi_B(k_B)\phi_T(k_T)\phi_B^*(\bar{k}_B)\phi_B^*(\bar{k}_T)$ is peaked at $k_B \sim \bar{k}_B \sim p_B$ and $k_T \sim \bar{k}_T \sim p_T$, so that only the solution $\bar{k}_B = k_B$, $\bar{k}_T = k_T$ contributes significatively. Neglecting thus the second solution, explicit integration then gives:

$$\begin{aligned}\int \delta^2(k_B^\perp - \bar{k}_B^\perp)\delta^4\left(\bar{k}_T + \bar{k}_B - k_T + k_B\right)\frac{d^3\bar{k}_B d^3\bar{k}_T}{4\bar{E}_B\bar{E}_T} &= \int dk_T^L dk_B^L \delta(\bar{k}_T^L + \bar{k}_B^L - p_f^L)\delta(\bar{E}_T + \bar{E}_B - E_f) \\ &= \int \frac{d\bar{k}_T^L}{4E_B E_T}\delta\left(\sqrt{\bar{k}_T^2 + m_T^2} + \sqrt{(k_B^L)^2 + \left(\bar{k}_T^L - p_f^L\right)^2 + m_B^2} - E_f\right) \\ &= \frac{1}{4E_B E_T |v_B^L - v_T^L|}\end{aligned} \tag{10.77}$$

where by L superscript on a vector indicates its component orthogonal to the $\vec{b}$-plane. In the last step we used $\int \delta(f(x))dx = \sum_i \frac{1}{|f'(x_i)|}$ where the x_i are zeros of f. We can then gather our results:

$$d\sigma = \prod_i^n d\Omega_i \int d\Omega_T d\Omega_B \left|\phi_T\left(k_T\right)\right|^2 \left|\phi_B\left(k_B\right)\right|^2 (2\pi)^4\delta^4\left(P_f - k_T - k_B\right) \mid \mathcal{M}(T, B \to f)|^2 \frac{1}{\left|v_T^L - v_B^L\right| 2E_T 2E_B} \tag{10.78}$$



$$F' = F\cos\theta$$

$$A' = \frac{A}{\cos\theta}$$

$$v'^{L}_{B} = v^{L}_{B}\cos\theta$$

$$\frac{1}{4E_T E_B v'^{L}_{B}} = \frac{1}{4E_T E_B v^{L}_{B}}\frac{1}{\cos\theta}$$

Figure 10.10: Here we schematically show two different choices of $\vec{b}$-plane operated respectively by observer $\mathcal{O}$ and observer $\mathcal{O}'$. Their planes are relatively rotated by an angle θ in the 13 plane (direction 2 is orthogonal to the picture). As shown, the flux factors transform precisely like the area intercepted by a beam moving in the 3 direction.

As we said we are interested in wave-functions peaked around some specific momentum, which we call $p_{B,T}$ for the beam and the target respectively. For instance we could take a suitably normalized Gaussian wave packet

$$d\Omega_T|\phi_T(k_T)|^2 = \frac{d^3k_T}{(2\pi)^{3/2}\Delta^3}e^{-\frac{(k_T-p_T)^2}{2\Delta^2}}\,. \tag{10.79}$$

and similalarly for ϕ_B. In the limit where Δ is much smaller than the momentum scales over which the rest of intergrand (10.78) varies significantly, the above measure behaves like a delta function

$$\frac{d^3k_T}{(2\pi)^{3/2}\Delta^3}e^{-\frac{(k_T-p_T)^2}{2\Delta^2}} \xrightarrow{\Delta\to 0} \delta^3(k_T-p_T)d^3k_T\,. \tag{10.80}$$

One may be worried that the integrand, as it still involves a $\delta^4\left(P_f-k_T-k_B\right)$, is not a smooth function. However after integrating over any finite region of the final phase space the result will be a smooth function, and, in pracxtice, our measured quantities always correspond to integrating over such finite region, given the resolution with which the final momenta are measured is finite. We can thus safely make the replacement expressed by eq. (**??**) for both ϕ_T and ϕ_B. This greatly simplifies the expression for the differential cross-section, which we can write as the product of three factors

$$d\sigma = \left[\frac{1}{\left|v^L_T - v^L_B\right| 2E_T 2E_B}\right]\left[|\mathcal{M}(T,B\to f)|^2\right]\left[(2\pi)^4\delta^4\left(P_f-k_T-k_B\right)\prod_f d\Omega_f\right] \tag{10.81}$$

The third factor in square brackets, usually called the "phase-space", is Lorentz invariant, like the second factor, the squared matrix element. However the first factor, called the flux factor, is clearly not Lorentz invariant. In order to see that, we can for instance choose the L direction aligned with the third axis of our reference frame. We have then

$$\left|v^L_T - v^L_B\right| 2E_T 2E_B = 4|E_B p^3_T - E_T p^3_B| = |\varepsilon^{12\mu\nu}p_{B,\mu}p_{T,\nu}|\,, \tag{10.82}$$

which shows that the flux factor coincides with the 12 component C_{12} of the Lorentz tensor $C_{\alpha\beta} = \varepsilon_{\alpha\beta\mu\nu}p_{B,\mu}p_{T,\nu}$. This result clarifies that this factor is only invariant under a subgroup of the Lorentz group: boosts along 3 and rotations in the plane (1,2). In particular the flux factor is not invariant under rotations that do not leave the $\vec{b}$ plane invariant. One can easily check that, under such rotations, the flux factor precisely transforms like an area, as one should have expected. A hopefully self-explanatory example is presented in Fig. (10.10). However we can define a Lorentz invariant cross-section by picking the $\vec{b}$ plane and a convenient class of reference frames where

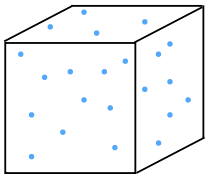


Figure 10.11

the velocities of both target and beam are aligned with the longitudinal direction: $\vec{v}_B \parallel \vec{v}_T \parallel 3$, or equivalently $\vec{v}_B^{\perp} = \vec{v}_T^{\perp} = 0$. In that case the only non-zero component of the tensor $C_{\alpha\beta}$ is the (1,2), so that we can write

$$\left|v_T^L - v_B^L\right| 2E_T 2E_B = 4|E_B p_T^3 - E_T p_B^3| = 4\sqrt{\frac{C_{\alpha\beta}C^{\alpha\beta}}{2}} = 4\sqrt{(p_B \cdot p_T)^2 - m_T^2 m_B^2} \tag{10.83}$$

This leads to the expression of the invariant differential cross-section:

$$\boxed{d\sigma = \frac{1}{4\sqrt{(p_B \cdot p_T)^2 - m_T^2 m_B^2}} |M(T, B \to f)|^2 \, d\Phi^{(n)}} \tag{10.84}$$

Where by $d\Phi^{(n)}$ we succintly indicate the **Lorentz invariant phase space** factor (LIPS)

$$\boxed{d\Phi^{(n)} = \prod_f d\Omega_f (2\pi)^4 \delta^{(4)}(P_f - P_i)} \tag{10.85}$$

Obviously eq. (10.84) can be computed in any frame, but its proper interpretation as a cross section, according to the definition in eq. (10.73), only works in the above mentioned special subclass of frames (and choice of $\vec{b}$ plane). For any other choice the cross section is given by the invariant result in eq. (10.84) times a correcting factor $\sqrt{C_{\alpha\beta}C^{\alpha\beta}}/\sqrt{2}|C_{12}|$. In the class of frames where eq. (10.84) two choices are of practical relevance. One is the is the center of mass frame, where $\vec{p}_T + \vec{p}_B = 0$, and the other is the rest frame of one of the two particles, say the target T.

10.3.2 Lifetime and decay rate

The existence of unstable particles is an empirical fact, which we should be able to describe in QFT. Notice though that such particle will never be observed as outgoing in the far future ($t \to +\infty$) because it will disentegrate before then. So it is not, strictly speaking, an asymptotic state that we can describe with the S-matrix. Let us proceed nonetheless, we will later find a fix to this objection.

Empirically the phenomenon of particle decay works as follows. We imagine to have reservoir containing at a certain time t a number $N_D(t)$ of unstable (decaying) particles D (see fig.10.11). The decay rate Γ and lifetime τ are then defined as

$$\Gamma = \frac{\#\text{decays per unit time}}{\#\text{D particles in sample}}, \qquad \Gamma \equiv \frac{1}{\tau}. \tag{10.86}$$

For sufficient dilution of the sample, which is practically always the case, the process of decay of each particle is independent of the presence of the others. In that situation Γ measures the probability of decay per unit time,

and in system satisfying time translation it should therefore be time independent. The number of particles in the sample $N_D(t)$ will then satisfy

$$\frac{dN_D(t)}{dt} = -\Gamma N_D(t) \qquad \Rightarrow \qquad N_D(t) = e^{-\Gamma t} N_D(0) \tag{10.87}$$

Like for particle collisions, different channels can contribute to the decay. In full analogy with the case of the cross section, we can then write

$$\Gamma = \frac{1}{N_D}\frac{dN_D}{dt} = \frac{1}{N_D}\sum_a \frac{dN_D^a}{dt} \equiv \sum_a \Gamma_a \tag{10.88}$$

where the Γ^a respresent the partial rates in each different a-channel. Again, like before, we can define the differential rates according to

$$\Gamma_a = \int d\Gamma_a = \int \frac{d\Gamma(p_1, \dots, p_{n_a})}{d\Omega_1 \dots d\Omega_{n_a}} d\Omega_1 \dots d\Omega_{n_a} \,. \tag{10.89}$$

The partial rate is tied to the branching ratio $\mathrm{B}r(a) \equiv \Gamma_a/\Gamma$, which measures the probability for the particle to deacy in a specific channel. Indeed, given an initial sample of $N_D(0)$ D-particles, the number of decays in the a-channel at each time is $\dot{N}_a(t) = \Gamma_a N_D(t)$, which when integrated gives

$$N_a\Big|_{\text{total}} = \int_0^\infty \dot{N}_a(t)dt = \frac{\Gamma_a}{\Gamma} N_D(0) \equiv \mathrm{Br}(a) N_D(0) \,. \tag{10.90}$$

Let us now try to connect the differential decay rate to S-matrix elements, like we did for the cross-section. The initial state will be a properly normalized ($\langle \Phi_D | \phi_D \rangle = 1$) wave packet state with a single D-type particle

$$|\phi_D\rangle = \int d\Omega_D \phi_D(k_D) |k_D\rangle \,. \tag{10.91}$$

In the same spirit of our previous discussion, the amplitude for decay in a generic state with particles is given by the S-matrix element

$$\langle p_1 \dots p_n | S | \phi_D \rangle \tag{10.92}$$

so that the differential decay probability is

$$\begin{aligned} dP(D \to f) &= \prod_i^n d\Omega_i \, |\langle p_1, \dots, p_n | S | \phi_D \rangle|^2 \\ dP(D \to f) &= \prod_i^n d\Omega_i \int d\Omega_D d\bar{\Omega}_D (2\pi)^8 \delta^{(4)}(P_f - k_D) \delta^{(4)}(P_f - \bar{k}_D) \mathcal{M}(D \to f) \mathcal{M}(\bar{D} \to f)^* \phi_D(k_D) \phi_D^*(\bar{k}_D) \,, \end{aligned} \tag{10.93}$$

where in the second equation we have used the reduced matrix element. Of course, our approach here seems, at best, heuristic, given, as we already said, that a decaying particle is not an asymptotic state. Let us continue nonetheless.

When integrating over the 3-momentum $\bar{k}_D$, the 3-momentum delta's fix $\vec{\bar{k}}_D = \vec{k}_D = \vec{P}_f$. Given the energy of the D particle is fixed by its 3-momentum according to $E_D = \sqrt{m_D^2 + k_D^2}$, the equality of the spacial momenta implies $E_D = \bar{E}_D$, so that in the end we obtain

$$dP(D \to f) = \prod_i^n d\Omega_i \int \frac{d\Omega_D}{2E_D} (2\pi)^3 \delta^{(3)}(\vec{P}_f - \vec{k}_D) \underbrace{(2\pi)^2 \delta(E_f - E_D)^2}_{= 2\pi\delta(E_f - E_D) 2\pi\delta(0)} |\mathcal{M}(D \to f)|^2 |\phi_D(k_D)|^2 \tag{10.94}$$

What does $2\pi\delta(0)$ mean? The appearance of this infinite factor is the price we are paying because we blindly applied the S-matrix formalism to a particle that decays, which was not meant to be. How should we interpret this infinite factor? Can we fix this apparent paradox? In fact we can, as we now show. Even though it may not be manifest at this stage, a singular term like that can only come from some divergent integral. Indeed when defining the S-matrix we let the initial and final times to go to infinity, and, in particular, a delta of energy differences can only come from integrals of the form

$$2\pi\delta(E_f - E_D) = \int_{-\infty}^{+\infty} dt \; e^{i(E_f - E_D)t} \,. \tag{10.95}$$

Working over a finite time span T we can formally write

$$2\pi\delta(0) = \lim_{E_f\to E_D} 2\pi\delta(E_f - E_D) = \lim_{E_f\to E_D}\lim_{T\to\infty}\int_{T/2}^{T/2} dt\; e^{i(E_f-E_D)t} \sim \lim_{T\to\infty} T \tag{10.96}$$

so that at finite time we can identify $2\pi\delta(0) = T$ (notice that this is fully analogous to what done in section 4.2.1, where we identified $(2\pi)^3\delta^3(0) = V$). We can then proceed, admittedly heuristically, as follows. As the D is unstable, we cannot take the $T\to\infty$ limit. Intuitively we can only treat D as a genuine, though approximate, asymptotic state for times that are much shorter than its lifetime τ_D. we must then assume $T \ll \tau_D$. Working at finite T it thus seems fair to replace $2\pi\delta(0) \to T$ in eq. (10.94). Yet we have the other $\delta(E_f - E_D)$ factor in that equation which at finite T is actually equal to

$$2\pi\delta_T(E_f - E_D) = \int_{T/2}^{T/2} dt\; e^{i(E_f-E_D)t} = \frac{sin[(E_f - E_D)T]}{E_f - E_D}\,. \tag{10.97}$$

Now we would like to still effectively treat this term as $2\pi\delta(E_f - E_D)$. That is possible, as long as the integrand in eq. (10.94) is a sufficiently smooth function. More precisely, indicating by ΔE_{dyn} the energy scale, typical of the dynamics and encoded in the matrix element and phase space, over which the integrand changes, we can make the identification $\delta_T(E_f - E_D) \to \delta(E_f - E_D)$ as long as $T \gg 1/\Delta E_{dyn}$. A natural interpretation of $1/\Delta E_{dyn}$ is as the time scale that roughly controls the creation of a D-particle. For T in the range

$$\frac{1}{\Delta E_{dyn}} \ll T \ll \tau_D \tag{10.98}$$

we can thus interpret eq. (10.94) as a time dependent decay probability and write

$$dP(D\to f)(T) = \int \frac{d\Omega_D}{2E_D}|\phi_D(k_D)|^2\mathcal{M}(D\to f)|^2\prod_i^n d\Omega_i(2\pi)^4\delta^{(4)}(P_f - k_D)\;\times T\,. \tag{10.99}$$

As the time dependence is purely linear, the obvious intepretation is that this equation gives the differential decay rate $dP(D\to f)(T)/T$. Taking the initial wavepacket localized around a 3-momentum p_D like we did in eq. (10.80), we then arrive at our final result for the partial decay rate

$$\boxed{d\Gamma(D\to f) = \frac{|\mathcal{M}(D\to f)|^2}{2E_D}d\Phi^{(n)}}\,. \tag{10.100}$$

We should stress that in order for the above line of reasoning to make sense it is necessary that $\frac{1}{\Delta E_{dyn}} \ll \tau_D$: only for long lived particles can we really talk about decay rates as well as successfully adapt the S-matrix formalism to compute. Notice that the discussion we made here is precisely the one which is given in ordinary QM when deriving Fermi's Golden Rule.

In order to determine the total decay rate Γ we must finally integrate over all possible final states and sum over all channels. Schematically the result will have the form

$$\Gamma = \frac{1}{2E_D}\times(\text{Lorentz invariant term}) \tag{10.101}$$

As the lifetime satisfies $\tau = 1/\Gamma$ we then have

$$\tau = 2E_D\times(\text{Lorentz invariant term}) = 2m_D\gamma_D\times(\text{Lorentz invariant term}) \tag{10.102}$$

where $\gamma_D = 1/\sqrt{(1-v_D^2)}$ is the relativistic γ-factor that expresses the dilation of time intervals between the rest frame of the particle and the lab frame where the particle moves with velocity v_D. Our result then shows that the lifetime depends on the reference frame precisely as dictated by special relativity.

10.4 Amplitudes in Perturbation Theory

We saw in the last section that quantities of interest in quantum field theory are all dependent on the amplitude squared extracted from the corresponding S-Matrix element. We derive now a perturbative expression for those amplitudes.

10.4.1 Perturbative Expansion

In order to understand why we use perturbation theory, let us start with an example. Consider a real scalar field Lagrangian which has $\mathbb{Z}_2$ symmetry[4]:

$$\mathcal{L} = \underbrace{\frac{1}{2}(\partial_\mu\phi)^2 - \frac{m_0^2}{2}\phi^2}_{\mathcal{L}_0} \underbrace{-\frac{\lambda_0}{4!}\phi^4}_{\mathcal{L}_I} \tag{10.103}$$

The Hamiltonian of the system is:

$$H = \underbrace{\frac{1}{2}\pi^2 + \frac{1}{2}(\vec{\nabla}\phi)^2 + \frac{m_0^2}{2}\phi^2}_{H_0} + \underbrace{\frac{\lambda_0}{4!}\phi^4}_{H_I} \tag{10.104}$$

At leading order in λ_0, this is a good choice for the interaction Hamiltonian. The system is weakly coupled as long as:

$$\frac{\lambda_0}{8\pi^2} \ll 1 \tag{10.105}$$

In this regime, we make the hypothesis that we can work in perturbation theory in λ_0, so that we can write for the S-Matrix operator:

$$S = 1 + \lambda_0 S^{(1)} + \lambda_0^2 S^{(2)} + \dots \tag{10.106}$$

Some processes come at the leading order (LO) (at order λ_0 in the Lagrangian or more precisely at order λ_0^2 in the S-Matrix element as we square the amplitude of the process), and at that order we see that the splitting is good. However, at next to leading order (NLO) the splitting is not good! Indeed, interactions lead to corrections to the so-called *bare* parameters of the Lagrangian (m_0, λ_0) as well as to a renormalisation of the field through a parameter called the *field renormalisation strength* ($\phi \to Z^{1/2}\phi$). This implies for example that the *physical* mass will get a dependence on the coupling constant:

$$m^2 = m_0^2(1 + a\lambda_0 + b\lambda_0^2 + \dots) \tag{10.107}$$

And therefore to obtain a good splitting of our Lagrangian, we define $Z = 1+\delta_Z$, $\quad \delta_m = m_0^2 Z - m^2$, $\quad \delta_\lambda = \lambda_0 Z^2 - \lambda$ where m and λ are the observable quantities. The Lagrangian becomes:

$$\mathcal{L} = \underbrace{\frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2}_{\mathcal{L}_0} \underbrace{-\frac{\lambda}{4!}\phi^4 + \frac{1}{2}\delta_Z\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}\delta_m\phi^2 - \frac{\delta_\lambda}{4!}\phi^4}_{\mathcal{L}_I} \tag{10.108}$$

Once this is done, we can proceed with our calculations in perturbation theory. Notice that for all practical purposes of these lectures we will work at LO so that $\delta_Z = \delta_m = \delta_\lambda = 0$. From now on, let us assume that we have a good splitting of our theory so that:

$$H = H_0 + H_I \tag{10.109}$$

[4]It is not necessary to add higher order terms as they arise in a renormalisation procedure.

10.4.2 Dyson Series Representation of the S-Matrix

As we saw, the S-Matrix operator is related to the limit of an operator $U(t,t')$ involving both the free and interaction Hamiltonian:

$$U(t,t') = e^{iH_0(t-t_0)}e^{-iH(t-t')}e^{-iH_0(t'-t_0)} \tag{10.110}$$

This expression allows us to easily prove important properties of the evolution operator:

$$U(t_1,t)U(t,t_2) = U(t_1,t_2) \tag{10.111}$$
$$U^\dagger(t_1,t_2) = U^{-1}(t_1,t_2) = U(t_2,t_1) \tag{10.112}$$

We shall now find a perturbative solution to the latter. In order to do this, we can find a differential equation that $U(t,t')$ must satisfy and solve it with appropriate initial conditions. Being an evolution operator, we choose as initial conditions:

$$U(t',t') = 1 \tag{10.113}$$

Now, a good guess would be to take its time derivative and see if we can recognise some solvable equation. Thus:

$$\begin{aligned} i\frac{\partial}{\partial t}U(t,t') &= e^{iH_0(t-t_0)}(H-H_0)e^{-iH(t-t')}e^{-iH_0(t'-t_0)} \\ &= \underbrace{e^{iH_0(t-t_0)}H_I e^{-iH_0(t-t_0)}}_{H_I(t)} e^{iH_0(t-t_0)}e^{-iH(t-t')}e^{-iH_0(t'-t_0)} \\ &= H_I(t)U(t,t') \end{aligned} \tag{10.114}$$

where $H_I(t)$ is the interaction Hamiltonian in the interaction picture. More simply:

$$\boxed{i\frac{\partial}{\partial t}U(t,t') = H_I(t)U(t,t')} \tag{10.115}$$

This is nothing but the Schrödinger equation for the interaction picture propagator! If the interaction picture Hamiltonian was time independent we would simply have the usual:

$$U(t,t') = e^{-iH_I\Delta t} \tag{10.116}$$

However the situation is here more delicate. We will solve the problem pertubatively in $H_I(t)$ by writing:

$$U(t,t') = \sum_n U^{(n)}(t,t') \tag{10.117}$$

and to solve for $U^{(n)}(t,t')$ we iterate n times the resolution of the differential equation starting from the zeroth order. For the latter:

$$i\frac{\partial}{\partial t}U^{(0)}(t,t') = 0 \tag{10.118}$$

Integrating this equation from t' to t and recalling the chosen initial condition, we have:

$$\int_{t'}^{t} d\tau \frac{\partial}{\partial \tau} U^{(0)}(\tau, t') = U^{(0)}(t, t') - U^{(0)}(t', t') = 0 \Rightarrow U^{(0)}(t, t') = 1 \tag{10.119}$$

At first order, we have:

$$i\frac{\partial}{\partial t} U^{(1)}(t, t') = H_I(t) U^{(0)}(t, t') = H_I(t) \tag{10.120}$$

Since we already used our initial condition, we set:

$$U^{(n)}(t', t') = 0, \quad \forall\ n > 0 \tag{10.121}$$

and therefore:

$$U^{(1)}(t, t') = -i \int_{t'}^{t} d\tau H_I(\tau) \tag{10.122}$$

The first non-trivial step comes at second order:

$$U^{(2)}(t, t') = -i \int_{t'}^{t} dt_1 \left(H_I(t_1)(-i) \int_{t'}^{t_1} dt_2 H_I(t_2) \right) = (-i)^2 \int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2 H_I(t_1) H_I(t_2) \tag{10.123}$$

We would like to find an equivalent but simpler expression for this integral. In order to do so, notice that this double integral, when represented on the $t_1 - t_2$ plane, covers only the lower-half of the square of size $(t-t') \times (t-t')$ cut by its diagonal. In other words, the region in the square that satisfies $t_2 < t_1$. Mathematically, this can be represented by a θ-function:

$$(-i)^2 \int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2 H_I(t_1) H_I(t_2) \equiv (-i)^2 \int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2\ \theta(t_1 - t_2) H_I(t_1) H_I(t_2) \tag{10.124}$$

Written with that θ function, the integral over t_2 can be extended to any time beyond t_1 as the value of the integrand is zero for times beyond the latter. In particular we can extend it up to t:

$$(-i)^2 \int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2\ \theta(t_1 - t_2) H_I(t_1) H_I(t_2) \equiv (-i)^2 \int_{t'}^{t} dt_1 \int_{t'}^{t} dt_2\ \theta(t_1 - t_2) H_I(t_1) H_I(t_2) \tag{10.125}$$

The last trick consists in noticing that written in this way we have that the integrals $dt_1 dt_2$ are symmetric and hence only the symmetric part of the integrand will contribute:

$$\begin{aligned} (-i)^2 \int_{t'}^{t} dt_1 \int_{t'}^{t} dt_2\ \theta(t_1 - t_2) H_I(t_1) H_I(t_2) &\equiv \frac{(-i)^2}{2} \int_{t'}^{t} dt_1 dt_2\ \left(\theta(t_1 - t_2) H_I(t_1) H_I(t_2) + \theta(t_2 - t_1) H_I(t_2) H_I(t_1)\right) \\ &= \frac{(-i)^2}{2} \int_{t'}^{t} dt_1 dt_2\ T\left\{H_I(t_1) H_I(t_2)\right\} \end{aligned} \tag{10.126}$$

We finally arrived at a suitable expression for the second order value of the evolution operator:

$$U^{(2)}(t,t') = \frac{(-i)^2}{2}\int_{t'}^{t} dt_1 dt_2\; T\left\{H_I(t_1)H_I(t_2)\right\} \tag{10.127}$$

This result is now straightforwadly generalised for the nth order. Using the same trick, *i.e.* inserting $n-1$ θ-functions to extend all the integrals up to time t and fully symmetrising the integrand, we obtain:

$$U^{(n)}(t,t') = \frac{(-i)^n}{n!}\int_{t'}^{t} T\left\{\prod_{i=1}^{n} dt_i H_I(t_i)\right\} \tag{10.128}$$

In the end, the expression, in perturbation theory, for the evolution operator is:

$$\boxed{U(t,t') = 1 + \sum_{n=1}^{\infty}\frac{(-i)^n}{n!}\int_{t'}^{t} T\left\{\prod_{i=1}^{n} dt_i H_I(t_i)\right\} \equiv T e^{-i\int_{t'}^{t} dt' H_I(t')}} \tag{10.129}$$

where the last expression must be understood as the Taylor expansion of the exponential with the time-ordered product inside the multiple integrals and is just a notational short-hand. Finally, recalling the definition of the S-Matrix operator:

$$\boxed{S = \Omega^{\dagger}(+\infty)\Omega(-\infty) = U(+\infty,-\infty) = T e^{-i\int_{-\infty}^{\infty} dt H_I(t)}} \tag{10.130}$$

This is the so-called *Dyson series representation of the S-Matrix*. Notice that the S-Matrix operator is independent of the time origin t_0. Indeed, the latter can be changed by a simple shift in the integral inside the exponetial.

We would like to study the Lorentz-invariance of the S-Matrix using the expression we just derived. Tha Hamiltonian formalism is in this case not suitable as it requires a splitting of space and time. In order to do so, we should go back to a Lagrangian formulation. Recall the Legendre transform which brings us from the latter to the former:

$$\mathcal{H} = \dot{\phi}\frac{\partial\mathcal{L}}{\partial\dot{\phi}} - \mathcal{L} = \underbrace{\dot{\phi}\frac{\partial\mathcal{L}_0}{\partial\dot{\phi}} - \mathcal{L}_0}_{\mathcal{H}_0} + \underbrace{\dot{\phi}\frac{\partial\mathcal{L}_I}{\partial\dot{\phi}} - \mathcal{L}_I}_{\mathcal{H}_I} \tag{10.131}$$

We see that we simply have:

$$\mathcal{H}_I = -\mathcal{L}_I \quad \Leftrightarrow \quad \text{No derivative couplings} \tag{10.132}$$

This is the case in many theories but it is not always true! One simple example is the case of scalar electrodynamics with a local $U(1)$ symmetry:

$$\mathcal{L} = -\frac{1}{4}(F_{\mu\nu})^2 + |D_\mu\phi|^2 - m^2\phi\phi^*, \quad D_\mu\phi = \partial_\mu\phi + ieqA_\mu\phi \tag{10.133}$$

In most of the cases we are interested in, we have:

$$\boxed{S = Te^{i\int d^4x \mathcal{L}_I(x)}} \tag{10.134}$$

Now, all parts of this expression are manisfestly Lorentz invariant, except the time-ordering operator. However, the time-ordering of two spacetime points is Lorentz invariant unless their separation is spacelike. Indeed, if the points are spacelike separated, we can always go to a frame where their separation is of the form $(0, \Delta\vec{x})$, so that the notion of ordering is not Lorentz invariant. Nevertheless, we are working with *causal* theories, *i.e*:

$$[\mathcal{L}(x), \mathcal{L}(y)] = 0, \quad |x-y|^2 < 0 \tag{10.135}$$

and since the Lagrangians commute when the points are spacelike separated, the time ordering is unnecessary in that case. In short, our formalism is Lorentz invariant provided *causality*![5]

A further comment should be made. In our formalism, we are working with the S-Matrix *operator* that connects free states. We never make use of the S-Matrix whose elements are the scalar products between asymptotic states. In the same spirit, the fields entering $\mathcal{L}_I(x)$ are the free Heisenberg fields which evolve naturally with the free Hamiltonian and admit a plane-waves expansion.

10.4.3 Wick's Theorem

We have now the basic ingredients: an explicit power series expansion in $\mathcal{L}_I$ to compute the S-Matrix operator S as well as the interacting part of the Lagrangian itself $\mathcal{L}_I(x) \equiv \mathcal{L}_I[\phi(x)]$ with $\phi(x)$ the Heisenberg picture fields of the free theory and:

$$\mathcal{L}_I(x) = e^{iH_0 t}\mathcal{L}_I(t=0,\vec{x})e^{-iH_0 t} \equiv e^{iH_0 t}\mathcal{L}_I[\phi(t=0,\vec{x})]e^{-iH_0 t} = \mathcal{L}_I[\phi(x)] \tag{10.136}$$

We just need an efficient way to compute the T-product of Lagrangians in the Dyson series and its matrix elements. The key to this is to realize that correlators of free fields can always be decomposed as a sum and product of two point functions. This is Wick's theorem, which will be derived in this section. This is at the foundation of Feynman diagrams. We will develop this method step by step starting with some useful definitions. Let us start for simplicity with a single real scalar field, we will treat the fermionic case at the end of the section. We define:

$$\phi(x) = \int d\Omega_p \left(e^{-ipx}a_p + e^{ipx}a_p^\dagger\right) \equiv \phi^-(x) + \phi^+(x) \tag{10.137}$$

with $\phi^-(x)$, $\phi^+(x)$ the destruction and annihilation local fields respectively[6]. They satisfy the following commutation relations:

$$\left[\phi^-(x), \phi^+(y)\right] = \int d\Omega_p e^{-ip(x-y)} = D(x-y)\cdot\mathbb{1} \tag{10.138}$$

$$\left[\phi^\pm(x), \phi^\pm(y)\right] = 0 \tag{10.139}$$

where $\mathbb{1}$ is the identity operator. Any functional of the field ϕ can be written in terms of these two fields as:

$$\mathcal{O}[\phi] \equiv \mathcal{O}[\phi^+, \phi^-] = \sum_n \sum_m \int F(x_1,\dots,x_n;y_1,\dots,y_m)\prod_{i=1}^n d^4x_i\phi^+(x_i)\prod_{j=1}^m d^4y_j\phi^-(y_j) \tag{10.140}$$

i.e. with all creation fields to the left of the annihilation fields. In particular, the S-Matrix operator can also be written in this form. This form is very useful when the functional is evaluated on the vacuum state. Indeed by the properties of raising and lowering operators on the vacuum state, $\langle 0|\mathcal{O}[\phi]|0\rangle$ only non-vanishing contribution comes from the term in the sum where $m = n = 0$. We can generalize this to a product of several bosonic operators, leading to the following definiton:

[5] For a more rigorous proof of the Lorentz invariance of the S-Matrix, read *The Quantum Theory of Fields, Vol I* (Weinberg) Section 3.5

[6] We follow here a different convention than Peskin and Schroeder.

Definition. *The **normal ordered** product of bosonic operators $\mathcal{O}_1, \dots, \mathcal{O}_n$ is denoted as:*

$$N\{\mathcal{O}_1 \cdots \mathcal{O}_n\} \equiv : \mathcal{O}_1 \cdots \mathcal{O}_n : \tag{10.141}$$

and is defined as the product in which all creation operators are to the left of all destruction operators.

Examples. *1. Identity operator: normal ordering is trivial*

$$: \mathbb{1} := \mathbb{1} \tag{10.142}$$

2. One scalar field:

$$: \phi :=: \phi^- + \phi^+ := \phi^- + \phi^+ = \phi \tag{10.143}$$

3. Two scalar fields:

$$: \phi_1(x)\phi_2(y) := \phi_1^+(x)\phi_2^+(y) + \phi_1^+(x)\phi_2^-(y) + \phi_2^+(y)\phi_1^-(x) + \phi_1^-(x)\phi_2^-(y) \tag{10.144}$$

The normal ordered product is related to the ordinary one according to

$$\begin{aligned} \phi_1(x)\phi_2(y) &= (\phi_1^-(x) + \phi_1^+(x))(\phi_2^-(y) + \phi_2^+(y)) \\ &= \phi_1^+(x)\phi_2^+(y) + \phi_1^+(x)\phi_2^-(y) + \phi_2^+(y)\phi_1^-(x) + \phi_1^-(x)\phi_2^-(y) + \left[\phi_1^-(x), \phi_2^+(y)\right] \end{aligned} . \tag{10.145}$$

When ϕ_1 and ϕ_2 are commuting fields the commutator vanishes and the two notions of product coincide. For identical fields, $\phi_1 = \phi_2 = \phi$ we have instead

$$\phi(x)\phi(y) =: \phi(x)\phi(y) : + D(x-y) \cdot \mathbb{1} . \tag{10.146}$$

Let us give some properties of the normal ordered product.

Properties. *Let us consider a set of bosonic operators $\{\mathcal{O}_n \equiv \mathcal{O}_n[\phi_n]\}$ and of c-numbers $a_n \in \mathbb{C}$.*

1. The normal ordered product is linear:

$$: \sum_n a_n \mathcal{O}_n := \sum_n a_n : \mathcal{O}_n : \tag{10.147}$$

2. Since the creation fields commute among them and so do the annihilation fields, we have for any permutation π:

$$: \phi_1(x_1) \cdots \phi_n(x_n) :=: \phi_{\pi(1)}(x_{\pi(1)}) \cdots \phi_{\pi(n)}(x_{\pi(n)}) : \tag{10.148}$$

so that:

$$: \mathcal{O}_1 \cdots \mathcal{O}_n :=: \mathcal{O}_{\pi(1)} \cdots \mathcal{O}_{\pi(n)} : \tag{10.149}$$

3. The vacuum expectation value (VEV) of a normal ordered product vanishes:

$$\langle 0| : \mathcal{O}_1 \cdots \mathcal{O}_n : |0\rangle = 0 \tag{10.150}$$

Generalizing equation 10.146, we can rewrite ordinary (or time-ordered) products of operators as a sum of normal-ordered products and two-point functions. Then using property 10.150, we can compute the VEV, in which all normal ordered products will vanish.

For that purpose, we define contractions, the operation that take two operators and give a two point function, which is simply c-number. In our case, there is two useful ways to do that.

Definition. *The **contraction** of two operators $\mathcal{O}_1$ and $\mathcal{O}_2$ is:*

$$\mathcal{O}_1\underbrace{\mathcal{O}_2} \equiv \langle 0|\mathcal{O}_1\mathcal{O}_2|0\rangle \tag{10.151}$$

*Similarly, the **chronological contraction** is defined as:*

$$\overbrace{\mathcal{O}_1\mathcal{O}_2} \equiv \langle 0|T\mathcal{O}_1\mathcal{O}_2|0\rangle \tag{10.152}$$

Examples. *In the case of free fields we have:*

$$\phi_1(x)\phi_2(y) =: \phi_1(x)\phi_2(y) : + D(x-y)\cdot \mathbb{1} \tag{10.153}$$

where we drop from now on the indices on the propagator as it is clear from the context if it is zero or non-zero. This implies:

1. *Their contraction is given by:*

$$\phi_1\underbrace{(x)\phi_2}(y) \equiv D(x-y) \tag{10.154}$$

2. *Their chronological contraction is given by:*

$$\phi_1\overbrace{(x)\phi_2}(y) \equiv D_F(x-y) \tag{10.155}$$

where $D_F(x-y)$ is the Feynman propagator and is given by:

$$D_F(x-y) = \theta(x^0-y^0)D(x-y) + \theta(y^0-x^0)D(y-x) \tag{10.156}$$

Finally, let us give a property of the normal ordering of contracted operators:

Properties. *Let us consider a set of bosonic operators $\{\mathcal{O}_n \equiv \mathcal{O}_n[\phi_n]\}$, then:*

$$\begin{aligned} :\mathcal{O}_1\cdots\mathcal{O}_{i-1}\underbrace{\mathcal{O}_i\mathcal{O}_{i+1}\cdots\mathcal{O}_{j-1}}\mathcal{O}_j\mathcal{O}_{j+1}\cdots\mathcal{O}_n : &= : \overbrace{\mathcal{O}_i\mathcal{O}_j}^{c\text{-}number}\ \mathcal{O}_1\cdots\mathcal{O}_{i-1}\mathcal{O}_{i+1}\cdots\mathcal{O}_{j-1}\mathcal{O}_{j+1}\cdots\mathcal{O}_n : \\ &\equiv \mathcal{O}_i\mathcal{O}_j : \mathcal{O}_1\cdots\mathcal{O}_{i-1}\mathcal{O}_{i+1}\cdots\mathcal{O}_{j-1}\mathcal{O}_{j+1}\cdots\mathcal{O}_n : \end{aligned} \tag{10.157}$$

Theorem. *We have now all the tools to state **Wick's theorem**.*

Let us consider a set of free fields $\{\phi_n \equiv \phi_n(x_n)\}$. Then we have the following three results:

1. *The time-ordering of n fields can be written as:*

$$T\phi_1\cdots\phi_n =: \phi_1\cdots\phi_n : + : \{\textit{all possible chronological contractions}\} : \tag{10.158}$$

2. *The product of n fields can be written as:*

$$\phi_1\cdots\phi_n =: \phi_1\cdots\phi_n : + : \{\textit{all possible contractions}\} : \tag{10.159}$$

3. *The product of two normal-ordered products of $n = p+q$ fields can be written as:*

$$: \phi_1\cdots\phi_p :: \phi_{p+1}\cdots\phi_n :=: \phi_1\cdots\phi_n : + : \{\textit{all possible contractions}\} : \tag{10.160}$$

where (multiple-)contractions arise only between pairs of fields with one field from the first p and the other from the last q. Contractions between pairs of fields exclusively in the first p or last q do not arise.

A proof by induction of the theorem is given in Appendix A.

We will now use this theorem to find a convenient expression for the S-Matrix elements. They are of the form:

$$\langle 0|a_{p_1}\cdots a_{p_i}T\left\{\phi_1\cdots\phi_n\right\}a^\dagger_{k_1}\cdots a^\dagger_{k_j}|0\rangle \tag{10.161}$$

We consider the case for which $p_i \neq k_j\ \forall i,j$. Using 10.158, we obtain:

$$\begin{aligned}\langle 0|a_{p_1}\cdots a_{p_i}T\left\{\phi_1\cdots\phi_n\right\}a^\dagger_{k_1}\cdots a^\dagger_{k_j}|0\rangle &= \langle 0|a_{p_1}\cdots a_{p_i}:\phi_1\cdots\phi_n:a^\dagger_{k_1}\cdots a^\dagger_{k_j}|0\rangle\\ &+\langle 0|a_{p_1}\cdots a_{p_i}:\{\text{all possible chronological contractions}\}:a^\dagger_{k_1}\cdots a^\dagger_{k_j}|0\rangle\\ &=\langle 0|:a_{p_1}\cdots a_{p_i}::\phi_1\cdots\phi_n::a^\dagger_{k_1}\cdots a^\dagger_{k_j}:|0\rangle\\ &+\langle 0|:a_{p_1}\cdots a_{p_i}::\{\text{all possible chronological contractions}\}::a^\dagger_{k_1}\cdots a^\dagger_{k_j}:|0\rangle\end{aligned} \tag{10.162}$$

where in the last step, we noticed that the ladder operators are already normal ordered and thus using 10.160, we obtain our final result:

$$\boxed{\langle 0|a_{p_1}\cdots a_{p_i}T\left\{\phi_1\cdots\phi_n\right\}a^\dagger_{k_1}\cdots a^\dagger_{k_j}|0\rangle = \sum_{\substack{\text{all (chrono-)}\\ \text{contractions}}}\langle 0|:a_{p_1}\cdots\underbrace{a_{p_i}\phi_1}\cdots\overbrace{\phi_n a^\dagger_{k_1}}\cdots a^\dagger_{k_j}:|0\rangle} \tag{10.163}$$

where only terms in which everything is contracted contribute since the VEV of normal ordered products vanishes. The following contractions appear in the sum:

$$\underbrace{a_p\phi}(x) = \langle 0|a_p\phi(x)|0\rangle = \langle p|\phi(x)|0\rangle = e^{ipx} \tag{10.164}$$

$$\underbrace{\phi(x)a^\dagger_p} = \langle 0|\phi(x)a^\dagger_p|0\rangle = \langle 0|\phi(x)|p\rangle = e^{-ipx} \tag{10.165}$$

$$\overbrace{\phi(x)\phi(y)} = \langle 0|T\phi(x)\phi(y)|0\rangle = D_F(x-y) \tag{10.166}$$

Finally, let's comment on the generalization to fermions. Wick's theorem for fermions can be stated identically to the bosonic case on the condition that we adapt the definition of normal ordering to be self-consistent for fermions. Very similarly to bosons, we can separate destruction and creation operators in the definition of the fermionic fields:

$$\psi = \psi^+ + \psi^-, \quad \psi^\dagger = (\psi^\dagger)^+ + (\psi^\dagger)^- \tag{10.167}$$

where we have the anticommutators

$$\left\{\psi^-(x),(\psi^\dagger)^+(y)\right\} = \int d\Omega_p(\not{p}+m)e^{-ip(x-y)} = S(x-y)\cdot\mathbb{1} \tag{10.168}$$

$$\left\{\psi^\pm(x),(\psi^\dagger)^\pm(y)\right\} = \left\{\psi^\pm(x),\psi^\pm(y)\right\} = 0 \tag{10.169}$$

However, as for the time ordering, one needs to take into account the anticommuting properties of the fermionic variable by adding an additional minus sign depending on the ordering of the fields.

Definition. *The **normal ordered** product of fermionic operators $\mathcal{O}_1,\ldots,\mathcal{O}_n$ is denoted as $:\mathcal{O}_1\cdots\mathcal{O}_n:$ and is defined as the product in which all creation operators are to the left of all destruction operators to the right, times the sign $(-1)^{n_\pi}$ of the permutation π necessary to achieve the normal ordering.*

Examples. *1. One fermion:*

$$:\psi:=:\psi^-+\psi^+:=\psi^-+\psi^+=\psi \tag{10.170}$$

2. Two fermions:

$$:\psi_1(x)\psi_2(y):=\psi_1^+(x)\psi_2^+(y)+\psi_1^+(x)\psi_2^-(y)-\psi_2^+(y)\psi_1^-(x)+\psi_1^-(x)\psi_2^-(y) \tag{10.171}$$

Where one should notice the minus sign in the third term from exchanging ψ_1^- and ψ_2^+.

3. *n fermions:*

$$:\psi_1...\psi_n:=\ \psi_1^+...\psi_n^+ + \psi_1^+...\psi_{n-1}^+\psi_n^- - \psi_1^+...\psi_{n-2}^+\psi_n^+\psi_{n-1}^- + \psi_1^+...\psi_{n-1}^+\psi_n^+\psi_{n-2}^- - ... \tag{10.172}$$

Properties. *The normal product for fermionic operators respects similar properties to the bosonic case*

1.

$$\psi_1(x)\psi_2(y) =: \psi_1(x)\psi_2(y) : +\{\psi_1^-(x),\psi_2^+(y)\} =: \psi_1(x)\psi_2(y) : +\langle 0|\psi_1^-(x)\psi_2^+(y)|0\rangle\cdot\mathbb{1} \tag{10.173}$$

2. *Since the creation fields anti-commute among themself and so do the annihilation fields, we have for any permutation π:*

$$:\psi_1(x_1)\cdots\psi_n(x_n):=(-1)^{n_\pi}\ :\psi_{\pi(1)}(x_{\pi(1)})\cdots\psi_{\pi(n)}(x_{\pi(n)}):\tag{10.174}$$

3. *Contractions are defined the same way for bosonic operators and fermionic operators. However, because of the anti-commuting properties of fermions, the rule for contractions for a set of bosonic and fermionic operators $\{\mathcal{O}_n\equiv\mathcal{O}_n[\psi_n]\}$ inside a normal ordered product becomes:*

$$:\mathcal{O}_1\cdots\mathcal{O}_{i-1}\underbrace{\mathcal{O}_i\mathcal{O}_{i+1}\cdots\mathcal{O}_{j-1}}\mathcal{O}_j\mathcal{O}_{j+1}\cdots\mathcal{O}_n:\ =(-1)^{n_{\pi_{ff}}}\underbrace{\mathcal{O}_i\mathcal{O}_j}:\mathcal{O}_1\cdots\mathcal{O}_{i-1}\mathcal{O}_{i+1}\cdots\mathcal{O}_{j-1}\mathcal{O}_{j+1}\cdots\mathcal{O}_n:\tag{10.175}$$

Where $n_{\pi_{ff}}$ is the order of the fermionic operators permutation needed to move $\mathcal{O}_i$ and $\mathcal{O}_j$ from their initial to their final position.

Examples. *We take a set of bosonic ϕ_i and fermionic ψ_i operators.*

$$:\psi_1\underbrace{\psi_2\psi_3}\psi_4:=(-1)^3\underbrace{\psi_2\psi_4}:\psi_1\psi_3:\tag{10.176}$$

$$:\phi_1\underbrace{\psi_2\phi_3}\psi_4:=(-1)^0\underbrace{\psi_2\psi_4}:\phi_1\phi_3:\tag{10.177}$$

$$:\psi_1\underbrace{\psi_2\phi_3}\psi_4:=(-1)^2\underbrace{\psi_2\psi_4}:\psi_1\phi_3:\tag{10.178}$$

4. *Chronological contractions work in the same way:*

$$:\mathcal{O}_1\cdots\mathcal{O}_{i-1}\overbrace{\mathcal{O}_i\mathcal{O}_{i+1}\cdots\mathcal{O}_{j-1}}\mathcal{O}_j\mathcal{O}_{j+1}\cdots\mathcal{O}_n:\ =(-1)^{n_{\pi_{ff}}}\overbrace{\mathcal{O}_i\mathcal{O}_j}:\mathcal{O}_1\cdots\mathcal{O}_{i-1}\mathcal{O}_{i+1}\cdots\mathcal{O}_{j-1}\mathcal{O}_{j+1}\cdots\mathcal{O}_n:\tag{10.179}$$

Using these definitions for normal ordering, contractions, and chronological contractions of fermionic operators, Wick's theorem becomes valid for both bosonic and fermionic operators.

10.4.4 Feynman Rules and Feynman Diagrams

At this point we have defined all ingredients to introduce the Feynman rules. Let us sketch the basic idea:

- The S-Matrix operator is a series of time-ordered products of free fields.
- The *in* and *out* states correspond to applications of creation and annihiliation operators to the free vacuum.
- By Wick's theorem, we can write the time-ordered products in terms of normal ordered products which are already in a suitable form to take matrix elements with *in* and *out* states.
- The set of Wick contractions within normal ordered matrix elements is effectively described by a set of diagrammatic rules: *Feynman diagrams.*

Feynman diagrams are built out of three type of elements:

1. External lines from $a_{k_1}^\dagger\cdots a_{k_i}^\dagger|0\rangle$ and $\langle 0|a_{p_1}\cdots a_{p_j}$.

2. Internal lines or propagators: $\overbrace{\phi_i \phi_j}$.
3. Interaction vertices: interaction Lagrangian.

In order to illustrate the derivation of the Feynman rules as well as the computation of Feynman diagrams, let us start with a simple example, the $\lambda\phi^3$ Lagrangian:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{m^2}{2}\phi^2 - \frac{\lambda}{3!}\phi^3 \tag{10.180}$$

The S-matrix operator admits an perturbative expansion (Dyson series):

$$S = \sum_n S^{(n)} \tag{10.181}$$

$$S^{(n)} = \frac{1}{n!}\int d^4x_1 \cdots d^4x_n \left(-i\frac{\lambda}{3!}\right)^n T\left\{\phi(x_1)^3 \cdots \phi(x_n)^3\right\} \tag{10.182}$$

The most general contribution to the matrix element:

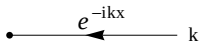
$$iT \equiv \langle p_1 \cdots p_j | S | k_1 \cdots k_i \rangle \tag{10.183}$$

has the form:

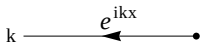
$$\frac{1}{n!}\int d^4x_1 \cdots d^4x_n \left(-i\frac{\lambda}{3!}\right)^n \langle 0 | a_{p_1} \cdots a_{p_j} T\left\{\phi(x_1)^3 \cdots \phi(x_n)^3\right\} a^\dagger_{k_1} \cdots a^\dagger_{k_i} | 0 \rangle \tag{10.184}$$

and therefore we will have three types of contractions to which are added the interaction vertices:

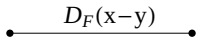
- Incoming lines: $\phi \underbrace{(x) a^\dagger_k}$



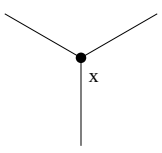
- Outgoing lines: $a_k \underbrace{\phi(x)}$



- Propagators: $\phi(x)\phi(y)$



- Interaction vertices: $-i\frac{\lambda}{3!}\int d^4x$



Each possible Wick contraction corresponds to a diagram where the n vertices are connected among each other by internal lines and to *in* and *out* states (external lines). The amplitude is thus the *sum* over all possible *topologies* (each topology being weighted by its own combinatorical factor). Notice that physically, summing over all topologies is equivalent to summing over all possible histories of the incoming particles before being scattered into the outgoing particles. We see here the close link to the path integral formulation of quantum mechanics. The rules we just derived are the *Feynman rules in position space*. We will derive shortly similar rules in momentum space, but before doing so, let us warm ourselves up with a simple example.

Example. *Let us compute the amplitude associated to the following Feynman diagram in $\lambda\phi^3$ theory:*

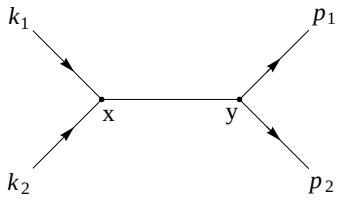


Figure 10.12: A Feynman diagram contributing to the 2 to 2 scattering in $\lambda\phi^3$.

From left to right, we have:

- *Two external lines with momenta k_1 and k_2 arriving at the x vertex giving a factor $e^{-ix(k_1+k_2)}$*
- *The x vertex itself. Notice that the vertices have a factor $1/3!$, but there are exactly $3!$ different ways of contracting three lines and a vertex, hence exactly cancelling the factorial[7]. Therefore, a contracted vertex effectively corresponds to the Feynman rule $-i\lambda\int d^4x$.*
- *The propagator (internal line) $D_F(x-y)$.*
- *The second vertex and the last two external lines give a similar expression to previous one.*
- *Finally, the diagram comes at second order in λ, there is therefore an extra factor $1/2$. However, we see that it is symmetric in x and y. Indeed, we could have chosen to contract the incoming lines with the y vertex and the outgoing lines with the x vertex giving rise to the same diagram. It is a general feature that the $1/n!$ of the Dyson series exactly cancels the symmetry which exchanges vertices, we can therefore drop these two factors of the computation.*

Putting everything together, we obtain:

$$(-i\lambda)^2\int d^4x d^4y D_F(x-y)e^{-ix(k_1+k_2)+iy(p_1+p_2)} \tag{10.185}$$

It is however often simpler to compute the amplitude and to express the Feynman rules in momentum space. This is done simply by performing the position integrals. For example a generic vertex with three incoming momenta corresponds to the integration:

$$\int d^4x e^{-ix(p_1+p_2+p_3)} = (2\pi)^4\delta^{(4)}(p_1+p_2+p_3) \tag{10.186}$$

Therefore, momentum flow is conserved at each vertex. Electric circuits analogy give a useful intuitive picture of the set of constraints impsoed by the delta-functions (Kirchhof's laws). The remaining rules are as follows:

- External lines: The exponential in momentum space becomes the identity operator. For a scalar field:

$$\text{External lines} \;=\; 1 \tag{10.187}$$

[7] In the Lagrangian, a ϕ^n interaction is generally normalised with a $1/n!$ combinatorical factor for this reason.

- Propagators: From the Lorentz invariant expression of D_F we have:

$$D_F(x) = \int \frac{d^4p}{(2\pi)^4} \frac{ie^{-ipx}}{p^2 - m^2} \tag{10.188}$$

The exponential was previously included in the vertex integration leaving us with a factor:

$$\int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2} \tag{10.189}$$

We therefore need to integrate over all internal lines. Let us consider V vertice, I internal lines and $I-V+1 = L$ loops. Thus, schematically:

$$\int \frac{d^4p_1}{(2\pi)^4} \cdots \frac{d^4p_I}{(2\pi)^4} \underbrace{\delta^{(4)} \cdots \delta^{(4)}}_{V \text{ vertices}} = (2\pi)^4 \delta^{(4)}(\sum_{in} k - \sum_{out} p) \int \frac{d^4p_1}{(2\pi)^4} \cdots \frac{d^4p_L}{(2\pi)^4} \tag{10.190}$$

So that we are left with an integration over the set of remaining free internal momenta that run in the loops.

- Interaction vertices: After the integration is done, we are left with the coupling constant and the combinatorical factor. As we previously argued, the number of topologically equivalent diagrams exactly cancels the $1/3!$ so that for each vertex we simply associate a factor of $(-i\lambda)$.

We can synthesise the result in a table of Feynman rules. We also add the position space rules for completeness. The algorithm is as follows:

1. Draw all possible diagrams contributing to the process you consider.
2. Choose whether to work in momentum or position space.
3. For each diagram, using the following Feynman rules, compute its amplitude:

	Position space	Momentum space
Ext. lines	$e^{-ix(\pm p)}$	1
Int. lines	$D_F(x-y)$	$D_F(p)$
Vertices	$-i\lambda \int d^4x$	$-i\lambda$

(10.191)

In addition, in momentum space, apply momentum conservation at each vertex, integrate over the momenta running in the loops and factor out an overall delta-function $(2\pi)^4\delta^{(4)}(\sum_{in} k - \sum_{out} p)$.

4. Most of combinatorics factors have been treated with this algorithm, but there is an extra symmetry factor that can arise when we have closed loops inside the diagram. Compute it and divide the amplitude by this number[8].

What remains, is the contribution to the *reduced* amplitude $i\mathcal{M}$:

$$S(\{k\} \to \{p\}) = \mathbb{1} + (2\pi)^4 \delta^{(4)}(\sum_{in} k - \sum_{out} p) i\mathcal{M}(\{k\} \to \{p\}) \tag{10.192}$$

[8]This symmetry factor arises because, for example, in a closed loop, it does not matter how we connect internal lines inside it. For further details, see Peskin and Schroeder Section 4.4.

Application:
Consider the following process in $\lambda\phi^3$:

$$k_1, k_2 \longrightarrow p_1, p_2 \tag{10.193}$$

This is an *elastic scattering* process. Let us first compute explicitly all contractions and after that see how the algorithm we derived above leads us to the same result in a much faster way. We first bother doing it the brave way because it is important to understand very well all steps of the full computation. The algorithm can then be applied efficiently and, in the case of complicated diagrams, no symmetry factor will be left behind!

The *in* and *out* states are simply:

$$|in\rangle = a^\dagger_{k_1} a^\dagger_{k_2} |0\rangle \tag{10.194}$$
$$\langle out| = \langle 0| a_{p_1} a_{p_2} \tag{10.195}$$

The S-Matrix operator is given, up to second order, by:

$$S - 1 = iT = -i\frac{\lambda}{3!}\int d^4x T\phi^3(x) + \frac{1}{2}\left(-i\frac{\lambda}{3!}\right)^2 \int d^4x d^4y T\phi^3(x)\phi^3(y) + \mathcal{O}(\lambda^3) \tag{10.196}$$

By Wick's theorem, only the fully contracted products contribute to the process. Therefore, at first order, there is no contribution (we cannot fully contract an odd number of operators!)[9]. The first contribution comes at second order. What are all possible contractions and what are their multiplicity? The matrix element we consider is:

$$\frac{1}{2}\left(-i\frac{\lambda}{3!}\right)^2 \int d^4x d^4y \sum_{\substack{\text{all (chrono-)}\\ \text{contractions}}} \langle 0| : \underbrace{a_{p_1} a_{p_2} \phi^3}(x) \overbrace{\phi^3}(y) \underbrace{a^\dagger_{k_1}} a^\dagger_{k_2} : |0\rangle \tag{10.197}$$

As there are four ladder operators and six fields, there must be one chronological contraction. Let us first consider the terms involving a chronological contraction of the form $\phi(x)\phi(x) = D_F(0) = \phi(y)\phi(y)$. Such a term would be for example:

$$\sim \int d^4x d^4y D_F(x-x) e^{-ixk_1 - iy(k_2 - p_1 - p_2)} \sim \delta^{(4)}(k_1)\delta^{(4)}(k_2 - p_1 - p_2) = 0 \tag{10.198}$$

This vanishes as by assumption $k_1 \neq 0$. Thus, only terms with the chronological contraction $\phi(x)\phi(y) = D_F(x-y)$ contribute.

Let us start with the internal line. How many way of contracting one $\phi(x)$ with one $\phi(y)$ do we have? We first must choose one $\phi(x)$: three possible choices. Then must contract it with one $\phi(y)$: three choices as well. The multiplicity for the contraction of the internal line is then $3 \times 3 = 9$. We are left with:

$$9 \times \frac{1}{2}\left(-i\frac{\lambda}{3!}\right)^2 \int d^4x d^4y D_F(x-y) \sum_{\substack{\text{all}\\ \text{contractions}}} \langle 0| : \underbrace{a_{p_1} a_{p_2} \phi}(x)\phi(x)\phi\underbrace{(y)\phi(y) a^\dagger_{k_1}} a^\dagger_{k_2} : |0\rangle \tag{10.199}$$

[9] It is a general feature that for a theory with a ϕ^n interaction, if n is even all orders of the Dyson series contribute to any elastic scattering and if n is odd only odd orders. What about a $2 \to 3$ inelastic scattering? $1 \to 3$ decay?

Let us now notice that since $D_F(x-y) = D_F(y-x)$ (as it only depends on the causal distance) the full expression is $x \leftrightarrow y$ symmetric. Therefore, in the sum, we can remove all contractions which differ by an exchange $x \leftrightarrow y$ together with the factor $1/2$ of the Dyson series. We previously explained that the Taylor coefficient $1/n!$ at nth order canceled exactly the combinatorical factor arising from the exchange of the n vertices. We see here an explicit example. Thus:

$$9 \times \left(-i\frac{\lambda}{3!}\right)^2 \int d^4x d^4y D_F(x-y) \sum_{\substack{\text{all} \\ \text{contractions} \\ x \not\leftrightarrow y}} \langle 0| : a_{\underbrace{p_1} a_{\underbrace{p_2}} \phi(\underbrace{x})\phi(x)\phi(\underbrace{y})\phi(y) a^\dagger_{\underbrace{k_1}} a^\dagger_{k_2} : |0\rangle \tag{10.200}$$

where, by $x \not\leftrightarrow y$, we mean that we do not count contraction that are equivalent under the exchange of the vertices. What are the possible contractions left?

1. Contract a_{p_1} and a_{p_2} with $\phi(x)$ and $a^\dagger_{k_1}$ and $a^\dagger_{k_2}$ with $\phi(y)$. For the first contraction, we have two possibilities among $\phi(x)$s, then the second annihilation operator is constrained to the remaining field. Similarly, for the first creation operator with start with two possibilities and the remaining one is constrained. This gives rise to a multiplicity $2 \times 2 = 4$ for these contractions. We choose one and we compute:

$$9 \times 4 \times \left(-i\frac{\lambda}{3!}\right)^2 \int d^4x d^4y D_F(x-y) e^{ix(p_1+p_2)-iy(k_1+k_2)} \tag{10.201}$$

   Here, the contraction of the annihilition operators with the $\phi(y)$s and of the creation operators with the $\phi(x)$s falls in the class of equivalent diagrams under the exchange $x \leftrightarrow y$ and thus have already been accounted for (notice how the integral is invariant under the echange of the two variables).

2. Contract a_{p_1} and $a^\dagger_{k_1}$ with $\phi(x)$ and a_{p_2} and $a^\dagger_{k_2}$ with $\phi(y)$. Again the multiplicity is $2 \times 2 = 4$. The result is:

$$9 \times 4 \times \left(-i\frac{\lambda}{3!}\right)^2 \int d^4x d^4y D_F(x-y) e^{ix(p_1-k_1)-iy(k_2-p_2)} \tag{10.202}$$

3. The last possible contraction is a_{p_1} and $a^\dagger_{k_2}$ with $\phi(x)$ and a_{p_2} and $a^\dagger_{k_1}$ with $\phi(y)$. Again the multiplicity is $2 \times 2 = 4$. The result is:

$$9 \times 4 \times \left(-i\frac{\lambda}{3!}\right)^2 \int d^4x d^4y D_F(x-y) e^{ix(p_1-k_2)-iy(k_1-p_2)} \tag{10.203}$$

We see that in each of the three different contractions a factor $9 \times 4 = (3!)^2$ arises. This cancels exactly the normalisation of the interaction Lagrangian and corresponds to the fact that a vertex has 3! possible ways to be contracted with three lines.

The sum of all possible contractions gives:

$$(-i\lambda)^2 \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2-m^2} \int d^4x d^4y e^{-ip(x-y)} \left(e^{ix(p_1+p_2)-iy(k_1+k_2)} + e^{ix(p_1-k_1)-iy(k_2-p_2)} + e^{ix(p_1-k_2)-iy(k_1-p_2)}\right) \tag{10.204}$$

where we rewrote explicitly the propagator in terms of the four-dimensional momentum integral. Let us compute the first term:

$$\begin{aligned}\int \frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2}\int d^4x d^4y e^{-ip(x-y)}e^{ix(p_1+p_2)-iy(k_1+k_2)} &= \int \frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2}\int d^4x e^{ix(p_1+p_2-p)}\int d^4y e^{-iy(k_1+k_2-p)} \\ &= \int \frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2}(2\pi)^4\delta^{(4)}(p_1+p_2-p)(2\pi)^4\delta^{(4)}(k_1+k_2-p) \\ &= \frac{i}{(k_1+k_2)^2-m^2}(2\pi)^4\delta^{(4)}(p_1+p_2-k_1-k_2)\end{aligned} \tag{10.205}$$

and similarly for the other terms. Finally our result is:

$$iT = (2\pi)^4\delta^{(4)}(p_1+p_2-k_1-k_2)(-i\lambda)^2\left(\frac{i}{(k_1+k_2)^2-m^2}+\frac{i}{(k_1-p_1)^2-m^2}+\frac{i}{(k_1-p_2)^2-m^2}\right) \tag{10.206}$$

so that:

$$\boxed{i\mathcal{M} = (-i\lambda)^2\left(\frac{i}{(k_1+k_2)^2-m^2}+\frac{i}{(k_1-p_1)^2-m^2}+\frac{i}{(k_1-p_2)^2-m^2}\right)} \tag{10.207}$$

The computation was long and tedious but we arrived to the desired result: we have an expression for the reduced amplitude for the elastic scattering of two particles at second order in $\lambda\phi^3$. Notice that we never talked nor used Feynman rules/diagrams as these are just tools to compute these expressions more effectively but are not *per se* necessary. Nevertheless, let us now see how a powerful tool they are!

How do we proceed? We have a well defined algorithm (set of Feynman rules) which we can follow (almost) blindly to compute any amplitude we wish.

1. *Draw all Feynman diagrams that contribute to the considered process at the desired order.* At first order, we must join a 3-vertex to 4 external lines: topologically impossible. At second order, we have three possibilities:

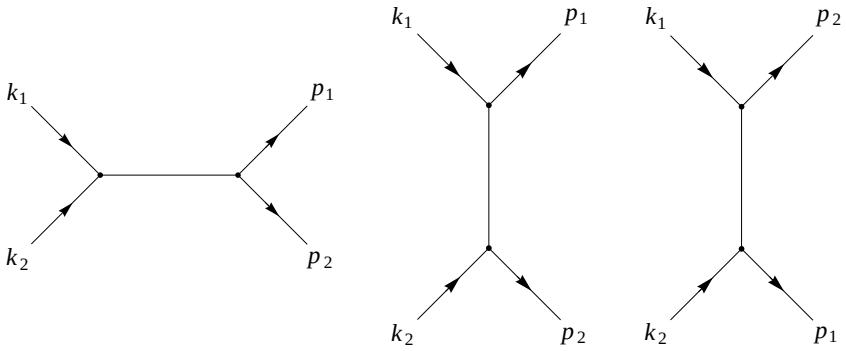


Figure 10.13: The three Feynman diagrams contributing to the 2 to 2 scattering in $\lambda\phi^3$.

2. *Choose whether to work in momentum or position space.* Notice that we have two vertices one internal line and no loops. Since there are no loops, the Feynman rules in momentum space are straightforward (no integration whatsoever). Hence, we choose the latter (even with loops the calculations are often simpler to carry out in momentum space).

3. *Use the Feynman rules.* Call p the momentum flowing in the internal line, then the three diagrams will be given by the expression:

$$(-i\lambda)^2\frac{i}{p^2-m^2} \tag{10.208}$$

where for each diagram we need to apply a different momentum conservation law at the vertices and we must sum the three diagrams. We must also add an overall four-dimensional delta-function and consider the symmetry factor. There is no particular symmetry factor for these diagrams (no internal loop or any other fancy topology) and therefore, we obtain:

$$\boxed{i\mathcal{M} = (-i\lambda)^2 \left(\frac{i}{(k_1+k_2)^2 - m^2} + \frac{i}{(k_1-p_1)^2 - m^2} + \frac{i}{(k_1-p_2)^2 - m^2} \right)} \tag{10.209}$$

In a few straightforward steps we could derive the same result as we did before struggling with contractions and combinatorics. We this simple example we saw how the set of Feynman rules is a powerful algorithm to compute scattering processes. To conclude this subsection, we give a few definitions and properties of the reduced amplitude.

We studied here the two-to-two elastic scattering of the same particle. Imagine that we considered the four particles as incoming, then they would be completely undistinguishable one from another. Since the matrix element is a function of the four incoming momenta:

$$i\mathcal{M} = f(k_1, k_2, k_3, k_4), \qquad k_1 + k_2 + k_3 + k_4 = 0 \tag{10.210}$$

it must be symmetric under the exchange of any two momenta. Moreover, it must be Lorentz invariant. Thus, we seek three independent Lorentz invariant quantities that can parametrise the amplitude, and under whose exchange the latter remains invariant. They are given by:

$$\boxed{\begin{aligned} s &= (k_1+k_2)^2 \\ t &= (k_1+k_3)^2 \\ u &= (k_1+k_4)^2 \end{aligned}} \tag{10.211}$$

These are the *Mandelstam variables*. In a two-to-two scattering, they also exist and are given by the same formulas, replacing $k_3 \to -p_1$ and $k_4 \to -p_2$. When $p_i^2 = k_i^2 = m^2$, we have:

$$\boxed{s + t + u = 4m^2} \tag{10.212}$$

As we saw, in a two-to-two eleastic scattering we have three Feynman diagrams. As a function of the Mandelstam variables, we have:

$$i\mathcal{M}(s,t,u) = (-i\lambda)^2 \left(\frac{i}{s-m^2} + \frac{i}{t-m^2} + \frac{i}{u-m^2} \right) \tag{10.213}$$

Each of the three diagrams corresponds to a different *channel* and we see how the amplitude is (as we deal with four identical particles) the sum of the three channels. Knowing this, it would have been enough to compute, say, the s-channel and add the t and u contributions simply by changing channel. Indeed the t-channel diagram's amplitude corresponds to the s-channel amplitude with s and t exchanged and u left invariant. Similarly we obtain the u-channel amplitude.

Let us compute the values of the three variables and see to what they correspond physically. Let us choose the center-of-mass (CM) frame of the incoming particles. Then, from the definition of the CM and the conservation of four-momentum:

$$\begin{aligned} k_1 &= (E, \vec{k}) \\ k_2 &= (E, -\vec{k}) \\ p_1 &= (E, \vec{p}) \\ p_2 &= (E, -\vec{p}), \qquad |\vec{k}| = |\vec{p}| \end{aligned} \tag{10.214}$$

Then:

$$\begin{aligned} s &= 4E^2 = E_{CM}^2 \\ t &= -\frac{1}{2}(s - 4m^2)(1 - \cos\theta) \\ u &= -\frac{1}{2}(s - 4m^2)(1 + \cos\theta) \end{aligned} \tag{10.215}$$

where E_{CM} is the energy of the CM and θ is the angle between $\vec{k}$ and $\vec{p}$. We see that s really measures the energy of the CM and is a "fixed" quantity, whereas t and u measure roughly the forward and backward transfer of momentum respectively and carry all the angular dependence of the process.

Now that we have all necessary ingredients, we can compute the cross-section for this process. Recalling Eq. 10.84, we have in terms of the Mandelstam variables:

$$d\sigma_{CM} = \frac{|\mathcal{M}(s,t,u)|^2}{2s\sqrt{s^2 - 4m^2}} d\Phi^{(2)} \tag{10.216}$$

When all the particles have the same mass, the two-body LIPS simplifies greatly and we obtain:

$$d\sigma_{CM} = \frac{|\mathcal{M}(s,t,u)|^2}{32\pi s} d(\cos\theta) \tag{10.217}$$

where we used the ϕ independence of the matrix element to integrate over the azimutal angle.

To summarise this subsection, we developped the tools necessary to compute scattering amplitudes and using a simple example we could carry out our first complete QFT computation!

Chapter 11

Quantum Electrodynamics

11.1 Feynman rules of QED

The Lagrangian of QED reads:

$$\mathcal{L}_{QED} = -\frac{1}{4}F_{\mu\nu}^2 + \bar{\psi}(i\not{D} - m)\psi, \quad \not{D} = \gamma^\mu(\partial_\mu + ieA_\mu) \tag{11.1}$$

It can be split into two parts:

$$\mathcal{L}_0 = -\frac{1}{4}F_{\mu\nu}^2 + \bar{\psi}(i\not{\partial} - m)\psi, \tag{11.2}$$

$$\mathcal{L}_{int} = -eA^\mu\bar{\psi}\gamma_\mu\psi. \tag{11.3}$$

The fine structure constant of QED is $\alpha = e^2/4\pi \approx 1/137$.

- Photon propagator: $\overbrace{A_\mu(x)A_\nu}(y) = \langle 0|T(A_\mu(x)A_\nu(y))|0\rangle \cdot \mathbb{1}$

But recall:

$$A_\mu = \int d\Omega_k \left(e^{-ikx}a_\mu(k) + e^{ikx}a_\mu^\dagger(k)\right), \tag{11.4}$$

$$\langle 0|a_\mu(k)a_\nu^\dagger(k')|0\rangle = -\eta_{\mu\nu}(2\pi)^3 2k^0\delta^3(\vec{k} - \vec{k'}). \tag{11.5}$$

Hence the photon propagator can be written in position space:

$$\overbrace{A_\mu(x)A_\nu}(y) = -\eta_{\mu\nu}G_F(x - y, m = 0) = \int \frac{d^4k}{(2\pi)^4}\frac{-i\eta_{\mu\nu}e^{-ik(x-y)}}{k^2 + i\varepsilon}, \tag{11.6}$$

and in momentum space:

$$\mu \overset{\overset{k}{\leftarrow}}{\sim\sim} \nu \quad = \quad \frac{-i\eta_{\mu\nu}}{k^2 + i\varepsilon} \tag{11.7}$$

- Fermion propagator: $\psi_\alpha\overbrace{(x)\bar{\psi}_\beta}(y) = S_{\alpha\beta}(x - y) = \langle 0|T(\psi_\alpha(x)\bar{\psi}_\beta(y))|0\rangle$

In set 23 it is proven that:

$$(i\partial\!\!\!/_x - m)_{\alpha\beta} S_{\beta\gamma}(x-y) = i\delta^4(x-y)\delta_{\alpha\gamma}, \tag{11.8}$$

$$S = i(i\partial\!\!\!/ - m)^{-1} \tag{11.9}$$

Thus the propagator reads in position space:

$$S_{\alpha\beta} = \int \frac{d^4p}{(2\pi)^4} \left(\frac{i}{p\!\!\!/ - m + i\varepsilon} \right)_{\alpha\beta} e^{-ip(x-y)} \tag{11.10}$$

$$= \int \frac{d^4p}{(2\pi)^4} \frac{i(p\!\!\!/ + m)_{\alpha\beta}}{p^2 - m^2 + i\varepsilon} e^{-ip(x-y)} \tag{11.11}$$

$$\tag{11.12}$$

And in momentum space:

$$\alpha \xleftarrow{p} \beta \quad = \quad \frac{i(p\!\!\!/ + m)_{\alpha\beta}}{p^2 - m^2 + i\varepsilon} \tag{11.13}$$

- Photon external lines:

A photon state is characterized by momentum and spin. A general states reads:

$$|k, \varepsilon\rangle \equiv \varepsilon^\mu a_\mu^\dagger(k)|0\rangle, \tag{11.14}$$

where the polarization respecting:

$$\begin{cases} \varepsilon_\mu^* \varepsilon^\mu = -1 \\ k^\mu \varepsilon_\mu = 0 \end{cases} \tag{11.15}$$

- Incoming photon: $\overbrace{A_\mu(x)|k,\varepsilon\rangle} \equiv \langle 0|A_\mu(x)|k,\varepsilon\rangle = e^{-ikx}\varepsilon_\mu(k) =$ 
- Outgoing photon: $\langle k,\varepsilon|\overbrace{A_\mu(x)} \equiv \langle k,\varepsilon|A_\mu(x)|0\rangle = e^{-ikx}\varepsilon_\mu(k) =$ 
- Fermion external lines:

A fermion is characterized by its momentum and spin.

$$\psi_\alpha(x) = \int d\Omega_p \sum_{r=1}^{2} \left(e^{-ipx} u_\alpha(r,p) a(r,p) + e^{ipx} v_\alpha(r,p) b^\dagger(r,p) \right) \tag{11.16}$$

$$\text{Particle state: } |p,r\rangle = a^\dagger(r,p)|0\rangle \tag{11.17}$$

$$\text{Anti-particle state: } |\bar{p},\bar{r}\rangle = b^\dagger(\bar{r},\bar{p})|0\rangle \tag{11.18}$$

- Incoming fermion: $\psi_\alpha(x)|p,r\rangle \equiv \langle 0|\psi_\alpha(x)|p,r\rangle = e^{-ipx}u_\alpha(r,p) =$ 
- Outgoing fermion: $\langle p,r|\bar{\psi}_\alpha(x) \equiv \langle p,r|\bar{\psi}_\alpha(x)|0\rangle = e^{ipx}\bar{u}_\alpha(r,p) =$ 

- Incoming anti-fermion: $\bar{\psi}_\alpha(x)|\bar{p},\bar{r}\rangle \equiv \langle 0|\bar{\psi}_\alpha(x)|\bar{p},\bar{r}\rangle = e^{-ipx}\bar{v}_\alpha(\bar{r},\bar{p}) =$ 

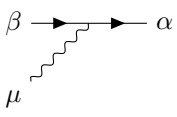
- Outgoing anti-fermion: $\langle\bar{p},\bar{r}|\psi_\alpha(x) \equiv \langle\bar{p},\bar{r}|\psi_\alpha(x)|0\rangle = e^{ipx}v_\alpha(\bar{r},\bar{p}) =$ 

- Vertex:

The lagrangian of interaction in QED has the form:

$$\mathcal{L}_{int} = -eA_\mu\bar{\psi}\gamma^\mu\psi \equiv -e\left(\gamma^\mu\right)_{\beta\alpha}A_\mu\bar{\psi}_\beta\psi_\alpha \tag{11.19}$$

which follows in the following vertex rule:



$$= -ie\gamma^\mu_{\beta\alpha} \tag{11.20}$$

Finally, here are all the QED Feynman rules in momentum space.

External lines

- Incoming photon: 

$= \varepsilon_\mu(k)$

- Outgoing photon: 

$= \varepsilon^*_\mu(k)$

- Incoming fermion: 

$= u_\alpha(p,r)$

- Outgoing fermion: 

$= \bar{u}_\alpha(p,r)$

- Incoming anti-fermion: 

$= \bar{v}_\alpha(\bar{p},\bar{r})$

- Outgoing anti-fermion: 

$= v_\alpha(\bar{p},\bar{r})$

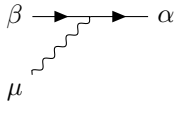
Internal lines

- Photon propagator: 

$= \frac{-i\eta_{\mu\nu}}{k^2+i\varepsilon}$

- Fermion propagator: 

$= \frac{i(\not{p}+m)_{\alpha\beta}}{p^2-m^2+i\varepsilon}$

- Vertex: 

$= -ie\gamma^\mu_{\beta\alpha}$

The Feynman rules listed above can be used to compute any process in QED. To compute a process, momentum conservation must be imposed at each vertex. Thus, a generic factor of $(2\pi)^4\delta^4(P_f - P_i)$ can be factored out.

A general in-state is written is the form:

$$|\text{in}\rangle = |\underbrace{(k_1,\varepsilon_1),...(k_n,\varepsilon_n)}_{\text{photons}};\underbrace{(p_1,r_1),...(p_q,r_q)}_{\text{electrons}};\underbrace{(\bar{p}_1,\bar{r}_1),...(\bar{p}_q,\bar{r}_q)}_{\text{positrons}}\rangle \tag{11.21}$$

Out states are written in a similar way. The S-matrix can be written using Dyson series, that is $S = \sum_n S_n$. The n^{th}-therm of the serie reads:

$$S_n = \frac{1}{n!}(-ie)^n \int d^4x_1...d^4x_n T(A_{\mu_1}\bar{\psi}\gamma^{\mu_1}\psi(x_1)...A_{\mu_n}\bar{\psi}\gamma^{\mu_n}\psi(x_n)) \tag{11.22}$$

Therefore, any matrix element can be computed by contracting the terms of the Dyson serie with the in and out-state, recalling that all terms must be contracted. The matrix element will then be proportional to:

$$\frac{1}{n!}(-ie)^n \int d^4x_1...d^4x_n \langle \text{ out } |T(A_{\mu_1}\bar{\psi}\gamma^{\mu_1}\psi(x_1)...A_{\mu_n}\bar{\psi}\gamma^{\mu_n}\psi(x_n))| \text{ in } \rangle. \tag{11.23}$$

Furthermore, residual loop momenta must be integrated over, by adding an integral of the type:

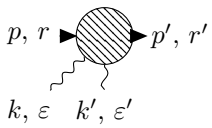
$$\int \prod_{i=1}^{N_L} \frac{d^4k_i}{(2\pi)^4}, \quad N_L = N_p - N_v + 1, \tag{11.24}$$

with N_l the number of loops, N_p the number of propagators and N_v the number of vertices. Note also that each closed fermionic loop will add an additional factor of (-1) due to the anti-commutation of fermions.

11.2 Compton scattering

A first computation that can be performed using the Feynman rules listed above is the Compton scattering, which describes the interaction between a photon and an electron.

$$\langle (k',\varepsilon'),(p',r')|S|(k,\varepsilon),(p,r)\rangle \;=\; \tag{11.25}$$



There are two lowest order diagrams, called s and t-channels. The difference between the different channels is the value of the momentum of the propagator. It can be obtained by conservation of momentum at each vertex.

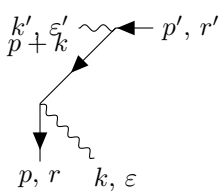


Figure 11.1: s-channel

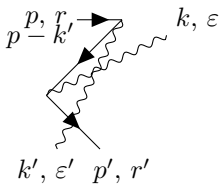


Figure 11.2: t-channel

These two diagrams are related by the crossing symmetry $(k,\varepsilon) \leftrightarrow (-k',\varepsilon'^*)$. Let us first have a look at the s-channel first. Using the Feynman rules for QED, one can identify each external lines, vertices and propagators and write the matrix element as:

$$i\mathcal{M}_s = \varepsilon'^*_\nu(k')\bar{u}(p',r')(-ie\gamma^\nu)\frac{i}{\not p + \not k - m}(-ie\gamma^\mu)u(p,r)\varepsilon_\mu(k). \tag{11.26}$$

The t-channel admits a different configuration of external line linked to each vertex, and will read:

$$i\mathcal{M}_t = \varepsilon_\nu(k)\bar{u}(p',r')(-ie\gamma^\nu)\frac{i}{\not p - \not k' - m}(-ie\gamma^\mu)u(p,r)\varepsilon'^*_\mu(k'). \tag{11.27}$$

One would have obtained the same result if one had considered the Wick contraction of the interaction Lagrangian with external legs at first order in perturbation theory. However, Feynman diagrams allow a much more efficient way to write such processes.

The total matrix element of Compton scattering is the sum of the s and t-channels, that is:

$$i\mathcal{M} = -ie^2\varepsilon'^*_\mu(k')\varepsilon_\nu(k)\left[\bar{u}(p',r')\left(\gamma^\mu\frac{i}{\not p + \not k - m}\gamma^\nu + \gamma^\nu\frac{i}{\not p - \not k' - m}\gamma^\mu\right)u(p,r)\right]. \tag{11.28}$$

This matrix element is manifestly Lorentz invariant since all Lorentz indices are contracted.

11.2.1 Ward identity

Recall the properties of the photon Hilbert space:

1. $|k,\varepsilon\rangle \equiv \varepsilon a^{\mu\dagger}(k)|0\rangle$, with $k^\mu\varepsilon_\mu = 0$.
2. There exists a relation of equivalence $\varepsilon_\mu(k) \sim \varepsilon_\mu(k) + bk_\mu$.

These two properties imply that there exists only two physical polarizations for the photon. Furthermore, b should not affect any physical quantity due to the equivalence relation. Hence the S-matrix should be unaffected by $\varepsilon_\mu(k) \to \varepsilon_\mu(k) + bk_\mu$, i.e.:

$$\mathcal{M}(\varepsilon,\varepsilon') = \mathcal{M}(\varepsilon + bk, \varepsilon' + b'k') = (\varepsilon_\mu + bk_\mu)(\varepsilon'^*_\nu + b'k'_\nu)\mathcal{M}^{\mu\nu} \tag{11.29}$$

In Eq. 11.28, one had found a result of the form $\mathcal{M} = \varepsilon'^*_\mu\varepsilon_\nu\mathcal{M}^{\mu\nu}$. Thus, a sufficient condition to satisfy Eq. 11.29 would be that:

$$k'_\mu\mathcal{M}^{\mu\nu} = k_\nu\mathcal{M}^{\mu\nu} = 0. \tag{11.30}$$

Recalling our result, and noting that by momentum conservation $p - k' = p' - k$, one has:

$$\mathcal{M}^{\mu\nu} = -e^2\bar{u}(p',r')\left(\gamma^\mu\frac{i}{\not p + \not k - m}\gamma^\nu + \gamma^\nu\frac{i}{\not p' - \not k - m}\gamma^\mu\right)u(p,r). \tag{11.31}$$

Let us try to contract this with k_ν. One will also use that $(\not p - m)u(p,r) = 0$ and $\bar{u}(\not p' - m) = 0$.

$$\begin{aligned}
k_\nu\mathcal{M}^{\mu\nu} &= -e^2\bar{u}\left[\gamma^\mu\frac{i}{\not p + \not k - m}\not k + \not k\frac{i}{\not p' - \not k - m}\gamma^\mu\right]u && (11.32)\\
&= -e^2\bar{u}\left[\gamma^\mu\frac{i}{\not p + \not k - m}(\not p + \not k - m - \not p + m) + (-\not p' + \not k + m + \not p' - m)\frac{i}{\not p' - \not k - m}\gamma^\mu\right]u && (11.33)\\
&= -e^2\bar{u}\left[\gamma^\mu - \gamma^\mu\right]u && (11.34)\\
&= 0. && (11.35)
\end{aligned}$$

Eq. 11.30 holds for number of external photon to all order in perturbation theory. It is known as the Ward identity. It is a quantum manifestation of the continuity equation $\partial_\mu J^\mu = 0$. Indeed, recall the gauge invariance:

$$\varepsilon_\mu \to \varepsilon_\mu + b k_\mu \;\Leftrightarrow\; A_\mu \to A_\mu + \partial_\mu \alpha \tag{11.36}$$

$$\int \mathcal{L}_I = -e \int A_\mu J^\mu \to -e \int A_\mu J^\mu - e \int \alpha \partial_\mu J^\mu \tag{11.37}$$

Hence the invariance of the Lagrangian implies that $\partial_\mu J^\mu = 0$.

11.2.2 Classical approximation of Compton scattering

Applying eq. (10.84) to the scattering amplitude in eq. (11.28), one derives the differential cross-section for Compton scattering. This computation is done in one of the exercise sessions. The result, known as the Klein-Nishina formula, for an electron initially at rest $p = (m_e, \vec{0})$ reads

$$\frac{d\sigma}{d\cos\theta} = \frac{\pi\alpha^2}{m_e^2}\left(\frac{\omega'}{\omega}\right)^2 \left[\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \sin^2\theta\right] \tag{11.38}$$

where $\alpha = e^2/4\pi \simeq 1/137$ is the fine-structure constant, ω and ω' are respectively the initial and final energies of the photon and θ is its deflection angle: $\vec{k}\cdot\vec{k}'/\omega\omega' = \cos\theta$. In the soft limit $\omega \to 0$ one has $\omega'/\omega \to 1$ and the result reduces to the well known Thomson cross section of classical electrodynamics

$$\left.\frac{d\sigma}{d\cos\theta}\right|_{\text{Thomson}} = \frac{\pi\alpha^2}{m_e^2}\left[1 + \cos^2\theta\right] . \tag{11.39}$$

This limiting result invites us to better understand the conditions for the applicability of classical physics in the description of light matter scattering.

One condition is that the electron must be localized in a wave packet of size $\Delta x \ll \lambda \equiv k_\gamma^{-1}$ during the passage of the electromagnetic wave. Furthermore, by uncertainty principle, $\Delta x \Delta k_e \sim 1$. Thus, one has the condition:

$$\Delta k_e \gg k_\gamma. \tag{11.40}$$

However, this property of localization must be maintained at least for the time $\omega_\gamma^{-1} \equiv k_\gamma^{-1}$, corresponding to one period of oscillation of the electromagnetic field. A very narrow wave packet will diffuse and spread very fast. If $\Delta k_e \gtrsim m_e$, the spread velocity is ~ 1. Thus, over a time ω_γ^{-1}, it will spread over the whole wavelength. One must then also impose $\Delta k_e \ll m_e$. In the end, the two conditions become:

$$k_\gamma \ll \Delta k_e \ll m_e\,, \tag{11.41}$$

corresponding to the scattering of a non-relativistic with a very soft photon.

The above conditions are necessary, but are they really sufficient? In principle they are not, given we should also ask for the electron motion to be semi-classical. The condition of semi-classicality normally corresponds to the request that the 'action' be much greater than 1. But which 'action'? Reasonably that should be the action corresponding to one full swing of the electron at the passage of the wave. A simple estimate gives

$$S \sim \int \frac{P_e^2}{2m_e} dt \sim \frac{e^2E^2}{m_e\omega_\gamma^3}, \tag{11.42}$$

with E the electric field and we used $P_e \sim eE/\omega_\gamma$ integrating over a time $\sim 1/\omega_\gamma$. This leads to the condition:

$$\frac{e^2E^2}{m_e\omega_\gamma^3} \gg 1, \tag{11.43}$$

which corresponds to the WKB condition for the applicability of the semi-classical approximation. Now indicating by $r_e = e^2/m_e$ the classical electron radius (which controls the Thomson scattering cross-section) and using the relation $E^2 = n_\gamma\omega_\gamma$ with n_γ the photon density, the above relation can be rewritten ($\lambda_\gamma = 1/\omega_\gamma$)

$$n_\gamma \lambda_\gamma^2 r_e \gg 1, \tag{11.44}$$

which corresponds to some large photon density.

However, when explicit computations are performed,one finds that the result does not depend on the photon number density, at least at lowest order in the expansion in α Thus the classical result holds simply when Eq. 11.41 is satisfied. Thus:

$$\sigma_{\text{Compton}}\big|_{\omega_\gamma \to 0} \equiv \sigma_{\text{Thomson}}. \tag{11.45}$$

Chapter 12

The Standard Model

12.1 Yang-Mills theory

QED is based on a $U(1)$ gauge group. In this case, the gauge transformation of the photon and fermion fields are the following:

$$\begin{cases} A'_\mu = A_\mu - \partial_\mu \alpha \\ \psi' = e^{iq\alpha}\psi \end{cases} \tag{12.1}$$

which can be written using a matrix $U \equiv e^{iq\alpha} \in U(1)$ as:

$$\begin{cases} A'_\mu = A_\mu - iU\partial_\mu U^{-1} \\ \psi' = U^q \psi \end{cases} \tag{12.2}$$

The lagrangian of QED is constructed with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and the covariant derivative $D_\mu \psi = (\partial_\mu + iqA_\mu)\psi$.

The particularity of $U(1)$ is that it is an abelian group. It is well known that the gauge group of the standard model is $SU(3) \times SU(2) \times U(1)$. Hence, one needs a generalization of the QED lagrangian for non-abelian groups in order for the standard model to contain strong and week interactions. Let us therefore have a look at non-abelian groups.

Let $U = e^{i\alpha_A T_A}$ an element of the group, with T_A the abstract generators of the Lie group. T_A are hermitian for compact groups. In this case, $U^\dagger = U^{-1}$. For a field in a representation r, a gauge transformation reads:

$$\psi'_r = e^{i\alpha_A T_A^{(r)}} \psi_r. \tag{12.3}$$

This time, $T_A^{(r)}$ is an explicit matrix representation of the Lie algebra. Furthermore, in order to have a local transformation, the Lie parameter depends on the position, i.e. $\alpha_A \equiv \alpha_A(x)$.

An example of such gauge transformation is for the gauge group $G = SU(2)$, with ψ lying in the fundamental representation, i.e. it is a doublet. In this case, the gauge transformation reads:

$$\psi' = e^{i\alpha_A \sigma_A}\psi, \tag{12.4}$$

with σ_A the Pauli matrices, the generators of $SU(2)$ in the fundamental representation.

In order to construct the covariant derivative, one needs to introduce vector field(s) $A_\mu \equiv A_\mu^A T_A$. A_μ is an element of the Lie algebra, and A_μ^A lives in the adjoint representation. There are therefore not only one gauge field, such as the photon for a $U(1)$ gauge theory, but several of them, in fact as many as the size of the adjoint representation of the group. There are three of them for $SU(2)$ and eight for $SU(3)$ for instance.

To find how the gauge fields transform, Recall that:

$$[T_A, T_B] = if_{ABC}T_C, \tag{12.5}$$

$$(T_C^{adj})_{AB} \equiv if_{ABC} \tag{12.6}$$

Thus one defined the gauge transformation as:

$$A'_\mu \equiv UA_\mu U^{-1} - iU\partial_\mu U^{-1}. \tag{12.7}$$

Note that one retrieves QED for an abelian group. The covariant derivative for a fiend ψ in the representation r of the group can be written as:

$$D_\mu \psi_r \equiv (\partial_\mu + iA_\mu^A T_A^{(r)})\psi_r \equiv (\partial_\mu + iA_\mu)\psi_r. \tag{12.8}$$

Let us check that this transforms covariantly:

$$\begin{aligned} D_\mu\psi &\to (\partial_\mu + iUA_\mu U^{-1} + U\partial_\mu U^{-1})U\psi && (12.9)\\ &= U\left[\partial_\mu + iA_\mu + \partial_\mu U^{-1}U + U^{-1}\partial_\mu U\right]\psi && (12.10)\\ &= U\left[\partial_\mu + iA_\mu + \partial_\mu (U^{-1}U)\right]\psi && (12.11)\\ &= UD_\mu\psi, && (12.12) \end{aligned}$$

where we used that $\partial_\mu(U^{-1}U) = \partial_\mu \mathbb{1} = 0$.

One also need to construct the field strength tensor for the vector fields. It is defined as:

$$\begin{aligned} iF_{\mu\nu} &\equiv (\partial_\mu + iA_\mu)(\partial_\nu + iA_\nu) - (\partial_\nu + iA_\nu)(\partial_\mu + iA_\mu) && (12.13)\\ &= i(\partial_\mu A_\nu - \partial_\nu A_\mu) - [A_\mu, A_\nu] && (12.14) \end{aligned}$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu) + i\left[A_\mu, A_\nu\right] \equiv F_{\mu\nu}^A T_A. \tag{12.15}$$

We can see that it corresponds to the $U(1)$ field strengt tensor for an abelian gauge theory, as expected. $F_{\mu\nu}$ also belongs to the Lie algebra. Indeed, under a gauge transformation, one can check that:

$$F_{\mu\nu} \to F'_{\mu\nu} = UF_{\mu\nu}U^{-1}. \tag{12.16}$$

Thus $F_{\mu\nu}$ is covariant.

Note also that one can always choose a basis of the Lie algebra for which $\mathrm{Tr}(T_A^{(r)}T_B^{(r)}) = t_r\delta_{AB}$, with t_r the Dynkin index. One can now look at the invariant terms we can construct:

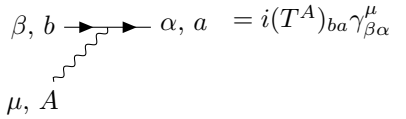
- $\bar{\psi}D_\mu\psi \to \bar{\psi}U^\dagger UD_\mu\psi = \bar{\psi}D_\mu\psi$
- $\mathrm{Tr}\left(F_{\mu\nu}F^{\mu\nu}\right) \to \mathrm{Tr}\left(UF_{\mu\nu}U^\dagger UF^{\mu\nu}U^\dagger\right) = \mathrm{Tr}\left(F_{\mu\nu}F^{\mu\nu}\right)$
- $F_{\mu\nu}^A F^{A\mu\nu} = \frac{1}{t_r}\mathrm{Tr}\left(F_{\mu\nu}F^{\mu\nu}\right)$

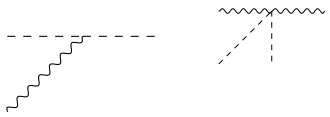
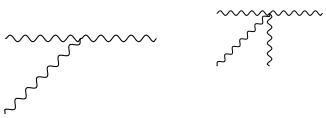
Hence the most general lagrangian containing a fermion and a scalar field reads:

$$\mathcal{L} = -\frac{1}{4g^2}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\gamma^\mu D_\mu\psi + (D_\mu\phi)^\dagger(D^\mu\phi) + \{\text{all possible gauge invariant interactions of } \psi \text{ and } \phi\} \tag{12.17}$$

Finally, one wants to have a properly normalized kinetic term. Thus, one can rescale the field $A_\mu \to gA_\mu$. Therefore, the kinetic term becomes $-\frac{1}{4g^2}F_{\mu\nu}F^{\mu\nu} \to -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, and the covariant derivative $D_\mu = \partial_\mu + iA_\mu \to \partial_\mu + igA_\mu$.

Let us now write the Feynman rules for the vertices of the theory. Recall that the indices α and β denote Lorentz indices, while a and b denote gauge indices.

- $\bar{\psi} D_\mu \psi \longrightarrow$ 

- $(D_\mu \phi)^\dagger (D^\mu \phi) \longrightarrow$ 
- $F_{\mu\nu} F^{\mu\nu} \longrightarrow$ 

Hence, the bosons are themselves charged. The introduction of a non abelian gauge theory allowed a self interaction between the gauge bosons, which does not occur with an abelian gauge theory such as QED.

12.2 Basics of the Standard Model

12.2.1 Gauge group of the Standard Model

The Standard model is a theory based on the gauge group:

$$G = \underbrace{SU(3)}_{\text{color}} \times \underbrace{SU(2)}_{\text{weak isospin}} \times \underbrace{U(1)}_{\text{hypercharge}} \tag{12.18}$$

The $SU(3)$ sector describes strong interactions. This sector of the Standard model is called Quantum Chromodynamics (QCD). The $SU(2) \times U(1)$ sector describes electro-weak interactions. It contains both weak interactions and QCD.

- $SU(3)$ has 8 generators, thus there will be 8 gluons $G^a_\mu, a = 1...8$
- $SU(2)$ has 3 generators, thus there will be 3 vector bosons $W^I_\mu, I = 1...3$
- $UU(1)$ has 1 generator, thus there will be 1 vector bosons B_μ

Each spinor or scalar field belongs to an irrep of G. The general covariant derivative reads:

$$D_\mu \psi = (\partial_\mu + i G^a_\mu X^a + i W^I_\mu T^I + i B_\mu \frac{Y}{2}) \psi, \tag{12.19}$$

with X^a in the irrep of $SU(3)$, T^I in the irrep of $SU(2)$ and Y the hypercharge of ψ. The factor $1/2$ multiplying the hypercharge is conventional, and could be absorbed in the definition of Y.

One has now all the tools to construct the lagrangian of the SM. One now need to specify the field content of the model.

12.2.2 Field content of the Standard Model

Quarks

In the Standard Model, quarks are either in a $SU(2)$ doublet belonging to the $(1/2, 0)$ representation of the Lorentz group, or in a $SU(2)$ singlet belonging to the $(0, 1/2)$ representation of the Lorentz group.

- $q_i^L = (\underline{3}, \underline{2}, 1/3)$
- $u_i^R = (\underline{3}, \underline{1}, 4/3)$
- $d_i^R = (\underline{3}, \underline{1}, -2/3)$

The underlined numbers denote the representation in which the quark lie in $SU(3)$ and $SU(2)$ respectively, while the third number denotes the hypercharge of the quark. Hence all quarks belong to the fundamental representation of $SU(3)$. They can therefore interact via strong interaction.

The index i denotes the family. There are three exact copies of these quarks, differentiated only by their mass. The families are:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix}. \tag{12.20}$$

Leptons

The leptons are either in a $SU(2)$ doublet or in a $SU(2)$ singlet.

- $l_i^L = (\underline{1}, \underline{2}, -1)$
- $e_i^R = (\underline{1}, \underline{1}, -2)$

Note that contrary to the quarks, there are no right-handed neutrinos in the Standard Model. Again, there are three families:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} \tag{12.21}$$

The covariant derivatives will depend on the representation in which the quark lies.

- $e_R = (\underline{1}, \underline{1}, -2) \quad D_\mu e_R = (\partial_\mu - iB_\mu)e_R$
- $l_L = (\underline{1}, \underline{2}, -1) \quad D_\mu l_L = (\partial_\mu + \frac{i}{2}\sigma_I W_\mu^I - \frac{i}{2}B_\mu)l_L$
- $u_R = (\underline{3}, \underline{1}, 4/3) \quad D_\mu u_R = (\partial_\mu + \frac{i}{2}\lambda_a G_\mu^a + i\frac{2}{3}B_\mu)u_R$
- $d_R = (\underline{3}, \underline{1}, -2/3) \quad D_\mu d_R = (\partial_\mu + \frac{i}{2}\lambda_a G_\mu^a - \frac{i}{3}B_\mu)d_R$
- $q_L = (\underline{3}, \underline{1}, -2/3) \quad D_\mu q_L = (\partial_\mu + \frac{i}{2}\lambda_a G_\mu^a + \frac{i}{2}\lambda_a G_\mu^a + \frac{i}{6}B_\mu)q_L$

The λ_a are the Gell-Mann matrices, which are the generators of $SU(3)$ in the fundamental representation. They are 3×3 matrices, the $SU(3)$ equivalent of the Pauli matrices for $SU(2)$.

Scalars

The scalar of the Standard Model is the Higgs, a complex $SU(2)$ doublet:

$$H = (\underline{1}, \underline{2}, 1) = \begin{pmatrix} H_+ \\ H_0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{12.22}$$

$$D_\mu H = (\partial_\mu + \frac{i}{2}\sigma_I W_\mu^I + \frac{i}{2}B_\mu)H \tag{12.23}$$

12.2.3 Lagrangian of the Standard Model

Using all the invariant terms that can be obtained from the content of the Standard model, the full lagrangian reads:

$$\begin{aligned}\mathcal{L}_{SM} = &-\frac{1}{4g_s^2}G^a_{\mu\nu}G^{\mu\nu a} - \frac{1}{4g^2}W^I_{\mu\nu}W^{\mu\nu I} - -\frac{1}{4g_Y^2}B_{\mu\nu}B^{\mu\nu} + \sum_{i=1}^{3}\left[i\bar{q}_i\not{D}q_i + i\bar{u}_i\not{D}u_i + i\bar{d}_i\not{D}d_i + i\bar{l}_i\not{D}l_i + i\bar{e}_i\not{D}e_i\right] \\ &+ (D_\mu H)^\dagger(D^\mu H) - \left[-m^2H^\dagger H + \frac{\lambda}{2}(H^\dagger H)^2\right] + Y^u_{ij}(\bar{q}_i\tilde{H})u_j + Y^d_{ij}(\bar{q}_iH)d_j + Y^e_{ij}(\bar{l}_i\tilde{H})e_j,\end{aligned} \tag{12.24}$$

with $\tilde{H} = \varepsilon H^*$. Note that $\tilde{H}$ and H transform the same under $SU(2)$ since $\varepsilon^{-1}\sigma_I^*\varepsilon = -\sigma_I$.

12.3 Higgs mechanism

There is no explicit mass term for both the fermions and the vectors. Such terms would violate the gauge theory required. Hence, at first sight, all particles are massless, in disagreement with observations. One will see they will acquire a mass through the Higgs mechanism. The only field having an explicit mass term is the Higgs scalar, but it has the "wrong" sing. The full Higgs potential reads:

$$V(H) = -m^2H^\dagger H + \frac{\lambda}{2}(H^\dagger H)^2. \tag{12.25}$$

The correct vacuum is not well described by expanding around $H = 0$. One must expand the theory around the minimum of $V(H)$. It has an infinite set of minima satisfying:

$$H^\dagger H = \frac{m^2}{\lambda} \equiv v^2, \tag{12.26}$$

where v is defined as the vacuum expectation value (VEV) of the Higgs doublet. It is measured to be around $v \approx 174GeV$. By performing a $SU(2)_L$ rotation, one can choose the vacuum expectation value to be real and in the lower component. At the minimum, the Higgs doublet becomes:

$$\langle H\rangle = \begin{pmatrix}0\\ v\end{pmatrix}, \tag{12.27}$$

One can now expand the Higgs doublet in Eq. 12.24 around the minimum of the potential. This process will give an explicit mass term to all the particles, except to the neutrinos, which is expected!

12.3.1 Fermion masses

The term in the Standard Model Lagrangian that will give rise to the mass of the fermions is the following:

$$\mathcal{L}_{fermions} = Y^u_{ij}(\bar{q}_i\tilde{H})u_j + Y^d_{ij}(\bar{q}_iH)d_j + Y^e_{ij}(\bar{l}_i\tilde{H})e_j \tag{12.28}$$

$$= Y^u_{ij}v\bar{u}_{iL}u_{jR} + Y^d_{ij}v\bar{d}_{iL}d_{jR} + Y^e_{ij}v\bar{e}_{iL}e_{jR} + \text{Interaction with Higgs} \tag{12.29}$$

Thus one can identify the masses of the fermions:

$$M^u_{ij} = Y^u_{ij}v, \quad M^d_{ij} = Y^d_{ij}v, \quad M^e_{ij} = Y^e_{ij}v \tag{12.30}$$

This mass matrix is diagonal in the family basis:

$$M^u_{diag} = \begin{pmatrix} y_u & 0 & 0 \\ 0 & y_c & 0 \\ 0 & 0 & y_t \end{pmatrix} v = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix}, \; M^d_{diag} = \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix}, \; M^e_{diag} = \begin{pmatrix} m_e & 0 & 0 \\ 0 & m_\mu & 0 \\ 0 & 0 & m_\tau \end{pmatrix} \tag{12.31}$$

What about the neutrinos mass? It cannot be obtained by Higgs mechanism. Indeed:

$$l_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, \quad l_L \langle H \rangle = v e_L, \tag{12.32}$$

Thus ν_L is projected out due to the absence of right-handed neutrinos. Hence $m_\mu = 0$ at this stage, which approximate nicely the reality.

12.3.2 Vector boson masses

The vacuum expectation value of the Higgs boson breaks $SU(2) \times U(1)_Y$. Indeed, for T_1, T_2 and T_3 the generators of $SU(2)$:

$$T_1 \langle H \rangle = \frac{\sigma_1}{2} \begin{pmatrix} 0 \\ v \end{pmatrix} = \begin{pmatrix} v/2 \\ 0 \end{pmatrix}, \quad T_2 \langle H \rangle = \frac{\sigma_2}{2} \begin{pmatrix} 0 \\ v \end{pmatrix} = \begin{pmatrix} -iv/2 \\ 0 \end{pmatrix}, \quad T_3 \langle H \rangle = \frac{\sigma_3}{2} \begin{pmatrix} 0 \\ -v/2 \end{pmatrix} = \begin{pmatrix} -iv/2 \\ 0 \end{pmatrix} \tag{12.33}$$

$$\frac{Y}{2} \langle H \rangle = \frac{1}{2} \begin{pmatrix} 0 \\ -v/2 \end{pmatrix} = \begin{pmatrix} 0 \\ v/2 \end{pmatrix} \tag{12.34}$$

Hence, all generators are broken by the vacuum. However, there is a residual symmetry- Indeed, $Q = T_3 + Y/2$ remains unbroken.

$$Q \langle H \rangle = \frac{\sigma_3 + 1}{2} \begin{pmatrix} 0 \\ v \end{pmatrix} = 0 \tag{12.35}$$

Q is the generator of the electric charge. By applying it the the particles of the Standard Model, one can obtain the value of their electric charge:

$$Ql = \frac{\sigma_3 - 1}{2} \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} = \begin{pmatrix} 0 \\ -e_L \end{pmatrix}, \quad Ql = \frac{\sigma_3 + 1/3}{2} \begin{pmatrix} u_L \\ d_L \end{pmatrix} = \begin{pmatrix} \frac{2}{3} u_L \\ -\frac{1}{3} d_L \end{pmatrix}, \tag{12.36}$$

Hence $q_\nu = 0, q_e = -1, q_u = 2/3$ and $q_d = -1/3$ as expected. A proton, which is formed of the quarks uud, will therefore have an electric charge $q_p = 1$, while a neutron, composed of udd, will have $q_n = 0$.

Let us now have a look at excitations of the Higgs boson around the vacuum. One can write them as:

By preforming an appropriate gauge transformation on the fields, one can choose the Higgs field in the unitary gauge, i.e.:

$$H = \begin{pmatrix} 0 \\ v \end{pmatrix} + \begin{pmatrix} \pi_1 + i\pi_2 \\ h + i\pi_3 \end{pmatrix}, \tag{12.37}$$

where h and π_i are the field fluctuations around $\langle H \rangle$. One can find the transformation of these fields under $SU(2) \times U(1)$:

$$\begin{aligned} H &\to e^{i\frac{\sigma_I\alpha_I}{2}} e^{i\frac{\alpha_Y}{2}} H \\ &\approx (1 + i\frac{\sigma_I\alpha_I}{2} + i\frac{\alpha_Y}{2}) \left[\begin{pmatrix} 0 \\ v \end{pmatrix} + \begin{pmatrix} \pi_1 + i\pi_2 \\ h + i\pi_3 \end{pmatrix} \right] \\ &= \begin{pmatrix} 0 \\ v \end{pmatrix} + \begin{pmatrix} \pi_1 + i\pi_2 \\ h + i\pi_3 \end{pmatrix} + \frac{iv}{2} \begin{pmatrix} \alpha_1 + i\alpha_2 \\ -\alpha_3 + \alpha_Y \end{pmatrix} + \mathcal{O}(\alpha^2) \end{aligned} \tag{12.38}$$

Therefore:

$$\begin{aligned} \delta(\pi_1 + i\pi_2) &= \frac{v}{2}(\alpha_1 + i\alpha_2) + \mathcal{O}(\alpha^2), \\ \delta(\pi_3) &= \frac{v}{2}(-\alpha_3 + \alpha_Y) + \mathcal{O}(\alpha^2), \\ \delta(h) &= 0 + \mathcal{O}(\alpha^2) \end{aligned} \tag{12.39}$$

Hence, all π_i shift under a gauge symmetry, while h remains unchanged. The π_i behave exactly like the scalar ϕ in the massive vector model. Thus, one can choose a suitable gauge fixing to set them to 0. Exactly like the ϕ previously, the three π_i are eaten by three corresponding vector bosons which acquire a mass. However, h cannot be eliminated by gauge fixing. It is physical: it is the Higgs boson.

Let us perform a gauge transformation with $\alpha_1 = \alpha_2 = 0$, and $\alpha_3 = \alpha_Y \equiv \alpha$. In this case:

$$U = e^{i\frac{\sigma_3\alpha}{2}} e^{i\frac{\alpha}{2}} = e^{i\alpha(T_3 + Y/2)} = e^{i\alpha Q} \tag{12.40}$$

By applying U to H, one deduces that $Q_{\pi_1} = Q_{\pi_2} = 1$ and $Q_h = Q_{\pi_3} = 0$. Now let us fix $\pi_i = 0$. This choice of gauge is called the unitary gauge. The vector boson masses will emerge from the therm $(D_\mu H)^\dagger (D^\mu H)$ of the Standard Model Lagrangian:

$$(D_\mu H)^\dagger (D^\mu H) = (\partial_\mu h)^2 + (v + h)(0, 1) \left[\left(\frac{1}{2}\sigma_I W_\mu^I + \frac{1}{2} B_\mu \right)^2 \right] \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{12.41}$$

$$= (\partial_\mu h)^2 + \frac{(v + h)}{4} \left[(W_\mu^1)^2 + (W_\mu^2)^2 + (W_\mu^3 - B_\mu)^2 \right] \tag{12.42}$$

Te gauge bosons W_μ^1 and W_μ^2 acquire a mass. Furthermore, the linear combination $W_\mu^3 + B_\mu$ remains massless. This boson will become the photon. The linear combination orthogonal to the photon will give rise to the Z boson. In the photon field direction, one requires $W_\mu^3 - B_\mu = 0$, i.e. $W_\mu^3 = B_\mu = A_\mu$. Thus the kinetic term becomes:

$$-\frac{1}{4g^2} W_{\mu\nu}^3 W^{3\mu nu} - \frac{1}{4g_Y^2} B_{\mu\nu} B^{\mu\nu} \to -\frac{1}{4} \left(\frac{1}{g^2} + \frac{1}{g_Y^2} \right) A_{\mu\nu} A^{\mu\nu}. \tag{12.43}$$

Thus one can find the relation between the electric charge and the couplings g and g_Y:

$$\frac{1}{e^2} = \frac{1}{g^2} + \frac{1}{g_Y^2}. \tag{12.44}$$

The electric charge coupling is a well known measurable parameter of the Standard Model, and its value is approximately $e^2/4\pi \approx 1/137$.

One can now rescale the fields in order to have a canonically normalized kinetic term for W_μ^I and B_μ:

$$W_\mu^I \to g W_\mu^I, \quad B_\mu \to g_Y B_\mu \tag{12.45}$$

The kinetic part of the Standard Model Lagrangian becomes:

$$\mathcal{L}_{kin} = -\frac{1}{4} W_{\mu\nu}^I W^{I\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} + \left(\frac{v}{2} \right)^2 \left[g^2 (W_\mu^1)^2 + g^2 (W_\mu^2)^2 + (g W_\mu^3 - g_Y B_\mu)^2 \right] \tag{12.46}$$

One can write the mass matrix for the four vector bosons, which contains off diagonal terms:

$$M_v^2 = \frac{v^2}{2}\begin{pmatrix} g^2 & 0 & 0 & 0 \\ 0 & g^2 & 0 & 0 \\ 0 & 0 & g^2 & -gg_Y \\ 0 & 0 & -gg_Y & g_Y^2 \end{pmatrix} \tag{12.47}$$

The physical gauge bosons are the eigenstates of this matrix. Furthermore, a special linear combination of W_μ^1 and W_μ^2, which are already eigenstates with the same mass, will have an electric charge of ± 1. The four physical bosons are therefore:

$$W_\mu^\pm = \frac{W_\mu^1 \mp W_\mu^2}{\sqrt{2}}, \tag{12.48}$$

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu, \tag{12.49}$$

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu, \tag{12.50}$$

where we introduced the weak mixing angle:

$$\sin\theta_W \equiv \frac{g_Y}{\sqrt{g^2 + g_Y^2}}. \tag{12.51}$$

The masses of the bosons are :

$$m_W^2 = \frac{g^2 v^2}{2}, \quad m_Z^2 = \frac{(g^2 + g_Y^2)v^2}{2}, \quad m_A^2 = 0. \tag{12.52}$$

As expected, one has three massive gauge bosons of electric charge ± 1 and 0, and one massless gauge boson of electric charge 0. Note also the relations:

$$e^2 = \frac{g_y^2 g^2}{g^2 + g_Y^2} = g^2 \sin^2\theta_W, \tag{12.53}$$

$$\frac{m_W^2}{m_Z^2} = \frac{g^2}{g^2 + g_Y^2} = \cos\theta_W. \tag{12.54}$$

12.4 A few physical implications

From the covariant derivative of fermions:

$$\mathcal{L}_{current} = gW_\mu^+ J^{-\mu} + \frac{g}{\cos\theta_W} Z_\mu J_Z^\mu + eA_\mu J_A^\mu, \tag{12.55}$$

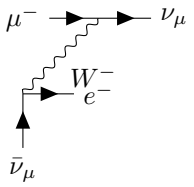
with the currents given by:

$$J_\mu^- = \frac{1}{\sqrt{2}}(\bar{\nu}_L^i \bar{\sigma}_\mu \nu_L^i - c_e \bar{e}_L^i \bar{\sigma}_\mu e_L^i - c_u \bar{u}_L^i \bar{\sigma}_\mu u_L^i + ...), \tag{12.56}$$

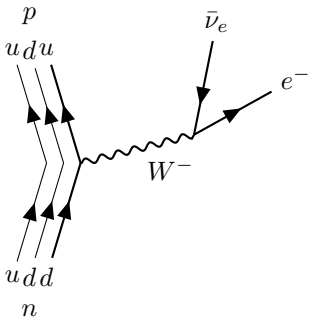
with c_e and c_u the coupling of the Z-boson with the fermions, which depends on the representation of each fermion. One can now compute processes using the Feynman rules, such as the muon decay or the neutron decay.

The masses of the different particles can be measured experimentally and one gets:

$$m_W \sim 80\text{GeV}, \quad m_n \sim 938\text{MeV}, \quad m_\nu \sim 106\text{MeV}. \tag{12.57}$$



Muon decay $\mu \to \nu_\mu\ e\ \bar{\nu}_e$



Neutron decay $n \to p\ e\ \bar{\nu}_e$

Thus, $m_W \gg m_n, m_\mu$. Hence these interactions appear weak due to the propagator of the W-boson. Indeed, in the low energy limit $q^2 \ll m_W^2$:

$$\langle W_\mu^+ W_\nu^- \rangle = \frac{-i\left(\eta_{\mu\nu} - \frac{q_\mu q_\nu}{m_W^2}\right)}{q^2 - m_W^2} \simeq \frac{i\eta_{\mu\nu}}{m_W^2}. \tag{12.58}$$

Thus the matrix element of the muon decay is approximately:

$$\mathcal{M} \sim (ig)^2 \left(\frac{i\eta_{\mu\nu}}{m_W^2}\right) J_\mu^- J_\nu^+ = -i\frac{g^2}{m_W^2} J_\mu^- J_\nu^+ \tag{12.59}$$

One can see this interaction as an effective theory with the effective Lagrangian:

$$\mathcal{L}_{eff} = -\frac{g^2}{m_W^2} J_\mu^- J_\nu^+ \tag{12.60}$$

This theory is known as the Fermi Theory of the β-decay, proposed by Fermi in 1932, well before the discovery of the W-boson. The coupling constant of this effective theory is the Fermi constant G_F, defined as:

$$\frac{g^2}{m_W^2} = \frac{2}{v^2} \equiv 4\sqrt{2}G_F, \quad G_F^{-1} = 2\sqrt{2}v^2 = (*00\ \text{GeV})^2, \tag{12.61}$$

with v the vev of the Higgs.

Appendix A

Proof of Wick's Theorem

A.1 First Wick Theorem

This Appendix will prove the first Wick Theorem, stated as follows. Let us consider a set of free (fermionic or bosonic) fields $\{\phi_n \equiv \phi_n(x_n)\}$. Then the time-ordered product of n fields can be written as:

$$T\phi_1 \cdots \phi_n =: \phi_1 \cdots \phi_n : + : \{\text{all possible chronological contractions}\} : \tag{A.1}$$

The proof goes by induction. Let's first consider the bosonic case. The first non-trivial case is $n = 2$. There, we have explicitly

$$\begin{aligned} T\phi_1(x_1)\phi_2(x_2) =& \theta(x_1^0 - x_2^0)\phi_1(x_1)\phi_2(x_2) + \theta(x_2^0 - x_1^0)\phi_2(x_2)\phi_1(x_1) \\ =& \theta(x_1^0 - x_2^0)\,[: \phi_1(x_1)\phi_2(x_2) : + \langle 0|\phi_1(x_1)\phi_2(x_2)|0\rangle \cdot \mathbb{1}] \\ &+ \theta(x_2^0 - x_1^0)\,[: \phi_2(x_2)\phi_1(x_1) : + \langle 0|\phi_2(x_2)\phi_1(x_1)|0\rangle \cdot \mathbb{1}] \\ =& : \phi_1(x_1)\phi_2(x_2) : + \langle 0|T\phi_1(x_1)\phi_2(x_2)|0\rangle \cdot \mathbb{1} \\ =& : \phi_1(x_1)\phi_2(x_2) : + : \phi_1(\overbrace{x_1)\phi_2(x_2}) : \end{aligned} \tag{A.2}$$

Therefore the theorem is true for $n = 2$. Let's now assume that the theorem is true for $n-1$ fields and let's prove it for n fields. Without loss of generality let's take $x_1^0 \geq x_2^0 \cdots \geq x_n^0$, such that:

$$T\phi_1(x_1) \cdots \phi_n(x_n) = \phi_1(x_1) \cdots \phi_n(x_n) = \left(\phi_1^+(x_1) + \phi_1^-(x_1)\right) T\phi_2(x_2) \cdots \phi_n(x_n) \tag{A.3}$$

where we decomposed the first field in terms of creation ϕ_1^+ and destruction ϕ_1^- operators. We can then treat them separately. The easiest one is ϕ_1^+ as the chronological contraction with any field vanishes by the properties of the vacuum state $|0\rangle$:

$$\overbrace{\phi_1^+ \phi_k} = \langle 0|\phi_1^+\phi_k|0\rangle \cdot \mathbb{1} = 0 \tag{A.4}$$

Knowing that Wick theorem is satisfied for $n-1$ fields, it follows that

$$T\phi_1^+\phi_2 \cdots \phi_n =: \phi_1^+ \cdots \phi_n : + : \{\text{all possible chronological contractions}\} : \tag{A.5}$$

For ϕ_1^-, the proof is done by commuting it with all the other fields:

$$\begin{aligned} T\phi_1^-\phi_2 \cdots \phi_n =& \phi_2 \cdots \phi_n\phi_1^- + [\phi_1^-, \phi_2 \cdots \phi_n] \\ =& \phi_2 \cdots \phi_n\phi_1^- + [\phi_1^-, \phi_2]\phi_3 \cdots \phi_n + \phi_2[\phi_1^-, \phi_3]\phi_4 \cdots \phi_n + \cdots \end{aligned} \tag{A.6}$$

Now, notice that

$$[\phi_1^-, \phi_k] = \langle 0|[\phi_1^-, \phi_k]|0\rangle \cdot \mathbb{1} = \langle 0|\phi_1^-\phi_k|0\rangle \cdot \mathbb{1} = \overbrace{\phi_1^- \phi_k} \tag{A.7}$$

Therefore

$$\begin{aligned} T\phi_1^-\phi_2\cdots\phi_n =&\phi_2\cdots\phi_n\phi_1^- + \overbrace{\phi_1^-\phi_2}\phi_3\cdots\phi_n + \cdots + \overbrace{\phi_1^-\phi_n}\phi_2\cdots\phi_{n-1} \\ =&T\{\phi_2\cdots\phi_n\}\phi_1^- + \overbrace{\phi_1^-\phi_2}T\{\phi_3\cdots\phi_n\} + \cdots + \overbrace{\phi_1^-\phi_n}T\{\phi_2\cdots\phi_{n-1}\} \end{aligned} \tag{A.8}$$

Now applying Wick theorem for $n-1$ fields, we get

$$T\phi_1^-\phi_2\cdots\phi_n =: \phi_2\cdots\phi_n :\phi_1^- + \sum_{\substack{\text{all possible}\\ \text{chronological}\\ \text{contractions}}} : \phi_2\cdots\phi_n :\phi_1^- + \sum_{\substack{\text{all possible}\\ \text{chronological}\\ \text{contractions}\\ \text{involving } \phi_1^-}} : \phi_2\cdots\phi_n\phi_1^- : \tag{A.9}$$

and using that $: \phi_2\cdots\phi_n :\phi_1^- =: \phi_2\cdots\phi_n\phi_1^- :=: \phi_1^-\phi_2\cdots\phi_n :$, we get

$$T\phi_1^-\phi_2\cdots\phi_n =: \phi_1^-\phi_2\cdots\phi_n : + \sum_{\substack{\text{all possible}\\ \text{chronological}\\ \text{contractions}}} : \phi_1^-\phi_2\cdots\phi_n : \tag{A.10}$$

Adding up the contribution from ϕ_1^+ and ϕ_1^-, we get Wick's theorem for n fields:

$$T\phi_1\phi_2\cdots\phi_n =: \phi_1\phi_2\cdots\phi_n : + \sum_{\substack{\text{all possible}\\ \text{chronological}\\ \text{contractions}}} : \phi_1\phi_2\cdots\phi_n : \tag{A.11}$$

concluding the proof by induction.

In the fermionic case, the proof follows the same method, keeping track of the sign from the anticommutation of the operator, as exemplified for n=2:

$$\begin{aligned} T\psi_1(x_1)\psi_2(x_2) =&\theta(x_1^0-x_2^0)\psi_1(x_1)\psi_2(x_2) - \theta(x_2^0-x_1^0)\psi_2(x_2)\psi_1(x_1) \\ =&\theta(x_1^0-x_2^0)\,[: \psi_1(x_1)\psi_2(x_2) : + \langle 0|\psi_1(x_1)\psi_2(x_2)|0\rangle\cdot\mathbb{1}] \\ &-\theta(x_2^0-x_1^0)\,[- : \psi_1(x_2)\psi_2(x_1) : + \langle 0|\psi_2(x_2)\psi_1(x_1)|0\rangle\cdot\mathbb{1}] \\ =&: \psi_1(x_1)\psi_2(x_2) : + \langle 0|T\psi_1(x_1)\psi_2(x_2)|0\rangle\cdot\mathbb{1} \\ =&: \psi_1(x_1)\psi_2(x_2) : + : \psi_1\overbrace{(x_1)\psi_2(x_2)} : \end{aligned} \tag{A.12}$$

To prove the theorem for n fields using the theorem for (n-1) fields we then again split the fermionic field $\psi_1 = \psi_1^+ + \psi_1^-$. For ψ_1^+, the proof does not feature any differences than in the bosonic case. For ψ_1^-, we this time anti-commute it with all the other fields:

$$\begin{aligned} T\psi_1^-\psi_2\cdots\psi_n =&(-1)^{n-1}\psi_2\cdots\psi_n\psi_1^- + \{\psi_1^-,\psi_2\}\psi_3\cdots\psi_n - \psi_2\{\psi_1^-,\psi_3\}\psi_4\cdots\psi_n + \cdots \\ =&(-1)^{n-1}\psi_2\cdots\psi_n\psi_1^- + \overbrace{\psi_1^-\psi_2}\psi_3\cdots\psi_n - \overbrace{\psi_1^-\psi_3}\psi_2\psi_4\cdots\psi_n + \cdots \end{aligned} \tag{A.13}$$

As in the bosonic case, we then apply Wick theorem for n-1 fields on the first term, giving the following final result

$$T\psi_1\psi_2\cdots\psi_n =: \psi_1\psi_2\cdots\psi_n : + \sum_{\substack{\text{all possible}\\ \text{chronological}\\ \text{contractions}}} : \psi_1\psi_2\cdots\psi_n : \tag{A.14}$$

where all the minus signs are hidden in the definition of chronological contractions.

$$: \psi_1\cdots\psi_{i-1}\overbrace{\psi_i\psi_{i+1}\cdots\psi_{j-1}}\psi_j\psi_{j+1}\cdots\psi_n : \; = (-1)^{n_{\pi_{ff}}}\overbrace{\psi_i\psi_j} : \psi_1\cdots\psi_{i-1}\psi_{i+1}\cdots\psi_{j-1}\psi_{j+1}\cdots\psi_n : \tag{A.15}$$

where $n_{\pi_{ff}}$ is the order of the permutation needed to move ψ_i and ψ_j from their initial to their final position.

The proof of the second Wick theorem follows the same procedure, but dropping the time ordering and replacing chronological contractions with standard contractions.

A.2 Third Wick Theorem

We want to show that

$$: \phi_{a_1} \cdots \phi_{a_n} :: \phi_{b_1} \cdots \phi_{b_m} : \; =: \phi_{a_1} \cdots \phi_{a_n} \phi_{b_1} \cdots \phi_{b_m} : + \sum_{a,b} : \phi_{a_1} \cdots \underbrace{\phi_{a_n} \phi_{b_1}} \cdots \phi_{b_m} : \;, \tag{A.16}$$

where the sum runs over all the contractions (meaning also multiple contractions) between some ϕ_a's and some ϕ_b's (note that no contractions between ϕ_a's and ϕ_a's or ϕ_b's and ϕ_b's appear).
Let's show this by induction.
We consider as the step 0 the one in which $n = m = 1$ (for either $n = 0$ or $m = 0$ the statement is trivially true):

$$: \phi_a :: \phi_b : \; = \; \phi_a^+ \phi_b^+ + \phi_a^+ \phi_b^- + \phi_a^- \phi_b^+ + \phi_a^- \phi_b^- \; = \; : \phi_a \phi_b : + : \underbrace{\phi_a \phi_b} : . \tag{A.17}$$

We recall the following properties

$$\underbrace{\phi_a^+ \phi_b} \; = \; 0, \tag{A.18}$$

$$\underbrace{\phi_a^- \phi_b} \; = \; \underbrace{\phi_a \phi_b}. \tag{A.19}$$

To complete the proof, we now suppose the theorem to hold for $n - 1 + m$ fields, and want to induce its validity for $n + m$. So we take as true

$$: \phi_{a_2} \cdots \phi_{a_n} :: \phi_{b_1} \cdots \phi_{b_m} : \; = \; : \phi_{a_2} \cdots \phi_{a_n} \phi_{b_1} \cdots \phi_{b_m} : + \sum_{a_{>1},b} : \phi_{a_2} \cdots \underbrace{\phi_{a_n} \phi_{b_1}} \cdots \phi_{b_m} : \quad . \tag{A.20}$$

Now we can write

$$\begin{aligned} : \phi_{a_1} \cdots \phi_{a_n} :: \phi_{b_1} \cdots \phi_{b_m} : &= \phi_{a_1}^+ : \phi_{a_2} \cdots \phi_{a_n} :: \phi_{b_1} \cdots \phi_{b_m} : + : \phi_{a_2} \cdots \phi_{a_n} : \phi_{a_1}^- : \phi_{b_1} \cdots \phi_{b_m} : \\ &= \phi_{a_1}^+ : \phi_{a_2} \cdots \phi_{a_n} :: \phi_{b_1} \cdots \phi_{b_m} : + : \phi_{a_2} \cdots \phi_{a_n} :: \phi_{b_1} \cdots \phi_{b_m} : \phi_{a_1}^- \\ &+ \sum_{a_1,b_i} : \phi_{a_2} \cdots \phi_{a_n} :: \phi_{b_1} \cdots \underbrace{\phi_{a_1} \phi_{b_i}} \cdots \phi_{b_m} : \quad . \end{aligned} \tag{A.21}$$

In the first line of the previous equation we have used the relation

$$: \phi_{a_1} \cdots \phi_{a_n} : \; = \; \phi_{a_1}^+ : \phi_{a_2} \cdots \phi_{a_n} : + : \phi_{a_2} \cdots \phi_{a_n} : \phi_{a_1}^- . \tag{A.22}$$

This can be deduced by simply rewriting

$$: \phi_{a_1} \cdots \phi_{a_n} : \; = \; : \phi_{a_1}^+ \cdots \phi_{a_n} : + : \phi_{a_1}^- \cdots \phi_{a_n} :, \tag{A.23}$$

and considering that the the field $\phi_{a_1}^+$ is already on the far left, so the normal ordering has no effect on it; as for the field $\phi_{a_1}^-$, the normal ordering will bring it always at the right of all the $\phi_{a_i}^+$'s, and since it commutes with all the $\phi_{a_i}^-$'s, this means that in every term of the expansion of $: \phi_{a_1}^- \cdots \phi_{a_n} :$ in fields $\phi_{a_i}^+$ and $\phi_{a_i}^-$ it can be commuted and appear on the far right. These two considerations, combined together, give the relation above. Another way of proving this relation is by means of Wick 3 applied to n fields (this is possible since we have supposed it to be valid up to $n - 1 + m > n$ fields).

Coming back to the decomposition for $n + m$ fields, with the relation for $n - 1 + m$ fields one rewrites it as

$$\begin{aligned} : \phi_{a_1} \cdots \phi_{a_n} :: \phi_{b_1} \cdots \phi_{b_m} : &= \phi_{a_1}^+ : \phi_{a_2} \cdots \phi_{a_n} \phi_{b_1} \cdots \phi_{b_m} : + \sum_{a_{>1},b} \phi_{a_1}^+ : \phi_{a_2} \underbrace{\cdots \phi_{a_n} \phi_{b_1}} \cdots \phi_{b_m} : \\ &+ : \phi_{a_2} \cdots \phi_{a_n} \phi_{b_1} \cdots \phi_{b_m} : \phi_{a_1}^- + \sum_{a_{>1},b} : \phi_{a_2} \underbrace{\cdots \phi_{a_n} \phi_{b_1}} \cdots \phi_{b_m} : \phi_{a_1}^- \\ &+ \sum_{a_1,b_i} : \phi_{a_2} \cdots \phi_{a_n} :: \phi_{b_1} \cdots \underbrace{\phi_{a_1} \phi_{b_i}} \cdots \phi_{b_m} : \end{aligned} \tag{A.24}$$

The fifth term can be split in two contributions:

$$\begin{aligned} \sum_{a_1,b_i} : \phi_{a_2} \cdots \phi_{a_n} :: \phi_{b_1} \cdots \underbrace{\phi_{a_1} \phi_{b_i}} \cdots \phi_{b_m} : \; &= \sum_{a_1,b_i} : \phi_{a_2} \cdots \phi_{a_n} \phi_{b_1} \cdots \underbrace{\phi_{a_1} \phi_{b_i}} \cdots \phi_{b_m} : \\ &+ \sum_{a_1,b_i} \sum_{a_{>1},b} : \phi_{a_2} \cdots \underbrace{\phi_{a_n} \phi_{b_1}} \cdots \cdots \underbrace{\phi_{a_1} \phi_{b_i}} \cdots \phi_{b_m} : \;, \end{aligned} \tag{A.25}$$

so that the overall expression is

$$
\begin{aligned}
:\phi_{a_1}\cdots\phi_{a_n}::\phi_{b_1}\cdots\phi_{b_m}: &= \phi^+_{a_1}:\phi_{a_2}\cdots\phi_{a_n}\phi_{b_1}\cdots\phi_{b_m}: + \sum_{a>1,b}\phi^+_{a_1}:\phi_{a_2}\underbrace{\cdots\phi_{a_n}\phi_{b_1}}\cdots\phi_{b_m}: \\
&+ :\phi_{a_2}\cdots\phi_{a_n}\phi_{b_1}\cdots\phi_{b_m}:\phi^-_{a_1} + \sum_{a>1,b}:\phi_{a_2}\underbrace{\cdots\phi_{a_n}\phi_{b_1}}\cdots\phi_{b_m}:\phi^-_{a_1} \\
&+ \sum_{a_1,b_i}:\phi_{a_2}\cdots\phi_{a_n}\phi_{b_1}\cdots\underbrace{\phi_{a_1}\phi_{b_i}}\cdots\phi_{b_m}: \\
&+ \sum_{a_1,b_i}\sum_{a>1,b}:\phi_{a_2}\underbrace{\cdots\phi_{a_n}\phi_{b_1}}\cdots\cdots\underbrace{\phi_{a_1}\phi_{b_i}}\cdots\phi_{b_m}: \quad .
\end{aligned}
\tag{A.26}
$$

The sum of the first and the third term is:

$$
:\phi_{a_1}\cdots\phi_{b_n}: \tag{A.27}
$$

The sum of the second and fourth term is

$$
\sum_{a,b}:\phi_{a_1}\cdots\underbrace{\phi_{a_n}\phi_{b_1}}\cdots\phi_{b_m}: \tag{A.28}
$$

where a_1 is not involved. The sum of the last two terms is:

$$
\sum_{a,b}:\phi_{a_1}\cdots\underbrace{\phi_{a_n}\phi_{b_1}}\cdots\phi_{b_m}: \tag{A.29}
$$

where a_1 is involved. Thus the sum of the second, fourth, fifth and sixth term is:

$$
\sum_{a,b}:\phi_{a_1}\cdots\underbrace{\phi_{a_n}\phi_{b_1}}\cdots\phi_{b_m}: \tag{A.30}
$$

as defined in equation (A.16).

Since step 0 is true and step $n-1+m$ implies step $n+m$, then the induction is complete, and the theorem (A.16) is proved. □