Notes on QED

Version of December 17, 2020

Lectures
University of Pavia

Alessandro Bacchetta

"No sensible man would insist that these things are as I have described them."

Plato, *Dialogues, Phaedo*

Contents

1

Introduction

This lecture notes are for the Quantum Electrodynamics course of the University of Pavia. This course represents the first introduction to the topic of Quantum Field Theory. It is a one-semester course and it is meant to be followed by a second course on Quantum Field Theory.

Why do we need a Quantum Field Theory? The goal is to have a theory that is compatible both with Special Relativity and Quantum Mechanics, i.e., a theory for "Relativistic Quantum Mechanics." It turns out that this requires the use of a field theory. In such a theory, the "particles" are identified with oscillations of a field.

Classical nonrel. particles	Classical rel. particles
Quantum nonrel. particles	~~Quantum rel. particles~~

Classical nonrel. fields	Classical rel. fields
Quantum nonrel. fields	**Quantum rel. fields**

The "problem" of merging Quantum Mechanics and Special Relativity can be intuitively seen in the following way. Consider a state with a measurable lifetime $\Delta\tau$ and a measurable energy. The standard deviation of the energy of the state can be related to its

lifetime via Heisenberg principle.[1]

$$\Delta E \Delta \tau \geq \hbar/2. \tag{1.1}$$

This relation can be interpreted by saying that a state with extremely short lifetime can have an extremely large energy dispersion. For time intervals of the order of $\hbar/(4m_e c^2) \approx 10^{-22}$ s, the energy uncertainty can be as large as the rest energy of an electron and a positron. This opens up the possibility that a state composed by an electron and a positron is created out of nothing, as long as it decays in a short-enough time. Thus, the vacuum is not really empty. Most of this relentless activity in the vacuum has no measurable effect. However, it has to be taken into account when we carry out high-precision experiments at subatomic level.

Nowadays, it is common wisdom that particles can be created out of energy. By colliding an electron and a positron at high energies, we can produce all sorts of other particles, including Higgs bosons. However, it is not possible to describe this process within particle mechanics, because in this case particles are considered to be there in the first place, they cannot be created nor destroyed. A similar problem occurs with spontaneous emission of photons from atoms. Richard Feynman told this story about this problem:

> My father once asked me: "I understand that they say that light is emitted from an atom when it goes from one state to another, from an excited state to a state of lower energy."
> I said, "That's right."
> "And light is a kind of particle, a photon, I think they call it."
> "Yes."
> "So if the photon comes out of the atom when it goes from the excited to the lower state, the photon must have been in the atom in the excited state."
> I said, "Well, no."
> He said, "Well, how do you look at it so you can think of a particle photon coming out without it having been in there in the excited state?"
> I thought a few minutes, and I said, "I'm sorry; I don't know. I can't explain it to you."
> He was very disappointed after all these years and years of trying to teach me something, that it came out with such poor results.

This anecdote tells you how difficult can be to discuss the creation and annihilation process with the language of "particles." The situation is more flexible with field oscillations: they can be created or suppressed, corresponding to the creation or destruction of particles. The electromagnetic field serves as a good example: we are not so suprised about the fact that electromagnetic radiation can be absorbed or created by atomic transitions. In fact, it turns out that the calculation of the coefficients of absorption, spontaneous

[1] Note that this is *not* a relation between energy and external time, but between energy and the lifetime of the state

emission and stimulated emission can be carried out within quantum field theory without resorting to any extra assumption (see discussion in Sec. 1.2 and 1.3 of [31]).

From nonrelativistic classical particle mechanics, we know how to obtain nonrelativistic quantum particle mechanics and also relativistic classical particle mechanics. We will try to build a theory of relativistic quantum particle mechanics, but encounter some problems. The first part of the course will be devoted to this.

We can turn for inspiration to the description of the ElectroMagnetic (EM) field. It is a classical field theory that is from the beginning relativistic. It turns out that it is possible to quantize this theory, thereby obtaining a theory which is relativistic and quantistic, containing field quanta that can be identified with "particles" (photons). When the EM field interacts with particles at the atomic level, but whose momenta are not large, it is still possible to treat the EM field as a quantized field and the particles as nonrelativistic quantum particles: this is the approach of Quantum Optics. But the ideas underpinning the quantization of EM fields can be applied also to fields of different types. Their field quanta can be identified with other particles. In this approach, particles are like ripples in fields that fill the whole space. The second part of the course will then be devoted to understanding relativistic classical field theories and how to quantize them.

The conceptual steps of field quantization are needed just to describe free particles, i.e., free fields. But we want of course also to be able to describe the interaction of different particles/fields. We are going to study scattering processes, where we assume that we start from an initial state consisting of free fields, the fields then undergo an interaction and end up into a final state again consisting of free fields. A classical analogy may be when a sound wave hits a membrane, it is partially reflected and partially absorbed giving origin to a vibration of the membrane.

In a scattering process we always start with some particle (i.e., some excitation of the fields) and end up with some other particles, after some interaction has occurred. The strength of the interaction is determined by the "coupling constant." In principle, the particles can interact (infinitely) many times, but if the coupling constant is small, we can expand the computation in a perturbative series and stop at a certain number of interactions. For instance, in QED we have the coupling constant (in rationalized Gaussian units, also called Lorentz-Heaviside units)

$$\alpha = \frac{e^2}{4\pi\hbar c} \approx \frac{1}{137}. \tag{1.2}$$

It makes then sense to stop the calculation at the lowest order in α, since the corrections will be of the order of 1%. For highly precise computations and/or for theories where the coupling constant is larger it is essential to go beyond.

The contributions to the scattering at a given order of the coupling constant can be conveniently described by the so-called "Feynman diagrams." For instance, if we want to compute the scattering of a photon and an electron (Compton scattering), we have to take into consideration diagrams of the type depicted in Fig. 1.1.

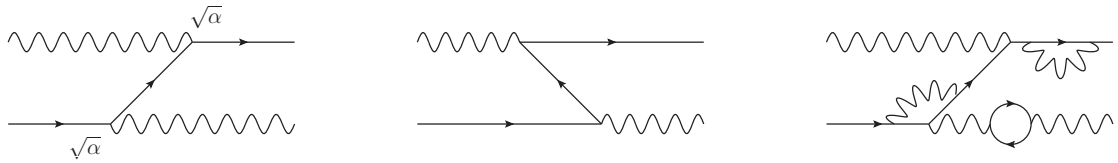


Figure 1.1: Feynman diagrams contributing to the process of photon-electron scattering (Compton scattering). The first two diagrams represent two different "time-ordered" contributions to the same Feynman diagram.

In this course, we will stop at the lowest order in the coupling constant. We will consider only the so-called "tree-level diagrams" as opposed to diagrams containing loops. At the end of the course, you should *in principle* have the tools to compute any tree-level diagram in QED and other theories.

For loop diagrams and all the fundamental issues related to them, you should attend the Quantum Field Theory course in the second semester.

QED is probably the theory that has been tested to the highest accuracy in experiments. For instance, the theoretical computation of the electron anomalous magnetic moment agrees with the experimental measurement at the 10^{-9} level. However, to reach this kind of precision QED calculations have to be carried out to order α^4, involving 891 Feynman diagrams [24] (calculations reaching order α^5 have also been performed [4]). By the end of the Quantum Field Theory course, you will be able to compute the order α correction and reach an agreement with the experimental value of within a few percents [35].

$$a_e\big(\mathcal{O}(\alpha)\big) = 0.00116 \tag{1.3}$$

$$a_e\big(\mathcal{O}(\alpha^4)\big) = 0.0011596521817(8) \tag{1.4}$$

$$a_e|_{\text{exp}} = 0.00115965218091(26) \tag{1.5}$$

Most of the applications of Quantum Field Theory have to do with Elementary Particle Physics, i.e., the study of what we at present consider the fundamental constituents of matter. It is ironic that what we call "Particle Physics" is based on the notion that particles do not exist, but have to be replaced by field excitations! Carlo Rovelli uses these inspiring words to describe the world of quantum field theory, in his book "Seven brief lessons on Physics"

> Just as the calmest sea looked at closely sways and trembles, however slightly, so the fields that form the world are subject to minute fluctuations.
>
> [...]
>
> We have arrived very far from the mechanical world of Newton and Laplace, where minute cold stones eternally travelled on long precise trajectories in geometrically immutable space. Quantum mechanics and experiments with particles have taught us that the world is a continuous, restless swarming of things; a continuous coming to light and disappearance of ephemeral entities. A set of vibrations, as in the switched-on hippy world of the 1960s. A world of happenings, not of things.

However, the ideas of field quantization can be used also to study nonrelativistic quantum many-body systems, which are extremely important also in Solid State Physics.

1.1 Recommended resources

As with all courses at the M.Sc. level, you are encouraged to look at more than one reference to gain a deeper knowledge of the subject.

- These lecture notes are largely based on older lecture notes by prof. F. Miglietta (available in the library).
- The calculations of tree-level QED processes are presented in detail in the lecture notes of prof. F. Piccinini (available in the library).
- The course basically covers the first eight chapters and App. A of "Quantum Field Theory" by Mandl and Shaw [26].
- The first part of the course is treated in detail in Ch. 2, 3, 4 of the book by Ryder [29].
- Two recent textbooks on the topic are the ones by Maggiore [25] (more concise) and Schwartz [31] (more extended).
- The book by Aitchison and Hey [3], Vol. 1, presents similar topics with a pedagogical and accessible style.
- The book by Peskin and Schroeder [28] has become a "classic" but it is more advanced and more phenomenology-oriented than the previous ones. In any case, the first five chapters of the book roughly correspond to what is covered in the present course.
- The book by Halzen-Martin [21] is lower-level than the previous ones, but it may be useful to understand the concepts without diving too deep into the theory.
- A classic book, a "bible", is that by Weinberg [33], difficult but very complete.
- A well-documented history of the birth of Quantum Field Theory can be found in the book "Inward Bound" by A. Pais [27], in particular Ch. 15.

2

Relativistic wave equations

Qunatum Field Theory textbooks often skip the discussion of relativistic single-particle equations and start straight from the language of field theory. The discussion is important not only historically, but also to understand some features of relativistic quantum theories. It is also the occasion to introduce some technicalities that are anyway needed in the following chapters. Apart from the lecture notes of prof. Miglietta, the topics covered in this chapter are treated in a very concise way in App. A of Mandl–Shaw [26] and in Ch. 3 of Peskin–Schroeder [28]. A more extended treatment can be found in Ch. 2 of Ryder [29].

2.1 The Klein–Gordon equation

The Schrödinger equation is one of the "axioms" of Quantum Mechanics and reads

$$i\hbar\partial_t\psi(t,\vec{x}) = H\psi(t,\vec{x}), \tag{2.1}$$

where we used the notation $\partial_t = \frac{\partial}{\partial t}$.

For a free particle of mass m, we have the nonrelativistic relation

$$E = \frac{\vec{p}^2}{2m} \tag{2.2}$$

We can then obtain the Schrödinger equation by the replacement

$$E \to i\hbar\partial_t \qquad\qquad \vec{p} \to -i\hbar\vec{\nabla} \tag{2.3}$$

leading to

$$i\hbar\partial_t\psi(t,\vec{x}) = -\frac{\hbar^2\vec{\nabla}^2}{2m}\psi(t,\vec{x}). \tag{2.4}$$

An "obvious" guess to go from a nonrelativistic to relativistic description is to start from the relation

$$E^2 = m^2c^4 + \vec{p}^2c^2 \tag{2.5}$$

The nonrelativistic case is recovered in the limit $c \to \infty$

$$E = mc^2\sqrt{1+\frac{\vec{p}^2}{m^2c^2}} \overset{c\to\infty}{\approx} mc^2 + \frac{\vec{p}^2}{2m}. \tag{2.6}$$

Now we use the replacements (2.3)

$$\begin{aligned} -\partial_t^2\phi(t,\vec{x}) &= \left(\frac{m^2c^4}{\hbar^2} - \vec{\nabla}^2c^2\right)\phi(t,\vec{x}) \\ \left(\frac{\partial_t^2}{c^2} - \vec{\nabla}^2\right)\phi(t,\vec{x}) &= -\frac{m^2c^2}{\hbar^2}\phi(t,\vec{x}) \end{aligned} \tag{2.7}$$

and we finally obtain the Klein–Gordon equation

$$\boxed{\left(\partial^\mu\partial_\mu + \mu^2\right)\phi(x) = 0} \tag{2.8}$$

where we made use of relativistic notations and introduced $\mu = mc/\hbar$. The function is called Klein–Gordon, even though at least six authors (including Schrödinger) independently stated it in 1926 [27].

2.1.1 Excursus: relativistic notation

A concise discussion of relativistic notation can be found in Sec. 2.1 of Mandl–Shaw [26].

Lorentz transformations include rotations and boosts in all directions. Recall that the effect of a Lorentz boost in the x direction can be written as

$$ct' = \gamma(ct + \beta\, x) \tag{2.9}$$
$$x' = \gamma(x + \beta\, ct) \tag{2.10}$$
$$y' = y \tag{2.11}$$
$$z' = z \tag{2.12}$$

where $\gamma = 1/\sqrt{1-\beta^2}$ and $\beta = v/c$.

A more convenient notation for our purposes is to write the Lorentz boost in matrix form

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cosh\epsilon & \sinh\epsilon & 0 & 0 \\ \sinh\epsilon & \cosh\epsilon & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{2.13}$$

where we introduced the rapidity $\epsilon = \ln[\gamma(1+\beta)] = \ln\sqrt{(1+v/c)/(1-v/c)}$. Note that the inverse transformation requires only a sign change of velocity or, equivalently, a sign change of rapidity.

A Lorentz invariant is

$$c^2t^2 - x^2 - y^2 - z^2 = c^2t'^2 - x'^2 - y'^2 - z'^2. \tag{2.14}$$

The best way to formally deal with Relativity is to introduce a space of four-vectors (Minkowski space) with a definition of a scalar product (a metric) that is invariant under Lorentz transformations. To do this, we need to introduce two types of vectors, connected by a metric tensor. A (contravariant) four-vector in Minkowski space is denoted as

$$a^\mu = (a^0, a^1, a^2, a^3) = (a^0, \vec{a}) \tag{2.15}$$

The coordinate vector in c.g.s. units would be

$$x^\mu = (ct, x, y, z)|_{\text{c.g.s.}} \tag{2.16}$$

In natural units (see next section) it is simply $x^\mu = (t, x, y, z)$.

The metric tensor is

$$g^{\mu\nu} = g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.17}$$

Note that

$$g^{\mu\lambda}g_{\lambda\nu} = \delta^\mu_\nu. \tag{2.18}$$

Covariant vectors are

$$a_\mu = g_{\mu\nu}a^\nu = (a^0, -a^1, -a^2, -a^3) = (a^0, -\vec{a}), \tag{2.19}$$

Scalar products can then be defined as

$$a \cdot b = a^\mu b_\nu = a^0 b^0 - \vec{a}\cdot\vec{b}. \tag{2.20}$$

Lorentz transformations will be denoted by $L^\mu{}_\nu$. The effect of a Lorentz transformation is to modify the vectors in this way

$$a^\mu \longrightarrow a'^\mu = L^\mu{}_\nu a^\nu. \tag{2.21}$$

By definition, Lorentz transformations leave scalar products unchanged

$$\begin{aligned} a^\mu a_\mu = a'^\mu a'_\mu &= L^\mu{}_\nu a^\nu L_\mu{}^\sigma a_\sigma \\ &\Rightarrow L^\mu{}_\nu L_\mu{}^\sigma = \delta^\sigma_\nu \\ &\Rightarrow L^{\mu\nu}L_{\mu\sigma} = \delta^\nu_\sigma \end{aligned} \tag{2.22}$$

Any four-component object that transforms like a^μ under a Lorentz transformation is a Lorentz four-vector. Apart from the example of the coordinate vector, we can also consider, e.g.,

$$p^\mu = (E/c, p_x, p_y, p_z), \qquad k^\mu = (\omega/c, k_x, k_y, k_z), \tag{2.23}$$

Important: the four-dimensional "nabla" or "del" operator is defined as

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right) = (\frac{\partial_t}{c}, \vec{\nabla}). \tag{2.24}$$

Even if it is a covariant vector, the relation with the three-dimensional nabla operator is different from that of a normal vector. This guarantees that the differential of a scalar function $\phi(x)$ is a scalar

$$\delta\phi(x) = \partial_t\phi(t,\vec{x})\delta t + \vec{\nabla}\phi(t,\vec{x})\cdot\delta\vec{x} = \partial_\mu\phi(x)\delta x^\mu. \tag{2.25}$$

A useful operator is the so-called d'Alembertian

$$\partial^\mu\partial_\mu = \frac{\partial_t^2}{c^2} - \vec{\nabla}^2 \tag{2.26}$$

often denoted by $\Box$, which generalizes the three-dimensional Laplacian operator.

2.1.2 Excursus: natural units

A nice discussion is presented in Sec. 6.1 of Mandl–Shaw [26]. When dealing with Quantum Mechanics and Relativity the fundamental quantities we need are

$$\hbar \approx 1.05 \times 10^{-34}\ \text{kg m}^2/\text{s}, \qquad c \approx 3.00 \times 10^8\ \text{m/s}. \tag{2.27}$$

Natural units are defined by the choice

$$\hbar = c = 1. \tag{2.28}$$

In natural units, all quantities can be expressed in terms of a power of M or, equivalently, of energy. Connecting natural-unit expressions to c.g.s. units[1] is not difficult, it is sufficient to multiply by the appropriate combinations of $\hbar$ and c factors.

[1]With c.g.s. we mean rationalized Gaussian units, or Lorentz–Heaviside units, where $\epsilon_0 = 1$. Please note that several equations, including Maxwell's equations, may look different in different systems. For instance, in the c.g.s. system we have

$$\vec{\nabla}\cdot\vec{E} = \rho, \qquad \vec{E} = -\vec{\nabla}\Phi - \frac{1}{c}\frac{\partial\vec{A}}{\partial t}, \qquad \vec{F}_{\text{Lorentz}} = \frac{q}{c}\vec{v}\times\vec{B} \tag{2.29}$$

For instance, in c.g.s. the proton weights $m_P = 1.67 \times 10^{-27}$ kg. In natural units, we can choose to express it as an energy. Since the units of energy are J = kg m^2/s^2 we need to multiply the value by a combination of factors that has dimension m^2/s^2, i.e., c^2.

$$m_P = 1.67 \times 10^{-27}\ \text{kg}\Big|_{\text{c.g.s.}} \longrightarrow 1.67 \times 10^{-27}\ \text{c}^2\ \text{kg}\Big|_{\text{n.u.}} = 1.50 \times 10^{-10}\ \text{J}\Big|_{\text{n.u.}} \tag{2.30}$$

For subnuclear physics, the preferred unit for energy is GeV. The conversion between Joules and GeV is

$$1\ \text{GeV} = 10^9\ \text{eV} = 10^9 \times 1.60 \times 10^{-19}\ \text{C V} = 1.60 \times 10^{-10}\ \text{J}. \tag{2.31}$$

From this we conclude that

$$m_P = 0.938\ \text{GeV}\Big|_{\text{n.u.}}. \tag{2.32}$$

A second and a meter in c.g.s. units can be expressed in terms of energy in n.u.

$$1\,\text{s}\Big|_{\text{c.g.s.}} \longrightarrow \frac{1\text{s}}{\hbar}\Big|_{\text{n.u.}} = 0.952 \times 10^{34}\ \text{J}^{-1}\Big|_{\text{n.u.}} = 1.52 \times 10^{24}\ \text{GeV}^{-1}\Big|_{\text{n.u.}}, \tag{2.33}$$

$$1\,\text{m}\Big|_{\text{c.g.s.}} \longrightarrow \frac{1\text{m}}{\hbar c}\Big|_{\text{n.u.}} = 0.317 \times 10^{26}\ \text{J}^{-1}\Big|_{\text{n.u.}} = 0.507 \times 10^{16}\ \text{GeV}^{-1}\Big|_{\text{n.u.}}. \tag{2.34}$$

This means that in n.u. time and space have the dimensions of the inverse of an energy (or the inverse of a mass). Note that in natural units the typical size of a nucleon is 1 fm = 10^{-15} m $\approx 1/(200\ \text{MeV})$.

The electron charge squared in the SI units is equal to 2.56×10^{-38} C^2. In c.g.s. units it has the dimensions of kg m^3/s^2. To convert to c.g.s. units we use the value of the dielectric constant $\varepsilon_0 = 8.85 \times 10^{-12}$ C^{2}s^2/(kg m^3). We obtain in the c.g.s. $e^2 = 2.89 \times 10^{-25}$ kg m^3/s^2.

The fine structure constant has no dimensions in c.g.s. units and therefore has no dimensions also in natural units. The definition is

$$\alpha = \frac{e^2}{4\pi\hbar c}\Big|_{\text{c.g.s.}} = \frac{e^2}{4\pi}\Big|_{\text{n.u.}} \approx \frac{1}{137} \tag{2.35}$$

In natural units, the typical size of an atom can be written in terms of Bohr's radius $r_B \approx 0.5 \times 10^{-10}$ m $= 1/(4\ \text{KeV})$ and the corresponding potential energy is $\alpha/r_B \approx 30$ eV, which is the order of magnitude of the energies involved in atomic phenomena.

As another exercise, we can consider cross sections. In natural units they are given often in GeV^{-2}. In c.g.s. units they have the dimension of a surface, therefore they should be given in m^2 (or more commonly in barn$= 10^{-28}$ m^2). The conversion is

$$\begin{aligned}\sigma = N\ \text{GeV}^{-2}\big|_{\text{n.u.}} \longrightarrow N\ \text{GeV}^{-2}(\hbar c)^2\big|_{\text{c.g.s.}} &= N\ \text{GeV}^{-2}(1.973 \times 10^{-16})^2\ \text{GeV}^2\text{m}^2\big|_{\text{c.g.s.}} \\ &= N\ 0.389\ \text{mbarn}\big|_{\text{c.g.s.}}\end{aligned} \tag{2.36}$$

2.1.3 Covariance of the Klein-Gordon equation

We define as "Lorentz invariant" a quantity that does not change under Lorentz transformations, e.g., m, $a_\mu a^\mu$, $F_{\mu\nu}F^{\mu\nu}$... Lorentz invariants can be constructed not only with four-vectors, but also with spinors, as we shall see later. All Lorentz scalars are Lorentz invariants[2]. A scalar quantity (i.e., a quantity expressed by a number) is not necessarily a Lorentz scalar and therefore not necessarily Lorentz invariant: for instance, energy is a scalar quantity, but is is a component of a four vector and changes under Lorentz transformations.

We use the term "Lorentz covariance" to indicate that a certain relation has the same form under Lorentz transformations (i.e., it is the same in all inertial reference frames). In other words, two observers in two different inertial frames should agree on the validity of the relation. It would be more meaningful to call it "form invariance." A relation involving Lorentz scalars is (manifestly) Lorentz covariant (e.g., $m^2 = p_\mu p^\mu$). The same relation written in terms of components ($E^2 = m^2 + \vec{p}^2$) is covariant, but not manifestly so. The nonrelativistic relation for energy $E = \vec{p}^2/(2m)$ is *not* Lorentz covariant (it is covariant under rotations only).

In the Klein–Gordon equation, we have m which is clearly a Lorentz invariant, and we have the operator $\partial_\mu \partial^\mu$, which is also a Lorentz invariant since

$$\partial^\mu \partial_\mu \to \partial'^\mu \partial'_\mu = L^\mu{}_\nu \partial^\nu L_\mu{}^\sigma \partial_\sigma = \partial^\mu \partial_\mu \tag{2.37}$$

For the Klein–Gordon equation to be Lorentz covariant, we need the function ϕ to be a "scalar function", i.e.,

$$\phi'(x') = \phi(x). \tag{2.38}$$

For instance, a function like $f(x) = k_\mu x^\mu$ is a scalar function: if we apply a Lorentz transformation, the four-vector x^μ will change into $x'^\mu = L^\mu_\nu x^\nu$, but its change will be compensated by a change of the four-vector k_μ. The electrostatic potential $V(t, \vec{x}) = q/(4\pi|\vec{x}|)$ is *not* a Lorentz scalar field.

2.1.4 Nonrelativistic limit

We want to study the limit for $c \to \infty$. We obviously need to write things in the c.g.s. system. We first rewrite

$$\phi = e^{-i\frac{mc^2}{\hbar}t}\phi' \tag{2.39}$$

which serves the purpose of isolating the rest energy mc^2.

[2]Pseudoscalars are also invariants, but they change sign under parity

We need to use

$$\begin{aligned}\partial_t^2\left(e^{-i\frac{mc^2}{\hbar}t}\phi'\right) &= \partial_t\left(-\frac{imc^2}{\hbar}e^{-i\frac{mc^2}{\hbar}t}\phi' + e^{-i\frac{mc^2}{\hbar}t}\partial_t\phi'\right)\\ &= \left(-\frac{m^2c^4}{\hbar^2}e^{-i\frac{mc^2}{\hbar}t}\phi' - \frac{imc^2}{\hbar}e^{-i\frac{mc^2}{\hbar}t}\partial_t\phi' - \frac{imc^2}{\hbar}e^{-i\frac{mc^2}{\hbar}t}\partial_t\phi' + e^{-i\frac{mc^2}{\hbar}t}\partial_t^2\phi'\right)\end{aligned} \tag{2.40}$$

Starting from the Klein–Gordon equation we can perform the following steps:

$$\begin{aligned}0 &= \left(\frac{\partial_t^2}{c^2} - \vec{\nabla}^2 + \frac{m^2c^2}{\hbar^2}\right)e^{-i\frac{mc^2}{\hbar}t}\phi'\\ &= e^{-i\frac{mc^2}{\hbar}t}\left(\frac{\partial_t^2}{c^2} - \frac{2im}{\hbar}\partial_t - \cancel{\frac{m^2c^2}{\hbar^2}} - \vec{\nabla}^2 + \cancel{\frac{m^2c^2}{\hbar^2}}\right)\phi'\\ &\approx -\frac{2m}{\hbar^2}e^{-i\frac{mc^2}{\hbar}t}\left(i\hbar\partial_t + \frac{\hbar^2\vec{\nabla}^2}{2m}\right)\phi' = 0.\end{aligned} \tag{2.41}$$

which corresponds to Schrödinger equation. The solutions of the Schrödinger equation for a free particle are of the form

$$\psi(t,\vec{x}) = e^{-i\omega_k t + i\vec{k}\cdot\vec{x}} \tag{2.42}$$

with $\omega_k = \hbar\vec{k}^2/(2m)$.

2.1.5 Solutions of the Klein-Gordon equation

The starting equation is

$$\left(\partial^\mu\partial_\mu + m^2\right)\phi(x) = 0 \tag{2.43}$$

We seek a solution of the equation in the form of plane waves, i.e.,

$$\phi(x) = \phi(k)e^{-ik\cdot x} = \phi(k)e^{-ik_\mu x^\mu} = \phi(\omega,k)e^{-i\omega t + i\vec{k}\cdot\vec{x}} \tag{2.44}$$

The four-vector k in the c.g.s. system would be defined as

$$k^\mu = (\omega/c,\vec{k})|_{\text{c.g.s.}} \tag{2.45}$$

where ω is the frequency of the wave (units $1/\text{s}$) and $\vec{k}$ is the wave number (units $1/\text{m}$) . The four-momentum vector would be

$$p^\mu = (E/c,\vec{p})|_{\text{c.g.s.}} = (\omega\hbar/c,\hbar\vec{k})|_{\text{c.g.s.}} \tag{2.46}$$

where E is the energy (units kg m^2/s^2), p is the momentum (units kg m/s). In the natural units system, however, these two vectors are just the same. Talking about frequency or

energy of a plane wave is the same, and talking about wave-vector or momentum is the same too.

Plugging (2.44) into (2.43) we obtain

$$\begin{aligned}\partial^\mu\partial_\mu\phi(x) &= \phi(k)\partial^\mu\left(\frac{\partial}{\partial x^\mu}e^{-ik_\mu x^\mu}\right)\\ &= \phi(k)\partial^\mu\left(-ik_\mu\, e^{-ik_\mu x^\mu}\right)\\ &= -k^\mu k_\mu\, \phi(k)e^{-ik_\mu x^\mu}\end{aligned} \tag{2.47}$$

The Klein–Gordon equation in this case becomes

$$(k^2 - m^2)\phi(x) = 0 \tag{2.48}$$

which is solved for $k^2 = m^2$. When this condition is fulfilled, we say that the particle is "on mass shell" or simply "on shell." The on-shell condition implies that

$$\begin{aligned}\omega^2 &= \vec{k}^2 + m^2\\ \Rightarrow \omega &= \pm\omega_k = \pm\sqrt{\vec{k}^2 + m^2}\end{aligned} \tag{2.49}$$

We can check what this means in c.g.s. units.

$$\begin{aligned}\frac{\omega^2}{c^2} &= \vec{k}^2 + \frac{m^2c^2}{\hbar^2}\\ \hbar^2\omega^2 &= \hbar^2c^2\vec{k}^2 + m^2c^4\\ E^2 &= \vec{p}^2c^2 + m^2c^4\end{aligned} \tag{2.50}$$

which is the correct relativistic expression for the energy we started from. The above relations are also known as "dispersion relations," a term coming from wave mechanics.

Let us for a moment focus on the positive energy or positive frequency solutions. What is the "mass"? From the mathematical point of view, it is just a parameter in the wave equation. It is however connected to the energy and the momentum of the free wave in the same way as the mass of a particle. In a sense, this could be taken as a definition of what the mass is.

2.1.6 Problems with the Klein–Gordon equation

The result of the energy poses a first problem, namely that there are also solutions with a negative energy and their energy has no lower limit. This is due essentially to the fact that we started from a quadratic expression for E, which does not exclude negative-energy values.

It seems almost as there may exist two identical versions of the Klein–Gordon wave, with positive and negative energy respectively.

There is another problem with the Klein–Gordon equation. In Quantum Mechanics we can define the "probability density" and a "current density" in the following way

$$\rho(t,\vec{x}) = \psi^*(t,\vec{x})\psi(t,\vec{x}) \tag{2.51}$$

$$\vec{j}(t,\vec{x}) = -\frac{i\hbar}{2m}\Big(\psi^*(t,\vec{x})\vec{\nabla}\psi(t,\vec{x}) - \psi(t,\vec{x})\vec{\nabla}\psi^*(t,\vec{x})\Big) \tag{2.52}$$

Due to the validity of the Schrödinger equation, the two quantities are connected by the continuity equation

$$\partial_t\rho + \vec{\nabla}\cdot\vec{j} = 0 \tag{2.53}$$

which means also that the integral of the probability density over all space is constant (use divergence theorem)

$$\int d^3x\rho(t,\vec{x}) = \text{const.} \tag{2.54}$$

For the Klein–Gordon case, we can introduce a four-current

$$j^\mu(x) = \frac{i}{2m}\Big(\phi^*(x)\partial^\mu\phi(x) - \phi(x)\partial^\mu\phi^*(x)\Big). \tag{2.55}$$

Indeed, this four-current satisfies a continuity equation. Using the Klein–Gordon equation, we can check that (we drop the argument of the ϕ function for convenience)

$$\begin{aligned}
\partial_\mu j^\mu &= 0 \\
\partial_\mu j^\mu &= \frac{i}{2m}\Big(\big(\partial_\mu\phi^*\big)\partial^\mu\phi + \phi^*\partial_\mu\partial^\mu\phi \\
&\qquad - \big(\partial_\mu\phi\big)\partial^\mu\phi^* - \phi\partial_\mu\partial^\mu\phi^*\Big) \\
&= \frac{i}{2m}\Big(\phi^*\partial_\mu\partial^\mu\phi - \phi\partial_\mu\partial^\mu\phi^*\Big) \\
&= \frac{i}{2m}\Big(\phi^* m^2\phi - \phi m^2\phi^*\Big) = 0.
\end{aligned} \tag{2.56}$$

If we isolate the density contribution, i.e., the 0th component, we obtain

$$\rho(x) = \frac{i}{2m}\Big(\phi^*(x)\partial_t\phi(x) - \phi(x)\partial_t\phi^*(x)\Big). \tag{2.57}$$

If we apply this definition to the solution with a negative ω (i.e., $e^{i\omega_k t + i\vec{k}\cdot\vec{x}}$) we obtain

$$\rho(x) = \frac{i}{2m}2i\omega_k\phi^*(x)\phi(x) = -\frac{\omega_k}{m}\phi^*(x)\phi(x), \tag{2.58}$$

which is negative and is impossible to interpret as a probability density.

2.1.7 Coupling to electromagnetic field

In classical mechanics, the coupling of a charge particle to the electromagnetic field can can be obtained by the so-called "minimal substitution" (or "minimal coupling"), i.e.,

$$E \to E - q\Phi, \qquad \vec{p} \to \vec{p} - q\vec{A}, \tag{2.59}$$

where q is the charge of the particle, Φ is the scalar potential and $\vec{A}$ is the vector potential of the elm field

Analogously, in the case of quantum mechanics, we use the substitution

$$i\hbar\partial_t \to i\hbar\partial_t - q\Phi, \qquad -i\hbar\vec{\nabla} \to -i\hbar\vec{\nabla} - q\vec{A}, \tag{2.60}$$

In four-vector notation, the minimal substitution can be written in a remarkably simple way

$$\partial_\mu \to D_\mu = \partial_\mu + \frac{iq}{\hbar c}A_\mu \tag{2.61}$$

where $A^\mu = (\Phi, \vec{A})$ is the four-potential of the elm field.

It is possible to compute the solutions of the K–G equation in the presence of an electrostatic field an try to apply this approach to the study of the electron in the hydrogen atom. In the nonrelativistic limit, we would obtain the same results as for the Schrödinger equation. However, the relativistic corrections would lead to wrong results.

2.2 The Dirac equation

Dirac is certainly the most influential scientist in the development of QED. Nobel prize winner in 1933. One of the greatest theorists of the XX century. He gave contributions to Quantum Mechanics (the bra and ket notation, for instance) and to Mathematics (the Dirac delta and the theory of distributions). But his most important achievement was the equation that bears his name, published in 1928 [16] and engraved on his memorial stone in Westminster Abbey.

One of the problematic points of the Klein–Gordon equation is the fact that it is deduced from a relation involving E^2 and thus contains a second-order time derivative (together with second-order space derivatives). Dirac tried to look for an equation with a first-order time (and space) derivatives. He demanded that the solution of this equation were also solutions of the Klein–Gordon equation, which guarantees also the right dispersion relation (i.e., the right relation between energy and momentum, or frequency and wave number).

We start from the following expression of the Dirac equation

$$\left(i\gamma^\mu\partial_\mu - m\right)\psi(x) = 0 \tag{2.62}$$

For the moment, we don't know what the factors γ^μ mean.

We now multiply the equation to the left by a similar operator in order to get as close as possible to the Klein–Gordon equation

$$\Big(i\gamma^\nu\partial_\nu + m\Big)\Big(i\gamma^\mu\partial_\mu - m\Big)\psi(x) = 0. \tag{2.63}$$

If the following condition holds

$$\left(\gamma^\nu\partial_\nu\right)\left(\gamma^\mu\partial_\mu\right) \stackrel{?}{=} \partial^\mu\partial_\mu \tag{2.64}$$

then indeed we can recover the Klein–Gordon equation. If we write the above condition explicitly

$$\left(\gamma^0\partial^0 - \gamma^1\partial^1 - \gamma^2\partial^2 - \gamma^3\partial^3\right)\left(\gamma^0\partial^0 - \gamma^1\partial^1 - \gamma^2\partial^2 - \gamma^3\partial^3\right) = \partial^0\partial^0 - \partial^1\partial^1 - \partial^2\partial^2 - \partial^3\partial^3 \tag{2.65}$$

we realize that we need

$$(\gamma^0)^2 = 1 \qquad (\gamma^i)^2 = -1 \qquad \gamma^0\gamma^i + \gamma^i\gamma^0 = 0 \text{ etc.} \tag{2.66}$$

In a condensed form, we need

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = \{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \tag{2.67}$$

The γs cannot be numbers, i.e., commuting objects, they must be matrices. They are called Dirac matrices. Since we are talking about matrices, it is more appropriate to write their anticommutation relation as

$$\boxed{\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = \{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbb{1}.} \tag{2.68}$$

All the properties of the Dirac matrices descend from (2.68). The Dirac matrices are said to generate a Clifford algebra. Usually, the Dirac matrices are chosen with the extra condition (called "Hermiticity condition")

$$\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0. \tag{2.69}$$

Finally, it is very useful to use the condensed notation

$$\not{a} = \gamma^\mu a_\mu, \tag{2.70}$$

which allows us to write the Dirac equation as

$$\boxed{\left(i\not{\partial} - m\right)\psi(x) = 0.} \tag{2.71}$$

2.2.1 Dirac matrices

Due to their importance also in practical calculations, we need to spend some time on studying the properties of Dirac matrices. It is important to keep in mind that Dirac matrices have one Lorentz index and two Dirac indices. When needed to avoid confusion, we will explicitly write the Dirac indices with capital letters, i.e.,

$$\gamma^{\mu}_{AB} \tag{2.72}$$

but most of the time this indices are dropped, which may be a source of confusion at the beginning.

Dirac matrices must be traceless

$$\begin{aligned} \mathrm{Tr}[\gamma^{\nu}] &= \frac{1}{g^{\mu\mu}}\mathrm{Tr}[\gamma^{\mu}\gamma^{\mu}\gamma^{\nu}] \qquad (\text{choosing } \mu \neq \nu) \\ &= \frac{1}{g^{\mu\mu}}\mathrm{Tr}[\gamma^{\mu}\gamma^{\nu}\gamma^{\mu}] = -\frac{1}{g^{\mu\mu}}\mathrm{Tr}[\gamma^{\mu}\gamma^{\mu}\gamma^{\nu}] = -\mathrm{Tr}[\gamma^{\nu}] = 0. \end{aligned} \tag{2.73}$$

Due to the fact that $(\gamma^0)^2 = \mathbb{1}$, its eigenvalues can be only ± 1, while from $(\gamma^i)^2 = -\mathbb{1}$, their eigenvalues can be only $\pm i$. Traceless matrices with these eigenvalues can only be even dimensional. But we need at least four independent matrices, so we cannot use two dimensional matrices (there are only three independent ones plus the identiy). The minimum dimension for Dirac matrices is thus 4×4.

It is useful also to introduce the fifth matrix

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{2.74}$$

with the property

$$\gamma_5\gamma^{\mu} = -\gamma^{\mu}\gamma_5 \tag{2.75}$$

Due to the anticommutation relations, we can also write

$$\gamma_5 = -\frac{i}{4!}\varepsilon_{\mu\nu\rho\sigma}\gamma^{\mu}\gamma^{\nu}\gamma^{\rho}\gamma^{\sigma} \tag{2.76}$$

where we introduced the totally antisymmetric tensor with the convention[3]

$$\varepsilon^{0123} = +1, \qquad \varepsilon_{0123} = -1. \tag{2.77}$$

The Hermitian conjugates

$$\gamma^{0\dagger} = \gamma^0, \qquad \gamma^{i\dagger} = -\gamma^i, \qquad \gamma_5^{\dagger} = \gamma_5. \tag{2.78}$$

[3]This convention is often a source of confusion. Our convention is consistent with Peskin and Schroeder, see Eq. 3.68 in [28], but not with Mandl–Shaw, see Eq. A.13 in [26].

Other interesting properties (what is interesting is also to understand how to prove them and get used to manipulating γ matrices)

$$\gamma^\mu\gamma_\mu = \frac{1}{2}(g_{\mu\nu}\gamma^\mu\gamma^\nu + g_{\mu\nu}\gamma^\nu\gamma^\mu) = g_{\mu\nu}g^{\mu\nu} = \delta^\mu_\mu = 4, \tag{2.79}$$

$$\gamma_\mu\gamma^\alpha\gamma^\mu = \gamma_\mu(2g^{\alpha\mu} - \gamma^\mu\gamma^\alpha) = 2\gamma^\alpha - 4\gamma^\alpha = -2\gamma^\alpha, \tag{2.80}$$

$$\gamma^\beta\gamma^\mu\gamma^\nu\gamma_\beta = 4g^{\mu\nu}. \tag{2.81}$$

$$\gamma^\alpha\gamma^\mu\gamma^\beta\gamma^\nu\gamma_\alpha = -2\gamma^\nu\gamma^\beta\gamma^\mu, \tag{2.82}$$

We shall see that traces of γ matrices will be very important in the calculation of Feynman diagrams. Some useful "trace theorems" are

$$\mathrm{Tr}[\gamma^\alpha\gamma^\beta] = \frac{1}{2}\Big(\mathrm{Tr}[\gamma^\alpha\gamma^\beta] + \mathrm{Tr}[\gamma^\alpha\gamma^\beta]\Big) = g^{\alpha\beta}\mathrm{Tr}[\mathbb{1}] = 4g^{\alpha\beta}, \tag{2.83}$$

$$\mathrm{Tr}[\gamma^\alpha\gamma^\beta\gamma^\rho\gamma^\sigma] = 4\big(g^{\alpha\beta}g^{\rho\sigma} + g^{\alpha\sigma}g^{\beta\rho} - g^{\alpha\rho}g^{\beta\sigma}\big), \tag{2.84}$$

$$\mathrm{Tr}[\gamma_5] = \mathrm{Tr}[\gamma^0\gamma^0\gamma_5] = \mathrm{Tr}[\gamma^0\gamma_5\gamma^0] = -\mathrm{Tr}[\gamma^0\gamma^0\gamma_5] = 0, \tag{2.85}$$

$$\mathrm{Tr}[\text{odd num. of } \gamma] = \mathrm{Tr}[\gamma^\alpha\gamma^\beta\ldots\gamma^\sigma\gamma_5\gamma_5] = \mathrm{Tr}[\gamma_5\gamma^\alpha\gamma^\beta\ldots\gamma^\sigma\gamma_5] \tag{2.86}$$

$$= -\mathrm{Tr}[\gamma^\alpha\gamma^\beta\ldots\gamma^\sigma\gamma_5\gamma_5] = 0, \tag{2.87}$$

$$\mathrm{Tr}[\gamma_5\gamma^\alpha\gamma^\beta] = 0, \tag{2.88}$$

$$\mathrm{Tr}[\gamma_5\gamma^\alpha\gamma^\beta\gamma^\rho\gamma^\sigma] = -4i\varepsilon^{\alpha\beta\rho\sigma}. \tag{2.89}$$

Another very useful set of relations (they are all different version of the same relation), which descends quite easily from the anticommutation relations, is

$$\gamma^\mu\not{p} = 2p^\mu - \not{p}\gamma^\mu, \tag{2.90}$$

$$\not{k}\not{p} = 2p\cdot k - \not{p}\not{k}, \tag{2.91}$$

$$\not{p}^2 = p^2. \tag{2.92}$$

2.2.2 Representations of the Dirac matrices

The properties discussed above are independent of the specific representation of the matrices we choose. However, sometimes specific calculations require that we write down explicitly the matrix elements.

The so-called standard or Dirac representation is explicitly

$$\gamma^0 \overset{\text{SR}}{=} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \gamma^1 \overset{\text{SR}}{=} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \tag{2.93}$$

$$\gamma^2 \overset{\text{SR}}{=} \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \qquad \gamma^3 \overset{\text{SR}}{=} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{2.94}$$

It is of course possible to check the properties of the γ matrices directly with this representation. The notation can be considerably condensed by using the Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.95}$$

whereby we obtain

- Standard or Dirac representation

$$\gamma^0 \overset{\text{SR}}{=} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad \gamma^i \overset{\text{SR}}{=} \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \qquad \gamma_5 \overset{\text{SR}}{=} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}. \tag{2.96}$$

- Chiral representation

$$\gamma^0 \overset{\text{CR}}{=} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \qquad \gamma^i \overset{\text{CR}}{=} \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \qquad \gamma_5 \overset{\text{CR}}{=} \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}. \tag{2.97}$$

  This version is in accordance with Peskin-Schroeder. Ryder, for instance, has an opposite sign for γ^i and γ_5.

The chiral representation can be written in an even more compact way introducing the notation

$$\sigma^\mu = (\mathbb{1}, \sigma^i), \qquad \overline{\sigma}^\mu = (\mathbb{1}, -\sigma^i), \tag{2.98}$$

so that

$$\gamma^\mu \overset{\text{CR}}{=} \begin{pmatrix} 0 & \sigma^\mu \\ \overline{\sigma}^\mu & 0 \end{pmatrix}. \tag{2.99}$$

A different representation can be always obtained by means of a unitary transformation on the matrices

$$\gamma^\mu|_{\text{rep. 2}} = U\gamma^\mu|_{\text{rep. 1}}U^\dagger, \qquad UU^\dagger = 1. \tag{2.100}$$

For instance, the unitary transformation that connects the standard representation to the chiral representation is

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1} & -\mathbb{1} \\ \mathbb{1} & \mathbb{1} \end{pmatrix} = \frac{1}{\sqrt{2}} \Big(\mathbb{1} - \gamma^0\gamma_5 \Big). \tag{2.101}$$

2.2.3 The Dirac Hamiltonian

From the Dirac equation, we want to obtain an expression similar to Schrödinger equation with an Hamiltonian operator. We can do this by multiplying the equation to the left by γ^0

$$\begin{aligned}
&\big(i\gamma^\mu\partial_\mu - m\big)\psi(x) = 0\\
&\gamma^0\big(i\gamma^\mu\partial_\mu - m\big)\psi(x) = 0\\
&\big(i\partial_t\mathbb{1} + i\gamma^0\vec{\gamma}\cdot\vec{\nabla} - \gamma^0 m\big)\psi(x) = 0\\
&i\partial_t\mathbb{1}\psi(x) = H_D\psi(x)
\end{aligned} \tag{2.102}$$

with

$$H_D = -i\vec{\alpha}\vec{\nabla} + \beta m, \qquad \vec{\alpha} = \gamma^0\vec{\gamma} \qquad \beta = \gamma^0. \tag{2.103}$$

In c.g.s. units

$$i\partial_t\hbar\psi(x) = \Big[-i\hbar c\vec{\alpha}\cdot\vec{\nabla} + \beta mc^2\Big]\psi(x). \tag{2.104}$$

The Dirac Hamiltonian is an Hermitian operator because $\vec{\alpha}$ and β are Hermitian, on the basis of Eq. 2.69.[4].

2.2.4 Energy of plane-wave solutions

We seek solution of the the form

$$\psi_A(x) = f_A(k)\, e^{-ik\cdot x} = f_A(\omega, k)\, e^{-i\omega t + i\vec{k}\cdot\vec{x}} \tag{2.105}$$

note the presence of a Dirac index A: ψ_A is now not just a function, but a vector with the same indices as the Dirac matrices. It is called a Dirac spinor. The Dirac equation is a collection of four equations, for each component of the Dirac spinor.

Let us now apply the Dirac equation, in the form (2.102) to our trial solution

$$\omega\mathbb{1}f_A(k)e^{-ik\cdot x} = \big[\vec{\alpha}\cdot\vec{k} + \beta m\big]f_A(k)e^{-ik\cdot x}. \tag{2.106}$$

The expression $\vec{\alpha}\cdot\vec{k} + \beta m$ corresponds to the Hamiltonian operator in momentum space.

To make the situation clearer, we write the equation in standard representation where

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \qquad \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}. \tag{2.107}$$

We obtain

$$\begin{pmatrix} (m-\omega)\mathbb{1} & \vec{\sigma}\cdot\vec{k} \\ \vec{\sigma}\cdot\vec{k} & -(m+\omega)\mathbb{1} \end{pmatrix} f_A(k) = 0 \tag{2.108}$$

[4]Remember that the Hermitian of ∂_μ is $-\partial_\mu$

We choose $\vec{k}$ in the z direction ($\vec{k} = (0, 0, |\vec{k}|)$) and we explicitly obtain the system of equations

$$\begin{pmatrix} m-\omega & 0 & |\vec{k}| & 0 \\ 0 & m-\omega & 0 & -|\vec{k}| \\ |\vec{k}| & 0 & -m-\omega & 0 \\ 0 & -|\vec{k}| & 0 & -m-\omega \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{pmatrix} = 0. \tag{2.109}$$

The eigenvalues equations give us

$$\begin{aligned} &(m-\omega)^2(m+\omega)^2 + \vec{k}^4 + 2\vec{k}^2(m^2-\omega^2) = 0 \\ &(m^2-\omega^2)^2 + 2\vec{k}^2(m^2-\omega^2) + \vec{k}^4 = 0 \\ &\left(m^2 - \omega^2 + \vec{k}^2\right)^2 = 0 \\ &\omega = \pm\omega_k = \pm\sqrt{\vec{k}^2 + m^2} \end{aligned} \tag{2.110}$$

Note that each of the solutions is twice degenerate (of the four eigenvalues, two are $+\omega_k$ and two are $-\omega_k$). Interestingly, Dirac started from a first-order partial equation in the hope of removing the negative energy states, and he ended up with four possible solutions, two of which are negative. It turns out that this degeneracy is due to the fact that the Dirac equation describes spin-half particles, which can have spin up or down. The two spin states have the same energy in the case of a free particle.

Similarly to the Klein–Gordon case: there exist states where the energy is negative and not bounded from below. Dirac tried to solve this issue in a paper in 1930 [17]. He hypothesized that the vacuum is in reality filled by a "sea" of infinitely many particles. If the particles are electrons, due to Pauli exclusion principle they fill the energy levels up to what can be defined as the zero-point energy and no electron can fall in an energy state lower than that (this would not be possible for a integer-spin particle). However, an electron of the sea can be promoted to a higher level of energy, above zero. This at the same time creates a positive-energy electron and leaves behind a "hole" in the sea. A hole in the sea produces the same effects as a particle with the same characteristics of the electron, but opposite charge. This idea led to the prediction of the existence of the anti-electron, i.e., the positron. At the beginning, Dirac thought that the proton could represent the antiparticle of the electron (1929). However, it was shown by H. Weyl that negative-energy solutions must have the same mass as the positive-energy ones (1930). Moreover, Oppenheimer and Tamm proved that the possibility of annihilation of protons and electrons would quickly lead to instability of ordinary matter (1930). Therefore, Dirac predicted the existence of the positron in 1931. The positron was eventually observed by Anderson in 1932. Heisenberg wrote: "I think that the discovery of antimeatter was perhaps the biggest jump of all the big jumps in physics in our century."

In spite of this achievement, the theory of holes has some drawbacks. Here is an incomplete list of problems: it works only for fermions (implying that no relativistic theory for bosons could exist); the infinite collections of sea fermions should have an infinite charge and infinite mass; there could be interactions among sea fermions.

2.2.5 The Dirac four-current

It is possible to define also in the case of the Dirac equation a conserved four-current

$$j^\mu = \overline{\psi}\gamma^\mu\psi \tag{2.111}$$

where we defined

$$\overline{\psi} = \psi^\dagger\gamma^0 \tag{2.112}$$

It is possible to derive this "strange" formula in a way that we shall see in Sec. 2.2.11. It is also possible to check that it has indeed the properties of a four-vector, as we shall see in one of the next subsections. For the moment, let's give the formula for granted and let's check if its four-divergence is zero and what is the associated conserved quantity.

We first need to write the "conjugate" Dirac equation for $\overline{\psi}$. We take the Dirac equation, transpose and conjugate it, insert a $\gamma^0\gamma^0 = 1$ and multiply to the right by γ^0 (remember that $\gamma^0\gamma^0 = 1$):

$$\begin{aligned} & i\partial_\mu\gamma^\mu\psi - m\psi = 0 \\ \Rightarrow & -i\partial_\mu\psi^\dagger\gamma^{\mu\dagger} - m\psi^\dagger = 0 \\ & -i\partial_\mu\psi^\dagger\gamma^0\gamma^0\gamma^{\mu\dagger}\gamma^0 - m\psi^\dagger\gamma^0 = 0 \\ & i\partial_\mu\overline{\psi}\gamma^\mu + m\overline{\psi} = 0 \end{aligned} \tag{2.113}$$

Often, this equation is written as

$$\begin{aligned} \overline{\psi}\Big(i\overleftarrow{\partial}_\mu\gamma^\mu + m\Big) = 0 \\ \overline{\psi}\Big(i\overleftarrow{\not{\partial}} + m\Big) = 0 \end{aligned} \tag{2.114}$$

The arrow simply means that the derivative operator acts on the function to its left. It should not be confused with a vector. The reason for introducing this notation is just that the equation looks better.

Let us now check the four-divergence of the current

$$\begin{aligned} \partial_\mu j^\mu &= \partial_\mu\big(\overline{\psi}\gamma^\mu\psi\big) \\ &= \big(\partial_\mu\overline{\psi}\big)\gamma^\mu\psi + \overline{\psi}\gamma^\mu\big(\partial_\mu\psi\big) \\ &= im\overline{\psi}\psi - im\overline{\psi}\psi = 0 \end{aligned} \tag{2.115}$$

Now, we can observe that the "density" j^0 is

$$\rho = j^0 = \overline{\psi}\gamma^0\psi = \psi^\dagger\gamma^0\gamma^0\psi = \psi^\dagger\psi \tag{2.116}$$

We see in this case that the density is positive definite, contrary to the Klein–Gordon case.

2.2.6 Coupling to electromagnetic field and nonrelativistic limit

As in the case of classical mechanics or the Schrödinger equation, the coupling to the electromagnetic field can be obtained by the so-called "minimal substitution", i.e.,

$$\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu \tag{2.117}$$

where $A^\mu = (\Phi, \vec{A})$ is the four-potential of the elm field. If we want to recover the correct expression including explicit $\hbar$ and c factors, we need to check that qA^μ has the dimensions of an energy, ML^2/T^2, while ∂_μ as the dimensions of $1/L$. Therefore, we have to divide the second term by $\hbar c$, which has dimensions ML^3/T^2, i.e.,

$$\partial_\mu \to D_\mu = \partial_\mu + \frac{iq}{\hbar c}A_\mu \quad \text{(c.g.s.)} \tag{2.118}$$

The Dirac equation for an electron with $q = -e$ becomes

$$\left[i\hbar\left(\partial\!\!\!/ - \frac{ie}{\hbar c}A\!\!\!/\right) - mc\right]\psi = 0 \tag{2.119}$$

This can be rewritten (multiplying to the left by γ^0 and remembering that the $\partial^0 = \partial_t/c$

$$\begin{aligned} &\gamma^0\left[i\hbar\gamma^0\frac{\partial_t}{c} + \frac{e}{c}\gamma^0\Phi - \frac{e}{c}\vec{\gamma}\cdot\vec{A} + i\hbar\vec{\gamma}\cdot\vec{\nabla} - mc\right]\psi = 0 \\ &i\hbar\mathbb{1}\partial_t\psi = \left[-ic\hbar\vec{\alpha}\cdot\vec{\nabla} + mc^2\beta - e\Phi\mathbb{1} + e\vec{\alpha}\cdot\vec{A}\right]\psi \end{aligned} \tag{2.120}$$

To study the nonrelativistic limit, it is convenient to first pose

$$\psi = e^{-i\frac{mc^2}{\hbar}t}\psi' \tag{2.121}$$

so that the equation becomes

$$i\hbar\mathbb{1}\partial_t\psi' = \left[c\vec{\alpha}\cdot\left(-i\hbar\vec{\nabla} + \frac{e}{c}\vec{A}\right) - e\Phi\mathbb{1} + mc^2(\beta - \mathbb{1})\right]\psi' \tag{2.122}$$

work in standard representation and split the (four-dimensional) Dirac spinor in two components (each one two-dimensional)

$$\psi' = \begin{pmatrix}\varphi \\ \chi\end{pmatrix}. \tag{2.123}$$

The splitting could be done also using the projector operators

$$\mathcal{P}^+ = \frac{\mathbb{1} + \gamma^0}{2}, \qquad \mathcal{P}^- = \frac{\mathbb{1} - \gamma^0}{2}. \tag{2.124}$$

Then, we obtain two (two-dimensional) equations

$$\begin{cases} i\hbar\mathbb{1}\partial_t\varphi = c\vec{\sigma}\cdot\left(-i\hbar\vec{\nabla}+\frac{e}{c}\vec{A}\right)\chi - e\Phi\mathbb{1}\varphi \\ i\hbar\mathbb{1}\partial_t\chi = c\vec{\sigma}\cdot\left(-i\hbar\vec{\nabla}+\frac{e}{c}\vec{A}\right)\phi - e\Phi\mathbb{1}\chi - 2mc^2\mathbb{1}\chi \end{cases} \tag{2.125}$$

The identity matrices (now in two dimensions) are a bit redundant, but they make it explicitly clear that each equation is still two-dimensional. We shall drop them in the following.

$$\begin{cases} \left(i\hbar\partial_t + e\Phi\right)\varphi = c\vec{\sigma}\cdot\left(-i\hbar\vec{\nabla}+\frac{e}{c}\vec{A}\right)\chi \\ \left(i\hbar\partial_t + e\Phi + 2mc^2\right)\chi = c\vec{\sigma}\cdot\left(-i\hbar\vec{\nabla}+\frac{e}{c}\vec{A}\right)\varphi \end{cases} \tag{2.126}$$

The nonrelativistic approximation amounts to neglecting the terms $\partial_t + e\Phi$ compared to $2mc^2$ in the second equation, leading to

$$\chi \approx \frac{1}{2mc}\vec{\sigma}\cdot\left(-i\hbar\vec{\nabla}+\frac{e}{c}\vec{A}\right)\varphi \tag{2.127}$$

In practice, it also means that $\chi \ll \varphi$. We then obtain an equation for φ

$$\left(i\hbar\partial_t + e\Phi\right)\varphi \approx \frac{1}{2m}\left[\vec{\sigma}\cdot\left(-i\hbar\vec{\nabla}+\frac{e}{c}\vec{A}\right)\right]^2\varphi \tag{2.128}$$

On the r.h.s. we have an expression of this type (from now on, all expressions are in Euclidean metric, i.e., the upper or lower position of the indices does not matter)

$$\begin{aligned} (\vec{\sigma}\cdot\vec{a})(\vec{\sigma}\cdot\vec{b}) &= \sigma_i\sigma_j a^i b^j \\ &= \left(\frac{1}{2}\{\sigma_i,\sigma_j\} + \frac{1}{2}[\sigma_i,\sigma_j]\right)a^i b^j \\ &= \left(\frac{1}{2}2\delta_{ij}\mathbb{1} + \frac{1}{2}2i\varepsilon_{ijk}\sigma^k\right)a^i b^j \\ &= \vec{a}\cdot\vec{b}\mathbb{1} + i\vec{\sigma}\cdot(\vec{a}\times\vec{b}) \end{aligned} \tag{2.129}$$

We have to be careful when considering the second term because instead of a and b we have operators acting on the φ. In fact

$$\varepsilon_{ijk}\left(-i\hbar\partial^i + eA^i/c\right)\left(-i\hbar\partial^j + eA^j/c\right)\varphi = \varepsilon_{ijk}\left(-\hbar^2\partial^i\partial^j - i\hbar\frac{e}{c}\left(\partial^i A^j + A^i\partial^j\right) + \frac{e^2}{c^2}A^iA^j\right)\varphi \tag{2.130}$$

The first and last terms drop because ε is antisymmetric. The surviving term is

$$\partial^i A^j\varphi = (\partial^i A^j)\varphi + A^j\partial^i\varphi \tag{2.131}$$

Therefore, our expression becomes

$$-\varepsilon_{ijk} i\hbar \frac{e}{c}\left[(\partial^i A^j) + A^j\partial^i + A^i\partial^j\right]\varphi = -\varepsilon_{ijk} i\hbar \frac{e}{c}(\partial^i A^j)\varphi = -i\hbar\frac{e}{c}\left(\vec{\nabla}\times\vec{A}\right)_k\varphi = -i\hbar\frac{e}{c}B_k\varphi \tag{2.132}$$

In conclusion, our nonrelativistic equation becomes

$$i\hbar\partial_t\varphi \approx \left[\frac{1}{2m}\left(-i\hbar\vec{\nabla} + \frac{e}{c}\vec{A}\right)^2 - e\Phi + \frac{\hbar}{2m}\frac{e}{c}\vec{B}\cdot\vec{\sigma}\right]\varphi \tag{2.133}$$

In practice, we obtain the Schrödinger–Pauli equation with the standard coupling to an elm field (first part) plus a term that couples the magnetic field to the spin.

Not only: usually a magnetic moment μ couples to the magnetic field with $\vec{B}\cdot\vec{\mu}$. The magnetic moment is connected to the angular momentum by

$$\vec{\mu} = \frac{q}{2mc}\vec{L} \tag{2.134}$$

Naively, we could assume that the same relation applies to spin $\vec{S} = \frac{\hbar}{2}\vec{\sigma}$. But in reality, we have to introduce a g factor

$$\vec{\mu} = g\frac{q}{2mc}\vec{S} \tag{2.135}$$

For the Dirac particle (i.e., electrons and in general spin-half particles), inspection of the Schrödinger–Pauli equation tells us that $g = 2$. This result was obtained by Pauli out of phenomenological considerations and Dirac obtained it from theoretial principles.

It is also possible to solve the Dirac equation when a Coulomb potential is present ($\Phi = Ze/r$ and $\vec{A} = 0$). We should start from Eq. (2.103), replace the time derivative with $\partial_t/c \longrightarrow \partial_t/c + \frac{i}{\hbar c}\frac{Ze^2}{r}$ to obtain the Dirac Hamiltonian in a Coulomb potential

$$H'_D = -i\hbar c\vec{\alpha}\cdot\vec{\nabla} + \beta mc^2 - \frac{Ze^2}{r}. \tag{2.136}$$

The eigenvalues of the Hamiltonian can be obtained (by means of several steps, see, e.g., Sec. 3.7 of [25]) and correspond to the energy levels of a hydrogen atom, with the correct spin-orbit coupling and all relativistic corrections in place (Darwin term, Thomas precession factor), i.e.,

$$\begin{aligned} E &= mc^2\left[1 + \left(\frac{Z\alpha}{n - \left(j+\frac{1}{2}\right) + \left[\left(j+\frac{1}{2}\right)^2 - Z^2\alpha^2\right]^{\frac{1}{2}}}\right)^2\right]^{-\frac{1}{2}} \\ &\approx mc^2\left[1 - \frac{Z^2\alpha^2}{2n^2} - \frac{Z^4\alpha^4}{2n^4}\left(\frac{n}{j+\frac{1}{2}} - \frac{3}{4}\right)\right]. \end{aligned} \tag{2.137}$$

2.2.7 Excursus: rotations and Pauli spinors

Before addressing the general question of Lorentz transformations and Dirac spinors, let us first remind ourselves some useful issues concerning three-dimensional rotations and Pauli spinors.

Rotations are orthogonal matrices with determinant 1

$$R^T = R^{-1}, \qquad \det R = +1 \tag{2.138}$$

Matrices with such properties form a group, called the Special Orthogonal group in n dimensions, $SO(n)$. The group is non-Abelian because the elements of the group do not commute. The combination of these transformations with space reflections forms the general orthogonal group, $O(n)$, which encompasses also improper rotations, with $\det R = -1$. In general, a group is an abstract mathematical construction that is independent of its specific *representation*. For instance, rotation matrices in a three-dimensional space are a representation of the group $SO(3)$. They act on three-dimensional vectors, which constitute the basis for that representation. The dimension of the representation corresponds to the dimension of the base space (in this case 3).[5] A (completely) reducible representation can be written in block-diagonal form and its basis vectors can be split into subsets that do not mix with each other.

We know very well how to write down a rotation matrix in three dimensions. There are three of them, let us consider a rotation by an angle θ about the z axis

$$\begin{aligned} R^3{}_{ij} \equiv R^{(12)}{}_{ij} &= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \overset{\theta\to 0}{\approx} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \theta \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \delta_{ij} + \theta\left(-\delta^1_i\delta^2_j + \delta^2_i\delta^1_j\right) = \mathbb{1} + i\theta J^3, \end{aligned} \tag{2.139}$$

where in the second step we have already written the infinitesimal form of the rotation matrix, which is just a small variation of the identity. Since the elements of the group can be defined by means of parameters that take continuous values (the angle θ and the other angles in the other directions), the group is called a Lie group. The operator J^3 is defined as the generator of rotations around the z axis. The group is characterized by three such generators, which fulfill the following commutation relations (the algebra of the Lie group)

$$[J^i, J^j] = i\varepsilon_{ijk}J^k. \tag{2.140}$$

The commutation relations are the same as for the components of angular momentum. In fact, angular momentum operators are the generators of rotation. The values of ε_{ijk} are called the structure constants. The generators look different in different representations of the group, but they follow always the same algebra and the structure constants remain the same.

[5] The dimensions of the representation do not have to correspond to the n index denoting the group.

A rotation of a finite angle about a direction $\vec{n}$ can be written as

$$R = e^{i\theta \vec{J}\cdot\vec{n}} \tag{2.141}$$

and so forth for the other components. The exponential of a matrix is defined by its power series expansion, i.e.:

$$e^M = \sum_{k=0}^{\infty} \frac{1}{k!} M^k. \tag{2.142}$$

For instance, in the case of R^3 we have written an infinitesimal rotation as in Eq. (2.139) and we can check that we can write a finite rotation as

$$\begin{aligned} R^3 = e^{i\theta J^3} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \theta \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{\theta^2}{2} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{\theta^3}{6} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \dots \\ &= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \tag{2.143}$$

We can give a different proof of the above by finding the unitary matrix that diagonalizes iJ^3, which is

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} i & -i & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}. \tag{2.144}$$

In fact

$$U^{-1}(iJ^3)U = \begin{pmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.145}$$

Then

$$R^3 = e^{i\theta J^3} = U e^{\theta U^{-1}(iJ^3)U} U^{-1} = U \begin{pmatrix} e^{i\theta} & 0 & 0 \\ 0 & e^{-i\theta} & 0 \\ 0 & 0 & 1 \end{pmatrix} U^{-1} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.146}$$

Suppose we want to check that non-relativistic quantum mechanics, is invariant under rotations. For a free particle, for instance, we ought to check that the Schrödinger equation

$$i\hbar\partial_t \psi(t,\vec{x}) = -\frac{\hbar^2 \vec{\nabla}^2}{2m} \psi(t,\vec{x}) \tag{2.147}$$

does not change under rotations. On the wavefunction, the rotation has the following effect

$$\psi'(t,\vec{x}' = R\vec{x}) = \psi(t,\vec{x}), \tag{2.148}$$

i.e., the value of the rotated wavefunction at the rotated position $\vec{x}'$ is the same as the value of the original wavefunction at $\vec{x}$. The wavefunction in this case is invariant under rotations. This is naturally the case whenever the function depends on scalar products between vectors, for instance on $\vec{k} \cdot \vec{x}$.

To be sure about the invariance of the Schrödinger equation, we have finally to check that

$$\left(R(\vec{\nabla})\right)^2 = \vec{\nabla} R^T \cdot R\vec{\nabla} = \vec{\nabla}^2 \tag{2.149}$$

(the time-derivative term is also trivial).

This case is admittedly quite simple. But the situation is not trivial when it comes to the addition of spin. The Schrödinger–Pauli equation contains an operator $\vec{B} \cdot \vec{\sigma}$ and the wavefuntion is now represented by a two-component Pauli spinor. We would like our equation to retain its form after rotations, i.e., to be rotation covariant. Since a rotation changes the vector $\vec{B}$, covariance can be guaranteed only if rotations have an effect also on the spinors. Ideed, in quantum mechanics we learn that a spinor oriented along the z axis is different from a spinor oriented along the x axis and the two can be connected by a rotation matrix U in spinor space. A rotation then leads to the following changes

$$\vec{B} \to R\vec{B}, \qquad \psi \to U\psi, \qquad \vec{B} \cdot \vec{\sigma}\, \psi \to U\big(\vec{B} \cdot \vec{\sigma}\, \psi\big). \tag{2.150}$$

where the rotation in Pauli spinor space can be described by means of a 2×2 unitary matrix ($SU(2)$ group).

Rotational covariance means that

$$(R\vec{B}) \cdot \vec{\sigma}\, U\psi = U\big(\vec{B} \cdot \vec{\sigma}\, \psi\big) \tag{2.151}$$

that is, acting with the rotated operator after rotating the spinor is the same as rotating the spinor after acting with the operator. The above requirement leads to

$$\Rightarrow U^{-1}(R\vec{B}) \cdot \vec{\sigma}\, U\psi = \vec{B} \cdot \vec{\sigma}\, \psi \qquad \Rightarrow (R\vec{B}) \cdot (U^{-1}\vec{\sigma}\, U)\psi = \vec{B} \cdot \vec{\sigma}\, \psi \tag{2.152}$$

which is satisfied if

$$\boxed{U^{-1}\sigma^i\, U = R^i{}_j \sigma^j} \tag{2.153}$$

Having this result in our hands, we can conclude that the expectation value of the spin operator transforms as a normal three-dimensional vector under rotations, i.e.,

$$\psi'^{\dagger}\vec{\sigma}\psi' = R(\psi^{\dagger}\vec{\sigma}\psi) \tag{2.154}$$

This behavior is far from trivial. Using the notation $\vec{\sigma}$ makes sense only because of this property.

Explicitly, Pauli spinor rotations can be written

$$\psi' = U\psi = e^{-i\,\theta\,\vec{\sigma}\cdot\vec{n}/2}\psi. \tag{2.155}$$

For a rotation about the z axis we can write

$$\begin{aligned} U^3 \equiv U^{(12)} \equiv e^{-i\theta,\sigma^3/2} &= \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix} = \begin{pmatrix} \cos\frac{\theta}{2} - i\sin\frac{\theta}{2} & 0 \\ 0 & \cos\frac{\theta}{2} + i\sin\frac{\theta}{2} \end{pmatrix} \\ &= \mathbb{1}\cos\frac{\theta}{2} - i\sigma^3\sin\frac{\theta}{2} \approx \mathbb{1} - i\,\theta\,\frac{\sigma^3}{2} = \mathbb{1} - \theta\Big[\frac{\sigma^1}{2}, \frac{\sigma^2}{2}\Big] \end{aligned} \tag{2.156}$$

For an infinitesimal rotation about the z axis we can check that

$$\left(\mathbb{1} + i\theta\frac{\sigma^3}{2}\right)\sigma_i\left(\mathbb{1} - i\theta\frac{\sigma^3}{2}\right) = \Big[\delta_{ij} + \theta\big(-\delta_i^1\delta_j^2 + \delta_i^2\delta_j^1\big)\Big]\,\sigma^j = \big(\mathbb{1} + i\theta J_3\big)_{ij}\sigma^j. \tag{2.157}$$

This result is in agreement with Eq. (2.153).

In summary, rotational covariance of the Schrödinger-Pauli equation implies specific behaviors of Pauli spinors under rotations. We will see an analogous discussion in the case of Dirac spinors.

2.2.8 Excursus: Lorentz transformations

We have already reminded ourselves that Lorentz transformations enjoy the property of leaving four-dimensional scalar products unchanged. In particular, we have written Eq. (2.22), which can be rewritten in several equivalent ways

$$L^\mu{}_\nu L_\mu{}^\sigma = \delta_\nu^\sigma \tag{2.158}$$

$$L^{\mu\nu} L_\mu{}^\sigma = g^{\nu\sigma} \tag{2.159}$$

$$L^{\mu\nu} L_{\mu\sigma} = \delta_\sigma^\nu \tag{2.160}$$

$$\big(L^{\nu\mu}\big)^T L_{\mu\sigma} = \delta_\sigma^\nu \qquad \Rightarrow L^T = L^{-1} \tag{2.161}$$

Yet another equivalent way is

$$\boxed{g_{\mu\nu} = g^{\rho\sigma} L_{\rho\mu} L_{\sigma\nu} \qquad (g = L^T g L)} \tag{2.162}$$

Due to the above properties, Lorentz transformations must have

$$|\det L| = 1, \qquad\qquad |L_0^0| \geq 1. \tag{2.163}$$

(The second relation follows from Eq. (2.158) with $\nu = 0$ and $\sigma = 0$)

Lorentz transformations form a group, the Lorentz group, denoted as $O(1,3)$. Transformations with $\det L = +1$ and $L_0^0 \geq 1$ form a continous subgroup named the restricted Lorentz group, $SO^+(1,3)$. Addition of the space inversion and time inversion tranformations complete the group.

Transformations of the restricted Lorentz group can be written in an infinitesimal form, which is given by the following equation

$$L^\mu{}_\nu \approx \delta_\nu^\mu + \sum_{\alpha=0}^{3}\sum_{\beta>\alpha}^{3} \epsilon_{\alpha\beta} M^{(\alpha\beta)\mu}{}_\nu \tag{2.164}$$

The indices α and β denote if we are making a boost (one of the two indices must be 0) or a rotation (both indices must be spatial). For instance, $(\alpha\beta) = (1\ 2)$ denotes a rotation about the z axis.

Any finite Lorentz transformation can be seen as a combination of several succesive infinitesimal transformations and can be written in an exponential form as (fixing α and β)

$$L^{(\alpha\beta)\mu}{}_{\nu} = e^{\epsilon M^{(\alpha\beta)\mu}{}_{\nu}}. \tag{2.165}$$

In order for transformation (2.164) to be a Lorentz transformation, we need to check that, for any α and β,

$$\begin{aligned} g_{\rho\sigma}L^{\rho}{}_{\mu}L^{\sigma}{}_{\nu} &= g_{\mu\nu} \\ g_{\rho\sigma}L^{\rho}{}_{\mu}L^{\sigma}{}_{\nu} &= g_{\rho\sigma}\left(\delta^{\rho}_{\mu} + \epsilon M^{\rho}{}_{\mu}\right)\left(\delta^{\sigma}_{\nu} + \epsilon M^{\sigma}{}_{\nu}\right) \\ &\approx g_{\mu\nu} + \epsilon\left(g_{\mu\sigma}M^{\sigma}{}_{\nu} + g_{\rho\nu}M^{\rho}{}_{\mu}\right) \\ &= g_{\mu\nu} + \epsilon\left(M_{\mu\nu} + M_{\nu\mu}\right) \end{aligned} \tag{2.166}$$

This means that the tensor M must be antisymmetric in the μ and ν indices, i.e.,

$$M_{\mu\nu} = -M_{\nu\mu}. \tag{2.167}$$

The "generators" of the infinitesimal transformations can be written down explicitly as

$$M^{(\alpha\beta)\mu}{}_{\nu} = g^{\alpha\mu}\delta^{\beta}_{\nu} - g^{\beta\mu}\delta^{\alpha}_{\nu} \tag{2.168}$$

$$M^{(\alpha\beta)}{}_{\mu\nu} = \delta^{\alpha}_{\mu}\delta^{\beta}_{\nu} - \delta^{\beta}_{\mu}\delta^{\alpha}_{\nu} \tag{2.169}$$

This expression is in agreement with (2.166). We can also check that the expression is correct starting from the explicit expressions of some Lorentz transformation. Let's start from a rotation about the z axis by an angle ϵ

$$L^{(12)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\epsilon & -\sin\epsilon & 0 \\ 0 & \sin\epsilon & \cos\epsilon & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \overset{\epsilon\to 0}{\approx} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \epsilon \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.170}$$

The last matrix in fact corresponds to

$$M^{(12)\mu}{}_{\nu} = g^{1\mu}\delta^{2}_{\nu} - g^{2\mu}\delta^{1}_{\nu} \tag{2.171}$$

(the only nonzero elements are for $\mu = 1$ and $\nu = 2$ and the value is -1, or for $\mu = 2$ and $\nu = 1$ and the value is $+1$).

We can check also a boost in the x direction[6]

$$L^{(01)} = \begin{pmatrix} \cosh\epsilon & \sinh\epsilon & 0 & 0 \\ \sinh\epsilon & \cosh\epsilon & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \overset{\epsilon\to 0}{\approx} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \epsilon \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.172}$$

[6]Remeber that ϵ in this case represents the rapidity $\epsilon = \ln[\gamma(1+\beta)] = \ln\sqrt{(1+v/c)/(1-v/c)}$.

The last matrix in fact corresponds to

$$M^{(01)\mu}{}_{\nu} = g^{0\mu}\delta^1_\nu - g^{1\mu}\delta^0_\nu \tag{2.173}$$

(the only nonzero elements are for $\mu = 0$ and $\nu = 1$ and the value is 1, or for $\mu = 1$ and $\nu = 0$ and the value is $+1$). The matrix is not antisymmetric, but this is because we have used it in the form with a contravariant and a covariant index. As soon as we lower the contravariant index, we obtain an antisymmetric matrix.

We can write down the algebra of the generators of Lorentz transformations

$$\left[M^{(\alpha\beta)}, M^{(\gamma\delta)}\right] = -g^{\alpha\gamma}M^{(\beta\delta)} - g^{\beta\delta}M^{(\alpha\gamma)} + g^{\alpha\delta}M^{(\beta\gamma)} + g^{\beta\gamma}M^{(\alpha\delta)}. \tag{2.174}$$

This is what fundamentally defines the Lorentz group. The standard transformations in four-dimensional space are just one possible *representation* of this algebra (we could call it the vector representation). It is possible to find other representations that fulfill this algebra. Note that we used the term "representation" with two different meanings: when we speak about a group representation, we mean a set of transformations (acting on some vector space with certain dimensions) that satisfies the algebra of the group. When we spoke about the standard or chiral representations of Dirac matrices, we simply meant two different ways of writing the Dirac matrices, corresponding to two different bases in the Dirac spinor space. When necessary to avoid confusion, we will speak about a "standard basis" and a "chiral basis."

2.2.9 Dirac equation and Lorentz transformations

The Dirac equation reads

$$(i\gamma^\mu\partial_\mu - m)\psi(x) = 0 \tag{2.175}$$

The application of Lorentz transformations leads to

$$(i\gamma^\mu\partial'_\mu - m)\psi'(x') \stackrel{?}{=} 0. \tag{2.176}$$

where

$$x'^\mu = L^\mu{}_\nu x^\nu, \qquad \partial'_\mu = L_\mu{}^\rho\partial_\rho \tag{2.177}$$

In order for the Dirac equations to be relativistically covariant, the spinor has to transform in the appropriate way under Lorentz transformations. We have to find a matrix in Dirac space that describes this transformation

$$\psi'(x') = \Lambda\psi(x) \tag{2.178}$$

and for which (compare with Eq. (2.152))

$$\Lambda^{-1}(i\gamma^\mu\partial'_\mu - m)\Lambda\psi(x) = (i\gamma^\mu\partial_\mu - m)\psi(x) = 0. \tag{2.179}$$

In practice, the above identity is guaranteed if there exist a matrix for which (compare with Eq. (2.153))

$$\boxed{\Lambda^{-1}\gamma^{\mu}\Lambda = L^{\mu}{}_{\nu}\gamma^{\nu},} \tag{2.180}$$

i.e., the effect of Λ is to transform the γ matrices formally as the components of a four-vector (justifying a posteriori why we denote them with an index μ).

We assume that the infinitesimal form of the matrix we are looking for can be written as

$$\Lambda \approx \mathbb{1} - \frac{i}{2}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}\epsilon_{\alpha\beta}\sigma^{\alpha\beta} \tag{2.181}$$

Then, we can write the infinitesimal transformations (we focus for the moment on a transformation with a specific α and β)

$$\begin{aligned}\Lambda^{-1}\gamma^{\mu}\Lambda &\approx \Big(\mathbb{1} + \frac{i}{2}\epsilon\sigma^{\alpha\beta}\Big)\gamma^{\mu}\Big(\mathbb{1} - \frac{i}{2}\epsilon\sigma^{\alpha\beta}\Big) = \gamma^{\mu} + \frac{i}{2}\epsilon\sigma^{\alpha\beta}\gamma^{\mu} - \frac{i}{2}\epsilon\gamma^{\mu}\sigma^{\alpha\beta} \\ &= \gamma^{\mu} + \frac{i}{2}\epsilon\big[\sigma^{\alpha\beta},\gamma^{\mu}\big]\end{aligned} \tag{2.182}$$

$$L^{\mu}{}_{\nu}\gamma^{\nu} \approx \gamma^{\mu} + \epsilon M^{(\alpha\beta)\mu}{}_{\nu}\gamma^{\nu} \tag{2.183}$$

and the covariance condition (2.180) can be written as

$$\frac{i}{2}\big[\sigma^{\alpha\beta},\gamma^{\mu}\big] = M^{(\alpha\beta)\mu}{}_{\nu}\gamma^{\nu} = g^{\alpha\mu}\gamma^{\beta} - g^{\beta\mu}\gamma^{\alpha}. \tag{2.184}$$

The above condition is solved by the following explicit combination

$$\boxed{\sigma^{\alpha\beta} = \frac{i}{2}\big[\gamma^{\alpha},\gamma^{\beta}\big].} \tag{2.185}$$

Let us check

$$\frac{i}{2}\big[\sigma^{\alpha\beta},\gamma^{\mu}\big] = -\frac{1}{4}\Big[\big[\gamma^{\alpha},\gamma^{\beta}\big],\gamma^{\mu}\Big] = -\frac{1}{4}\Big[\gamma^{\alpha}\gamma^{\beta},\gamma^{\mu}\Big] + \frac{1}{4}\Big[\gamma^{\beta}\gamma^{\alpha},\gamma^{\mu}\Big] \tag{2.186}$$

using

$$[ab,c] = a\{b,c\} - \{c,a\}b \tag{2.187}$$

we obtain

$$\begin{aligned}\frac{i}{2}\big[\sigma^{\alpha\beta},\gamma^{\mu}\big] &= -\frac{1}{2}\gamma^{\alpha}g^{\beta\mu} + \frac{1}{2}g^{\alpha\mu}\gamma^{\beta} + \frac{1}{2}\gamma^{\beta}g^{\alpha\mu} - \frac{1}{2}g^{\beta\mu}\gamma^{\nu} \\ &= g^{\alpha\mu}\gamma^{\beta} - g^{\beta\mu}\gamma^{\alpha} = M^{(\alpha\beta)\mu}{}_{\nu}\gamma^{\nu}.\end{aligned} \tag{2.188}$$

We could check that the $\sigma^{\alpha\beta}$ generators respect the same algebra as the $M^{(\alpha\beta)}$ generators. They form therefore another representation of the Lorentz group, which we can call the spinor representation.

We can write the generators of the Λ transformations explicitly in chiral basis

$$\sigma^{ij} = \frac{i}{2}\left[\gamma^i, \gamma^j\right] \overset{\text{CR}}{=} \varepsilon^{ijk} \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix}, \tag{2.189}$$

$$\sigma^{0i} = \frac{i}{2}\left[\gamma^0, \gamma^i\right] \overset{\text{CR}}{=} i \begin{pmatrix} -\sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}. \tag{2.190}$$

In standard basis

$$\sigma^{ij} \overset{\text{SR}}{=} \varepsilon^{ijk} \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix}, \qquad \sigma^{0i} \overset{\text{SR}}{=} i \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix}. \tag{2.191}$$

Writing the Dirac spinors in chiral basis as

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}, \tag{2.192}$$

from the explicit (block-diagonal) form of the generators we can conclude that Lorentz transformations act separately on the upper and lower components of the spinor (in chiral basis). Rotations transform upper and lower components in the same way, while boost act differently on upper and lower components. The fact that the generators can be written in block-diagonal form is an indication that the spinor representation we are using is in fact *reducible*, and we could wonder if it is possible to find simpler representations. We will not investigate further this topic, but this leads to the introduction of Weyl spinors (see, e.g., Sec. 2.3 and 2.7 of Ryder, or Sec. 2.5 and 2.6 of Maggiore).

The explicit form of a finite boost transformation in chiral representation is

$$\Lambda^{(0i)} = e^{-\frac{i}{2}\epsilon\sigma^{0i}} \overset{\text{CR}}{=} \begin{pmatrix} e^{-\epsilon\sigma^i/2} & 0 \\ 0 & e^{\epsilon\sigma^i/2} \end{pmatrix}. \tag{2.193}$$

In particular, a boost in the z direction can be written as

$$\Lambda^{(03)} = e^{-\frac{i}{2}\epsilon\sigma^{03}} \overset{\text{CR}}{=} \begin{pmatrix} e^{-\epsilon\sigma^3/2} & 0 \\ 0 & e^{\epsilon\sigma^3/2} \end{pmatrix} = \begin{pmatrix} \cosh\frac{\epsilon}{2} - \sigma^3 \sinh\frac{\epsilon}{2} & 0 \\ 0 & \cosh\frac{\epsilon}{2} + \sigma^3 \sinh\frac{\epsilon}{2} \end{pmatrix}. \tag{2.194}$$

For a generic boost in the $\vec{n}$ direction we can use a vector of boost parameters $\epsilon\vec{n}$. The final result is simple (although the steps to obtain it are not so simple):

$$\Lambda = e^{-\frac{i}{2}\epsilon n^i \cdot \sigma^{0i}} \overset{\text{CR}}{=} \begin{pmatrix} e^{-\epsilon\vec{n}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{\epsilon\vec{n}\cdot\vec{\sigma}/2} \end{pmatrix} = \begin{pmatrix} \cosh\frac{\epsilon}{2} - \vec{n}\cdot\vec{\sigma} \sinh\frac{\epsilon}{2} & 0 \\ 0 & \cosh\frac{\epsilon}{2} + \vec{n}\cdot\vec{\sigma} \sinh\frac{\epsilon}{2} \end{pmatrix}. \tag{2.195}$$

In standard representation

$$\Lambda = e^{-\frac{i}{2}\epsilon n^i \cdot \sigma^{0i}} \overset{\text{SR}}{=} \begin{pmatrix} \cosh\frac{\epsilon}{2} & \vec{n}\cdot\vec{\sigma} \sinh\frac{\epsilon}{2} \\ \vec{n}\cdot\vec{\sigma} \sinh\frac{\epsilon}{2} & \cosh\frac{\epsilon}{2} \end{pmatrix}. \tag{2.196}$$

It is useful to explicitly consider boosts from the rest-frame of a particle, where the momentum is $k_{\mathrm{RF}} = (m, 0, 0, 0)$, to a frame where the momentum is $k = (E, k^1, k^2, k^3)$, with $|\vec{k}| = \sqrt{E^2 - m^2}$. According to what is more convenient, we can express the boosts in terms of E and m, or $|\vec{k}|$ and m. The rapidity (the boost parameter) must be

$$\epsilon = \frac{1}{2} \ln \frac{E + |\vec{k}|}{E - |\vec{k}|}, \tag{2.197}$$

which means also

$$\cosh \frac{\epsilon}{2} = \sqrt{\frac{E+m}{2m}} = \frac{E+m}{\sqrt{2m(E+m)}} \qquad \sinh \frac{\epsilon}{2} = \sqrt{\frac{E-m}{2m}} = \frac{|\vec{k}|}{\sqrt{2m(E+m)}}. \tag{2.198}$$

The boost matrices, for any final momentum k, can be written in standard representation as [29, p. 52]

$$\Lambda(k) \overset{SR}{=} \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 & 0 & \frac{k^3}{E+m} & \frac{k^1 - ik^2}{E+m} \\ 0 & 1 & \frac{k^1 + ik^2}{E+m} & \frac{-k^3}{E+m} \\ \frac{k^3}{E+m} & \frac{k^1 - ik^2}{E+m} & 1 & 0 \\ \frac{k^1 + ik^2}{E+m} & \frac{-k^3}{E+m} & 0 & 1 \end{pmatrix} = \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 & \frac{\vec{k}\cdot\vec{\sigma}}{E+m} \\ \frac{\vec{k}\cdot\vec{\sigma}}{E+m} & 1 \end{pmatrix}, \tag{2.199}$$

while in chiral representation, they can be written as [28, p. 46] [7]

$$\Lambda(k) \overset{CR}{=} \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 - \frac{\vec{k}\cdot\vec{\sigma}}{E+m} & 0 \\ 0 & 1 + \frac{\vec{k}\cdot\vec{\sigma}}{E+m} \end{pmatrix}. \tag{2.200}$$

2.2.10 Parity transformation

What about the action of parity transformations (i.e., space inversion)? The corresponding matrix is

$$P^{\mu}{}_{\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{2.201}$$

As done before with the restricted Lorentz transformations, we have to find the appropriate Dirac transformation that guarantees

$$\Lambda_P^{-1}(i\gamma^{\mu}\partial'_{\mu} - m)\Lambda_P \psi(x) = (i\gamma^{\mu}\partial_{\mu} - m)\psi(x). \tag{2.202}$$

In practice, we need to fulfill the conditions

$$\Lambda_P^{-1}\gamma^0\Lambda_P = \gamma^0, \qquad -\Lambda_P^{-1}\gamma^i\Lambda_P = \gamma^i. \tag{2.203}$$

[7] To see the connection with the notation of Peskin–Schroeder, it is convenient to use relations such as $\sqrt{(E+|k|)} + \sqrt{(E-|k|)} = \sqrt{2(E+m)}$ and $\sqrt{(E+|k|)} - \sqrt{(E-|k|)} = \sqrt{2(E-m)}$.

The solution is

$$\Lambda_P = \kappa\gamma^0 \tag{2.204}$$

with $\kappa = \pm 1, \pm i$. A spinor will change in the following way (leaving the imaginary solutions aside)

$$\psi' = \pm\gamma^0\psi. \tag{2.205}$$

The two opposite signs correspond to spinors with opposite intrinsic parity.

2.2.11 Spinor bilinear expressions

The transformation matrix we obtained to describe the effect of Lorentz transformations on spinors have the property

$$\Lambda^{-1} = \gamma^0\Lambda^\dagger\gamma^0. \tag{2.206}$$

The relation holds also for Λ_P. Note that this property clearly shows that the Λ transformation matrices are not unitary, which is also related to the fact that the $\sigma^{\alpha\beta}$ generators are not Hermitian.

We can explicitly check the above statement for infinitesimal transformations

$$\begin{aligned}\gamma^0\Lambda^{(\alpha\beta)\dagger}\gamma^0 &\approx \gamma^0\Big(\mathbb{1} - \frac{i}{2}\epsilon\sigma^{\alpha\beta}\Big)^\dagger\gamma^0 \\ &= \gamma^0\gamma^0 + \frac{i}{2}\epsilon\gamma^0\sigma^{\alpha\beta\dagger}\gamma^0 \\ &= \mathbb{1} + \frac{i}{2}\epsilon\gamma^0\sigma^{\alpha\beta\dagger}\gamma^0 \\ &= \mathbb{1} + \frac{i}{2}\epsilon\sigma^{\alpha\beta} \approx \Lambda^{(\alpha\beta)-1}\end{aligned} \tag{2.207}$$

In the last step, we used

$$\begin{aligned}\gamma^0\sigma^{\alpha\beta\dagger}\gamma^0 &= -\frac{i}{2}\gamma^0\big[\gamma^{\beta\dagger},\gamma^{\alpha\dagger}\big]\gamma^0 = -\frac{i}{2}\gamma^0\big(\gamma^{\beta\dagger}\gamma^{\alpha\dagger} - \gamma^{\alpha\dagger}\gamma^{\beta\dagger}\big)\gamma^0 \\ &= -\frac{i}{2}\Big(\gamma^0\gamma^{\beta\dagger}\gamma^0\gamma^0\gamma^{\alpha\dagger}\gamma^0 - \gamma^0\gamma^{\alpha\dagger}\gamma^0\gamma^0\gamma^{\beta\dagger}\gamma^0\Big) = -\frac{i}{2}\Big(\gamma^\beta\gamma^\alpha - \gamma^\alpha\gamma^\beta\Big) \\ &= \sigma^{\alpha\beta}\end{aligned} \tag{2.208}$$

We are now in the position to check the nature of the following bilinear combinations under Lorentz transformations

$\overline{\psi}\psi$	Lorentz scalar	(2.209)
$\overline{\psi}\gamma_5\psi$	Lorentz pseudo-scalar	(2.210)
$\overline{\psi}\gamma^\mu\psi$	Lorentz vector	(2.211)
$\overline{\psi}\gamma_5\gamma^\mu\psi$	Lorentz pseudo-vector	(2.212)

For instance, in the first case

$$\overline{\psi}\psi = \psi^\dagger\gamma^0\psi \longrightarrow \psi^\dagger\Lambda^\dagger\gamma^0\Lambda\psi = \psi^\dagger\gamma^0\gamma^0\Lambda^\dagger\gamma^0\Lambda\psi = \overline{\psi}\Lambda^{-1}\Lambda\psi = \overline{\psi}\psi. \tag{2.213}$$

The above relation works also with Λ_P.

The pseudoscalar combination works in the same way, since Λ contains an even number of γ matrices and therefore commutes with γ_5

$$\overline{\psi}\gamma_5\psi \longrightarrow \psi^\dagger\Lambda^\dagger\gamma^0\gamma_5\Lambda\psi = \psi^\dagger\gamma^0\gamma^0\Lambda^\dagger\gamma^0\gamma_5\Lambda\psi = \overline{\psi}\Lambda^{-1}\gamma_5\Lambda\psi = \overline{\psi}\gamma_5\psi. \tag{2.214}$$

but for parity transformations there is a difference, since $\Lambda_P = \gamma_0$ and therefore anticommutes with the γ_5

$$\overline{\psi}\gamma_5\psi \overset{\text{parity}}{\longrightarrow} \psi^\dagger\Lambda_P^\dagger\gamma^0\gamma_5\Lambda_P\psi = \psi^\dagger\gamma^0\gamma^0\Lambda_P^\dagger\gamma^0\gamma_5\Lambda_P\psi = \overline{\psi}\Lambda_P^{-1}\gamma_5\Lambda_P\psi = -\overline{\psi}\gamma_5\psi. \tag{2.215}$$

Checking the vector combination is relatively easy, once we know (2.180)

$$\overline{\psi}\gamma^\mu\psi \longrightarrow \psi^\dagger\Lambda^\dagger\gamma^0\gamma^\mu\Lambda\psi = \psi^\dagger\gamma^0\gamma^0\Lambda^\dagger\gamma^0\gamma^\mu\Lambda\psi = \overline{\psi}\Lambda^{-1}\gamma^\mu\Lambda\psi = L^\mu{}_\nu\overline{\psi}\gamma^\nu\psi. \tag{2.216}$$

We are now in the position of justifying why we used exactly this kind of combination to introduce the concept of a Dirac four-current in Sec. 2.2.5.

2.2.12 Plane-wave solutions

We know that the Dirac equation possesses four plane-wave solutions, two with positive and two with negative energy. We now want to obtain the explicit expression for these solutions in the chiral or standard representation.

We first focus on the positive-energy solutions. We can study the solutions in the rest-frame of the particle, where $k_{\text{RF}} = (m, 0, 0, 0)$, and then boost to a frame where the particle has a generic momentum k, using the explicit expression for boosts in Dirac space.

The expression for our positive-energy plane-wave solution is (for the moment we do not worry about the normalization)

$$\psi(x) = u(k)e^{-ikx} \tag{2.217}$$

The Dirac equation implies for the spinor u

$$(\not{k} - m)u = 0 \tag{2.218}$$

In the CM frame this becomes simply

$$(m\gamma^0 - m)u(0) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2m & 0 \\ 0 & 0 & 0 & -2m \end{pmatrix} u(0) = 0. \tag{2.219}$$

The solutions can be written as

$$u_1(0) \overset{SR}{=} \sqrt{2m}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \qquad u_2(0) \overset{SR}{=} \sqrt{2m}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}. \tag{2.220}$$

The specific choice of spinors is arbitrary: in this case we have taken eigenvectors of σ^3, but we could have chosen any generic direction. Careful: there are two different common choices for the normalization of the spinors. Here, we follow the conventions of Peskin–Schroeder, although other books, e.g., Mandl–Shaw and Ryder, prefer to normalize the spinor to 1, i.e., they have an extra $1/\sqrt{2m}$ in front.

We can redo the calculation in chiral representation, or just apply the unitary transformation (2.101) to change from one to the other. In any case

$$u_1(0) \overset{CR}{=} \sqrt{m}\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \qquad u_2(0) \overset{CR}{=} \sqrt{m}\begin{pmatrix}0\\1\\0\\1\end{pmatrix}. \tag{2.221}$$

Looking at Eq. (2.199), we realize that boosting a u spinors in standard representation (not in chiral representation) is the same as applying these transformations

$$u(\vec{k}) \overset{SR}{=} \frac{(\not{k}+m)}{\sqrt{2m(E+m)}}u(0), \qquad \overline{u}(\vec{k}) \overset{SR}{=} \overline{u}(0)\frac{(\not{k}+m)}{\sqrt{2m(E+m)}}. \tag{2.222}$$

With these ingredients, we can then compute the explicit form of a $u(\vec{k})$ spinor

$$u_1(\vec{k}) \overset{SR}{=} \frac{1}{\sqrt{E+m}}\begin{pmatrix}E+m\\0\\k^3\\k^1+ik^2\end{pmatrix}, \qquad u_2(\vec{k}) \overset{SR}{=} \frac{1}{\sqrt{E+m}}\begin{pmatrix}0\\E+m\\k^1-ik^2\\-k^3\end{pmatrix}. \tag{2.223}$$

Note that often the spinor is written as a function of the four-momentum k. However, due to the on-shell condition $k^2 = m^2$, in reality the spinor can be seen as depending only on the three-momentum $\vec{k}$.

We can do similar steps for the chiral representation and obtain

$$u_1(\vec{k}) \overset{CR}{=} \frac{1}{\sqrt{2(E+m)}}\begin{pmatrix}E+m-k^3\\-k^1-ik^2\\E+m+k^3\\k^1+ik^2\end{pmatrix}, \quad u_2(\vec{k}) \overset{CR}{=} \frac{1}{\sqrt{2(E+m)}}\begin{pmatrix}-k^1+ik^2\\E+m+k^3\\k^1-ik^2\\E+m-k^3\end{pmatrix}. \tag{2.224}$$

The representation above can be useful to discuss the so-called *ultrarelativistic* limit, where masses are neglected and thus $|k| = E$. If we consider a particle with momentum in the z direction, we obtain

$$u_1(\vec{k}) \overset{CR}{\approx} \sqrt{2E}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \qquad u_2(\vec{k}) \overset{CR}{\approx} \sqrt{2E}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}. \tag{2.225}$$

The above expressions describe the result of boosting with a momentum k in a generic direction a spinor which has spin oriented along z in its rest frame. We could build other expressions where the spin is oriented along the direction of momentum.

For the negative-energy solutions, we choose the following strategy [32]. We want to describe a plane-wave solution of the Dirac equation with $k^2 = m^2$ and $k^0 < 0$. If we choose $p = -k$, we need a solution with $p^2 = m^2$ and $p^0 > 0$. In practice, instead of having a negative-energy solution moving with momentum $\vec{k}$, we look for a positive-energy solution moving with momentum $-\vec{k}$. Our trial solution is then

$$\psi(x) = u(\vec{k})e^{-ikx} = v(\vec{p})e^{+ipx} \tag{2.226}$$

The spinor v has to solve the equation

$$(\not{p} + m)v(\vec{p}) = 0 \tag{2.227}$$

(note the sign change for the mass term).

To solve this equation, we use a trick. We make use of a matrix with the following properties

$$C^{-1}\gamma^\mu C = -\gamma^{\mu *}, \tag{2.228}$$

explicitly, in both our standard and chiral representations C corresponds to

$$C = -i\gamma^2 \tag{2.229}$$

with the properties

$$C = C^{-1} = C^\dagger. \tag{2.230}$$

We can check that

$$v(\vec{p}) = Cu(\vec{p})^* \tag{2.231}$$

is indeed a solution of the Dirac equation for the negative-energy states

$$\not{p}v(\vec{p}) = CC^{-1}p_\mu\gamma^\mu Cu(\vec{p})^* = C\left[-\not{p}u(\vec{p})\right]^* = C\left[-mu(\vec{p})^*\right] = -mv(\vec{p}). \tag{2.232}$$

Therefore, to obtain the v spinors that represent a solution of the Dirac equation for negative-energy states, we simply take the positive-energy solutions and apply the tranformation (2.231). This transformation is called charge conjugation because it turns particles into antiparticles. To uphold this interpretation of the charge conjugation transformation, we can also note that if ψ solves the Dirac equation with coupling to an elm field,

Eq. (2.119), then $\psi_C = C\psi^*$ solves the equation

$$\left[i\left(\partial\!\!\!/ + ieA\!\!\!/\right) - m\right]\psi_C = 0, \tag{2.233}$$

i.e., the Dirac equation with coupling to an elm field, but with an opposite charge.

The rest-frame solutions look like

$$v_1(0) \stackrel{SR}{=} \sqrt{2m}\begin{pmatrix}0\\0\\0\\-1\end{pmatrix}, \qquad v_2(0) \stackrel{SR}{=} \sqrt{2m}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}. \tag{2.234}$$

The rest-frame solutions are orthogonal and normalized to $2m$, i.e.,

$$u_i^\dagger(0)u_j(0) = 2m\,\delta_{ij}, \qquad v_i^\dagger(0)v_j(0) = 2m\,\delta_{ij}, \qquad u_i(0)^\dagger v_j(0) = 0. \tag{2.235}$$

We can obtain the solutions for a momentum p by boosting or by applying charge conjugation to the positive-energy solutions. In any case, the final outcome is

$$v_1(\vec{p}) \stackrel{SR}{=} -\frac{1}{\sqrt{E+m}}\begin{pmatrix}p^1 - ip^2\\-p^3\\0\\E+m\end{pmatrix}, \qquad v_2(\vec{p}) \stackrel{SR}{=} \frac{1}{\sqrt{E+m}}\begin{pmatrix}p^3\\p^1+ip^2\\E+m\\0\end{pmatrix}. \tag{2.236}$$

Note that boosting a v spinors in standard representation (not in chiral representation) is the same as acting with

$$v(\vec{p}) \stackrel{SR}{=} \frac{(-p\!\!\!/ + m)}{\sqrt{2m(E+m)}}v(0), \qquad \overline{v}(\vec{p}) \stackrel{SR}{=} \overline{v}(0)\frac{(-p\!\!\!/ + m)}{\sqrt{2m(E+m)}}. \tag{2.237}$$

With these results in our hands, we can check several properties of the spinors. First of all, we have already shown that combinations $\overline{u}u$ are Lorentz scalars, therefore we can compute them in the rest frame and obtain

$$\overline{u}_i(\vec{k})u_j(\vec{k}) = \overline{u}_i(0)u_j(0) = u_i^\dagger(0)\gamma^0 u_j(0) = 2m\,\delta_{ij}, \tag{2.238}$$

$$\overline{v}_i(\vec{k})v_j(\vec{k}) = \overline{v}_i(0)v_j(0) = v_i^\dagger(0)\gamma^0 v_j(0) = -2m\,\delta_{ij}, \tag{2.239}$$

$$\overline{u}_i(\vec{k})v_j(\vec{k}) = 0. \tag{2.240}$$

Note that these combinations, where the Dirac indices are contracted, must give the same result in any representation.

Next, we can also compute combinations of the type $u^\dagger u$ (and check that they are not scalars, but rather correspond to the zeroth component of a four-vector). Again, we can

choose any representation to do the computation. For instance, we work in standard representation and use Eq. (2.222) to apply a boost. We will need the relation

$$\begin{aligned}(\not{k}+m)\gamma^0(\not{k}+m) &= (2k^0-\gamma^0\not{k}+\gamma^0 m)(\not{k}+m)\\ &= 2k^0(\not{k}+m)+\gamma^0(m^2-k^2) = 2k^0(\not{k}+m)\end{aligned} \tag{2.241}$$

and also

$$\overline{u}_r(0)(\not{k}+m)u_s(0) \overset{SR}{=} \overline{u}_r(0)(k^0\gamma^0+m)u_s(0) = 2m(E+m)\delta_{rs} \tag{2.242}$$

where the last step can be conveniently computed in standard representation. We can obtain

$$\begin{aligned}u_r^\dagger(\vec{k})u_s(\vec{k}) &= \frac{1}{2m(E+m)}\Big((\not{k}+m)u_r(0)\Big)^\dagger(\not{k}+m)u_s(0)\\ &= \frac{1}{2m(E+m)}u_r(0)^\dagger(\not{k}+m)^\dagger(\not{k}+m)u_s(0)\\ &= \frac{1}{2m(E+m)}u_r(0)^\dagger\gamma^0\gamma^0(\not{k}+m)^\dagger\gamma^0\gamma^0(\not{k}+m)u_s(0)\\ &= \frac{1}{2m(E+m)}\overline{u}_r(0)(\not{k}+m)\gamma^0(\not{k}+m)u_s(0)\\ &= \frac{1}{2m(E+m)}2k^0\overline{u}_r(0)(\not{k}+m)u_s(0) = 2k^0\delta_{rs}.\end{aligned} \tag{2.243}$$

For the v spinors, we can either check the computation in a similar way, or use charge conjugation to show that

$$v_r^\dagger(\vec{k})v_s(\vec{k}) = u_r^{\dagger *}(\vec{k})C^\dagger C u_s(\vec{k})^* = \big(u_r^\dagger(\vec{k})u_s(\vec{k})\big)^* = 2k^0\delta_{rs}. \tag{2.244}$$

We can also show that

$$u_r^\dagger(\vec{k})v_s(-\vec{k}) = v_r^\dagger(\vec{k})u_s(-\vec{k}) = 0. \tag{2.245}$$

More generally, we can obtain the expression for $\overline{u}\gamma^\mu u$ using the effect of a Lorentz boost

$$\begin{aligned}\overline{u}_r(\vec{k})\gamma^\mu u_s(\vec{k}) &= \overline{u}_r(0)\Lambda^{-1}(k)\gamma^\mu\Lambda(k)u_s(0)\\ &= L^\mu{}_\nu\overline{u}_r(0)\gamma^\nu u_s(0)\\ &= L^\mu{}_\nu\overline{u}_r(0)\frac{k^\nu_{\mathrm{RF}}}{m}u_s(0) = 2k^\mu\delta_{rs}.\end{aligned} \tag{2.246}$$

which is a vector as expected. In the last line, we made use of the fact that we can compute explicitly the expression $\overline{u}_r(0)\gamma^\mu u_s(0)$ in the rest frame: the γ^0 term gives the same result as Eq. 2.243, the γ^i terms give 0, corresponding to the components of k_{RF} The result for $\overline{v}\gamma^\mu v$ is identical.

It is sometimes useful to know the equations satisfied by the barred versions of the spinors. We remind ourselves that the standard spinors fulfill

$$(\not{k}-m)u = 0, \qquad (\not{k}+m)v = 0. \tag{2.247}$$

For the barred spinors we can obtain (it is sufficent to take the barred version of the previous equations):

$$\overline{u}(\not{k} - m) = 0, \qquad \overline{v}(\not{k} + m) = 0. \tag{2.248}$$

Last but not least, we need to introduce the so-called positive and negative energy projectors. They are defined as

$$\Lambda^+_{AB} = \sum_r u_{rA}(\vec{k})\overline{u}_{rB}(\vec{k}), \qquad \Lambda^-_{AB} = -\sum_r v_{rA}(\vec{k})\overline{v}_{rB}(\vec{k}). \tag{2.249}$$

They are Dirac matrices (outer products of two Dirac spinors).

We can check that they are projectors since

$$\left(\Lambda^+\right)^2 = \sum_r \sum_s u_r(\vec{k})\overline{u}_r(\vec{k})u_s(\vec{k})\overline{u}_s(\vec{k}) = 2m\sum_r \sum_s u_r(\vec{k})\delta_{rs}\overline{u}_s(\vec{k}) = 2m\Lambda^+, \tag{2.250}$$

$$\left(\Lambda^-\right)^2 = \sum_r \sum_s v_r(\vec{k})\overline{v}_r(\vec{k})v_s(\vec{k})\overline{v}_s(\vec{k}) = -2m\sum_r \sum_s v_r(\vec{k})\delta_{rs}\overline{v}_s(\vec{k}) = 2m\Lambda^-. \tag{2.251}$$

We can check that the projectors are orthogonal and complete

$$\Lambda^\pm_{AB}\Lambda^\mp_{AB} = 0, \qquad \Lambda^+_{AB} + \Lambda^-_{AB} = 2m\mathbb{1}_{AB}. \tag{2.252}$$

Due to Dirac's equation, we know that

$$(\not{k} - m)\Lambda^+ = 0 \qquad \Rightarrow \qquad \Lambda^+ = A(\not{k} + m), \tag{2.253}$$

$$(\not{k} + m)\Lambda^- = 0 \qquad \Rightarrow \qquad \Lambda^- = B(\not{k} - m). \tag{2.254}$$

To fulfill Eqs. (2.250), (2.251), (2.252), we need $A = 1$ and $B = -1$.

In conclusion, the result is

$$\Lambda^+ = \sum_r u_r(\vec{k})\overline{u}_r(\vec{k}) = \not{k} + m, \tag{2.255}$$

$$\Lambda^- = -\sum_r v_r(\vec{k})\overline{v}_r(\vec{k}) = -\not{k} + m. \tag{2.256}$$

2.2.13 Spin and helicity [optional]

In our analysis of the Dirac equation, we have identified two degenerate positive-energy and two negative-energy states. If we confine ourselves to the rest-frame of the particle (which is possible only for massive particles), we can introduce the Pauli spin operator (see also Eq. (2.191))

$$\Sigma_k = \frac{1}{4}\epsilon_{ijk}\sigma^{ij} \overset{SR}{=} \frac{1}{2}\begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix}. \tag{2.257}$$

The last expression is valid also in chiral representation, see Eq. (2.189). Spinors can be eigenvectors of this operator with eigenvalues $+1$ and -1. We have seen that the two states correspond to different energies in the presence of a magnetic field, which leads us

to the identification of the two states as "spin up" and "spin down." The operators have the property

$$\Big[\Sigma_i, \Sigma_j] = i\epsilon_{ijk}\Sigma_k \tag{2.258}$$

which is what we expect for a spin-1/2 operator.

This discussion is limited to the rest frame of the particle. In spite of the fact that spin is essentially tied to relativistic invariance, it is nontrivial to build a spin four-vector. In order to distinguish the two independent states for u and v spinors in any frame, the most common choice is to introduce the helicity operator of a particle with momentum $\vec{k}$, defined as

$$\text{helicity} = h(\vec{k}) = \hat{k} \cdot \vec{\Sigma} \tag{2.259}$$

where $\hat{k}$ denotes the unit vector in the direction of the particle's momentum. In other words, helicity corresponds to the projection of "spin" along the direction of motion.

The helicity operators commutes with the Hamiltonian $\vec{a} \cdot \vec{k} + \beta m$. Let us consider the spinors in the form of Eqs. (2.223) and (2.236) and a momentum $\vec{k}_z$ with $k_1 = k_2 = 0$. Applying the helicity operator we can explicitly check in standard representation that

$$h(\vec{k}_z)u_1(\vec{k}_z) = u_1(\vec{k}_z), \qquad h(\vec{k}_z)u_2(\vec{k}_z) = -u_2(\vec{k}_z), \tag{2.260}$$

$$h(\vec{k}_z)v_1(\vec{k}_z) = -v_1(\vec{k}_z), \qquad h(\vec{k}_z)v_2(\vec{k}_z) = v_2(\vec{k}_z). \tag{2.261}$$

Our choice of spinors 1 and 2 corresponds to helicity eignestates if the momentum of the particle is in the z direction. It is possible to use spinors with helicity along a generic k. The only crucial feature is that there are obviously always two degenerate states.

There is still an asymmetry in the labeling of u and v spinors for what concerns the different sign in the helicity eigenvalue. The final justification will be given in Sec. 3.5.3.

2.2.14 Feynman interpretation of negative-energy solutions [optional]

Before closing this chapter, let us discuss again the question of negative-energy solutions (see [3]). Although the formalism of Quantum Field Theory better clarifies the issue, it is possible to introduce an interpretation of the negative-energy solutions, due to Feynman, that to a certain extent provides a bridge between single-particle wave equations and quantum fields.

This interpretation is more easily discussed in the Klein–Gordon case. We have found that we can have solutions of this form

$$\phi^{\pm}(x) = e^{\mp i\omega_k t + i\vec{k}\cdot\vec{x}} \tag{2.262}$$

with positive and negative energies. We could say that the negative energy solutions are equivalent to positive energy solutions propagating backward in time,

$$\phi^{-}(x) = e^{-i\omega_k(-t) + i\vec{k}\cdot\vec{x}}, \tag{2.263}$$

but it is not clear if this has any physical meaning. Suppose, however, that we take a positive-energy solution and take its complex conjugate

$$\phi^{+*}(x) = e^{i\omega_k t - i\vec{k}\cdot\vec{x}}, \tag{2.264}$$

This is equivalent to a positive energy solution, propagating backward in time and with opposite $\vec{k}$. This solution has a physical meaning. The Klein–Gordon equation coupled to an elm field is

$$\Big[(\partial^\mu + iqA^\mu)(\partial_\mu + iqA_\mu) + m^2\Big]\phi(x) = 0. \tag{2.265}$$

Taking the complex conjugate we obtain

$$\Big[(\partial^\mu - iqA^\mu)(\partial_\mu - iqA_\mu) + m^2\Big]\phi^*(x) = 0. \tag{2.266}$$

Therefore, ϕ^* describes an *antiparticle*, i.e., the same as a particle but with charge $(-q)$.

This observation can be generalized to the effect of negative-energy solutions on any physical process and leads to the following rule

> A negative-energy Klein–Gordon particle with momentum k^μ propagating forward in time is equivalent to a positive-energy Klein–Gordon antiparticle with momentum $-k^\mu$ propagating backward in time.

This means that instead of working with negative-energy particles, we can conveniently switch to positive-energy antiparticles. Suppose a system absorbs a negative-energy particle with (positive) charge q and momentum $k^\mu = (-\omega_k, \vec{k})$, its charge increases, its energy *decreases*, its momentum increases. This is the same effect as emitting a positive-energy antiparticle with momentum $(\omega_k, -\vec{k})$.

This leads also to a clarification of the problem of the the Klein–Gordon current, definded in Eq. (2.55). Instead of interpreting it as a "probability current," let us interpret it as a charge current, multiplying it by the charge of the particle, i.e.,

$$j^\mu(x) = \frac{iq}{2m}\Big(\phi^*(x)\partial^\mu\phi(x) - \phi(x)\partial^\mu\phi^*(x)\Big) = q\frac{k^\mu}{m}\phi^*(x)\phi(x). \tag{2.267}$$

With this simple change, we can stop worrying about the presence of negative densities, since we are now talking about charge densities, not probability densities.

Not only. We can also observe that the current generated by negative-energy solutions with $k^\mu = (-\omega_k, \vec{k})$ is equal to the current generated by a particle with $p^\mu = -k^\mu = (\omega_k, -\vec{k})$, but with charge $-q$, i.e., the current of an antiparticle with positive energy and momentum $-\vec{k}$.

For fermions, the situation is less straightforward because of the presence of spin and because the current density was positive also for the negative-energy solutions.

Similar to the Klein–Gordon case, we note that the solutions were of the form

$$\psi^+(x) = u_{1,2}(k)e^{-i\omega_k t + i\vec{k}\cdot\vec{x}} \tag{2.268}$$

$$\psi^-(x) = u_{3,4}(k)e^{i\omega_k t + i\vec{k}\cdot\vec{x}} \tag{2.269}$$

with positive and negative energies. The negative energy solutions are equivalent to positive energy solutions propagating backward in time

$$\psi^{-}(x) = u_{3,4}(\vec{k})e^{-i\omega_k(-t)+i\vec{k}\cdot\vec{x}}. \tag{2.270}$$

To match this into something with physical meaning, we consider $C\psi^*$, which solves the Dirac equation for a particle with an opposite charge (and opposite spin), i.e., it describes an antiparticle. Therefore, we consider the function

$$C\psi^{+*}(x) = v_{1,2}(\vec{k})e^{i\omega_k t-i\vec{k}\cdot\vec{x}}. \tag{2.271}$$

which describes an antiparticle with momentum $\vec{k}$ and a certain spin in the rest frame. This is equivalent to a positive-energy solution, propagating backward in time, with momentum $-\vec{k}$, and with opposite spin in the rest frame.

Feynman's hypothesis can in this case can be phrased as follows

> A negative-energy fermion with momentum k^μ and spin projection s_z propagating forward in time is equivalent to a positive-energy anti-fermion with momentum $-k^\mu$ and spin projection $-s_z$ propagating backward in time.

There is only a further subtlety: for fermions, sometimes we must take care of extra minus signs due to the fact that they must obey antisymmetrization rules. This should become clear only later in the course. This has an effect on our definition of the current. The charge current generated by positive energies is

$$j^\mu = q\bar{u}(\vec{k})\gamma^\mu u(\vec{k}) = 2q\left(\omega_k, \quad \vec{k}\right) \tag{2.272}$$

Adding the extra minus sign due to anticommutation rules, the current generated by negative energies is

$$j^\mu = -q\bar{v}(-\vec{k})\gamma^\mu v(-\vec{k}) = -2q\left(\omega_k, \quad -\vec{k}\right) \tag{2.273}$$

which corresponds to the current of an antiparticle with momentum $-\vec{k}$, as for the Klein–Gordon case.

2.3 Conclusions

In this chapter, we have dealt with relativistic wave equations that try to generalize the nonrelativistic Schrödinger equation. In the next chapter we will see that the formalism of field equations is more appropriate to deal with relativistic quantum mechanics.

Typical questions that can come out during the exam:

1. Effects of boosts on Dirac spinors;
2. Prove that $\overline{u}_r u_s$ is a scalar and compute it;
3. Discuss the positive or negative energy plane-wave solutions of the Dirac equation;
4. Compute the positive or negative energy projectors.

3

Free quantum fields

3.1 Classical field theory

Apart from the lecture notes of prof. Miglietta, the topics covered in this section are based on Goldstein [20], Ch. 11, and Mandl–Shaw [26], Ch. 2.

3.1.1 The principle of least action

We remember that the action is defined as the time integral of the Lagrangian

$$S \equiv \int_{t_0}^{t_1} dt\; L(q, \dot{q}; t). \tag{3.1}$$

The action depends on the "path" chosen to go from a state at time t_0 to another state at time t_1. Here path means a function $q(t)$, i.e., the time dependence of the generalized coordinate q. We can define an action for any of these paths. Technically, we say that the action is a functional of the path (i.e., it is a function that depends on another function). We can compute what is the difference in the action for a variation of a given path, keeping the starting and ending points fixed. For instance, if the path is described by some function $q(t)$, we can take

$$q'(t) = q(t) + \delta q(t) \qquad \text{with} \quad \delta q(t_0) = \delta q(t_1) = 0. \tag{3.2}$$

The corresponding change in the action is called a "functional derivative" and is mathematically indicated as $\delta S[q]/\delta q(t)$.

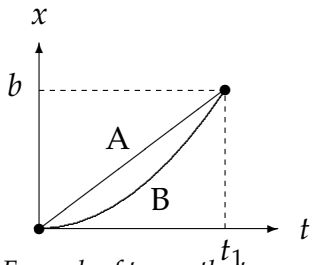


Figure 3.1: Example of two paths to compute the action

In classical mechanics, the path that the system follows is the one that corresponds to a stationary action ("least-action principle"). This means that if we try to compute the action for a different path, obtained by some small variation of the first, we should get a vanishing variation of the action. In other words, the action for the actual path is a minimum (or maximum). In mathematical notation this is written as

$$\frac{\delta S[q]}{\delta q(t)} = 0. \tag{3.3}$$

For our purposes, the functional derivative can be defined as

$$\frac{\delta G[f(x')]}{\delta f(x)} = \lim_{\epsilon \to 0} \frac{G[f(x') + \epsilon \delta(x - x')] - G[f(x')]}{\epsilon}. \tag{3.4}$$

The condition of stationary action leads to (assuming there is no explicit dependence of the Lagrangian on time)

$$\begin{aligned} \delta S &= \int_{t_0}^{t_1} dt \left\{ \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right\} = \int_{t_0}^{t_1} dt \left\{ \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} d_t(\delta q) \right\} \\ &= \int_{t_0}^{t_1} dt \; \delta q \left\{ \frac{\partial L}{\partial q} - d_t \frac{\partial L}{\partial \dot{q}} \right\} + \int_{t_0}^{t_1} dt \; d_t \left(\frac{\partial L}{\partial \dot{q}} \delta q \right) \\ &= \int_{t_0}^{t_1} dt \; \delta q \left\{ \frac{\partial L}{\partial q} - d_t \frac{\partial L}{\partial \dot{q}} \right\} + \cancel{\left(\frac{\partial L}{\partial \dot{q}} \delta q \right) \Big|_{t_1}} - \cancel{\left(\frac{\partial L}{\partial \dot{q}} \delta q \right) \Big|_{t_0}}. \end{aligned} \tag{3.5}$$

In conclusion, to have $\delta S = 0$ for any variation δq, we need to impose the Euler–Lagrange equations (or equations of motion):

$$\frac{\partial L}{\partial q_i} - d_t \frac{\partial L}{\partial \dot{q}_i} = 0. \tag{3.6}$$

Let us make a super-simple example. Suppose we have a free particle with mass $m = 2$ that moves along the x coordinate from $x_0 = 0$ at $t_0 = 0$ to $x_1 = b$ at t_1. Our generalized coordinate corresponds in this case to the standard coordinate $q = x$. The Lagrangian is

$$L = \dot{q}^2. \tag{3.7}$$

We try two different paths A and B: a straight line and a parabolic line. Their equations are

$$q_A(t) = \frac{b}{t_1} t, \qquad\qquad q_B(t) = \frac{b}{t_1^2} t^2, \tag{3.8}$$

and their derivatives

$$\dot{q}_A(t) = \frac{b}{t_1}, \qquad \dot{q}_B(t) = 2\frac{b}{t_1^2}t. \tag{3.9}$$

We obtain for the respective actions

$$S_A = \frac{b^2}{t_1^2}\int_0^{t_1} dt = \frac{b^2}{t_1}, \qquad S_B = 4\frac{b^2}{t_1^4}\int_0^{t_1} dt\, t^2 = \frac{4}{3}\frac{b^2}{t_1}. \tag{3.10}$$

We checked that path A leads to a smaller action than path B. We can compute what is the path implied by the Euler–Lagrange equations:

$$\frac{\partial L}{\partial q} = 0, \qquad d_t\frac{\partial L}{\partial \dot{q}} = d_t(2\dot{q}) = 2\ddot{q}. \tag{3.11}$$

Not surprisingly, the principle of least action tells us that $\ddot{q} = 0$ (corresponding to the fact that this system has no acceleration), which integrates to

$$q(t) = \frac{b}{t_1}t \tag{3.12}$$

if we want to fulfill the initial and final conditions.

The principle of least action can be used as the fundamental postulate to derive all classical mechanics. Feynman extended the use of the principle as the fundamental starting point of his formulation of Quantum Mechanics based on the "path-integral formalism." This particular formalism will turn out to be useful to study Quantum Field Theory, but will not be used in this introductory course.

3.1.2 Hamiltonian

Let us briefly review the concept of Hamiltonian. We start by taking the derivative of the Lagrangian with respect to time

$$d_t L = \frac{\partial L}{\partial q}d_t q + \frac{\partial L}{\partial \dot{q}}d_t\dot{q} + \partial_t L \tag{3.13}$$

If the Lagrangian does not depend explicitly on time (i.e., any time translation leaves it unchanged), we have

$$\partial_t L = 0, \tag{3.14}$$

and from the equations of motion

$$\frac{\partial L}{\partial q} = d_t\frac{\partial L}{\partial \dot{q}}. \tag{3.15}$$

In conclusion

$$d_t L = d_t\left(\frac{\partial L}{\partial \dot{q}}\dot{q}\right) \Rightarrow d_t(p\dot{q} - L) = d_t H = 0. \tag{3.16}$$

From this result we conclude that the Hamiltonian is conserved. This is one particular example of the general result of Nöther's theorem, which we will review more completely in Sec. 3.1.4. The Hamiltonian is the quantity that is conserved if the Lagrangian is independent of time translations and corresponds physically to the total energy of the system.

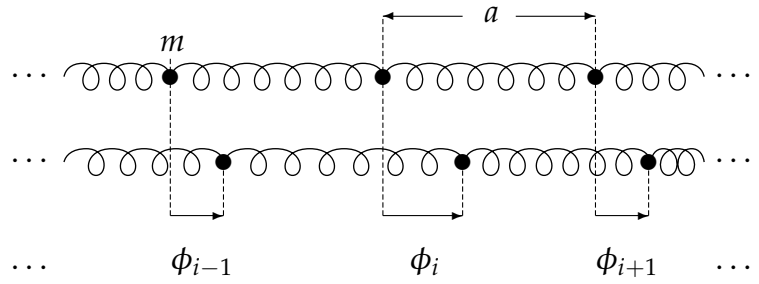


Figure 3.2: Schematic drawing of a discrete system of n particles with equal mass m connected by springs with length a (and elastic constant κ). The variables ϕ_i denote the displacements (drawn at the bottom) from the equilibrium position (drawn at the top).

3.1.3 From discrete to continuous systems

We start with considering a system of n particles with equal mass m connected by springs with length a and elastic constant κ that can move only in the longitudinal direction. With ϕ_i we denote the displacement of the particle from the equilibrium position. The kinetic energy of each particle can be written as

$$T_i = \frac{1}{2} m (d_t \phi_i)^2 \tag{3.17}$$

The total kinetic energy of the system is obviously

$$T = \frac{1}{2} m \sum_i (d_t \phi_i)^2 \tag{3.18}$$

The potential energy of the single particle can be written as

$$V_i = \frac{1}{2} \kappa (\phi_{i+1} - \phi_i)^2. \tag{3.19}$$

The total potential energy is

$$V = \frac{1}{2} \kappa \sum_i (\phi_{i+1} - \phi_i)^2 = \ldots + \frac{1}{2} \kappa (\phi_i - \phi_{i-1})^2 + \frac{1}{2} \kappa (\phi_{i+1} - \phi_i)^2 + \ldots \tag{3.20}$$

The force acting on particle i can be derived by $F_i = -\partial V / \partial \phi_i$. We can obtain

$$F = -\kappa(\phi_i - \phi_{i-1}) + \kappa(\phi_{i+1} - \phi_i), \tag{3.21}$$

which corresponds to what we would expect from the balance of the forces of the two springs acting on particle i.

The Lagrangian of the system is

$$\begin{aligned} L \equiv T - V &= \frac{1}{2} a \sum_i \Big[\frac{m}{a} (d_t \phi_i)^2 - \kappa a \Big(\frac{\phi_{i+1} - \phi_i}{a} \Big)^2 \Big] \\ &= \frac{1}{2} a \sum_i \Big[\mu \dot{\phi}_i^2 - Y \Big(\frac{\phi_{i+1} - \phi_i}{a} \Big)^2 \Big]. \end{aligned} \tag{3.22}$$

We have in mind to go to the continuum limit by taking $a \to 0$ with μ (mass per unit length) and Y (Young modulus) constant.

The Euler–Lagrange equations are

$$\frac{\partial L}{\partial q_i} - d_t \frac{\partial L}{\partial \dot{q}_i} = 0. \tag{3.23}$$

Applied to our system, we obtain (pay attention to the fact that the coordinate ϕ_i occurs in the term i and $(i-1)$ of the total Lagrangian)

$$\begin{aligned} -d_t \frac{\partial L}{\partial \dot{\phi}_i} &= -\frac{a}{2}\, d_t\big(2\mu\dot{\phi}_i\big) = -\frac{a}{2}\, 2\mu\ddot{\phi}_i, \\ \frac{\partial L}{\partial \phi_i} &= \frac{a}{2}\frac{\partial}{\partial \phi_i}\Big[-Y\Big(\frac{\phi_{i+1}-\phi_i}{a}\Big)^2 - Y\Big(\frac{\phi_i-\phi_{i-1}}{a}\Big)^2\Big] \\ &= \frac{a}{2}\Big[2Y\frac{\phi_{i+1}-\phi_i}{a^2} - 2Y\frac{\phi_i-\phi_{i-1}}{a^2}\Big] \end{aligned} \tag{3.24}$$

Now we go to the continuum limit. In this case, the generalized coordinates ϕ_i become a continuous function of the position x, therefore

$$\phi_i(t) \to \phi(t,x), \qquad \frac{\phi_{i+1}-\phi_i}{a} = \partial_x \phi(t,x) \tag{3.25}$$

and

$$\frac{\frac{\phi_{i+1}-\phi_i}{a} - \frac{\phi_i-\phi_{i-1}}{a}}{a} = \partial_x^2 \phi(t,x). \tag{3.26}$$

The function ϕ are what we call "fields."

The Euler–Lagrange equation becomes

$$\mu\, \partial_t^2 \phi(t,x) - Y\, \partial_x^2 \phi(t,x) = 0. \tag{3.27}$$

The discrete system was characterized by a large numbers of equations, depending only on t. Now we have a single equation, but with partial derivatives. The Lagrangian becomes

$$L = \int dx \frac{1}{2}\Big(\mu\, (\partial_t \phi(t,x))^2 - Y\, (\partial_x \phi(t,x))^2\Big) \equiv \int dx \mathcal{L}. \tag{3.28}$$

The integrand is called the "Lagrangian density."

With three spatial dimensions, we would have had to start with three independent displacements (one longitudinal, two transverse). We could have labeled them with an index r, i.e., $\phi_{ri}(t)$. This is often referred to as a polarization index. In the continuum limit, the set of coordinates becomes a continuous function of the position, i.e.,

$$\phi_{r\,i}(t) \to \phi_r(t, \vec{x}). \tag{3.29}$$

The Lagrangian in this case becomes

$$L = \int d^3x\, \mathcal{L} \tag{3.30}$$

where the Lagrangian density can in general depend on ϕ_r,$\partial_\mu\phi_r$, x, where ϕ_r denote now a generic set of fields and x is four-dimensional (the fact that we use a four-dimensional notation for x does not mean that the theory is relativistic).

We now turn the attention to the action. For a continuous system, it becomes

$$S \equiv \int_\Omega d^4x\, \mathcal{L}(\phi_r, \partial_\mu\phi_r; x). \tag{3.31}$$

Here, Ω denotes a region of space and time and Γ its boundary.

Note that in natural units the action has no dimensions, d^4x has dimensions $[M]^{-4}$ (see Eqs. (2.33) and (2.34)). Therefore the Lagrangian density must have dimensions $[M]^4$. This is consistent also with the fact that in c.g.s. units the Lagrangian density is expressed in $\mathrm{J/m^3}$, which means that in n.u. it can be expressed as J^4.

The principle of least action in this case can be formulated as follows: δS must be zero under a transformation of the fields

$$\phi'_r(x) = \phi_r(x) + \delta\phi_r(x) \qquad \text{with} \quad \delta\phi_r(x)\big|_\Gamma = 0. \tag{3.32}$$

An extension of the standard calculation for the discrete case leads to the Euler–Lagrange equations for the fields (we assume that there is no explicit dependence on x)

$$\begin{aligned}
\delta S &= \int_\Omega d^4x\, \delta\mathcal{L} \\
&= \int_\Omega d^4x \left\{ \frac{\partial\mathcal{L}}{\partial\phi_r}\delta\phi_r + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\partial_\mu(\delta\phi_r) \right\} \\
&= \int_\Omega d^4x\, \delta\phi_r \left\{ \frac{\partial\mathcal{L}}{\partial\phi_r} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)} \right\} + \int_\Omega d^4x\, \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\delta\phi_r \right) \\
&= \int_\Omega d^4x\, \delta\phi_r \left\{ \frac{\partial\mathcal{L}}{\partial\phi_r} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)} \right\} + \int_\Gamma d\Gamma_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\delta\phi_r \right) \\
&= \int_\Omega d^4x\, \delta\phi_r \left\{ \frac{\partial\mathcal{L}}{\partial\phi_r} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)} \right\}.
\end{aligned} \tag{3.33}$$

The last step is due to the condition that the variation $\delta\phi_r$ vanishes on the boundary Γ. Therefore, the Euler–Lagrange equations (or equations of motion) for the fields are:

$$\boxed{\frac{\partial\mathcal{L}}{\partial\phi_r} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)} = 0}. \tag{3.34}$$

Adding to the Lagrangian density a four-divergence ($\partial_\mu f^\mu$) causes no change to the equations of motions. If such a term is added, the action changes by a surface integral over the boundary Γ. This contribution either does not depend on the fields, or, if it does, gives no contribution to the variation of the action because $\delta\phi_r$ vanishes on the boundary Γ. Therefore, any two Lagrangians that differ by a four-divergence are equivalent.

In case we include also an explicit dependence of the Lagrangian on the coordinates, already in the second line we have to add a term $\partial_\mu(\mathcal{L}\delta x^\mu)$ (see App. 3.A). The variation of the action should be done so that the variation of the fields *and* the variation of the coordinates should vanish at the boundary

$$\delta S = \int_\Omega d^4x\, \delta\phi_r \left\{ \frac{\partial \mathcal{L}}{\partial \phi_r} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)} \right\} + \int_\Omega d^4x\, \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)} \delta\phi_r + \mathcal{L}\, \delta x^\mu \right) \tag{3.35}$$

Similarly to the case of discrete systems, also in field theories we can introduce a concept analogous to the generalized momentum conjugate to the generalized coordinates. In this case, we speak about a conjugate field defined as

$$\boxed{\pi_r(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_r(x)}}. \tag{3.36}$$

We can define a Hamiltonian density by

$$\boxed{\mathcal{H}(x) = \pi_r(x)\dot{\phi}_r(x) - \mathcal{L}}. \tag{3.37}$$

A summation over the repeated field indices is implied. The Hamiltonian is the integral over d^3x of the Hamiltonian density.

3.1.4 Symmetry transformations and Nöther's theorem.

If the Lagrangian density is invariant under (continuous) symmetry transformations, we can construct conserved quantities using Nöther's theorem.

Let us first focus on the so-called internal symmetries. In this case, we consider a transformation where the coordinates are kept unchanged, but the fields undergo a transformation of the type

$$\phi_r(x) \longrightarrow \phi'_r(x) = \phi_r(x) + \delta\phi_r(x). \tag{3.38}$$

For our continuous symmetries, the δ can be considered proportional to some infinitesimal parameter ϵ. The change in the Lagrangian density due to the above transformation can be written as

$$\begin{aligned} \delta\mathcal{L} &= \frac{\partial \mathcal{L}}{\partial \phi_r}\delta\phi_r + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)}\delta(\partial_\mu \phi_r) \\ &= \cancel{\left\{ \frac{\partial \mathcal{L}}{\partial \phi_r} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)} \right\}} \delta\phi_r + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)} \delta\phi_r \right) \\ &= \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)} \delta\phi_r \right). \end{aligned} \tag{3.39}$$

In the last step we used the equations of motion.[1] If the Lagrangian is invariant under this transformation, i.e., $\delta\mathcal{L} = 0$, then we can introduce a the four-current

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)} \frac{\delta \phi_r}{\epsilon} \tag{3.40}$$

which is "conserved," in the sense that its four-divergence is zero. The definition of the current contains a $1/\epsilon$, which denotes the parameter of the infinitesimal transformation, to give a definition of the current that does not depend on the "size" of the transformation (examples should clarify this point better). As reminded in the previous chapter, a vanishing four-divergence means that if we consider a three-dimensional volume V with surface S, the rate of change of the integral of j^0 in the volume is equal to the flux of $\vec{j}$ through the surface. Integrating j^0 over the full space, we obtain a conserved quantity. In other words, if we define

$$Q = \int d^3x j^0(x) \tag{3.41}$$

then $d_t Q = 0$.

We now consider also transformations that involve coordinate transformations, i.e.,

$$\begin{cases} x^\mu \to x'^\mu = x^\mu + \delta x^\mu, \\ \phi_r(x) \to \phi'_r(x') = \phi_r(x) + \Delta\phi_r(x). \end{cases} \tag{3.42}$$

We can express the total variation of the field as (omitting the field index)

$$\Delta\phi = \phi'(x') - \phi(x) = \big(\phi'(x') - \phi(x')\big) + \big(\phi(x') - \phi(x)\big) = \delta\phi + (\partial_\nu \phi)\, \delta x^\nu \tag{3.43}$$

The term $\delta\phi$ is usually referred to as "variation in form" and is not related to the change of coordinates.

The variation of the Lagrangian becomes now (see the steps in App. 3.A)

$$\begin{aligned} \delta\mathcal{L} &= \frac{\partial \mathcal{L}}{\partial \phi_r}\delta\phi_r + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)}\delta(\partial_\mu \phi_r) + \partial_\mu(\mathcal{L}\delta x^\mu) \\ &= \cancel{\left\{\frac{\partial \mathcal{L}}{\partial \phi_r} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)}\right\}}\delta\phi_r + \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)}\delta\phi_r\right) + \partial_\mu(\mathcal{L}\delta x^\mu) \\ &= \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)}\delta\phi_r + \mathcal{L}\delta x^\mu\right). \end{aligned} \tag{3.44}$$

The conclusion is that the conserved current takes the form

$$j^\mu = \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)}\delta\phi_r + \mathcal{L}\ \delta x^\mu\right)/\epsilon \tag{3.45}$$

[1] Note that the transformations we are considering now have nothing to do with the transformations considered before to obtain the equation of motions. The latter vanish on the boundary.

Inserting now the total variation of the fields, Eq. (3.43), we obtain

$$\begin{aligned} j^\mu &= \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)}\Big(\Delta\phi_r - (\partial_\nu \phi)\ \delta x^\nu\Big) + \mathcal{L}\ \delta x^\mu\right]/\epsilon \\ &= \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)}\Delta\phi_r - \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)}\partial_\nu \phi - \mathcal{L}\delta^\mu_\nu\right)\delta x^\nu\right]/\epsilon \\ &= \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)}\Delta\phi_r - \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)}\partial^\nu \phi - \mathcal{L}\, g^{\mu\nu}\right)\delta x_\nu\right]/\epsilon \end{aligned} \tag{3.46}$$

Let us define

$$\boxed{T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\partial^\nu \phi - g^{\mu\nu}\mathcal{L}.} \tag{3.47}$$

we can rewrite Nöther's current as

$$\boxed{j^\mu = \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_r)}\Delta\phi_r - T^{\mu\nu}\delta x_\nu\right)/\epsilon} \tag{3.48}$$

Let us now see two important examples of symmetries and their associated Nöther's current. We start with invariance under translations (in time and space). In this case, the transformations are

$$\begin{cases} x^\mu \to x'^\mu = x^\mu + \epsilon \hat{n}^\mu, \\ \phi_r(x) \to \phi'_r(x') = \phi_r(x) \Rightarrow \Delta\phi_r(x) = 0. \end{cases} \tag{3.49}$$

Using these results to build the conserved current we obtain

$$j^\mu = -T^{\mu\nu}\hat{n}_\nu. \tag{3.50}$$

Since the unit vector $\hat{n}$ is arbitrary, we see that the conservation of Nöther's current corresponds to

$$\boxed{\partial_\mu T^{\mu\nu} = 0.} \tag{3.51}$$

The tensor $T^{\mu\nu}$ is called the energy-momentum tensor. The continuity equations are four (one for each index ν) and correspond in fact to conservation of energy and conservation of momentum in each spatial direction. The conserved quantities are

$$P^\nu(t) = \int d^3x\ T^{0\nu}(x). \tag{3.52}$$

Let us check some components of the energy-momentum tensor (see Fig. 3.3)

$$T^{00} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}}\dot{\phi} - \mathcal{L} = \mathcal{H}. \tag{3.53}$$

This corresponds to the Hamiltonian density, i.e., the energy density. The associated current T^{i0} is the "energy current density" and generalizes the concept of the Poynting vector

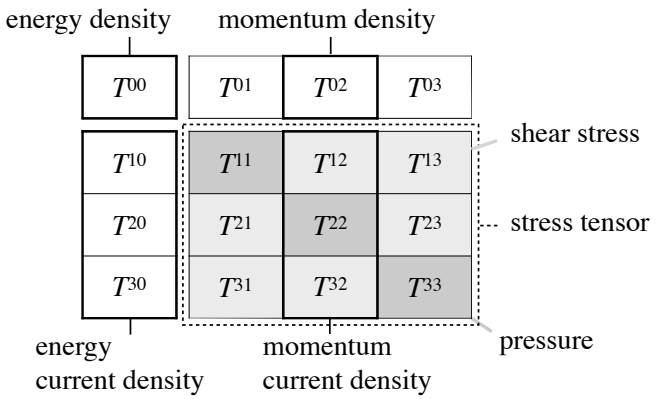


Figure 3.3: Components of the energy-momentum tensor. Note that in the literature often the tensor is defined with inverted μ and ν indices, therefore also the interpretation of the individual elements changes accordingly.

of the electromagnetic field. The continuity equation with $\nu = 0$ corresponds then to the conservation of energy $P^0 = H$. For $\nu = 1, 2, 3$ we obtain the conservation laws for the three components of the momentum. The components T^{0i} correspond to the "momentum density" and the components T^{ji} represent "momentum current density." More appropriately they form the so-called "stress tensor."

The energy-momentum tensor as defined above is usually called the "canonical" energy-momentum tensor and is in general not symmetric. However, we can always build a symmetrized version of the tensor, called the Belinfante energy-momentum tensor.[2] The point is that if we add a term $\partial_\lambda f^{\lambda\mu\nu}$ with $f^{\lambda\mu\nu} = -f^{\mu\lambda\nu}$, then this implies no change in the conservation laws, i.e., if we choose

$$T'^{\mu\nu} = T^{\mu\nu} + \partial_\lambda f^{\lambda\mu\nu} \tag{3.54}$$

it is still true that

$$\partial_\mu T'^{\mu\nu} = \partial_\mu T^{\mu\nu} = 0 \tag{3.55}$$

Invariance under rotations, instead, not surprisingly leads to conservation of angular momentum, although the definition of what we mean with angular momentum for a field is nontrivial. The transformation we are dealing with are Lorentz transformations. In this case, we have to distinguish the type of field we are considering. We define scalar fields those that do not change under Lorentz transformations. Fermion fields transform according to what we studied for the solutions of the Dirac equation. Vector fields transform as vectors.

For a scalar field, Lorentz transformations imply [see Eq. (2.164)]

$$\begin{cases} x^\mu \to x'^\mu = x^\mu + \epsilon_{\alpha\beta} M^{(\alpha\beta)\mu\nu} x_\nu, \\ \phi_r(x) \to \phi'_r(x') = \phi_r(x). \end{cases} \tag{3.56}$$

[2]This is useful, e.g., in the context of General Relativity, where Einstein's equations are written in terms of a symmetric energy-momentum tensor

Nöther's current, Eq. (3.48), takes the form ($\delta\phi = 0$)

$$j^{\mu(\alpha\beta)} \equiv \mathcal{M}^{\mu\alpha\beta} = -T^{\mu\nu} M^{(\alpha\beta)}_{\nu\sigma} x^{\sigma} = -T^{\mu\alpha} x^{\beta} + T^{\mu\beta} x^{\alpha}, \tag{3.57}$$

where we used the explicit form of the Lorentz generators in Eq. (2.169). The ensuing conserved quantities form the following tensor

$$J^{\alpha\beta} = \int d^3x \, \mathcal{M}^{0\alpha\beta}. \tag{3.58}$$

This tensor is antisymmetric, therefore we cannot have the same indices α and β. If we have two spatial indices (invariance under rotations), e.g., 1 and 2, the conserved quantity corresponds to the component 3 of the vector product of the position x and the momentum density T^{0i}, i.e., to angular momentum in the 3 direction. If we choose a time and spatial index (invariance under boosts), the conserved quantity turns out to be the position of the center-of-mass.

For a fermion field, Lorentz transformations induce a change also on the fields, according to Eq. (2.178)

$$\begin{cases} x^{\mu} \to x'^{\mu} = x^{\mu} + \epsilon_{\alpha\beta} M^{(\alpha\beta)\mu\nu} x_{\nu}, \\ \phi_r(x) \to \phi'_r(x') = \phi_r(x) - \frac{i}{2} \epsilon_{\alpha\beta} \sigma^{\alpha\beta} \phi_r(x). \end{cases} \tag{3.59}$$

The corresponding angular momentum tensor becomes

$$\mathcal{M}^{\mu\alpha\beta} \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \left(-\frac{i}{2} \sigma^{\alpha\beta} \right) \phi - T^{\mu\alpha} x^{\beta} + T^{\mu\beta} x^{\alpha}. \tag{3.60}$$

For spatial α and β, the total angular momentum is composed now by two parts: one related to the "internal" change of the fields upon rotations (absent in scalar fields) and one related to the rotation of coordinates in the physical space. This second term is identified with orbital angular momentum, the first is identified with spin. In fact, the operator σ^{ij} is related to the spin operator in the direction k, cf. Eqs. (2.189) and (2.191). Angular momentum and spin are not conserved separately: only total angular momentum is.

3.1.5 Global gauge invariance and charge conservation

Let us consider a particularly simple example of symmetry: suppose we are dealing with a complex field and the Lagrangian is invariant under a global phase transformation of the fields, i.e.,

$$\begin{cases} \phi(x) \to \phi'(x) = e^{i\alpha} \phi(x) \approx \phi + i\alpha\phi \\ \phi^*(x) \to \phi^{*\prime}(x) = e^{-i\alpha} \phi^*(x) \approx \phi^* - i\alpha\phi^* \end{cases} \tag{3.61}$$

It is clearly an internal symmetry, with no change of the coordinates. The Lagrangian can be anything built with combinations of $\phi^*\phi$. If the Lagrangian is chosen to be real, the invariance under the above symmetry is guaranteed.

Let us then check what we find for the conserved current. For an infinitesimal transformation ($\alpha \to 0$, where α represents then the infinitesimal parameter of our transformation) we have

$$\Delta\phi = \delta\phi \sim i\alpha\phi, \qquad \Delta\phi^* = \delta\phi^* \sim -i\alpha\phi^*. \tag{3.62}$$

The conserved current is obtained from Eq. (3.40). The field and its complex conjugate have to be treated as independent fields, i.e., if they were two fields with different index r

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\frac{\delta\phi}{\alpha} + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^*)}\frac{\delta\phi^*}{\alpha} = i\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\phi - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^*)}\phi^*\right). \tag{3.63}$$

We can identify a conserved quantity

$$Q = \int d^3x j^0 = i\int d^3x\left(\frac{\partial\mathcal{L}}{\partial(\partial_0\phi)}\phi - \frac{\partial\mathcal{L}}{\partial(\partial_0\phi^*)}\phi^*\right). \tag{3.64}$$

We note that for real fields, the conserved current would be identically zero. On the other hand, for complex fields (and whenever the Lagrangian is invariant under the above transformation) we can always define the above conserved quantity. We call it "charge," but is does not necessarily mean electric charge. It could be for instance hypercharge, baryonic number, leptonic number, flavor... Real fields, on the other hand, are not appropriate to describe fields with conservation of some global charge. For instance, the electromagnetic field is real and there is no global charge associated to it.

The phase transformation we have taken into consideration here is an example of Abelian transformation. Two phase transformations of this kind always commute (since in this case, the phase is just a number). There can be examples also of non-Abelian global gauge transformations. Consider for instance a field which is composed of two components $\phi = (\phi_u, \phi_d)$. We could also introduce the concept of a rotation in this two-dimensional space, which is formally similar to the rotations of Pauli spinors. Rotations of this kind are in fact usually called isospin rotations: the rotation would be described by a 2×2 unitary matrix V, with $\det(\mathrm{V}) = +1$ if we want them to describe continuous rotations connected to the identity. Therefore, an isospin rotation can be written as[3]

$$\phi' = e^{-i\,\theta\,\vec{\sigma}\cdot\vec{n}/2}\phi, \qquad \Rightarrow \delta\phi = -i\,\theta\,\frac{\vec{\sigma}\cdot\vec{n}}{2}\phi. \tag{3.65}$$

If our Lagrangian is invariant under this kind of transformation (for instance, if it is just the sum of two identical Lagrangians, one for ϕ_u and one for ϕ_d), the corresponding Nöther's current is the isospin current (there are three of them, since the axis of rotation is arbitrary)

$$j^{\mu i} = -i\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\frac{\sigma^i}{2}\phi. \tag{3.66}$$

[3]Often the Pauli matrices σ^i are denoted as τ^i when they are used to describe isospin rotations, but they are the same.

3.1.6 A classical example: the electromagnetic field

We will treat the quantized EM field later in the chapter, but let us consider for the moment this important example in the classical case. A good treatment of this topic can be found in the book by V. Barone [6].

We expect to recover Maxwell's equations. In the razionalized Gaussian units they can be written (in vacuum and with $c = 1$) as

$$\begin{aligned} \vec{\nabla} \cdot \vec{E} &= \rho, && \text{(Max I)} \\ \vec{\nabla} \cdot \vec{B} &= 0, && \text{(Max II)} \\ \vec{\nabla} \times \vec{E} &= -\partial_t \vec{B}, && \text{(Max III)} \\ \vec{\nabla} \times \vec{B} &= \vec{j} + \partial_t \vec{E}, && \text{(Max IV)} \end{aligned} \tag{3.67}$$

The first and last equations (Gauss's theorem and Ampére–Maxwell law) are usually referred to as inhomogeneous Maxwell's equations, since they contain terms related to the electric currents and charges. The second and third (Gauss's theorem for the magnetic field and Faraday–Neumann–Lenz's equation) are the homogeneous ones.

We have to first of all determine the Lagrangian density of the field. It turns out to be

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - j^\mu A_\mu \tag{3.68}$$

The last term is there only in the presence of a current. Remember that $j^\mu = (\rho,\ \vec{j})$. The EM tensor is defined as

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{3.69}$$

and is antisymmetric.[4] The A^μ are our fields: they are the four components of a vector field. Remember that $A^\mu = (\Phi,\ \vec{A})$.

We remember that the connection of the EM tensor and of the four-potential to the electric and magnetic fields is

$$E^i = F^{i0} = \partial^i A^0 - \partial^0 A^i = -\nabla^i A^0 - \partial_t A^i \tag{3.70}$$

$$\varepsilon^{ijr} B^r = F^{ji} = \partial^j A^i - \partial^i A^j = -\nabla^j A^i + \nabla^i A^j. \tag{3.71}$$

corresponding to

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}. \tag{3.72}$$

[4] Note that there are two alternative definitions of the EM tensor: Mandl–Shaw uses a different sign.

The validity of the homogeneous Maxwell's equations are already guaranteed by the above definitions. In fact, multiplying Eq. (3.71) by ε_{ijk} and using the property $\varepsilon_{ijk}\varepsilon^{ijr} = 2\delta^r_k$ we obtain

$$\varepsilon^{ijk}\varepsilon^{ijr}B^r = 2B^k = 2\varepsilon^{ijk}\partial^j A^i = 2\varepsilon^{ijk}\nabla^i A^j = 2\Big(\vec{\nabla}\times\vec{A}\Big)^k \tag{3.73}$$

In conclusion, Eqs. (3.70) and (3.71) become

$$\vec{E} = -\vec{\nabla}\Phi - \partial_t\vec{A}, \tag{3.74}$$

$$\vec{B} = \vec{\nabla}\times\vec{A}, \tag{3.75}$$

From the previous equations, we can check that

$$\vec{\nabla}\cdot\vec{B} = 0 \tag{Max II}$$

$$\vec{\nabla}\times\vec{E} = -\partial_t\vec{B} \tag{Max III}$$

which correspond to the homogeneous Maxwell's equations.

The shortest way to write these two equations is

$$\varepsilon_{\mu\nu\rho\sigma}\partial^\rho F^{\mu\nu} = 0. \tag{3.76}$$

They are automatically satisfied by the fact that F is antisymmetric. We can introduce the so-called dual field tensor ${}^*F^{\mu\nu} = -\varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}$ and further simplify the above equation

$$\partial_\mu\, {}^*F^{\mu\nu} = 0. \tag{3.77}$$

The first two of Maxwell's equations can be obtained as Euler-Lagrange equations derived from our Lagrangian. Expanding the Lagrangian we find

$$\begin{aligned}\mathcal{L} &= -\frac{1}{4}\Big(\partial_\mu A_\nu - \partial_\nu A_\mu\Big)\Big(\partial^\mu A^\nu - \partial^\nu A^\mu\Big) - j^\mu A_\mu \\ &= -\frac{1}{4}\Big((\partial_\mu A_\nu)\partial^\mu A^\nu - (\partial_\mu A_\nu)\partial^\nu A^\mu - (\partial_\nu A_\mu)\partial^\mu A^\nu + (\partial_\nu A_\mu)\partial^\nu A^\mu\Big) - j^\mu A_\mu \\ &= -\frac{1}{2}\Big((\partial_\nu A_\mu)\partial^\nu A^\mu - (\partial^\mu A^\nu)\partial_\nu A_\mu\Big) - j^\mu A_\mu\end{aligned} \tag{3.78}$$

To obtain the Euler–Lagrange equations we need

$$\frac{\partial\mathcal{L}}{\partial A_\mu} = -j^\mu, \tag{3.79}$$

$$\frac{\partial\mathcal{L}}{\partial(\partial_\nu A_\mu)} = -\Big(\partial^\nu A^\mu - \partial^\mu A^\nu\Big) = -F^{\nu\mu}. \tag{3.80}$$

Note that in the last equation superficially there seems to be a factor $1/2$ missing. This is not an error. Suppose you have an expression $a_\rho a^\rho$ and you take the derivative with respect to a_μ. We need to do the calculation explicitly component by component to convince

ourselves

$$\frac{\partial}{\partial a_0}(a_\rho a^\rho) = \frac{\partial}{\partial a^0}(a^0 a^0 - a^i a^i) = 2a^0, \tag{3.81}$$

$$\frac{\partial}{\partial a_i}(a_\rho a^\rho) = -\frac{\partial}{\partial a^i}(a^0 a^0 - a^i a^i) = 2a^i. \tag{3.82}$$

The four-dimensional way to write this is

$$\frac{\partial}{\partial a_\mu}(a_\rho a^\rho) = \frac{\partial}{\partial a_\mu}\left(g^{\rho\sigma} a_\rho a_\sigma\right) = g^{\rho\sigma}\left(\delta^\mu_\rho a_\sigma + a_\rho \delta^\mu_\sigma\right) = 2a^\mu \tag{3.83}$$

The same kind of procedure can be applied to the calculation of

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\mu)} &= -\frac{1}{2} g^{\alpha\rho} g^{\beta\sigma}\left[\delta^\nu_\alpha \delta^\mu_\beta\, \partial_\rho A_\sigma + (\partial_\alpha A_\beta)\delta^\nu_\rho \delta^\mu_\sigma - \delta^\nu_\rho \delta^\mu_\sigma\, \partial_\beta A_\alpha - (\partial_\rho A_\sigma)\delta^\nu_\beta \delta^\mu_\alpha\right] \\ &= -\left(\partial^\nu A^\mu - \partial^\mu A^\nu\right) = -F^{\nu\mu}.\end{aligned} \tag{3.84}$$

In conclusion, the Euler–Lagrange equations for the electromagnetic field (including an electromagnetic current j^μ) are

$$\boxed{\partial_\nu F^{\nu\mu} = j^\mu.} \tag{3.85}$$

This is the shortest possible way to write down Maxwell's equations (actually, only the first two). It is a collection of four different equations, one for each index μ.

To obtain the first two of Maxwell's equations, let us consider separately the cases $\mu = 0,\ i$ in Eq. (3.85)

$$j^0 = \partial_\nu F^{\nu 0} = -\partial^i F^{i0} = \nabla^i F^{i0} \tag{3.86}$$

$$j^i = \partial_\nu F^{\nu i} = \partial_t F^{0i} + \nabla^j F^{ji} \tag{3.87}$$

which lead to

$$\vec{\nabla}\cdot\vec{E} = \rho, \tag{Max I}$$

$$\vec{\nabla}\times\vec{B} = \vec{j} + \partial_t \vec{E}. \tag{Max IV}$$

We can also compute the Lagrangian and the Hamiltonian in terms of the E and B fields (let us consider $j^\mu = 0$)

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}F^{\mu\nu} = -\frac{1}{4}\Big(- F^{0i}F^{0i} - F^{0i}F^{0i} + F^{ij}F^{ij}\Big) = \frac{1}{2}\left(\vec{E}^2 - \vec{B}^2\right) \tag{3.88}$$

To compute the Hamiltonian, we can compute also the energy-momentum tensor. However, we need to use a trick, i.e., the fact that we can add a term of the type $\partial_\rho f^{\rho\mu\nu}$ with $f^{\rho\mu\nu} = -f^{\mu\rho\nu}$. In this case, we choose

$$\begin{aligned}T'^{\mu\nu} &= T^{\mu\nu} - \partial_\rho(F^{\rho\mu}A^\nu) \\ &= F^{\rho\mu}\partial^\nu A_\rho + \cancel{\partial_\rho(F^{\rho\mu})A^\nu} - F^{\rho\mu}\partial_\rho A^\nu - g^{\mu\nu}\mathcal{L} \\ &= -F^{\rho\mu}(\partial_\rho A^\nu - \partial^\nu A_\rho) - g^{\mu\nu}\mathcal{L} = F^{\rho\mu}F^\nu{}_\rho - g^{\mu\nu}\mathcal{L}\end{aligned} \tag{3.89}$$

(the added term is correctly antisymmetric in the $\rho\mu$ indices). Now we can compute the Hamiltonian density

$$\begin{aligned} T'^{00} &= F^{\rho 0}F^{0}{}_{\rho} - \mathcal{L} = -F^{i0}F^{0i} - \mathcal{L} = F^{0i}F^{0i} - \mathcal{L} \\ &= \vec{E}^2 - \frac{1}{2}(\vec{E}^2 - \vec{B}^2) = \frac{1}{2}(\vec{E}^2 + \vec{B}^2) \end{aligned} \tag{3.90}$$

This is the correct form of the energy density of the electromagnetic field (radiation field).

We can also compute the EM field momentum density

$$T'^{0i} = F^{\rho 0}F^{i}{}_{\rho} = -F^{j0}F^{ij} = E^j\epsilon^{ijr}B^r = (\vec{E}\times\vec{B})^i, \tag{3.91}$$

which is the expression of the Poynting vector.

In the above discussion we avoided talking about a crucial property of the electromagnetic field, i.e., gauge invariance. In covariant notation, gauge invariance means that we can transform our fields (i.e., the four-potential) in this way

$$A^\mu(x) \longrightarrow A^{\mu\prime}(x) = A^\mu(x) + \partial^\mu f(x) \tag{3.92}$$

without affecting the physics. The Lagrangian in the absence of currents is unaffected by this gauge transformation. When currents are present, the Lagrangian changes upon gauge transformations. However, if the current is conserved, $\partial_\mu j^\mu = 0$, the Lagrangian changes by a four-divergence, which insures that the equations of motions are not changed. We will discuss gauge invariance more extensively when we will quantize the electromagnetic field.

3.2 Field quantization

3.2.1 The harmonic oscillator

It is useful to quickly review some concepts related to the quantized harmonic oscillator, as it will provide the basic ideas for field quantization.

The Lagrangian is

$$L = \frac{m\dot{x}^2}{2} - \frac{m\omega^2}{2}x^2. \tag{3.93}$$

In this case, the generalized coordinate is the standard coordinate x. The conjugate momentum is

$$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x} \tag{3.94}$$

The Hamiltonian can be written as

$$H = p\dot{x} - L = \frac{p^2}{2m} + \frac{m\omega^2}{2}x^2. \tag{3.95}$$

Quantum mechanics can be obtained by promoting the generalized coordinates and momenta to operators and imposing the so-called "canonical commutation relations"

$$\big[x, p\big] = i\hbar. \tag{3.96}$$

To do a relativistic theory, it is more natural to use Heisenberg picture, where the operators depend also on time. It should be clear that in Heisenberg picture the commutation relation are *equal-time* commutation relations.

Generally speaking, the canonical quantization procedure for standard quantum mechanics can be described by these two steps: i) promote any function of the generalized coordinates to operators, ii) replace Poisson's brackets with commutation rules

$$\big\{f, g\big\}_{\mathrm{PB}} \longrightarrow -\frac{i}{\hbar}\big[f, g\big]. \tag{3.97}$$

Poisson's brackets are defined as

$$\big\{f, g\big\}_{\mathrm{PB}} = \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}\right) \tag{3.98}$$

where the sum runs over all generalized coordinates and momenta. Therefore, for a system with several generalized coordinate we obtain

$$\big[q_i, p_j\big] = i\hbar\delta_{ij}, \qquad \big[q_i, q_j\big] = 0, \qquad \big[p_i, p_j\big] = 0. \tag{3.99}$$

The analysis of the harmonic oscillator can be done also using ladder operators (creation and annihilation operators). They can be introduced as

$$a = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega x + ip), \qquad a^\dagger = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega x - ip). \tag{3.100}$$

We can check that

$$\big[a, a^\dagger\big] = -\frac{i}{\hbar}\big[x, p\big] = 1. \tag{3.101}$$

The Hamiltonian can be written as

$$H = \frac{\hbar\omega}{2}\big(a^\dagger a + aa^\dagger\big) = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) \tag{3.102}$$

where we also see the occurrence of the number operator

$$N = a^\dagger a \tag{3.103}$$

We can use as a Hilbert space for our harmonic oscillator the eigenvectors of the number operator (the so-called Fock space). We denote them by $\big|n\big\rangle$. The following properties

hold

$$N|n\rangle = n|n\rangle \qquad \text{with} \quad n = 0, 1, 2, \ldots, \tag{3.104}$$

$$Na^\dagger|n\rangle = (n+1)a^\dagger|n\rangle, \qquad Na|n\rangle = (n-1)a|n\rangle, \tag{3.105}$$

$$|n+1\rangle = \frac{1}{\sqrt{n+1}}a^\dagger|n\rangle, \qquad |n-1\rangle = \frac{1}{\sqrt{n}}a|n\rangle, \tag{3.106}$$

$$|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle. \tag{3.107}$$

The vacuum state is the one for which

$$a|0\rangle = 0, \qquad \langle 0|0\rangle = 1. \tag{3.108}$$

The eigenvectors of N are also eigenvectors of the Hamiltonian with energy

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right). \tag{3.109}$$

The ladder operators in Heisenberg picture follow the equation of motion

$$i\hbar\frac{da(t)}{dt} = \big[a(t), H\big]. \tag{3.110}$$

with the solution

$$a(t) = ae^{-i\omega t}. \tag{3.111}$$

If we have a superposition of independent oscillators, each at discrete positions denoted by the indices $i,\ j$, with different wavevectors $k,\ k'$, the commutation relations will be valid for each oscillator independently

$$\big[x_i, p_j\big] = i\hbar\delta_{ij}, \qquad \text{or} \quad \big[a_k, a_{k'}^\dagger\big] = \delta_{kk'}. \tag{3.112}$$

All other commutators vanish

$$\big[a_k, a_{k'}\big] = \big[a_k^\dagger, a_{k'}^\dagger\big] = 0. \tag{3.113}$$

There can be also extra indices to denote the oscillation modes (polarizations) $r,\ s$.

$$\big[x_{ri}, p_{sj}\big] = i\hbar\delta_{rs}\delta_{ij}. \qquad \text{or} \quad \big[a_{rk}, a_{sk'}^\dagger\big] = \delta_{rs}\delta_{kk'}. \tag{3.114}$$

The Hamiltonian in this case can be written as

$$H = \sum_r\sum_k \hbar\omega_k\left(a_{rk}^\dagger a_{rk} + \frac{1}{2}\right) \tag{3.115}$$

We can introduce the occupation number operators for each oscillation modes (defined by wavevector and polarization)

$$N_{rk} = a_{rk}^\dagger a_{rk} \tag{3.116}$$

we can define its eigenvectors

$$|n\rangle = \frac{1}{\sqrt{n_{rk}!}}\left(a_{rk}^{\dagger}\right)^{n_{rk}}|0\rangle. \tag{3.117}$$

We may find it uncomfortable to have a zero-point energy. In the case of a single oscillator it was $\hbar\omega/2$. For n discrete oscillators it becomes $n\hbar\omega/2$. In the continuum case it becomes infinite. It does not really matter, because what matters are energy differences between one state an another. In this case, we can simply shift, or define, as zero the energy of the vacuum, the $|0\rangle$ state. This corresponds to simply dropping the $+1/2$ in the Hamiltonian.

The way to formally handle this problem is to introduce the concept of *normal ordering*. Normal ordering consists in writing all the annihilation operators to the right. For instance

$$N[aa^{\dagger}] = a^{\dagger}a, \qquad\qquad N[a^{\dagger}a] = a^{\dagger}a. \tag{3.118}$$

For classical mechanics, the ordering of the operators is irrelevant. The Lagrangian we started from did not have any specific ordering rule. But since we are working with non-commuting operators, the order of the operators matter and we have to specify a prescription about that: we can decide to start from a *normal-ordered* Lagrangian.

Instead of defining normal ordering as "putting all annihilation operators to the right," we can use a more general definition for any operator $\hat{O}$

$$N[\hat{O}] = \hat{O} - \langle 0|\hat{O}|0\rangle, \tag{3.119}$$

i.e., the operator minus its vacuum expectation value. The latter is not an operator anymore, but a scalar function. We can include it into the definition of our Lagrangian without causing any difference in the physics. Note that

$$\langle 0|aa^{\dagger}|0\rangle = \langle 0|aa^{\dagger}|0\rangle - \langle 0|a^{\dagger}a|0\rangle = \langle 0|[a,a^{\dagger}]|0\rangle = [a,a^{\dagger}]\langle 0|0\rangle = [a,a^{\dagger}] \tag{3.120}$$

therefore

$$N[aa^{\dagger}] = aa^{\dagger} - \langle 0|aa^{\dagger}|0\rangle = aa^{\dagger} - [a,a^{\dagger}] = a^{\dagger}a, \tag{3.121}$$

which shows the consistency of the two definitions of normal ordering given above.

The wavefunction of the harmonic oscillator in the $n = 1$ state can be defined in this way

$$\Psi(x) = \langle x|1\rangle = \langle x|a^{\dagger}|0\rangle. \tag{3.122}$$

and the wave function of two oscillators in the $n = 1$ state is

$$\Psi(x_1, x_2) = \langle x_1, x_2|a_1^{\dagger}a_2^{\dagger}|0\rangle. \tag{3.123}$$

Since the a operators commute, the function is symmetric.

3.2.2 The fermionic oscillator

The above discussion is appropriate for oscillators with integer spin (*bosonic* oscillators). The occupation number can have any integer value, which means that there can be a superposition of any number of quanta in each oscillation mode. This cannot be the case for a semi-integer spin oscillator (*fermionic*), where the occupation number can only be 1 or 0.

It turns out, somewhat surprisingly, that the treatment of a fermionic oscillator can be carried out simply by changing the commutation rules into anticommutation rules, i.e.,

$$\{f, g\}_{\text{PB}} \longrightarrow -\frac{i}{\hbar}\{f, g\}. \tag{3.124}$$

At the level of creation and annihilation operators, this leads to

$$\left\{a_{rk}, a^{\dagger}_{sk'}\right\} = \delta_{rs}\delta_{kk'} \qquad \text{fermions.} \tag{3.125}$$

All other anticommutators vanish

$$\left\{a_{rk}, a_{sk'}\right\} = \left\{a^{\dagger}_{rk}, a^{\dagger}_{sk'}\right\} = 0. \tag{3.126}$$

Also in this case we can introduce a number operator in the same way as for the bosonic oscillator. The difference is that the eigenvalues of that operator can now be only 1 and 0. We have in fact

$$a^2 = \left(a^{\dagger}\right)^2 = 0 \tag{3.127}$$

$$N^2 = a^{\dagger}aa^{\dagger}a = a^{\dagger}\left(1 - a^{\dagger}a\right)a = N \Rightarrow N(N-1) = 0 \tag{3.128}$$

For each oscillation mode, we have only two possible states $\left|0\right\rangle$ and $\left|1_r\right\rangle = a^{\dagger}_r\left|0\right\rangle$.

Note that if we have a state formed by two oscillations

$$\left|1_{rk}1_{sk'}\right\rangle = a^{\dagger}_{rk}a^{\dagger}_{sk'}\left|0\right\rangle = -a^{\dagger}_{sk'}a^{\dagger}_{rk}\left|0\right\rangle = -\left|1_{sk'}1_{rk}\right\rangle. \tag{3.129}$$

That is, we recover the property that fermion states must be antisymmetric. If $r = s$ and $k = k'$ we see that the state must be zero, i.e., there cannot be two oscillations with the same momentum and polarization.

The wavefunction of the fermionic oscillator in the $n = 1$ state can be defined in this way

$$\Psi(x) = \left\langle x\middle|1\right\rangle = \left\langle x\middle|a^{\dagger}\right|0\rangle. \tag{3.130}$$

and the wave function of two oscillators in the $n = 1$ state is

$$\Psi(x_1, x_2) = \left\langle x_1, x_2\middle|a^{\dagger}_1 a^{\dagger}_2\right|0\rangle. \tag{3.131}$$

Since the a operators *anticommute*, the function is *antisymmetric*.

3.2.3 Procedure to quantize fields

To build a theory of quantized fields, we closely follow what was done in the discrete case. To go from discrete mechanics to field mechanics, we replace the generalized coordinates with time-and-space-dependent fields. Then

- we promote the fields to operators;
- for bosons, we replace Poisson brackets with commutators multiplied by $(-i/\hbar)$;
- for fermions, we replace Poisson brackets with anticommutators multiplied by $(-i/\hbar)$.

Note that for a Poisson bracket involving fields we typically need this kind of relation (see also Eq. (3.4))

$$\begin{aligned}\{\phi_r(x), \pi_s(x')\}_{\mathrm{PB}} &= \sum_i \int d^3z \left(\frac{\partial \phi_r(x)}{\partial \phi_i(z)} \frac{\partial \pi_s(x')}{\partial \pi_i(z)} - \frac{\partial \pi_s(x')}{\partial \phi_i(z)} \frac{\partial \phi_r(x)}{\partial \pi_i(z)} \right) \Big|_{x^0 = x'^0 = z^0} \\ &= \sum_i \int d^3z \, \delta_{ri} \delta_{si} \delta^3(\vec{x} - \vec{z}) \delta^3(\vec{x}' - \vec{z}) = \delta_{rs} \delta^3(\vec{x} - \vec{x}').\end{aligned} \tag{3.132}$$

In practice, apart from some exceptions, the above procedure can also be translated into the following

- for bosons, we assume the validity of the following *equal-time* commutation relations [5]

$$\big[\phi_r(x), \pi_s(x')\big] = i\hbar \, \delta_{rs} \delta(\vec{x} - \vec{x}') \qquad \big[\phi, \phi\big] = \big[\pi, \pi\big] = 0 \qquad \text{for } t = t', \tag{3.133}$$

- for fermions, we assume the validity of the following *equal-time* anticommutation relations

$$\big\{\phi_r(x), \pi_s(x')\big\} = i\hbar \, \delta_{rs} \delta(\vec{x} - \vec{x}') \qquad \big\{\phi, \phi\big\} = \big\{\pi, \pi\big\} = 0 \qquad \text{for } t = t'. \tag{3.134}$$

This procedure is called in different ways: "field quantization," "canonical field quantization," or "second quantization." This last name is a bit deceiving. There is only one quantization, but it is applied to field operators. The first steps into the ideas of field quantization (for a vibrating string) appeared in the so-called *Dreimännerarbeit* of Born, Heisenberg and Jordan in 1926 [11]. The formalism of field quantization for bosonic fields was devoloped by Dirac [15] and by Jordan and Klein in 1927 [22]. The first discussion of quantization of fermionic fields was done by Jordan and Wigner in 1928 [23]. Jordan

[5] At different times the commutation relations are different.

should be rated as one of the most important pioneer of Quantum Field Theory (in spite of his political views in favor of Nazism).

The necessity of using anticommutation or commutation rules is assumed at this point as a postulate. However, it turns out that it is dictated by the so-called spin-statistics theorem, first formulated by Pauli in 1940. In essence, the theorems says that if we want a theory that is Lorentz invariant, has positive energies and positive norms, and respects causality, we need to quantize bosons with commutation rules and fermions with anticommutation rules. We will see some aspects of this requirements in the following.

In terms of ladder operators, the quantization procedure consists in imposing the commutation or anticommutation relations

$$\big[a_r(\vec{k}), a_s^\dagger(\vec{k}')\big] = (2\pi)^3 2\omega_k \delta_{rs}\delta(\vec{k}-\vec{k}') \quad \big[a,a\big] = \big[a^\dagger, a^\dagger\big] = 0 \quad \text{for bosons,} \tag{3.135}$$

$$\big\{a_r(\vec{k}), a_s^\dagger(\vec{k}')\big\} = (2\pi)^3 2\omega_k \delta_{rs}\delta(\vec{k}-\vec{k}') \quad \big\{a,a\big\} = \big\{a^\dagger, a^\dagger\big\} = 0 \quad \text{for fermions.} \tag{3.136}$$

Pay attention because there are different possible conventions concerning the above commutation relations, corresponding to different definitions of the ladder operators. Mandl–Shaw and the notes of Miglietta are written with discrete values of k because they rely on the box quantization and take the continuum limit in the end. The commutation relations are written then without the prefactor $(2\pi)^3 2\omega_k$ and replacing the Dirac delta with a Krönecker delta $\delta_{kk'}$. Peskin–Schroeder use a slightly different normalization (the correspondence is $a(k)|_{\text{P-S}} = a/\sqrt{2\omega_k}$)

The reason to introduce the prefactor $(2\pi)^3 2\omega_k$ can be explained in this way. We can use an integration over d^3k and show that

$$\int d^3k\; \delta(\vec{k}-\vec{k}') = 1 \tag{3.137}$$

but the integration measure is not invariant. On the other hand, we can use an integration measure which is manifestly Lorentz invariant (at least for orthochronous transformations)[6]

$$\begin{aligned}
&\int \frac{d^4k}{(2\pi)^4} 2\pi\; \delta(k^2-m^2)\theta(k^0) \\
&\quad = \int \frac{d^4k}{(2\pi)^3}\delta(k^2-m^2)\theta(k^0) = \int \frac{d^4k}{(2\pi)^3}\delta\big[(k^0)^2-\omega_k^2\big]\theta(k^0) \\
&\quad = \int \frac{d^4k}{(2\pi)^3}\frac{1}{2k^0}\big[\delta(k^0+\omega_k)+\delta(k^0-\omega_k)\big]\theta(k^0) = \int \frac{d^3k}{(2\pi)^3}\frac{dk^0}{2k^0}\delta(k^0-\omega_k) \\
&\quad = \int \frac{d^3k}{(2\pi)^3 2\omega_k}.
\end{aligned} \tag{3.138}$$

It is the choice of this integration measure that makes one prefer to use the above definition of commutation rules.

[6] The following property of the Dirac δ is necessary: $\delta\big(f(x)\big) = \sum_{x_0} \delta(x-x_0)/|f'(x_0)|$, where x_0 are the points for which $f(x_0) = 0$.

We should also mention that it is possible to find different forms of quantization: the standard one, where the commutation relations are written at equal times, is called "instant-form" quantization, but it is possible also to use "light-front" and "point-form" quantization [18]. The difference is that the commutation relations are written for

$$x^0 = \text{constant} \qquad \text{instant form,} \tag{3.139}$$

$$x^0 + x^3 = \text{constant} \qquad \text{light-front form,} \tag{3.140}$$

$$x_\mu x^\mu = \text{constant} \qquad \text{point-form.} \tag{3.141}$$

The explicit form of the commutation relations is different in the three cases.

3.3 The real Klein–Gordon field

3.3.1 Lagrangian density and equations of motion

The Lagrangian density for this field is

$$\boxed{\mathcal{L}_{\text{KG}} = \frac{1}{2}\Big(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2\Big).} \tag{3.142}$$

Note that since the dimensions of the Lagrangian density in n.u. are $[M]^4$, the dimension of the field ϕ must be $[M]$.

To derive the equation of motion we need

$$\frac{\partial\mathcal{L}}{\partial\phi} = -m^2\phi, \qquad \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} = \partial^\mu\phi. \tag{3.143}$$

Just to avoid confusions, let's check the last result

$$\begin{aligned}\frac{1}{2}\frac{\partial}{\partial(\partial_\mu\phi)}\Big(\partial_\rho\phi\partial^\rho\phi\Big) &= \frac{1}{2}\frac{\partial}{\partial(\partial_\mu\phi)}\Big(\partial_\rho\phi\, g^{\rho\sigma}\partial_\sigma\phi\Big)\\ &= \frac{1}{2}\Big(\delta^\mu_\rho\, g^{\rho\sigma}\partial_\sigma\phi + \partial_\rho\phi\, g^{\rho\sigma}\,\delta^\mu_\sigma\Big) = \partial^\mu\phi\end{aligned} \tag{3.144}$$

The equation of motion turns out to be precisely the Klein–Gordon equation

$$\Big(\partial^\mu\partial_\mu + m^2\Big)\phi = 0. \tag{3.145}$$

Some remarks about the Lagrangian density [30]

1. The Lagrangian density does not depend explicitly on x, therefore it is invariant under translations and leads to conservation of energy and momentum;

2. The Lagrangian density is a Lorentz scalar, therefore it is invariant under Lorentz transformations and leads to conservation of angular momentum and center-of-mass position;

3. The Lagrangian density depends on ϕ and $\partial_\mu\phi$. No higher derivatives can be included if we want to end up with a second-order differential equation for the equation of motion.

4. The Lagrangian density is not uniquely determined, since we can always add a four-divergence $\partial_\nu F^\nu$ without changing the action and the equations of motion.

As we already remarked when studying the invariance of the Klein–Gordon equation, the fields do not have to change under Lorentz transformations. They are called "scalar" fields. They have no spin: there is no need to introduce spinors, there is no internal degree of freedom that has to undergo some change due to Lorentz transformations.

In the case of the real Klein–Gordon field, there is no conserved global charge. Therefore, the real Klein–Gordon field can be appropriate for the description of scalar (spin zero) chargeless fields. This is the case of, e.g., the Higgs boson.

The conjugate fields are

$$\pi(x) = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \dot{\phi}(x). \tag{3.146}$$

The energy-momentum tensor can be written as

$$T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - g^{\mu\nu}\mathcal{L}. \tag{3.147}$$

In this case, it is symmetric.

The Hamiltonian density can be written as

$$\mathcal{H} = T^{00} = \dot{\phi}^2 - \mathcal{L} = \frac{1}{2}\Big(\dot{\phi}^2 + (\vec{\nabla}\phi)^2 + m^2\phi^2\Big). \tag{3.148}$$

The momentum density can be written as

$$\mathcal{P}^i = T^{0i} = \dot{\phi}\partial^i\phi \quad \Rightarrow \vec{\mathcal{P}} = -\dot{\phi}\vec{\nabla}\phi. \tag{3.149}$$

3.3.2 Solutions for the Klein–Gordon field

The general solution will have a form like

$$\phi(x) \sim \phi(\vec{k})e^{-ikx} \tag{3.150}$$

with $k^0 = \omega_k = \sqrt{|\vec{k}|^2 + m^2}$. However, we would like now to study all possible k and we would like to work with reasonably normalized waves. A way to look for the right normalization is to work in a box with boundary conditions and then let the box dimensions go to infinity. We are not going to repeat the details of the derivation, but the final outcome is that the functions have to be normalized by a factor $1/\sqrt{2V\omega}$. The continuum limit is obtained at the end using the relation

$$\frac{1}{V}\sum_{\vec{k}=-\infty}^{\infty} \longrightarrow \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty} d^3k. \tag{3.151}$$

We will work here with continuous values of k from the start. An important mathematical relation to remember is

$$\frac{1}{(2\pi)^3}\int d^3y\, e^{i\vec{y}\cdot\vec{r}} = \delta(\vec{r}). \tag{3.152}$$

The solutions of the Klein–Gordon equation can be written in the form[7]

$$\boxed{\phi(x) = \phi^+(x) + \phi^-(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\Big(a(\vec{k})e^{-ikx} + a^\dagger(\vec{k})e^{ikx}\Big).} \tag{3.153}$$

with $k^0 = \omega_k = \sqrt{m^2 + \vec{k}^2}$.

We can easily check that the above function represents a general solution of the Klein–Gordon equation. Moreover, the presence of $a^\dagger$ ensures that the solution is Hermitian ($\phi = \phi^\dagger$). Note that we are talking about Hermitian instead of simply real and we used the † instead of the simple conjugation because we are now working with *field operators*.

We can find the expression for the conjugate fields

$$\pi(x) = \dot{\phi}(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k}(-i\omega_k)\,\Big(a(\vec{k})e^{-ikx} - a^\dagger(\vec{k})e^{ikx}\Big)\Big|_{k^0=\omega_k}, \tag{3.154}$$

We now perform our *field quantization* and impose the equal-time commutation relations of Eq. (3.133) (in the following, it is understood that $k^0 = \omega_k$ and $k^{0\prime} = \omega_{k'}$)

$$\begin{aligned}\big[\phi(x), \pi(x')\big]\Big|_{x^0=x^{0\prime}} = &\int \frac{d^3k}{(2\pi)^3 2\omega_k}\int \frac{d^3k'}{(2\pi)^3 2\omega_{k'}}(-i\omega_{k'}) \\ &\times \Big(\big[a(\vec{k}), a(\vec{k}')\big]\, e^{-ikx}e^{-ik'x'} - \big[a(\vec{k}), a^\dagger(\vec{k}')\big]\, e^{-ikx}e^{ik'x'} \\ &+ \big[a^\dagger(\vec{k}), a(\vec{k}')\big]\, e^{ikx}e^{-ik'x'} - \big[a^\dagger(\vec{k}), a^\dagger(\vec{k}')\big]\, e^{ikx}e^{ik'x'}\Big)\Big|_{x^0=x^{0\prime}}.\end{aligned} \tag{3.155}$$

Not surprisingly, it turns out that we have to assume the validity of the commutation rules for the creation and annihilation operators, Eq. (3.135). Then

$$\begin{aligned}\big[\phi(x), \pi(x')\big]\Big|_{x^0=x^{0\prime}} &= \int \frac{d^3k}{(2\pi)^3 2\omega_k}\int \frac{d^3k'}{(2\pi)^3 2\omega_{k'}}(-i\omega_{k'})(2\pi)^3 \\ &\quad\times\Big(-2\omega_k\delta(\vec{k}-\vec{k}')\, e^{-ikx}e^{ik'x'} - 2\omega_k\delta(\vec{k}-\vec{k}')\, e^{ikx}e^{-ik'x'}\Big)\Big|_{x^0=x^{0\prime}} \\ &= \int \frac{d^3k}{(2\pi)^3}\frac{i}{2}\Big(e^{-ik(x-x')} + e^{ik(x-x')}\Big)\Big|_{x^0=x^{0\prime}} \\ &= \int \frac{d^3k}{(2\pi)^3}\frac{i}{2}\Big(e^{i\vec{k}\cdot(\vec{x}-\vec{x}')} + e^{-i\vec{k}\cdot(\vec{x}-\vec{x}')}\Big) \\ &= i\int \frac{d^3k}{(2\pi)^3}\, e^{i\vec{k}\cdot(\vec{x}-\vec{x}')} = i\delta(\vec{x}-\vec{x}').\end{aligned} \tag{3.156}$$

[7] This is a general statement from the theory of Partial Differential Equations.

In the second-last step, we made use of the fact that we can change the integration variable of the second term from $\vec{k}$ to $-\vec{k}$ without any effect.

If neede, one can invert Eq. (3.153) to obtain the creation and annihilation operators

$$a(\vec{k}) = \int d^3x\, e^{ikx} i \overleftrightarrow{\partial}_t \phi(x), \tag{3.157}$$

$$a^\dagger(\vec{k}) = \int d^3x\, \phi(x) i \overleftrightarrow{\partial}_t e^{-ikx}, \tag{3.158}$$

where we used the notation

$$A(x) \overleftrightarrow{\partial} B(x) = A(x)\big(\partial B(x)\big) - \big(\partial A(x)\big)B(x). \tag{3.159}$$

3.3.3 Hamiltonian and momentum of the Klein–Gordon field

We can compute the Hamiltonian starting from the definition Hamiltonian density in Eq. (3.148)

$$H = \int d^3x\, \mathcal{H} = \int d^3x\, \frac{1}{2}\Big(\dot{\phi}^2 + (\vec{\nabla}\phi)^2 + m^2\phi^2\Big). \tag{3.160}$$

We have

$$\dot{\phi}(x) = \pi(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k}(-i\omega_k)\,\Big(a(\vec{k})e^{-ikx} - a^\dagger(\vec{k})e^{ikx}\Big)\Big|_{k^0=\omega_k}, \tag{3.161}$$

$$\vec{\nabla}\phi(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k}(i\vec{k})\,\Big(a(\vec{k})e^{-ikx} - a^\dagger(\vec{k})e^{ikx}\Big)\Big|_{k^0=\omega_k}. \tag{3.162}$$

At this point, we can rely on the fact that the Hamiltonian is a constant, therefore we can conveniently compute it for $t = 0$. We can check (see, e.g., Eq. 3.26 of Srednicki) that indeed the full calculation with $t \neq 0$ gives the same result.

$$\begin{aligned}
\big(\phi(x)\big)^2\big|_{t=0} &= \int \frac{d^3k\, d^3k'}{(2\pi)^6\, 2\omega_k\, 2\omega_{k'}}\Big(a(\vec{k})e^{i\vec{k}\cdot\vec{x}} + a^\dagger(\vec{k})e^{-i\vec{k}\cdot\vec{x}}\Big)\Big(a(\vec{k}')e^{i\vec{k}'\cdot\vec{x}} + a^\dagger(\vec{k}')e^{-i\vec{k}'\cdot\vec{x}}\Big) \\
&= \int \frac{d^3k\, d^3k'}{(2\pi)^6\, 2\omega_k\, 2\omega_{k'}}\Big(a(\vec{k})a(\vec{k}')e^{i(\vec{k}+\vec{k}')\cdot\vec{x}} + a^\dagger(\vec{k})a^\dagger(\vec{k}')e^{-i(\vec{k}+\vec{k}')\cdot\vec{x}} \\
&\quad + a(\vec{k})a^\dagger(\vec{k}')e^{i(\vec{k}-\vec{k}')\cdot\vec{x}} + a^\dagger(\vec{k})a(\vec{k}')e^{-i(\vec{k}-\vec{k}')\cdot\vec{x}}\Big).
\end{aligned} \tag{3.163}$$

We now perform the d^3x integration needed to obtain the Hamiltonian. We make use of

the usual property (3.152) to obtain $\delta(\vec{k} \pm \vec{k}')$

$$\begin{aligned}\int d^3x \big(m\phi(x)\big)^2\big|_{t=0} &= \int \frac{d^3k\, d^3k'}{(2\pi)^3\, 2\omega_k\, 2\omega_{k'}}\, m^2 \Big[\Big(a(\vec{k})a(\vec{k}') + a^\dagger(\vec{k})a^\dagger(\vec{k}')\Big)\delta(\vec{k}+\vec{k}') \\ &\quad + \Big(a(\vec{k})a^\dagger(\vec{k}') + a^\dagger(\vec{k})a(\vec{k}')\Big)\delta(\vec{k}-\vec{k}')\Big] \\ &= \int \frac{d^3k}{(2\pi)^3\, 4\omega_k^2}\, m^2 \Big(a(\vec{k})a(-\vec{k}) + a^\dagger(\vec{k})a^\dagger(-\vec{k}) \\ &\quad + a(\vec{k})a^\dagger(\vec{k}) + a^\dagger(\vec{k})a(\vec{k})\Big).\end{aligned} \tag{3.164}$$

The other two terms in the Hamiltonian can be calculated in a similar manner. The only difference is the change of sign of the mixed terms and the presence of the extra factors $-i\omega_k$ and $i\vec{k}$:

$$\begin{aligned}\int d^3x \big(\dot{\phi}(x)\big)^2\big|_{t=0} &= \int \frac{d^3k\, d^3k'}{(2\pi)^3\, 2\omega_k\, 2\omega_{k'}}\, (-\omega_k\omega_{k'}) \Big[\Big(a(\vec{k})a(\vec{k}') + a^\dagger(\vec{k})a^\dagger(\vec{k}')\Big)\delta(\vec{k}+\vec{k}') \\ &\quad - \Big(a(\vec{k})a^\dagger(\vec{k}') + a^\dagger(\vec{k})a(\vec{k}')\Big)\delta(\vec{k}-\vec{k}')\Big] \\ &= \int \frac{d^3k}{(2\pi)^3\, 4\omega_k^2}(-\omega_k^2) \Big(a(\vec{k})a(-\vec{k}) + a^\dagger(\vec{k})a^\dagger(-\vec{k}) \\ &\quad - a(\vec{k})a^\dagger(\vec{k}) - a^\dagger(\vec{k})a(\vec{k})\Big).\end{aligned} \tag{3.165}$$

$$\begin{aligned}\int d^3x \big(\vec{\nabla}\phi(x)\big)^2\big|_{t=0} &= \int \frac{d^3k\, d^3k'}{(2\pi)^3\, 2\omega_k\, 2\omega_{k'}}\, (-\vec{k}\cdot\vec{k}') \Big[\Big(a(\vec{k})a(\vec{k}') + a^\dagger(\vec{k})a^\dagger(\vec{k}')\Big)\delta(\vec{k}+\vec{k}') \\ &\quad - \Big(a(\vec{k})a^\dagger(\vec{k}') + a^\dagger(\vec{k})a(\vec{k}')\Big)\delta(\vec{k}-\vec{k}')\Big] \\ &= \int \frac{d^3k}{(2\pi)^3\, 4\omega_k^2}\, \vec{k}^2 \Big(a(\vec{k})a(-\vec{k}) + a^\dagger(\vec{k})a^\dagger(-\vec{k}) \\ &\quad + a(\vec{k})a^\dagger(\vec{k}) + a^\dagger(\vec{k})a(\vec{k})\Big).\end{aligned} \tag{3.166}$$

In the last step, note that $-\vec{k}\cdot\vec{k}' = \vec{k}^2$ if $\vec{k} = -\vec{k}'$, and $-\vec{k}\cdot\vec{k}' = -\vec{k}^2$ if $\vec{k} = \vec{k}'$.

Putting things together and remembering that $\omega_k^2 = \vec{k}^2 + m^2$ we obtain

$$H = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3 2\omega_k}\, \omega_k \Big(a(\vec{k})a^\dagger(\vec{k}) + a^\dagger(\vec{k})a(\vec{k})\Big). \tag{3.167}$$

Up to now, we did not make use of the quantization conditions. The above result for the Hamiltonian could have been derived also in the context of a classical field theory (apart from the fact that we treated a as operators and used the †). But now we make use of the commutation relations and replace

$$a(\vec{k})a^\dagger(\vec{k}) = a^\dagger(\vec{k})a(\vec{k}) + (2\pi)^3 2\omega_k \delta(0) \tag{3.168}$$

we then obtain

$$H = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, \omega_k \Big(a^\dagger(\vec{k}) a(\vec{k}) \Big) + \frac{1}{2} \int d^3k\, \omega_k \delta(0). \tag{3.169}$$

Note that the Hamiltonian (i.e., the energy of the field) is positive definite. Its eigenvalues must always be positive. Note also that if we would have adopted anticommutation rules for the a and $a^\dagger$ operators, we would have obtained only the last term. This is a part of the spin-statistics theorem, which tells you that in order to have a sensible result for the energy, for bosons we need to impose commutation relations on the operators. We shall see how the situation changes for fermions.

The last term of the Hamiltonian corresponds to the zero-point energy of a (infinite) system of harmonic oscillators and gives a divergent result, but at this point we can introduce the concept of normal ordering, as we did for the quantum harmonic oscillator.

Note that, in the case of normal ordering of field operators, we have to take into consideration that ϕ^+ components contain annihilation operators and have to go to the right. We then have (see Eq. 3.18 of Mandl–Shaw)

$$\begin{aligned} N\big[\phi(x)\phi(y)\big] &= N\big[\big(\phi^+(x) + \phi^-(x)\big)\big(\phi^+(y) + \phi^-(y)\big)\big] \\ &= N\big[\phi^+(x)\phi^+(y) + \phi^+(x)\phi^-(y) + \phi^-(x)\phi^+(y) + \phi^-(x)\phi^-(y)\big] \\ &= \phi^+(x)\phi^+(y) + \phi^-(y)\phi^+(x) + \phi^-(x)\phi^+(y) + \phi^-(x)\phi^-(y) \\ &= \phi(x)\phi(y) - \big[\phi^+(x), \phi^-(y)\big]. \end{aligned} \tag{3.170}$$

Alternatively, we can use the definition of normal ordering in Eq. (3.119) and check that

$$\begin{aligned} \big\langle 0\big|\phi(x)\phi(y)\big|0\big\rangle &= \big\langle 0\big|\phi^+(x)\phi^-(y)\big|0\big\rangle \\ &= \big\langle 0\big|\phi^+(x)\phi^-(y)\big|0\big\rangle - \big\langle 0\big|\phi^-(y)\phi^+(x)\big|0\big\rangle = \big[\phi^+(x), \phi^-(y)\big], \end{aligned} \tag{3.171}$$

leading to

$$N\big[\phi(x)\phi(y)\big] = \phi(x)\phi(y) - \big\langle 0\big|\phi(x)\phi(y)\big|0\big\rangle = \phi(x)\phi(y) - \big[\phi^+(x), \phi^-(y)\big], \tag{3.172}$$

which is consistent with our previous calculation.

The analysis with the normal-ordered Lagrangian would be equal to the one we have done, except that the Hamiltonian would be now

$$\boxed{H = \frac{1}{2} \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, \omega_k N\Big[a(\vec{k})a^\dagger(\vec{k}) + a^\dagger(\vec{k})a(\vec{k})\Big] = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, \omega_k \Big(a^\dagger(\vec{k})a(\vec{k})\Big).} \tag{3.173}$$

In other words, the effect of introducing normal ordering gives the same results as before, without any zero-point energy (the action of the Hamiltonian on the vacuum would give

0). In the following, we will always assume that we should be starting with normal-ordered Lagrangians.[8]

In similar ways we can also compute the momentum associated to the field, starting from Eq. 3.149. Steps analogous to the Hamiltonian lead to

$$\vec{P} = \int \frac{d^3k}{(2\pi)^3 2\omega_k} \vec{k}\Big(a^\dagger(\vec{k})a(\vec{k})\Big). \tag{3.174}$$

We can check in different ways the fact that the operators $a^\dagger$ and a correspond to creation and annihilation operators. There is no fundamental difference compared to the single-particle harmonic oscillator. We can identify a "oscillation number" density operator $N \propto a^\dagger(\vec{k})a(\vec{k})$. We can check that it commutes with the Hamiltonian. We can express the state of the system in terms of eigenstates of the number operator (Fock space). The vacuum state is the one for which

$$a(\vec{k})\big|0\big\rangle = 0, \qquad\qquad \big\langle 0\big|0\big\rangle = 1. \tag{3.175}$$

A state with a single oscillation/particle with momentum k is given by

$$\big|1_k\big\rangle = a^\dagger(\vec{k})\big|0\big\rangle. \tag{3.176}$$

The normalization of the one-particle state is

$$\langle 1_k|1_{k'}\rangle = \big\langle 0\big|a(\vec{k})a^\dagger(\vec{k}')\big|0\big\rangle = \cancel{\big\langle 0\big|a^\dagger(\vec{k})a(\vec{k}')\big|0\big\rangle} + \big\langle 0\big|[a(\vec{k}), a^\dagger(\vec{k}')]\big|0\big\rangle = (2\pi)^3 2\omega_k \delta(\vec{k}-\vec{k}') \tag{3.177}$$

Fock states can be built by acting with creation operators on the vacuum, while coordinate space states are built acting on the vacuum with the full field operator

$$\big|x\big\rangle = \phi(x)\big|0\big\rangle. \tag{3.178}$$

The wavefunction of a generic state is defined as

$$\Psi(x) = \big\langle x\big|\Psi\big\rangle. \tag{3.179}$$

For instance, the wavefunction corresponding to one oscillation of momentum p can be written as

$$\begin{aligned}\Psi(x) = \big\langle 0\big|\phi^\dagger(x)\big|1_p\big\rangle &= \int \frac{d^3k}{(2\pi)^3 2\omega_k}\Big(\big\langle 0\big|a(\vec{k})\big|1_p\big\rangle e^{-ikx} + \cancel{\big\langle 0\big|a^\dagger(\vec{k})\big|1_p\big\rangle e^{ikx}}\Big) \\ &= \int \frac{d^3k}{(2\pi)^3 2\omega_k}\big\langle 0\big|a(\vec{k})a^\dagger(\vec{p})\big|0\big\rangle e^{-ikx} \\ &= \int \frac{d^3k}{(2\pi)^3 2\omega_k}(2\pi)^3 2\omega_k \delta(\vec{k}-\vec{p})e^{-ikx} = e^{-ipx}\end{aligned} \tag{3.180}$$

[8]There may be peculiar situations where the zero-point energy plays a role: if the fields are constrained inside a finite region of space, e.g., a box, the zero-point energy inside the box is different from the zero-point energy outside the box. This is because the oscillations inside the box must have well-defined wavelengths, while outside the box all wavelengths are admissible. The difference in zero-point energies can cause a pressure on the box. This is the so-called Casimir effect.

We can prove that

$$\big[H, a^\dagger(\vec{k})\big] = \omega_k\, a^\dagger(\vec{k}), \qquad \big[H, a(\vec{k})\big] = -\omega_k\, a(\vec{k}), \tag{3.181}$$

$$H\, a^\dagger(\vec{k})\big|n_k\big\rangle = (n+1)\omega_k\, a^\dagger(\vec{k})\big|n_k\big\rangle, \qquad H\, a(\vec{k})\big|n_k\big\rangle = (n-1)\omega_k\, a(\vec{k})\big|n_k\big\rangle, \tag{3.182}$$

$$\vec{P}\, a^\dagger(\vec{k})\big|n_k\big\rangle = (n+1)\vec{k}\, a^\dagger(\vec{k})\big|n_k\big\rangle, \qquad \vec{P}\, a(\vec{k})\big|n_k\big\rangle = (n-1)\vec{k}\, a(\vec{k})\big|n_k\big\rangle. \tag{3.183}$$

3.3.4 Causality

It is instructive to explore also the expression of the commutator between two fields at different times (dropping the vanishing commutators from the beginning):

$$\begin{aligned}\big[\phi(x), \phi(x')\big] &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} \int \frac{d^3k'}{(2\pi)^3 2\omega_{k'}} \\ &\quad \times \Big(\big[a(\vec{k}), a^\dagger(\vec{k'})\big]\, e^{-ikx} e^{ik'x'} + \big[a^\dagger(\vec{k}), a(\vec{k'})\big]\, e^{ikx} e^{-ik'x'} \Big) \\ &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} \int \frac{d^3k'}{(2\pi)^3 2\omega_{k'}} \\ &\quad \times \Big((2\pi)^3 2\omega_k \delta(\vec{k} - \vec{k'})\, e^{-ikx} e^{ik'x'} - (2\pi)^3 2\omega_k \delta(\vec{k} - \vec{k'})\, e^{ikx} e^{-ik'x'} \Big) \\ &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} \Big(e^{-ik(x-x')} - e^{ik(x-x')} \Big)\end{aligned} \tag{3.184}$$

We can make some important observations concerning this expression. It depends only on the difference $x - x'$, which is consistent with translational invariance. It is Lorentz invariant because the integration measure is Lorentz invariant and the rest is a function of four-dimensional scalar products. If we consider $t = t'$, the commutator gives 0, as it should

$$\begin{aligned}\big[\phi(x), \phi(x')\big]\big|_{t=t'} &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} \Big(e^{-ik(x-x')} - e^{ik(x-x')} \Big)\Big|_{x^0 = x^{0\prime}} \\ &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} \Big(e^{i\vec{k}\cdot(\vec{x}-\vec{x}')} - e^{-i\vec{k}\cdot(\vec{x}-\vec{x}')} \Big) = 0\end{aligned} \tag{3.185}$$

where we used again the possibility of changing sign to $\vec{k}$ in the second term due to the fact that we are integrating over d^3k.

This means that the commutator vanishes for any space-like separation $(x - x')^2 < 0$, due to the fact that we can always perform a boost to a frame where $t - t' = 0$ and calculate the commutator there (obtainining 0). In other words, fields at two points with space-like separation commute. Only if the separation is time-like (i.e., inside the light cone), the fields do not commute.

We could not have obtained this result choosing anticommutation rules instead of commutation rules, because no negative sign would have appeared from the second term in Eq. (3.184)

This result is intimately connected with *causality*. If causality has to be preserved, a measurement at x should not influence a measurent at x' outside the light-cone, i.e., if $(x-x')^2 < 0$. This means that any two operators representing observable quantities must commute outside the light-cone.

Suppose we wanted to measure the energy density of the field at two different space-time points: the Hamiltonian operators at x and x' must commute. To keep the discussion simple, let us consider only the term $m^2\phi^2$ in the Hamiltonian. For $(x-x')^2 < 0$ we require

$$\begin{aligned}\Big[\phi(x)\phi(x),\phi(x')\phi(x')\Big] &= \phi(x)\phi(x)\phi(x')\phi(x') - \phi(x')\phi(x')\phi(x)\phi(x) \\ &= \phi(x)\phi(x)\phi(x')\phi(x') - \phi(x)\phi(x)\phi(x')\phi(x') = 0\end{aligned} \tag{3.186}$$

The condition is fulfilled if $\phi(x)$ and $\phi(x')$ commute, which is what we demonstrated above.

The full calculation of the commutator gives the so-called Pauli–Jordan causal function, that we quote without derivation

$$\big[\phi(x),\phi(x')\big] = iD(x-x') \tag{3.187}$$

where

$$\begin{aligned}D(x) &= -i\int \frac{d^3k}{(2\pi)^3 2\omega_k}\Big(e^{-ikx} - e^{ikx}\Big) \\ &= -i\int \frac{d^4k}{(2\pi)^3}\Big(\theta(x^0)-\theta(-x^0)\Big)e^{-ikx} \\ &= \frac{1}{2\pi}\Big(\theta(x^0)-\theta(-x^0)\Big)\Big[\delta(x^2) - \frac{m}{2\sqrt{x^2}}J_1\big(m\sqrt{x^2}\big)\Big],\end{aligned} \tag{3.188}$$

and where J_1 is the first-order Bessel function of the first kind. Note that for a massless particle the commutator is nonzero only on the light cone. For a massive particle it is nonzero also inside the light cone.

3.4 The complex Klein–Gordon field

This "type" of field is very similar to the real one. The only difference is that the field operators do not have to be Hermitian. We need to consider ϕ and $\phi^\dagger$ as two independent fields.

3.4.1 Lagrangian density and equations of motion

The Lagrangian density for this field is

$$\boxed{\mathcal{L}_{\mathrm{KG}*} = N\Big[\partial_\mu\phi^\dagger\partial^\mu\phi - m^2\phi^\dagger\phi\Big].} \tag{3.189}$$

For the equations of motions we need

$$\frac{\partial \mathcal{L}}{\partial \phi} = -m^2 \phi^\dagger, \qquad \frac{\partial \mathcal{L}}{\partial \phi^\dagger} = -m^2 \phi, \tag{3.190}$$

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = \partial^\mu \phi^\dagger, \qquad \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^\dagger)} = \partial^\mu \phi. \tag{3.191}$$

They lead to

$$\left(\partial^\mu \partial_\mu + m^2\right)\phi = 0, \qquad \left(\partial^\mu \partial_\mu + m^2\right)\phi^\dagger = 0. \tag{3.192}$$

The conjugate fields are

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \dot{\phi}^\dagger(x), \qquad \pi^\dagger(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}^\dagger} = \dot{\phi}(x). \tag{3.193}$$

The energy-momentum tensor can be written as (normal ordering is understood)

$$T^{\mu\nu} = \partial^\mu \phi^\dagger \, \partial^\nu \phi + \partial^\nu \phi^\dagger \, \partial^\mu \phi - g^{\mu\nu}\mathcal{L}. \tag{3.194}$$

In this case, it is symmetric.

The Hamiltonian density can be written as

$$\mathcal{H} = T^{00} = \dot{\phi}^\dagger \dot{\phi} + \dot{\phi}^\dagger \, \dot{\phi} - \mathcal{L} = \dot{\phi}^\dagger \, \dot{\phi} + \vec{\nabla}\phi^\dagger \cdot \vec{\nabla}\phi + m^2 \phi^\dagger \phi. \tag{3.195}$$

The momentum density can be written as

$$\mathcal{P}^i = T^{0i} = \dot{\phi}\partial^i \phi^\dagger + \dot{\phi}^\dagger \partial^i \phi \quad \Rightarrow \vec{\mathcal{P}} = -\dot{\phi}\vec{\nabla}\phi^\dagger - \dot{\phi}^\dagger \vec{\nabla}\phi. \tag{3.196}$$

3.4.2 Solutions for the complex Klein–Gordon field

The general solutions can be written as

$$\phi(x) = \phi^+(x) + \phi^-(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\left(a(\vec{k})e^{-ikx} + b^\dagger(\vec{k})e^{ikx}\right), \tag{3.197}$$

$$\phi^\dagger(x) = \phi^{\dagger+}(x) + \phi^{\dagger-}(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\left(b(\vec{k})e^{-ikx} + a^\dagger(\vec{k})e^{ikx}\right). \tag{3.198}$$

with $k^0 = \omega_k = \sqrt{m^2 + \vec{k}^2}$.

The expression of the conjugate fields

$$\pi(x) = \dot{\phi}^\dagger = \int \frac{d^3k}{(2\pi)^3 2\omega_k}(-i\omega_k)\left(b(\vec{k})e^{-ikx} - a^\dagger(\vec{k})e^{ikx}\right)\Big|_{k^0=\omega_k}, \tag{3.199}$$

$$\pi^\dagger(x) = \dot{\phi} = \int \frac{d^3k}{(2\pi)^3 2\omega_k}(-i\omega_k)\left(a(\vec{k})e^{-ikx} - b^\dagger(\vec{k})e^{ikx}\right)\Big|_{k^0=\omega_k}. \tag{3.200}$$

There are two canonical commutation relations (all others are zero)

$$\left[\phi(x), \pi(x')\right]\big|_{t=t'} = i\,\delta(\vec{x}-\vec{x}'), \qquad \left[\phi^\dagger(x), \pi^\dagger(x')\right]\big|_{t=t'} = i\,\delta(\vec{x}-\vec{x}'). \tag{3.201}$$

They correspond to two commutation relations (all others are zero)

$$\left[a(\vec{k}), a^\dagger(\vec{k}')\right] = (2\pi)^3 2\omega_k \delta(\vec{k}-\vec{k}'), \qquad \left[b(\vec{k}), b^\dagger(\vec{k}')\right] = (2\pi)^3 2\omega_k \delta(\vec{k}-\vec{k}'). \tag{3.202}$$

Therefore, there are two types of creation and annihilation operators, the a and b ones. There are two number operators and a generic state can have a certain number of excitations of type a and b, obtained from the vacuum state by the application of creation operators.

3.4.3 Hamiltonian, momentum, and charge of the complex Klein–Gordon field

The Hamiltonian can be obtained in a way similar to the real case. Let us look in detail at one of the terms

$$\begin{aligned}\phi^\dagger(x)\phi(x)\big|_{t=0} &= \int \frac{d^3k\, d^3k'}{(2\pi)^6\, 2\omega_k\, 2\omega_{k'}} \left(b(\vec{k}')e^{i\vec{k}'\cdot\vec{x}} + a^\dagger(\vec{k}')e^{-i\vec{k}'\cdot\vec{x}}\right)\left(a(\vec{k})e^{i\vec{k}\cdot\vec{x}} + b^\dagger(\vec{k})e^{-i\vec{k}\cdot\vec{x}}\right) \\ &= \int \frac{d^3k\, d^3k'}{(2\pi)^6\, 2\omega_k\, 2\omega_{k'}} \Big(b(\vec{k})a(\vec{k}')e^{i(\vec{k}+\vec{k}')\cdot\vec{x}} + a^\dagger(\vec{k})b^\dagger(\vec{k}')e^{-i(\vec{k}+\vec{k}')\cdot\vec{x}} \\ &\quad + a^\dagger(\vec{k})a(\vec{k}')e^{i(\vec{k}-\vec{k}')\cdot\vec{x}} + b(\vec{k})b^\dagger(\vec{k}')e^{-i(\vec{k}-\vec{k}')\cdot\vec{x}}\Big).\end{aligned} \tag{3.203}$$

We now perform the d^3x integration needed to obtain the Hamiltonian:

$$\begin{aligned}\int d^3x\, m^2\phi^\dagger(x)\phi(x)\big|_{t=0} &= \int \frac{d^3k\, d^3k'}{(2\pi)^3\, 2\omega_k\, 2\omega_{k'}}\, m^2\Big[\Big(b(\vec{k})a(\vec{k}') + a^\dagger(\vec{k})b^\dagger(\vec{k}')\Big)\delta(\vec{k}+\vec{k}') \\ &\quad + \Big(b(\vec{k})b^\dagger(\vec{k}') + a^\dagger(\vec{k})a(\vec{k}')\Big)\delta(\vec{k}-\vec{k}')\Big] \\ &= \int \frac{d^3k}{(2\pi)^3\, 4\omega_k^2}\, m^2\Big(b(\vec{k})a(-\vec{k}) + a^\dagger(\vec{k})b^\dagger(-\vec{k}) \\ &\quad + b(\vec{k})b^\dagger(\vec{k}) + a^\dagger(\vec{k})a(\vec{k})\Big).\end{aligned} \tag{3.204}$$

As in the real case, we can compute also the other terms of the Hamiltonian. Without repeating all the steps, the outcome of the calculation gives in the end, after imposing normal ordering

$$\boxed{H = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, \omega_k\Big(a^\dagger(\vec{k})a(\vec{k}) + b^\dagger(\vec{k})b(\vec{k})\Big).} \tag{3.205}$$

The Hamiltonian is positive-definite, but receives two different contributions. To have a better understanding of what these two oscillation modes represent, let's look at Nöther's current associated to the global gauge transformation discussed in Sec. 3.1.5. The Lagrangian density we started from is indeed invariant under this kind of transformation. Therefore, we can identify a conserved charge as defined in Eq. (3.64) (remember that there is an arbitrariness in the normalization):

$$Q = iq \int d^3x \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}^\dagger} \phi^\dagger - \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \phi \right) = iq \int d^3x \left(\pi^\dagger \phi^\dagger - \pi \phi \right). \tag{3.206}$$

As for the Hamiltonian, we can compute the "charge" at $t = 0$. Using the explicit decompositions of the fields and their conjugate, the calculation runs in a way similar to the Hamiltonian. We obtain

$$\boxed{Q = q \int \frac{d^3k}{(2\pi)^3 2\omega_k} \left(a^\dagger(\vec{k}) a(\vec{k}) - b^\dagger(\vec{k}) b(\vec{k}) \right).} \tag{3.207}$$

where we already applied normal ordering. We observe that the two types of oscillations give opposite contributions to the total charge. They can be interpreted, therefore, as particles (a operators) and antiparticles (b operators). The complex Klein–Gordon field can be used to describe fields/particles with spin 0 and with a charge of some type, for instance π^+ and π^-.

Note that the "probability density" we discussed in the context of the single-particle Klein–Gordon equation was exactly the same as the present charge density for the fields. It was impossible to interpret it as a probability density due to the negative contributions. In the context of field theories, it is straightforward to interpret it as a charge density, not as a probability density.

3.5 The Dirac field

This topic is treated in, e.g., Ch. 4 of Mandl–Shaw, Sec. 4.3 of Ryder, Sec. 7.2 of Aitchison–Hey, Sec. 3.5 of Peskin–Schroeder.

3.5.1 Lagrangian density and equations of motion

The starting point is the following Lagrangian density. As done before, we include normal ordering, although it becomes relevant only upon quantization.

$$\boxed{\mathcal{L}_{\text{Dir}} = N\left[\overline{\psi} (i \partial\!\!\!/ - m) \psi \right].} \tag{3.208}$$

Note that since the dimensions of the Lagrangian density in n.u. are $[M]^4$, the dimension of the field ψ must be $[M]^{3/2}$.

To obtain the equations of motion we need

$$\frac{\partial \mathcal{L}}{\partial \overline{\psi}} = -m\overline{\psi}, \qquad \frac{\partial \mathcal{L}}{\partial \dot{\overline{\psi}}} = (i\partial\!\!\!/ - m)\psi, \tag{3.209}$$

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} = \overline{\psi}\, i\,\gamma^\mu, \qquad \frac{\partial \mathcal{L}}{\partial(\partial_\mu \overline{\psi})} = 0. \tag{3.210}$$

They lead to the Dirac equations (remember that $\overline{\psi}\overleftarrow{\partial\!\!\!/} = \partial_\mu \overline{\psi}\gamma^\mu$)

$$\overline{\psi}(i\overleftarrow{\partial\!\!\!/} + m) = 0, \qquad (i\partial\!\!\!/ - m)\psi = 0. \tag{3.211}$$

The conjugate fields are

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\psi^\dagger(x), \qquad \overline{\pi}(x) = \frac{\partial \mathcal{L}}{\partial \dot{\overline{\psi}}} = 0. \tag{3.212}$$

It is interesting to note that the field conjugate to $\overline{\psi}$ vanishes. However, note also that the field conjugate to ψ is proportional to $\overline{\psi}$. This is an indication that ψ and $\overline{\psi}$ are not dynamically independent fields and we should worry only about one of them.

It is possible to do the analysis in a symmetric way, just starting from a different version of the Lagrangian density (normal-ordering is understood)

$$\begin{aligned}\mathcal{L} &= \overline{\psi}\left(\frac{i}{2}\overleftrightarrow{\partial\!\!\!/} - m\right)\psi \equiv \overline{\psi}\left(\frac{i}{2}\overrightarrow{\partial\!\!\!/} - \frac{i}{2}\overleftarrow{\partial\!\!\!/} - m\right)\psi = \overline{\psi}\Big(\frac{i}{2}\partial\!\!\!/ - m\Big)\psi - \overline{\psi}\Big(\frac{i}{2}\overleftarrow{\partial\!\!\!/}\Big)\psi \\ &= \overline{\psi}\Big(\frac{i}{2}\partial\!\!\!/ - m\Big)\psi - \frac{i}{2}\partial_\mu(\overline{\psi}\gamma^\mu\psi) + \overline{\psi}\Big(\frac{i}{2}\partial\!\!\!/\Big)\psi = \overline{\psi}(i\partial\!\!\!/ - m)\psi - \frac{i}{2}\partial_\mu(\overline{\psi}\gamma^\mu\psi).\end{aligned} \tag{3.213}$$

The two versions of the Lagrangian differ only by a four divergence, with no physical consequences. The different Lagrangians lead to different definitions of the conjugate fields

$$\pi(x) = \frac{i}{2}\psi^\dagger(x), \qquad \overline{\pi}(x) = -\frac{i}{2}\gamma^0\psi(x), \tag{3.214}$$

and different forms of the Poisson brackets

$$\{\psi(x), \pi(x')\}_{\mathrm{PB}} = 2\delta^3(\vec{x} - \vec{x}') \tag{3.215}$$

The factor 2 comes from the fact that there is a contribution to the Poisson bracket from the term $(\partial\psi/\partial\overline{\pi})(\partial\pi/\partial\overline{\psi})$, since the fields and their conjugates are not independent. The anticommutation relations applied to the creation and annihilation operators turn out to be unchanged. We shall not pursue this direction, we will use the first version of the Lagrangian and consider only the anticommutation relations between ψ and π.

The energy-momentum tensor can be written as (normal ordering is understood)

$$T^{\mu\nu} = \overline{\psi}\, i\,\gamma^\mu\partial^\nu\psi - g^{\mu\nu}\overline{\psi}\cancel{(i\partial\!\!\!/ - m)\psi} = \overline{\psi}\, i\,\gamma^\mu\partial^\nu\psi. \tag{3.216}$$

The last step is due to the validity of the equations of motion. Note that the tensor is not symmetric in this case.

The Hamiltonian density can be written as

$$\begin{aligned}\mathcal{H} &= T^{00} = \overline{\psi}\, i\gamma^0\partial^0\psi \\ &= \psi^\dagger\big(-i\gamma^0\vec{\gamma}\cdot\vec{\nabla} + m\gamma^0\big)\psi = \psi^\dagger H_D\psi\end{aligned} \tag{3.217}$$

where in the second line we used the Dirac equation and we used the Dirac Hamiltonian of Eq. (2.103).

The momentum density can be written as

$$\mathcal{P}^i = T^{0i} = \psi^\dagger\, i\,\partial^i\psi. \tag{3.218}$$

3.5.2 Solutions for the Dirac field

In the case of the Dirac field, the solutions can be written in general as (A is a Dirac index, s is a polarization index)

$$\psi_A(x) = \psi_A^+(x) + \psi_A^-(x) = \sum_{s=1,2}\int\frac{d^3k}{(2\pi)^3 2\omega_k}\Big(c_s(\vec{k})u_{sA}(\vec{k})e^{-ikx} + d_s^\dagger(\vec{k})v_{sA}(\vec{k})e^{ikx}\Big), \tag{3.219}$$

$$\overline{\psi}_A(x) = \overline{\psi}_A^+(x) + \overline{\psi}_A^-(x) = \sum_{s=1,2}\int\frac{d^3k}{(2\pi)^3 2\omega_k}\Big(d_s(\vec{k})\bar{v}_{sA}(\vec{k})e^{-ikx} + c_s^\dagger(\vec{k})\bar{u}_{sA}(\vec{k})e^{ikx}\Big). \tag{3.220}$$

with $k^0 = \omega_k = \sqrt{m^2+\vec{k}^2}$ and the spinors u and v are the objects we studied in the context of the single-particle Dirac equation. For sake of brevity, let us drop the arrow on the three-vectors in the arguments of the spinors and the creation and annihilation operators.

The conjugate field (as mentioned before, we consider only the field conjugate to ψ) is

$$\pi(x) = i\psi^\dagger(x) = i\sum_{s=1,2}\int\frac{d^3k}{(2\pi)^3 2\omega_k}\Big(d_s(k)v_{sA}^\dagger(k)e^{-ikx} + c_s^\dagger(k)u_{sA}^\dagger(k)e^{ikx}\Big). \tag{3.221}$$

Since they represent spin-half particles, i.e., fermions, we adopt the anticommutation

relations. Let us calculate (see Ryder, Eq. 4.50)[9]

$$
\begin{aligned}
\{\psi_A(x), \pi_B(x')\}\Big|_{t=t'} &= \{\psi_A(x), i\psi_B^\dagger(x')\}\Big|_{t=t'} \\
&= i\sum_{r,s}\int \frac{d^3k}{(2\pi)^3 2\omega_k}\int \frac{d^3k'}{(2\pi)^3 2\omega_{k'}} \\
&\quad \times \Big(u_{rA}(k)u_{sB}^\dagger(k')\,\{c_r(k), c_s^\dagger(k')\}\,e^{-ikx+ik'x'} \\
&\quad + v_{rA}(k)v_{sB}^\dagger(k')\,\{d_r^\dagger(k), d_s(k')\}\,e^{ikx-ik'x'}\Big)\Big|_{t=t'} \\
&= i\sum_{r,s}\int \frac{d^3k}{(2\pi)^3 2\omega_k}\int \frac{d^3k'}{(2\pi)^3 2\omega_{k'}}\,(2\pi)^3 2\omega_k\delta_{rs}\delta(\vec{k}-\vec{k}') \\
&\quad \times \Big(u_{rA}(k)\bar{u}_{sC}(k')\gamma^0_{CB}\,e^{-ikx+ik'x'} + v_{rA}(k)\bar{v}_{sC}(k')\gamma^0_{CB}\,e^{ikx-ik'x'}\Big)\Big|_{t=t'} \\
&= i\int \frac{d^3k}{(2\pi)^3 2\omega_k}\Big((\not{k}+m)_{AC}\gamma^0_{CB}\,e^{i\vec{k}\cdot(\vec{x}-\vec{x}')} + (\not{k}-m)_{AC}\gamma^0_{CB}\,e^{-i\vec{k}\cdot(\vec{x}-\vec{x}')}\Big) \\
&= i\int \frac{d^3k}{(2\pi)^3 2\omega_k}\Big((k^0\gamma^0 - k^i\gamma^i + m)_{AC} + (k^0\gamma^0 + k^i\gamma^i - m)_{AC}\Big)\gamma^0_{CB}\,e^{i\vec{k}\cdot(\vec{x}-\vec{x}')} \\
&= i\int \frac{d^3k}{(2\pi)^3}\mathbb{1}_{AB}\,e^{i\vec{k}\cdot(\vec{x}-\vec{x}')} = i\,\mathbb{1}_{AB}\,\delta(\vec{x}-\vec{x}').
\end{aligned}
\tag{3.222}
$$

In the third-to-last step, we changed the integration variable of the second term, from $\vec{k}$ to $-\vec{k}$. This is the reason of the change of sign of the $k^i\gamma^i$ term (as usual, it is understood that $k^0 = \omega_k$ for the free field).

Note that up to this point there is no compelling reason to choose anticommutators instead of commutators. We could have obtained similar results even with the choice of commutation rules. However, it would have been necessary to switch d with $d^\dagger$ in the expansion of Eq. (3.219).

3.5.3 Hamiltonian, momentum, charge, and spin of the Dirac field

For the computation of the Hamiltonian we need

$$
\dot{\psi}_A(x) = \sum_{s=1,2}\int \frac{d^3k}{(2\pi)^3 2\omega_k}\,(-i\omega_k)\Big(c_s(\vec{k})u_{sA}(\vec{k})e^{-ikx} - d_s^\dagger(\vec{k})v_{sA}(\vec{k})e^{ikx}\Big). \tag{3.223}
$$

We can compute the Hamiltonian and momentum of the Dirac field (we do it at $t = 0$,

[9] In the first step, we avoid writing out the anticommutators that are already known to vanish.

we drop Dirac indices)

$$
\begin{aligned}
H &= \int d^3x\, i\psi^\dagger(x)\dot{\psi}(x) \\
&= \sum_{s,s'} \int d^3x \int \frac{d^3k\, d^3k'}{(2\pi)^6\, 2\omega_k\, 2\omega_{k'}}\, \omega_k \Big(d_s(\vec{k}) v_s^\dagger(\vec{k}) e^{-ikx} + c_s^\dagger(\vec{k}) u_s^\dagger(\vec{k}) e^{ikx} \Big) \\
&\quad \times \Big(c_{s'}(\vec{k}') u_{s'}(\vec{k}') e^{-ik'x} - d_{s'}^\dagger(\vec{k}') v_{s'}(\vec{k}') e^{ik'x} \Big) \\
&= \sum_{s,s'} \int \frac{d^3k\, d^3k'}{(2\pi)^3\, 2\omega_k\, 2\omega_{k'}}\, \omega_k \Big\{ \Big(d_s(\vec{k}) c_{s'}(\vec{k}') v_s^\dagger(\vec{k}) u_{s'}(\vec{k}') - c_s^\dagger(\vec{k}) d_{s'}^\dagger(\vec{k}') u_s^\dagger(\vec{k}) v_{s'}(\vec{k}') \Big) \delta(\vec{k}+\vec{k}') \\
&\quad + \Big(c_s^\dagger(\vec{k}) c_{s'}(\vec{k}') u_s^\dagger(\vec{k}) u_{s'}(\vec{k}') - d_s(\vec{k}) d_{s'}^\dagger(\vec{k}') v_s^\dagger(\vec{k}) v_{s'}(\vec{k}') \Big) \delta(\vec{k}-\vec{k}') \Big\} \\
&= \sum_{s,s'} \int \frac{d^3k}{(2\pi)^3\, 2\omega_k}\, \frac{1}{2} \Big(d_s(\vec{k}) c_{s'}(-\vec{k}) \cancel{v_s^\dagger(\vec{k}) u_{s'}(-\vec{k})} - c_s^\dagger(\vec{k}) d_{s'}^\dagger(-\vec{k}) \cancel{u_s^\dagger(\vec{k}) v_{s'}(-\vec{k})} \\
&\quad + c_s^\dagger(\vec{k}) c_{s'}(\vec{k}) u_s^\dagger(\vec{k}) u_{s'}(\vec{k}) - d_s(\vec{k}) d_{s'}^\dagger(\vec{k}) v_s^\dagger(\vec{k}) v_{s'}(\vec{k}) \Big) \\
&= \sum_{s,s'} \int \frac{d^3k}{(2\pi)^3\, 2\omega_k}\, \frac{1}{2} \Big(c_s^\dagger(\vec{k}) c_{s'}(\vec{k}) 2\omega_k \delta_{ss'} - d_s(\vec{k}) d_{s'}^\dagger(\vec{k}) 2\omega_k \delta_{ss'} \Big) \\
&= \sum_{s} \int \frac{d^3k}{(2\pi)^3\, 2\omega_k}\, \omega_k \Big(c_s^\dagger(\vec{k}) c_s(\vec{k}) - d_s(\vec{k}) d_s^\dagger(\vec{k}) \Big).
\end{aligned}
\tag{3.224}
$$

The result for the Dirac field can be compared with that of the complex Klein–Gordon field, where we had the combination $a^\dagger(\vec{k})a(\vec{k}) + b(\vec{k})b^\dagger(\vec{k})$. At this point, we make use of anticommutation rules on the ladder operators

$$
d_s(\vec{k}) d_s^\dagger(\vec{k}) = -d_s^\dagger(\vec{k}) d_s(\vec{k}) + (2\pi)^3 2\omega_k \delta(0). \tag{3.225}
$$

As before, we can either replace the operators as above and subtract an (infinite) constant energy, or we can specify from the beginning that we are considering operators in normal ordering. For fermionic fields, however, normal ordering is defined with an extra minus sign due to the anticommutation rules, i.e.,

$$
N[dd^\dagger] = -d^\dagger d \tag{3.226}
$$

This is consistent with the definition of normal ordering given in Eq. (3.119) since

$$
\langle 0|dd^\dagger|0\rangle = \langle 0|dd^\dagger|0\rangle + \langle 0|d^\dagger d|0\rangle = \langle 0|\{d, d^\dagger\}|0\rangle = \{d, d^\dagger\}\langle 0|0\rangle = \{d, d^\dagger\} \tag{3.227}
$$

therefore

$$
N[dd^\dagger] = dd^\dagger - \langle 0|dd^\dagger|0\rangle = dd^\dagger - \{d, d^\dagger\} = -d^\dagger d. \tag{3.228}
$$

The difference compared to the bosonic field is essential. In fact, the Hamiltonian with normal ordering turns out to be

$$
\boxed{H = \int d^3x\, N[i\psi^\dagger(x)\dot{\psi}(x)] = \sum_s \int \frac{d^3k}{(2\pi)^3\, 2\omega_k}\, \omega_k \Big(c_s^\dagger(\vec{k}) c_s(\vec{k}) + d_s^\dagger(\vec{k}) d_s(\vec{k}) \Big).} \tag{3.229}
$$

We can appreciate once again that the Hamiltonian is positive-definite. Note that if we decided to use commutation rules instead of anticommutation rules, this would have led us to a Hamiltonian not bound from below. This is another example of the validity of the spin-statistics theorem.

As for the bosonic fields, we can check what is the effect of normal-ordering on two fermionic fields (writing the Dirac indices explicitly)

$$\begin{aligned} N\big[\psi_A(x)\psi_B(y)\big] &= N\big[\psi_A^+(x)\psi_B^+(y) + \psi_A^+(x)\psi_B^-(y) + \psi_A^-(x)\psi_B^+(y) + \psi_A^-(x)\psi_B^-(y)\big] \\ &= \psi_A^+(x)\psi_B^+(y) - \psi_B^-(y)\psi_A^+(x) + \psi_A^-(x)\psi_B^+(y) + \psi_A^-(x)\psi_B^-(y) \\ &= \psi_A(x)\psi_B(y) - \big\{\psi_A^+(x), \psi_B^-(y)\big\}. \end{aligned} \tag{3.230}$$

This is again consistent (minus sign included) with the definition in Eq. (3.119).

As before, we can study the behavior of the Lagrangian density under a global gauge transformation. It is easy to see that it is indeed invariant under this kind of transformation. The calculation of the charge goes in much the same way as for the complex Klein–Gordon field. The Nöther current is

$$j^\mu = -iq\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\psi - \cancel{\frac{\partial\mathcal{L}}{\partial(\partial_\mu\overline{\psi})}}\overline{\psi}\right) = q\overline{\psi}\gamma^\mu\psi \tag{3.231}$$

which is the well-known current density we introduced while studying the Dirac equation.

The resulting conserved charge is

$$Q = q\sum_s \int \frac{d^3k}{(2\pi)^3 2\omega_k}\left(c_s^\dagger(\vec{k})c_s(\vec{k}) - d_s^\dagger(\vec{k})d_s(\vec{k})\right). \tag{3.232}$$

As before, we see that there are two contributions, from particles and antiparticles (each one has two contributions from the different spin states). Note that, at variance with the result obtained while studying the Dirac equation, the conserved charge is not positive definite. This is due to the fact that we have now introduced anticommutation rules.

Finally, let us consider the projection of the spin operator along the direction of the momentum, what we have called the helicity operator in Sec. 2.2.13. For the Dirac field we have, using Eqs. (2.260) and (2.261),

$$h = \int d^3x\psi^\dagger h(\vec{k})\psi = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\left(c_1^\dagger(\vec{k})c_1(\vec{k}) - c_2^\dagger(\vec{k})c_2(\vec{k}) + d_1^\dagger(\vec{k})d_1(\vec{k}) - d_2^\dagger(\vec{k})d_2(\vec{k})\right). \tag{3.233}$$

from which we see that the particle and antiparticle components with spin label 1 give a positive contribution to helicity, while the components with spin label 2 give a negative contribution. This justifies the asymmetry between the treatment of u and v spinors in Secs. 2.2.12 and 2.2.13.

3.5.4 Causality

An interesting discussion of this topic is done in Peskin–Schroeder, Sec. 3.5, where they start consider the possibility of postulating commutation rules instead of anticommutation rules for fermionic fields. With commutation rules, the Hamiltonian would receive negative contributions from the $d^\dagger d$ operators. To avoid the problem, we could argue that these negative energy contributions are for some reason forbidden, inaccessible. In this case, however, we would run into problems with causality, because negative energy states play an essential role in restoring causality.

Here, we limit ourselves to check if causality is preserved when anticommutation rules are imposed. As we observed for the Klein–Gordon field, a measurement at x should not influence a measurent at x' outside the light-cone, i.e., if $(x-x')^2 < 0$. This still means that any two operators representing observable quantities must commute outside the light-cone. For simplicity, let us consider the measurement of the energy density and let us consider the term $m\overline{\psi}\psi$ in the Hamiltonian. For $(x-x')^2 < 0$ we require

$$\begin{aligned}&\Big[\overline{\psi}_A(x)\psi_A(x), \overline{\psi}_B(x')\psi_B(x')\Big] \\ &\qquad = \overline{\psi}_A(x)\psi_A(x)\overline{\psi}_B(x')\psi_B(x') - \overline{\psi}_B(x')\psi_B(x')\overline{\psi}_A(x)\psi_A(x) = 0\end{aligned} \tag{3.234}$$

Interestingly, this condition is fulfilled if all fields operators either commute or anticommute (because the operator is composed of two such fields, so we have to switch the operators twice). Obviously, in the fermionic case we are interested in the anticommutation rules. Through steps similar to Eq. (3.222), we can check that

$$\begin{aligned}\big\{\psi_A(x), \overline{\psi}_B(x')\big\} &= \sum_{r,s}\int \frac{d^3k}{(2\pi)^3 2\omega_k}\int \frac{d^3k'}{(2\pi)^3 2\omega_{k'}} \\ &\quad \times \Big(u_{rA}(k)\bar{u}_{sB}(k')\,\{c_r(k), c_s^\dagger(k')\}\,e^{-ikx+ik'x'} \\ &\quad + v_{rA}(k)\bar{v}_{sB}(k')\,\{d_r^\dagger(k), d_s(k')\}\,e^{ikx-ik'x'}\Big) \\ &= \int \frac{d^3k}{(2\pi)^3 2\omega_k}\Big((\not{k}+m)_{AB}\,e^{-ik(x-x')} - (-\not{k}+m)_{AB}\,e^{ik(x-x')}\Big) \\ &= \big(i\not{\partial}_x + m\big)_{AB}\int \frac{d^3k}{(2\pi)^3 2\omega_k}\Big(e^{-ik(x-x')} - e^{ik(x-x')}\Big).\end{aligned} \tag{3.235}$$

Apart from the Dirac structure with the derivative at the beginning, the second part corresponds to the commutator between two scalar fields analyzed in Eq. (3.184): the same considerations applying there hold here and allow us to say that the anticommutator vanishes outside the light-cone, thus preserving causality also in the fermion-field case.

3.6 The electromagnetic field

Historically, the quantization of the electromagnetic field and its effect on black-body radiation was the phenomenon that marked the beginning of Quantum Mechanics. The formalization of field quantization for the electromagnetic field came only in 1929, with works by Heisenberg and Pauli. It was the first field to be quantized, even before the quantization of the Dirac or Klein–Gordon fields. However, due to the complications added by gauge invariance, we treat this case as the last one.

3.6.1 Lagrangian density and equations of motion

The electromagnetic field is massless and has spin-1. This kind of field has only two independent components, but is described by a four-vector A^μ. There are two alternatives, depending on how we choose to fix the gauge conditions: we can eliminate two of the components from the beginning, but we loose manifest covariance; or we can keep four components, of which two should however have no effect on the physics.

We have already seen how it is possible to define the Lagrangian density that leads to Maxwell's equations, which was

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - j^\mu A_\mu \tag{3.236}$$

The fields A^μ must have dimension $[M]$, in order for the Lagrangian density to have dimensions $[M]^4$..

The theory must be invariant under gauge transformations

$$A^\mu(x) \longrightarrow A^{\mu\prime}(x) = A^\mu(x) + \partial^\mu f(x). \tag{3.237}$$

The Lagrangian density is not invariant under this transformation, but the difference is just a four-divergence

$$\mathcal{L} \longrightarrow \mathcal{L}' = \mathcal{L} - \partial_\mu\big(j^\mu f(x)\big) \tag{3.238}$$

(if we assume that for an electric current the continuity equation holds, $\partial_\mu j^\mu = 0$).

Note that a Lagrangian with a mass term of the type $m^2 A^\mu A_\mu$ would not be invariant under gauge transformations: gauge symmetry requires our field to be massless.

Note that gauge invariance does not lead to any new conservation law, because it is not an additional symmetry that the Lagrangian fulfills, but rather it is a constraint that we have to impose on the Lagrangian from the beginning because we know that physical results should not depend on the gauge. This extra constraint originates from the fact that we prefer to work with the four-potential A^μ, but in reality we are introducing extra, nonphysical degrees of freedom (often called auxiliary fields).

Our starting Lagrangian, however, is not ideal for field quantization. This is due to the fact that the field conjugate to A^0 vanishes

$$\pi^0 = \frac{\partial\mathcal{L}}{\partial(\partial^0 A^0)} = -F^{00} = 0. \tag{3.239}$$

At this point, we make use of the freedom to choose the gauge. There are several possibilities, but the most important ones are

$$\vec{\nabla} \cdot \vec{A} = 0 \qquad \text{Coulomb gauge,} \tag{3.240}$$

$$\partial^\mu A_\mu = 0 \qquad \text{Lorenz gauge,} \tag{3.241}$$

$$n^\mu A_\mu = 0 \qquad \text{temporal, light-cone, or axial gauge.} \tag{3.242}$$

Note that the Lorenz condition above does not fix the gauge completely. Performing a gauge transformation with a function $f(x)$ for which

$$\partial_\mu \partial^\mu f(x) = 0, \tag{3.243}$$

we still obtain a field that satisfies the Lorenz gauge condition. The Lorenz gauge condition identifies a class of gauges. Effectively, the Lorenz gauge reduces the number of independent components of A^μ from four to three. But we can use the extra gauge freedom to further reduce the number of independent components to two. Consider the example of a free field with no charges and currents. If we further choose a function $f(x)$ for which

$$\partial^0 f(x) = -\Phi \tag{3.244}$$

the resulting A^0 component of the field vanishes, while it must be that $\vec{\nabla} \cdot \vec{A} = 0$. Therefore, we fall in the case of the Coulomb gauge.[10]

The Coulomb condition reduces the independent components of the fields to two (which is the case of the physical electromagnetic field). The condition is only one, so it would appear that there were still three independent components of the field. However it turns out that the equation of motions (Maxwell's equations) allow to further fix the scalar potential $A^0 = \Phi$. In case of the absence of charges, Φ can be fixed to 0. In case of the presence of charges, Φ is anyway fixed by the distribution of charges.

The advantage of the Coulomb gauge is that the A^0 field does not appear as a degree of freedom, therefore we do not have to worry about the absence of a conjugate field and we can do our quantization in a way similar to before (but with some important differences). The Coulomb gauge is, however, non-covariant. For instance, A^0 changes from frame to frame, therefore it may be cumbersome to deal with Lorentz invariance. The Lorenz gauge is appealing because it is explicitly Lorentz covariant, but it contains non-physical degrees of freedom.

Quantization in Coulomb gauge is presented, e.g., in the first Ch. 1 of Mandl–Shaw, in the first part of Sec. 4.4 of Ryder, Quantization in the Lorenz gauge is developed, e.g., in Ch. 5 of Mandl–Shaw, in the second part of Sec. 4.4 of Ryder, in Sec 7.3.2 of Aitchison-Hey. The lecture notes of prof. Miglietta present both approaches.

[10] To be precise, the Lorenz gauge with $A^0 = 0$ is called the "radiation gauge" and corresponds to the Coulomb gauge only in the absence of charges.

3.6.2 Lorenz gauge

Instead of working with the original Lagrangian, we use

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}(\partial^\mu A_\mu)^2 - j^\mu A_\mu. \tag{3.245}$$

This Lagrangian is not gauge invariant anymore. It corresponds to the original Lagrangian only in the Lorenz gauge. The additional term we included is called "gauge-fixing" term.

To obtain the equations of motion we need

$$\frac{\partial\mathcal{L}}{\partial A_\mu} = -j^\mu, \tag{3.246}$$

$$\frac{\partial\mathcal{L}}{\partial(\partial_\nu A_\mu)} = -\partial^\nu A^\mu + \partial^\mu A^\nu - \frac{1}{\xi}g^{\mu\nu}\partial^\rho A_\rho. \tag{3.247}$$

The equation of motions are

$$\partial_\rho\partial^\rho A^\mu - (1 - 1/\xi)\partial^\mu\partial_\rho A^\rho = j^\mu. \tag{3.248}$$

The above equations do not correspond to Maxwell's equations. They do only with the extra condition imposed by the Lorenz gauge. In other words, the new Lagrangian correctly describes the physics only in the Lorenz gauge.

In the following, let us choose $\xi = 1$ (this choice is called "Feynman gauge") and consider the field in the absence of external currents. We then assume as a starting point the Lagrangian

$$\mathcal{L}' = N\left[-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}(\partial^\mu A_\mu)^2\right]. \tag{3.249}$$

The equations of motions, in the absence of currents, now are

$$\partial_\rho\partial^\rho A^\mu = 0, \tag{3.250}$$

i.e., they are completely analogous to a set of four Klein-Gordon equations without masses.

The conjugate fields are

$$\pi^\nu = \frac{\partial\mathcal{L}}{\partial(\partial_0 A_\nu)} = -F^{0\nu} - g^{0\nu}\partial_\mu A^\mu. \tag{3.251}$$

For the π^0 field, we have potentially a nonzero result. However, if we assume the validity of the Lorenz gauge condition in a "strong sense," i.e., at the level of the operators, we would obtain again a vanishing field. The solution is to consider the condition in a "weak sense," i.e., that the expectation value of the operator $\partial_\mu A^\mu$ vanishes on any physical state

$$\langle\Psi|\partial_\mu A^\mu|\Psi\rangle = 0. \tag{3.252}$$

We will see the consequence of this condition at the end of the chapter.

As usual, it is possible to start from a Lagrangian that differs by a four divergence. For instance it is convenient to use

$$\boxed{\mathcal{L}_{\rm EM} = -\frac{1}{2}(\partial_\mu A_\nu)\,\partial^\mu A^\nu.} \tag{3.253}$$

The difference between the two Lagrangians is

$$\begin{aligned}\mathcal{L}_{\rm EM} - \mathcal{L}' &= -\frac{1}{2}\cancel{(\partial_\mu A_\nu)\,\partial^\mu A^\nu} + \frac{1}{2}\Big(\cancel{(\partial_\mu A_\nu)\partial^\mu A^\nu} - (\partial^\mu A^\nu)\partial_\nu A_\mu\Big) + \frac{1}{2}(\partial^\mu A_\mu)^2 \\ &= \frac{1}{2}\Big(- (\partial^\mu A^\nu)\partial_\nu A_\mu + \partial^\mu A_\mu \partial_\nu A^\nu\Big) \\ &= \frac{1}{2}\Big(- \partial^\mu\big(A^\nu \partial_\nu A_\mu\big) + \partial^\mu\big(A_\mu \partial_\nu A^\nu\big)\Big)\end{aligned} \tag{3.254}$$

The equations of motions are unchanged. The conjugate fields, with the last version of the Lagrangian, are simpler

$$\pi^\nu = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_\nu)} = -\partial^0 A^\nu. \tag{3.255}$$

3.6.3 Polarization vectors

As we know from classical electrodynamics, the electromagnetic fields need to be described using polarization vectors. In Minkowski space there are in general four possible polarization vectors.

To build explicitly the polarization vectors, we can start with introducing a time-like vector n^μ (for which $n^2 > 0$) and set[11]

$$\epsilon_0^\mu = \frac{n^\mu}{\sqrt{n^2}}. \tag{3.256}$$

Then we choose one of the other three polarization vectors along the component of k orthogonal to ϵ_0 and with modulus -1, i.e.,

$$\epsilon_3^\mu = \frac{1}{\sqrt{n^2(k\cdot n)^2 - n^4 k^2}}\big(n^2 k^\mu - (k\cdot n)n^\mu\big). \tag{3.257}$$

Finally, the other two polarization vectors are chosen to be orthogonal to both n and k and with modulus -1.

We normally call $\lambda = 0$ timelike or scalar polarization, $\lambda = 1, 2$ transverse polarizations, and $\lambda = 3$ longitudinal polarization.

[11] We could also start from a space-like vector, normalized to -1, with the due attention to changing signs where required. Using a light-like vector requires a different strategy because it cannot be normalized to ± 1.

The polarization vectors fulfill the following properties:

$$\epsilon^{\mu}_{\lambda}(\vec{k})\epsilon^{*}_{\lambda'\mu}(\vec{k}) = g_{\lambda\lambda'}, \tag{3.258}$$

$$\sum_{\lambda,\lambda'} g_{\lambda\lambda'}\,\epsilon^{\mu *}_{\lambda}(\vec{k})\epsilon^{\nu}_{\lambda'}(\vec{k}) = g^{\mu\nu}. \tag{3.259}$$

Remember that the indices λ and λ' refer to the polarization states, not to Lorentz indices, even though $g_{\lambda\lambda'}$ is formally equal to the Minkowski metric tensor. To make a clear distinction, in Mandl–Shaw the notation $\zeta_\lambda\delta_{\lambda\lambda'}$ is used, where $\zeta_0 = -1$ and $\zeta_i = 1$ (therefore, $-\zeta_\lambda\delta_{\lambda\lambda'} = g_{\lambda\lambda'}$).

It is useful to write down the result of summing over the transverse polarization states only

$$\begin{aligned}\sum_{\lambda=1,2} \epsilon^{\mu *}_{\lambda}(\vec{k})\epsilon^{\nu}_{\lambda}(\vec{k}) &= -g^{\mu\nu} + \epsilon^{\mu *}_{0}(\vec{k})\epsilon^{\nu}_{0}(\vec{k}) - \epsilon^{\mu *}_{3}(\vec{k})\epsilon^{\nu}_{3}(\vec{k}) \\ &= -g^{\mu\nu} + \frac{1}{(k\cdot n)^2 - n^2k^2}\Big[k\cdot n\Big(k^\mu\, n^\nu + k^\nu\, n^\mu\Big) - n^2\, k^\mu\, k^\nu - k^2\, n^\mu\, n^\nu\Big].\end{aligned} \tag{3.260}$$

If we impose from the beginning that $k^2 = 0$ the situation considerably simplifies. The sum of the transverse polarization states becomes

$$\sum_{\lambda=1,2} \epsilon^{\mu *}_{\lambda}(\vec{k})\epsilon^{\nu}_{\lambda}(\vec{k}) = -g^{\mu\nu} + \frac{1}{k\cdot n}\Big(k^\mu\, n^\nu + k^\nu\, n^\mu\Big) - \frac{n^2}{(k\cdot n)^2}\, k^\mu\, k^\nu. \tag{3.261}$$

The sum of the two unphysical polarization vectors is in this case proportional to k^μ

$$\epsilon^{\mu}_{0}(\vec{k}) + \epsilon^{\mu}_{3}(\vec{k}) = \frac{\sqrt{n^2}}{k\cdot n}k^\mu. \tag{3.262}$$

We have also

$$k_\mu\epsilon^{\mu}_{0}(\vec{k}) = -k_\mu\epsilon^{\mu}_{3}(\vec{k}) = \frac{k\cdot n}{\sqrt{n^2}}. \tag{3.263}$$

In a frame where the momentum k is in the z direction, i.e., $k = (|\vec{k}|, 0, 0, |\vec{k}|)$, the polarization vectors can be written as

$$\epsilon^{\mu}_{0}(\vec{k}) = (1,0,0,0), \tag{3.264}$$

$$\epsilon^{\mu}_{1}(\vec{k}) = (0,1,0,0), \qquad \epsilon^{\mu}_{2}(\vec{k}) = (0,0,1,0), \qquad \epsilon^{\mu}_{3}(\vec{k}) = (0,0,0,1). \tag{3.265}$$

The transverse polarizations above correspond to the so-called linear transverse polarizations. For the circular transverse polarizations, one can choose

$$\epsilon^{\mu}_{+}(\vec{k}) = -\tfrac{1}{\sqrt{2}}(0,1,i,0), \qquad \epsilon^{\mu}_{-}(\vec{k}) = \tfrac{1}{\sqrt{2}}(0,1,-i,0). \tag{3.266}$$

In all cases, $k_\mu\epsilon^{\mu}_{1} = k_\mu\epsilon^{\mu}_{2} = k_\mu\epsilon^{\mu}_{+} = k_\mu\epsilon^{\mu}_{-} = 0$.

3.6.4 Solutions for the electromagnetic field

The general solution has the following form

$$A^\mu(x) = A^{\mu+}(x) + A^{\mu-}(x) = \sum_{\lambda=0}^{3} \int \frac{d^3k}{(2\pi)^3 2\omega_k} \Big(\epsilon^\mu_\lambda(\vec{k}) a_\lambda(\vec{k}) e^{-ikx} + \epsilon^{\mu*}_\lambda(\vec{k}) a^\dagger_\lambda(\vec{k}) e^{ikx} \Big), \tag{3.267}$$

with $k^0 = \omega_k = |\vec{k}|$. The only differences with respect to a real Klein–Gordon field are the fact that the mass is zero and the presence of the polarization vectors ϵ. They appear because now the field is a vector field, with four Lorentz components and with, in principle, four different polarizations.

In order to check in detail the form of the commutation relations, we need to compute Poisson brackets

$$\begin{aligned} \big\{ A^\mu(x), \pi^\nu(x') \big\}_{\text{PB}} &= \int d^3z \left(\frac{\partial A^\mu(x)}{\partial A^\rho(z)} \frac{\partial \pi^\nu(x')}{\partial \pi_\rho(z)} - \frac{\partial \pi^\nu(x')}{\partial A^\rho(z)} \frac{\partial A^\mu(x)}{\partial \pi_\rho(z)} \right) \\ &= \int d^3z \, \delta^\mu_\rho g^{\nu\sigma} \delta^\rho_\sigma \delta(\vec{x} - \vec{z}) \delta(\vec{x}' - \vec{z}) = g^{\mu\nu} \delta(\vec{x} - \vec{x}'). \end{aligned} \tag{3.268}$$

We want now to check the usual canonical commutation relations (we are dealing with a spin-1 field): [12]

$$\begin{aligned} \big[A^\mu(x), \pi^\nu(x') \big] \Big|_{x^0 = x^{0\prime}} = \sum_{\lambda,\lambda'} \int \frac{d^3k}{(2\pi)^3 2\omega_k} \int \frac{d^3k'}{(2\pi)^3 2\omega_{k'}} (-i\omega_{k'}) \\ \times \Big(- \epsilon^\mu_\lambda(\vec{k}) \epsilon^{\nu*}_{\lambda'}(\vec{k}') \big[a_\lambda(\vec{k}), a^\dagger_{\lambda'}(\vec{k}') \big] \, e^{-ikx} e^{ik'x'} \\ + \epsilon^{\mu*}_\lambda(\vec{k}) \epsilon^\nu_{\lambda'}(\vec{k}') \big[a^\dagger_\lambda(\vec{k}), a_{\lambda'}(\vec{k}') \big] \, e^{ikx} e^{-ik'x'} \Big) \Big|_{x^0 = x^{0\prime}}. \end{aligned} \tag{3.269}$$

In order to obtain the canonical commutation relations, we have to impose

$$\big[a_\lambda(\vec{k}), a^\dagger_{\lambda'}(\vec{k}') \big] = -g_{\lambda\lambda'} (2\pi)^3 2\omega_k \delta(\vec{k} - \vec{k}'). \tag{3.270}$$

Note that this relation has a peculiarity: the commutation relations for transverse and longitudinal polarizations are the usual ones, but the time-like polarization has an extra minus sign. We shall investigate later some of the consequences of this choice.

The final result for the commutation relations is

$$\begin{aligned} \big[A^\mu(x), \pi^\nu(x') \big] \Big|_{x^0 = x^{0\prime}} &= \int \frac{d^3k}{(2\pi)^3} \frac{i}{2} \Big(e^{-ik(x-x')} + e^{ik(x-x')} \Big) \sum_{\lambda,\lambda'} g_{\lambda\lambda'} \, \epsilon^{\mu*}_\lambda(\vec{k}) \epsilon^\nu_{\lambda'}(\vec{k}) \Big|_{x^0 = x^{0\prime}} \\ &= i g^{\mu\nu} \delta(\vec{x} - \vec{x}'). \end{aligned} \tag{3.271}$$

[12] We remove vanishing commutators from the beginning.

We could have also started from the Lagrangian in the form of Eq. (3.249). In this case the conjugate fields have additional terms with spatial derivatives and we should check what happens to commutators of this type

$$\begin{aligned}\left[A^\mu(x),\partial^i A^\nu\right]\Big|_{x^0=x^{0\prime}} = \sum_{\lambda,\lambda'}\int\frac{d^3k}{(2\pi)^3 2\omega_k}\int\frac{d^3k'}{(2\pi)^3 2\omega_{k'}}(ik'^i)\\ \times\Big(\epsilon^\mu_\lambda(\vec{k})\epsilon^{\nu *}_{\lambda'}(\vec{k}')\left[a_\lambda(\vec{k}),a^\dagger_{\lambda'}(\vec{k}')\right]e^{-ikx}e^{ik'x'}\\ -\epsilon^{\mu *}_\lambda(\vec{k})\epsilon^\nu_{\lambda'}(\vec{k}')\left[a^\dagger_\lambda(\vec{k}),a_{\lambda'}(\vec{k}')\right]e^{ikx}e^{-ik'x'}\Big)\Big|_{x^0=x^{0\prime}}.\end{aligned} \tag{3.272}$$

By inserting the commutation rules of Eq. (3.270) into the previous equation we obtain that the contributions containing the spatial derivatives vanish,

$$\begin{aligned}\left[A^\mu(x),\partial^i A^\nu\right]\Big|_{x^0=x^{0\prime}} &= \int\frac{d^3k}{(2\pi)^3 2\omega_k}\frac{i}{2}\Big(k^i e^{i\vec{k}\cdot(\vec{x}-\vec{x}')}+k^i e^{-i\vec{k}\cdot(\vec{x}-\vec{x}')}\Big)\sum_{\lambda,\lambda'}g_{\lambda\lambda'}\,\epsilon^{\mu *}_\lambda(\vec{k})\epsilon^\nu_{\lambda'}(\vec{k})\\ &= g^{\mu\nu}\int\frac{d^3k}{(2\pi)^3 2\omega_k}\frac{ik^i}{2}\cancel{\Big(e^{i\vec{k}\cdot(\vec{x}-\vec{x}')}-e^{i\vec{k}\cdot(\vec{x}-\vec{x}')}\Big)}=0.\end{aligned} \tag{3.273}$$

The fact that the contribution with spatial derivatives vanish is in agreement with the fact that we can use both expressions for the Lagrangian.

3.6.5 Hamiltonian of the electromagnetic field

Using the Lagrangian in the form of Eq. (3.253), it is easy to compute the Hamiltonian as

$$\begin{aligned}H &= \int d^3x\left(\pi_\mu\dot{A}^\mu-\mathcal{L}\right)=\int d^3x\left(-\dot{A}_\mu\dot{A}^\mu+\frac{1}{2}\dot{A}_\mu\dot{A}^\mu-\frac{1}{2}\vec{\nabla}A_\mu\cdot\vec{\nabla}A^\mu.\right)\\ &= -\int d^3x\,\frac{1}{2}\Big(\dot{A}^2+(\vec{\nabla}A)^2\Big).\end{aligned} \tag{3.274}$$

Calculations similar to the Klein–Gordon field (cf. Sec. 3.3.3) lead to the following result for the Hamiltonian (after the application of normal ordering)

$$\begin{aligned}H &= \int\frac{d^3k}{(2\pi)^3 2\omega_k}\,\omega_k\sum_\lambda(-g_{\lambda\lambda})\,a^\dagger_\lambda(\vec{k})a_\lambda(\vec{k})\\ &= \int\frac{d^3k}{(2\pi)^3 2\omega_k}\,\omega_k\Big(-a^\dagger_0(\vec{k})a_0(\vec{k})+\sum_{\lambda=1}^{3}a^\dagger_\lambda(\vec{k})a_\lambda(\vec{k})\Big).\end{aligned} \tag{3.275}$$

Due to the presence of the contribution of the time-like polarizations, the Hamiltonian may have negative expectation values. This is due to the negative sign in the commutation rules. Moreover, the norm of a state with a time-like photon (or any odd number of time-like photons) is negative:

$$\langle 1_{0\,k}|1_{0\,k}\rangle = \big\langle 0\big|a_0(\vec{k})a^\dagger_0(\vec{k})\big|0\big\rangle = \cancel{\big\langle 0\big|a^\dagger_0(\vec{k})a_0(\vec{k})\big|0\big\rangle}+\big\langle 0\big|[a_0(\vec{k}),a^\dagger_0(\vec{k})]\big|0\big\rangle = -(2\pi)^3 2k^0\delta(0). \tag{3.276}$$

To overcome these problems, we adopt the so-called Gupta–Bleuler procedure. We know that scalar and longitudinal polarizations should not be physical and are somehow artifacts of our choice of a "non-physical" gauge. We therefore make use of the Lorenz-gauge condition (in its weak sense). The situation is similar to the case of the equations of motion, where we needed to make use of the Lorenz condition to recover Maxwell's equations. Imposing the condition

$$\partial_\mu A^\mu \big|\Psi\big\rangle = 0 \tag{3.277}$$

is already too strong a requirement: not even the vacuum satisfies this condition. However, it is sufficient to assume

$$\partial_\mu A^{\mu+} \big|\Psi\big\rangle = 0 \tag{3.278}$$

to guarantee that

$$\big\langle\Psi\big|\partial_\mu A^\mu\big|\Psi\big\rangle = \big\langle\Psi\big|\partial_\mu A^{\mu-} + \partial_\mu A^{\mu+}\big|\Psi\big\rangle = \big\langle\Psi\big|\partial_\mu A^{\mu-}\big|\Psi\big\rangle = \big\langle\Psi\big|\partial_\mu A^{\mu+}\big|\Psi\big\rangle^* = 0. \tag{3.279}$$

The condition (3.278) implies that

$$\sum_\lambda k_\mu \epsilon^\mu_\lambda(\vec{k}) a_\lambda(\vec{k}) \big|\Psi\big\rangle = 0. \tag{3.280}$$

The two transverse polarization states do not contribute since they are orthogonal to k. The only nonzero contributions come from time-like and longitudinal photons and since $k_\mu \epsilon^\mu_0 = -k_\mu \epsilon^\mu_3$ we obtain

$$\big[a_0(\vec{k}) - a_3(\vec{k})\big]\big|\Psi\big\rangle = 0, \qquad \text{i.e.,} \quad a_0(\vec{k})\big|\Psi\big\rangle = a_3(\vec{k})\big|\Psi\big\rangle \tag{3.281}$$

and its adjoint

$$\big\langle\Psi\big|\big[a_0^\dagger(\vec{k}) - a_3^\dagger(\vec{k})\big] = 0, \qquad \text{i.e.,} \quad \big\langle\Psi\big|a_0^\dagger(\vec{k}) = \big\langle\Psi\big|a_3^\dagger(\vec{k}), \tag{3.282}$$

An example of a physical state is

$$\big|\Psi\big\rangle = \big(-a_0^\dagger(q) + a_3^\dagger(q)\big)\big|0\big\rangle \tag{3.283}$$

Due to the conditions above, in any physical state the contribution from time-like photons to any observable should cancel that of longitudinal photons, leaving only the physical contributions.

Let us take for instance the case of the Hamiltonian. Due to the above conditions we have

$$\big\langle\Psi\big|a_0^\dagger(\vec{k})a_0(\vec{k}) - a_3^\dagger(\vec{k})a_3(\vec{k})\big|\Psi\big\rangle = \big\langle\Psi\big|a_0^\dagger(\vec{k})\big[a_0(\vec{k}) - a_3(\vec{k})\big]\big|\Psi\big\rangle = 0. \tag{3.284}$$

In the end, the net result is that the expectation value of the Hamiltonian on a physical state is

$$\boxed{\big\langle\Psi\big|H\big|\Psi\big\rangle = \big\langle\Psi\big|\int \frac{d^3k}{(2\pi)^3 2\omega_k}\,\omega_k \sum_{\lambda=1}^{2} a_\lambda^\dagger(\vec{k})a_\lambda(\vec{k})\big|\Psi\big\rangle,} \tag{3.285}$$

i.e., only the transverse polarizations give a contribution to the energy.

3.6.6 Coulomb gauge [optional]

The equations of motion (in the absence of currents) were

$$\partial_\nu F^{\nu\mu} = \partial_\nu \partial^\nu A^\mu - \partial^\mu \partial_\nu A^\nu = 0 \tag{3.286}$$

For $\mu = 0$

$$\cancel{\partial_t^2 \Phi} - \vec{\nabla}^2 \Phi - \partial_t (\cancel{\partial_t \Phi} - \vec{\nabla} \cdot \vec{A}) = -\vec{\nabla}^2 \Phi - \partial_t (-\cancel{\vec{\nabla} \cdot \vec{A}}) = 0. \tag{3.287}$$

This means that in Coulomb gauge the scalar potential has to fulfill Laplace's equation. If we choose the boundary condition $\Phi(t, \infty) = 0$ the solution is trivially $\Phi(x) = 0$. Due to the gauge condition, the time component of our field A^μ is not a dynamical field anymore. When charges are present, Φ may be different from zero, but it is entirely given by the distribution of charges and is not a dynamical field. It can be defined as an "auxiliary field." Auxiliary fields are used to deal with systems with constraints. In this case, our system is constrained by the Coulomb gauge condition.

We can stop considering the A^0 field and continue the discussion with the A^i fields only. The classical solutions for the potential are

$$\vec{A}(x) = \vec{A}^+(x) + \vec{A}^-(x) = \sum_{\lambda=1}^{2} \int \frac{d^3k}{(2\pi)^3 2\omega_k} \left(\vec{\epsilon}_\lambda(\vec{k}) a_\lambda(\vec{k}) e^{-ikx} + \vec{\epsilon}^*_\lambda(\vec{k}) a^\dagger_\lambda(\vec{k}) e^{ikx} \right), \tag{3.288}$$

with $k^0 = \omega_k = |\vec{k}|$.

In this case, we have only two possible polarization states. This is due to the fact that the choice of gauge imposes

$$\vec{k} \cdot \vec{\epsilon}_\lambda = 0. \tag{3.289}$$

The conjugate fields are

$$\pi^i = -F^{0i} = E^i. \tag{3.290}$$

We can proceed now with canonical quantization and impose the commutation rules on the a and $a^\dagger$ operators in the usual form. It turns out however that the commutation rules for the fields and their conjugate are nontrivial (see, e.g., Ryder)

$$\left[A^i(x), \pi^j(x') \right] \Big|_{x^0 = x^{0\prime}} = \int \frac{d^3k}{(2\pi)^3} \left(\delta^{ij} - \frac{k^i k^j}{|\vec{k}|^2} \right) e^{i\vec{k}\cdot\vec{x}}. \tag{3.291}$$

3.7 Conclusions

In this chapter, we have dealt with field theories and how to quantize them. To do this, we promoted the fields to operators and imposed commutation or anticommutation rules to the fields and their conjugate, or to the corresponding creation and annihilation operators. We have seen that scalar and vector fields (in general, bosonic fields) must be quantized using commutation relations and Dirac fields (in general, fermionic fields) must be

quantized using anticommutation relations. We have seen how to derive the fundamental properties of these free fields and in particular we have seen that they always lead to positive energies. In the next chapter, we will see out to deal with interactions between the fields and compute some the cross sections for simple scattering processes.

Typical questions that can come out during the exam:

1. Prove Nöther's theorem for internal symmetries;
2. Derive and discuss the energy-momentum tensor;
3. Discuss the quantization of bosonic and fermionic harmonic oscillators;
4. Describe the procedure of field quantization;
5. Derive the Hamiltonian of the Klein–Gordon field;
6. Derive the Hamiltonian of the Dirac field;
7. Discuss the quantization of the electromagnetic field.

3.A Variation of the action in presence of explicit dependence on coordinates

In Eq. (3.33) we assumed that the Lagrangian did not depend explicitly on the coordinates. If we relax this assumption, we need to include some extra terms (see, e.g., [29, Sec. 3.2]).

$$\delta S = \int_\Omega d^4x' \, \mathcal{L}(\phi', \partial_\mu \phi', x'^\mu) - \int_\Omega d^4x \, \mathcal{L}(\phi, \partial_\mu \phi, x^\mu) \tag{3.292}$$

We describe the coordinate transformation as

$$x^\mu \to x'^\mu = x^\mu + \delta x^\mu \tag{3.293}$$

We need to consider the Jacobian of this transformation

$$d^4x' = \left| \frac{\partial x'^\mu}{\partial x^\lambda} \right| d^4x = \left| \delta^\mu_\lambda + \partial_\lambda \delta x^\mu \right| d^4x \approx [1 + \partial_\mu(\delta x^\mu)] d^4x \tag{3.294}$$

To understand the last step, let us check a two-dimensional example

$$\det \begin{pmatrix} 1 + \partial_0 \delta x^0 & \partial_0 \delta x^1 \\ \partial_1 \delta x^0 & 1 + \partial_1 \delta x^1 \end{pmatrix} \approx 1 + \partial_0 \delta x^0 + \partial_1 \delta x^1 \tag{3.295}$$

all other terms of the determinant are higher-order in the variation.

Including the effect of the Jacobian, we obtain for the variation of the action

$$\delta S = \int_\Omega d^4x \Big\{ \mathcal{L}(\phi', \partial_\mu \phi', x'^\mu) - \mathcal{L}(\phi, \partial_\mu \phi, x^\mu) + \mathcal{L}(\phi', \partial_\mu \phi', x'^\mu) \partial_\mu (\delta x^\mu) \Big\} \tag{3.296}$$

writing the Lagrangian as

$$\mathcal{L}(\phi', \partial_\mu \phi', x'^\mu) = \mathcal{L}(\phi, \partial_\mu \phi, x^\mu) + \frac{\partial \mathcal{L}}{\partial \phi_r} \delta \phi_r + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_r)} \partial_\mu (\delta \phi_r) + \big(\partial_\mu \mathcal{L}\big) \delta x^\mu \tag{3.297}$$

and stopping at first order in the variations, we get

$$\begin{aligned} \delta S &= \int_\Omega d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \phi_r} \delta \phi_r + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_r)} \partial_\mu (\delta \phi_r) + \big(\partial_\mu \mathcal{L}\big) \delta x^\mu + \mathcal{L} \partial_\mu (\delta x^\mu) \right\} \\ &= \int_\Omega d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \phi_r} \delta \phi_r + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_r)} \partial_\mu (\delta \phi_r) + \partial_\mu \big(\mathcal{L} \delta x^\mu \big) \right\} \end{aligned} \tag{3.298}$$

which eventually leads to Eq. (3.35).

4

Interacting quantum fields

In the previous chapter we introduced the essential concepts in the description of free fields. However, fields with no interactions are quite dull. Interesting phenomena emerge when we put different fields together and we assume the presence of interactions among them. The resulting Lagrangian density will typically be the sum of the free Lagrangians plus an interaction Lagrangian

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_I. \tag{4.1}$$

The interaction Lagrangian contains a parameter that defines the strength of the interaction, the so-called "coupling constant."

Finding the analytic solutions of the corresponding equation of motions is normally unfeasible. Therefore, the path to follow is to consider the interaction as a perturbation of the free-field situation. In other words, the problem can be studied as an expansion in the coupling constants, truncated to a certain order.

4.1 Lagrangians with interactions

For an extended discussion, see Sec. 4.1 of Peskin–Schroeder.

The ways to include interactions between fields are restricted by the usual requirements that the Lagrangian density must be a Lorentz scalar and must be Hermitian. Normally, we also prefer to avoid introducing derivatives of the fields in the interaction part. This guarantees that the conjugate fields of the interacting theory correspond still to the

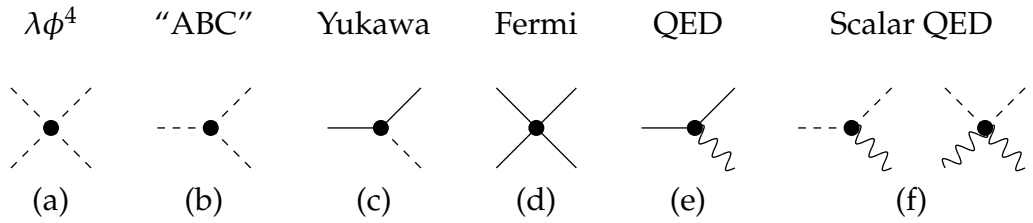


Figure 4.1: Interactions between fields can be represented diagrammatically in terms of vertices.

free-field theory. All canonical commutation or anticommutation rules are thus unaffected.

In spite of the constraints, the possibilities to build a Lagrangian with interactions are endless. A strong constrain is posed by the possibility of *renormalizing* the theory. The rule of thumb in this case is: renormalizable theories must have coupling constant with dimensions $[M]^d$ with $d \geq 0$ (see, e.g., Sec. 5.6 of Maggiore for a relatively concise discussion of this issue).

Apart from QED (which we are going to analyze in the next section), instructive examples are the so-called ϕ^4 *theory*, *"ABC" theory*, *Yukawa theory*,and *scalar QED*.

- In the ϕ^4 theory, a real scalar field interacts with itself. The Lagrangian is

$$\mathcal{L}_{\phi^4} = \mathcal{L}_{\mathrm{KG}} + \mathcal{L}_{\mathrm{I}} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4. \tag{4.2}$$

  The fields have dimensions $[M]$, therefore the coupling λ is adimensional and the theory is renormalizable. The interaction could be described by the symbolic *vertex* in Fig. 4.1 (a).

- The "ABC" theory (see Aitchison–Hey) is constructed with three distinct real scalar fields that interact together

$$\mathcal{L}_{ABC} = \mathcal{L}_{\mathrm{KGA}} + \mathcal{L}_{\mathrm{KGB}} + \mathcal{L}_{\mathrm{KGC}} + \mathcal{L}_{\mathrm{I}} = \sum_{i=A,B,C}\left(\frac{1}{2}\partial_\mu\phi_i\partial^\mu\phi_i - \frac{1}{2}m_i^2\phi_i^2\right) - \lambda\phi_A\phi_B\phi_C. \tag{4.3}$$

  The coupling constant has dimension $[M]^1$. The corresponding vertex is depicted in Fig. 4.1 (b).

- The Yukawa theory describes the interaction between a scalar field and a fermionic field

$$\mathcal{L}_{\mathrm{Yuk}} = \mathcal{L}_{\mathrm{Dir}} + \mathcal{L}_{\mathrm{KG}} + \mathcal{L}_{\mathrm{I}} = \overline{\psi}(i\partial\!\!\!/ - m)\psi + \frac{1}{2}\left(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2\right) + g\overline{\psi}\psi\phi. \tag{4.4}$$

  The field ϕ has dimension $[M]$, while the field ψ has dimension $[M]^{3/2}$, therefore the coupling constant is adimensional also in this case. The Yukawa theory can be used to describe the interactions between nucleons and pions. The corresponding vertex is depicted in Fig. 4.1 (c).

- We will shortly mention scalar QED at the end of the next section. There are in this case two vertices, depicted in Fig. 4.1 (f).

An example of nonrenormalizable theory is a theory containing a four-fermion interaction

$$\mathcal{L}_{\mathrm{I}} = G_F \overline{\psi}\psi\overline{\psi}\psi. \tag{4.5}$$

In this case, the coupling constant must have dimensions $[M]^{-2}$. This Lagrangian was used by Fermi described the production of β radiation inside the nucleus in 1933. It represented the first application of quantum field theory outside QED. Although valid only approximately, it gave account of the fact that the neutron can decay in a proton and an electron, without being necessarily constituents of the neutron, and in addition it includes the emission of a neutrino, the particle that Pauli proposed to resolve the problem of the missing energy in β decay. The corresponding vertex is depicted in Fig. 4.1 (d).

4.1.1 The QED Lagrangian

In this course, we are mainly interested in the QED case, where fermions interact with the electromagnetic field.

A way to derive the form of the QED Lagrangian is to invoke the principle of *local* gauge symmetry. We know that the Dirac Lagrangian is invariant under a global gauge transformation (i.e., changing the phase of the fields by a constant amount everywhere). What happens if we require the theory to be invariant under a local gauge transformation, i.e., a different phase change for any point in space? The transformation is

$$\psi(x) \to \psi'(x) = e^{-iq\chi(x)}\psi(x). \tag{4.6}$$

The original Dirac Lagrangian is not invariant under this transformation, because

$$i\partial_\mu\psi' = e^{-iq\chi(x)}i\partial_\mu\psi(x) + qe^{-iq\chi(x)}\psi(x)\partial_\mu\chi(x), \tag{4.7}$$

which means

$$\mathcal{L}'_{\mathrm{Dir}} - \mathcal{L}_{\mathrm{Dir}} = q\overline{\psi}\gamma^\mu\psi\partial_\mu\chi \tag{4.8}$$

To cancel the symmetry-breaking term, we are looking for an interaction term that transforms as

$$\mathcal{L}'_{\mathrm{I}} - \mathcal{L}_{\mathrm{I}} = -q\overline{\psi}\gamma^\mu\psi\partial_\mu\chi. \tag{4.9}$$

We discover that if we make the electron interact with the EM field

$$\mathcal{L}_{\mathrm{I}} - q\overline{\psi}\gamma^\mu\psi A_\mu \tag{4.10}$$

then the result is invariant under the combined transformation

$$A^\mu(x) \to A^{\mu\prime}(x) = A^\mu(x) + \partial^\mu\chi(x) \qquad \psi(x) \to \psi'(x) = e^{-iq\chi(x)}\psi(x). \tag{4.11}$$

The first transformation is the usual gauge transformation of the EM potential. This is the motivation to call the phase transformation applied to the Dirac fields a gauge transformation (global or local).

If we introduce the useful concept of the covariant derivative[1]

$$D_\mu = \partial_\mu + iqA_\mu(x) \tag{4.12}$$

where q is the charge of the particle (for the electron $q = -e$), we can write the Lagrangian as

$$\mathcal{L} = \overline{\psi}(i\not{D} - m)\psi. \tag{4.13}$$

The starting QED Lagrangian can be written as (for the electron $q = -e$),

$$\boxed{\mathcal{L}_{\text{QED}} = \mathcal{L}_{\text{Dir}} + \mathcal{L}_{\text{EM}} + \mathcal{L}_{\text{I}} = \overline{\psi}(i\not{\partial} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + e\overline{\psi}\gamma^\mu\psi A_\mu.} \tag{4.14}$$

The last term corresponds to the interaction of the electromagnetic field with a current $j^\mu = -e\overline{\psi}\gamma^\mu\psi$, which is the conserved current of the Dirac field associated to global gauge invariance. The parameter e represents the electric charge, which is also connected to the fine structure constant

$$\alpha = \frac{e^2}{4\pi} \approx \frac{1}{137}. \tag{4.15}$$

This number represents a measure of the strength of the interaction between the fields. It is a small number, which justifies a perturbative treatment.

Note that the interaction Lagrangian does not contain derivatives of the fields.

Using the covariant derivative we can write the QED Lagrangian as

$$\mathcal{L} = \overline{\psi}(i\not{D} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{4.16}$$

The equations or motion for the interacting theory are different. They turn out to be

$$(i\not{D} - m)\psi = 0, \tag{4.17}$$

$$\partial_\nu F^{\nu\mu} = -e\overline{\psi}\gamma^\mu\psi. \tag{4.18}$$

They correspond, respectively, to the Dirac equation coupled to the electromagnetic field by the minimal coupling prescription and to Maxwell's equation coupled to the electric current associated to the fermion field.

Let us briefly discuss also scalar QED. In this case, photons are coupled to a complex Klein–Gordon field. The recipe to obtain the interaction is similar to the QED case: it is sufficient to replace the normal derivatives of the Klein–Gordon Lagrangian with covariant derivatives. The Lagrangian is then gauge-invariant. When expanding the covariant

[1] Note that there can be different conventions for the sign in front of q in the covariant derivative. This is ultimately irrelevant, because it has to do with the choice of what we consider to be particles and antiparticles.

derivative, two types of interaction occur: a vertex with one photon field and two scalar fields and a vertex with two photons and two scalars.

$$\begin{aligned}\mathcal{L}_{\text{sc. QED}} &= \mathcal{L}_{\text{KG*}} + \mathcal{L}_{\text{EM}} + \mathcal{L}_{\text{I}} \\ &= \partial_\mu \phi^\dagger \partial^\mu \phi - m^2 \phi^\dagger \phi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - ieA_\mu \left(\phi^\dagger \partial^\mu \phi - \phi \partial^\mu \phi^\dagger\right) + e^2 A_\mu A^\mu \phi^\dagger \phi.\end{aligned} \tag{4.19}$$

It may seem problematic to have an interaction which contains the derivatives of the fields, but in reality this case is still treatable with standard techniques (see, e.g., the discussion at the end of Sec. 7.4 of Aitchison–Hey).

4.2 The scattering matrix

The material of this section is presented in Quantum Mechanics textbooks (see, e.g., the treatment of time-dependent perturbation theory in Sakurai, the treatment of the scattering matrix in the book by Boffi), but it is also reviewed in, e.g., Mandl–Shaw, Sec. 6.2, and Peskin-Schroeder, Sec. 4.5.

In quantum mechanics, the effect of perturbations can be described by means of the scattering matrix. The scattering matrix describes the probability to make a transition from an initial state $\big|i\big\rangle$, long before the scattering occurs, to a final state $\big|f\big\rangle$, long after the scattering occurs. The initial and final states are considered to be eigenstates of the unperturbed Hamiltonian. The definition of the scattering matrix can be taken as

$$\big|f\big\rangle = \big|\Psi(t=\infty)\big\rangle = S\big|\Psi(t=-\infty)\big\rangle = S\big|i\big\rangle. \tag{4.20}$$

The transition from $\big|i\big\rangle$ to $\big|f\big\rangle$ is then associated to the matrix elements of the scattering matrix

$$S_{fi} = \big\langle f\big|S\big|i\big\rangle. \tag{4.21}$$

The scattering matrix corresponds to the limit of the time-evolution operator in Dirac picture

$$S \equiv \lim_{\substack{t_i \to -\infty \\ t_f \to \infty}} U(t_f, t_i). \tag{4.22}$$

The scattering matrix is unitary, i.e.,

$$SS^\dagger = S^\dagger S = 1 \sum_f |S_{fi}|^2 = 1, \tag{4.23}$$

which is a mathematical way to express the conservation of probability.

Without deriving it, we give for granted the following expansion (Dyson's expansion) of the S matrix

$$\begin{aligned}S &= \sum_{n=0}^{\infty} (-i)^n \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 \ldots \int_{-\infty}^{t_{n-1}} dt_n H_I(t_1) H_I(t_2) \ldots H_I(t_n) \\ &= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \ldots \int_{-\infty}^{\infty} dt_n T\Big[H_I(t_1) H_I(t_2) \ldots H_I(t_n)\Big],\end{aligned} \tag{4.24}$$

where we have used the time-ordered product $T[\ldots]$, yet another form of operator ordering. The operators are ordered so that the later times stand to the left of earlier times.

Note that the S matrix can be written as an exponential, since its definition corresponds to the expansion of an exponential. Moreover, note that at "zeroth" order in the couplings the S matrix is just the identity, which means that nothing happens to the initial states.

We can use the above results almost directly. There are only two things we have to care about. First, we need to switch to Hamiltonian density, which is a trivial step. Second, we have to keep in mind that switching the order of fermionic operators involves adding an extra minus sign.

The time-ordered product for two operators can be written explicitly as

$$T\big[A(x_1)B(x_2)\big] = \theta(t_1 - t_2)A(x_1)B(x_2) + (-1)^P\theta(t_2 - t_1)B(x_2)A(x_1). \tag{4.25}$$

where P is the number of interchanges of neighboring fermion operators required to achieve the desired time ordering. For instance, $P = 0$ if A and B are two bosonic fields, and $P = 1$ if A and B are two fermionic fields. Note, however, that the fermion interchanges are essentially irrelevant when it comes to reordering interaction Hamiltonians, which in all cases of interest for us always contain two fermionic fields (in order to have a scalar Lagrangian density).

In summary, we can write

$$\boxed{S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} d^4x_1 \int_{-\infty}^{\infty} d^4x_2 \ldots \int_{-\infty}^{\infty} d^4x_n T\Big[\mathcal{H}_I(x_1)\mathcal{H}_I(x_2)\ldots\mathcal{H}_I(x_n)\Big].} \tag{4.26}$$

For convenience, we can rewrite the Dyson expansion in this way

$$S = \sum_{n=0}^{\infty} S^{(n)} \tag{4.27}$$

where

$$S^{(n)} = \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} d^4x_1 \int_{-\infty}^{\infty} d^4x_2 \ldots \int_{-\infty}^{\infty} d^4x_n T\Big[\mathcal{H}_I(x_1)\mathcal{H}_I(x_2)\ldots\mathcal{H}_I(x_n)\Big]. \tag{4.28}$$

4.3 Scattering matrix and Feynman diagrams

We now consider the problem of computing the scattering matrix and in particular the single terms of the Dyson expansion, see Eq. (4.28). We note that, if the interaction Lagrangian does not contain derivatives of the fields, we have

$$\mathcal{H}_I = -\mathcal{L}_I. \tag{4.29}$$

The zeroth order term, $S^{(0)}$, is trivially the identity. No scattering takes place in this case.

4.3.1 First-order contributions in the "ABC" theory

As an illustration of the calculation of a scattering amplitude, we first consider the "ABC" theory discussed in Aitchison–Hey, Sec. 6.3.

The interaction Hamiltonian reads

$$\begin{aligned}\mathcal{H}_I &= \lambda N\Big[\phi_A(x)\phi_B(x)\phi_C(x)\Big] \\ &= \lambda N\Big[\big(\phi_A^+(x)+\phi_A^-(x)\big)\big(\phi_B^+(x)+\phi_B^-(x)\big)\big(\phi_C^+(x)+\phi_C^-(x)\big)\Big].\end{aligned} \tag{4.30}$$

We remind ourselves the expression and meaning of the components of the field operators

$$\phi^+(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, a(\vec{k})e^{-ikx} \qquad \text{particle absorption,} \tag{4.31}$$

$$\phi^-(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, a^\dagger(\vec{k})e^{ikx} \qquad \text{particle creation.} \tag{4.32}$$

There are eight possible terms in the expansion of the above interaction, each corresponding to a different process. For illustration purposes, let's start from a process that starts with particle C and ends with particle A and B.

The initial state is (in the following, let us skip the vector notation inside the argument of the ladder operators)

$$\big|i\big\rangle = a_C^\dagger(p_C)\big|0\big\rangle, \tag{4.33}$$

the final state is

$$\big\langle f\big| = \big\langle 0\big|a_A(p_A)a_B(p_B). \tag{4.34}$$

The term of the scattering-matrix expansion under consideration is

$$S_{(a)}^{(1)} = -i\lambda\int_{-\infty}^{\infty} d^4x\; N\Big[\phi_A^-(x)\phi_B^-(x)\phi_C^+(x)\Big]. \tag{4.35}$$

No other terms can contribute to the transition we want to compute, i.e., the expectation value of any other combination of the fields would vanish between those particular initial and final states. Note that the operator is in this case already normal ordered, which is however not relevant because the A, B, and C field operators commute.

The transition amplitude from the initial to final states defined above is

$$\big\langle f\big|S_{(a)}^{(1)}\big|i\big\rangle = -i\lambda\int d^4x\,\big\langle 0\big|a_A(p_A)a_B(p_B)\Big(\phi_A^-(x)\phi_B^-(x)\phi_C^+(x)\Big)a_C^\dagger(p_C)\big|0\big\rangle. \tag{4.36}$$

We replace the usual expression for the field operators

$$\begin{aligned}\big\langle f\big|S_{(a)}^{(1)}\big|i\big\rangle = -i\lambda\int d^4x\,&\big\langle 0\big|a_A(p_A)a_B(p_B) \\ &\times\int\frac{d^3k}{(2\pi)^3 2\omega_k}a_A^\dagger(k)e^{ikx}\int\frac{d^3q}{(2\pi)^3 2\omega_q}a_B^\dagger(q)e^{iqx}\int\frac{d^3k'}{(2\pi)^3 2\omega_{k'}}a_C(k')e^{-ik'x} \\ &\times a_C^\dagger(p_C)\big|0\big\rangle.\end{aligned} \tag{4.37}$$

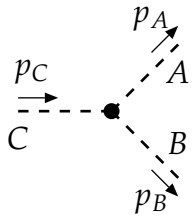


Figure 4.2: Feynman diagram corresponding to the decay $C \to AB$ at first order in the "ABC" theory containing three different scalar fields.

We have now expressions such as

$$a_C(k')a_C^\dagger(p_C)|0\rangle = [a_C(k'), a_C^\dagger(p_C)]|0\rangle + \cancel{a_C^\dagger(p_C)a_C(k')|0\rangle} = (2\pi)^3 2\omega_{k'}\delta(\vec{k}' - \vec{p}_C)|0\rangle, \tag{4.38}$$

and similarly for the A and B fields. We see here why there could not be contributions from other terms in the interaction Hamiltonian.

Replacing these results into the transition amplitude

$$\begin{aligned}\langle f|S_{(a)}^{(1)}|i\rangle &= -i\lambda \int d^4x\, e^{-i(p_C - p_B - p_A)x} \\ &= -(2\pi)^4 i\lambda\, \delta^4(p_C - p_B - p_A).\end{aligned} \tag{4.39}$$

The last delta function expresses the conservation of four-momentum.

We can rewrite the above result introducing the so-called Feynman amplitude $\mathcal{M}$

$$\langle f|S_{(a)}^{(1)}|i\rangle = (2\pi)^4\, \delta^4(p_C - p_B - p_A)\; i\mathcal{M}_{(a)}^{(1)}. \tag{4.40}$$

where

$$i\mathcal{M}_{(a)}^{(1)} = -i\lambda. \tag{4.41}$$

The scattering amplitude $\mathcal{M}$ is often defined with different prefactors, e.g., 1 or $-i$ instead of i. This detail is irrelevant in the end, since observables are related to the modulus squared of the amplitude.

The scattering amplitude can be described by the Feynman diagram drawn in Fig. 4.2. In this simple case, there is only one Feynman rule to take into consideration: to each vertex, we need to associate a factor $-i\lambda$.

What kind of observable can we compute with this result? We can study the decay rate of the process under study, and consequently the half-life. Without proving it, we give the formula for the decay rate of a particle with mass m and momentum p into n particles with momenta p_f (see, e.g., Peskin-Schroeder, end of Sec. 4.5)

$$d\Gamma = \frac{1}{2m}\,|\mathcal{M}|^2 \prod_f \frac{d^3p_f}{(2\pi)^3 2E_f}\,(2\pi)^4\delta^4(p - \Sigma p_f). \tag{4.42}$$

The formula is valid in the particle rest-frame, which is also the center-of-mass frame (CMF) of the decay products.

We can specialize the above formula for the decay $C \to AB$

$$d\Gamma = \frac{1}{2m_C}\,|\mathcal{M}|^2 \frac{d^3p_A}{(2\pi)^3 2E_A}\frac{d^3p_B}{(2\pi)^3 2E_B}\,(2\pi)^4\delta^4\big(p_C - p_A - p_B\big). \tag{4.43}$$

At this point, we have to decide with respect to which variables we want our decay rate to be differential. Clearly, the three components of momentum of particle A and the three components of particle B are redundant, because the four-dimensional delta function introduces four constraints. In principle, we can keep the decay rate differential with respect to two variables. We could for instance choose to study the angular distribution of the decay products in the center-of-mass of the system. The involved momenta are

$$p_C \overset{\text{CMF}}{=} \big(m_C,\ \ 0,\ \ 0,\ \ 0\big)\,, \tag{4.44}$$

$$p_A \overset{\text{CMF}}{=} \big(E_A,\ \ |\vec{p}_A|\sin\theta\cos\phi,\ \ |\vec{p}_A|\sin\theta\sin\phi,\ \ |\vec{p}_A|\cos\theta\big)\,, \tag{4.45}$$

$$p_B \overset{\text{CMF}}{=} \big(E_B,\ \ -|\vec{p}_A|\sin\theta\cos\phi,\ \ -|\vec{p}_A|\sin\theta\sin\phi,\ \ -|\vec{p}_A|\cos\theta\big)\,, \tag{4.46}$$

with $E_A = \sqrt{|\vec{p}_A|^2 + m_A^2}$ and similarly for E_B. On top of this, we must have $E_A + E_B = m_C$, which puts in turn the constraint

$$|\vec{p}_A| = \frac{1}{2m_C}\sqrt{\Big(m_C^2 - (m_A+m_B)^2\Big)\Big(m_C^2 - (m_A-m_B)^2\Big)}. \tag{4.47}$$

There are several ways to obtain the final result in a convenient way. For instance, we can follow these steps

$$\begin{aligned}
&\frac{d^3p_A}{(2\pi)^3 2E_A}\frac{d^3p_B}{(2\pi)^3 2E_B}\,(2\pi)^4\delta^4\big(p_C - p_A - p_B\big)\\
&\qquad= \frac{d^3p_A}{(2\pi)^3 2E_A}\frac{d^3p_B}{(2\pi)^3 2E_B}\,(2\pi)^4\delta^3\big(\vec{p}_C - \vec{p}_A - \vec{p}_B\big)\delta\big(E_C - E_A - E_B\big)\\
&\qquad= \frac{|\vec{p}_A|^2 d|\vec{p}_A| d\Omega}{(2\pi)^3 2E_A}\frac{1}{2E_B}\,(2\pi)\delta\big(E_C - E_A - E_B\big)\Big|_{\vec{p}_C = \vec{p}_A + \vec{p}_B}\\
&\qquad= \frac{|\vec{p}_A|^2 d\Omega}{(2\pi)^2 2E_A 2E_B}\left|\frac{\partial(E_A+E_B)}{\partial|\vec{p}_A|}\right|^{-1}\Big|_{p_C = p_A + p_B}.
\end{aligned} \tag{4.48}$$

In the last step, we used the properties of the delta function to change from a delta function expressed in terms of $E_A + E_B$ to an expression in terms of $|\vec{p}_A|$.

The above steps are general. If we want now to use the center-of-mass variables, we ought to write

$$\begin{aligned}
\frac{\partial(E_A+E_B)}{\partial|\vec{p}_A|} &= \frac{\partial}{\partial|\vec{p}_A|}\Big(\sqrt{|\vec{p}_A|^2 + m_A^2} + \sqrt{|\vec{p}_A|^2 + m_B^2}\Big) = |\vec{p}_A|\left(\frac{1}{E_A} + \frac{1}{E_B}\right)\\
&= \frac{|\vec{p}_A|\,m_c}{E_A\,E_B}.
\end{aligned} \tag{4.49}$$

Putting things together, including the calculation of $\mathcal{M}$, we obtain the following expression for the decay rate

$$d\Gamma = \frac{1}{32\pi^2}\frac{|\vec{p}_A|}{m_C^2}\lambda^2 d\Omega. \tag{4.50}$$

Note that the decay rate in this case is isotropic, i.e., independent of the angles. Breaking of isotropy can occur if we deal with particles with spin, but not with scalar particles. Note also that the decay rate would go to zero if $m_C \leq m_A + m_B$ (due to Eq. (4.47)).

Integration over the solid angle in this case gives just a 4π factor, i.e.,

$$\Gamma = \frac{1}{8\pi}\frac{|\vec{p}_A|}{m_C^2}\lambda^2. \tag{4.51}$$

The total decay rate has the correct dimensions of a mass or energy (due to the fact that λ in this theory has the dimension of a mass). In natural units, the inverse of the decay rate has the dimesions $[M]^{-1}$, which multiplied by $\hbar/c^2$ gives a time in standard units and corresponds to the lifetime of state C.

4.3.2 First-order contributions in QED

For QED, the interaction Hamiltonian reads

$$\begin{aligned}\mathcal{H}_I &= -eN\left[\overline{\psi}\not{A}\psi\right] \\ &= -eN\left[\left(\overline{\psi}^+ + \overline{\psi}^-\right)\left(\not{A}^+ + \not{A}^-\right)\left(\psi^+ + \psi^-\right)\right].\end{aligned} \tag{4.52}$$

We remind ourselves once more the meaning of the different components of the fields

$$\overline{\psi}^+(x) = \sum_s \int \frac{d^3k}{(2\pi)^3 2\omega_k} d_s(\vec{k})\overline{v}_s(\vec{k})e^{-ikx} \qquad \text{positron absorption,} \tag{4.53}$$

$$\overline{\psi}^-(x) = \sum_s \int \frac{d^3k}{(2\pi)^3 2\omega_k} c_s^\dagger(\vec{k})\overline{u}_s(\vec{k})e^{ikx} \qquad \text{electron creation,} \tag{4.54}$$

$$A^{\mu+}(x) = \sum_{\lambda=0}^{3} \int \frac{d^3k}{(2\pi)^3 2\omega_k} \epsilon_\lambda^\mu(\vec{k})a_\lambda(\vec{k})e^{-ikx} \qquad \text{photon absorption,} \tag{4.55}$$

$$A^{\mu-}(x) = \sum_{\lambda=0}^{3} \int \frac{d^3k}{(2\pi)^3 2\omega_k} \epsilon_\lambda^{\mu*}(\vec{k})a_\lambda^\dagger(\vec{k})e^{ikx} \qquad \text{photon creation,} \tag{4.56}$$

$$\psi^+(x) = \sum_s \int \frac{d^3k}{(2\pi)^3 2\omega_k} c_s(\vec{k})u_s(\vec{k})e^{-ikx} \qquad \text{electron absorption,} \tag{4.57}$$

$$\psi^-(x) = \sum_s \int \frac{d^3k}{(2\pi)^3 2\omega_k} d_s^\dagger(\vec{k})v_s(\vec{k})e^{ikx} \qquad \text{positron creation.} \tag{4.58}$$

There are eight possible terms in the expansion of the above interaction. For illustration purposes, let us start from the term $\overline{\psi}^-\not{A}^-\psi^+$. This term can be nonzero only if in

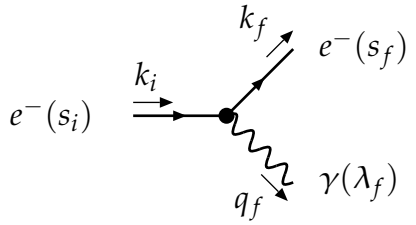


*Figure 4.3: Feynman diagram corresponding to the emission of a photon by an electron (*Bremsstrahlung*) at first order in QED.*

the initial state there is an electron (which is then destroyed by ψ^+) and the final state contains an electron and a photon. Therefore, it describes the emission of a photon from an electron. This process is depicted in the Feynman diagram in Fig. 4.3

The initial state is

$$\big|i\big\rangle = c^\dagger_{s_i}(k_i)\big|0\big\rangle, \tag{4.59}$$

the final state is

$$\big\langle f\big| = \big\langle 0\big|c_{s_f}(k_f)a_{\lambda_f}(q_f). \tag{4.60}$$

The term of the scattering-matrix expansion under consideration is

$$S^{(1)}_{(a)} = ie\int_{-\infty}^{\infty} d^4x\, N\Big[\overline{\psi}^-(x)\not{A}^-(x)\psi^+(x)\Big]. \tag{4.61}$$

In this case, the operators are already normal-ordered. The transition amplitude from the initial to final states defined above is

$$\big\langle f\big|S^{(1)}_{(a)}\big|i\big\rangle = ie\int d^4x\,\big\langle 0\big|c_{s_f}(k_f)a_{\lambda_f}(q_f)\Big(\overline{\psi}^-(x)\not{A}^-(x)\psi^+(x)\Big)c^\dagger_{s_i}(k_i)\big|0\big\rangle. \tag{4.62}$$

We replace the usual expression for the field operators

$$\begin{aligned}\big\langle f\big|S^{(1)}_{(a)}\big|i\big\rangle = ie\int d^4x\,\big\langle 0\big|c_{s_f}(k_f)a_{\lambda_f}(q_f)\\ \times\sum_s\int\frac{d^3k}{(2\pi)^3 2\omega_k}c^\dagger_s(k)\overline{u}_s(k)e^{ikx}\\ \times\sum_\lambda\int\frac{d^3q}{(2\pi)^3 2\omega_q}a^\dagger_\lambda(q)\not{\epsilon}^*_\lambda(q)e^{iqx}\\ \times\sum_{s'}\int\frac{d^3k'}{(2\pi)^3 2\omega_{k'}}c_s(k')u_s(k')e^{-ik'x}c^\dagger_{s_i}(k_i)\big|0\big\rangle.\end{aligned} \tag{4.63}$$

Nonzero results can be obtained only if the spins and momenta of the fields correspond to the spins and momenta of the initial or final states. Formally, as we as seen in the previous section, we can replace products such as $c_{s_f}(k_f)c^\dagger_s(k)$ with their normal-ordered version minus the anticommutators (or plus their commutator for the bosonic a operators), which give us the usual delta functions. There is only a subtlety concerning unphysical photon states. We have

$$a_{\lambda_f}(q_f)a^\dagger_\lambda(q)\big|0\big\rangle = [a_{\lambda_f}(q_f),a^\dagger_\lambda(q)]\big|0\big\rangle + \cancel{a^\dagger_\lambda(q)a_{\lambda_f}(q_f)\big|0\big\rangle} = (2\pi)^3 2\omega_q(-g_{\lambda\lambda_f})\delta(\vec{q}-\vec{q}_f)\big|0\big\rangle. \tag{4.64}$$

Since the final state must be a physical state, λ_f cannot be scalar or longitudinal, but only transverse, so that $-g_{\lambda\lambda_f} = \delta_{\lambda\lambda_f}$. Alternatively, we can say that a physical state can contain scalar and longitudinal polarizations, but only in equal numbers (cf. Eq. (3.281)), so that the contributions of the two polarizations exactly cancel.

Therefore

$$\begin{aligned}\langle f|S^{(1)}_{(a)}|i\rangle &= ie\int d^4x\, \overline{u}_{s_f}(k_f)\, \not{\epsilon}^*_{\lambda_f}(q_f)\, u_{s_i}(k_i)\, e^{-i(k_i-q_f-k_f)x} \\ &= (2\pi)^4 ie\, \overline{u}_{s_f}(k_f)\, \not{\epsilon}^*_{\lambda_f}(q_f)\, u_{s_i}(k_i)\, \delta^4\big(k_i-q_f-k_f\big).\end{aligned} \tag{4.65}$$

The last delta function expresses the conservation of four-momentum.

We can rewrite the above result introducing the so-called Feynman amplitude $\mathcal{M}$

$$\boxed{\langle f|S^{(1)}_{(a)}|i\rangle = (2\pi)^4\, \delta^4\big(k_i-q_f-k_f\big)\, i\mathcal{M}^{(1)}_{(a)}.} \tag{4.66}$$

where

$$\boxed{i\mathcal{M}^{(1)}_{(a)} = ie\, \overline{u}_{s_f}(k_f)\, \not{\epsilon}^*_{\lambda_f}(q_f)\, u_{s_i}(k_i).} \tag{4.67}$$

The scattering amplitude $\mathcal{M}$ is often defined with different prefactors, e.g., 1 or $-i$ instead of i. This detail is irrelevant in the end, since observables are related to the modulus squared of the amplitude.

The Feynman amplitude is a Dirac scalar. If we write the Dirac indices explicitly, we obtain

$$\begin{aligned}i\mathcal{M}^{(1)}_{(a)} &= ie\, \overline{u}_{s_fA}(k_f)\, (\not{\epsilon}^*_{\lambda_f}(q_f))_{AB}\, u_{s_iB}(k_i) \\ &= ie\, u_{s_iB}(k_i)\overline{u}_{s_fA}(k_f)(\not{\epsilon}^*_{\lambda_f}(q_f))_{AB} = ie\mathrm{Tr}\Big[\Big(u_{s_i}(k_i)\overline{u}_{s_f}(k_f)\Big)\Big(\not{\epsilon}^*_{\lambda_f}(q_f)\Big)\Big]\end{aligned} \tag{4.68}$$

The possibility of using traces in Dirac space is particularly useful in some practical calculations, as we will see in the next sections.

The transition we considered (photon emission from an electron), from the point of view of the QED interaction is in principle allowed and is of order $\sqrt{\alpha}$. However, in the scattering process we assume that the initial and final states correspond to free field configurations (eigenstates of the free field operators). As such, they have to fulfill the on-shellness condition $k^2 = m^2$ in the case of the electron/positron and $q^2 = 0$ in the case of the photon. Fulfilling this condition and at the same time the condition of conservation of four momentum is not possible

$$\begin{cases} k_i^2 = k_f^2 = m^2 \qquad q_f^2 = 0 \\ k_i^2 = (k_f+q_f)^2 = k_f^2 + 2k_f\cdot q_f + q_f^2 \end{cases} \quad \Rightarrow \quad q_f = 0. \tag{4.69}$$

This type of transition is possible only if one of the particles is off-shell, i.e., $k^2 \neq m^2$. Off-shell particles can exist only as intermediate states, never as initial or final states, and are called "virtual particles."

Other possible processes included in $S^{(1)}$ are:

- $N\big[\overline{\psi}^{-}\not{A}^{+}\psi^{+}\big]$: photon absorption by an electron;
- $N\big[\overline{\psi}^{+}\not{A}^{+}\psi^{-}\big]$: photon absorption by a positron (note that normal ordering will induce an overall change of sign for this choice) ;
- $N\big[\overline{\psi}^{+}\not{A}^{-}\psi^{-}\big]$: photon emission by a positron (note that normal ordering will induce an overall change of sign for this choice);
- $N\big[\overline{\psi}^{+}\not{A}^{-}\psi^{+}\big]$: electron-positron annihilation into a photon;
- $N\big[\overline{\psi}^{-}\not{A}^{+}\psi^{-}\big]$: electron-positron creation from a photon;
- $N\big[\overline{\psi}^{+}\not{A}^{+}\psi^{+}\big]$: annihilation of an electron, a positron, and a photon;
- $N\big[\overline{\psi}^{-}\not{A}^{-}\psi^{-}\big]$: creation of an electron, a positron, and a photon;

Just to see another example, let's see what happens to the electron-positron creation case. The initial and final states are

$$\big|i\big\rangle = a^{\dagger}_{\lambda_i}(q_i)\big|0\big\rangle, \qquad\qquad \big\langle f\big| = \big\langle 0\big|c_{s_f}(k_f)d_{r_f}(p_f). \tag{4.70}$$

The transition amplitude for this process is

$$\big\langle f\big|S^{(1)}_{(b)}\big|i\big\rangle = ie\int d^4x\,\big\langle 0\big|c_{s_f}(k_f)d_{r_f}(p_f)\Big(\overline{\psi}^{-}(x)\gamma^{\alpha}\psi^{-}(x)A^{+}_{\alpha}(x)\Big)a^{\dagger}_{\lambda_i}(q_i)\big|0\big\rangle. \tag{4.71}$$

The normal-ordering prescription requires us to move the A^+ field, containing an absorption operator, to the right. This is not a real issue in the present case, since the A field commutes with the ψ field. However, to make things more transparent we wrote the operator in the correct way.

Without explicitly repeating the calculations, we obtain

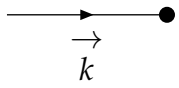
$$\begin{aligned}\big\langle f\big|S^{(1)}_{(b)}\big|i\big\rangle &= ie\int d^4x\,\overline{u}_{s_f}(k_f)\,\not{\epsilon}_{\lambda_i}(q_i)\,v_{r_f}(p_f)\,e^{-i(q_i-p_f-k_f)x}\\ &= (2\pi)^4 ie\,\overline{u}_{s_f}(k_f)\,\not{\epsilon}_{\lambda_i}(q_i)\,v_{r_f}(p_f)\,\delta^4\big(q_i-p_f-k_f\big).\end{aligned} \tag{4.72}$$

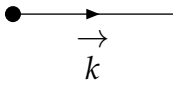
In general, the scattering matrix is connected to the Feynman amplitude via the relation

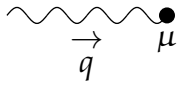
$$\boxed{\big\langle f\big|S\big|i\big\rangle = \mathbb{1} + (2\pi)^4\,\delta^4\big(\textstyle\sum k_i - \sum k_f\big)\,i\mathcal{M}(i\to f)}. \tag{4.73}$$

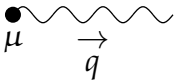
The explicit form of the Feynman amplitude at any order of the expansion in the coupling constant can be reconstructed using Feynman diagrams and Feynman rules. For instance, given the initial and final states considered above (e.g, electron going to electron and photon), there is only one possible diagram at order $\sqrt{\alpha}$ (the so-called "leading

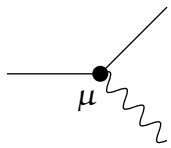
order"). We need simply to associate (dropping for simplicity the spin indices):

Incoming fermion:  $= u(k),$ (4.74)

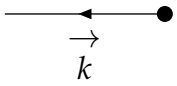
Outgoing fermion:  $= \overline{u}(k),$ (4.75)

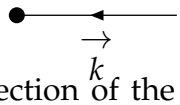
Incoming photon:  $= \epsilon_\mu(q)$ (4.76)

Outgoing photon:  $= \epsilon^*_\mu(q)$ (4.77)

Fermion-photon vertex:  $= ie\gamma^\mu$ (4.78)

The e in the vertex is consistent with the fact that the electron has charge $-e$, and it is what we conventionally identify as the "particle" and we associate with u spinors.[2] For a positron, its antiparticle, the vertex remains the same, but the incoming and outgoing positrons are denoted by

Incoming antifermion:  $= \overline{v}(k),$ (4.79)

Outgoing antifermion:  $= v(k).$ (4.80)

Note that for external antifermions, the direction of the momentum is normally taken opposite to the direction of the fermion line.

When computing a Feynman amplitude from Feynman rules, keeping track of the overall sign due to the inversion of fermion fields is relevant when several Feynman diagrams are contributing to a process. Only the relative sign between the Feynman diagrams matters, not their absolute sign. Whenever two diagrams are related by an interchange of two fermion lines, a negative sign must be added to one of the two.

4.4 Wick's theorem

The S matrix is expressed in terms of time-ordered products (T-products). Moreover, the interaction Hamiltonian is expressed in terms of normal products (N-products). To effi-

[2]Note that in different textbooks convetions may be different, as for the case of the covariant derivative. Ultimately, the absolute sign of the interaction does not change.

ciently compute the higher-order contributions to the S matrix, we should first consider how to simplify the time-ordered product of several operators.

The definition of normal product in Eq. (3.119) allows us to write

$$A(x_1)B(x_2) = N[A(x_1)B(x_2)] + \langle 0|A(x_1)B(x_2)|0\rangle. \tag{4.81}$$

Note also that

$$N\big[A(x_1)B(x_2)\big] = (-1)^P N\big[B(x_2)A(x_1)\big]. \tag{4.82}$$

Let us now assume A and B are two generic field operators and consider the time-ordered product (for $t_1 \neq t_2$)

$$\begin{aligned} T\big[A(x_1)B(x_2)\big] &= \theta(t_1-t_2)A(x_1)B(x_2) + (-1)^P\theta(t_2-t_1)B(x_2)A(x_1) \\ &= \theta(t_1-t_2)\Big\{N[A(x_1)B(x_2)] + \langle 0|A(x_1)B(x_2)|0\rangle\Big\} \\ &\quad + (-1)^P\theta(t_2-t_1)\Big\{N[B(x_2)A(x_1)] + \langle 0|B(x_2)A(x_1)|0\rangle\Big\} \\ &= N[A(x_1)B(x_2)] + \langle 0|T\big[A(x_1)B(x_2)\big]|0\rangle \end{aligned} \tag{4.83}$$

where in the last step we used (4.82). The last term of the above equation is the vacuum expectation value of a time-ordered product of two fields, also known as *contraction*. It can be nonzero only if they are two fields of the same type, in which case it is called the Feynman propagator.

The above relation can be generalized to any number of field operators. To do this, we first introduce the shorthand notations

$$\underbrace{A(x_1)B(x_2)} = \langle 0|T\big[A(x_1)B(x_2)\big]|0\rangle \tag{4.84}$$

and we further define

$$N\Big[\underbrace{A_1\underbrace{B_2C_3}D_4\ldots Y_m}\ldots Z_n\Big] = (-1)^P\ \underbrace{A_1C_3}\ \underbrace{B_2Y_m}\ N\Big[D_4\ldots Z_n\Big] \tag{4.85}$$

Eq. (4.83) can be generalized to any number of fields, in which case

$$T\big[A_1B_2\ldots Z_n\big] = N\Big[A_1B_2\ldots Z_n + \text{all possible contractions}\Big]. \tag{4.86}$$

This identity is known as Wick's theorem [34]. Eq. (4.83) is the particular case for two field operators. The proof by induction starting from the 2-field case is outlined, e.g., in Peskin–Schroeder, Sec. 4.3.

In the calculation of the S matrix, we need to consider a time-ordered product of interaction Hamiltonians, each one containing normal-ordered operators. In this case, Wick's theorem remains valid, but the contractions between operators at equal times (i.e., belonging to the same Hamiltonian operator) should not be considered. Consider for instance the simple case of an operator composed by two bosonic fields

$$T\Big[N\big[\phi(x)\phi(x)\big]\Big] = N\big[\phi(x)\phi(x)\big] + \underbrace{\phi(x)\phi(x)} \tag{4.87}$$

The second term must be omitted because the fields are already normal-ordered. This observation is generalized into the following version of Wick's theorem:

$$T\Big[N\big[A_1B_1\dots Z_1\big]\dots N\big[A_nB_n\dots Z_n\big]\Big] = N\Big[A_1B_1\dots Z_n + \text{all possible contractions, excluding equal-time contractions}\Big]. \tag{4.88}$$

4.5 Feynman propagators

As we have seen, the calculation of the scattering matrix requires the knowledge of contractions involving field operators. Only certain contractions can be nonzero, since they must involve the same number of creation and annihilation operators. As a matter of fact, we only need contractions between two fields of the same nature, which are the so-called Feynman propagators.

4.5.1 The scalar propagator

The Feynman propagator for a scalar field is defined as

$$\begin{aligned} i\Delta_F(x_1-x_2) &= \underbrace{\phi(x_1)\phi}(x_2) = \big\langle 0\big|T\big[\phi(x_1)\phi(x_2)\big]\big|0\big\rangle \\ &= \big\langle 0\big|\theta(t_1-t_2)\phi(x_1)\phi(x_2)+\theta(t_2-t_1)\phi(x_2)\phi(x_1)\big|0\big\rangle \end{aligned} \tag{4.89}$$

We split the field in positive and negative frequency parts and we remember that ϕ^+ contains an absorption operator and therefore gives zero when acting on the vacuum, while ϕ^- contains a creation operator and therefore gives zero when acting on the vacuum to its left. Therefore

$$i\Delta_F(x_1-x_2) = \big\langle 0\big|\theta(t_1-t_2)\phi^+(x_1)\phi^-(x_2)+\theta(t_2-t_1)\phi^+(x_2)\phi^-(x_1)\big|0\big\rangle \tag{4.90}$$

The Feynman propagator is composed by two contributions: when $t_1 > t_2$, a particle is created at x_2 and destroyed at x_1; when $t_2 > t_1$, a particle is created at x_1 and destroyed at x_2. We can represent the propagator graphically as in Fig. 4.4.

At this point, we observe that

$$\big\langle 0\big|\phi^+(x_1)\phi^-(x_2)\big|0\big\rangle = \big[\phi^+(x_1),\phi^-(x_2)\big]. \tag{4.91}$$

Therefore, we eventually obtain

$$i\Delta_F(x_1-x_2) = \theta(t_1-t_2)\big[\phi^+(x_1),\phi^-(x_2)\big]+\theta(t_2-t_1)\big[\phi^+(x_2),\phi^-(x_1)\big]. \tag{4.92}$$

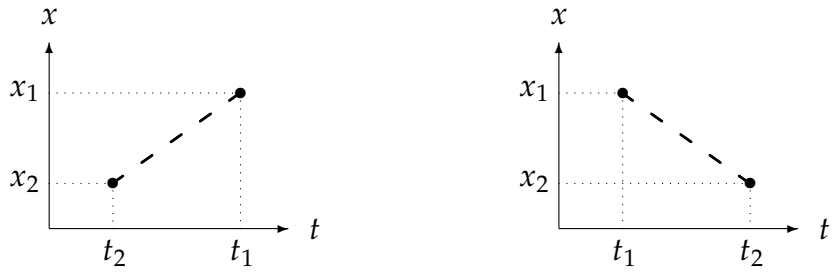


Figure 4.4: Graphical representation of the two time-ordered contributions to the Feynman propagator for a real scalar field.

The calculation of the commutator is not difficult. We use the usual expression of the field operators of Eq. (4.32) and then we carry out the calculation of the commutator

$$\begin{aligned}\left[\phi^+(x_1),\phi^-(x_2)\right] &= \int \frac{d^3k_1}{(2\pi)^3 2\omega_{k_1}} \int \frac{d^3k_2}{(2\pi)^3 2\omega_{k_2}} \left[a(\vec{k}_1), a^\dagger(\vec{k}_2)\right] e^{-ik_1x_1} e^{ik_2x_2} \\ &= \int \frac{d^3k_1}{(2\pi)^3 2\omega_{k_1}} \int \frac{d^3k_2}{(2\pi)^3 2\omega_{k_2}} (2\pi)^3 2\omega_{k_2}\, \delta(\vec{k}_1-\vec{k}_2)\, e^{-ik_1x_1} e^{ik_2x_2} \\ &= \int \frac{d^3k_1}{(2\pi)^3 2\omega_{k_1}} e^{-ik_1(x_1-x_2)}\end{aligned} \tag{4.93}$$

In conclusion, we get for the Feynman propagator (replacing the variables $x_1 - x_2 \to x$ and $k_1 \to k$)

$$\begin{aligned}i\Delta_F(x) &= \theta(t) \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{-ikx} + \theta(-t) \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{ikx} \\ &= \theta(t) \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{-i\omega_k t} e^{i\vec{k}\cdot\vec{x}} + \theta(-t) \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{i\omega_k t} e^{-i\vec{k}\cdot\vec{x}} \\ &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{i\vec{k}\cdot\vec{x}} \Big[\theta(t)\, e^{-i\omega_k t} + \theta(-t)\, e^{i\omega_k t}\Big].\end{aligned} \tag{4.94}$$

(In the last step we changed the integration variable from k to $-k$ in the second integral.) Here the two contributions for $t > 0$ and $t < 0$ are still separated. However, some mathematics allows us to write the Feynman propagator with a more compact formula, manifestly Lorentz invariant, where the distinction between the two contributions disappears, i.e.,

$$\boxed{i\Delta_F(x) = \int \frac{d^4k}{(2\pi)^4} \frac{i\, e^{-ikx}}{\left(k^2 - m^2 + i\epsilon\right)},} \tag{4.95}$$

where ϵ is positive and a limit for $\epsilon \to 0$ is understood. It is important to stress that in this formula k^0 *is not fixed to be equal to* ω_k, as we implicitly assumed in the previous equations.

Let us verify the correspondence between this expression and that of Eq. (4.94). We

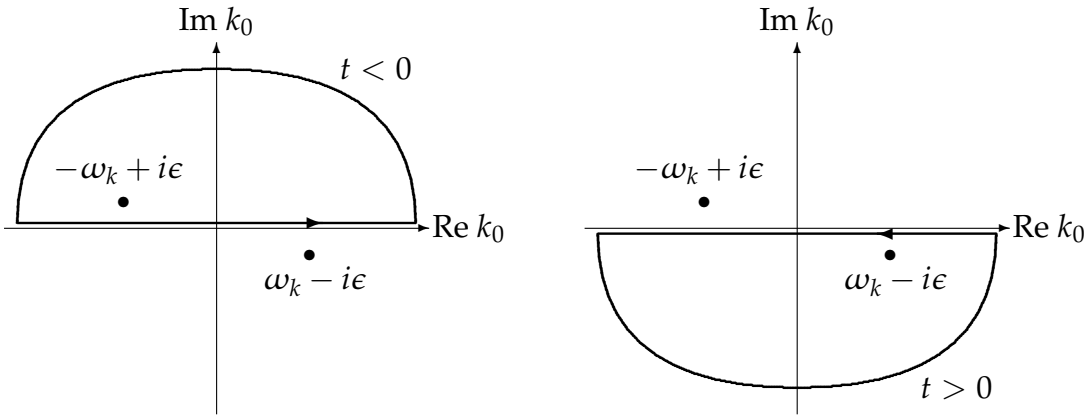


Figure 4.5: Pole positions and choice of contours relevant for the calculation of the scalar propagator.

need to observe that

$$\begin{aligned}\frac{1}{k^2 - m^2 + i\epsilon} &= \frac{1}{\left(k^0\right)^2 - |\vec{k}|^2 - m^2 + i\epsilon} \\ &= \frac{1}{\left(k^0 - \sqrt{|\vec{k}|^2 + m^2} + i\epsilon'\right)\left(k^0 + \sqrt{|\vec{k}|^2 + m^2} - i\epsilon'\right)} \\ &= \frac{1}{\left(k^0 - \omega_k + i\epsilon'\right)\left(k^0 + \omega_k - i\epsilon'\right)},\end{aligned} \tag{4.96}$$

where $\epsilon' = \epsilon/(2\omega_k)$ (always positive) and where we neglected terms of the order of ϵ^2.

We can write the Feynman propagator as

$$\begin{aligned}i\Delta_F(x) &= \int \frac{d^4k}{(2\pi)^4} \frac{i\, e^{-ikx}}{\left(k^0 - \omega_k + i\epsilon'\right)\left(k^0 + \omega_k - i\epsilon'\right)} \\ &= \int \frac{d^3k\, dk^0}{(2\pi)^4} \frac{i\, e^{-ik^0 t}\, e^{i\vec{k}\cdot\vec{x}}}{\left(k^0 - \omega_k + i\epsilon'\right)\left(k^0 + \omega_k - i\epsilon'\right)}\end{aligned} \tag{4.97}$$

We solve the integration over k^0 using Cauchy's theorem. We notice that the integrand has two poles: one for $k_0 = \omega_k - i\epsilon'$ and the second for $k_0 = -\omega_k + i\epsilon'$. We further note that for $t > 0$, the numerator vanishes for large negative imaginary values of k^0: to use Cauchy's theorem we need to close the contour in the lower half of the complex k^0 plane, picking up the first pole (and adding a minus sign because the contour runs clockwise). For $t < 0$, the numerator vanishes for large positive imaginary values of k^0: to use Cauchy's theorem we need to close the contour in the upper half of the complex k^0 plane, picking up the second pole. In conclusion, we obtain

$$i\Delta_F(x) = \int \frac{d^3k}{(2\pi)^4}\, e^{i\vec{k}\cdot\vec{x}} \left(\theta(t)(-2\pi i)\frac{i\, e^{-i\omega_k t}}{2\omega_k} + \theta(-t)(2\pi i)\frac{i\, e^{i\omega_k t}}{(-2\omega_k)}\right), \tag{4.98}$$

which corresponds to Eq. (4.94).

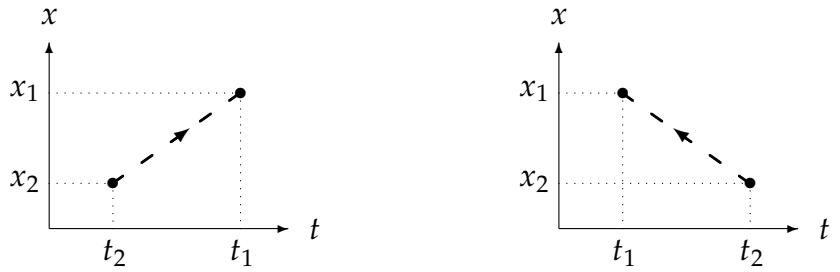


Figure 4.6: Graphical representation of the two time-ordered contributions to the Feynman propagator for a complex scalar field. The only difference with respect to the real field is the need to specify the orientation of the line in order to distinguish whether a particle or an antiparticle is produced.

In momentum space, the Feynman propagator for a scalar field (i.e., the Fourier transform of Eq. (4.95)) is then given by

$$\boxed{i\Delta_F(k) = \frac{i}{\left(k^2 - m^2 + i\epsilon\right)}.} \tag{4.99}$$

For a complex scalar field, the definition of the propagator is

$$\begin{aligned} i\Delta_F(x_1 - x_2) &= \left\langle 0\middle| T\left[\phi(x_1)\phi^\dagger(x_2)\right]\middle|0\right\rangle \\ &= \left\langle 0\middle|\theta(t_1 - t_2)\phi^+(x_1)\phi^{\dagger-}(x_2) + \theta(t_2 - t_1)\phi^{\dagger+}(x_2)\phi^-(x_1)\middle|0\right\rangle. \end{aligned} \tag{4.100}$$

The expression of the field operators is

$$\phi^+(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, a(\vec{k})e^{-ikx} \qquad \text{particle absorption,} \tag{4.101}$$

$$\phi^-(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, b^\dagger(\vec{k})e^{ikx} \qquad \text{antiparticle creation,} \tag{4.102}$$

$$\phi^{\dagger+}(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, b(\vec{k})e^{-ikx} \qquad \text{antiparticle absorption,} \tag{4.103}$$

$$\phi^{\dagger-}(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, a^\dagger(\vec{k})e^{ikx} \qquad \text{particle creation.} \tag{4.104}$$

There is no difference in the calculation with respect to the real case. However, the two contributions can be interpreted in a slightly different way: when $t_1 > t_2$, a particle is created at x_2 and destroyed at x_1; when $t_2 > t_1$, an *antiparticle* is created at x_1 and destroyed at x_2 (see Fig. 4.6). In the graphical representation, we need to orient the direction of our line, in order to distinguish the first case from the second.

The Feynman propagator is also a Green's function. It is in fact a solution of the inhomogeneous Klein–Gordon equation

$$\left(\partial_\mu \partial^\mu + m^2\right)\Delta_F(x) = -\delta^4(x). \tag{4.105}$$

We can check the above statement

$$\begin{aligned}\left(\partial_\mu\partial^\mu + m^2\right)\Delta_F(x) &= \left(\partial_\mu\partial^\mu + m^2\right)\int \frac{d^4k}{(2\pi)^4}\,\frac{e^{-ikx}}{\left(k^2 - m^2 + i\epsilon\right)} \\ &= \int \frac{d^4k}{(2\pi)^4}\,\frac{\left(-k^2 + m^2\right)e^{-ikx}}{\left(k^2 - m^2 + i\epsilon\right)} = -\delta^4(x).\end{aligned} \tag{4.106}$$

In the above, the position of the poles is irrelevant. Different Green's functions can be obtained using different prescriptions for the position of the poles. The one shown above, with $+i\epsilon$, is the one that defines the Feynman propagator and is correctly connected to the time-ordered product.

Green's functions are important for the solution of differential equations. The most common example is the solution of Poisson's equation for the electrostatic potential

$$\nabla^2\Phi(\vec{x}) = -\rho(\vec{x}). \tag{4.107}$$

It is easier to first solve the problem

$$\nabla^2 G(\vec{x} - \vec{x}') = -\delta^{(3)}(\vec{x} - \vec{x}'). \tag{4.108}$$

With the boundary condition that $G \to 0$ at large distances, the solution is

$$G(\vec{x} - \vec{x}') = \frac{1}{4\pi|\vec{x} - \vec{x}'|} \tag{4.109}$$

and corresponds to the potential generated by a pointlike source. Then, the solution for a generic charge distribution ρ is given by

$$\Phi(\vec{x}) = \int d^3x'\; G(\vec{x} - \vec{x}')\rho(\vec{x}'), \tag{4.110}$$

corresponding to the linear superposition of the fields generated by many pointlike sources.

4.5.2 The fermion propagator

The calculation of the Feynman propagator for the Dirac field starts again from the definition

$$\begin{aligned}i\left(S_F\right)_{AB}(x_1 - x_2) &= \left\langle 0\middle|T\left[\psi_A(x_1)\overline{\psi}_B(x_2)\right]\middle|0\right\rangle \\ &= \left\langle 0\middle|\theta(t_1 - t_2)\psi_A^+(x_1)\overline{\psi}_B^-(x_2) - \theta(t_2 - t_1)\overline{\psi}_B^+(x_2)\psi_A^-(x_1)\middle|0\right\rangle \\ &= \theta(t_1 - t_2)\left\{\psi_A^+(x_1), \overline{\psi}_B^-(x_2)\right\} - \theta(t_2 - t_1)\left\{\overline{\psi}_B^+(x_2), \psi_A^-(x_1)\right\}.\end{aligned} \tag{4.111}$$

The difference with respect to the scalar propagator is the presence of the anticommutators. The minus sign in the second term is due to the fermionic nature of the fields. The Dirac indices A and B are explicitly written. The two contributions can be interpreted in

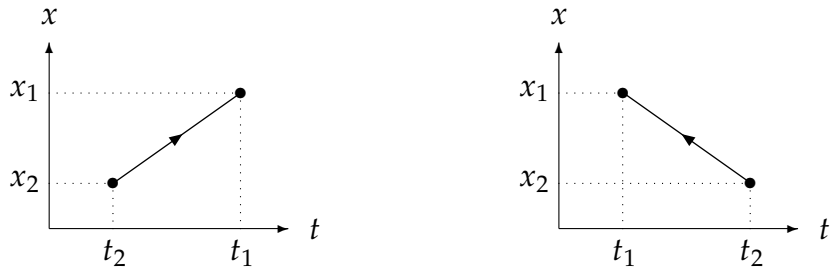


Figure 4.7: Graphical representation of the two time-ordered contributions to the Feynman propagator for a fermion field.

the same way as the complex scalar field: when $t_1 > t_2$, a particle is created at x_2 and destroyed at x_1; when $t_2 > t_1$, an antiparticle is created at x_1 and destroyed at x_2. The diagrammatic representation is given in Fig. 4.7

To calculate the anticommutators, we start from the usual expressions of the field operators of Eqs. (4.53) and following. Starting from the first anticommutator

$$\begin{aligned}
&\{\psi_A^+(x_1), \overline{\psi}_B^-(x_2)\} \\
&= \sum_{s_1,s_2} \int \frac{d^3k_1}{(2\pi)^3 2\omega_{k_1}} \int \frac{d^3k_2}{(2\pi)^3 2\omega_{k_2}} \{c_{s_1}(\vec{k}_1), c_{s_2}^\dagger(\vec{k}_2)\}\, u_{s_1A}(k_1)\overline{u}_{s_2B}(k_2)\, e^{-ik_1x_1} e^{ik_2x_2} \\
&= \sum_{s_1,s_2} \int \frac{d^3k_1}{(2\pi)^3 2\omega_{k_1}} \int \frac{d^3k_2}{(2\pi)^3 2\omega_{k_2}} (2\pi)^3 2\omega_{k_2} \delta(\vec{k}_1 - \vec{k}_2)\delta_{s_1s_2} u_{s_1A}(k_1)\overline{u}_{s_2B}(k_2)\, e^{-ik_1x_1} e^{ik_2x_2} \\
&= \int \frac{d^3k_1}{(2\pi)^3 2\omega_{k_1}} \sum_s u_{sA}(k_1)\overline{u}_{sB}(k_1)\, e^{-ik_1(x_1-x_2)} \\
&= \int \frac{d^3k_1}{(2\pi)^3 2\omega_{k_1}} (\not{k}_1 + m)_{AB}\, e^{-ik_1(x_1-x_2)}
\end{aligned} \tag{4.112}$$

The second anticommutator is similar

$$\begin{aligned}
&\{\overline{\psi}_B^+(x_2), \psi_A^-(x_1)\} \\
&= \sum_{s_1,s_2} \int \frac{d^3k_1}{(2\pi)^3 2\omega_{k_1}} \int \frac{d^3k_2}{(2\pi)^3 2\omega_{k_2}} \{d_{s_1}(\vec{k}_1), d_{s_2}^\dagger(\vec{k}_2)\}\, \overline{v}_{s_2B}(k_2) v_{s_1A}(k_1)\, e^{ik_1x_1} e^{-ik_2x_2} \\
&= \int \frac{d^3k_1}{(2\pi)^3 2\omega_{k_1}} \sum_s v_{sA}(k_1)\overline{v}_{sB}(k_1)\, e^{ik_1(x_1-x_2)} \\
&= \int \frac{d^3k_1}{(2\pi)^3 2\omega_{k_1}} (\not{k}_1 - m)_{AB}\, e^{ik_1(x_1-x_2)}
\end{aligned} \tag{4.113}$$

Inserting the two commutators in the definition of the propagator we obtain

$$\begin{aligned} iS_F(x) &= \theta(t)\, e^{-i\omega_k t} \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, (\not{k}+m)\, e^{i\vec{k}\cdot\vec{x}} - \theta(-t)\, e^{i\omega_k t} \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, (\not{k}-m)\, e^{-i\vec{k}\cdot\vec{x}} \\ &= \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, e^{i\vec{k}\cdot\vec{x}} \Big[\theta(t)\, (\omega_k\gamma^0 - \vec{k}\cdot\vec{\gamma} + m)\, e^{-i\omega_k t} \\ &\quad + \theta(-t)\, (-\omega_k\gamma^0 - \vec{k}\cdot\vec{\gamma} + m)\, e^{i\omega_k t}\Big]. \end{aligned} \tag{4.114}$$

Again, the propagator can be written as before in a more compact way

$$\boxed{iS_F(x) = \int \frac{d^4k}{(2\pi)^4}\, \frac{i\,(\not{k}+m)e^{-ikx}}{\left(k^2 - m^2 + i\epsilon\right)},} \tag{4.115}$$

and its Fourier transform is given by

$$\boxed{iS_F(k) = \frac{i\,(\not{k}+m)}{\left(k^2 - m^2 + i\epsilon\right)}.} \tag{4.116}$$

Again, we stress that in these formulas k^0 *is not fixed to be equal to* ω_k, as was implicitly assumed in the previous equations. With Cauchy theorem it is easy to repeat the proof that the above expressions lead to the same result as Eq. (4.114). Note that since the fermion propagator depends on k and not only on k^2, it is relevant to remember that k is the momentum flowing from x_2 to x_1, i.e., it has to follow the direction of the fermion line (in the complex scalar field case, we could have oriented the momentum opposite to the arrow with no consequence).

4.5.3 The photon propagator

The definition is the same as for the scalar propagator, apart from the presence of the Lorentz indices on the fields

$$iD_F^{\mu\nu}(x_1 - x_2) = \theta(t_1 - t_2)\left[A^{+\mu}(x_1), A^{-\nu}(x_2)\right] + \theta(t_2 - t_1)\left[A^{+\nu}(x_2), A^{-\mu}(x_1)\right]. \tag{4.117}$$

The general expression for the field operators is

$$A^{\mu+}(x) = \sum_{\lambda=0}^{3} \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, \epsilon_\lambda^\mu(\vec{k}) a_\lambda(\vec{k}) e^{-ikx}, \tag{4.118}$$

$$A^{\mu-}(x) = \sum_{\lambda=0}^{3} \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, \epsilon_\lambda^{\mu*}(\vec{k}) a_\lambda^\dagger(\vec{k}) e^{ikx}. \tag{4.119}$$

We can then compute the commutator with steps similar to the previous cases. The only difference is the need to take care of the vector part of the field represented by the polarization vectors ϵ and make use of Eq. (3.259):

$$\begin{aligned}
&\left[A^{+\mu}(x_1), A^{-\nu}(x_2)\right] \\
&= \sum_{\lambda_1,\lambda_2} \int \frac{d^3k_1}{(2\pi)^3 2\omega_{k_1}} \int \frac{d^3k_2}{(2\pi)^3 2\omega_{k_2}} \left[a_{\lambda_1}(\vec{k}_1), a^\dagger_{\lambda_2}(\vec{k}_2)\right] \epsilon^\mu_{\lambda_1}(k_1)\epsilon^{\nu *}_{\lambda_2}(k_2)\, e^{-ik_1x_1} e^{ik_2x_2} \\
&= \sum_{\lambda_1,\lambda_2} \int \frac{d^3k_1}{(2\pi)^3 2\omega_{k_1}} \int \frac{d^3k_2}{(2\pi)^3 2\omega_{k_2}} (-g_{\lambda_1\lambda_2})(2\pi)^3 2\omega_{k_2} \delta(\vec{k}_1 - \vec{k}_2)\, \epsilon^\mu_{\lambda_1}(k_1)\epsilon^{\nu *}_{\lambda_2}(k_2)\, e^{-ik_1x_1} e^{ik_2x_2} \\
&= -\int \frac{d^3k_1}{(2\pi)^3 2\omega_{k_1}} \sum_{\lambda_1,\lambda_2} g_{\lambda_1\lambda_2} \epsilon^\mu_{\lambda_1}(k_1)\epsilon^{\nu *}_{\lambda_2}(k_1)\, e^{-ik_1(x_1-x_2)} \\
&= -g^{\mu\nu} \int \frac{d^3k_1}{(2\pi)^3 2\omega_{k_1}}\, e^{-ik_1(x_1-x_2)}.
\end{aligned} \tag{4.120}$$

The only difference with respect to the scalar propagator is the presence of the extra $-g^{\mu\nu}$ (and the absence of a mass). Therefore, we can summarize the result of the propagator as

$$\boxed{iD_F^{\mu\nu}(x) = \int \frac{d^4k}{(2\pi)^4} \frac{-ig^{\mu\nu}\, e^{-ikx}}{\left(k^2 + i\epsilon\right)},} \tag{4.121}$$

and its Fourier transform is given by

$$\boxed{iD_F^{\mu\nu}(k) = \frac{-ig^{\mu\nu}}{\left(k^2 + i\epsilon\right)}.} \tag{4.122}$$

It is important to stress at this point that the above result is valid only in the Feynman gauge (i.e., Lorenz gauge with $\xi = 1$). What happens in other gauges? Let us first of all remind ourselves that the photon propagator is a Green's function. The photon propagator in Feynman gauge solves the equation (see Eq. (3.248))

$$\partial_\rho \partial^\rho D_F^{\mu\nu}(x) = -g^{\mu\nu}\delta^4(x). \tag{4.123}$$

More generally, in Lorenz gauges with generic ξ, the photon propagator turns out to be

$$iD_F^{\mu\nu}(k)|_{\mathrm{LG}} = \frac{i}{\left(k^2 + i\epsilon\right)} \left(-g^{\mu\nu} + (1-\xi)\frac{k^\mu k^\nu}{k^2} \right). \tag{4.124}$$

and solves the equation

$$\left[\partial_\rho \partial^\rho \delta^\mu_\sigma - (1 - 1/\xi)\partial^\mu \partial_\sigma\right] D_F^{\sigma\nu}(x) = -g^{\mu\nu}\delta^4(x). \tag{4.125}$$

To give another example, in axial gauges the propagator turns out to be (see next Section)

$$iD_F^{\mu\nu}(k)|_{\rm AG} = \frac{i}{\left(k^2+i\epsilon\right)}\left(-g^{\mu\nu}+\frac{k^\mu n^\nu + k^\nu n^\mu}{k\cdot n}-\frac{k^\mu k^\nu}{(k\cdot n)^2}\,n^2\right). \tag{4.126}$$

Without proving it, we stress that the differences between the various gauges should always cancel in physical results (the proof relies on Ward's identity, cf., e.g., Peskin–Schroeder, Sec. 7.4). We will check this statment in some examples in the next sections.

4.6 Photon propagator in axial and Coulomb gauge [optional]

The expression of the photon propagator in axial gauge can be obtained starting from a Lagrangian with the gauge-fixing term

$$\mathcal{L}_{\rm g.f.} = -\frac{1}{2\xi}(n_\mu A^\mu)^2 \tag{4.127}$$

The Green's function of the resulting equations of motion is

$$iD_F^{\mu\nu}(k)|_{\rm AG1} = \frac{i}{\left(k^2+i\epsilon\right)}\left(-g^{\mu\nu}+\frac{k^\mu n^\nu + k^\nu n^\mu}{k\cdot n}-\frac{k^\mu k^\nu}{(k\cdot n)^2}\,(n^2-\xi k^2)\right). \tag{4.128}$$

The standard choice is $\xi = 0$, corresponding to the result in Eq. (4.126).

Alternatively, we can check what happens to the expression of the polarization sum in the last steps of Eq. (4.120). Instead of using Eq. (3.259), we have to use Eq. (3.261), and we obtain again the result of Eq. (4.126).

The expression of the polarization sum was obtained by setting $k^2 = 0$. However, the covariant expression of the propagator implies that the photon is not necessarily on-shell, i.e., $k^2 \neq 0$. Therefore, in the axial-gauge case, to be complete we should use Eq. (3.260), which includes *four* terms.

$$\begin{aligned} iD_F^{\mu\nu}(k)|_{\rm AG2} = \frac{i}{\left(k^2+i\epsilon\right)}\Bigg(&-g^{\mu\nu}+\frac{k\cdot n}{(k\cdot n)^2-k^2}\left(k^\mu n^\nu + k^\nu n^\mu\right) \\ &-\frac{n^2}{(k\cdot n)^2-k^2}\,k^\mu k^\nu - \frac{k^2}{(k\cdot n)^2-k^2}\,n^\mu n^\nu\Bigg). \end{aligned} \tag{4.129}$$

It is possible to derive this result as the Green's function of the equations of motion for the free field, but in this case we need to add two gauge fixing terms $(n_\mu A^\mu)^2$ and $(\partial_\mu A^\mu)^2$.

It turns out that the last term in Eq.(4.129) can be omitted becasue it is cancelled by an analogous term appearing in the interaction Hamiltonian when adopting this gauge choice. This situation is similar in Coulomb gauge and this is the case more frequently discussed in textbooks. Let us see some more details.

The "standard" form of the Coulomb gauge propagator is

$$iD_F^{\mu\nu}(k)|_{\rm CG} = \frac{i}{(k^2+i\epsilon)}\left(-g^{\mu\nu}+\frac{k^0\left(k^\mu n^\nu + k^\nu n^\mu\right)}{|\vec{k}|^2}-\frac{k^\mu k^\nu}{|\vec{k}|^2}\right). \tag{4.130}$$

As for the axial gauge, we could obtain the propagator in Coulomb gauge as a Green's function. In this case we would need to include in the Lagrangian a gauge-fixing term of the type

$$\mathcal{L}_{\rm g.f.} = -\frac{1}{2\xi}(\vec{\nabla}\cdot\vec{A})^2 \tag{4.131}$$

from which we would obtain

$$iD_F^{\mu\nu}(k)|_{\rm CG1} = \frac{i}{(k^2+i\epsilon)}\left(-g^{\mu\nu}+\frac{k^0\left(k^\mu n^\nu + k^\nu n^\mu\right)}{|\vec{k}|^2}-\frac{k^\mu k^\nu}{|\vec{k}|^2}-\frac{\xi k^2 k^\mu k^\nu}{|\vec{k}|^4}\right). \tag{4.132}$$

The usual choice is $\xi = 0$, corresponding to Eq. (4.130).

Alternatively, we can use the properties of polarization vectors. In Coulomb gauge, we have effectively two transverse degrees of freedom only, orthogonal to k and to the pure time direction, which corresponds to using Eq. (3.260) with $n = (1,0,0,0)$. In practice, for the polarization sum in Coulomb gauge we obtain

$$iD_F^{\mu\nu}(k)|_{\rm CG2} = \frac{i}{(k^2+i\epsilon)}\left(-g^{\mu\nu}+\frac{k^0\left(k^\mu n^\nu + k^\nu n^\mu\right)}{|\vec{k}|^2}-\frac{k^\mu k^\nu}{|\vec{k}|^2}-\frac{k^2 n^\mu n^\nu}{|\vec{k}|^2}\right). \tag{4.133}$$

which corresponds to Eq. (4.130) only if we impose $k^2 = 0$, eliminating the last term.

Let us devote some attention to the very last term and add it to Eq. (4.121), we obtain an extra term

$$\int \frac{d^4k}{(2\pi)^4}\,\frac{i\cancel{k^2}\,n^\mu n^\nu\, e^{-ikx}}{\cancel{(k^2+i\epsilon)}}. \tag{4.134}$$

The integration over k^0 in this case gives a $\delta(t)$, which means that this is an "instantaneous term", also called a Coulomb term.

It looks like we have a term that potentially spoils the agreement with the Feynman-gauge calculation. However, this problem is resolved when we look at the form of the interaction Hamiltonian in Coulomb gauge, which contains the expression $\overline{\psi}(x)\not{A}(x)\psi(x)$. In Coulomb gauge, the A^0 component of the field is treated as an auxiliary field, has to fulfill Poisson's equation and is determined by the charge distribution at time t, i.e. (using Eq. (4.110) and the charge density associated to the Dirac field $\rho(x') = -e\overline{\psi}(x')\gamma^0\psi(x')$)

$$A^0(x) = \Phi(x) = -e\int d^3x'\frac{\overline{\psi}(x')\gamma^0\psi(x')}{4\pi|\vec{x}-\vec{x}'|} = -e\int d^4x'\frac{\overline{\psi}(x')\gamma^0\psi(x')}{4\pi|\vec{x}-\vec{x}'|}\delta(t-t'). \tag{4.135}$$

The contribution to the scattering matrix will then become

$$-e\int d^4x\left[\overline{\psi}(x)\gamma^0\psi(x)A^0(x)\right] = e^2\int d^4x\int d^4x'\,\overline{\psi}(x)\gamma^0\psi(x)\frac{\overline{\psi}(x')\gamma^0\psi(x')}{4\pi|\vec{x}-\vec{x}'|}\delta(t-t') \tag{4.136}$$

This is the form of the Coulomb interaction (also sometimes called "contact interaction") which is an instantaneous interaction, in the sense that a change in the charge density distribution at $\vec{x}'$ would be immediately felt by a charge at $\vec{x}$.

Coulomb's interaction is part of the first-order term in Dyson's expansion of the scattering matrix, *but* it is of order e^2 and thus it should be summed to the second-order contributions to S. It turns out that from the second-order contribution to S, the instantaneous term in the photon propagator *exactly cancels* the contribution from the instantaneous Coulomb interaction. In the words of Weinberg ([33], p. 355): "We se that the apparent violation of Lorentz invariance in the instantaneous Coulomb interaction is cancelled by another apparent violation of Lorentz invariance, that the fields A^μ are not four-vectors, and therefore have a non-covariant propagator."

From a practical point of view, it means that in non-covariant gauges we can avoid including the last term of the propagator, proportional to k^2, even if the photon is off-shell.

Where does the Coulomb interaction ends up, then? Does it disappear? No, because it is a physical effect that must be already included in the $-g^{\mu\nu}$ part of the propagator, which is the same in all gauges. In fact, if you look at the nonrelativistic limit of the Feynman gauge expression (following, e.g., p. 125 of Peskin–Schroeder), you recover Coulomb's potential.

4.7 Second-order contributions: some examples

4.7.1 "ABC" theory

Let us start from the "ABC" theory. The second-order expansion of the scattering matrix is

$$S^{(2)} = -\frac{\lambda^2}{2!}\int d^4x_1 d^4x_2\; T\Big[N\Big[\phi_A(x_1)\phi_B(x_1)\phi_C(x_1)\Big]\; N\Big[\phi_A(x_2)\phi_B(x_2)\phi_C(x_2)\Big]\Big]. \quad (4.137)$$

Applying Wick's theorem we can obtain in principle one term without contractions, 9 terms with one contraction, 9 terms with two contractions, one term with three contractions. However, only contractions between the *same fields* are nonzero, so that we obtain only 8 terms: one term without contractions, 3 terms with one contraction, 3 terms with two contractions, one term with three contractions, i.e., [3]

$$S^{(2)}_{(a)} = -\frac{\lambda^2}{2!}\int d^4x_1 d^4x_2\; N\Big[\big(\phi_A\phi_B\phi_C\big)_{x_1}\big(\phi_A\phi_B\phi_C\big)_{x_2}\Big], \quad (4.138)$$

$$S^{(2)}_{(b)} = -\frac{\lambda^2}{2!}\int d^4x_1 d^4x_2\; N\Big[\big(\underbrace{\phi_A}\phi_B\phi_C\big)_{x_1}\big(\underbrace{\phi_A}\phi_B\phi_C\big)_{x_2}\Big], \quad (4.139)$$

[3] As a shorthand, we use x_1 and x_2 as subscripts to refer to terms that are all evaluated at x_1 or x_2.

$$S^{(2)}_{(c)} = -\frac{\lambda^2}{2!}\int d^4x_1 d^4x_2\; N\Big[\big(\phi_A\phi_B\phi_C\big)_{x_1}\big(\phi_A\phi_B\phi_C\big)_{x_2}\Big], \tag{4.140}$$

$$S^{(2)}_{(d)} = -\frac{\lambda^2}{2!}\int d^4x_1 d^4x_2\; N\Big[\big(\phi_A\phi_B\phi_C\big)_{x_1}\big(\phi_A\phi_B\phi_C\big)_{x_2}\Big], \tag{4.141}$$

$$S^{(2)}_{(e)} = -\frac{\lambda^2}{2!}\int d^4x_1 d^4x_2\; N\Big[\big(\phi_A\phi_B\phi_C\big)_{x_1}\big(\phi_A\phi_B\phi_C\big)_{x_2}\Big], \tag{4.142}$$

$$S^{(2)}_{(f)} = -\frac{\lambda^2}{2!}\int d^4x_1 d^4x_2\; N\Big[\big(\phi_A\phi_B\phi_C\big)_{x_1}\big(\phi_A\phi_B\phi_C\big)_{x_2}\Big], \tag{4.143}$$

$$S^{(2)}_{(g)} = -\frac{\lambda^2}{2!}\int d^4x_1 d^4x_2\; N\Big[\big(\phi_A\phi_B\phi_C\big)_{x_1}\big(\phi_A\phi_B\phi_C\big)_{x_2}\Big], \tag{4.144}$$

$$S^{(2)}_{(h)} = -\frac{\lambda^2}{2!}\int d^4x_1 d^4x_2\; N\Big[\big(\phi_A\phi_B\phi_C\big)_{x_1}\big(\phi_A\phi_B\phi_C\big)_{x_2}\Big]. \tag{4.145}$$

We will discuss all the possibilities only in the section dedicated to QED. For the moment, let us focus on a specific choice of initial and final states. For instance, let us consider two A and B fields in the initial and final states, i.e.,

$$\big|i\big\rangle = a^\dagger_A(k_i)a^\dagger_B(p_i)\big|0\big\rangle, \qquad\qquad \big\langle f\big| = \big\langle 0\big|a_B(p_f)a_A(k_f). \tag{4.146}$$

We need to check if there is a possible term in the above expressions that can connect these states. It needs to contain two uncontracted A and two uncontracted B fields, therefore we need contribution $S^{(2)}_{(d)}$. In particular, we need one absorption and one creation operator from the A fields and similarly for the B fields. There are four possible contributions, corresponding to

$$\big(\phi^+_A\phi^-_B\phi_C\big)_{x_1}\big(\phi^-_A\phi^+_B\phi_C\big)_{x_2} \qquad\qquad \big(\phi^-_A\phi^+_B\phi_C\big)_{x_1}\big(\phi^+_A\phi^-_B\phi_C\big)_{x_2} \tag{4.147}$$

$$\big(\phi^+_A\phi^+_B\phi_C\big)_{x_1}\big(\phi^-_A\phi^-_B\phi_C\big)_{x_2} \qquad\qquad \big(\phi^-_A\phi^-_B\phi_C\big)_{x_1}\big(\phi^+_A\phi^+_B\phi_C\big)_{x_2} \tag{4.148}$$

The two terms in the first line are identical, since if we exchange the integration variables x_1 and x_2 we obtain the same expression. It is sufficient to consider only one of them, multiplying it by a factor 2 that cancels the 2! in the denominator. These terms correspond to diagram (ab1) in Fig. 4.8, since we have one creation and one absorption operator at each vertex. Similar considerations hold for the two terms in the second line: they are identical, leading to a cancellation of the 2! factor, and they correspond to diagram (ab2) in Fig. 4.8, since they contain two creation and two absorption operators at each vertex.

This is a case where there are *two* amplitudes contributing to the same process, with the same initial and final states. It is important to determine correctly the relative sign between these two amplitudes: in this case, only scalar fields are involved and no sign change occurs if the sequence of operators is modified due to normal ordering.

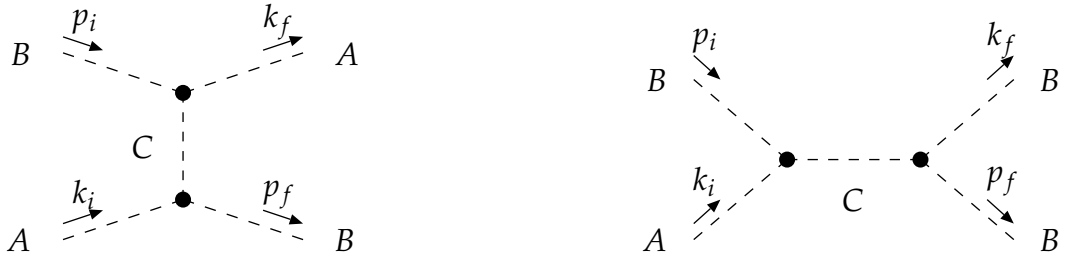


Figure 4.8: Feynman diagrams contributing to $AB \to AB$ scattering.

4.7.2 Electron-muon scattering

Before considering all processes occurring in QED with electrons and positrons, we first consider the interaction between an electron and a muon. Muons can be considered as identical to electrons, apart from having a higher mass ($m_\mathrm{m} = 106$ MeV, $m_\mathrm{e} = 0.511$ MeV). The Lagrangian is obtained just by duplicating the QED Lagrangian, except for the free photon field part, i.e.,

$$\mathcal{L}_{\mathrm{e-m}} = \overline{\psi}_\mathrm{e}\big(i\partial\!\!\!/ - m_\mathrm{e}\big)\psi_\mathrm{e} + \overline{\psi}_\mathrm{m}\big(i\partial\!\!\!/ - m_\mathrm{m}\big)\psi_\mathrm{m} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + e\overline{\psi}_\mathrm{e}\gamma^\mu\psi_\mathrm{e}A_\mu + e\overline{\psi}_\mathrm{m}\gamma^\mu\psi_\mathrm{m}A_\mu. \tag{4.149}$$

The second-order contributions to the scattering matrix are

$$\begin{aligned} S^{(2)} = -\frac{e^2}{2!}\int d^4x_1 d^4x_2\, T\Big[& N\Big[\overline{\psi}_\mathrm{e}(x_1)A\!\!\!/(x_1)\psi_\mathrm{e}(x_1) + \overline{\psi}_\mathrm{m}(x_1)A\!\!\!/(x_1)\psi_\mathrm{m}(x_1)\Big] \\ & N\Big[\overline{\psi}_\mathrm{e}(x_2)A\!\!\!/(x_1)\psi_\mathrm{e}(x_2) + \overline{\psi}_\mathrm{m}(x_2)A\!\!\!/(x_1)\psi_\mathrm{m}(x_2)\Big]\Big]. \end{aligned} \tag{4.150}$$

The Lagrangian is obviously more complicated than including only electrons. However, it allows us to study some simpler cases if specific initial and final states are chosen, as it will be clear in the following.

Let us consider the case of a scattering of an electron and a muon. The initial and final states are

$$\big|i\big\rangle = c^\dagger_{\mathrm{e},s_i}(k_i)c^\dagger_{\mathrm{m},r_i}(p_i)\big|0\big\rangle, \qquad \big\langle f\big| = \big\langle 0\big|c_{\mathrm{m},r_f}(p_f)c_{\mathrm{e},s_f}(k_f). \tag{4.151}$$

We need to have one absorption and one creation operator for the electron and the same for the muon. Therefore, we need to select the following term in the Wick expansion of the scattering matrix

$$\begin{aligned} S^{(2)}_{(e\mu)} = -\frac{e^2}{2!}\int d^4x_1 d^4x_2\, N\Big[& \big(\overline{\psi}_e\gamma^\alpha \underbrace{A_\alpha\psi_e}\big)_{x_1}\big(\overline{\psi}_\mathrm{m}\gamma^\beta \underbrace{A_\beta\psi_\mathrm{m}}\big)_{x_2} \\ & + \big(\overline{\psi}_\mathrm{m}\gamma^\alpha \underbrace{A_\alpha\psi_\mathrm{m}}\big)_{x_1}\big(\overline{\psi}_\mathrm{e}\gamma^\beta \underbrace{A_\beta\psi_\mathrm{e}}\big)_{x_2}\Big] \end{aligned} \tag{4.152}$$

The second term is identical to the first and can be combined with it to get rid of the factor 2 in the denominator.

Considering only the terms that can give a nonzero contribution and introducing the photon propagator we obtain

$$S^{(2)}_{(e\mu)} = -e^2 \int d^4x_1 d^4x_2 \left(\overline{\psi}^-_{\mathrm{e}} \gamma^\alpha \psi^+_{\mathrm{e}}\right)_{x_1} iD_{F,\alpha\beta}(x_1 - x_2) \left(\overline{\psi}^-_{\mathrm{m}} \gamma^\beta \psi^+_{\mathrm{m}}\right)_{x_2}. \tag{4.153}$$

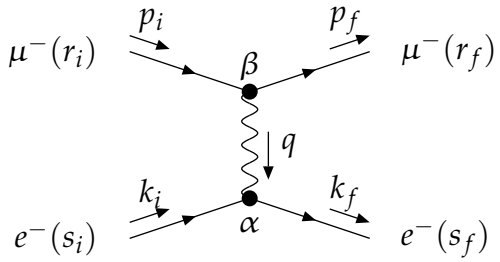


Figure 4.9: Feynman diagram contributing to electron-muon scattering.

Next, we have to insert the expressions of the field operators.

$$\begin{aligned}
\overline{\psi}^-_{\mathrm{e}}(x_1) &= \sum_{s'} \int \frac{d^3k'}{(2\pi)^3 2\omega_{k'}} c^\dagger_{\mathrm{e},s'}(\vec{k}') \overline{u}_{\mathrm{e},s'}(\vec{k}') e^{ik'x_1} && \text{electron creation,} \\
\psi^+_{\mathrm{e}}(x) &= \sum_{s''} \int \frac{d^3k''}{(2\pi)^3 2\omega_{k''}} c_{\mathrm{e},s''}(\vec{k}'') u_{\mathrm{e},s''}(\vec{k}'') e^{-ik''x_1} && \text{electron absorption,} \\
\overline{\psi}^-_{\mathrm{m}}(x_1) &= \sum_{r'} \int \frac{d^3p'}{(2\pi)^3 2\omega_{p'}} c^\dagger_{\mathrm{m},r'}(\vec{p}') \overline{u}_{\mathrm{m},r'}(\vec{p}') e^{ip'x_2} && \text{muon creation,} \\
\psi^+_{\mathrm{m}}(x) &= \sum_{r''} \int \frac{d^3p''}{(2\pi)^3 2\omega_{p''}} c_{\mathrm{m},r''}(\vec{p}') u_{\mathrm{m},r''}(\vec{p}'') e^{-ip''x_2} && \text{muon absorption.}
\end{aligned}$$

As a first step, let us focus on the involved creation and absorption operators. Our scattering matrix contains this expression

$$\begin{aligned}
&\left\langle 0 \middle| c_{\mathrm{m},r_f}(p_f) c_{\mathrm{e},s_f}(k_f) c^\dagger_{\mathrm{e},s'}(k') c_{\mathrm{e},s''}(k'') c^\dagger_{\mathrm{m},r'}(p') c_{\mathrm{m},r''}(p'') c^\dagger_{\mathrm{e},s_i}(k_i) c^\dagger_{\mathrm{m},r_i}(p_i) \middle| 0 \right\rangle \\
&= \left\langle 0 \middle| c_{\mathrm{e},s_f}(k_f) c^\dagger_{\mathrm{e},s'}(k') c_{\mathrm{e},s''}(k'') c^\dagger_{\mathrm{e},s_i}(k_i) c_{\mathrm{m},r_f}(p_f) c^\dagger_{\mathrm{m},r'}(p') c_{\mathrm{m},r''}(p'') c^\dagger_{\mathrm{m},r_i}(p_i) \middle| 0 \right\rangle \\
&= \left\langle 0 \middle| \left[c_{\mathrm{e},s_f}(k_f), c^\dagger_{\mathrm{e},s'}(k') \right] \left[c_{\mathrm{e},s''}(k''), c^\dagger_{\mathrm{e},s_i}(k_i) \right] \left[c_{\mathrm{m},r_f}(p_f), c^\dagger_{\mathrm{m},r'}(p') \right] \left[c_{\mathrm{m},r''}(p''), c^\dagger_{\mathrm{m},r_i}(p_i) \right] \middle| 0 \right\rangle \\
&= (2\pi)^3 2\omega_{k_f} \delta_{s_f s'} \delta(\vec{k}_f - \vec{k}') \, (2\pi)^3 2\omega_{k_i} \delta_{s_i s''} \delta(\vec{k}_i - \vec{k}'') \\
&\quad \times (2\pi)^3 2\omega_{p_f} \delta_{r_f r'} \delta(\vec{p}_f - \vec{p}') \, (2\pi)^3 2\omega_{p_i} \delta_{r_f r''} \delta(\vec{p}_i - \vec{p}'') \langle 0|0\rangle
\end{aligned} \tag{4.154}$$

where we just used anticommutation rules for the ladder operators. Once we have simplified this part of the scattering matrix, we can perform the sum of over the spins and

integrals over the momenta using the deltas and obtain

$$\begin{aligned}\langle f|S^{(2)}_{(e\mu)}|i\rangle = -e^2 \int d^4x_1 d^4x_2\, \overline{u}_{\mathrm{e},s_f}(\vec{k}_f) e^{ik_f x_1} \gamma^\alpha u_{\mathrm{e},s_i}(\vec{k}_i) e^{-ik_i x_1} iD_{F,\alpha\beta}(x_1-x_2) \\ \overline{u}_{\mathrm{m},r_f}(\vec{p}_f) e^{ip_f x_2} \gamma^\beta u_{\mathrm{m},r_i}(\vec{p}_i) e^{-ip_i x_2}\end{aligned} \tag{4.155}$$

We now replace the expression of the photon propagator and perform the integrations over x_1 and x_2

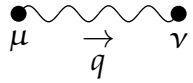
$$\begin{aligned}\langle f|S^{(2)}_{(e\mu)}|i\rangle &= -e^2 \int d^4x_1 d^4x_2\, \overline{u}_{\mathrm{e},s_f}(\vec{k}_f) e^{ik_f x_1} \gamma^\alpha u_{\mathrm{e},s_i}(\vec{k}_i) e^{-ik_i x_1} \\ &\quad \times \int \frac{d^4q}{(2\pi)^4} \frac{-ig_{\alpha\beta}\, e^{-iq(x_1-x_2)}}{(q^2+i\epsilon)} \overline{u}_{\mathrm{m},r_f}(\vec{p}_f) e^{ip_f x_2} \gamma^\beta u_{\mathrm{m},r_i}(\vec{p}_i) e^{-ip_i x_2} \\ &= -e^2 \int \frac{d^4q}{(2\pi)^4} \int d^4x_1 d^4x_2\, e^{-i(k_i-k_f+q)x_1}\, e^{-i(p_i-p_f-q)x_2} \\ &\quad \times \overline{u}_{\mathrm{e},s_f}(\vec{k}_f) \gamma^\alpha u_{\mathrm{e},s_i}(\vec{k}_i) \frac{-ig_{\alpha\beta}}{(q^2+i\epsilon)} \overline{u}_{\mathrm{m},r_f}(\vec{p}_f) \gamma^\beta u_{\mathrm{m},r_i}(\vec{p}_i) \\ &= -e^2 \int \frac{d^4q}{(2\pi)^4} (2\pi)^4 \delta^{(4)}(k_i-k_f+q)(2\pi)^4 \delta^{(4)}(p_i-p_f-q) \\ &\quad \times \overline{u}_{\mathrm{e},s_f}(\vec{k}_f) \gamma^\alpha u_{\mathrm{e},s_i}(\vec{k}_i) \frac{-ig_{\alpha\beta}}{(q^2+i\epsilon)} \overline{u}_{\mathrm{m},r_f}(\vec{p}_f) \gamma^\beta u_{\mathrm{m},r_i}(\vec{p}_i) \\ &= (2\pi)^4 \delta^{(4)}(k_i+p_i-k_f-p_f) \\ &\quad \times (-e^2)\, \overline{u}_{\mathrm{e},s_f}(\vec{k}_f) \gamma^\alpha u_{\mathrm{e},s_i}(\vec{k}_i) \left.\frac{-ig_{\alpha\beta}}{(q^2+i\epsilon)}\right|_{q=p_i-p_f} \overline{u}_{\mathrm{m},r_f}(\vec{p}_f) \gamma^\beta u_{\mathrm{m},r_i}(\vec{p}_i).\end{aligned} \tag{4.156}$$

The calculation leads to a simple formula for the scattering matrix, where we can again find conservation of total four-momentum, but also conservation of four momentum at each vertex of the Feynman diagram.

The Feynman amplitude is

$$i\mathcal{M}^{(2)}_{(e\mu)} = \overline{u}_{\mathrm{e},s_f}(\vec{k}_f)(-ie\gamma^\alpha) u_{\mathrm{e},s_i}(\vec{k}_i) \left.\frac{-ig_{\alpha\beta}}{(q^2+i\epsilon)}\right|_{q=p_i-p_f} \overline{u}_{\mathrm{m},r_f}(\vec{p}_f)(-ie\gamma^\beta) u_{\mathrm{m},r_i}(\vec{p}_i). \tag{4.157}$$

On top of the Feynman rules for the incoming and outgoing fermions, which we found already at order e, we can now add Feynman rule for photon propagators

Photon propagator (Feynman gauge): 

$$= \frac{-ig_{\mu\nu}}{q^2+i\epsilon} \tag{4.158}$$

At this point, we should worry about gauge choices and make an important remark: physical results cannot depend on the choice of gauge. The photon propagator depends

on gauge. However, we can explicitly check that the additional pieces appearing in Eq. (4.124) or Eq. (4.126) give no contribution. For instance, let us replace $g_{\alpha\beta}$ with $q_\alpha n_\beta$. In the first part of the Feynman amplitude we would obtain

$$\overline{u}_{\mathrm{e},s_f}(\vec{k}_f)\gamma^\alpha u_{\mathrm{e},s_i}(\vec{k}_i)q_\alpha\big|_{q=k_f-k_i} = \overline{u}_{\mathrm{e},s_f}(\vec{k}_f)(\not{k}_f - \not{k}_i)u_{\mathrm{e},s_i}(k_i) = 0. \tag{4.159}$$

The last equality holds due to Dirac's equation for the spinors

$$(\not{k}_i - m_e)u(k_i) = 0, \qquad \overline{u}(k_f)(\not{k}_f - m_e) = 0. \tag{4.160}$$

In conclusion, even if the photon propagator changes in different gauges, the scattering amplitude does not change.

4.7.3 Yukawa interaction [optional]

We now use the Yukawa Lagrangian to describe the interaction between protons. At second order we have

$$S^{(2)} = -\frac{g^2}{2!}\int d^4x_1 d^4x_2\; T\Big[N\Big[\overline{\psi}(x_1)\psi(x_1)\phi(x_1)\Big]\; N\Big[\overline{\psi}(x_2)\psi(x_2)\phi(x_2)\Big]\Big]. \tag{4.161}$$

Suppose we focus on the scattering between a proton and an antiproton, by selecting the following initial and final states

$$\big|i\big\rangle = d^\dagger_{s_i}(k_i)c^\dagger_{r_i}(p_i)\big|0\big\rangle, \qquad \big\langle f\big| = \big\langle 0\big|c_{r_f}(p_f)d_{s_f}(k_f). \tag{4.162}$$

Without repeating all considerations, we can deduce that the relevant parts of the scattering matrix will be

$$S^{(2)} = -g^2\int d^4x_1 d^4x_2\; N\Big[\underbrace{(\overline{\psi}^+\psi^-\phi)_{x_1}(\overline{\psi}^-\psi^+\phi)_{x_2}} + \underbrace{(\overline{\psi}^+\psi^+\phi)_{x_1}(\overline{\psi}^-\psi^-\phi)_{x_2}}\Big], \tag{4.163}$$

where we have already put together identical terms and simplified the 2! in the denominator.

In this case, since we are dealing with fermion fields and there are two contributions to the scattering process, we need to determine what is their relative sign. Focusing only on the ordering of the ladder operators, we have $N[dd^\dagger c^\dagger c] = d^\dagger c^\dagger dc$ for the first term and $N[dcc^\dagger d^\dagger] = -d^\dagger c^\dagger dc$, i.e., the two scattering amplitudes enter with the *opposite* sign.

The scattering amplitudes become

$$i\mathcal{M}_{(\mathrm{Y1})} = -g^2\overline{u}_{s_f}(\vec{k}_f)u_{s_i}(\vec{k}_i)\;\frac{i}{(q^2-M^2+i\epsilon)}\bigg|_{q=p_i-p_f}\overline{v}_{r_i}(\vec{p}_i)v_{r_f}(\vec{p}_f), \tag{4.164}$$

$$i\mathcal{M}_{(\mathrm{Y2})} = +g^2\overline{u}_{s_f}(\vec{k}_f)v_{r_f}(\vec{p}_f)\;\frac{i}{(q^2-M^2+i\epsilon)}\bigg|_{q=p_i+k_i}\overline{v}_{r_i}(\vec{p}_i)u_{s_i}(\vec{k}_i), \tag{4.165}$$

where M is the mass of the scalar field. Note that the relative sign between the two amplitudes can be determined also looking at their expressions: their are connected by a simple exchange of two spinors (u and v), which keeps memory of the way ladder operators were ordered in the interaction Hamiltonian.

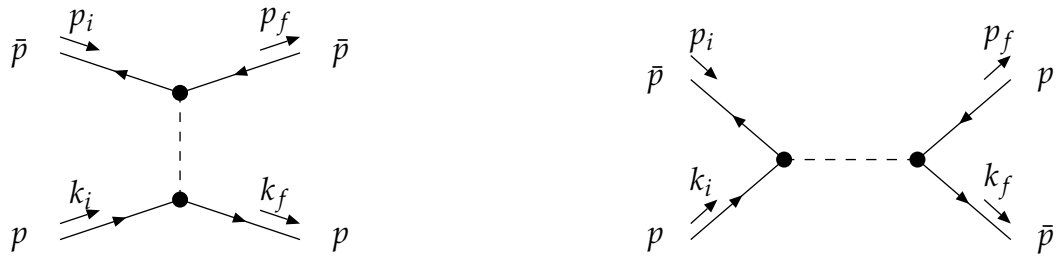


Figure 4.10: Feynman diagrams contributing to $p\bar{p}$ scattering in Yukawa theory.

4.7.4 Nonrelativistic approximation [optional]

It is instructive to study what happens if we apply a nonrelativistic approximation to the scattering amplitudes of the electron-muon and of the proton-antiproton scattering obtained in the previous sections in the nonrelativistic approximation. This is discusse, e.g., in Peskin-Schroeder, Ch. 4 and in Mandl-Shaw, Sec. 8.7.

Let us suppose that one of the particle scatters off a static potential produced by the other particle, which we consider the target. In the nonrelativistic limit, we can do these approximations at the target level

$$p_i = (m, \vec{p}_i), \qquad p_f = (m, \vec{p}_f), \tag{4.166}$$

and

$$(p_i - p_f)^2 \approx -|\vec{p}_i - \vec{p}_f|^2 = -|\vec{k}_f - \vec{k}_i|^2. \tag{4.167}$$

For what concerns spinors, if we use the standard representation we can approximate them as

$$u_1(\vec{p}) \approx u_1(0) = \sqrt{2m}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \qquad u_2(\vec{p}) \approx u_2(0) = \sqrt{2m}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \tag{4.168}$$

$$v_1(\vec{p}) \approx v_1(0) = -\sqrt{2m}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}, \qquad v_2(\vec{p}) \approx v_2(0) = \sqrt{2m}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}. \tag{4.169}$$

Let us consider the electron scattering off an approximately static muon target. The relevant part of the interaction in the nonrelativistic limit comes from $\beta = 0$ in Eq. (4.157) (only $\bar{u}(0)\gamma^0 u(0)$ survives), which implies in its turn $\alpha = 0$. Then we get

$$i\mathcal{M}_{(e\mu)} \approx \frac{-ie^2}{|\vec{k}_f - \vec{k}_i|^2}\overline{u}_{\mathrm{e},s_f}(\vec{k}_f)\gamma^0 u_{\mathrm{e},s_i}(\vec{k}_i)\overline{u}_{\mathrm{m},r_f}(0)\gamma^0 u_{\mathrm{m},r_i}(0). \tag{4.170}$$

The scattering off an external EM field can be described by the following Feynman amplitude (cf. Eq. (8.88) of Mandl-Shaw)

$$i\mathcal{M} = \frac{ie}{2m}\overline{u}_{\mathrm{e},s_f}(\vec{k}_f)\, \not{A}_{\mathrm{ext}}(\vec{k}_f - \vec{k}_i)\, u_{\mathrm{e},s_i}(\vec{k}_i). \tag{4.171}$$

In our case, the external EM field is the one generated by the nonrelativistic approximation of the interaction with the muon target, wich turn out to be, by comparison with Eq. (4.170)

$$A^{\rho}_{\mathrm{ext}}(\vec{q}) = \left(-\frac{e}{|\vec{q}|^2},\, 0,\, 0,\, 0 \right). \tag{4.172}$$

This corresponds to the electrostatic potential generated by a negative charge. Its Fourier transform is

$$\Phi(r) = -\frac{e}{r}, \tag{4.173}$$

which is precisely the Coulomb potential of a negative charge (in razionalized Gaussian units).

What happens if we consider the scattering off an antimuon, instead of a muon? we will have a sequence of ladder operators $N[c^{\dagger}_{\mathrm{e}} c_{\mathrm{e}} d_{\mathrm{m}} d^{\dagger}_{\mathrm{m}}] = -c^{\dagger}_{\mathrm{e}} c_{\mathrm{e}} d^{\dagger}_{\mathrm{m}} d_{\mathrm{m}}$, which introduces a minus sign with respect to the muon case. The rest remains exactly the same and the corresponding potential becomes that of a *positive* charge. In other words, the change of sign between a particle-particle or particle-antiparticle scattering (from repulsive to attractive) is due to normal ordering. The change of sign is typical of interactions mediated by a vector particle (the photon).

Let us analyze now the Yukawa case. Diagram (Y2), Eq. (4.165), will not contribute, because spinor products $\bar{u}v$ and $\bar{v}u$ are suppressed. This means that particle-antiparticle annihilation is not possible in the nonrelativistic limit. Diagram (Y1), Eq. (4.164), gives

$$i\mathcal{M}_{(\mathrm{Y1})} \approx \frac{ig^2}{|\vec{k}_f - \vec{k}_i|^2 + M^2}\overline{u}_{s_f}(\vec{k}_f) u_{s_i}(\vec{k}_i)\, \overline{v}_{r_i}(0) v_{r_f}(0) \tag{4.174}$$

In this case, the scattering off a static external scalar field would just be

$$i\mathcal{M} = \frac{ig}{2m}\overline{u}_{\mathrm{e},s_f}(\vec{k}_f)\, u_{\mathrm{e},s_i}(\vec{k}_i)\phi_{\mathrm{ext}}(\vec{k}_f - \vec{k}_i), \tag{4.175}$$

which leads to the identification

$$\phi_{\mathrm{ext}}(\vec{q}) = \frac{g}{|\vec{q}|^2 + M^2} \tag{4.176}$$

and a corresponding potential

$$V(r) \propto \frac{g}{r} e^{-Mr} \tag{4.177}$$

which is indeed the short-ranged Yukawa potential. Notice that the potential is positive. The particle-antiparticle interaction is attractive.

What happens if we scatter a proton off another proton, instead of an antiproton? First, we will have a sequence of ladder operators $N[c^\dagger c c^\dagger c] = -c^\dagger c^\dagger c c$, which introduces a minus sign with respect to the previous case. This is similar to the electron-muon case. However, we have a product $\overline{u}u$ instead of $\overline{v}v$, which introduces another minus sign. Overall, the sign of the scattering amplitude does not change from proton-proton to proton-antiproton and the potential is attractive in both cases. This is typical of an interaction mediated by a scalar particle.

4.8 Second-order contribution in QED: general case

We now proceed to considering the second-order terms in the expansion of the scattering matrix for QED:

$$S^{(2)} = -\frac{e^2}{2!}\int d^4x_1 d^4x_2\; T\Big[N\Big[\overline{\psi}(x_1)\not{A}(x_1)\psi(x_1)\Big]\; N\Big[\overline{\psi}(x_2)\not{A}(x_2)\psi(x_2)\Big]\Big]. \tag{4.178}$$

Application of Wick's theorem leads to six possible independent terms. Remember that we should not take into account equal-time contractions and that the nonzero contractions are only the ones between $\overline{\psi}\psi$, $\psi\overline{\psi}$, and AA).

$$S^{(2)}_{(a)} = -\frac{e^2}{2!}\int d^4x_1 d^4x_2\; N\Big[\big(\overline{\psi}\gamma^\alpha A_\alpha\psi\big)_{x_1}\big(\overline{\psi}\gamma^\beta A_\beta\psi\big)_{x_2}\Big], \tag{4.179}$$

$$S^{(2)}_{(b)} = -\frac{e^2}{2!}\int d^4x_1 d^4x_2\; N\Big[\big(\overline{\psi}\gamma^\alpha A_\alpha\underbrace{\psi}\big)_{x_1}\big(\underbrace{\overline{\psi}}\gamma^\beta A_\beta\psi\big)_{x_2} + \big(\underbrace{\overline{\psi}\gamma^\alpha A_\alpha\psi}\big)_{x_1}\big(\overline{\psi}\gamma^\beta A_\beta\psi\big)_{x_2}\Big], \tag{4.180}$$

$$S^{(2)}_{(c)} = -\frac{e^2}{2!}\int d^4x_1 d^4x_2\; N\Big[\big(\overline{\psi}\gamma^\alpha\underbrace{A_\alpha\psi}\big)_{x_1}\big(\overline{\psi}\gamma^\beta A_\beta\psi\big)_{x_2}\Big] \tag{4.181}$$

$$S^{(2)}_{(d)} = -\frac{e^2}{2!}\int d^4x_1 d^4x_2\; N\Big[\big(\overline{\psi}\gamma^\alpha\underbrace{A_\alpha\psi}\big)_{x_1}\big(\overline{\psi}\gamma^\beta A_\beta\psi\big)_{x_2} + \big(\overline{\psi}\gamma^\alpha A_\alpha\psi\big)_{x_1}\big(\overline{\psi}\gamma^\beta A_\beta\psi\big)_{x_2}\Big], \tag{4.182}$$

$$S^{(2)}_{(e)} = -\frac{e^2}{2!}\int d^4x_1 d^4x_2\; N\Big[\big(\overline{\psi}\gamma^\alpha A_\alpha\psi\big)_{x_1}\big(\overline{\psi}\gamma^\beta A_\beta\psi\big)_{x_2}\Big], \tag{4.183}$$

$$S^{(2)}_{(f)} = -\frac{e^2}{2!}\int d^4x_1 d^4x_2\; N\Big[\big(\overline{\psi}\gamma^\alpha A_\alpha\psi\big)_{x_1}\big(\overline{\psi}\gamma^\beta A_\beta\psi\big)_{x_2}\Big], \tag{4.184}$$

$S_{(a)}$ is easy to analyze because it consists of two independent terms equal to the first-order one. They correspond to the combination of two of the processes described in the previous section and not physically possible.

Vacuum bubbles

$S_{(f)}$ is also easy to analyze: it contains three propagators and nothing else, therefore it describes no transition and is also called a vacuum diagram or vacuum bubble. This kind of diagrams do not give a contribution to scattering processes (it can be combined with any diagram where particles go through unscattered) and should be omitted in the calculation of the scattering matrix. Also at higher orders, it turns out that these kinds of contributions (called "disconnected diagrams") can be omitted in the calculation of the scattering matrix.

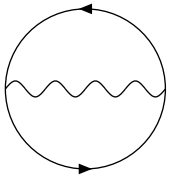


Figure 4.11: Graphical representation of the operator $S^{(2)}_{(f)}$, the so-called vacuum bubble.

Self-energies

$S_{(e)}$ contains two uncontracted photon operators. Note that

$$\begin{aligned} & N\Big[\big(\overline{\psi}\gamma^\alpha A_\alpha \psi\big)_{x_1}\big(\overline{\psi}\gamma^\beta A_\beta \psi\big)_{x_2}\Big] \\ &\quad = \overline{\psi}_A(x_1)\psi_B(x_2)\psi_C(x_1)\overline{\psi}_D(x_2)\gamma^\alpha_{AC}\gamma^\beta_{DB}N\big[A_\alpha(x_1)A_\beta(x_2)\big] \\ &\quad = (-1)\psi_B(x_2)\overline{\psi}_A(x_1)\psi_C(x_1)\overline{\psi}_D(x_2)\gamma^\alpha_{AC}\gamma^\beta_{DB}N\big[A_\alpha(x_1)A_\beta(x_2)\big] \\ &\quad = (-1)iS_{F,BA}(x_2-x_1)\gamma^\alpha_{AC}iS_{F,CD}(x_1-x_2)\gamma^\beta_{DB}N\big[A_\alpha(x_1)A_\beta(x_2)\big] \\ &\quad = (-1)\mathrm{Tr}\big[iS_F(x_2-x_1)\gamma^\alpha iS_F(x_1-x_2)\gamma^\beta\big]N\big[A_\alpha(x_1)A_\beta(x_2)\big]. \end{aligned} \tag{4.185}$$

We can choose $+$ or $-$ components of the two A fields: there are four possible combinations (see Fig. 4.12). The two fermion propagators form a "loop," and this is in fact called a *loop diagram* (vacuum bubbles are also loop diagrams). The occurrence of the $-$ sign in Eq. (4.185) is typical of fermion loops. Of the four combinations in Fig. 4.12, only two are physical, in the sense that they can be consistent with momentum conservation. The two diagrams are actually equivalent, as can be seen upon exchange of integration variables, and contribute to the so-called photon self-energy. At the order we are considering, the photon self-energy does not contribute to a scattering process. At higher orders, however, it must be taken into account and induces a modification of the properties of the photon itself. The momentum integrals included in the fermion propagators generate divergences that have to be dealt with using renormalization.

$S_{(d)}$ contains a fermion and a photon propagator, leaving two fermion fields to act on the initial and final states. First of all, we check that the two different terms in the expression of $S_{(d)}$ are the same. In the second term, we can permute the two blocks $(\overline{\psi}_A \not{A}_{AB}\psi_B)$

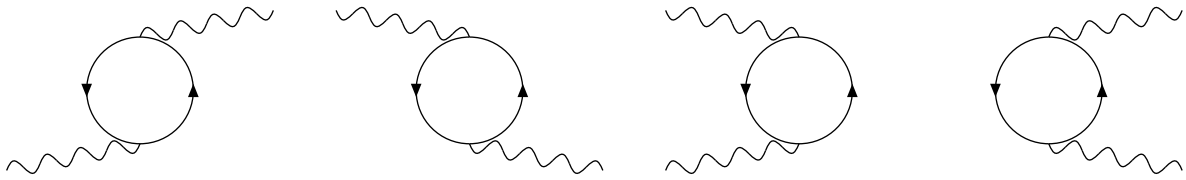


Figure 4.12: Graphical representation of the various terms contained in the operator $S^{(2)}_{(e)}$. The first two terms are equivalent and represent the photon self-energy, which however does not give a contribution to a scattering process at the order we are considering. The last two contributions are unphysical.

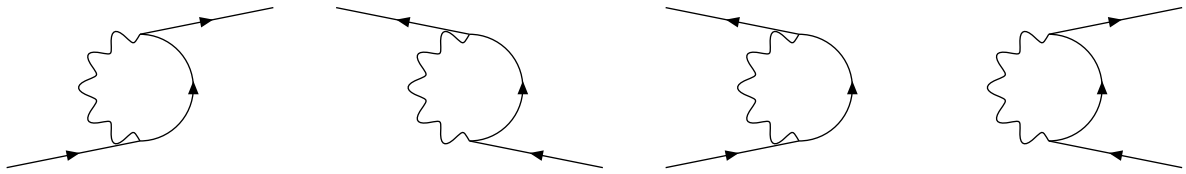


Figure 4.13: Graphical representation of the various terms contained in the operator $S^{(2)}_{(d)}$. The first term is the electron self-energy diagram, the second is the positron self-energy diagram. They do not give a contribution to a scattering process at the order we are considering. The last two contributions are unphysical.

(no sign change is required since the block contains a pair of fermion fields and no problem arises from the Dirac structure); then we exchange the integration variable $x_1 \leftrightarrow x_2$, to obtain the same thing as the first term. Therefore, the $S_{(d)}$ term can be rewritten in a shorter way (note that the disappearance of the 2! factor in the denominator)

$$S^{(2)}_{(d)} = -e^2 \int d^4x_1 d^4x_2 \, N\Big[\underbrace{(\overline{\psi}\gamma^\alpha A_\alpha \psi)_{x_1} (\overline{\psi}\gamma^\beta A_\beta \psi)_{x_2}}\Big], \tag{4.186}$$

Similar to before, there are in principle four possibilities depending on the choice of $+$ and $-$ components of the uncontracted fermion fields. They are illustrated in Fig. 4.13 and are again examples of loop diagrams. Only two of them are physical: a particle can be absorbed and then created, in which case we are considering a contribution to a process where there is a particle in the initial and final state, or the same can happen with an antiparticle. These processes correspond to the so-called electron and positron self-energy diagrams.

Two-to-two processes

$S_{(b)}$ contains a fermion propagator, two fermion fields and two photon fields. As for $S_{(d)}$, the two different terms are the same and we can write simply

$$S^{(2)}_{(b)} = -e^2 \int d^4x_1 d^4x_2 \, N\Big[\underbrace{(\overline{\psi}\gamma^\alpha A_\alpha \psi)_{x_1}} (\overline{\psi}\gamma^\beta A_\beta \psi)_{x_2}\Big], \tag{4.187}$$

There are in principle 16 different combinations of $+$ and $-$ components of the two fermion fields and two photon fields. Many of them are nonphysical: for instance, a combination with all $+$ components means that we are absorbing two fermions and two

photons and creating nothing. The physical combinations (i.e., the ones that can be consistent with momentum conservation) contain two particles in the initial state and two in the final state and are only six (corresponding to four different processes), i.e.,

$$\begin{aligned} S^{(2)}_{(b1)} &= -e^2 \int d^4x_1 d^4x_2 \; \overline{\psi}^-_A(x_1)\gamma^\alpha_{AC} A^-_\alpha(x_1) i S_{F,CD}(x_1-x_2)\gamma^\beta_{DB} A^+_\beta(x_2)\psi^+_B(x_2) \\ &= -e^2 \int d^4x_1 d^4x_2 \; \left(\gamma^\alpha i S_F(x_1-x_2)\gamma^\beta\right)_{AB} \overline{\psi}^-_A(x_1) A^-_\alpha(x_1) A^+_\beta(x_2)\psi^+_B(x_2), \end{aligned} \tag{4.188}$$

$$S^{(2)}_{(b2)} = -e^2 \int d^4x_1 d^4x_2 \; \left(\gamma^\alpha i S_F(x_1-x_2)\gamma^\beta\right)_{AB} \overline{\psi}^-_A(x_1) A^-_\beta(x_2) A^+_\alpha(x_1)\psi^+_B(x_2), \tag{4.189}$$

both corresponding to the process $\gamma e^- \to \gamma e^-$ (Compton scattering by electrons)

$$S^{(2)}_{(b3)} = +e^2 \int d^4x_1 d^4x_2 \; \left(\gamma^\alpha i S_F(x_1-x_2)\gamma^\beta\right)_{AB} \psi^-_B(x_2) A^-_\alpha(x_1) A^+_\beta(x_2)\overline{\psi}^+_A(x_1), \tag{4.190}$$

$$S^{(2)}_{(b4)} = +e^2 \int d^4x_1 d^4x_2 \; \left(\gamma^\alpha i S_F(x_1-x_2)\gamma^\beta\right)_{AB} \psi^-_B(x_2) A^-_\beta(x_2) A^+_\alpha(x_1)\overline{\psi}^+_A(x_1), \tag{4.191}$$

both corresponding to the process $\gamma e^+ \to \gamma e^+$ (Compton scattering by positrons). Note the change of sign due to the fact that we inverted two anticommuting fermion fields.

$$S^{(2)}_{(b5)} = -e^2 \int d^4x_1 d^4x_2 \; \left(\gamma^\alpha i S_F(x_1-x_2)\gamma^\beta\right)_{AB} \overline{\psi}^-_A(x_1)\psi^-_B(x_2) A^+_\beta(x_2) A^+_\alpha(x_1), \tag{4.192}$$

corresponding to the process $\gamma\gamma \to e^+e^-$ (pair creation)

$$S^{(2)}_{(b6)} = -e^2 \int d^4x_1 d^4x_2 \; \left(\gamma^\alpha i S_F(x_1-x_2)\gamma^\beta\right)_{AB} A^-_\beta(x_2) A^-_\alpha(x_1)\overline{\psi}^+_A(x_1)\psi^+_B(x_2), \tag{4.193}$$

corresponding to the process $e^+e^- \to \gamma\gamma$ (pair annihilation).

$S_{(c)}$ contains a photon propagator and four fermion fields acting on the initial and final states. Of all the possible combinations of creation and absorption operators, the only ones that are physical are the combinations describing $e^-e^+ \to e^+e^-$ (Bhabha scattering), $e^-e^- \to e^-e^-$ or $e^+e^+ \to e^+e^+$ (Møller scattering).

$$S^{(2)}_{(c1)} = -\frac{e^2}{2!} \int d^4x_1 d^4x_2 \; \left(\gamma^\alpha_{AB} i D_{F,\alpha\beta}(x_1-x_2)\gamma^\beta_{CD}\right) \overline{\psi}^-_A(x_1)\psi^-_B(x_1)\overline{\psi}^+_C(x_2)\psi^+_D(x_2), \tag{4.194}$$

$$S^{(2)}_{(c2)} = -\frac{e^2}{2!} \int d^4x_1 d^4x_2 \; \left(\gamma^\alpha_{AB} i D_{F,\alpha\beta}(x_1-x_2)\gamma^\beta_{CD}\right) \overline{\psi}^-_C(x_2)\psi^-_D(x_2)\overline{\psi}^+_A(x_1)\psi^+_B(x_1), \tag{4.195}$$

$$S^{(2)}_{(c3)} = +\frac{e^2}{2!} \int d^4x_1 d^4x_2 \; \left(\gamma^\alpha_{AB} i D_{F,\alpha\beta}(x_1-x_2)\gamma^\beta_{CD}\right) \overline{\psi}^-_A(x_1)\psi^-_D(x_2)\overline{\psi}^+_C(x_2)\psi^+_B(x_1), \tag{4.196}$$

$$S^{(2)}_{(c4)} = +\frac{e^2}{2!} \int d^4x_1 d^4x_2 \; \left(\gamma^\alpha_{AB} i D_{F,\alpha\beta}(x_1-x_2)\gamma^\beta_{CD}\right) \overline{\psi}^-_C(x_2)\psi^-_B(x_1)\overline{\psi}^+_A(x_1)\psi^+_D(x_2), \tag{4.197}$$

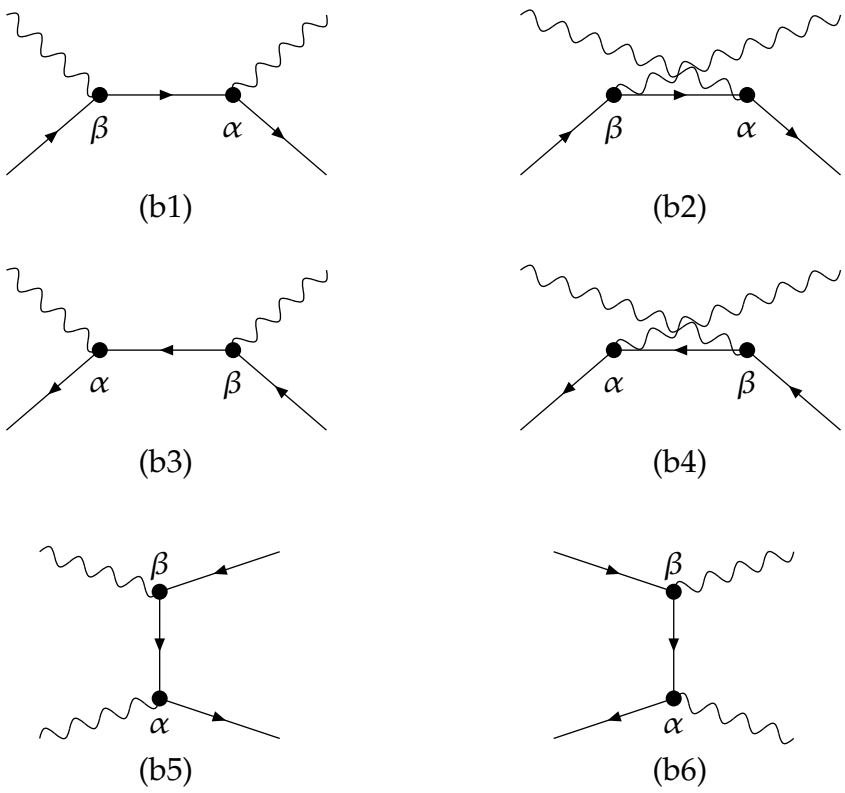


Figure 4.14: Graphical representations of the terms contained in $S^{(2)}_{(b)}$. First row: Compton scattering by electrons; second row: Compton scattering by positron; last row: $\gamma\gamma \to e^+e^-$ and $e^+e^- \to \gamma\gamma$.

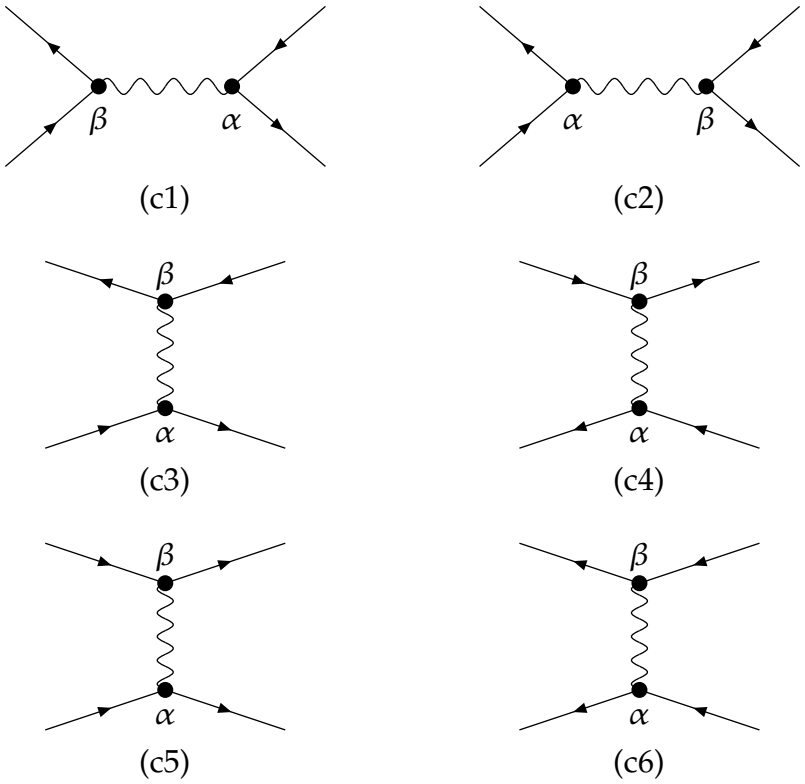


Figure 4.15: Graphical representations of the terms contained in $S^{(2)}_{(c)}$. First two rows: $e^+e^- \to e^+e^-$ (Bhabha scattering); the two contributions in each row are equivalent. Last row: Møller scattering by electrons or positrons.

The above expressions all contribute to the process $e^+e^- \to e^+e^-$ (Bhabha scattering). Note the sign change of the last two contributions, $c3$ and $c4$. Permuting the $\overline{\psi}\gamma\psi$ blocks and interchanging the integration variables $x_1 \leftrightarrow x_2$ leads to the conclusion that the first two expressions ($c1$ and $c2$) and the last two expressions ($c3$ and $c4$) are the same (which allows us to consider only one of them and remove the 2!).

$$S^{(2)}_{(c5)} = -\frac{e^2}{2!}\int d^4x_1 d^4x_2 \left(\gamma^\alpha_{AB} iD_{F,\alpha\beta}(x_1 - x_2)\gamma^\beta_{CD}\right)\overline{\psi}^-_A(x_1)\overline{\psi}^-_C(x_2)\psi^+_B(x_1)\psi^+_D(x_2), \quad (4.198)$$

corresponding to the process $e^-e^- \to e^-e^-$ (electron Møller scattering),

$$S^{(2)}_{(c6)} = -\frac{e^2}{2!}\int d^4x_1 d^4x_2 \left(\gamma^\alpha_{AB} iD_{F,\alpha\beta}(x_1 - x_2)\gamma^\beta_{CD}\right)\psi^-_B(x_1)\psi^-_D(x_2)\overline{\psi}^+_A(x_1)\overline{\psi}^+_C(x_2), \quad (4.199)$$

corresponding to the process $e^+e^+ \to e^+e^+$ (positron Møller scattering).

The different contributions to $S^{(2)}_{(b)}$ and $S^{(2)}_{(c)}$ are all examples of *tree-level diagrams*, e.g., diagrams with no loops.

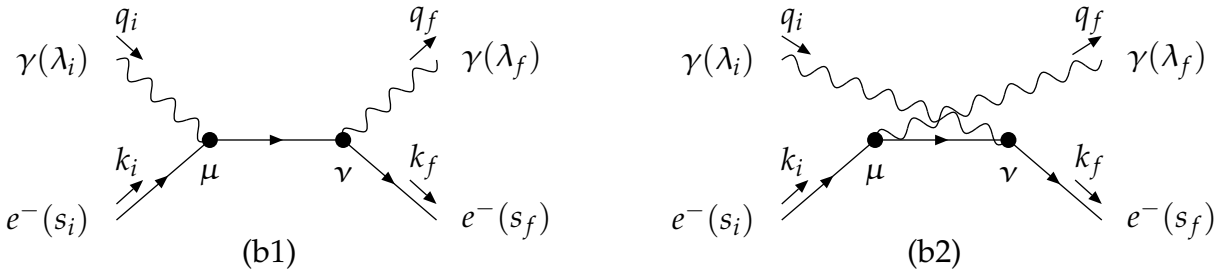


Figure 4.16: Feynman diagrams contributing to Compton scattering.

4.8.1 Example: Compton scattering

We now have to see what happens when applying each term of the $S^{(2)}$ scattering matrix to the appropriate initial and final states. Let us consider the case of electron Compton scattering. The initial and final states are in this case

$$|i\rangle = a^\dagger_{\lambda_i}(q_i)c^\dagger_{s_i}(k_i)|0\rangle, \qquad \langle f| = \langle 0|c_{s_f}(k_f)a_{\lambda_f}(q_f). \tag{4.200}$$

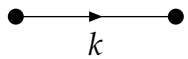
To compute the result of acting with the $S^{(2)}_{(b1)}$ matrix we need to insert the expressions of the fields and of the propagator. Without repeating all the steps, we already know that we can use the commutation or anticommutation rules for the ladder operators and identify the only nonvanishing contributions, i.e., the ones where the momenta and spins of the field operators correspond to the momenta and fields of the external particles. We obtain

$$\begin{aligned}
\langle f|S^{(2)}_{(b1)}|i\rangle &= -e^2 \int d^4x_1 d^4x_2\, \overline{u}_{s_f}(k_f)\, e^{ik_f x_1}\gamma_\nu \epsilon^{*\nu}_{\lambda_f}(q_f) e^{iq_f x_1} \\
&\quad \times \int \frac{d^4k}{(2\pi)^4}\, \frac{i\,(\not{k}+m)e^{-ik(x_1-x_2)}}{\left(k^2-m^2+i\epsilon\right)}\, \gamma_\mu \epsilon^\mu_{\lambda_i}(q_i) e^{-iq_i x_2}\, u_{s_i}(k_i)\, e^{-ik_i x_2} \\
&= -e^2 \int \frac{d^4k}{(2\pi)^4}\, d^4x_1 d^4x_2\, e^{-i(k_i+q_i-k)x_2}\, e^{-i(k-k_f-q_f)x_1} \\
&\quad \times \overline{u}_{s_f}(k_f)\, \gamma_\nu \epsilon^{*\nu}_{\lambda_f}(q_f)\, \frac{i\,(\not{k}+m)}{\left(k^2-m^2+i\epsilon\right)}\, \gamma_\mu \epsilon^\mu_{\lambda_i}(q_i)\, u_{s_i}(k_i) \\
&= -e^2 \int \frac{d^4k}{(2\pi)^4}(2\pi)^4\delta(k_i+q_i-k)(2\pi)^4\delta(k-k_f-q_f) \\
&\quad \times \overline{u}_{s_f}(k_f)\, \gamma_\nu \epsilon^{*\nu}_{\lambda_f}(q_f)\, \frac{i\,(\not{k}+m)}{\left(k^2-m^2+i\epsilon\right)}\, \gamma_\mu \epsilon^\mu_{\lambda_i}(q_i)\, u_{s_i}(k_i) \\
&= (2\pi)^4\delta(k_i+q_i-k_f-q_f) \\
&\quad \times (-e^2)\overline{u}_{s_f}(k_f)\, \gamma_\nu \epsilon^{*\nu}_{\lambda_f}(q_f)\, \frac{i\,(\not{k}+m)}{\left(k^2-m^2+i\epsilon\right)}\bigg|_{k=k_i+q_i} \gamma_\mu \epsilon^\mu_{\lambda_i}(q_i)\, u_{s_i}(k_i).
\end{aligned} \tag{4.201}$$

Also in this case, conservation of momentum holds true, not only for the sum of the initial and final momenta, but also at each interaction vertex separately.

As for the first-order case, we can identify the Feynman amplitude, which can be directly obtained by drawing the Feynman diagram of Fig. 4.16 (b1) and applying Feynman

rules. To the ones already described before we have to add the rule for the fermion propagator

Fermion propagator:  $= \dfrac{i(\not{k}+m)}{k^2-m^2+i\epsilon}$ (4.202)

The expression for the Feynman amplitude turns out to be

$$i\mathcal{M}_{(b1)} = -e^2\,\overline{u}_{s_f}(k_f)\,\gamma_\nu \epsilon^{*\nu}_{\lambda_f}(q_f)\,\frac{i\,(\not{k}+m)}{\left(k^2-m^2+i\epsilon\right)}\bigg|_{k=k_i+q_i}\gamma_\mu \epsilon^{\mu}_{\lambda_i}(q_i)\,u_{s_i}(k_i). \tag{4.203}$$

The scattering matrix term $S_{(b2)}$ is very similar to the previous calculation, except for the fact that the incoming and outgoing photon are attached to different vertices. This in turn changes how the momentum flowing in the propagator is fixed by momentum conservation. From the Feynman diagram of Fig. 4.16 (b2), we can directly obtain

$$i\mathcal{M}_{(b2)} = -e^2\,\overline{u}_{s_f}(k_f)\,\gamma_\nu \epsilon^{\nu}_{\lambda_i}(q_i)\,\frac{i\,(\not{k}+m)}{\left(k^2-m^2+i\epsilon\right)}\bigg|_{k=k_i-q_f}\gamma_\mu \epsilon^{*\mu}_{\lambda_f}(q_f)\,u_{s_i}(k_i). \tag{4.204}$$

Compton scattering by positron (terms $S_{(b3)}$ and $S_{(b4)}$) behave in a very similar way, except for the *overall sign that has to be reversed* due to the inversion of fermion fields, as shown in Eq. (4.191). This overall sign is however irrelevant in the calculation of physical observables.

A remark on gauge invariance and unphysical polarization states is in order here. In the above expressions, we should consider only physical photon polarization states. However, it is possible to show that unphysical states give a vanishing contribution. Remembering Eq. (3.262), we need to check that no contribution arises if we replace either or both $\epsilon^{*\nu}_{\lambda_f}(q_f) \to q_f^\nu$ and $\epsilon^{\mu}_{\lambda_i}(q_i) \to q_i^\mu$. Let us check in fact what happens if we do the first replacement. We need to use the following identities

$$k_i + q_i = k_f + q_f, \qquad \not{q}_f \not{k}_f = -\not{k}_f \not{q}_f + 2k_f\cdot q_f, \qquad \overline{u}(k_f)(\not{k}_f - m) = 0. \tag{4.205}$$

to obtain

$$\begin{aligned} i\mathcal{M}_{(b1)} &= -e^2\,\overline{u}_{s_f}(k_f)\,\not{q}_f\,\frac{i\,(\not{k}_f+\not{q}_f+m)}{\left(2k_f\cdot q_f+i\epsilon\right)}\,\gamma_\mu \epsilon^{\mu}_{\lambda_i}(q_i)\,u_{s_i}(k_i) \\ &= -e^2\,\overline{u}_{s_f}(k_f)\not{\epsilon}_{\lambda_i}(q_i)\,u_{s_i}(k_i) \end{aligned} \tag{4.206}$$

The amplitude does not vanish by itself, however we still have to sum it to the contribution $(b2)$. Using

$$\not{q}_f \not{k}_i = -\not{k}_i \not{q}_f + 2k_i\cdot q_f, \qquad\qquad (\not{k}_i - m)u(k_i) = 0. \tag{4.207}$$

we obtain

$$\begin{aligned} i\mathcal{M}_{(b2)} &= -e^2\, \overline{u}_{s_f}(k_f)\, \gamma_\nu \epsilon^\nu_{\lambda_i}(q_i)\, \frac{i\,(\not{k}_i - \not{q}_f + m)}{(\,-2k_i\cdot q_f + i\epsilon)}\, \not{\epsilon}_f\, u_{s_i}(k_i). \\ &= e^2\, \overline{u}_{s_f}(k_f) \not{\epsilon}_{\lambda_i}(q_i)\, u_{s_i}(k_i). \end{aligned} \tag{4.208}$$

Therefore, we can conclude that in this process contributions from unphysical polarization states cancel in the sum of the amplitudes. This is a specific example of Ward's identity.

4.8.2 Relation to time-ordered perturbation theory [optional]

In Eq. (4.201) we obtained that four-momentum is conserved at each vertex. It is interesting to note that this depends on the fact that we used the propagator in its manifestly covariant form, Eq. (4.115). Suppose we use the time-ordered version of Eq. (4.114), which is equivalent to integrating the scattering matrix expression over the k^0 variable. We obtain

$$\begin{aligned} \langle f|S^{(2)}_{(b1)}|i\rangle &= -e^2 \int d^4x_1 d^4x_2\, \overline{u}_{s_f}(k_f)\, e^{ik_f x_1} \gamma_\nu \epsilon^{*\nu}_{\lambda_f}(q_f) e^{iq_f x_1} \theta(t_1 - t_2)\, e^{-i\omega_k(t_1-t_2)} \\ &\quad \times \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, (\not{k} + m)\Big|_{k^0=\omega_k}\, e^{i\vec{k}\cdot(\vec{x}_1-\vec{x}_2)}\, \gamma_\mu \epsilon^\mu_{\lambda_i}(q_i) e^{-iq_i x_2}\, u_{s_i}(k_i)\, e^{-ik_i x_2} \\ &\quad - e^2 \int d^4x_1 d^4x_2\, \overline{u}_{s_f}(k_f)\, e^{ik_f x_1} \gamma_\nu \epsilon^{*\nu}_{\lambda_f}(q_f) e^{iq_f x_1} \theta(t_2 - t_1)\, e^{-i\omega_k(t_2-t_1)} \\ &\quad \times \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, (-\not{k} + m)\Big|_{k^0=\omega_k}\, e^{-i\vec{k}\cdot(\vec{x}_1-\vec{x}_2)}\, \gamma_\mu \epsilon^\mu_{\lambda_i}(q_i) e^{-iq_i x_2}\, u_{s_i}(k_i)\, e^{-ik_i x_2}. \end{aligned} \tag{4.209}$$

To keep the discussion short, let us focus only on the first term. Let us write explicitly the time and three-vector components of each four-vector, then integrate over the space components of x_1 and x_2 and finally integrate over the space components of k:

$$\begin{aligned} \langle f|S^{(2)}_{(b1)}|i\rangle &= -e^2 \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, dt_1 d^3x_1 dt_2 d^3x_2\, e^{-i(\omega_{k_i}+\omega_{q_i}-\omega_k)t_2}\, e^{i(\vec{k}_i+\vec{q}_i-\vec{k})\cdot\vec{x}_2} \\ &\quad \times e^{-i(\omega_k-\omega_{k_f}-\omega_{q_f})t_1}\, e^{i(\vec{k}-\vec{k}_f-\vec{q}_f)\cdot\vec{x}_1}\, \theta(t_1-t_2) \\ &\quad \times \overline{u}_{s_f}(\vec{k}_f)\, \gamma_\nu \epsilon^{*\nu}_{\lambda_f}(\vec{q}_f)\, (\omega_k\gamma^0 - \vec{k}\cdot\vec{\gamma} + m)\, \gamma_\mu \epsilon^\mu_{\lambda_i}(\vec{q}_i)\, u_{s_i}(\vec{k}_i) + \ldots \\ &= -e^2 \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, dt_1 dt_2\, e^{-i(\omega_{k_i}+\omega_{q_i}-\omega_k)t_2} (2\pi)^3 \delta^3(\vec{k}_i + \vec{q}_i - \vec{k}) \\ &\quad \times e^{-i(\omega_k-\omega_{k_f}-\omega_{q_f})t_1} (2\pi)^3 \delta^3(\vec{k} - \vec{k}_f - \vec{q}_f)\, \theta(t_1-t_2) \\ &\quad \times \overline{u}_{s_f}(\vec{k}_f)\, \gamma_\nu \epsilon^{*\nu}_{\lambda_f}(\vec{q}_f)\, (\omega_k\gamma^0 - \vec{k}\cdot\vec{\gamma} + m)\, \gamma_\mu \epsilon^\mu_{\lambda_i}(\vec{q}_i)\, u_{s_i}(\vec{k}_i) + \ldots \\ &= -e^2 (2\pi)^3 \delta^3(\vec{k}_i + \vec{q}_i - \vec{k}_f - \vec{q}_f) \int dt_1 dt_2\, e^{-i(\omega_{k_i}+\omega_{q_i}-\omega_k)t_2} \end{aligned}$$

$$\times e^{-i(\omega_k-\omega_{k_f}-\omega_{q_f})t_1}\,\theta(t_1-t_2)$$
$$\times\,\overline{u}_{s_f}(\vec{k}_f)\,\gamma_\nu\epsilon^{*\nu}_{\lambda_f}(\vec{q}_f)\frac{(\omega_k\gamma^0-\vec{k}\cdot\vec{\gamma}+m)}{2\omega_k}\bigg|_{\vec{k}=\vec{k}_i+\vec{q}_i}\gamma_\mu\epsilon^{\mu}_{\lambda_i}(\vec{q}_i)\,u_{s_i}(\vec{k}_i)+\dots \tag{4.210}$$

We are still left with the two integrations over the time components of x_1 and x_2. In the covariant calculation, these two integrations where done before the k^0 integration and led to the conservation of energy at each vertex of the diagram. Here, it is convenient to perform the change of variable $t_1 = t + t_2$:

$$\langle f|S^{(2)}_{(b1)}|i\rangle = -e^2(2\pi)^3\delta^3(\vec{k}_i+\vec{q}_i-\vec{k}_f-\vec{q}_f)\int dt\,dt'\,e^{-i(\omega_{k_i}+\omega_{q_i}-\omega_{k_f}-\omega_{q_f})t_2}e^{-i(\omega_k-\omega_{k_f}-\omega_{q_f})t}$$
$$\times\,\theta(t)\,\overline{u}_{s_f}(\vec{k}_f)\,\gamma_\nu\epsilon^{*\nu}_{\lambda_f}(\vec{q}_f)\frac{(\omega_k\gamma^0-\vec{k}\cdot\vec{\gamma}+m)}{2\omega_k}\bigg|_{\vec{k}=\vec{k}_i+\vec{q}_i}\gamma_\mu\epsilon^{\mu}_{\lambda_i}(\vec{q}_i)\,u_{s_i}(\vec{k}_i)+\dots \tag{4.211}$$

The t_2 integration corresponds to the Fourier transform of a constant, giving us a delta function:

$$\int dt_2\,e^{-i(\omega_{k_i}+\omega_{q_i}-\omega_{k_f}-\omega_{q_f})t_2} = 2\pi\delta(\omega_{k_i}+\omega_{q_i}-\omega_{k_f}-\omega_{q_f}), \tag{4.212}$$

and a Fourier transform of a step function that gives us (considering already that the previous delta function implies $\omega_{k_i}+\omega_{q_i} = \omega_{k_f}+\omega_{q_f}$)

$$\int dt\,e^{-i(\omega_k-\omega_{k_i}-\omega_{q_i})t}\,\theta(t) = \frac{-i}{(\omega_k-\omega_{k_i}-\omega_{q_i})-i\epsilon}. \tag{4.213}$$

The last equality can be checked by anti-Fourier transforming and using Cauchy's theorem.

In conclusion, reinserting explicitly the contribution from the second time-ordered diagram, we obtain the following expression

$$\langle f|S^{(2)}_{(b1)}|i\rangle = (2\pi)^4\delta^4(k_i+q_i-k_f-q_f)\,e^2$$
$$\times\left(\overline{u}_{s_f}(k_f)\,\gamma_\nu\epsilon^{*\nu}_{\lambda_f}(q_f)\frac{i(\omega_k\gamma^0-\vec{k}\cdot\vec{\gamma}+m)}{2\omega_k(\omega_k-\omega_{k_i}-\omega_{q_i}-i\epsilon)}\bigg|_{\vec{k}=\vec{k}_i+\vec{q}_i}\gamma_\mu\epsilon^{\mu}_{\lambda_i}(q_i)\,u_{s_i}(k_i)\right.$$
$$\left.+\,\overline{u}_{s_f}(k_f)\,\gamma_\nu\epsilon^{*\nu}_{\lambda_f}(q_f)\frac{i(\omega_k\gamma^0-\vec{k}\cdot\vec{\gamma}-m)}{2\omega_k(\omega_k+\omega_{k_i}+\omega_{q_i}-i\epsilon)}\bigg|_{\vec{k}=-(\vec{k}_i+\vec{q}_i)}\gamma_\mu\epsilon^{\mu}_{\lambda_i}(q_i)\,u_{s_i}(k_i)\right). \tag{4.214}$$

Comparing this last result with the expression in Eq. (4.201) we can highlight similarities and differences between the "covariant" version and the "time-ordered" version of the calculation. They are equivalent, but can be given different interpretations.

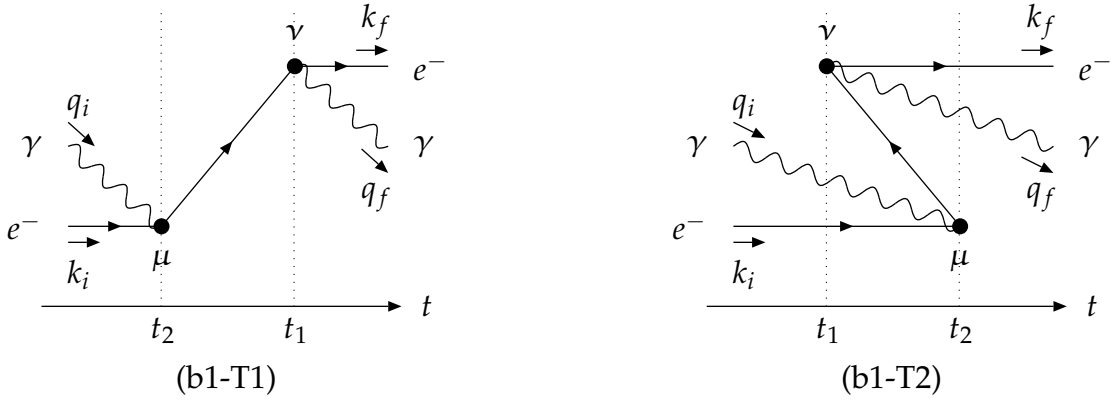


Figure 4.17: Time-ordered Feynman diagrams contributing to Compton scattering in the s-channel. Their sum corresponds to Fig. 4.16 b1. (Spin and polarization indices are not written.)

In both cases, the *total four-momentum is conserved*.

In the covariant treatment, Eq. (4.201), energy and momentum are conserved at each vertex, but the intermediate state is off-shell, i.e., it is a virtual particle with $k^0 \neq \omega_k$. In fact, as we have shown for the first-order calculations, it would be impossible to guarantee four-momentum conservation at each vertex with on-shell particles. Note that, due to the denominator of the propagator, particles with higher virtuality contribute less to the amplitude.

In the time-ordered treatment, Eq. (4.214), all the contributions from all possible ways to order the vertices have to be taken into considerations. With n vertices, there are $n!$ possible time-ordered diagrams. In this example there are only two, depicted in Fig. 4.17. The left panel shows a process where an electron absorbs a photon and radiates a photon at a later time. The right panel shows a process where an electron, a positron, and a photon are produced from the vacuum and at a later time the positron annihilates by combining with an electron and a photon.

Finally, in the time-ordered treatment the intermediate state is an on-shell particle or antiparticle (both states contribute) with $k^0 = \omega_k$, but *energy is not conserved at each vertex*. There is a violation of energy-conservation, but only for the intermediate states. Note that, due to the denominator of the propagator, contributions with a larger violation of energy conservation are less relevant.

4.8.3 Example: Møller scattering

Let us see another example for illustration purposes, without repeating all the calculations. Let us consider the term $S_{(c5)}$ corresponding to Møller scattering, i.e.,

$$S^{(2)}_{(c5)} = -\frac{e^2}{2!}\int d^4x_1 d^4x_2 \left(\gamma^\alpha_{AB} i D_{F,\alpha\beta}(x_1 - x_2)\gamma^\beta_{CD}\right)\overline{\psi}^-_A(x_1)\overline{\psi}^-_C(x_2)\psi^+_B(x_1)\psi^+_D(x_2). \quad (4.215)$$

The initial and final states are

$$|i\rangle = c^\dagger_{s_i}(k_i)c^\dagger_{r_i}(p_i)|0\rangle, \qquad \langle f| = \langle 0|c_{s_f}(k_f)c_{r_f}(p_f). \quad (4.216)$$

When acting, for instance, with the $\psi^+(x_1)$ field, there are two possibilities: we can destroy the electron with momentum p_i or the one with momentum k_i. The same happens with the final-state electrons. There are in total four possibilities, but only two of them are distinct. Two of them are connected by the interchange of the integration variables x_1 and x_2 and simply combine together to get rid of the 2! denominator. The corresponding two Feynman diagrams are depicted in Fig. 4.18.

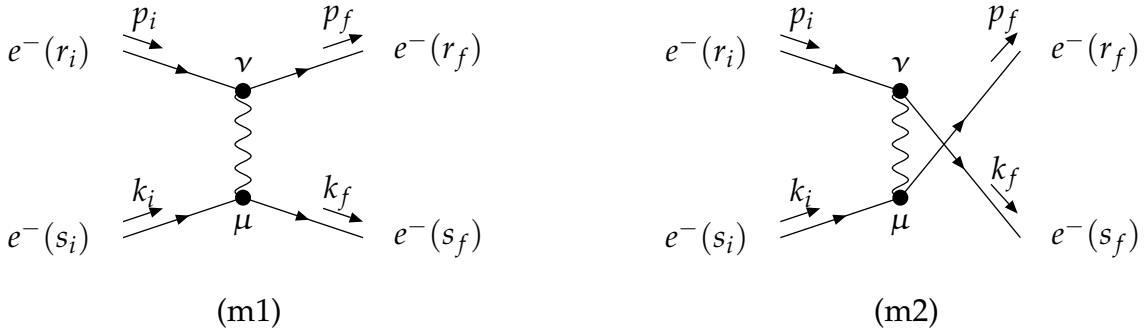


Figure 4.18: Feynman diagrams contributing to Møller scattering.

The two distinct possibilities lead to these final results

$$i\mathcal{M}_{(m1)} = -e^2\,\overline{u}_{s_f}(k_f)\gamma^\mu u_{s_i}(k_i)\,\frac{-ig_{\mu\nu}}{(q^2+i\epsilon)}\bigg|_{q=k_i-k_f}\overline{u}_{r_f}(p_f)\gamma^\nu u_{r_i}(p_i), \tag{4.217}$$

$$i\mathcal{M}_{(m2)} = +e^2\,\overline{u}_{r_f}(p_f)\gamma^\mu u_{s_i}(k_i)\,\frac{-ig_{\mu\nu}}{(q^2+i\epsilon)}\bigg|_{q=k_i-p_f}\overline{u}_{s_f}(k_f)\gamma^\nu u_{r_i}(p_i). \tag{4.218}$$

The *difference in sign* is very important: it is due to the switching of the two final-state fermion fields and it is an example of what was discussed at the end of the previous section (i.e., the fact that only the relative sign between two contributions to the same process matter, and it can be fixed by checking if there is an interchange of two fermion fields).

4.9 Cross section calculations

We resort to scattering experiments to test the behavior of our theory and the outcome of the calculations of the scattering matrix. Usually, two beams of particles are made to collide, or a beam is scattered off a target. The observable quantity is the number of events in the detector ($N_{\rm ev}$). This quantity depends however on the characteristics of the experiment: intensity of the beam, density of the target, relative velocity, efficiency, etc. These parameters define the so-called "luminosity" ($\mathcal{L}$) of the experiment. The luminosity can be formally defined as the number of incident particles ($N_{\rm inc}$) per unit surface (S) times the number of scattering centers ($N_{\rm cen}$) and has the dimensions of $[L]^{-2}$, or $[M]^2$ in

natural units.[4] The number of scattering events is related to the product of the luminosity times the cross section. The cross section can be defined as

$$\sigma = \frac{N_{\text{ev}}}{\mathcal{L}} \equiv \frac{N_{\text{ev}}}{N_{\text{inc}}\, N_{\text{cen}}/S} \tag{4.219}$$

therefore an observable quantity (number of events/luminosity), which is characteristic of the process under consideration and independent of the specific experiment where the process is measured.

A discussion of how the cross section is connected to the scattering matrix and the Feynman amplitude is presented, e.g., in Sec. 4.5 of Peskin–Schroeder, in Sec. 8.1 of Mandl–Shaw, in Sec. 4.3 of Halzen–Martin. Here, we consider only the outcome of the discussion.

Consider a process where two particles A and B scatter and produce a certain number of final-state particles n. The general formula expressing the relation between a differential cross section and the S matrix is

$$\boxed{d\sigma = \frac{1}{4\sqrt{(p_A \cdot p_B)^2 - m_A^2 m_B^2}}\, |\mathcal{M}|^2 \prod_f \frac{d^3 p_f}{(2\pi)^3 2E_f}\, (2\pi)^4 \delta^4\big(p_A + p_B - \textstyle\sum p_f\big).} \tag{4.220}$$

The denominator of the first term is often called the flux factor, since it is connected to the flux of particles impinging on a target. The formula is valid in any frame where the two initial particles are collinear.

The final part of the cross-section formula (from the product symbol) is called the Lorentz-invariant phase space (LIPS). Using four-momentum conservation we can simplify the phase-space factors. The starting formula is differential in the $3n$ components of the n final-state momenta. However, we have four conditions coming from the conservation of four-momentum, which means that in the end the cross section can be kept differential in only $3n - 4$ variables. The choice of which variables to consider depends on taste and, of course, on the need to compare to specific experimental measurements.

We will consider in our calculations only $2 \to 2$ processes. Therefore, we can keep the cross section differential in two variables. A common choice are the azimuthal and polar angles of one of the outgoing particles. We can therefore use the delta of conservation of momentum to fix the three-momentum of one of the outgoing particles and the modulus

[4] Normally, the luminosity is actually determined using a process whose cross-section is very well known and dividing the number of events by the cross section.

of the momentum of the other particle.

$$\begin{aligned}
&\frac{d^3p_C}{(2\pi)^3 2E_C}\frac{d^3p_D}{(2\pi)^3 2E_D}\,(2\pi)^4\delta^4\big(p_A+p_B-p_C-p_D\big)\\
&\qquad=\frac{d^3p_C}{(2\pi)^3 2E_C}\frac{d^3p_D}{(2\pi)^3 2E_D}\,(2\pi)^4\delta^3\big(\vec{p}_A+\vec{p}_B-\vec{p}_C-\vec{p}_D\big)\delta\big(E_A+E_B-E_C-E_D\big)\\
&\qquad=\frac{|\vec{p}_C|^2 d|\vec{p}_C|d\Omega}{(2\pi)^3 2E_C}\frac{1}{2E_D}\,(2\pi)\delta\big(E_A+E_B-E_C-E_D\big)\Big|_{\vec{p}_A+\vec{p}_B=\vec{p}_C+\vec{p}_D}\\
&\qquad=\frac{|\vec{p}_C|^2 d\Omega}{(2\pi)^2 2E_C 2E_D}\left|\frac{\partial(E_C+E_D)}{\partial|\vec{p}_C|}\right|^{-1}\Big|_{p_A+p_B=p_C+p_D}.
\end{aligned}\tag{4.221}$$

In the last step, we used the properties of the delta function to change from a delta function expressed in terms of E_C+E_D to an expression in terms of $|\vec{p}_C|$.

Center-of-mass frame

The two most commonly used frames of reference are the center-of-mass frame (CMF) and the target rest frame (TRF). In the former, we can write the momenta of the two initial particles as

$$p_A \overset{\text{CMF}}{=} \big(E_A,\ \ 0,\ \ 0,\ \ |\vec{p}_A|\big)\,, \qquad\qquad p_B \overset{\text{CMF}}{=} \big(E_B,\ \ 0,\ \ 0,\ \ -|\vec{p}_A|\big) \tag{4.222}$$

with $E_x=\sqrt{|\vec{p}_x|^2+m_x^2}$. Therefore

$$\begin{aligned}
(p_A\cdot p_B)^2-m_A^2m_B^2 &\overset{\text{CMF}}{=} E_A^2E_B^2+|\vec{p}_A|^4+2E_AE_B|\vec{p}_A|^2-m_A^2m_B^2\\
&=\big(|\vec{p}_A|^2+m_A^2\big)\big(|\vec{p}_A|^2+m_B^2\big)+|\vec{p}_A|^4+2E_AE_B|\vec{p}_A|^2-m_A^2m_B^2\\
&=|\vec{p}_A|^4+|\vec{p}_A|^2\big(m_A^2+m_B^2\big)+|\vec{p}_A|^4+2E_AE_B|\vec{p}_A|^2\\
&=|\vec{p}_A|^2\big(2|\vec{p}_A|^2+m_A^2+m_B^2+2E_AE_B\big)\\
&=|\vec{p}_A|^2\big(E_A^2+E_B^2+2E_AE_B\big)=|\vec{p}_A|^2\big(E_A+E_B\big)^2.
\end{aligned}\tag{4.223}$$

The flux factor can be written of course in terms of other variables.

For what concerns the phase space, we need to write

$$p_C \overset{\text{CMF}}{=} \big(E_C,\ \ |\vec{p}_C|\sin\theta\cos\phi,\ \ |\vec{p}_C|\sin\theta\sin\phi,\ \ |\vec{p}_C|\cos\theta\big)\,, \tag{4.224}$$

$$p_D \overset{\text{CMF}}{=} \big(E_D,\ \ -|\vec{p}_C|\sin\theta\cos\phi,\ \ -|\vec{p}_C|\sin\theta\sin\phi,\ \ -|\vec{p}_C|\cos\theta\big)\,, \tag{4.225}$$

with $E_C=\sqrt{|\vec{p}_C|^2+m_C^2}$ and similarly for E_D. We obtain

$$\frac{\partial(E_C+E_D)}{\partial|\vec{p}_C|}=\frac{\partial}{\partial|\vec{p}_C|}\Big(\sqrt{|\vec{p}_C|^2+m_C^2}+\sqrt{|\vec{p}_C|^2+m_D^2}\Big)=|\vec{p}_C|\Big(\frac{1}{E_C}+\frac{1}{E_D}\Big). \tag{4.226}$$

Putting everything together we arrive at the following compact formula for a $AB \to CD$ scattering in the CMF

$$\begin{aligned} d\sigma &= \frac{1}{4|\vec{p}_A|\left(E_A+E_B\right)}\ |\mathcal{M}|^2\ \frac{|\vec{p}_C|^2 d\Omega}{(2\pi)^2 2E_C 2E_D}\frac{1}{|\vec{p}_C|}\ \frac{E_C E_D}{E_C+E_D} \\ &= \frac{d\Omega}{64\pi^2\left(E_A+E_B\right)^2}\ \frac{|\vec{p}_C|}{|\vec{p}_A|}\ |\mathcal{M}|^2. \end{aligned} \tag{4.227}$$

Particularly simple is the case when all four particles have the same mass (including when all masses are neglected). Under these conditions, we have $E_a = E_B \equiv E$ and $|\vec{p}_C| = |\vec{p}_A|$. The formula reduces to

$$\boxed{d\sigma = \frac{d\Omega}{64\pi^2 4E^2}\ |\mathcal{M}|^2 \qquad \text{if } m_a = m_b = m_c = m_d.} \tag{4.228}$$

Target rest-frame

Let us analyze the formula in the target rest frame. For the calculation of the flux factor, we have

$$p_A \overset{\text{TRF}}{=} \left(E_A,\ \ 0,\ \ 0,\ \ |\vec{p}|_A\right), \quad p_B \overset{\text{TRF}}{=} \left(m_B,\ \ 0,\ \ 0,\ \ 0\right) \quad \left(E_A = \sqrt{|\vec{p}|_A^2 + m_A^2}\right) \tag{4.229}$$

(although we use the same notation, E_A and $\vec{p}_A$ are obviously different in the CMF and in the TRF). Therefore

$$(p_A \cdot p_B)^2 - m_A^2 m_B^2 \overset{\text{TRF}}{=} E_A^2 m_B^2 - m_A^2 m_B^2 = |\vec{p}|^2 m_B^2. \tag{4.230}$$

For the calculation of the phase space, for convenience let us neglect m_A and m_C and set $m_B = m_D$. We have

$$p_C \overset{\text{TRF}}{=} \left(|\vec{p}_C|,\ \ |\vec{p}_C|\sin\theta\cos\phi,\ \ |\vec{p}_C|\sin\theta\sin\phi,\ \ |\vec{p}_C|\cos\theta\right), \tag{4.231}$$

$$p_D \overset{\text{TRF}}{=} \left(E_D,\ \ |\vec{p}_D|\sin\theta_D\cos\phi,\ \ |\vec{p}_D|\sin\theta_D\sin\phi,\ \ |\vec{p}_D|\cos\theta_D\right), \tag{4.232}$$

with $E_D = \sqrt{|\vec{p}_D|^2 + m_D^2}$. Conservation of three-momentum requires

$$|\vec{p}_D|\sin\theta_D = -|\vec{p}_C|\sin\theta, \qquad |\vec{p}_D|\cos\theta_D = E_A - |\vec{p}_C|\cos\theta. \tag{4.233}$$

Squaring and summing the above:

$$|\vec{p}_D|^2 = |\vec{p}_C|^2 + E_A^2 - 2E_A|\vec{p}_C|\cos\theta \tag{4.234}$$

from which

$$E_D = \sqrt{|\vec{p}_C|^2 + E_A^2 - 2E_A|\vec{p}_C|\cos\theta + m_D^2}. \tag{4.235}$$

We can then write

$$\begin{aligned}\frac{\partial(E_C+E_D)}{\partial|\vec{p}_C|} &= \frac{\partial}{\partial|\vec{p}_C|}\Big(|\vec{p}_C| + \sqrt{|\vec{p}_C|^2+E_A^2-2E_A|\vec{p}_C|\cos\theta+m_D^2}\Big) \\ &= 1+\frac{|\vec{p}_C|-E_A\cos\theta}{\sqrt{|\vec{p}_C|^2+E_A^2-2E_A|\vec{p}_C|\cos\theta+m_D^2}} = 1+\frac{|\vec{p}_C|-E_A\cos\theta}{E_D}.\end{aligned} \tag{4.236}$$

Conservation of energy implies (it takes some steps to prove it)

$$E_A+m_B-|\vec{p}_C| = E_D \qquad \Rightarrow \qquad 1+\frac{|\vec{p}_C|-E_A\cos\theta}{E_D} = \frac{m_B E_A}{|\vec{p}_C|E_D}. \tag{4.237}$$

In conclusion, a compact formula for a $AB\to CD$ scattering in the TRF, neglecting the mass of particles A and C is (keeping in mind that $E_A=|\vec{p}_A|$ and $E_C=|\vec{p}_C|$)

$$\begin{aligned}d\sigma &= \frac{1}{4|\vec{p}_A|m_B}\,|\mathcal{M}|^2\,\frac{|\vec{p}_C|^2 d\Omega}{(2\pi)^2 2E_C 2E_D}\frac{|\vec{p}_C|E_D}{m_B E_A} \\ &= \frac{d\Omega}{64\pi^2 m_B^2}\,\frac{E_C^2}{E_A^2}\,|\mathcal{M}|^2.\end{aligned} \tag{4.238}$$

4.9.1 Mandelstam variables

In $AB\to CD$ processes, it is often useful to introduce the Lorentz invariants

$$s=(p_A+p_B)^2=(p_C+p_D)^2=m_A^2+m_B^2+2p_A\cdot p_B=m_C^2+m_D^2+2p_C\cdot p_D, \tag{4.239}$$

$$t=(p_A-p_C)^2=(p_B-p_D)^2=m_A^2+m_C^2-2p_A\cdot p_C=m_B^2+m_D^2-2p_B\cdot p_D, \tag{4.240}$$

$$u=(p_A-p_D)^2=(p_B-p_C)^2=m_A^2+m_D^2-2p_A\cdot p_D=m_B^2+m_C^2-2p_B\cdot p_C. \tag{4.241}$$

The Mandelstam variables are not independent

$$\begin{aligned}s+t+u &= 3m_A^2+m_B^2+m_C^2+m_D^2+2p_A\cdot(p_B-p_C-p_D) \\ &= 3m_A^2+m_B^2+m_C^2+m_D^2-2p_A\cdot p_A = m_A^2+m_B^2+m_C^2+m_D^2.\end{aligned} \tag{4.242}$$

It is useful to express Mandelstam variables in terms of CMF observables, especially when $m_A=m_B$ and $m_C=m_D$ respectively. Using Eqs. (4.222), (4.224), (4.225) we obtain

$$s=(E_A+E_B)^2=(E_C+E_D)^2=4E^2, \tag{4.243}$$

$$\begin{aligned}t &= -|\vec{p}_A|^2-|\vec{p}_C|^2+2|\vec{p}_A||\vec{p}_C|\cos\theta \\ &= -2E^2\left(1-\sqrt{1-\frac{m_A^2}{E^2}}\sqrt{1-\frac{m_C^2}{E^2}}\cos\theta-\frac{m_A^2+m_B^2}{2E^2}\right),\end{aligned} \tag{4.244}$$

$$u=-2E^2\left(1+\sqrt{1-\frac{m_A^2}{E^2}}\sqrt{1-\frac{m_C^2}{E^2}}\cos\theta-\frac{m_A^2+m_B^2}{2E^2}\right). \tag{4.245}$$

Particularly useful for our purposes is the case when all particles are massless (ultra-relativistic approximation), for which we have

$$s = 2p_A \cdot p_B = 2p_C \cdot p_D = 4E^2, \tag{4.246}$$

$$t = -2p_A \cdot p_C = -2p_B \cdot p_D = -2E^2(1 - \cos\theta), \tag{4.247}$$

$$u = -2p_A \cdot p_D = -2p_B \cdot p_C = -2E^2(1 + \cos\theta), \tag{4.248}$$

$$s + t + u = 0. \tag{4.249}$$

4.10 Examples of cross section calculations

Detailed calculations of the cross sections for the basic elementary processes of QED are presented in the lecture notes of prof. F. Piccinini (available in the library), in Ch. 8 of Mandl–Shaw, in Ch. 8 of Aitchison–Hey, Ch. 5 of Peskin–Schroeder.

4.10.1 Example 1: $e^+e^- \to \mu^+\mu^-$

Fig. 4.19 shows the Feynman diagram for the process under consideration at order α. This is the *leading order* diagram for this process, because there is no contribution at order e (i.e., order $\sqrt{\alpha}$). This is also a tree-level diagram, because it contains no loops. Finally, this is also called a *s channel* diagram, because the momentum squared flowing into the intermediate propagator is equal to the Mandelstam variable s.

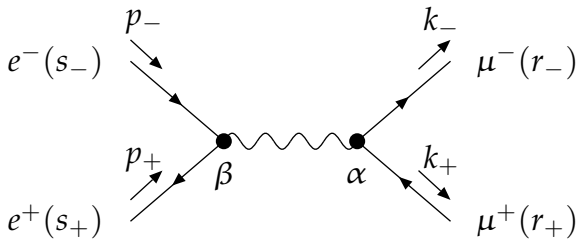


Figure 4.19: Feynman diagram for the process $e^+e^- \to \mu^+\mu^-$.

Using Feynman rules, we can immediately write the corresponding Feynman amplitude

$$i\mathcal{M} = e^2\, \overline{v}_{\mathrm{e},s_+}(p_+)\gamma_\beta u_{\mathrm{e},s_-}(p_-)\, \frac{i\, g^{\beta\alpha}}{(q^2 + i\epsilon)}\bigg|_{q=p_+ + p_-} \overline{u}_{\mathrm{m},r_-}(k_-)\gamma_\alpha v_{\mathrm{m},r_+}(k_+) \tag{4.250}$$

and its complex conjugate

$$-i\mathcal{M}^* = -e^2\, \overline{v}_{\mathrm{m},r_+}(k_+)\gamma_\rho u_{\mathrm{m},r_-}(k_-)\, \frac{i\, g^{\rho\sigma}}{(q^2 - i\epsilon)}\bigg|_{q=p_+ + p_-} \overline{u}_{\mathrm{e},s_-}(p_-)\gamma_\sigma v_{\mathrm{e},s_+}(p_+). \tag{4.251}$$

The modulus squared of the amplitude is (the dependence on spin and momenta is temporarily dropped for convenience)

$$\begin{aligned}|\mathcal{M}|^2 &= \frac{e^4}{s^2}\left(\overline{v}_{\mathrm{e}^+}\gamma_\alpha u_{\mathrm{e}^-}\right)\left(\overline{u}_{\mathrm{m}^-}\gamma^\alpha v_{\mathrm{m}^+}\right)\left(\overline{v}_{\mathrm{m}^+}\gamma^\sigma u_{\mathrm{m}^-}\right)\left(\overline{u}_{\mathrm{e}^-}\gamma_\sigma v_{\mathrm{e}^+}\right)\\ &= \frac{e^4}{s^2}\left(\overline{u}_{\mathrm{e}^-}\gamma_\sigma v_{\mathrm{e}^+}\right)\left(\overline{v}_{\mathrm{e}^+}\gamma_\alpha u_{\mathrm{e}^-}\right)\left(\overline{u}_{\mathrm{m}^-}\gamma^\alpha v_{\mathrm{m}^+}\right)\left(\overline{v}_{\mathrm{m}^+}\gamma^\sigma u_{\mathrm{m}^-}\right).\end{aligned} \tag{4.252}$$

In the last step, we made use of the fact that each term in parenthesis is a Dirac scalar and we are allowed to move it in a different position. Moreover, if we write explicitly the Dirac indices we are free to rearrange all terms and move the spinors in order to reconstruct structures like $u\overline{u}$ and $v\overline{v}$ and

$$\begin{aligned}|\mathcal{M}|^2 &= \frac{e^4}{s^2}\left(\overline{u}_{\mathrm{e}^-}\right)_E\left(\gamma_\sigma\right)_{EF}\left(v_{\mathrm{e}^+}\overline{v}_{\mathrm{e}^+}\right)_{FG}\left(\gamma_\alpha\right)_{GH}\left(u_{\mathrm{e}^-}\right)_H\\ &\quad\times\left(\overline{u}_{\mathrm{m}^-}\right)_A\left(\gamma^\alpha\right)_{AB}\left(v_{\mathrm{m}^+}\overline{v}_{\mathrm{m}^+}\right)_{BC}\left(\gamma^\sigma\right)_{CD}\left(u_{\mathrm{m}^-}\right)_D\\ &= \frac{e^4}{s^2}\left(\gamma_\sigma\right)_{EF}\left(v_{\mathrm{e}^+}\overline{v}_{\mathrm{e}^+}\right)_{FG}\left(\gamma_\alpha\right)_{GH}\left(u_{\mathrm{e}^-}\overline{u}_{\mathrm{e}^-}\right)_{HE}\\ &\quad\times\left(\gamma^\alpha\right)_{AB}\left(v_{\mathrm{m}^+}\overline{v}_{\mathrm{m}^+}\right)_{BC}\left(\gamma^\sigma\right)_{CD}\left(u_{\mathrm{m}^-}\overline{u}_{\mathrm{m}^-}\right)_{DA}\\ &= \frac{e^4}{s^2}\mathrm{Tr}\left[\gamma_\sigma\, v_{\mathrm{e}^+}\overline{v}_{\mathrm{e}^+}\gamma_\alpha\, u_{\mathrm{e}^-}\overline{u}_{\mathrm{e}^-}\right]\mathrm{Tr}\left[\gamma^\alpha\, v_{\mathrm{m}^+}\overline{v}_{\mathrm{m}^+}\,\gamma^\sigma\, u_{\mathrm{m}^-}\overline{u}_{\mathrm{m}^-}\right].\end{aligned} \tag{4.253}$$

Note that we reduced the expression to a contraction of two Lorentz tensors, one referring to the electron-positron side and one referring to the muon-antimuon side. Each one of these tensors, called *leptonic tensor*, is written as a trace of a product of Dirac matrices in the Dirac space of the electron and of the muon fields, respectively.

We choose not to consider the polarization of the external particles: we average over initial state and sum over final state polarization. We obtain

$$\begin{aligned}\frac{1}{4}\sum_{s_+,s_-,r_+,r_-}|\mathcal{M}|^2 &= \frac{e^4}{4s^2}\,\mathrm{Tr}\left[\gamma_\sigma\left(\sum_{s_+}v_{\mathrm{e}^+}\overline{v}_{\mathrm{e}^+}\right)\gamma_\alpha\left(\sum_{s_-}u_{\mathrm{e}^-}\overline{u}_{\mathrm{e}^-}\right)\right]\\ &\quad\times\mathrm{Tr}\left[\gamma^\alpha\left(\sum_{r_+}v_{\mathrm{m}^+}\overline{v}_{\mathrm{m}^+}\right)\gamma^\sigma\left(\sum_{r_-}u_{\mathrm{m}^-}\overline{u}_{\mathrm{m}^-}\right)\right]\\ &= \frac{e^4}{4s^2}\,\mathrm{Tr}\left[\gamma_\sigma\left(\not{p}_+ - m_{\mathrm{e}}\right)\gamma_\alpha\left(\not{p}_- + m_{\mathrm{e}}\right)\right]\\ &\quad\times\mathrm{Tr}\left[\gamma^\alpha\left(\not{k}_+ - m_{\mathrm{m}}\right)\gamma^\sigma\left(\not{k}_- + m_{\mathrm{m}}\right)\right].\end{aligned} \tag{4.254}$$

After the calculation of the traces, we obtain

$$\begin{aligned}\frac{1}{4}\sum_{\text{spin}}|\mathcal{M}|^2 &= \frac{e^4}{4s^2}\,4\Big[p_{-\sigma}p_{+\alpha}+p_{+\sigma}p_{-\alpha}-(p_-\cdot p_+)g_{\sigma\alpha}-m_e^2 g_{\sigma\alpha}\Big]\\ &\quad\times 4\Big[k_-^\sigma k_+^\alpha+k_+^\sigma k_-^\alpha-(k_-\cdot k_+)g^{\sigma\alpha}-m_m^2 g^{\sigma\alpha}\Big]\\ &= \frac{4e^4}{s^2}\Big[2(p_-\cdot k_+)(p_+\cdot k_-)+2(p_-\cdot k_-)(p_+\cdot k_+)+\dots\Big]\\ &= \frac{2e^4}{s^2}\Big[2(p_-\cdot k_+)2(p_+\cdot k_-)+2(p_-\cdot k_-)2(p_+\cdot k_+)+\dots\Big]\\ &= \frac{2e^4}{s^2}\Big[u^2+t^2+\dots\Big]\approx 32\pi^2\alpha^2\,\frac{u^2+t^2}{s^2}.\end{aligned} \tag{4.255}$$

From the second step, we neglected the masses of the electron and of the muon. This is a good approximation if s is much larger than m_m.

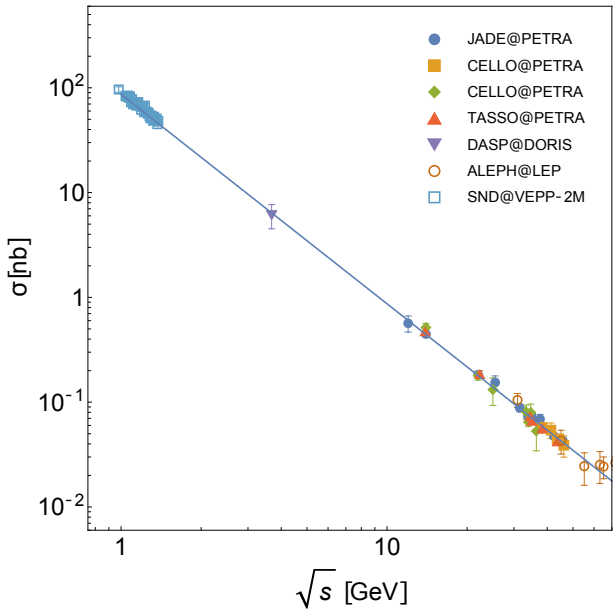


Figure 4.20: Experimental data for the total cross section of the process $e^+e^- \to \mu^+mu^-$ compared with Eq. (4.258). Data are available online at `https://www.hepdata.net/` *and are from the following experiments: DASP@DORIS [12], CELLO@PETRA [9, 10], JADE@PETRA [8], TASSO@PETRA [13], ALEPH@LEP [5], SND@VEPP-2M [2].*

To write the cross section, we express the Mandelstam invariants in terms of the CMS polar angle

$$\frac{u^2+t^2}{s^2}=\frac{1}{4}(1+\cos\theta)^2+\frac{1}{4}(1-\cos\theta)^2+=\frac{1}{2}(1+\cos^2\theta) \tag{4.256}$$

Replacing into Eq. (4.228) we obtain the differential cross section

$$\frac{d\sigma}{d\Omega}=\frac{\alpha^2}{4s}(1+\cos^2\theta). \tag{4.257}$$

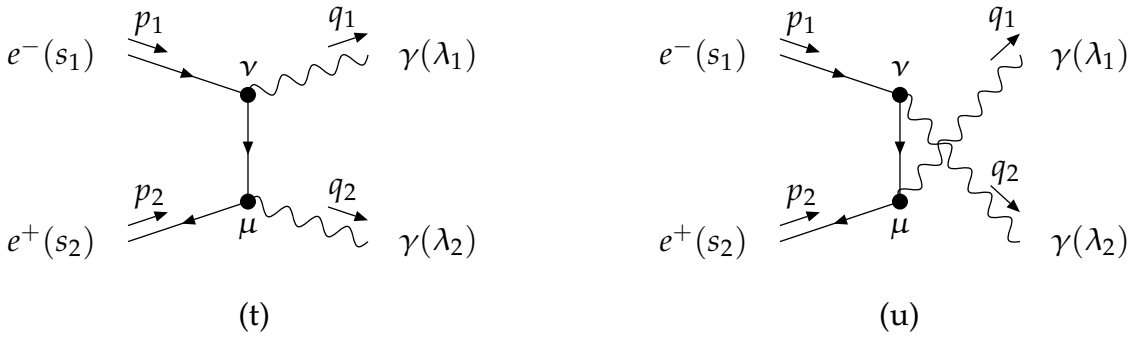


Figure 4.21: Feynman diagrams contributing to the process $e^-e^+ \to \gamma\gamma$.

This differential cross section can be integrated over $\cos\theta$ to give

$$\sigma = \frac{4\pi\alpha^2}{3s}. \tag{4.258}$$

The result for the cross section can be compared with experimental measurements from colliders operating at different values of the C.M.F. energy, $\sqrt{s}$, shown in Fig. 4.20. The agreement is excellent.

4.10.2 Example 2: $e^+e^- \to \gamma\gamma$

Here we consider an example that is not included in Piccinini's notes: the process $e^+e^- \to \gamma\gamma$. There are two "channels" for this process. The relevant Feynman diagrams are depicted in Fig. 4.21. The Feynman amplitudes for the two channels are

$$i\mathcal{M}_{(t)} = -e^2\, \overline{v}_{s_2}(p_2)\, \gamma_\mu \epsilon^{*\mu}_{\lambda_2}(q_2)\, \frac{i\,(\not{k}+m)}{(k^2-m^2+i\epsilon)}\bigg|_{k=p_1-q_1} \gamma_\nu \epsilon^{*\nu}_{\lambda_1}(q_1)\, u_{s_1}(p_1), \tag{4.259}$$

$$i\mathcal{M}_{(u)} = -e^2\, \overline{v}_{s_2}(p_2)\, \gamma_\mu \epsilon^{*\mu}_{\lambda_1}(q_1)\, \frac{i\,(\not{k}+m)}{(k^2-m^2+i\epsilon)}\bigg|_{k=p_1-q_2} \gamma_\nu \epsilon^{*\nu}_{\lambda_2}(q_2)\, u_{s_1}(p_1). \tag{4.260}$$

The first diagram is called a t-channel diagram, since the momentum k flowing in the propagator fulfills $k^2 = t$. The second diagram is called a u-channel diagram, since in this case $k^2 = u$. The two amplitudes have the same sign, because they are connected by an exchange of bosonic fields.

The complex-conjugate amplitudes are

$$-i\mathcal{M}^*_{(t)} = e^2\, \overline{u}_{s_1}(p_1)\, \not{\epsilon}_{\lambda_1}(q_1)\, \frac{i\,(\not{k}+m)}{(k^2-m^2-i\epsilon)}\bigg|_{k=p_1-q_1} \not{\epsilon}_{\lambda_2}(q_2)\, v_{s_2}(p_2), \tag{4.261}$$

$$-i\mathcal{M}^*_{(u)} = e^2\, \overline{u}_{s_1}(p_1)\, \not{\epsilon}_{\lambda_2}(q_2)\, \frac{i\,(\not{k}+m)}{(k^2-m^2-i\epsilon)}\bigg|_{k=p_1-q_2} \not{\epsilon}_{\lambda_1}(q_1)\, v_{s_2}(p_2). \tag{4.262}$$

We do not consider the polarization of the external particles: we average over initial

state and sum over final state polarization. We obtain

$$\frac{1}{4}\sum_{s_1,s_2,\lambda_1,\lambda_2}|\mathcal{M}|^2=\frac{1}{4}\sum_{s_1,s_2,\lambda_1,\lambda_2}\Big(|\mathcal{M}_t|^2+\mathcal{M}_t\mathcal{M}_u^*+\mathcal{M}_u\mathcal{M}_t^*+|\mathcal{M}_u|^2\Big). \tag{4.263}$$

We start analyzing the first term. First of all, we turn the product of Dirac spinors and matrices into a trace of Dirac matrices, similarly to what we have done in Eq. (4.68). To follow this step, we consider the Dirac structure of the scattering amplitude and write explicitly the Dirac indices

$$\begin{aligned}
&\overline{v}_{s_2A}(p_2)\Big(\not{\epsilon}^*_{\lambda_2}(q_2)\;(\not{p}_1-\not{q}_1+m)\;\not{\epsilon}^*_{\lambda_1}(q_1)\Big)_{AB}\,u_{s_1B}(p_1)\\
&\qquad\times\overline{u}_{s_1C}(p_1)\Big(\not{\epsilon}_{\lambda_1}(q_1)\;(\not{p}_1-\not{q}_1+m)\;\not{\epsilon}_{\lambda_2}(q_2)\Big)_{CD}\,v_{s_2D}(p_2)\\
&=v_{s_2D}(p_2)\overline{v}_{s_2A}(p_2)\Big(\not{\epsilon}^*_{\lambda_2}(q_2)\;(\not{p}_1-\not{q}_1+m)\;\not{\epsilon}^*_{\lambda_1}(q_1)\Big)_{AB}\,u_{s_1B}(p_1)\\
&\qquad\times\overline{u}_{s_1C}(p_1)\Big(\not{\epsilon}_{\lambda_1}(q_1)\;(\not{p}_1-\not{q}_1+m)\;\not{\epsilon}_{\lambda_2}(q_2)\Big)_{CD}\\
&=\mathrm{Tr}\Big[v_{s_2}(p_2)\overline{v}_{s_2}(p_2)\not{\epsilon}^*_{\lambda_2}(q_2)\;(\not{p}_1-\not{q}_1+m)\;\not{\epsilon}^*_{\lambda_1}(q_1)\,u_{s_1}(p_1)\\
&\qquad\times\overline{u}_{s_1}(p_1)\;\not{\epsilon}_{\lambda_1}(q_1)\;(\not{p}_1-\not{q}_1+m)\;\not{\epsilon}_{\lambda_2}(q_2)\Big]
\end{aligned} \tag{4.264}$$

Inserting this result into the t-channel squared amplitude we obtain

$$\begin{aligned}
\frac{1}{4}\sum_{s\ldots\lambda\ldots}|\mathcal{M}_t|^2&=\frac{e^4}{\big(t-m^2\big)^2}\frac{1}{4}\sum_{\lambda_1,\lambda_2}\mathrm{Tr}\Big[\sum_{s_2}v_{s_2}(p_2)\overline{v}_{s_2}(p_2)\not{\epsilon}^*_{\lambda_2}(q_2)\;(\not{p}_1-\not{q}_1+m)\\
&\qquad\times\not{\epsilon}^*_{\lambda_1}(q_1)\sum_{s_1}u_{s_1}(p_1)\overline{u}_{s_1}(p_1)\not{\epsilon}_{\lambda_1}(q_1)\;(\not{p}_1-\not{q}_1+m)\;\not{\epsilon}_{\lambda_2}(q_2)\Big]\\
&=\frac{e^4}{4\big(t-m^2\big)^2}\sum_{\lambda_2}\epsilon^{\mu*}_{\lambda_2}(q_2)\epsilon^{\beta}_{\lambda_2}(q_2)\sum_{\lambda_1}\epsilon^{\gamma*}_{\lambda_1}(q_1)\epsilon^{\alpha}_{\lambda_1}(q_1)\\
&\qquad\times\mathrm{Tr}\Big[(\not{p}_2-m)\gamma_\mu(\not{p}_1-\not{q}_1+m)\gamma_\nu(\not{p}_1+m)\gamma_\alpha(\not{p}_1-\not{q}_1+m)\gamma_\beta\Big].
\end{aligned} \tag{4.265}$$

In the second step, we made use of the definition of positive and negative energy projectors, Eqs. (2.255) and (2.256).[5] At this point, we have to worry about the summation over photon polarizations. In principle, we should consider here only physical polarization states of the photon, i.e., use the expression in Eq. (3.260). However, we can check that the addition of unphysical states is irrelevant in the sum $\mathcal{M}_{(t)}+\mathcal{M}_{(u)}$, with similar steps as discussed for Compton scattering in Sec. 4.8.1. This is another example of the application of Ward's identity. Because of this, we can replace

$$\boxed{\sum_\lambda\epsilon^{\mu*}_\lambda(q)\epsilon^{\alpha}_\lambda(q)\to-g^{\mu\alpha}.} \tag{4.266}$$

[5]Note that Piccinini and Mandl–Shaw have different conventions concerning the normalization of spinors, leading to a $1/2m$ factor in the projectors.

Using the above replacement, we can write

$$\begin{aligned}\frac{1}{4}\sum_{s\ldots\lambda\ldots}|\mathcal{M}_t|^2 &= \frac{e^4}{4\big(t-m^2\big)^2}\mathrm{Tr}\Big[(\not p_2-m)\gamma^\beta(\not p_1-\not q_1+m)\gamma_\alpha(\not p_1+m)\gamma^\alpha(\not p_1-\not q_1+m)\gamma_\beta\Big]\\ &= \frac{e^4}{4\big(t-m^2\big)^2}\mathrm{Tr}\Big[(-2\not p_2+4m)(\not p_1-\not q_1+m)(-2\not p_1+4m)(\not p_1-\not q_1+m)\Big].\end{aligned} \tag{4.267}$$

In the second step we made use of property (2.80). To simplify the calculations, from now on we neglect particle masses and obtain

$$\begin{aligned}\frac{1}{4}\sum_{s\ldots\lambda\ldots}|\mathcal{M}_t|^2 &\approx \frac{e^4}{t^2}\,\mathrm{Tr}\Big[\not p_2(\not p_1-\not q_1)\not p_1(\not p_1-\not q_1)\Big]\\ &= \frac{e^4}{t^2}4\Big[2\Big(p_2\cdot(p_1-q_1)\Big)\Big(p_1\cdot(p_1-q_1)\Big)-(p_2\cdot p_1)\Big((p_1-q_1)\cdot(p_1-q_1)\Big)\Big]\\ &= 2\frac{e^4}{t^2}\Big[-(2p_2\cdot p_1)(2p_1\cdot q_1)+(2p_2\cdot q_1)(2p_1\cdot q_1)-2p_2\cdot p_1\,(p_1-q_1)^2\Big]\Big]\\ &= 2\frac{e^4}{t^2}\big(st+ut-st\big)=2e^4\frac{u}{t}.\end{aligned} \tag{4.268}$$

Similar calculations yield (it is sufficient to replace $q_1\to q_2$ in the t-channel calculations)

$$\frac{1}{4}\sum_{s\ldots\lambda\ldots}|\mathcal{M}_u|^2\approx 2e^4\frac{t}{u}. \tag{4.269}$$

The interference terms are a bit more complex:

$$\begin{aligned}\frac{1}{4}\sum_{s\ldots\lambda\ldots}\mathcal{M}_t\mathcal{M}_u^* &= e^4\frac{1}{4}\sum_{s\ldots\lambda\ldots}\overline{v}_{s_2}(p_2)\not\epsilon^*_{\lambda_2}(q_2)\,\frac{(\not p_1-\not q_1+m)}{\big(t-m^2\big)}\,\not\epsilon^*_{\lambda_1}(q_1)\,u_{s_1}(p_1)\overline{u}_{s_1}(p_1)\\ &\quad\times\not\epsilon_{\lambda_2}(q_2)\,\frac{(\not p_1-\not q_2+m)}{\big(u-m^2\big)}\,\not\epsilon_{\lambda_1}(q_1)\,v_{s_2}(p_2)\\ &= \frac{e^4}{4\big(t-m^2\big)\big(u-m^2\big)}\\ &\quad\times\mathrm{Tr}\Big[(\not p_2-m)\gamma^\alpha(\not p_1-\not q_1+m)\gamma^\beta(\not p_1+m)\gamma_\alpha(\not p_1-\not q_2+m)\gamma_\beta\Big]\end{aligned} \tag{4.270}$$

To simplify the trace, we make use of Eqs. (2.82) followed by (2.81). Neglecting masses

from now on

$$\begin{aligned}\frac{1}{4}\sum_{s\ldots\lambda\ldots}\mathcal{M}_t\mathcal{M}_u^* &\approx \frac{e^4}{4tu}\mathrm{Tr}\Big[-2\not{p}_2\not{p}_1\gamma^\beta(\not{p}_1-\not{q}_1)(\not{p}_1-\not{q}_2)\gamma_\beta\Big]\\ &= -\frac{e^4}{tu}\,8(p_1\cdot p_2)\big((p_1-q_1)\cdot(p_1-q_2)\big)\\ &= -\frac{e^4}{tu}\,2(2p_1\cdot p_2)\big(-(2p_1\cdot q_2)-(2p_1\cdot q_1)+(2q_1\cdot q_2)\big)\\ &= -2\frac{e^4}{tu}\,s(u+t+s)=0.\end{aligned} \tag{4.271}$$

The final result for the squared scattering matrix is

$$\frac{1}{4}\sum_{s_1,s_2,\lambda_1,\lambda_2}|\mathcal{M}|^2 \approx 2e^4\left(\frac{u}{t}+\frac{t}{u}\right) = 32\pi^2\alpha^2\left(\frac{u}{t}+\frac{t}{u}\right). \tag{4.272}$$

Writing Mandelstam variables in terms of CMS angles and replacing into Eq. (4.228) we obtain the differential cross section

$$\boxed{\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E^2}\left(\frac{1+\cos^2\theta}{\sin^2\theta}\right) = \frac{\alpha^2}{s}\left(\frac{1+\cos^2\theta}{\sin^2\theta}\right).} \tag{4.273}$$

We can test the validity of this (lowest order) QED calculation against experimental data. Fig. 4.22 shows results from four experiments at different values of s. The agreement is very good ($\chi^2/\text{d.o.f.} = 1.2$, without taking into account systematic errors).

4.11 Conclusions

In this chapter we have considered interactions between different fields. We reviewed the concept of a scattering matrix and its expansion in terms of the interaction coupling constant. We have seen how to compute the scattering matrix at the lowest order in the coupling constant and how to relate it to cross sections, which are experimentally measurable.

Typical questions that can come out during the exam:

1. Derive and discuss the QED Lagrangian;
2. Derive Wick's theorem for a product of two field operators;
3. Compute any of the Feynman propagators discussed in the lectures;
4. Obtain an expression of the scattering matrix for QED at order e for some initial or final state and discuss the relation with Feynman rules;
5. Obtain an expression of the scattering matrix for QED at order e^2 for some initial or final state and discuss the relation with Feynman rules.

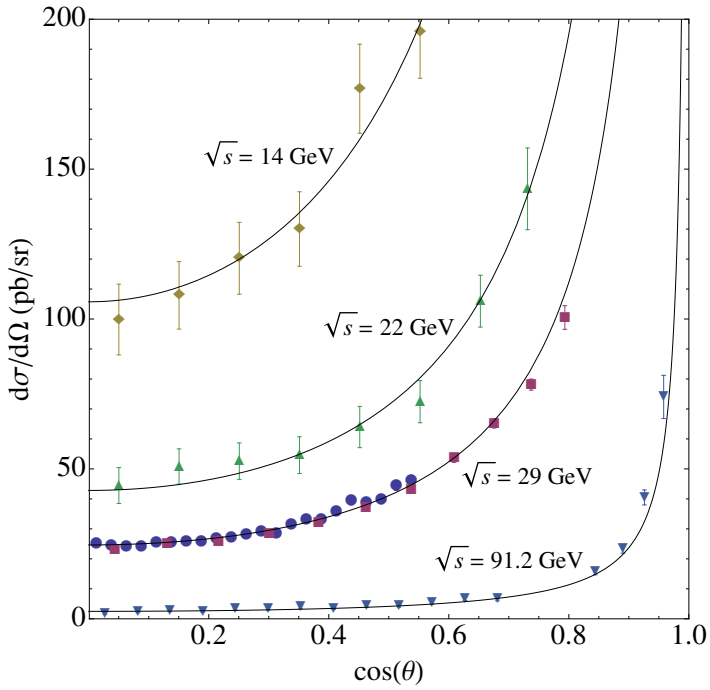


Figure 4.22: Experimental data for the process $e^+e^- \to \gamma\gamma$ compared with Eq. (4.273)*. Data are available online at* `https://www.hepdata.net/` *and are from the following experiments: JADE@PETRA [7], HSR@SLAC [14], MAC@SLAC [19], L3@LEP [1].*

Bibliography

[1] L3 Collaboration, M. Acciarri *et al.*, "Tests of QED at LEP energies using $e^+e^- \to \gamma\gamma(\gamma)$ and $e^+e^- \to l^+l^-\gamma\gamma$," Phys.Lett. **B353** (1995) 136–144.

[2] M. Achasov *et al.*, "Study of the process e+e-\to \mu+\mu- in the energy region \sqrt{s}=980, 1040 – 1380 MeV," Phys. Rev. D **79** (2009) 112012, [0810.0571].

[3] I. Aitchison and A. Hey, *Gauge Theories in Particle Physics: A Practical Introduction, Volume 1: From Relativistic Quantum Mechanics to QED, Fourth Edition*. Gauge Theories in Particle Physics: A Practical Introduction. CRC Press, 2012.

[4] T. Aoyama, M. Hayakawa, T. Kinoshita, and M. Nio, "Tenth-Order Electron Anomalous Magnetic Moment — Contribution of Diagrams without Closed Lepton Loops," Phys. Rev. D **91** (2015), no. 3, 033006, [1412.8284]. [Erratum: Phys.Rev.D 96, 019901 (2017)].

[5] ALEPH Collaboration, R. Barate *et al.*, "Study of the muon pair production at center-of-mass energies from 20-GeV to 136-GeV with the ALEPH detector," Phys. Lett. B **399** (1997) 329–341.

[6] V. Barone, *Relatività. Principi e applicazioni*. Bollati Boringhieri, 2004.

[7] JADE Collaboration, W. Bartel *et al.*, "Measurement of the Processes $e^+e^- \to e^+e^-$ and $e^+e^- \to \gamma\gamma$ at PETRA," Z.Phys. **C19** (1983) 197.

[8] JADE Collaboration, W. Bartel *et al.*, "New Results on $e^+e^- \to \mu^+\mu^-$ From the Jade Detector at {PETRA}," Z. Phys. C **26** (1985) 507.

[9] CELLO Collaboration, H. Behrend *et al.*, "Measurement of the Reaction $e^+e^- \to \mu^+\mu^-$ for 14-{GeV} $\leq \sqrt{s} \leq$ 36.4-{GeV}," Z. Phys. C **14** (1982) 283.

[10] CELLO Collaboration, H. Behrend *et al.*, "A Measurement of the Muon Pair Production in e^+e^- Annihilation at 38.3-{GeV} $\leq \sqrt{s} \leq$ 46.8-{GeV}," Phys. Lett. B **191** (1987) 209–216.

[11] M. Born, W. Heisenberg, and P. Jordan, "On Quantum Mechanics II," Z. Phys. **35** (1926) 557.

[12] DASP Collaboration, R. Brandelik *et al.*, "RESULTS FROM DASP ON E+ E- ANNIHILATION BETWEEN 3.1-GEV AND 5.2-GEV," Z. Phys. C **1** (1979) 233–256.

[13] TASSO Collaboration, W. Braunschweig *et al.*, "A Measurement of Muon Pair Production in e^+e^- Annihilation at Center-of-mass Energies 35-{GeV} $\leq \sqrt{s} \leq$ 46.8-{GeV}," Z. Phys. C **40** (1988) 163–170. [Erratum: Z.Phys.C 42, 348 (1989)].

[14] M. Derrick, K. Gan, P. Kooijman, J. Loos, B. Musgrave, *et al.*, "Experimental Study of the Reactions $e^+e^- \to e^+e^-$ and $e^+e^- \to \gamma\gamma$ at 29-GeV," Phys.Rev. **D34** (1986) 3286.

[15] P. A. M. Dirac, "The Quantum Theory of the Emission and Absorption of Radiation," Proceedings of the Royal Society of London **A 114** (1927) 243.

[16] P. A. M. Dirac, "The Quantum Theory of the Electron," Proceedings of the Royal Society of London **A 117** (1928) 610.

[17] P. A. M. Dirac, "A Theory of Electrons and Protons," Proceedings of the Royal Society of London **126** (1930), no. 801, pp. 360–365.

[18] P. A. Dirac, "Forms of Relativistic Dynamics," Rev.Mod.Phys. **21** (1949) 392–399.

[19] E. Fernandez, W. T. Ford, N. Qi, A. L. Read, J. G. Smith, *et al.*, "Tests of Quantum Electrodynamics With 2, 3 and 4 Photon Final States From e^+e^- Annihilation at $\sqrt{s} =$ 29-GeV," Phys.Rev. **D35** (1987) 1.

[20] H. Goldstein, C. Poole, J. Safko, and N. Cocca, *Meccanica classica*. Zanichelli, 2005.

[21] F. Halzen and A. Martin, *Quarks and leptons: an introductory course in modern particle physics*. Wiley, 1984.

[22] P. Jordan and O. Klein, "Zum Mehrkörperproblem in der Quantentheorie," Z. Phys. **45** (1927) 751.

[23] P. Jordan and E. Wigner, "Über das Paulische Äquivalenzverbot," Z. Phys. **47** (1928) 631.

[24] S. Laporta, "High-precision calculation of the 4-loop contribution to the electron g-2 in QED," Phys. Lett. B **772** (2017) 232–238, [1704.06996].

[25] M. Maggiore, *A Modern Introduction to Quantum Field Theory*. EBSCO ebook academic collection. Oxford University Press, 2005.

[26] F. Mandl and G. Shaw, *Quantum Field Theory*. John Wiley & Sons, 2010.

[27] A. Pais, *Inward Bound: Of Matter and Forces in the Physical World*. Oxford paperbacks. Clarendon Press, 1988.

[28] M. Peskin and D. Schroeder, *An Introduction To Quantum Field Theory*. Frontiers in Physics. Westview Press, 1995.

[29] L. Ryder, *Quantum Field Theory*. Cambridge University Press, 1996.

[30] F. Schwabl, R. Hilton, and A. Lahee, *Advanced Quantum Mechanics*. Advanced Texts in Physics Series. Springer, 2008.

[31] M. Schwartz, *Quantum Field Theory and the Standard Model*. Quantum Field Theory and the Standard Model. Cambridge University Press, 2014.

[32] D. Soper, "Quantum Field Theory." `http://physics.uoregon.edu/~soper/QFT/`, 2000.

[33] S. Weinberg, *The Quantum Theory of Fields. Vol. 1.* Cambridge University Press, 1996.

[34] G. C. Wick, "The evalutation of the collision matrix," Phys.Rev. **80** (1950) 268–272.

[35] Particle Data Group Collaboration, P. Zyla *et al.*, "Review of Particle Physics," PTEP **2020** (2020), no. 8, 083C01.