Lecture Notes on Quantum Field Theory I

Pilar Hernández and David Albandea

These lectures notes are based on the material covered in the course on Quantum Field Theory I in the Master in Advanced Physics at the University of Valencia, delivered in the years 2017-2021. The following references have been important sources in the preparation of these lectures:

- S. Weinberg, The Quantum Theory of Fields I, Cambridge University Press.
- M. E. Peskin and D.V. Schroeder, An Introduction to Quantum Field Theory, Taylor & Francis Group.
- C. Itzykson and J-B Zuber, Quantum field Theory, McGraw-Hill International Editions Physics Series.
- D. Tong, Lectures on Quantum Field Theory, http://www.damtp.cam.ac.uk/user/tong/qft.html.
- S. Coleman, Notes from Sydney Coleman's Physics 253a, arXiv:1110.5013.
- M. D. Schwartz, Quantum Field Theory and the Standard Model, Cambridge University Press.

I. Introduction: the need for a quantum theory of fields

Quantum field theory (QFT) is the formalism that unifies quantum mechanics (QM) and special relativity (SR), and enables us to understand what matter is made of.

Various types of experiments are used to explore the fundamental constituents of matter. In *fixed-target experiments* a beam of particles is focused towards a piece of material (target), see Fig. 1. Particles in the beam can be deflected, and even change in number and/or nature if the energy is high enough. From the characterization of the scattered particles and their angular distribution, we can infer the properties of the interactions of the beam particles with the target ones. These experiments mimic in the laboratory what happens when cosmic rays (protons or nuclei) hit the atmosphere (target) and produce showers as depicted in Fig. 2.

The birth of particle physics can be traced back to the study of the structure of the atom by Rutherford and others in fixed-target experiments, and to the discovery of new particles in cosmic rays. These experiments enabled us to establish the substructure of the atom, the atomic nucleus, and its constituents (i.e. the proton and the neutron). Furthermore they revealed the existence of many exotic and short-lived particles beyond those present in ordinary matter.

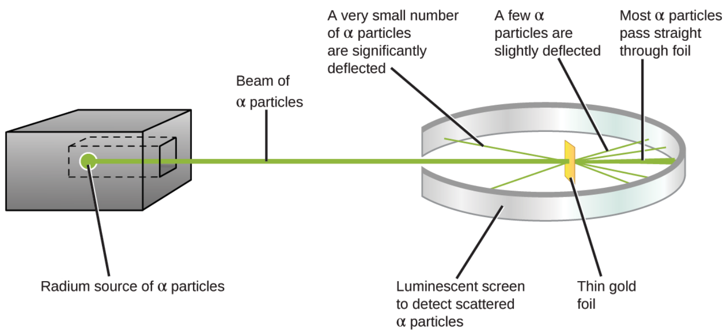


Figure 1: Classic experiment by Lord Rutherford that demonstrated the structure of the atom.

In *collider experiments* (eg. LHC or LEP) two beams of *simple* particles are accelerated in opposite directions and collide in the center of a detector. These experiments are easier to interpret because the complex substructure of the target disappears: we smash *simple* particles into *simple* particles. They are also more

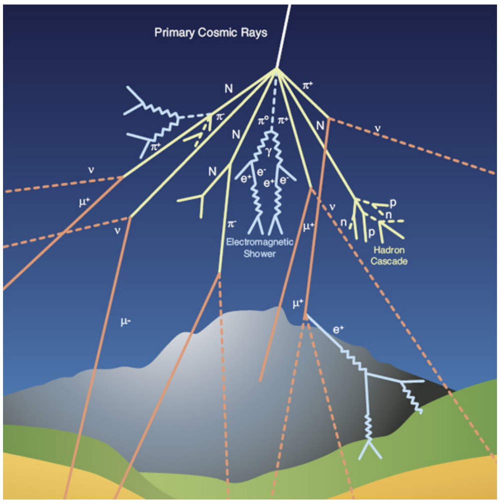


Figure 2: Cosmic Rays

efficient in reaching higher energies.

The type of measurements we can do, and would like to predict accurately, are:

- The nature of the scattered particles
- The number of particles per unit incoming flux that get deflected as a function of the scattering angle, θ.

The theoretical prediction of these observables requires that we understand quantum mechanics in the relativistic limit. QFT achieves this goal and allows us to understand and predict the dynamics of particles in the subatomic domain.

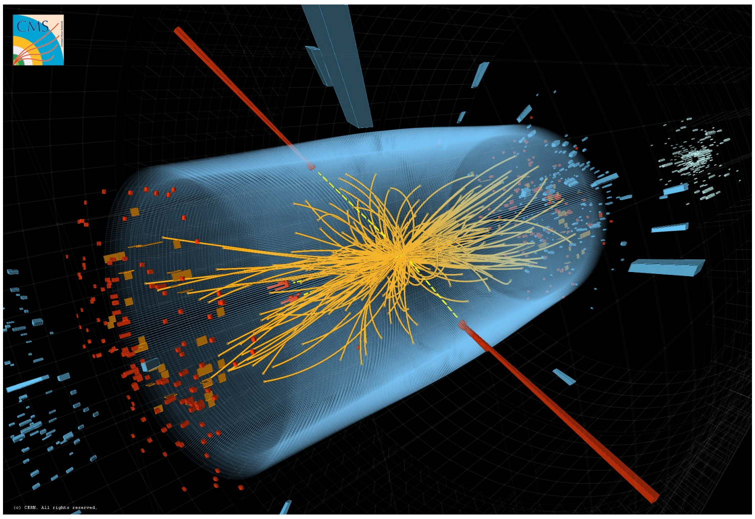


Figure 3: LHC proton-proton collision

1.1 Duality in quantum mechanics

As we know in quantum mechanics (QM) there is a duality between particles (e.g. an electron) and waves (e.g. the electromagnetic field). Particles of momentum p behave as de Broglie waves with wavelength

$$\lambda = \frac{h}{p}. \tag{1.2}$$

Similarly electromagnetic waves of frequency ν behave as bunches of quanta in the photoelectric effect or in Compton scattering. These are Einstein's photons, γ, that have energies given by

$$E = h\nu. \tag{1.3}$$

However a full unification of both concepts, particles and waves, is not achieved in QM, where the treatment of photons is completely different to that of electrons. Particles are described in terms of wave functions that satisfy the Schrödinger equation, but the electromagnetic field is usually treated classically.

In processes such as the photoelectric effect, the electromagnetic field is a classical wave that induces a time-dependent potential leading to transitions of electrons between atomic energy levels.

Instead, in Compton scattering, the shift in wavelength of light scattered at an angle θ on electrons at rest

$$\Delta\lambda = \frac{h}{m_e c}(1 - \cos\theta), \tag{1.4}$$

can be easily derived from simple particle kinematics assuming elastic scattering of an electron and a photon: $e + \gamma \to e + \gamma$, which clearly assumes a particle interpretation of the photon involved. However QM does not provide a method to quantify the probability of this process, which would require computing a transition amplitude between two states of one electron and one photon.

The same situation occurs in the case of spontaneous emission in which an electron in an excited atomic state decays to the ground state emitting a photon of a fixed frequency in the absence of any perturbation. In contrast with the photoelectric effect, there is no classical electromagnetic radiation present. The kinematics of the process can be easily understood with the same particle picture of the photon, but the probability of this process, which is related to the lifetime of an atomic state, requires a mathematical formulation of the photon as a particle, which is only achieved by the quantization of the electromagnetic field.

1.2 Quantum mechanics of relativistic particles

From a different perspective, the successful quantum treatment of particles in the non-relativistic regime does not extrapolate straightforwardly to the relativistic domain either. Several paradoxes demonstrate a clash with causality when we try to localize particles of mass m within distances smaller than their Compton wavelength

$$\lambda_{\mathrm{C}} = \frac{h}{mc}. \tag{1.5}$$

The origin of these paradoxes lies in the fact that the number of particles is not conserved in this regime. Let us suppose we manage to localize a particle within a distance of the order of half of the reduced Compton wavelength:

$$\Delta x \leq \frac{\lambda_{\mathrm{C}}}{4\pi}. \tag{1.6}$$

The uncertainty principle implies

$$\Delta x \Delta p \geq \frac{\hbar}{2} \to \Delta p \geq \frac{\hbar}{2\Delta x} \geq \frac{h}{\lambda_{\mathrm{C}}} \geq mc, \tag{1.7}$$

for a relativistic particle $E \simeq pc$ and therefore

$$\Delta E = c\Delta p \geq mc^2, \tag{1.8}$$

which is the rest energy of a particle, and therefore particles can be created in the process.

Surprisingly the solution to these paradoxes brings us to the same point as the need for a particle picture, or quantization, of the electromagnetic field, i.e.

a theory of quantum fields. QFT is the mathematical framework that achieves a complete unification of the classical concepts of particles and fields in relativistic and local quantum fields (see Fig. 4) which represent operators that create/annihilate particles.

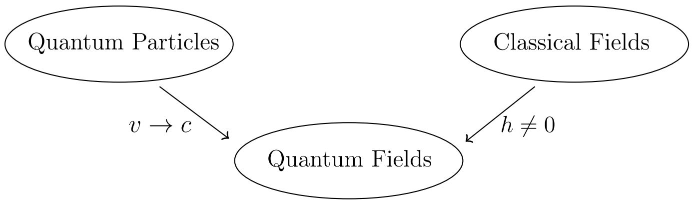


Figure 4: Quantum theory of fields

We will see that in QFT observables are not arbitrary Hermitian operators as in quantum mechanics: they must be constructed out of *local* quantum fields. Causality will not derive from the localization of particles but from the localization of interactions. We will see that quantum fields ensure relativistic invariance and an essential property of the laws of physics called cluster decomposition. This refers to the fact that measurements performed in well separated regions of space do not interfere, e.g. we can perform measurements locally in our lab, and they will not be affected by what is happening in the Moon.

QFT will deal with basic processes that we can measure in the lab. We would like to describe in a precise and quantitative way the two key ingredients in a particle collision:

- Asymptotic states: what are stable states of free particles?

  E.g. what are their masses, spin and charges

- Spectrum of unstable particles

  E.g. what are their masses, spin, charges and lifetime

- Scattering processes: what is the probability that a number of particles in some prepared incoming state arrive to an interaction region and "turns into" a different state of particles (same particle in a different state or even different particles)

We will see that QFT can answer these questions quantitatively and extremely accurately in some cases, up to the highest energy scales, we have explored so far, $E \sim 10^{13}$eV, that correspond to the smallest distances, $\sim 10^{-19}$m. There is little doubt that this abstract mathematical framework describes nature!

In the rest of this chapter we are going to introduce quantum fields as a way to solve apparent paradoxes that arise when trying to mix quantum physics and special relativity. We will start by briefly recalling some basic facts about both quantum mechanics and special relativity – and discuss why a naive extension of quantum mechanics to the relativistic domain fails, and why the new concept of quantum field is needed.

We are going to employ natural units throughout:

$$\hbar = c = 1. \tag{1.9}$$

In some cases, for aesthetical reasons some of these factors will be kept.

1.3 Non-relativistic QM in a nutshell

Quantum mechanics modifies in an essential way the classical concepts of the state of a system, of observable, and measurement.

• One-particle quantum states are vectors in a Hilbert space (the space of quadratically integrable complex functions, $L_2(C)$) that evolve in time

$$|\Psi(t)\ \rangle, \langle\Psi|\Psi\rangle = 1 \tag{1.10}$$

When expressed in the so-called complete position basis, $|\mathbf{x}\rangle$, the components of the state vector in this basis constitute the Schrödinger wave function:

$$\langle\mathbf{x}|\Psi(t)\rangle = \psi(\mathbf{x}, t). \tag{1.11}$$

The position basis is a complete basis, even if the basis states are not square-normalizable:

$$\int d\mathbf{x}\ |\mathbf{x}\rangle\langle\mathbf{x}| = 1, \quad \langle\mathbf{x}|\mathbf{x}'\rangle = \delta(\mathbf{x} - \mathbf{x}'). \tag{1.12}$$

Another important basis that is often used is the momentum basis, $|\mathbf{p}\rangle$:

$$\langle\mathbf{x}|\mathbf{p}\rangle = e^{i\mathbf{p}\cdot\mathbf{x}}. \tag{1.13}$$

$$\langle\mathbf{p}|\mathbf{p}'\rangle = \int d^3x \langle\mathbf{p}|\mathbf{x}\rangle\langle\mathbf{x}|\mathbf{p}'\rangle = \int d^3x\ e^{i(\mathbf{p}'-\mathbf{p})\cdot\mathbf{x}} = (2\pi)^3\delta(\mathbf{p}' - \mathbf{p}), \tag{1.14}$$

and the completeness condition[1] is

$$\int \frac{d^3p}{(2\pi)^3}|\mathbf{p}\rangle\langle\mathbf{p}| = 1. \tag{1.15}$$

[1] We could have equivalently normalized the states to eliminate the $(2\pi)^3$ in eqs. (1.14) and (1.15) and included a factor $(2\pi)^{-3/2}$ in eq. (1.13).

• Observables are Hermitian operators in Hilbert space

$$\hat{O}, \quad \langle \Phi | \hat{O} \Psi \rangle = \langle \hat{O} \Phi | \Psi \rangle. \tag{1.16}$$

As in finite vector spaces, these operators have real eigenvalues and the basis of eigenvectors forms a complete basis of physical states.

• The possible results of a measurement of an observable are exclusively the eigenvalues of the corresponding operator, and the probability of measuring one of these eigenvalues, e.g. o_n, when performing a measurement on a state $|\Psi\rangle$ is

$$\mathrm{Prob}(o_n) = |\langle o_n | \Psi \rangle|^2, \tag{1.17}$$

where $|o_n\rangle$ is the eigenvector corresponding to the eigenvalue o_n, ie.

$$\hat{O}|o_n\rangle = o_n |o_n\rangle. \tag{1.18}$$

Three observables/operators have a prominent role in NRQM: position, $\hat{X}$, momentum $\hat{P}_i$ and the energy or Hamiltonian, $\hat{H}$. Their eigenbasis are respectively, the position states, the momentum states and the stationary states:

$$\hat{X}|\mathbf{x}\rangle = \mathbf{x}|\mathbf{x}\rangle, \quad \hat{P}|\mathbf{p}\rangle = \mathbf{p}|\mathbf{p}\rangle, \; \hat{H}|E\rangle = E|E\rangle. \tag{1.19}$$

Position and momentum operators do not commute:

$$[\hat{X}_i, \hat{P}_j] = i\delta_{ij}, \tag{1.20}$$

which implies the Heisenberg uncertainty relation on any state:

$$\Delta X_i \Delta P_i \geq \frac{1}{2}, \tag{1.21}$$

where $\Delta O \equiv \langle \Psi | O^2 | \Psi \rangle - \langle \Psi | O | \Psi \rangle^2$.

• States evolve in time according to the Schrödinger equation, which in the position basis is a wave equation:

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{x}, t) = \left(-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{x}) \right) \psi(\mathbf{x}, t), \tag{1.22}$$

or in a basis-independent form:

$$i\hbar \frac{d}{dt} |\Psi\rangle = \hat{H} |\Psi\rangle, \tag{1.23}$$

where $\hat{H}$ is the quantum Hamiltonian operator, which for a non-relativistic particle of mass m in a potential $V(x)$ is:

$$\hat{H} = \frac{\hat{\mathbf{P}}^2}{2m} + V(\hat{X}). \tag{1.24}$$

It is therefore the classical Hamiltonian with the substitution

$$\mathbf{p} \to \hat{\mathbf{P}} = -i\hbar\nabla, \quad \mathbf{x} \to \hat{\mathbf{X}}. \tag{1.25}$$

The Schrödinger equation establishes that the time evolution of a quantum state is controlled by the Hamiltonian. It is easy to formally solve eq. (1.23) for Hamiltonians that do not depend explicitely on time:

$$|\Psi(t)\rangle = e^{-i\hat{H}(t-t_0)/\hbar}|\Psi(t_0)\rangle. \tag{1.26}$$

The quantum time evolution operator is therefore

$$U(t, t_0) \equiv e^{-i\hat{H}(t-t_0)/\hbar}. \tag{1.27}$$

• There are equivalent pictures in NRQM. We can equivalently say that states change with time and operators do not or the opposite. The first is the Schrödinger picture, while the second is the Heisenberg picture. Let us see how we pass from one to the other.

Let us consider the expectation value of the operator $\hat{O}$ in a state $|\Psi(t)\rangle$ in the Schrödinger picture:

$$\langle\Psi(t)|\hat{O}|\Psi(t)\rangle = \langle\Psi(0)|e^{i\hat{H}t}\hat{O}e^{-i\hat{H}t}|\Psi(0)\rangle = \langle\Psi_H(0)|\hat{O}_H(t)|\Psi_H(0)\rangle. \tag{1.28}$$

The expectation value is the same as in the Schrödinger picture but the Heisenberg operator depends on time while the states are fixed at $t = 0$, where they coincide with the Schrödinger ones.

In the Heisenberg picture, the Schrödinger equation is substituted by the operator equation (since only operators evolve in time):

$$\frac{d}{dt}\hat{O}_H(t) = i[\hat{H}, \hat{O}_H]. \tag{1.29}$$

An intermediate picture which is useful in the context of perturbation theory is called the interaction picture. If the Hamiltonian can be separated as $\hat{H} = \hat{H}_0 + \hat{H}_{\rm int}$, where $\hat{H}_0$ is the free Hamiltonian and $\hat{H}_{\rm int}$ includes the interaction, both states and operators evolve

$$\langle\Psi(t)|\hat{O}|\Psi(t)\rangle = \langle e^{i\hat{H}_0 t}\Psi(t)|e^{i\hat{H}_0 t}\hat{O}e^{-i\hat{H}_0 t}|e^{i\hat{H}_0 t}\Psi(t)\rangle = \langle\Psi_I(t)|\hat{O}_I(t)|\Psi_I(t)\rangle. \tag{1.30}$$

The operators satisfy therefore

$$\frac{d}{dt}\hat{O}_I(t) = i[\hat{H}_0, \hat{O}_I], \tag{1.31}$$

while the states

$$i\frac{d}{dt}|\Psi_I(t)\rangle = -\hat{H}_0|\Psi_I(t)\rangle + e^{i\hat{H}_0 t}He^{-i\hat{H}_0 t}|\Psi_I(t)\rangle = \hat{H}_I(t)|\Psi_I(t)\rangle. \tag{1.32}$$

1.4 Quantum harmonic oscillator

Since your first course in QM you have become familiar with the quantization of the system of a particle of mass m attached to a spring, or harmonic oscillator. The classical Hamiltonian is given by:

$$H(x) = \frac{\mathbf{p}^2}{2m} + \frac{1}{2} m\omega^2 (x - x_0)^2. \tag{1.33}$$

The canonical quantization of this system proceeds by constructing the quantum Hamiltonian according to eqs. (1.25).

The stationary states can be found by solving the Schrödinger equation, but it can be done much more easily by an algebraic method [2] defining the raising/lowering operators as combinations of the momentum and position operators:

$$\hat{a} \equiv \frac{1}{\sqrt{2\hbar\omega m}} \left(i\hat{P} + m\omega\hat{X} \right), \ \hat{a}^\dagger \equiv \frac{1}{\sqrt{2\hbar\omega m}} \left(-i\hat{P} + m\omega\hat{X} \right). \tag{1.34}$$

The Hamiltonian can be written as

$$\hat{H} = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right), \tag{1.35}$$

and using eq. (1.20),

$$[\hat{a}, \hat{a}^\dagger] = 1, \ [\hat{a}^\dagger, \hat{a}^\dagger] = 0, \ [\hat{a}, \hat{a}] = 0. \tag{1.36}$$

The eigenvalues can be shown to be quantized according to

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right), \tag{1.37}$$

with n any integer. The operator $\hat{a}^\dagger$ acting on the eigenstate corresponding to integer n, $|n\rangle$, produces another eigenstate with integer $n+1$, while the operator $\hat{a}$ transforms the state $|n\rangle \to |n-1\rangle$. More precisely,

$$\hat{a}^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle, \quad \hat{a}|n\rangle = \sqrt{n} |n-1\rangle. \tag{1.38}$$

We will see that there is a different interpretation to the rising and lowering operators of the harmonic oscillator. We can imagine that when we move from energy level $E_n \to E_{n\pm1}$ we are adding or subtracting a quantum of energy $\hbar\omega$. Such quantum of energy can be viewed as a particle like Einstein's photon. The operators $\hat{a}^\dagger$ and $\hat{a}$ in this picture are particle creation and annihilation operators. We will see that this interpretation is precisely what provides the particle interpretation of the quantized electromagnetic field, since the canonical quantization of the electromagnetic field will lead to a infinite set of harmonic oscillators.

[2] For those not familiar with this formalism, please read Chapter 2 of Griffith's book on Quantum Mechanics.

It is not surprising that when we move to the relativistic domain, it becomes necessary to extend the one-particle Hilbert space to include states with different number of particles, since interactions in this regime will generically modify the number of particles when the energy available is greater than the rest energy of one particle, $E \gg mc^2$.

A state of N free identical particles can be described by a state which is a superposition of the direct product of one-particle states in Hilbert space:

$$|\Psi_N\rangle = |\Psi_{\alpha_1}\rangle \otimes |\Psi_{\alpha_2}\rangle \otimes ....|\Psi_{\alpha_N}\rangle \in \mathcal{H}_{(N)}, \tag{1.39}$$

where each of the factors, $|\Psi_{\alpha_k}\rangle$, represents the state of particle k.

It is a postulate of quantum mechanics (that becomes the spin-statistics theorem in QFT) that not all possible states of this form are physical. Only the combinations of such states that are either completely symmetric (bosonic) or completely antisymmetric (fermionic) under the exchange of any two identical particles $i \leftrightarrow j$ are physically acceptable states. If we define the permutation operator $\hat{P}_{ij}$ of the i-th and j-th particles, physical states of N identical particles must be combinations of states of $\mathcal{H}_{(N)}$ that satisfy

$$\hat{P}_{ij}|\Psi_N\rangle = \pm|\Psi_N\rangle, \tag{1.40}$$

for any i, j. The $\pm$ correspond to bosons/fermions and whether it is one or the other depends on whether the spin of the one-particle state is integer or half-integer.

The full symmetrization or antisymmetrization of states of N particles implies that any such state is fully determined by the occupation numbers, N_i:

$$N_i = \#\text{number of particles in state } |\Psi_\text{i}\rangle, \tag{1.41}$$

and whether the particles are bosons or fermions. Up to a normalization:

$$|\Psi_N\rangle_{B/F} \sim \sum_{\text{perm}} (-1)^{\text{perm}} \overbrace{|\Psi_1\rangle \otimes ..|\Psi_1\rangle}^{N_1 \ terms} \otimes ... \overbrace{|\Psi_k\rangle \otimes ... \otimes |\Psi_k\rangle}^{N_k \ terms} \tag{1.42}$$

with

$$\sum_{i=1,..k} N_i = N. \tag{1.43}$$

We can therefore denote the state by

$$|N_1, ..., N_k\rangle_{F/B}, \ \sum_{i=1,k} N_i = N. \tag{1.44}$$

We will omit the sub-index F/B in the following.

When in a process we can change the number of particles, we need to enlarge this space to include all states with an arbitrary number of particles. This is called Fock space, which is defined as the direct sum of the Hilbert spaces of any number of particles

$$\mathcal{F} = \bigoplus_{N=0}^{\infty} \mathcal{H}_{(N)}, \quad \mathcal{H}_{(0)} \equiv |0\rangle, \tag{1.45}$$

where the state with no particles is called the vacuum state, $|0\rangle$.

We can define operators that transform states with different number of particles, in particular operators of creation $a_i^\dagger$ and annihilation a_i of a particle in a one-particle state of energy E_i:

$$\begin{aligned} \hat{a}_i^\dagger &: \mathcal{H}_{(N)} \to \mathcal{H}_{(N+1)} \\ \hat{a}_i &: \mathcal{H}_{(N)} \to \mathcal{H}_{(N-1)}. \end{aligned} \tag{1.46}$$

For bosons we define the operators of creation and annihilation as:

$$\begin{aligned} \hat{a}_i|N_1,...,N_i,...N_n\rangle &= \sqrt{N_i}|N_1,...,N_i-1,...N_n\rangle, \\ \hat{a}_i^\dagger|N_1,...,N_i,...N_n\rangle &= \sqrt{N_i+1}|N_1,...,N_i+1,...N_n\rangle \end{aligned} \tag{1.47}$$

It is easy to check the following properties:

- $\hat{N}_i \equiv \hat{a}_i^\dagger \hat{a}_i$ is the $i-th$ number operator since

$$\hat{N}_i|N_1,...,N_i,...,N_n\rangle = N_i|N_1,...,N_i,...,N_n\rangle_B \tag{1.48}$$

- they satisfy the following commutation relations:

$$[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}, \quad [\hat{a}_i, \hat{a}_j] = 0, \quad [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0, \tag{1.49}$$

- any state in Fock space can be constructed from the vacuum by acting with creation and annihilation operators

$$|N_1,....N_n\rangle_B = \frac{1}{\sqrt{N_1!}...\sqrt{N_n!}}(\hat{a}_1^\dagger)^{N_1}....(\hat{a}_n^\dagger)^{N_n}|0\rangle. \tag{1.50}$$

  indeed this state can be shown to be properly normalized and completely symmetric.

In the case of fermions, the occupation numbers can only be $0, 1$. The creation/annihilation operators are defined as:

$$\begin{aligned} \hat{a}_i|N_1, ..., N_i, ...N_n\rangle &= (-1)^{\nu_i} N_i |N_1, ..., 1-N_i, ...N_n\rangle, \\ \hat{a}_i^\dagger|N_1, ..., N_i, ...N_n\rangle &= (-1)^{\nu_i} (1-N_i) |N_1, ..., 1-N_i, ...N_n\rangle_F, \end{aligned} \tag{1.51}$$

with $\nu_i \equiv \sum_{k=0}^{i-1} N_k$. As before the number operator is $\hat{N}_i = \hat{a}_i^\dagger \hat{a}_i$, while the conmutation relations in this case are found to be

$$\{\hat{a}_i, \hat{a}_j^\dagger\} = \delta_{ij}, \quad \{\hat{a}_i, \hat{a}_j\} = 0, \quad \{\hat{a}_i^\dagger, \hat{a}_j^\dagger\} = 0, \tag{1.52}$$

and any state can be reconstructed from the vacuum by

$$|N_1, ...N_n\rangle_F = (\hat{a}_1^\dagger)^{N_1} ....(\hat{a}_n^\dagger)^{N_n} |0\rangle. \tag{1.53}$$

Although this formalism of creation/annihilation operators is extremely useful to discuss states with many particles, it can also be applied to the Hilbert space of one-particle states, and in this context we talk of the *second quantization* formalism.

	1st quan	2nd quan
States	$\|\Psi\rangle$	$\hat{a}_\Psi^\dagger \|0\rangle$
	$\langle\Psi\|$	$\langle 0\| \hat{a}_\Psi$
	$\|\Psi\rangle = \int_{\mathbf{p}} \|\mathbf{p}\rangle\langle\mathbf{p}\|\Psi\rangle$	$\hat{a}_\Psi^\dagger = \int_{\mathbf{p}} \langle\mathbf{p}\|\Psi\rangle \hat{a}_{\mathbf{p}}^\dagger$
Operators	$\hat{O} = \sum_{\alpha,\beta} \|\alpha\rangle\langle\alpha\|\hat{O}\|\beta\rangle\langle\beta\|$	$\hat{O} = \sum_{\alpha,\beta} \hat{a}_\alpha^\dagger \hat{a}_\beta \langle\alpha\|\hat{O}\|\beta\rangle$

Among the creation/annihilation operators, a special case are those corresponding to the position basis, that is the operators that create a particle at position $\mathbf{x}$, also called field operators:

$$\hat{\Psi}_{\mathbf{x}} \equiv \hat{a}_{\mathbf{x}}^\dagger, \quad \hat{a}_{\mathbf{x}}^\dagger |0\rangle = |\mathbf{x}\rangle. \tag{1.54}$$

It is very important to note that despite the similar notation, the field operators must not be mistaken with wave functions, as they are operators and not complex numbers (i.e. the components of the state vector in the position basis). Note that the field operator can also be written as

$$\Psi_{\mathbf{x}} = \int_{\mathbf{p}} \langle\mathbf{p}|\mathbf{x}\rangle \; \hat{a}_{\mathbf{p}}^\dagger = \int_{\mathbf{p}} e^{-i\mathbf{p}\mathbf{x}} \; \hat{a}_{\mathbf{p}}^\dagger. \tag{1.55}$$

We will see that these are the basic building blocks in the construction of quantum field theories. The quantization of the electromagnetic field is equivalent to that of a set of quantum harmonic oscillators, where the raising/lowering operators can then be interpreted as momentum basis creation/annihilation operators in a Fock space, and this is what gives a precise definition of what we mean by a photon.

Causality as we will see will be ensured by the locality of the field operator.

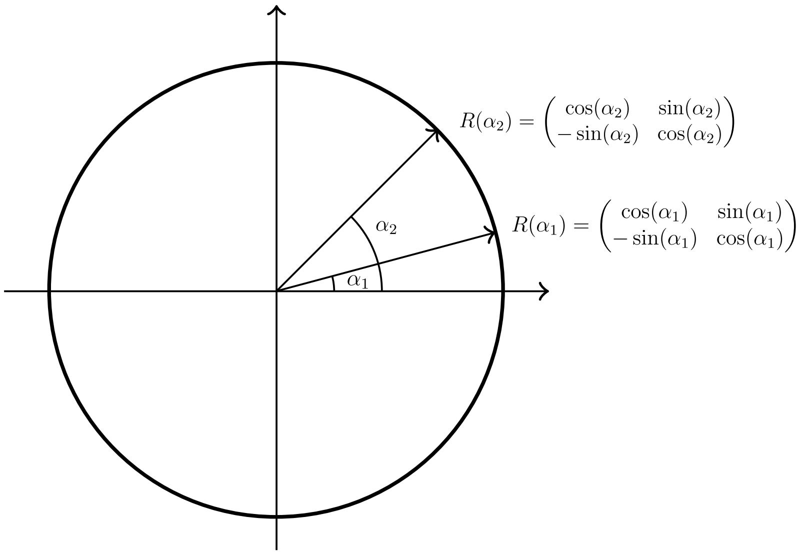


Figure 5: The parameter space of the rotation group in 2D is isomorphic to a unit circle. The parameter α_1 parameterizes the group element $R(\alpha_1)$, which is the well-known rotation in 2D. $R(\alpha_i) \equiv R(S(\alpha_i))$ is a matrix representation of the group element $S(\alpha_1)$, which lives in the abstract, mathematical space of group elements.

1.6 Symmetries in QM

Symmetries in quantum mechanics are transformations of state vectors that leave observables and probabilities invariant. Symmetries under translations, rotations and boosts are expected from the principle of relativity. Other internal symmetries such as gauge symmetries are fundamental in the dynamics of elementary particles as we will see. Symmetry transformations have a group structure since:

- there is an identity transformation (i.e. to do nothing)
- the product can be defined as the ordered composition of two transformations that also belong to the group

$$S_2 S_1 = \text{apply first transformation } S_1 \text{ and afterwards } S_2. \tag{1.56}$$

- an inverse can be defined (i.e. undo the transformation)
- associative product

$$S_3(S_2 S_1) = (S_3 S_2) S_1. \tag{1.57}$$

All the examples of continuous symmetries in physics that we will be dealing with are Lie groups[3], where the elements are functions of a set of real and continuous parameters, α, and the multiplication rule depends smoothly on them, meaning that there is some notion of "closeness" in the group: if two elements $S(\alpha_1)$ and $S(\alpha_2)$ of a group S are close together, then the parameters α_1 and α_2 are also close. Take as an example the group $U(1)$ of rotations in a 2D plane depicted in fig. 5.

Wigner's theorem establishes that symmetry transformations in Hilbert space are implemented by either unitary and linear operators (in most of the cases) or antiunitary/anti-linear operators (in the case of time inversion). We will be concerned with symmetries that are represented by unitary operators:

$$S(\alpha) : |\Psi\rangle \to U(S(\alpha))|\Psi\rangle, \quad U(S(\alpha))^\dagger U(S(\alpha)) = 1. \tag{1.58}$$

We say that these operators in Hilbert space are a unitary representation of the symmetry group.

When we consider infinitesimal transformations, very close to the identity (the element corresponding to $\alpha = \mathbf{0}$, the unitary operators representing a Lie group can be approximated by a Taylor expansion in the parameters:

$$U(S(\alpha)) \simeq 1 + i\alpha_a T^a + O(\alpha^2). \tag{1.59}$$

It is easy to show that the operators $T_a^\dagger = T_a$ are Hermitian and are called *group generators* which form an algebra:

$$[T_a, T_b] = i f_{abc} T_c. \tag{1.60}$$

These generators and their commutation relations are all we need to describe the group elements and group multiplication close to the identity.

The generators of the familiar symmetry operations:

- Spacial translations: under a translation of the origin of coordinates by $-\mathbf{a}$, a particle localized at $\mathbf{x}$ will appear localized at the point $\mathbf{x} + a$, therefore

  $$U(T(\mathbf{a}))|\mathbf{x}\rangle = |\mathbf{x} + a\rangle. \tag{1.61}$$

  Using this relation, it is easy to show that the generators of infinitesimal translations are the three components of the momentum operator:

  $$T_i = i\frac{\partial}{\partial x_i} = -P^i, \tag{1.62}$$

  and the unitary operator that represents finite translations is

  $$U(T(\mathbf{a})) = e^{-i\mathbf{a}\cdot\hat{\mathbf{P}}}. \tag{1.63}$$

[3] An excellent group theory reference for particle physicists: *Lie Algebras in Particle Physics*, by H. Georgi [].

- Time translations: under a translation of the time origin $-t_0$

$$U(T(t_0))|\Psi(t)\rangle = |\Psi(t+t_0)\rangle, \tag{1.64}$$

and using the Schrödinger equation it follows that the generator of time translations is the Hamiltonian. A finite time translation is implemented by the evolution operator

$$U(T(t_0)) = e^{-i\hat{H}t_0}. \tag{1.65}$$

This is of course the expected result according to eq. (1.27).

- Rotations: under a generic rotation R^{-1} of the coordinate system, the position space basis vectors change as

$$U(R)|\mathbf{x}\rangle = |R\mathbf{x}\rangle. \tag{1.66}$$

Using this property it is easy to show that the generators of rotations around the coordinate axes are the angular momentum quantum operators, $\hat{L}_i$, that satisfy the algebra:

$$[\hat{L}_i, \hat{L}_j] = i\epsilon_{ijk}\hat{L}_k. \tag{1.67}$$

We say that a dynamical system is invariant under a symmetry if the Hamiltonian commutes with the unitary operator that represents the group in Hilbert space:

$$[\hat{H}, U(S)] = 0, \tag{1.68}$$

which immediately implies that it commutes with all the generators

$$[\hat{H}, \hat{T}_i] = 0. \tag{1.69}$$

It then follows that the expectation values of the generators (which are Hermitian and therefore valid observables) are conserved quantities:

$$\frac{d}{dt}\langle\Psi|\hat{T}_i|\Psi\rangle = 0. \tag{1.70}$$

1.7 Special relativity in the quantum domain

The group of space-time symmetries in relativity is the Lorentz group, $SO(3,1)$ (or the Poincaré group when translations are also included). This is the group of linear coordinate transformations that preserve the Minkowski metric:

$$x'^{\mu} = \Lambda^{\mu}_{\nu}x^{\nu}, \quad g_{\mu\nu}\Lambda^{\mu}_{\sigma}\Lambda^{\nu}_{\tau} = g_{\sigma\tau} \tag{1.71}$$

where

$$g_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}. \tag{1.72}$$

It is easy to see that this change of coordinates preserve the inner product in Minkowski space:

$$x^{\mu} x^{\nu} g_{\mu\nu} = x'^{\mu} x'^{\nu} g_{\mu\nu}. \tag{1.73}$$

Property 1: the speed of light is constant

Let us consider $(t, \mathbf{x})$ as the coordinate separating two events. If these two events correspond to a light ray being emitted at the origin at $t = 0$ and reaching $\mathbf{x}$ at t, since the light speed is one in natural units:

$$|\mathbf{x}| = t, \tag{1.74}$$

therefore the norm in eq. (1.73) is $t^2 - \mathbf{x}^2 = 0$ and since this norm is invariant according to (1.73) it should also be the same in all reference frames. In a different reference frame where the point has coordinates $(t', \mathbf{x}')$, the norm $t'^2 - \mathbf{x}'^2 = 0$ and therefore in this new frame the light ray also moves at the speed of light.

Property 2: the speed of light is the maximum speed

In Fig. 6 we show on the plane $(t, |\mathbf{x}|)$ the so called light-cone, that is the region satisfying

$$t^2 - |\mathbf{x}|^2 > 0, \tag{1.75}$$

corresponding to a signal emitted at the origin at $t = 0$ with a velocity below than one and reaching $|\mathbf{x}|$ at time t. This is a time-like event. The limit of the light-cone corresponds to the light rays, with $t = |\mathbf{x}|$.

The region outside the light cone satisfies

$$t^2 - |\mathbf{x}|^2 < 0, \tag{1.76}$$

and corresponds therefore to a velocity above one, these are the space-like events.

Let's now consider a Lorentz transformation: $(t, |\mathbf{x}|) \to (t', |\mathbf{x}'|)$ for time-like and space-like events. They must be necessarily on the curve:

$$t'^2 - |\mathbf{x}'|^2 = a, \tag{1.77}$$

with $a > 0$ for time-like and $a < 0$ for space-like events. The curves are shown in Fig. 6.

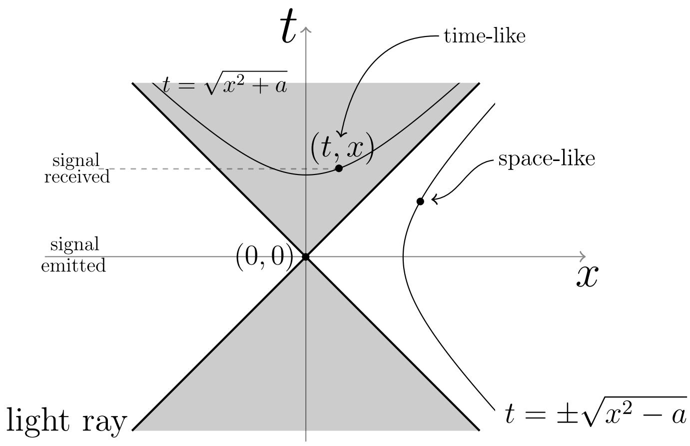


Figure 6: Light-cone, time-like and space-like events.

For time-like events, if $t > 0$ in one frame, $t' > 0$ in any frame. For space-like events, this is not the case and two events that are causally connected (one occurs after the other) might occur in the opposite order in a different frame. Causality is therefore not preserved for space-like events.

The velocity of light is therefore a maximum if we want to preserve causality.

Generators of the Lorentz group

There are six generators of Lorentz transformations. If we consider an infinitesimal transformation around the identity:

$$\Lambda^{\mu}{}_{\nu} \simeq \delta^{\mu}_{\nu} + \omega^{\mu}{}_{\nu} + \mathcal{O}(\omega^2), \tag{1.78}$$

and substitute in eq. (1.71), we obtain the relation

$$\omega^{\mu}{}_{\sigma} g_{\mu\tau} = -\omega^{\nu}{}_{\tau} g_{\sigma\nu} \to \omega_{\tau\sigma} = -\omega_{\sigma\tau}. \tag{1.79}$$

$\omega_{\mu\nu}$ is therefore an asymmetric real matrix and it depends on six real parameters that correspond to the three space rotations and boosts.

When defining a one-particle state we need to know how it transforms under Lorentz/Poincaré transformations. It might do so trivially and then we talk about a scalar object, its wave function, $\langle \mathbf{x}|\Psi\rangle$, would just be a complex number such that under a Lorentz Transformation

$$U(\Lambda)|\Psi\rangle = |\Psi'\rangle \to \Psi'(\mathbf{x}) = \langle \mathbf{x}|\Psi'\rangle = \langle \mathbf{x}|U(\Lambda)\Psi\rangle = \langle \Lambda^{-1}\mathbf{x}|\Psi\rangle = \Psi(\Lambda^{-1}\mathbf{x}). \tag{1.80}$$

However it is an experimental fact that there exists one-particle states that transform non-trivially under the Lorentz group. It is possible to classify all the irreducible representations of the Lorentz group – as we will see later. In this case the wave function is not fixed by one complex number but has another index running up to the dimension of the representation such that

$$\Psi'_\alpha(\mathbf{x}) = \langle \mathbf{x} | U(\Lambda)\Psi\rangle = R(\Lambda)_\alpha^\beta \Psi_\beta(\Lambda^{-1}\mathbf{x}). \tag{1.81}$$

We will consider several of these non-trivial representations later on.

1.8 *Relativistic QM versus causality*

Let us consider a free particle in position $\mathbf{x}_0$ at time $t = 0$:

$$|\Psi(0)\rangle = |\mathbf{x}_0\rangle. \tag{1.82}$$

We want to calculate the probability that the particle is found at time t at position $\mathbf{x}$, which is given by the square of the amplitude:

$$\begin{aligned} \langle \mathbf{x}|e^{-i\hat{H}t}|\mathbf{x}_0\rangle &= \int \frac{d\mathbf{p}}{(2\pi)^3} \langle \mathbf{x}|\mathbf{p}\rangle\langle \mathbf{p}|\mathbf{x}_0\rangle e^{-it\sqrt{\mathbf{p}^2+m^2}} \\ &= \int \frac{d\mathbf{p}}{(2\pi)^3}\, e^{i\mathbf{p}(\mathbf{x}-\mathbf{x}_0)} e^{-it\sqrt{\mathbf{p}^2+m^2}}, \end{aligned} \tag{1.83}$$

where we have assumed that the relativistic relation $E = \sqrt{\mathbf{p}^2 + m^2}$ holds at the operator level. We are interested in the answer when the points $(t, \mathbf{x})$ and $(0, \mathbf{x}_0)$ are causally disconnected, ie.

$$t^2 < (\mathbf{x} - \mathbf{x}_0)^2. \tag{1.84}$$

Assuming this condition holds, the integral can be expressed in terms of Bessel functions:

$$\begin{aligned} \text{eq. (1.83)} &= \frac{1}{(2\pi)^3}\int_0^\infty p^2 dp \int_{-1}^1 d\cos\theta \int_0^{2\pi} d\phi\; e^{ip\cos\theta|\mathbf{x}-\mathbf{x}_0|} e^{-it\sqrt{p^2+m^2}} \\ &= \frac{1}{2\pi^2}\int_0^\infty p dp\; \frac{\sin(p|\mathbf{x}-\mathbf{x}_0|)}{|\mathbf{x}-\mathbf{x}_0|} e^{-it\sqrt{p^2+m^2}} \\ &= -\frac{imt}{2\pi^2 r^2}\frac{\partial}{\partial r}\left(\frac{K_1(m\sqrt{r^2-t^2})}{\sqrt{r^2-t^2}}\right)\Bigg|_{r=|\mathbf{x}-\mathbf{x}_0|}. \end{aligned} \tag{1.85}$$

The integral is different from zero which implies that there is a non-zero probability that the particle has moved faster than light if it has reached $\mathbf{x}$ in time t. It is clear that the concept of causality clashes with the description of particles in terms of wave functions, when there is no restriction for these wave functions to extend beyond causally connected regions. There is a loophole in this argument and it is that we have assumed we could localize a quantum particle as much as we wished and it would still be described in terms of a one-particle state. This is of course not possible since the more we localize the particle the greater the momentum, and at some point it is possible to create more particles.

However, fields in classical physics, such as the electromagnetic field, extend generically over causally disconnected regions, and this does not violate causality provided information cannot be exchanged between these regions. In QFT this is assured by the restriction of observables to being local [4]. Causality will then ensured if local measurements in causally disconnected regions cannot interfere with each other. This implies that the set of observables localized in causally disconnected regions of space-time must commute:

$$[\hat{O}(x), \hat{O}(y)] = 0, \qquad (x-y)^2 < 0. \tag{1.86}$$

The construction of observables such as the Hamiltonian in terms of local quantum fields, satisfying eq. (1.86), underlies three fundamental properties of nature:

- Spin-statistics connection: half-integer spin particles have antisymmetric wave functions while integer spin particles have symmetric wave functons
- CPT is a symmetry
- The existence of antiparticles

[4]One cannot measure in the lab a property of an electron in the Andromeda galaxy, so all observables must be local to some extent.

II. Canonical Quantization of Relativistic Scalar Fields

Historically the first derivation of quantum fields came from the canonical quantization of the electromagnetic field. We will consider first the canonical quantization of a simpler field, the Klein-Gordon field, and postpone to a later chapter the extra complication of dealing with gauge invariance. We start with a quick reminder of basic facts of classical fields.

2.1 Classical field theory

A classical field theory is just a mechanical system that is described in terms of a magnitude or set of magnitudes, $\phi(x)$, that need to be defined for each point of space at some fixed time. For example a classical field is the density of a fluid, $\rho(\mathbf{x})$, or the electric field $\mathbf{E}(\mathbf{x})$. Such fields usually depend on time as dictated by some classical equations that can be derived from a least-action principle, where the action is defined in terms of a Lagrangian density, a functional of the field and its partial derivatives:

$$S = \int d^4x \; \mathcal{L}[\phi(x), \partial_\mu \phi(x)]. \tag{2.2}$$

The least action principle states that the fields should be such that the action remains invariant under arbitrary infinitesimal variations of the fields, $\phi \to \phi + \delta\phi$:

$$\begin{aligned} 0 = \delta S &= \int d^4x \left[\left.\frac{\partial \mathcal{L}}{\partial \phi}\right|_{\delta\phi=0} \delta\phi + \left.\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right|_{\delta\phi=0} \partial_\mu \delta\phi \right] \\ &= \int d^4x \left\{ \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) \right] \delta\phi + \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi \right] \right\}. \end{aligned} \tag{2.3}$$

The integral of the second term depends on the values of the fields in the asymptotic boundaries that we will assume to vanish. The first term must vanish for arbitrary field variations and this is only possible if

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) = 0. \tag{2.4}$$

These are the Euler-Lagrange equations.

Let us consider the simplest field theory, where the field is a scalar $\phi(x)$, and a Lorentz-invariant Lagrangian

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \; g^{\mu\nu} \partial_\nu \phi - \frac{1}{2} m^2 \phi^2. \tag{2.5}$$

It is easy to check that the Lagrangian remains invariant under a Lorentz transformation:

$$x' = \Lambda x, \quad \phi'(x') = \phi(x) = \phi(\Lambda^{-1} x'), \tag{2.6}$$

or

$$\phi'(x) = \phi(\Lambda^{-1}x). \tag{2.7}$$

The Euler-Lagrange equation is the Klein-Gordon equation:

$$(\partial_\mu \partial^\mu + m^2)\phi = 0. \tag{2.8}$$

Performing a Fourier transform of ϕ in space,

$$\phi(t, \mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} e^{i\mathbf{p}\mathbf{x}} \tilde{\phi}(t, \mathbf{p}), \tag{2.9}$$

and substituting back in the KG equation we find

$$\frac{d^2}{dt^2}\tilde{\phi}(t, \mathbf{p}) = -(\mathbf{p}^2 + m^2)\tilde{\phi}(t, \mathbf{p}), \tag{2.10}$$

which is the equation of motion of a classical harmonic oscillator with frequency $\omega = \sqrt{\mathbf{p}^2 + m^2}$. As we will see, the canonical quantization of the KG field will give a superposition of quantum harmonic oscillators.

2.2 Symmetries and Conservation Laws

As we know, symmetries in classical particle mechanics lead to conservation laws. This is also the case for classical field theories and it is the content of the famous Noether's theorem. Let us consider the following infinitesimal field transformation:

$$\phi(x) \to \phi(x) + \alpha \Delta\phi(x), \tag{2.11}$$

that leaves the Lagrangian invariant or is changed by a total derivative[5]

$$\mathcal{L}(x) \to \mathcal{L}(x) + \alpha \partial_\mu \mathcal{J}^\mu, \quad \Delta\mathcal{L}(x) = \alpha \partial_\mu \mathcal{J}^\mu. \tag{2.12}$$

We can also write the variation of the Lagrangian as:

$$\Delta\mathcal{L}(x) \to \alpha \Delta\phi \left(\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) + \alpha \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Delta\phi \right), \tag{2.13}$$

the first term is zero because of the equations of motion. Using eqs. (2.12) and (2.13) we obtain

$$\partial_\mu \left(\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) \Delta\phi \right) = \partial_\mu \mathcal{J}^\mu, \tag{2.14}$$

[5] A total derivate would not contribute to the action.

and therefore a conserved current exists:

$$\partial_\mu j^\mu = 0, \; j^\mu(x) \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Delta\phi - \mathcal{J}^\mu, \tag{2.15}$$

which in turn implies a conserved charge:

$$\frac{dQ}{dt} \equiv \frac{d}{dt} \int d^3x j^0 = - \int d^3x \; \nabla \mathbf{j} = 0. \tag{2.16}$$

Important examples are the invariance under a global rephasing or the symmetry under space-time translations.

Rephasing invariance
Let us consider a KG field that is complex $\phi(x)$. The Lagrangian in this case is

$$\mathcal{L} = \partial_\mu \phi^\dagger g^{\mu\nu} \partial_\nu \phi - m^2 \phi^\dagger \phi. \tag{2.17}$$

It is easy to see that this Lagrangian is invariant under the transformation:

$$\phi(x) \to e^{i\alpha} \phi(x). \tag{2.18}$$

For an infinitesimal α, eq. (2.11) implies $\Delta\phi = i\phi$. Noting that we now have two fields, ϕ and $\phi^\dagger$, the previous derivation of the conserved current of eq. (2.15) results in a summation over the two fields,

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \Delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^\dagger)} \Delta\phi^\dagger, \tag{2.19}$$

which simplifies to

$$j^\mu = i \left[(\partial^\mu \phi^\dagger)\phi - (\partial^\mu \phi)\phi^\dagger \right]. \tag{2.20}$$

When this complex KG field is coupled to the electromagnetic field we will see that the associated conserved charge is nothing but the electromagnetic charge.

Space-time translations
Let's now consider space-time translations

$$x_\mu \to x'_\mu = x_\mu - \alpha_\mu, \quad \phi'(x') = \phi(x) = \phi(x' + \alpha), \tag{2.21}$$

such that if we consider infinitesimal translations in the μ direction,

$$\phi(x) \to \phi(x + \alpha) = \phi(x) + \alpha^\mu \partial_\mu \phi(x) + ...$$

we find that the field variation is

$$\alpha^\mu \Delta_\mu \phi(x) = \phi'(x) - \phi(x) = \alpha^\mu \partial_\mu \phi \to \Delta_\mu \phi = \partial_\mu \phi. \tag{2.22}$$

This transformation law applies to any field and, in particular, for the Lagrangian itself, which is a scalar:

$$\Delta\mathcal{L} = \alpha^\mu \partial_\mu \mathcal{L} = \alpha^\mu \partial_\nu (\delta^\nu_\mu \mathcal{L}), \tag{2.23}$$

which means that

$$\mathcal{J}^\nu_\mu = \delta^\nu_\mu \mathcal{L}. \tag{2.24}$$

The conserved current, eq. (2.12):

$$T^\nu_\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\nu\phi)}\Delta_\mu\phi - \mathcal{J}^\nu_\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\nu\phi)}\partial_\mu\phi - \delta^\nu_\mu\mathcal{L}, \tag{2.25}$$

is the energy-momentum tensor.[6] The conserved charges associated with time and space translations are respectively the Hamiltonian and the three momentum:

$$H = \int d^3x\, T^{00}, \quad P^i = \int d^3x\, T^{0i}. \tag{2.26}$$

Exercise: Derive eqs. (2.20) and (2.25). Show using the equations of motion that they are conserved currents.

2.3 Canonical quantization of the real KG field

By canonical quantization we refer to a procedure of quantization which is parallel to what we do for the system of one particle. There, we promote the position, q, and momentum, p, to operators in Hilbert space:

$$q, p \to \hat{q}, \hat{p}, \tag{2.27}$$

that satisfy the canonical commutation relations:

$$[\hat{q}, \hat{p}] = i. \tag{2.28}$$

The Hamiltonian that controls the temporal evolution of the system is a function of p, q and therefore also an operator.

$$H(q,p) \to \hat{H} = H(\hat{q}, \hat{p}). \tag{2.29}$$

For a continuous system, in order to determine the classical state of the system at any fixed time, we need to specify a field value at all space points, therefore a single position becomes an infinite set of field values:

$$q(t) \to \phi(t, \mathbf{x}). \tag{2.30}$$

[6] The index ν is the index of the current, while μ is the label of the translation direction.

The corresponding momenta are obtained in analogy with the discrete system:

$$\pi(t,\mathbf{x}) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(t,\mathbf{x})}, \quad \dot{\phi}(t,\mathbf{x}) = \frac{\partial}{\partial t}\phi(t,\mathbf{x}), \tag{2.31}$$

where $\mathcal{L}$ is the Lagrangian density. The Hamiltonian density is therefore

$$\mathcal{H} = \phi(t,\mathbf{x})\dot{\phi}(t,\mathbf{x}) - \mathcal{L}. \tag{2.32}$$

For the Klein-Gordon field in eq. (2.5) we have

$$\mathcal{H} = \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2. \tag{2.33}$$

Quantization proceeds by promoting the classical field and the momenta to operators:

$$\phi(t,\mathbf{x}), \pi(t,\mathbf{x}) \to \hat{\phi}(t,\mathbf{x}), \hat{\pi}(t,\mathbf{x}), \tag{2.34}$$

satisfying the equal-time commutation relations:

$$[\hat{\phi}(t,\mathbf{x}), \hat{\pi}(t,\mathbf{y})] = i\delta^{(3)}(\mathbf{x}-\mathbf{y}),\ [\hat{\phi}(t,\mathbf{x}), \hat{\phi}(t,\mathbf{y})] = 0,\ \ [\hat{\pi}(t,\mathbf{x}), \hat{\pi}(t,\mathbf{y})] = 0. \tag{2.35}$$

The Hamiltonian is therefore the operator:

$$\hat{H} = \int d^3x \left(\frac{1}{2}\hat{\pi}^2 + \frac{1}{2}(\nabla\hat{\phi})^2 + \frac{1}{2}m^2\hat{\phi}^2\right). \tag{2.36}$$

From here onwards we eliminate hats to refer to operators.

Let us consider the Fourier transform:

$$\phi(t,\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}}\tilde{\phi}(t,\mathbf{p}), \quad \pi(t,\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}}\tilde{\pi}(t,\mathbf{p}). \tag{2.37}$$

Substituting in eq. (2.36) we find

$$H = \int \frac{d^3p}{(2\pi)^3}\left[\frac{1}{2}\tilde{\pi}(t,-\mathbf{p})\tilde{\pi}(t,\mathbf{p}) + \frac{1}{2}(\mathbf{p}^2+m^2)\tilde{\phi}(t,-\mathbf{p})\tilde{\phi}(t,\mathbf{p})\right]. \tag{2.38}$$

Since $\phi = \phi^\dagger, \pi = \pi^\dagger$,

$$\tilde{\phi}(t,-\mathbf{p}) = \tilde{\phi}(t,\mathbf{p})^*, \quad \tilde{\pi}(t,-\mathbf{p}) = \tilde{\pi}(t,\mathbf{p})^*. \tag{2.39}$$

We can make the connection to the harmonic oscillator, see eq. (1.34), more precise by defining

$$a_{\mathbf{p}} \equiv \sqrt{\frac{\omega_p}{2}}\left(\tilde{\phi}(t,\mathbf{p}) + i\frac{\tilde{\pi}(t,\mathbf{p})}{\omega_p}\right), \tag{2.40}$$

with

$$\omega_p \equiv \sqrt{\mathbf{p}^2 + m^2}, \tag{2.41}$$

Using eq. (2.39) it is easy to see that

$$\tilde{\phi}(t, \mathbf{p}) = \frac{1}{\sqrt{2\omega_p}}(a_{\mathbf{p}} + a^\dagger_{-\mathbf{p}}), \quad \tilde{\pi}(t, \mathbf{p}) = -i\sqrt{\frac{\omega_p}{2}}(a_{\mathbf{p}} - a^\dagger_{-\mathbf{p}}). \tag{2.42}$$

From eq. (2.35), the commutation relations in momentum space read

$$[\tilde{\phi}(t, \mathbf{p}), \tilde{\pi}(t, \mathbf{q})] \quad = \quad i(2\pi)^3\delta(\mathbf{p} + \mathbf{q}), \tag{2.43}$$

$$[\tilde{\phi}(t, \mathbf{p}), \tilde{\phi}(t, \mathbf{q})] \quad = \quad [\tilde{\pi}(t, \mathbf{p}), \tilde{\pi}(t, \mathbf{q})] = 0 \tag{2.44}$$

and from these, it is easy to derive:

$$[a_{\mathbf{p}}, a^\dagger_{\mathbf{p}'}] = (2\pi)^3\delta^{(3)}(\mathbf{p} - \mathbf{p}'), \quad [a_{\mathbf{p}}, a_{\mathbf{p}'}] = 0, \quad [a^\dagger_{\mathbf{p}}, a^\dagger_{\mathbf{p}'}] = 0. \tag{2.45}$$

It is easy to show that in terms of the new operators, the Hamiltonian takes the simple form:

$$H = \int \frac{d^3p}{(2\pi)^3}\omega_p \left(a^\dagger_{\mathbf{p}} a_{\mathbf{p}} + \frac{1}{2}[a_{\mathbf{p}}, a^\dagger_{\mathbf{p}}]\right) = \int \frac{d^3p}{(2\pi)^3}\omega_p \left(a^\dagger_{\mathbf{p}} a_{\mathbf{p}} + \frac{1}{2}(2\pi)^3\delta^{(3)}(\mathbf{0})\right). \tag{2.46}$$

The second term is an infinite constant, often called the vacuum energy, which is normally unobservable (at least if we ignore gravity) and so it can be substracted. We eliminate it by normal-ordering the Hamiltonian,

$$:H: \equiv \int \frac{d^3p}{(2\pi)^3}\omega_p \; a^\dagger_{\mathbf{p}} a_{\mathbf{p}}, \tag{2.47}$$

i.e. reordering the lowering operators always to the right of the raising operators. In this form, it is clear that the system corresponds to a sum of harmonic oscillators, see eq. (1.35), one for each $\mathbf{p}$, with energy, ω_p.

We can also interpret the raising/lowering operators as creation/annihilation operators in a Fock space of identical particles of energy ω_p. From the commutation relations, eq. (2.45):

$$[H, a^\dagger_{\mathbf{p}}] = \omega_p a^\dagger_{\mathbf{p}}, \quad [H, a_{\mathbf{p}}] = -\omega_p a_{\mathbf{p}}, \tag{2.48}$$

it follows that for any eigenstate of H with energy E, the state $a^\dagger_{\mathbf{p}}|E\rangle$ is also an eigenstate of H with energy $E + \omega_p$:

$$H a^\dagger_{\mathbf{p}}|E\rangle = \left([H, a^\dagger_{\mathbf{p}}] + a^\dagger_{\mathbf{p}} H\right)|E\rangle = (\omega_p + E)a^\dagger_{\mathbf{p}}|E\rangle. \tag{2.49}$$

Similarly

$$Ha_{\mathbf{p}}|E\rangle = (-\omega_p + E)a_{\mathbf{p}}|E\rangle, \tag{2.50}$$

which is consistent with the Fock space picture in which $a^\dagger_{\mathbf{p}}$ creates a particle with energy ω_p and $a_{\mathbf{p}}$ destroys a particle with the same energy.

Exercise: from eq. (2.35) show eqs. (2.43) and eqs. (2.45).
Exercise: show eq. (2.46) and the commutation relations eq. (2.48).

To further characterize the particle, we can evaluate the momentum, using eq. (2.26). Going to Fourier space and changing variables to $a, a^\dagger$, we find, after normal ordering that the total momentum is given by:

$$:P^i: = -:\int d^3x\ \pi\partial_i\phi: = \int \frac{d^3p}{(2\pi)^3} p^i a^\dagger_{\mathbf{p}} a_{\mathbf{p}}. \tag{2.51}$$

The commutation relation

$$[:P^i:, a^\dagger_{\mathbf{p}}] = p^i a^\dagger_{\mathbf{p}}, \tag{2.52}$$

implies that the particle created by $a^\dagger_{\mathbf{p}}$ has momentum $\mathbf{p}$, since then

$$:P^i:|\mathbf{p}\rangle = :P^i:a^\dagger_{\mathbf{p}}|0\rangle = p^i a^\dagger_{\mathbf{p}}|0\rangle = p^i|\mathbf{p}\rangle, \tag{2.53}$$

and therefore satisfies

$$E^2 = \omega^2_{\mathbf{p}} = \mathbf{p}^2 + m^2, \tag{2.54}$$

the relativistic dispersion relation of a free particle with mass m.

Furthermore the commutation relation in eq. (2.45) implies that the particles created by the operators $a^\dagger_{\mathbf{p}}$ obey Bose statistics: the wave function must be symmetric, this is because

$$a^\dagger_{\mathbf{p}} a^\dagger_{\mathbf{p}'}|0\rangle = a^\dagger_{\mathbf{p}'} a^\dagger_{\mathbf{p}}|0\rangle. \tag{2.55}$$

A one-particle state in this Fock space is therefore proportional to $a^\dagger_{\mathbf{p}}|0\rangle$. What is the correct normalization?

2.4 Relativistic normalization of one-particle states

In non-relativistic quantum mechanics, the one-particle momentum states are normalized as

$$\langle\mathbf{p}|\mathbf{p}'\rangle = (2\pi)^3\delta(\mathbf{p} - \mathbf{p}'), \tag{2.56}$$

and the completeness relation of the momentum basis reads

$$I = \int \frac{d^3p}{(2\pi)^3} |\mathbf{p}\rangle\langle\mathbf{p}|. \tag{2.57}$$

However this normalization is not Lorentz invariant.

Defining

$$|p\rangle \equiv \sqrt{2\omega_p}|\mathbf{p}\rangle, \tag{2.58}$$

we can rewrite the

$$\begin{aligned} I &= \int \frac{d^3p}{(2\pi)^3 2\omega_p} \sqrt{2\omega_p}|\mathbf{p}\rangle\langle\mathbf{p}|\sqrt{2\omega_p} = \int \frac{d^3p}{(2\pi)^3 2\omega_p} |p\rangle\langle p| \\ &= \int \frac{d^4p}{(2\pi)^3} \delta(p^2 - m^2)\theta(p_0)|p\rangle\langle p|, \end{aligned} \tag{2.59}$$

where we have used the following relation

$$\int dp_0 \delta(p^2 - m^2)\theta(p_0) = \int dp_0 \left(\frac{\delta(p_0 - \omega_p)}{2\omega_p} + \frac{\delta(p_0 + \omega_p)}{2\omega_p} \right) \theta(p_0) = \frac{1}{2\omega_p}. \tag{2.60}$$

The measure in the last equation in eq. (2.59),

$$\frac{d^4p}{(2\pi)^3} \delta(p^2 - m^2)\theta(p_0), \tag{2.61}$$

is manifestly Lorentz invariant under $p \to p' = \Lambda p$ since, $p^2 = p'^2$, $\mathrm{sign}(p_0) = \mathrm{sign}(p'_0)$ and $|\det(\Lambda)| = 1$.

The Lorentz invariant normalization of the one-particle states is therefore

$$|p\rangle = \sqrt{2\omega_p} a^\dagger_{\mathbf{p}} |0\rangle. \tag{2.62}$$

2.5 Quantum fields

We can now reconstruct the quantum field from eqs. (2.37) and (2.42) in terms of the creation/annihilation operators:

$$\phi(t, \mathbf{x}) = \int \frac{d^3p}{(2\pi)^3 \sqrt{2\omega_p}} \left[a_{\mathbf{p}}(t) e^{i\mathbf{p}\mathbf{x}} + a^\dagger_{\mathbf{p}}(t) e^{-i\mathbf{p}\mathbf{x}} \right]. \tag{2.63}$$

Finally we can solve for the time evolution. We consider the Heisenberg picture, where only operators change in time, and states do not change. The operators satisfy the Heisenberg equation

$$\frac{d}{dt} a_{\mathbf{p}} = i[H, a_{\mathbf{p}}] = -i\omega_p a_{\mathbf{p}}, \qquad \frac{d}{dt} a^\dagger_{\mathbf{p}} = i[H, a^\dagger_{\mathbf{p}}] = i\omega_p a^\dagger_{\mathbf{p}}, \tag{2.64}$$

which can be solved

$$a_{\mathbf{p}}(t) = e^{-i\omega_p t} a_{\mathbf{p}}(0), \qquad a^\dagger_{\mathbf{p}}(t) = e^{i\omega_p t} a^\dagger_{\mathbf{p}}(0), \tag{2.65}$$

therefore

$$\phi(x) = \phi(t, \mathbf{x}) = \int \frac{d^3p}{(2\pi)^3\sqrt{2\omega_p}} \left[a_{\mathbf{p}}(0) e^{-ipx} + a^\dagger_{\mathbf{p}}(0) e^{ipx} \right], \tag{2.66}$$

where $px = \omega_p t - \mathbf{px}$.

Summarizing, we have found that the canonical quantization of the real KG field results in a quantum field operator that creates/annihilates relativistic and bosonic particles of mass m localized at point $\mathbf{x}$.

Exercise: according to the particle interpretation, $\phi(0, \mathbf{x})$ *acting on the vacuum is the state* $|\mathbf{x}\rangle$*. Show that the state*

$$\langle 0|\phi(0, \mathbf{x})|p\rangle = \langle \mathbf{x}|\mathbf{p}\rangle. \tag{2.67}$$

Exercise: demonstrate that the quantum field satisfies the KG equation starting from the Heisenberg equation:

$$\frac{\partial \hat{\phi}}{\partial t} = i[\hat{H}, \hat{\phi}]. \tag{2.68}$$

Exercise: demonstrate that under space-time translation by a:

$$e^{i\mathbf{Pa}} \phi(x) e^{-i\mathbf{Pa}} = \phi(x + a). \tag{2.69}$$

2.5.1 Causality

The quantum field $\phi(x)$ is Hermitian and local. It is therefore an observable. We can show that it satisfies the property of eq. (1.86):

$$\begin{aligned}
[\phi(x), \phi(y)] &= \int \frac{d^3p}{(2\pi)^3\sqrt{2\omega_p}} \int \frac{d^3q}{(2\pi)^3\sqrt{2\omega_q}} \left(e^{-ipx} e^{iqy} [a_{\mathbf{p}}, a^\dagger_{\mathbf{q}}] + e^{ipx} e^{-iqy} [a^\dagger_{\mathbf{p}}, a_{\mathbf{q}}] \right) \\
&= \int \frac{d^3p}{(2\pi)^3\sqrt{2\omega_p}} \int \frac{d^3q}{(2\pi)^3\sqrt{2\omega_q}} \left(e^{-ipx} e^{iqy} - e^{ipx} e^{-iqy} \right) (2\pi)^3 \delta(\mathbf{p} - \mathbf{q}) \\
&= \int \frac{d^3p}{(2\pi)^3 2\omega_p} \left(e^{-ip(x-y)} - e^{ip(x-y)} \right) \equiv \Delta(x - y) - \Delta(y - x). \qquad (2.70)
\end{aligned}$$

There are several ways to show that this integral vanishes for space-like separations of x, y: $(x-y)^2 < 0$. One way to show this is using eq. (2.60) which shows that $\Delta(x)$

is Lorentz invariant under proper orthochronous Lorentz transformations. However if $(x-y)^2 < 0$ we can always choose a reference frame where $x'_0 = y'_0$. In this frame, it is trivial to show that $\Delta(x-y) = \Delta(y-x)$ and therefore $[\phi(x), \phi(y)] = 0$. But since this commutator is Lorentz invariant, the same result should hold in all reference frames.

Microcausality ensures that any observable constructed as a function of this causal field is also causal: there is no interference between measurements of this local observable in causally disconnected regions of space-time.

2.6 Canonical quantization of the complex KG field

Let us consider now a complex scalar field $\phi(x) \neq \phi^\dagger(x)$. The Lagrangian is

$$\mathcal{L} = \partial^\mu \phi^\dagger(x) g_{\mu\nu} \partial^\nu \phi(x) - m^2 \phi(x)^\dagger \phi(x), \tag{2.71}$$

and the Euler-Lagrange equations

$$(\nabla^2 + m^2)\phi = 0, \quad (\nabla^2 + m^2)\phi^\dagger = 0. \tag{2.72}$$

We have seen in Section 2.2 that this Lagrangian is invariant under a rephasing of the field $\phi(x) = e^{i\alpha}\phi(x)$, which implies that there is a conserved current:

$$j^\mu = i\left(\partial_\mu \phi^\dagger \phi - \phi^\dagger \partial^\mu \phi\right). \tag{2.73}$$

We can proceed with the canonical method to quantize this theory. [7] The momenta associated with the ϕ and $\phi^\dagger$ fields are:

$$\pi(t, \mathbf{x}) = \frac{\partial L}{\partial \dot{\phi}(t, \mathbf{x})} = \dot{\phi}^\dagger(t, \mathbf{x}), \qquad \pi^\dagger(t, \mathbf{x}) = \frac{\partial L}{\partial \dot{\phi}^\dagger(t, \mathbf{x})} = \dot{\phi}(t, \mathbf{x}), \tag{2.74}$$

and the equal time commutation relations:

$$[\phi(t, \mathbf{x}), \pi(t, \mathbf{x}')] = i\delta(\mathbf{x} - \mathbf{x}'), \qquad [\phi^\dagger(t, \mathbf{x}), \pi^\dagger(t, \mathbf{x}')] = i\delta(\mathbf{x} - \mathbf{x}'), \tag{2.75}$$

all the commutators of any two $(\pi, \pi^\dagger, \phi, \phi^\dagger)$ vanish. We can go to Fourier space and write:

$$\phi(t, \mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}} \tilde{\phi}(t, \mathbf{p}), \qquad \pi(t, \mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}} \tilde{\pi}(t, \mathbf{p}), \tag{2.76}$$

$$\phi^\dagger(t, \mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}} \tilde{\phi}^\dagger(t, -\mathbf{p}), \qquad \pi^\dagger(t, \mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}} \tilde{\pi}^\dagger(t, -\mathbf{p}). \tag{2.77}$$

[7] It is possible to separate the real and imaginary parts and quantize them independently as real KG fields (leading to $a_{R\mathbf{p}}, a_{I\mathbf{p}}$), with the identification $a_\mathbf{p} \equiv a_{R\mathbf{p}} + i a_{I\mathbf{p}}$, $b_\mathbf{p} \equiv a_{R\mathbf{p}} - i a_{I\mathbf{p}}$ and noting that $\pi = \pi_R - i\pi_I$.

The non-vanishing commutation relations in Fourier space are

$$[\tilde{\phi}(t,\mathbf{p}),\tilde{\pi}(t,-\mathbf{p}')] = i(2\pi)^3\delta(\mathbf{p}-\mathbf{p}'), \quad [\tilde{\phi}^\dagger(t,\mathbf{p}),\tilde{\pi}^\dagger(t,-\mathbf{p}')] = i(2\pi)^3\delta(\mathbf{p}-\mathbf{p}') \tag{2.78}$$

We can again take the Hamiltonian to a familiar form by the following change of variables:

$$a_\mathbf{p} \equiv \frac{1}{\sqrt{2\omega_p}}\left[i\tilde{\pi}^\dagger(t,-\mathbf{p}) + \omega_p\tilde{\phi}(t,\mathbf{p})\right], \quad b_\mathbf{p} \equiv \frac{1}{\sqrt{2\omega_p}}\left[i\tilde{\pi}(t,\mathbf{p}) + \omega_p\tilde{\phi}^\dagger(t,-\mathbf{p})\right]. \tag{2.79}$$

From which we have

$$\tilde{\phi}(t,\mathbf{p}) = \frac{1}{\sqrt{2\omega_p}}(a_\mathbf{p}(t) + b^\dagger_{-\mathbf{p}}(t)), \quad \tilde{\pi}(t,\mathbf{p}) = -i\sqrt{\frac{\omega_p}{2}}(b_\mathbf{p}(t) - a^\dagger_{-\mathbf{p}}(t)), \tag{2.80}$$

which reduce to the expressions eqs.(2.42) for a real field if $b = a$. It is easy to show that these operators satisfy the raising/lowering operator commutation relations

$$[a_\mathbf{p}, a^\dagger_{\mathbf{p}'}] = (2\pi)^3\delta(\mathbf{p}-\mathbf{p}'), \quad [b_\mathbf{p}, b^\dagger_{\mathbf{p}'}] = (2\pi)^3\delta(\mathbf{p}-\mathbf{p}'), \tag{2.81}$$

while all other combinations commute. The Hamiltonian simplifies to:

$$:\hat{H}: = \int \frac{d^3p}{(2\pi)^3}\sqrt{\mathbf{p}^2+m^2}\left(a^\dagger_\mathbf{p}a_\mathbf{p} + b^\dagger_\mathbf{p}b_\mathbf{p}\right), \tag{2.82}$$

where we have subtracted the vacuum energy terms. The Hamiltonian therefore corresponds to two infinite sets of harmonic oscillators that have the same particle interpretation as in the real case, but now there are two particles with the same energy $\sqrt{\mathbf{p}^2+m^2}$, one created by $a^\dagger_\mathbf{p}$ and the other by $b^\dagger_\mathbf{p}$. Both are bosons, since wave functions of two or more particles are symmetric, but they are not the same particles. This can be seen by looking at the conserved charge:

$$\begin{aligned}\hat{Q} &= \int d^3x j^0 = \int d^3x i\left(\pi\phi - \pi^\dagger\phi^\dagger\right) \qquad (2.83)\\ &= i\int\frac{d^3p}{(2\pi)^3}\left(\tilde{\pi}(t,-\mathbf{p})\tilde{\phi}(t,\mathbf{p}) - \tilde{\pi}^\dagger(t,-\mathbf{p})\tilde{\phi}^\dagger(t,\mathbf{p})\right) = \int\frac{d^3p}{(2\pi)^3}\left(b^\dagger_\mathbf{p}b_\mathbf{p} - a^\dagger_\mathbf{p}a_\mathbf{p}\right)\end{aligned}$$

We see that the total charge of a state counts the number of b particles minus the number of a particles, therefore the charge of the a particles is opposite to that of the b particles. The b particles are therefore antiparticles of the a particles and viceversa: they have the same mass and momentum but opposite charge. We can construct one-particle and antiparticle states by:

$$|p\rangle_+ = \sqrt{2E_p}a^\dagger_\mathbf{p}|0\rangle, \qquad |p\rangle_- = \sqrt{2E_p}b^\dagger_\mathbf{p}|0\rangle. \tag{2.84}$$

Two observations are in order. We will see that when a charged scalar particle is coupled to the electromagnetic field, the electric charge coincides with the one we just found. That antiparticles must have opposite charge applies to the electromagnetic charge and to all conserved charges (gauge or global). The existence of antiparticles, before they were discovered,[8] was first formulated by Dirac for spin 1/2 particles

[8] The first antiparticle seen was the positron by C.D. Anderson.

and was one of the most spectacular and deep implications of the combination of relativity and quantum mechanics.

Exercise: show that $Qa_{\mathbf{p}}^\dagger|0\rangle = a_{\mathbf{p}}^\dagger|0\rangle$*,* $Qb_{\mathbf{p}}^\dagger|0\rangle = -b_{\mathbf{p}}^\dagger|0\rangle$*, which means the states* $a_{\mathbf{p}}^\dagger|0\rangle$ *and* $b_{\mathbf{p}}^\dagger|0\rangle$ *have well defined and oposite charges.*

The time dependence of the a, b operators can be derived as in the real case from the eqs. (2.65) and the Heisenberg quantum field is finally:

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3\sqrt{2E_p}} \left[a_{\mathbf{p}}(0)e^{-ipx} + b_{\mathbf{p}}^\dagger(0)e^{ipx}\right]. \tag{2.85}$$

Therefore ϕ creates a anti-particle or destroys an particle, localized at $\mathbf{x}$, and carries therefore charge -1, while $\phi^\dagger(x)$ does the oposite and carries charge +1.

Exercise: show that $[Q, \phi(x)] = -\phi(x)$ *and* $[Q, \phi^\dagger(x)] = \phi^\dagger(x)$ *and use this to show that given any state with a well-defined charge* $Q|\Psi_q\rangle = q|\Psi_q\rangle$*, the state* $\phi^\dagger(x)|\Psi_q\rangle$ *has charge* $q = +1$*.*

As in the real case, the existence of two terms is fundamental to ensure micro-causality – that is the commutation of the fields in causally disconnected points:

$$[\phi(x), \phi^\dagger(y)] = 0, \quad (x-y)^2 < 0 \tag{2.86}$$

Exercise: demonstrate eq. (2.86).

Note that the field is not observable in this case since it is not Hermitian, but relation eq. (2.86) ensures that the same will be true for any Hermitian operator built out of the quantum field and its complex conjugate.

The quantization of the KG field has resulted in a Fock space with bosonic and scalar particles, meaning that they transform trivially under Lorentz transformations, i.e. only the coordinates of the fields change. Particles of this type include all the mesonic states (formed by quark-antiquark bound states) such as pions and kaons. The neutral pion, π^0, is its own antiparticle and is represented by a real KG field. The recently discovered Higgs particle might be the first elementary particle of this type. The charged pions π^+, π^- are a particle/antiparticle pair represented by a complex KG field. The neutral kaon, K^0, is not its own antiparticle, because it carries an almost conserved charge, i.e. strangeness, and therefore must have an antiparticle with opposite strangeness: the $\bar{K}^0$ meson. Both are therefore represented by complex KG fields.

The field theory we have quantized has led us to an infinite set of harmonic oscillators that have enabled us to understand what is a one-particle state of energy and momentum $(E, \mathbf{p})$ associated to this field: namely it is the state created from the vacuum by the creation operator of frequency, $\omega_p = E$. The theory can be solved exactly because it is a theory of free particles. The Lagrangian cannot be as simple as eq. (2.5) if it has to represent particle interactions.

III. Interacting Quantum Fields: Particle scattering

The theory of the real KG field generically will contain self-interactions. There is no reason to consider only a quadratic Lagrangian of the form of eq. (2.5),

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\, g^{\mu\nu}\partial_\nu\phi - \frac{1}{2}m^2\phi^2.$$

We could add terms that contain arbitrary powers of ϕ and $\partial_\mu\phi g^{\mu\nu}\partial_\nu\phi$, which are also Lorentz-invariant:

$$-\Delta\mathcal{L} = \text{Polynomial of}[\phi(x), \partial_\mu\phi\partial^\mu\phi] = \sum_{n=3}^{\infty}\lambda_n\phi^n + \sum_{m=2}^{\infty}\alpha_m(\partial_\mu\phi\partial^\mu\phi)^m + ... \quad (3.2)$$

Obviously, in the limit of small couplings $\lambda_n, \alpha_m \to 0$ we recover the free theory.

The action, that is the integral of the Lagrangian density over space-time has the same units as $\hbar$ and therefore has no units in our natural system of units. This means that the Lagrangian density must have units of energy^4,

$$[\mathcal{L}] = 4. \quad (3.3)$$

Since $[\partial_\mu] = 1$, the first quadratic term in eq. (2.5) then implies

$$[\phi] = 1, \quad (3.4)$$

and therefore $[m] = 1$ and

$$[\lambda_n] = 4 - n, [\alpha_m] = 4 - 4m, ... \quad (3.5)$$

The terms that contain couplings that have a dimension of energy to a positive power are called *relevant*, the only one in the polynomial above is that corresponding to $n = 3$. The effect of these terms in scattering amplitudes are expected to be of order $\mathcal{O}(\frac{\lambda_3}{E})$, where E is the typical energy of the process. One would then expect that they become irrelevant at high enough energies, $E \gg \lambda_3$ and relevant only at low energies. For $E > m$, we can ensure that they are small provided $\lambda_3 \ll m < E$. The terms with dimensionless couplings are called *marginal*, ie. the case of $n = 4$ above. The effect of such a term will be small provided $\lambda_4 \ll 1$, independent of the energy. Finally there are many terms with couplings that have negative energy dimensions ($n > 4, m > 1$ above). They are the *irrelevant* terms. On dimensional grounds, those terms are expected to contribute to the scattering amplitudes as $\mathcal{O}(\lambda_n E^{n-4}, \alpha_m E^{4m-4})$ so they will be significant at high energies and irrelevant at low energies, ie. for $E^{n-4} < \lambda_n^{-1}, E^{4m-4} < \alpha_m^{-1}$.

We will see in the final chapter the Wilsonian interpretation of these irrelevant terms: they represent the effect at low energies of new particles with a characteristic mass scale, Λ, such that $\Lambda^{n-4}\lambda_n = O(1)$. At energies $E < \Lambda$, these new particles cannot be produced in the scattering of the ϕ particles, because they are too heavy, but they can modify the interactions of the light particles represented by ϕ in the form of irrelevant terms. At sufficiently low energies, $E \ll \Lambda$ we can nevertheless neglect all irrelevant terms and keep only the relevant and marginal terms. A theory that includes only those terms is called *renormalizable*.

Let us set $\lambda_3 = 0$ (for example this could be justified by assuming a discrete symmetry $\phi \to -\phi$) an consider the only renormalizable marginal interaction possible in this theory:

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi(x) g^{\mu\nu} \partial_\nu \phi(x) - \frac{1}{2} m^2 \phi(x)^2 - \frac{\lambda}{4!}\phi^4, \tag{3.6}$$

where we have added a 4! which is a convenient normalization.

The theory including this term is no longer a set of decoupled harmonic oscillators. The Hamiltonian has an extra term given by

$$\hat{H} = \hat{H}_0 + \hat{H}_{\text{int}}, \qquad \hat{H}_{\text{int}} = \frac{\lambda}{4!}\int d^3x \hat{\phi}(x)^4. \tag{3.7}$$

Still the exact solutions in the limit $\lambda \to 0$ are the starting point of perturbation theory.

Firstly, we will set up the formulation of scattering processes. This will involve the following steps:

1. Define the S-matrix, that is the time-evolution operator in the interaction picture, that contains all the physical information about a scattering process
2. Find a perturbative approximation to the S-matrix: the so-called Dyson series
3. Establish a methodology to evaluate the terms in this expansion using the Wick's theorem and Feynman rules
4. Relation of the S-matrix elements with observables: cross-sections and decay widths

3.1 Particle scattering: the S-matrix

The scattering process we want to describe is the evolution of a system described by the canonically quantized field theory of eq. (3.6) from some initial time $t = -T$ when the system consists of a bunch of one-particle states very far apart so that we can assume they are free, see Fig. 7. Obviously this requires them to be wave-packets and not just momentum states, since otherwise they will be completely delocalized. The particles approach a region where they interact and scatter. At some large asymptotic time $t = T$ they have evolved again into bunch of particles sufficiently far apart so that they can be considered free again. For the time being we will ignore

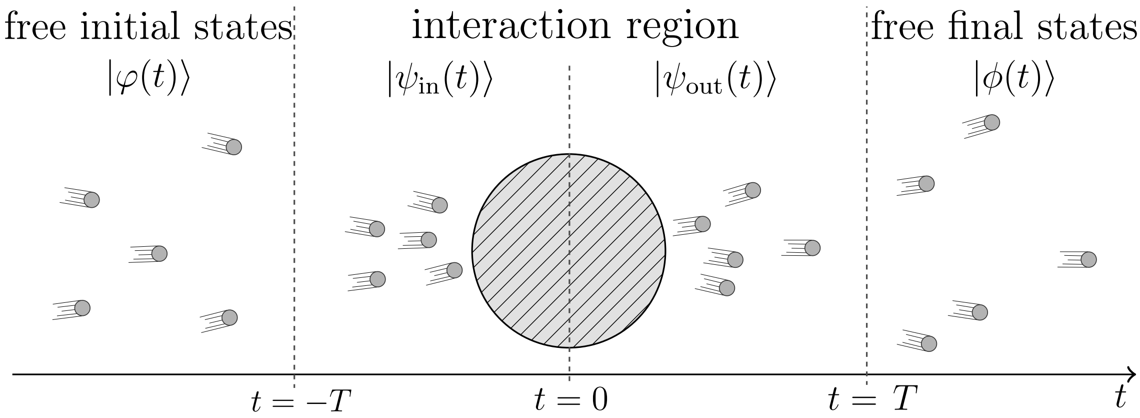


Figure 7: Scattering event

the difficulty of precisely justifying the assumption of particles being free when they are far apart and assume it is the case.

The *in-state* is therefore a solution of the full Schrödinger equation with the boundary condition:

$$i\frac{d}{dt}|\psi_{in}\rangle = \hat{H}|\psi_{in}\rangle, \quad \lim_{t\to -T}|\psi_{in}(t)\rangle = |\varphi(-T)\rangle, \tag{3.8}$$

and similarly

$$i\frac{d}{dt}|\psi_{out}\rangle = \hat{H}|\psi_{out}\rangle, \quad \lim_{t\to T}|\psi_{out}(t)\rangle = |\phi(T)\rangle, \tag{3.9}$$

$|\varphi(t)\rangle$ and $|\phi(t)\rangle$ for large $|t|$ are solutions of the free theory:

$$i\frac{d}{dt}|\varphi\rangle = \hat{H}_0|\varphi\rangle, \quad i\frac{d}{dt}|\phi\rangle = \hat{H}_0|\phi\rangle. \tag{3.10}$$

The probability that the in-state becomes the out-state is therefore

$$|\langle\psi_{out}|\psi_{in}\rangle|^2, \tag{3.11}$$

that can be evaluated at any time for example at $t=0$:

$$|\langle\psi_{out}(0)|\psi_{in}(0)\rangle|^2. \tag{3.12}$$

Using the time-evolution operator

$$\begin{aligned}\psi_{in}(0) &= U(0,-T)\psi_{in}(-T),\\ \psi_{out}(0) &= U(0,T)\psi_{out}(T),\end{aligned} \tag{3.13}$$

and since $U(t,t') = U^\dagger(t',t)$ and $U(t,t')U(t',t'') = U(t,t'')$ we have

$$\begin{aligned}\langle\psi_{out}(0)|\psi_{in}(0)\rangle &= \langle\psi_{out}(T)|U(T,0)U(0,-T)|\psi_{in}(-T)\rangle \\ &= \langle\phi(T)|U(T,-T)|\varphi(-T)\rangle. \end{aligned} \tag{3.14}$$

We can finally rewrite this amplitude in terms of the free-Hamiltonian Heisenberg-picture free states:

$$|\phi\rangle_H = |\phi(0)\rangle, \quad |\varphi\rangle_H = |\varphi(0)\rangle. \tag{3.15}$$

Using now the free time-evolution operator, $U_0(t,t')$, that evolves the free fields:

$$\begin{aligned}|\varphi(-T)\rangle &= U_0(-T,0)|\varphi\rangle_H, \\ |\phi(T)\rangle &= U_0(T,0)|\phi\rangle_H, \end{aligned} \tag{3.16}$$

we arrive at

$$\begin{aligned}\langle\psi_{out}(0)|\psi_{in}(0)\rangle &= \langle\phi_H|U_0(0,T)U(T,-T)U_0(-T,0)|\varphi_H\rangle| = \langle\phi_H|U_I(T,-T)|\varphi_H\rangle \\ &\equiv \langle\phi_H|S|\varphi_H\rangle. \end{aligned} \tag{3.17}$$

The S-matrix is the operator $U_I(T,-T)$, that is the time-evolution operator in the interaction picture:

$$U_I(t,t') = U_0(0,t)U(t,t')U_0(t',0) = e^{iH_0t}U(t,t')e^{-iH_0t'}. \tag{3.18}$$

By comparing eq. (3.17) with eq. (3.11), it is clear what we have achieved. Instead of in and out states, we can express the transition probability in terms of the matrix elements of a complex unitary operator on the basis of simple free-theory Heisenberg states that could represent, for example, two colliding particles with momenta $\mathbf{p}_1, \mathbf{p}_2$:

$$|\psi_H\rangle = |\mathbf{p}_1, \mathbf{p}_2\rangle = \sqrt{2\omega_{p_1}}\sqrt{2\omega_{p_2}}a^\dagger_{\mathbf{p}_1}a^\dagger_{\mathbf{p}_1}|0\rangle. \tag{3.19}$$

3.2 Dyson series

From eq. (3.18), it is easy to show that the time-evolution operator in the interaction picture satisfies the equation:

$$i\frac{d}{dt}U_I(t,t') = \hat{H}_I(t)U_I(t,t'), \tag{3.20}$$

where

$$\hat{H}_I(t) = e^{i\hat{H}_0t}\hat{H}_{\rm int}e^{-i\hat{H}_0t}, \tag{3.21}$$

is the Hamiltonian in the interaction picture. The initial condition is $U_I(t', t') = 1$. Let us assume that the parameter λ is small so that we can write down the solution to eq. (3.20) in a Taylor expansion in λ:

$$U_I(t, t') = \sum_n U_I^{(n)}(t, t'), \tag{3.22}$$

where $U_I^{(n)} = \mathcal{O}(\lambda^n)$.

Substituting this series in eq. (3.20) and taking into account that $\hat{H}_I$ is $\mathcal{O}(\lambda)$ we obtain order by order the following equations:

$$\begin{aligned}
i\frac{d}{dt}U_I^{(0)}(t, t') &= 0, \\
i\frac{d}{dt}U_I^{(1)}(t, t') &= \hat{H}_I(t)U_I^{(0)}(t, t'), \\
i\frac{d}{dt}U_I^{(2)}(t, t') &= \hat{H}_I(t)U_I^{(1)}(t, t'), \\
&\vdots
\end{aligned} \tag{3.23}$$

and $U_I^{(0)}(t', t') = 1$ and $U_I^{(n)}(t', t') = 0$ for $n > 0$.

The equations can be easily integrated:

$$\begin{aligned}
U_I^{(0)}(t, t') &= 1, \\
U_I^{(1)}(t, t') &= -i\int_{t'}^{t} dt_1 \hat{H}_I(t_1), \\
U_I^{(2)}(t, t') &= (-i)^2 \int_{t'}^{t} dt_1 \hat{H}_I(t_1) \int_{t'}^{t_1} dt_2 \hat{H}_I(t_2), \\
&\vdots
\end{aligned}$$

We define the time-ordered product, $T()$, as the operation that takes a product of fields at different times in order of decreasing times. Therefore

$$T(\hat{H}_I(t_1)\hat{H}_I(t_2)) = \theta(t_1 - t_2)\hat{H}_I(t_1)\hat{H}_I(t_2)) + \theta(t_2 - t_1)\hat{H}_I(t_2)\hat{H}_I(t_1). \tag{3.24}$$

We can rewrite

$$\begin{aligned}
\int_{t'}^{t} dt_1 \hat{H}_I(t_1) \int_{t'}^{t_1} dt_2 \hat{H}_I(t_2) &= \int_{t'}^{t} dt_1 \int_{t'}^{t} dt_2 \hat{H}_I(t_1)\hat{H}_I(t_2)\theta(t_1 - t_2) \\
&= \frac{1}{2}\int_{t'}^{t} dt_1 \int_{t'}^{t} dt_2 T(\hat{H}_I(t_1)\hat{H}_I(t_2)),
\end{aligned} \tag{3.25}$$

and more generally

$$\int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2 ... \int_{t'}^{t_{n-1}} dt_n \hat{H}_I(t_1)\hat{H}_I(t_2)...\hat{H}_I(t_n) = \frac{1}{n!}\int_{t'}^{t} dt_1 ... \int_{t'}^{t} dt_n T(\hat{H}_I(t_1)\hat{H}_I(t_2)...\hat{H}_I(t_n)). \tag{3.26}$$

Finally,

$$U_I(t,t') = \sum_n \frac{(-i)^n}{n!} \int_{t'}^{t} dt_1 ... \int_{t'}^{t} dt_n T(\hat{H}_I(t_1)\hat{H}_I(t_2)...\hat{H}_I(t_n)) \equiv Texp\left(-i\int_{t'}^{t} H_I(z)dz\right). \tag{3.27}$$

This is Dyson's series of the S-matrix.

3.3 Wick's theorem

Using Dyson's series, the evaluation of the S-matrix element requires that we compute

$$\langle \phi_H | T(\hat{H}_I(t_1)\hat{H}_I(t_2)...\hat{H}_I(t_n)) | \psi_H \rangle, \tag{3.28}$$

where the initial,$|\psi_H\rangle$, and final state, $|\phi_H\rangle$, will be multiparticle free Heisenberg picture states constructed from the vacuum with creation operators. For example if the initial state are two particles approaching each other with momenta, $\mathbf{p}_1$ and $\mathbf{p}_2$:

$$|\psi_H\rangle = \sqrt{2\omega_{p_1}}\sqrt{2\omega_{p_2}} a^\dagger_{\mathbf{p_1}} a^\dagger_{\mathbf{p_2}} |0\rangle. \tag{3.29}$$

The Hamiltonians are functions of the quantum fields and will also bring in powers of creation and annihilation operators. As usual, this computation of the expectation values of eq. (3.28) can be simplified if we take all annihilation operators to the right (that is to a normal order). Wick's theorem provides a useful relation between time-ordered and normal-ordered products of operators.

3.3.1 Feynman's propagator

Let's first consider the time-ordered product of two free quantum fields:

$$T(\hat{\phi}(x)\hat{\phi}(y)). \tag{3.30}$$

Let us call $\hat{\phi} = \phi^+ + \phi^-$, where ϕ^+ contains the $\hat{a}$ term and ϕ^- contains the $\hat{a}^\dagger$, that is

$$\phi^+(x) \equiv \int \frac{d^3p}{(2\pi)^2\sqrt{2\omega_p}} a_p e^{-ipx}, \quad \phi^-(x) \equiv \int \frac{d^3p}{(2\pi)^2\sqrt{2\omega_p}} a^\dagger_p e^{ipx}. \tag{3.31}$$

Since

$$T(\hat{\phi}(x)\hat{\phi}(y)) \;=\; \theta(x_0-y_0)\hat{\phi}(x)\hat{\phi}(y) + \theta(y_0-x_0)\hat{\phi}(y)\hat{\phi}(x), \tag{3.32}$$

we can use

$$\begin{aligned}\hat{\phi}(x)\hat{\phi}(y) &= \phi^+(x)\phi^+(y) + \phi^-(x)\phi^-(y) + \phi^-(x)\phi^+(y) + \phi^-(y)\phi^+(x) + [\phi^+(x),\phi^-(y)] \\ &= \,:\phi(x)\phi(y): + [\phi^+(x),\phi^-(y)],\end{aligned} \tag{3.33}$$

and the fact that fields under the normal-ordered symbol commute, i.e. $:\hat{\phi}(x)\hat{\phi}(y): = :\hat{\phi}(y)\hat{\phi}(x):$. Combining the regions $x_0 > y_0$ and $y_0 > x_0$ it follows

$$\begin{aligned} T(\hat{\phi}(x)\hat{\phi}(y)) =: \hat{\phi}(x)\hat{\phi}(y): \;&+\; \theta(x_0 - y_0)[\phi^+(x), \phi^-(y)] \\ &+\; \theta(y_0 - x_0)[\phi^+(y), \phi^-(x)] \end{aligned} \tag{3.34}$$

We define:

$$\contraction{}{\hat{\phi}}{(x)}{\hat{\phi}}\hat{\phi}(x)\hat{\phi}(y) \equiv \theta(x_0 - y_0)[\phi^+(x), \phi^-(y)] + \theta(y_0 - x_0)[\phi^+(y), \phi^-(x)]. \tag{3.35}$$

From eqs. (3.31), we obtain

$$\begin{aligned} [\phi^+(x), \phi^-(y)] &= \int \frac{d^3p}{(2\pi)^3\sqrt{2\omega_p}} \int \frac{d^3q}{(2\pi)^3\sqrt{2\omega_q}} e^{-ipx} e^{iqy} [a_p, a_q^\dagger] \\ &= \int \frac{d^3p}{(2\pi)^3 2\omega_p} e^{-ip(x-y)}, \end{aligned} \tag{3.36}$$

therefore

$$\begin{aligned} \contraction{}{\hat{\phi}}{(x)}{\hat{\phi}}\hat{\phi}(x)\hat{\phi}(y) &= \theta(x_0 - y_0) \int \frac{d^3p}{(2\pi)^3 2\omega_p} e^{-ip(x-y)} + \theta(y_0 - x_0) \int \frac{d^3p}{(2\pi)^3 2\omega_p} e^{ip(x-y)}, \\ &= \lim_{\epsilon \to 0} \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\epsilon} e^{-ip(x-y)}, \end{aligned} \tag{3.37}$$

where the final equality is easy to derive from a complex contour integration of p_0 shown in Fig. 8: in the lower half-plane for $x_0 > y_0$ and the upper half for $y_0 > x_0$. The pole $p_0 = \omega_p - i\epsilon$ is inside the integration region for the former and $p_0 = -\omega + i\epsilon$ for the latter.

The vacuum expectation value of the time-ordered product of two free fields is called Feynman's propagator:

$$\Delta_F(x - y) \equiv \langle 0|T(\hat{\phi}(x)\hat{\phi}(y))|0\rangle = \contraction{}{\hat{\phi}}{(x)}{\hat{\phi}}\hat{\phi}(x)\hat{\phi}(y). \tag{3.38}$$

It is easy to show that Feynman's propagator is the Green's function of the Klein-Gordon equation:

$$\left(\partial_\mu \partial^\mu + m^2\right) \Delta_F(x - y) = -i\delta(x - y). \tag{3.39}$$

Feynman's propagator is the building block in more complex expectation values of T-ordered products. Wick's theorem establishes that a general T-product of fields $\hat{\phi}_i = \hat{\phi}(x_i)$ can be written as

$$\begin{aligned} T\hat{\phi}_1....\hat{\phi}_n =: \hat{\phi}_1...\hat{\phi}_n : \;+\;& \text{all possible terms with pairs of fields} \\ & \text{contracted and normal ordered otherwise,} \end{aligned} \tag{3.40}$$

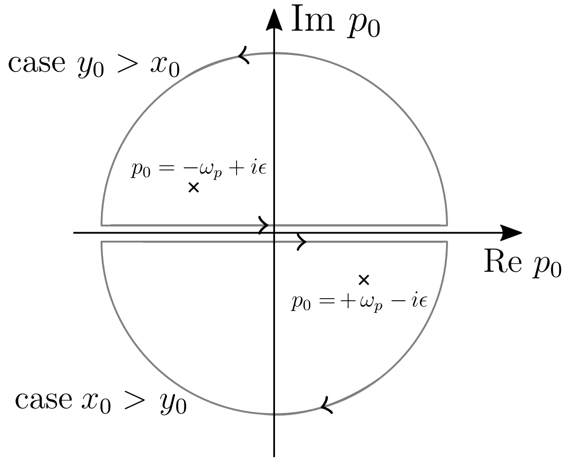


Figure 8: Contour integration of p_0.

where the terms with just one contraction are for example

$$\overbracket{\hat{\phi}_1\hat{\phi}_2} : \hat{\phi}_3....\hat{\phi}_n : + \overbracket{\hat{\phi}_1\hat{\phi}_3} : \hat{\phi}_2...\hat{\phi}_n : + ... \tag{3.41}$$

Operators on the right-hand side are all normal-ordered, and therefore will cancel the vacuum and give no contribution to the vacuum expectation value unless they are the identity, ie. the contribution corresponding to all fields being contracted two by two.

The proof can be derived by induction. Let us assume that it holds for a product with $n-1$ fields, then assuming without loss of generality that $x_1^0 > x_i^0$:

$$T\hat{\phi}_1....\hat{\phi}_n = \hat{\phi}_1 T\hat{\phi}_2....\hat{\phi}_n = \phi_1 W, \qquad W =: \hat{\phi}_2...\hat{\phi}_n : + \text{contracted terms without } \hat{\phi}_1 \tag{3.42}$$

Since W is normal-ordered:

$$\hat{\phi}_1 W = \phi_1^+ W + \phi_1^- W = W\phi_1^+ + [\phi_1^+, W] + \phi_1^- W =: \phi_1 W : + [\phi_1^+, W], \tag{3.43}$$

and $[\phi_1^+, W]$ contains all the contracted terms where $\hat{\phi}_1$ is contracted, which completes the proof.

Example: Consider three fields

$$\begin{aligned} \hat{\phi}_1 T\hat{\phi}_2\hat{\phi}_3 &= \hat{\phi}_1 : \hat{\phi}_2\hat{\phi}_3 : + \hat{\phi}_1 \overbracket{\hat{\phi}_2\hat{\phi}_3} \\ &= [\phi_1^+, : \hat{\phi}_2\hat{\phi}_3 :] + : \hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 : + : \hat{\phi}_1 : \overbracket{\hat{\phi}_2\hat{\phi}_3} \end{aligned} \tag{3.44}$$

Using $[A, BC] = [A, B]C + B[A, C]$

$$\begin{aligned}[\phi_1^+, :\hat{\phi}_2\hat{\phi}_3:] &= [\phi_1^+, \phi_2^+\phi_3^+] + [\phi_1^+, \phi_2^-\phi_3^+] + [\phi_1^+, \phi_3^-\phi_2^+] + [\phi_1^+, \phi_2^-\phi_3^-] \\ &= [\phi_1^+, \phi_2^-]\phi_3^+ + [\phi_1^+, \phi_3^-]\phi_2^+ + [\phi_1^+, \phi_2^-]\phi_3^- + [\phi_1^+, \phi_3^-]\phi_2^- \\ &= [\phi_1^+, \phi_2^-]\phi_3 + [\phi_1^+, \phi_3^-]\phi_2 = \overbrace{\hat{\phi}_1\hat{\phi}_2}\,\hat{\phi}_3 + \overbrace{\hat{\phi}_1\hat{\phi}_3}\,\hat{\phi}_2. \qquad (3.45)\end{aligned}$$

The total is

$$T\hat{\phi}_1\hat{\phi}_2\hat{\phi}_3 = :\hat{\phi}_1\hat{\phi}_2\hat{\phi}_3: + :\hat{\phi}_1: \overbrace{\hat{\phi}_2\hat{\phi}_3} + :\hat{\phi}_2: \overbrace{\hat{\phi}_1\hat{\phi}_3} + :\hat{\phi}_3: \overbrace{\hat{\phi}_1\hat{\phi}_2}\,. \; q.e.d. \qquad (3.46)$$

Wick's theorem shows that the vacuum expectation value of a T-ordered product of operators is therefore the sum of terms each of which is a product of Feynman propagators.

We will also be interested in initial and final states that are not the vacuum. For example we might consider an in state corresponding to k free particles, with momenta $\mathbf{p}_1, ...\mathbf{p}_k$, that scatter into j particles with momenta $\mathbf{q}_1, ...\mathbf{q}_m$. A Dyson series element of this amplitude will be of the form

$$\begin{aligned}\langle \mathbf{q}_1..\mathbf{q}_m|T(\hat{\phi}_1...\hat{\phi}_n)|\mathbf{p}_1...\mathbf{p}_k\rangle = \prod_{j=1,m} \sqrt{2\omega_{q_j}} \qquad (3.47) \\ \times \prod_{i=1,..,k} \sqrt{2\omega_{p_i}}\langle 0|a_{q_1}...a_{g_m}T(\hat{\phi}_1....\hat{\phi}_n)a^\dagger_{p_1}...a^\dagger_{p_k}|0\rangle.\end{aligned}$$

Also in this case, the trick of commuting all the annihilation operators to the right, gives rise to a modified Wick's theorem, where the result of vacuum expectation value is a sum of terms where all fields or creation/annihilation operators are contracted in pairs and no field or operator is left unpaired. Each contraction is the commutator of the contracted objects.

We have now new types of contractions besides those between two fields, eq. (3.37): between a field and an operator a_q, a field and an operator $a^\dagger_p$ or a_q and $a^\dagger_p$:

$$\begin{aligned}\overbrace{a_q\hat{\phi}}(x) \equiv [a_q, \hat{\phi}(x)] &= \frac{1}{\sqrt{2\omega_q}}e^{iqx}, \\ \overbrace{\hat{\phi}\,(x)a^\dagger_p} \equiv [\hat{\phi}(x), a^\dagger_p] &= \frac{1}{\sqrt{2\omega_p}}e^{-ipx}, \\ \overbrace{a_q a^\dagger_p} = [a_q, a^\dagger_p] &= (2\pi)^3\delta(\mathbf{q} - \mathbf{p}). \qquad (3.48)\end{aligned}$$

3.4 Feynman rules

Feynman rules are a series of diagrammatic rules that simplify the computation of the Dyson series expansion of S-matrix elements by making use of the Wick's theorem. The starting point is

$$\langle\beta|T\exp\left(-i\int_{-T}^{T}dtH_I(t)\right)|\alpha\rangle, \tag{3.49}$$

with

$$H_I(t)=e^{iH_0t}\int d\mathbf{x}\ \frac{\lambda}{4!}\hat{\phi}(0,\mathbf{x})^4e^{-iH_0t}=\int d\mathbf{x}\ \frac{\lambda}{4!}\hat{\phi}_I(x)^4, \tag{3.50}$$

where $\hat{\phi}_I(x)$ is the field in the interaction picture:

$$\hat{\phi}_I(x)=e^{iH_0t}\phi(0,\mathbf{x})e^{-iH_0t}. \tag{3.51}$$

Therefore

$$\langle\beta|\hat{S}|\alpha\rangle=\sum_n\frac{1}{n!}(-i)^n\left(\frac{\lambda}{4!}\right)^n\int d^4z_1...\int d^4z_n\langle\beta|T(\hat{\phi}(z_1)^4...\hat{\phi}(z_n)^4)|\alpha\rangle. \tag{3.52}$$

For the external states we take

$$|\alpha\rangle=|\mathbf{p}_1,....,\mathbf{p}_k\rangle,\ \ |\beta\rangle=|\mathbf{q}_1,....,\mathbf{q}_m\rangle. \tag{3.53}$$

The non-vanishing contributions can be shown diagrammatically as the sum of Feynman diagrams constructed as follows:

- Draw the n vertices, z_i, $i=1,...,n$, as points with four lines attached.
- For each external particle draw a line: incoming to the left, outgoing to the right. Each has an associated momentum.
- Link each external particle lines to a line coming out of a vertex so that each external line ends in a vertex. Once all the external lines are linked to a vertex, link the remaining open lines two by two, so that each of those lines connect two vertices.
- Calculate the amplitude of the diagram A_D and the symmetry factor S_D. The contribution of this diagram to the total result is

$$\frac{A_D}{S_D}. \tag{3.54}$$

- Repeat with as many distinct contractions of lines as possible, each giving a different diagram.

Once we have classified all the distinct diagrams, we can calculate the whole amplitude as

$$\sum_D \frac{A_D}{S_D}. \tag{3.55}$$

Feynman rules tell us how to calculate A_D:

- A factor $(-i\lambda)\int d^4 z_i$ for each vertex.
- A Feynman propagator $\Delta_F(z_i - z_j)$ for each line from vertex z_i to z_j.
- A factor $e^{-ip_j z_i}$ for each incoming external line with momenta p_j and linked to the vertex z_i.
- A factor $e^{iq_j z_i}$ for each outgoing external line with momenta q_j and linked to the vertex z_i.

The symmetry factor is given by

$$S_D = \frac{(4!)^n \; n!}{\#\text{contractions leading to the same diagram } D}. \tag{3.56}$$

It is clear that the number of contractions will always include an $n!$ that comes from exchanging vertices since they are all identical.

Example: Let us consider the scattering involving one incoming particle and one outgoing particle

At 0-th order

$$\langle \mathbf{q} | \mathbf{p} \rangle = (2\pi)^3 2\omega_p \delta(\mathbf{q} - \mathbf{p}). \tag{3.57}$$

At 1-st order

$$\frac{-i\lambda}{4!} \int d^4 z \langle \mathbf{q} | T(\hat{\phi}_I(z)^4) | \mathbf{p} \rangle \tag{3.58}$$

According to Wick's theorem, only the fully connected terms contribute since all normal ordered products vanish in the vacuum expectation value. All the distinct possibilities are the following, represented by the diagrams in the

- $3 \times \overbracket{a_q a_p^\dagger}\, \overbracket{\hat{\phi}_z \hat{\phi}_z}\, \overbracket{\hat{\phi}_z \hat{\phi}_z}$, depicted in Fig. 9.

The number of contractions is easy to understand: the first operator at z can be connected to either of the three remaining, while the two left must necessarily be connected. The contribution is

$$A_{D_1} = -i\lambda \cdot (2\pi)^3 2\omega_p \delta(\mathbf{q} - \mathbf{p}) \int d^4 z \; \Delta_F(z - z)^2, \;\; S_{D_1} = \frac{4!}{3} = 8. \tag{3.59}$$

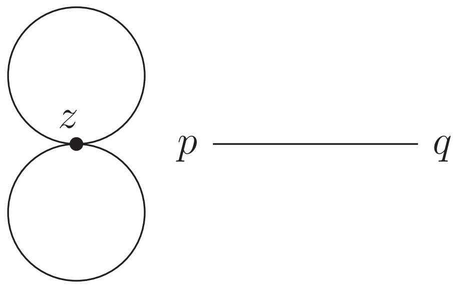


Figure 9: Contraction D_1.

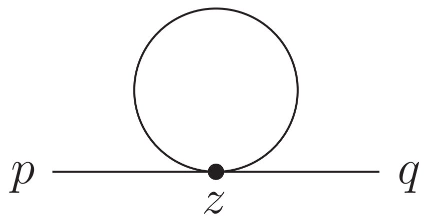


Figure 10: Contraction D_2.

- $12 \times a_q\hat{\phi}_z\,\hat{\phi}_z\hat{\phi}_z\,\hat{\phi}_z a_p^{\dagger}$, depicted in Fig. 10.

In this case, the incoming line of momentum p can be connected to one of the four fields at z, the outgoing line is then connected to one of the remaining three, while the two remaining again need to be connected among themselves. Therefore we have $4 \cdot 3 = 12$ possibilities. The contribution of these terms is

$$A_{D_2} = -i\lambda \cdot \int d^4z \;\; e^{-i(p-q)z}\Delta_F(z-z), \;\; S_{D_2} = \frac{4!}{12} = 2. \tag{3.60}$$

3.5 Cross sections and decay widths

The result of a scattering experiment is usually measured in terms of the cross section, σ.

Let us consider a beam of particles of type A that collide with a target made of particles of type B, see Fig. 11. When the beam crosses, the particles of type A become dispersed due to their interaction with particles B. We can measure the

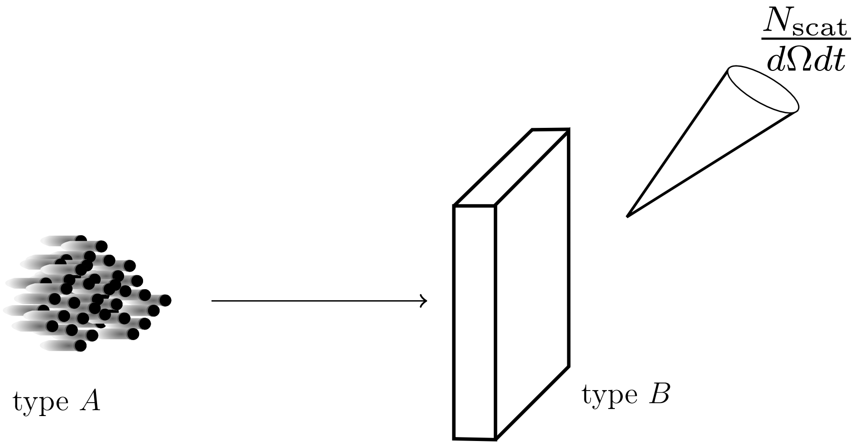


Figure 11: Scattering experiment.

number of particles that are dispersed per unit time

$$\frac{dN_{\text{scat}}}{dt} = \# \text{ of particles dispersed per unit time.} \tag{3.61}$$

We might as well measure the number of particles dispersed into a solid angle, $d\Omega$:

$$\frac{dN_{\text{scat}}}{dt d\Omega}. \tag{3.62}$$

We expect that this number is proportional to the flux of incoming particles A, Φ_A, and the number of target particles B, N_B, on which they can scatter. If we compute the number per unit target volume we would expect

$$\frac{dN_{\text{scat}}}{dV dt} \propto \Phi_A \rho_B, \tag{3.63}$$

where ρ_B is the density of target particles. The proportionality factor is what we call the *cross section*:

$$\sigma = \frac{dN_{\text{scat}}}{dV dt} (\Phi_A \rho_B)^{-1}. \tag{3.64}$$

Since the flux is the number of particles crossing a unit area in a unit time, the dimensions of the cross-section is that of an area, or in natural units, E^{-2}.

We expect that the number of scattering events is proportional to the probability of this process:

$$N_{\text{scat}} = \sum_f |\langle f|\hat{S}|i\rangle|^2, \tag{3.65}$$

where

Let us consider an in-state corresponding to two particles of types A and B with momenta, p_A and p_B. If we want the particles to be far apart at $t = -T$, they cannot have perfectly well-defined momenta because if that was the case they would be completely delocalized. Instead we should consider that they are wavepackets of the form

$$|i\rangle = |p_A, p_B\rangle = \int \frac{d^3p_1}{(2\pi)^3 2\omega_{p_1}} \int \frac{d^3p_2}{(2\pi)^3 2\omega_{p_2}} f_A(p_1) f_B(p_2) |p_1, p_2\rangle, \tag{3.66}$$

where the functions $f_A(p)$ and $f_B(p)$ are for example Gaussians centered at the average momenta, p_A and p_B respectively, i.e.

$$f_{A/B}(p) \propto e^{-(p-p_{A/B})^2/2\sigma^2}. \tag{3.67}$$

We assume the states to be properly normalized.

$$\langle p_A, p_B | p_A, p_B \rangle = 1 \to \int d^3p \; |f_{A/B}(p)|^2 = 1. \tag{3.68}$$

The number of scattered events resulting from this in-state would be

$$N_{\text{scat}} = \sum_f |\langle f|\hat{S}|i\rangle|^2, \tag{3.69}$$

where the sum is over the final states.[9]

The final state corresponds to any state that we define as a scattering event. For example if we consider all the events in which the particles of type A and B get scattered, that is they change their momenta with respect to p_A and p_B, then the final state would be for example $|f\rangle = |q_A, q_B\rangle$ for all q_A, q_B distinct from p_A, p_B. But it could also be more complicated and we also count scattering events where the particles A, B disappear and produce particles of a different nature, for example. The final state would have to be defined accordingly. To find the relation we are after, we do not need to specify $|f\rangle$.

On general grounds, the S-matrix can be written as

$$\hat{S} = I + i\hat{T}, \tag{3.70}$$

where I is the identity, since time evolution will always include the possibility that the state remains intact. Obviously the first term does not contribute if $|i\rangle \neq |f\rangle$.

The probability for the process is then

$$\begin{aligned} &\left|\langle f|\hat{T}|i\rangle\right|^2 = \langle f|\hat{T}|p_A, p_B\rangle\langle p_A, p_B|\hat{T}^\dagger|f\rangle = \\ &\int_{q_1}\int_{q_2}\int_{p_1}\int_{p_2} f_A^*(q_1) f_A(p_1) f_B^*(q_2) f_B(p_2) \langle f|\hat{T}|p_1, p_2\rangle\langle q_1, q_2|\hat{T}^\dagger|f\rangle, \end{aligned} \tag{3.71}$$

[9]Note that final states are a continuum and therefore the sum over final states is really an integral.

where

$$\int_p \equiv \int \frac{d^3p}{(2\pi)^3 2\omega_p}. \tag{3.72}$$

The amplitude will always have a delta function ensuring the conservation of energy momentum in the process, that is

$$\langle f|\hat{T}|p_1, p_2\rangle = (2\pi)^4 \delta(p_f - p_1 - p_2)\mathcal{M}(p_f; p_1, p_2). \tag{3.73}$$

Using this we arrive to

$$\begin{aligned}
|\langle f|\hat{T}|p_A, p_B\rangle|^2 &= \int_{q_1}\int_{q_2}\int_{p_1}\int_{p_2} f_A^*(q_1) f_A(p_1) f_B^*(q_2) f_B(p_2) \\
&\qquad \times \langle f|\hat{T}|p_1, p_2\rangle\langle q_1, q_2|\hat{T}^\dagger|f\rangle \\
&= \int_{q_1}\int_{q_2}\int_{p_1}\int_{p_2} f_A^*(q_1) f_A(p_1) f_B^*(q_2) f_B(p_2) \\
&\quad \times (2\pi)^4 \delta(p_F - p_1 - p_2)(2\pi)^4 \delta(p_1 + p_2 - q_1 - q_2) \\
&\quad \times \mathcal{M}(p_f; p_1, p_2)\mathcal{M}^*(p_f; q_1, q_2).
\end{aligned} \tag{3.74}$$

Using the relation

$$(2\pi)^4 \delta(q_1 + q_2 - p_1 - p_2) = \int d^4x \; e^{i(q_1+q_2-p_1-p_2)x}, \tag{3.75}$$

and since $p_1, q_1 \simeq p_A$ and $p_2, q_2 \simeq p_B$, we can approximate

$$\begin{aligned}
|\langle f|\hat{T}|p_A, p_B\rangle|^2 &= (2\pi)^4 \delta(p_f - p_A - p_B)|\mathcal{M}(p_f; p_A, p_B)|^2 \qquad (3.76) \\
&\quad \int_{q_1}\int_{q_2}\int_{p_1}\int_{p_2}\int d^4x \; e^{i(q_1+q_2-p_1-p_2)x} f_A^*(q_1) f_A(p_1) f_B^*(q_2) f_B(p_2) \\
&= (2\pi)^4 \delta(p_f - p_A - p_B)|\mathcal{M}(p_f; p_A, p_B)|^2 \int d^4x |f_A(x)|^2 |f_B(x)|^2,
\end{aligned}$$

where

$$\tilde{f}_{A/B}(x) \equiv \int_p e^{-ipx} f_{A/B}(p). \tag{3.77}$$

Since we want the number of scattering events per unit time and volume, we should consider the integrand of the space-time integral in x and we finally arrive at:

$$\frac{dN_{\text{scat}}}{dV dt} = (2\pi)^4 \delta(p_f - p_A - p_B)|\mathcal{M}(p_f; p_A, p_B)|^2 |\tilde{f}_A(x)|^2 |\tilde{f}_B(x)|^2. \tag{3.78}$$

The wave-function of the one-particle states

$$\langle \mathbf{x}|p_A\rangle(t) = \int_p f_A(p)\sqrt{2\omega_p} e^{-i\omega_p t}\langle \mathbf{x}|\mathbf{p}\rangle \simeq \sqrt{2\omega_{p_A}} \int_p f_A(p) e^{-ipx} = \sqrt{2\omega_{p_A}} \tilde{f}_A(x), \tag{3.79}$$

and similarly for the wave function of the in-state of particle B. Therefore the number density of the particle B, which is at rest, is

$$\rho_B = |\langle \mathbf{x}|p_B\rangle(t)|^2 = 2m_B|\tilde{f}_B(x)|^2. \tag{3.80}$$

The flux of particles A, moving at the speed v_A, can be obtained as

$$\Phi_A = \rho_A v_A = \rho_A \frac{|\mathbf{p}_A|}{\omega_{p_A}} = 2|\mathbf{p}_A||\tilde{f}_A(x)|^2. \tag{3.81}$$

The cross-section is then obtained normalizing eq.(3.78) by the A particle flux and the B number density. If we consider a generic final state with k particles of momenta, $|f\rangle = |q_1, ...q_k\rangle$, final states for all possible values of these momenta will contribute to the total cross-section. Therefore the total cross-section is given by

$$\frac{\sigma = \int \prod_{j=1,...k} \frac{d^3q_j}{(2\pi)^3 2\omega_{q_j}} \ (2\pi)^4 |\mathcal{M}(q_1, .., q_k; p_A, p_B)|^2 \delta^{(4)}(q_1 + ...q_k - p_A - p_B)}{4m_B|\mathbf{p}_A|,} \tag{3.82}$$

where we have used $p_B = (m_B, \mathbf{0})$. This is the result in the B-particle rest frame. We can write the result in a different frame noting

$$m_B|\mathbf{p}_A| = \sqrt{(p_A \cdot p_B)^2 - m_A^2 m_B^2} = \frac{1}{2}\sqrt{\lambda(s, m_A^2, m_B^2)}, \tag{3.83}$$

where

$$\lambda(a, b, c) \equiv a^2 + b^2 + c^2 - 2ab - 2ac - 2bc, \tag{3.84}$$

and

$$s \equiv (p_A + p_B)^2. \tag{3.85}$$

Another important quantity for unstable states is the *decay width*. The initial state in this case corresponds to a one particle state of type A, which is in a wavepacket:

$$|i\rangle = |p_A\rangle = \int_p f_A(p)|p\rangle. \tag{3.86}$$

This particle decays into a final state $|f\rangle$. The decay width is defined as the number of decays per unit volume and time, divided by the density of A particles:

$$\Gamma = \frac{dN_{\text{decay}}/(dtdV)}{\rho_A}, \tag{3.87}$$

As in the previous case,

$$N_{\text{decay}} = |\langle f|\hat{T}|p_A\rangle|^2, \tag{3.88}$$

and a similar analysis as before leads to

$$\frac{dN_{\rm decay}}{dtdV} = (2\pi)^4|\mathcal{M}(p_f;p_A)|^2\delta(p_f-p_A)|f_A(x)|^2, \tag{3.89}$$

while $\rho_A = 2m_A|f_A(x)|^2$ in the rest frame. The total decay width is obtained integrating over all the possible final states. If the decay is to a set of particles with momenta, $q_1, ...q_k$, we obtain

$$\Gamma = \frac{1}{2m_A}\int\prod_{j=1,..,k}\frac{d^3q_j}{(2\pi)^3 2\omega_{q_j}}|\mathcal{M}(q_1,...,q_k;p_A)|^2(2\pi)^4\delta(q_1+...+q_k-p_A). \tag{3.90}$$

The integration over final momenta in eqs. (3.82) and (3.90) is refered to as integration over phase space. In the case of two momenta, the integration can be done easily.

Phase space integration of two particles

Let us consider the collision of two particles A and B in the center of mass frame, that is

$$\mathbf{p}_A + \mathbf{p}_B = 0. \tag{3.91}$$

The final particles have masses m_1 and m_2 and momenta q_1 and q_2. Defining

$$\omega_{q_1} \equiv \sqrt{m_1^2+\mathbf{q_1}^2}, \quad \omega_{q_2} \equiv \sqrt{m_2^2+\mathbf{q_1}^2}, \tag{3.92}$$

then

$$\begin{aligned}
&\int\frac{d^3q_1}{(2\pi)^3 2\sqrt{m_1^2+\mathbf{q_1}^2}}\int\frac{d^3q_2}{(2\pi)^3 2\sqrt{m_2^2+\mathbf{q_2}^2}}(2\pi)^4\delta^{(4)}(q_1+q_2-p_A-p_B)\\
=&\int\frac{d^3q_1}{(2\pi)^2 4\omega_{q_1}\omega_{q_2}}\delta\left(\omega_{q_1}+\omega_{q_2}-E_A-E_B\right) = \int\frac{|\mathbf{q}_1|^2 d|\mathbf{q}_1|d\Omega_{q_1}}{(2\pi)^2 4\omega_{q_1}\omega_{q_2}}\delta\left(\omega_{q_1}+\omega_{q_2}-\sqrt{s}\right)\\
=&\frac{1}{32\pi^2}\frac{\sqrt{\lambda(s,m_1^2,m_2^2)}}{s},
\end{aligned} \tag{3.93}$$

where we have used the relation

$$\delta(\omega_{q_1}+\omega_{q_2}-\sqrt{s}) = \frac{1}{|\mathbf{q}_1|}\frac{\omega_{q_1}\omega_{q_2}}{\omega_{q_1}+\omega_{q_2}}\delta\left(|\mathbf{q}_1|-\sqrt{\frac{\lambda(s,m_1^2,m_2^2)}{4s}}\right). \tag{3.94}$$

The differential cross-section in the center of mass of the process $p_A+p_B \to q_1+q_2$ is, then,

$$\frac{d\sigma}{d\Omega_{q_1}} = \frac{1}{64\pi^2 s}\frac{\sqrt{\lambda(s,m_1^2,m_2^2)}}{\sqrt{\lambda(s,m_A^2,m_B^2)}}\,|\mathcal{M}(q_1,q_2;p_A,p_B)|^2. \tag{3.95}$$

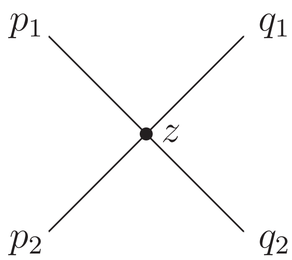


Figure 12: Feynman diagram for two scalar particles scattering.

Example: two-particle scattering in $\lambda\phi^4$ theory

In the scalar theory of eq. (3.6) we can consider the elastic scattering cross-section of two scalar particles of momenta p_1, p_2 to two particles with different momenta q_1, q_2, see Fig. 12. We need to compute the amplitude

$$\langle q_1, q_2|\hat{T}|p_1, p_2\rangle \tag{3.96}$$

The first order in the Dyson series contains just one vertex, and we have two incoming external lines and two outgoing lines. According to the Feynman rules in Section 3.4, we therefore obtain the amplitude

$$\begin{aligned}\langle q_1, q_2|\hat{T}|p_1, p_2\rangle &= (-i\lambda)\int d^4z e^{-ip_1z}e^{-ip_2z}e^{iq_1z}e^{iq_2z} \\ &= -i\lambda(2\pi)^4\delta^{(4)}(p_1+p_2-q_1-q_2),\end{aligned} \tag{3.97}$$

since the symmetry factor of this diagram is 1. Therefore

$$\mathcal{M}(q_1, q_2; p_1, p_2) = -i\lambda, \tag{3.98}$$

and the total cross section, integrating over phase space, is according to eq. (3.95) ($m_A = m_B = m_1 = m_2 = m$):

$$\frac{1}{2}\int d\Omega_{q_1}\frac{d\sigma}{d\Omega_{q_1}} = \frac{\lambda^2}{32\pi s}. \tag{3.99}$$

The factor 1/2 is necessary because in this case the final state particles are identical and so we would be double-counting if we considered as distinct final states with final momenta q_1, q_2 or q_2, q_1. More generally, it is necessary to divide the phase space integration by $1/k!$, when the number of identical particles in the final state is k.

IV. Spin 1/2 Quantum Fields

In previous chapters we have considered scalar fields that transform trivially under Lorentz transformations:

$$x \to x' = \Lambda x, \quad \phi(x) \to \phi'(x') = \phi(x), \tag{4.2}$$

which implies

$$\phi(x) \to \phi(\Lambda^{-1} x). \tag{4.3}$$

These fields cannot represent particles with spin (intrinsic angular momentum), such as the electron or the photon. These particles must be represented by quantum fields that transform non trivially under Lorentz transformations, generically as

$$\phi^a(x) \to D[\Lambda]^a_b \phi^b(\Lambda^{-1} x), \tag{4.4}$$

where $D[\Lambda]$ is some finite-dimensional representation of the Lorentz group, $O(1,3)$. It satisfies the properties:

$$\begin{aligned} D[\Lambda_1 \Lambda_2] &= D[\Lambda_1] D[\Lambda_2], \\ D[\Lambda^{-1}] &= D[\Lambda]^{-1}, \\ D[I] &= I. \end{aligned} \tag{4.5}$$

Being a Lie group, any representation of an infinitesimal transformation can be writen in terms of the generators T^i,

$$D[\Lambda] = I + i \sum_i \alpha_i T^i, \tag{4.6}$$

where α_i are the real continuous parameters that define the Lie group. The generators form the algebra, $\mathfrak{so}(1,3)$.

Particles in nature and their associated quantum fields are *irreducible representations of the Poincaré symmetry* group (the group including Lorentz transformations and translations). An irreducible representation, $D[\Lambda]$, is such that no subspace of states transforms *only* among themselves under all group elements. Irreducible representations are the basic building blocks of arbitrary representations of the group. We can say that a particle/field has the necessary components to transform irreducibly under Poincaré symmetry transformations.

Rotational symmetry and the concept of spin provide simpler examples of how symmetries dictate the nature of particles. Spin, j, is nothing but the half-integer label of the irreducible unitary representations of the rotation group, which have dimension $d = 2j + 1$. The smallest non-trivial representation is therefore spin

$j = 1/2$, which is two-dimensional and whose generators are the Pauli matrices σ_i. A general representation of an element of the group is given by the exponential

$$U = e^{i\frac{\theta_i \sigma_i}{2}}, \tag{4.7}$$

where θ_i are the rotations around the axis i. States with spin 1/2 are described by two component vectors, or spinors, that transform as

$$\Psi \to e^{i\frac{\theta_i \sigma_i}{2}} \Psi. \tag{4.8}$$

Particles with this property include the electron, and all the elementary fermions in the Standard Model.

The three-dimensional representation corresponding to $j = 1$ transforms like a vector under spatial rotations. All elementary interactions involve a bosonic field which transform as vectors under rotations, including the electromagnetic field that mediates electromagnetic interaction, the $W^{\pm}, Z^0$ bosons that mediate the weak interaction, and the gluons that mediate the strong force. All the elementary particles in the Standard Model are irreducible representations of the rotation group corresponding to $j = 0$, $j = 1/2$ and $j = 1$. The quantization of gravity, which is not yet satisfactory, would imply the existence of the graviton, which would be a particle of $j = 2$.

In the context of quantum field theory, we need to consider the representations of the full Lorentz group.

4.1 Irreducible representations of the Lorentz group

A particular example of a representation of the Lorentz group is the four-dimensional representation in space-time coordinate space,

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu, \qquad \mu = 0, 1, 2, 3, \tag{4.9}$$

where the four-dimensional matrices Λ leave the Minkowski metric invariant,

$$\Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu g_{\alpha\beta} = g_{\mu\nu}. \tag{4.10}$$

These matrices Λ, close to the identity, depend on six real parameters. Indeed, a general infinitesimal transformation can be written as

$$\Lambda^{\ \mu}{}_\nu = \delta^\mu_\nu + \omega^\mu{}_\nu + \mathcal{O}(\omega^2). \tag{4.11}$$

Substituting in eq. (4.10), we obtain, neglecting terms of $\mathcal{O}(\omega^2)$:

$$\omega_{\mu\nu} + \omega_{\nu\mu} = 0, \quad \omega_{\mu\nu} \equiv \omega^\alpha{}_\nu g_{\alpha\mu}. \tag{4.12}$$

ω is therefore a four dimensional antisymmetric real matrix which has six independent elements. We can choose these parameters as the elements above the diagonal: ω_{0i} and ω_{ij} with $i < j$ and $i, j = 1, 2, 3$. These parameters are easy to identify with boosts and rotations.

For example an infinitesimal rotation around x axis is implemented by the matrix:

$$\Lambda(R_x(\theta)) = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & & \theta \\ & & -\theta & \end{pmatrix}, \tag{4.13}$$

which corresponds to $\omega^2{}_3 = -\omega_{23} = -\omega^3{}_2 = \omega_{32} = \theta$.

An infinitesimal boost of rapidity β in the x direction,

$$\Lambda(B_x(\beta)) = \begin{pmatrix} 1 & \beta & & \\ \beta & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}, \tag{4.14}$$

corresponds to $\omega^0{}_1 = \omega_{01} = \omega^1{}_0 = -\omega_{10} = \beta$. Therefore we identify the six parameters with three rapidities, β_i, and the Euler angles, θ_i as:

$$\omega_{0i} = -\omega_{i0} = \beta_i, \quad \omega_{ij} = -\epsilon_{ijk}\theta_k. \tag{4.15}$$

We can easily find the generators $J^{\alpha\beta}$ associated to each parameter $\omega_{\alpha\beta}$, since by definition

$$\Lambda(\omega) = I + i \sum_{\alpha\beta=\{01,02,03,12,13,23\}} \omega_{\alpha\beta} J^{\alpha\beta} + \mathcal{O}(\omega^2). \tag{4.16}$$

In this expression $\omega_{\alpha\beta}$ are just numbers, i.e. the parameters of the transformation, not a matrix. It is easy to check that the generators

$$(J^{\alpha\beta})_{\mu\nu} = -i(\delta^\alpha_\mu \delta^\beta_\nu - \delta^\beta_\mu \delta^\alpha_\nu) \tag{4.17}$$

reduce eq. (4.16) to eq. (4.11).

Then

$$\Lambda(\omega) = I + i\beta_i J^{0i} - \frac{i}{2}\theta_k \epsilon_{ijk} J^{ij} \equiv I + i\beta_i K^i + i\theta_k J^k, \tag{4.18}$$

where the sum in $i, j = 1, 2, 3$ and

$$K^i \equiv J^{0i}, \quad J^k \equiv -\frac{1}{2}\epsilon_{ijk} J^{ij}. \tag{4.19}$$

It is straightforward, using eq. (4.19) and (4.17), to show that the generators satisfy the following commutation relations:

$$[J^i, J^j] = i\epsilon_{ijk}J^k, \quad [J^i, K^j] = i\epsilon_{ijk}K^k, \quad [K^i, K^j] = -i\epsilon_{ijk}J^k. \tag{4.20}$$

This is the Lorentz algebra, $\mathfrak{so}(1,3)$.

Exercise: Using eqs. (4.17) and (4.19) find explicit expressions for the generators J_i and K_i and prove the commutation relations.

To classify all the irreducible representations of this algebra, we notice that there are two commuting subalgebras. Defining

$$J^i_+ \equiv \frac{1}{2}\left(J^i + iK^i\right), \qquad J^i_- \equiv \frac{1}{2}\left(J^i - iK^i\right), \tag{4.21}$$

it is easy to see that these combinations satisfy

$$[J^i_+, J^j_+] = i\epsilon_{ijk}J^k_+, \quad [J^i_-, J^j_-] = i\epsilon_{ijk}J^k_-, \quad [J^i_+, J^j_-] = 0. \tag{4.22}$$

Exercise: show that eqs.(4.22) derives from eqs.(4.20).

The set of generators therefore contains two commuting subalgebras that coincide with that of the rotation group, $\mathfrak{so}(3)$ or $\mathfrak{su}(2)$ algebras. The finite-dimesional irreducible representations of the algebra of space rotations are well known. They are labeled by spin, a half-integer, j, and have dimension $d = 2j+1$. From this result we can easily obtain all those of the Lorentz group, as labeled by two half-integers (j_-, j_+) corresponding to the two subalgebras. The dimension is therefore $(2j_- + 1)(2j_+ + 1)$. The standard rotation generators are obtained as $J_i = J_i^+ + J_i^-$ and therefore the standard spin can take any value between $j = |j_+ - j_-|, ..., j_+ + j_-$ as established by the usual Clebsch-Gordan decomposition, as shown in Table 1.

4.2 Spin 1/2 representations

Interestingly we see that there are two different two-dimensional representations of spin 1/2 particles: $(1/2, 0)$ and $(0, 1/2)$. Each is called a *Weyl representation.* The two-dimensional matrices representing the generators of rotations and boosts are the Pauli matrices:

$$\begin{aligned} \left(0, \frac{1}{2}\right) \quad &: \quad J^i_+ = \frac{\sigma_i}{2}, J^i_- = 0 \to J_i = \frac{1}{2}\sigma_i, \quad K_i = \frac{-i}{2}\sigma_i, \\ \left(\frac{1}{2}, 0\right) \quad &: \quad J^i_- = \frac{\sigma_i}{2}, J^i_+ = 0 \to J_i = \frac{1}{2}\sigma_i, \; K_i = \frac{i}{2}\sigma_i. \end{aligned} \tag{4.23}$$

$\mathfrak{so}(1,3)$						
(j_-, j_+)	$(0,0)$	$(\frac{1}{2},0)$	$(0,\frac{1}{2})$	$(\frac{1}{2},\frac{1}{2})$	$(1,0)$	...
$\mathfrak{so}(3)$						
j	0	$\frac{1}{2}$	$\frac{1}{2}$	$1\oplus 0$	1	...

Table 1: Irreducible representations of the Lorentz algebra and the rotation group.

The two-component vectors are called *Weyl spinors*: the $(0,1/2)$ is a right-handed spinor, ψ_R, and the $(1/2,0)$ is a left-handed spinor, ψ_L, which transform under a Lorentz transformation as

$$\psi_R \to e^{\frac{1}{2}(i\theta_j\sigma_j+\beta_j\sigma_j)}\psi_R, \qquad \psi_L \to e^{\frac{1}{2}(i\theta_j\sigma_j-\beta_j\sigma_j)}\psi_L. \tag{4.24}$$

Let us consider fields in these representations. Under a Lorentz transformation:

$$\begin{aligned}\psi_L(x) &\to e^{\frac{1}{2}(i\theta_j\sigma_j-\beta_j\sigma_j)}\psi_L(\Lambda^{-1}x),\\ \psi_R(x) &\to e^{\frac{1}{2}(i\theta_j\sigma_j+\beta_j\sigma_j)}\psi_R(\Lambda^{-1}x).\end{aligned} \tag{4.25}$$

By recalling the way a four vector V^μ transforms under infinitesimal boost β_i and rotation θ_i,

$$\begin{aligned}V^0 &\to V^0+\omega^0{}_iV^i = V^0+\omega_{0i}V^i = V^0+\beta_iV^i,\\ V^i &\to V^i+\omega^i{}_0V^0+\omega^i{}_jV^j = V^0+\omega_{0i}V^0-\omega_{ij}V^j = V^0+\beta_iV^0+\epsilon_{ijk}\theta_kV^j,\end{aligned} \tag{4.26}$$

it is easy to check that the following bilinears transform as four-vectors:

$$\begin{aligned}&(\psi_L^\dagger\psi_L, -\psi_L^\dagger\vec{\sigma}\psi_L),\\ &(\psi_R^\dagger\psi_R, \psi_R^\dagger\vec{\sigma}\psi_R),\end{aligned} \tag{4.27}$$

Defining

$$\sigma^\mu \equiv (1,\vec{\sigma}), \qquad \overline{\sigma}^\mu = (1,-\vec{\sigma}), \tag{4.28}$$

we can write the four vectors as

$$\psi_L^\dagger\overline{\sigma}^\mu\psi_L, \quad \psi_R^\dagger\sigma^\mu\psi_R. \tag{4.29}$$

Exercise: Check that under an infinitesimal Lorentz transformation, Λ:

$$\psi_L^\dagger \overline{\sigma}^\mu \psi_L \to \Lambda^\mu_\nu \psi_L^\dagger \overline{\sigma}^\nu \psi_L, \tag{4.30}$$

while

$$i\psi_L^\dagger \overline{\sigma}^\mu \partial_\mu \psi_L, \tag{4.31}$$

is invariant.

Exercise: Check that the combination $\psi_L^\dagger \psi_L$ or $\psi_R^\dagger \psi_R$ are not Lorentz invariant.

We can also write an invariant term from the bilinear combination of the two representations:

$$\psi_L^\dagger \psi_R. \tag{4.32}$$

4.3 The Dirac equation

We can then write a Lorentz-invariant and real quadratic Lagrangian as

$$\mathcal{L} = i\psi_L^\dagger \overline{\sigma}^\mu \partial_\mu \psi_L + i\psi_R^\dagger \sigma^\mu \partial_\mu \psi_R - m(\psi_L^\dagger \psi_R + \psi_R^\dagger \psi_L), \tag{4.33}$$

where m is some coupling to be determined, note that the coefficients of the first two terms, if they are different from one, can be set to one by a redefinition of the fields. By neglecting total derivatives, it is easy to see that the Lagrangian is real.

We can rewrite this Lagrangian in a more familiar form by defining the four-component spinors

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}, \quad \overline{\psi} \equiv \left(\psi_R^\dagger, \psi_L^\dagger\right), \tag{4.34}$$

or what we call a *Dirac spinor*, and by defining the Dirac gamma matrices

$$\gamma^\mu \equiv \begin{pmatrix} 0 & \sigma^\mu \\ \overline{\sigma}^\mu & 0 \end{pmatrix}, \tag{4.35}$$

the Lagrangian obtains the familiar form

$$\mathcal{L} = i\overline{\psi}\gamma^\mu \partial_\mu \psi - m\overline{\psi}\psi. \tag{4.36}$$

The equation of motion can be easily derived by considering ψ and $\overline{\psi}$ as independent fields,

$$i\partial_\mu \gamma^\mu \psi - m\psi = 0, \quad -i\partial_\mu \overline{\psi}\gamma^\mu - m\overline{\psi} = 0, \tag{4.37}$$

which is simply the *Dirac equation* if we interpret m as the mass.

Dirac obtained this equation for the first time when trying to produce a first-order equation in a time-derivative, like the Schrödinger equation, that satisfied the relativistic relation between energy and momentum. It is easy to see that solutions to the Dirac equation also satisfy the Klein-Gordon equation, since

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

where we have used the basic properties of the gamma matrices:

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \tag{4.39}$$

It is interesting to note that we have arrived at the same point from the assumption of the quantum field transforming in irreducible representations of the Lorentz group, and writing the most general Lorentz-invariant and real Lagrangian density quadratic in the fields.

It is easy to rewrite eqs. (4.25) in terms of generators in the Dirac representation:

$$\psi(x) \to \psi'(x') = D(\Lambda)\psi(\Lambda^{-1}x') = e^{-\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}}\psi(\Lambda^{-1}x'), \tag{4.40}$$

where

$$S^{\mu\nu} = \frac{\sigma^{\mu\nu}}{2} \equiv \frac{i}{4}[\gamma^\mu, \gamma^\nu], \tag{4.41}$$

and therefore

$$S^{0i} = \frac{i}{4}[\gamma^0, \gamma^i] = -\frac{i}{2}\begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix}, \quad S^{ij} = -\frac{i}{4}[\sigma_i, \sigma_j] = \frac{1}{2}\epsilon_{ijk}\sigma_k \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.42}$$

Using eq. (4.15), eq. (4.42) and eq. (4.40), it is easy to reproduce eqs. (4.25).

Some useful properties of the γ matrices are:

- $(\gamma^i)^2 = -I$, $(\gamma_0)^2 = I$.
- "Transform" as vectors:

$$D(\Lambda)^{-1}\gamma^\mu D(\Lambda) = \Lambda^\mu_\nu \gamma^\nu \tag{4.43}$$

- Hermiticity: $(\gamma^0)^\dagger = \gamma^0$, $(\gamma^i)^\dagger = -\gamma^i$, which can also be writen as $(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0$

Before proceeding to the quantization of the Dirac field, let us find the solutions to the classical Dirac equation.

4.3.1 Solutions to the Dirac equation

We want to find the solutions to the Dirac equation. We consider plane wave solutions of the form:

$$\psi(x) = u(p)e^{-ipx}, v(p)e^{ipx} \tag{4.44}$$

with $p^2 = m^2$. The spinor $u(p)$ and $v(p)$ must satisfy

$$(\gamma^\mu p_\mu - m)u(p) = 0, \ (\gamma^\mu p_\mu + m)v(p) = 0. \tag{4.45}$$

Let us consider the solutions in the rest frame, $\bar{p} \equiv (m, \mathbf{0})$:

$$\begin{aligned} m(\gamma^0 - 1)u(\bar{p}) &= 0 \longrightarrow \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} u(\bar{p}) = 0, \quad u(\bar{p}) = \begin{pmatrix} \xi \\ \xi \end{pmatrix}, \\ m(\gamma^0 + 1)v(\bar{p}) &= 0 \longrightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} v(\bar{p}) = 0, \quad v(\bar{p}) = \begin{pmatrix} \eta \\ -\eta \end{pmatrix}. \end{aligned} \tag{4.46}$$

There are two independent solutions for the two-dimensional vectors ξ and η that can be taken as $\xi, \eta = (1,0), (0,1)$. The four independent and orthogonal solutions are:

$$u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, v_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}. \tag{4.47}$$

From these solutions in the rest frame we can reconstruct the solutions for arbitrary p, using eq. (4.43). Let us define $\Lambda(p)$ as the Lorentz boost that takes momentum in the rest frame to $p = (E, p^1, p^2, p^3)$,

$$D(\Lambda(p)) = e^{-i\omega_{0i}S^{0i}}, \tag{4.48}$$

where $\omega_{0i} = \frac{-p^i}{|\mathbf{p}|}\eta$ and $S^{0i} = \frac{i}{4}[\gamma^0, \gamma^i]$ and η is the rapidity:

$$\eta = \log\left(\frac{E + |\mathbf{p}|}{m}\right). \tag{4.49}$$

If $\bar{p} = (m, \mathbf{0})$:

$$D(\Lambda(p))^{-1}(\bar{p}_\mu\gamma^\mu - m)D(\Lambda(p)) = \bar{p}_\mu\Lambda^\mu_\nu\gamma^\nu - m = p_\nu\gamma^\nu - m, \tag{4.50}$$

and therefore

$$\begin{aligned} D(\Lambda(p))^{-1}(\bar{p}_\mu\gamma^\mu - m)D(\Lambda(p))D(\Lambda(p))^{-1}u(\bar{p}) &= 0, \\ (p_\mu\gamma^\mu - m)D(\Lambda(p))^{-1}u(\bar{p}) &= 0, \end{aligned} \tag{4.51}$$

which means that $D(\Lambda(p))^{-1}u(\bar{p})$ is a solution of the Dirac equation with momentum p:

$$u_s(p) = D(\Lambda(p))^{-1}u_s(\bar{p}) = \exp\begin{pmatrix} -\frac{\vec{p}\cdot\vec{\sigma}}{|\vec{p}|}\frac{\eta}{2} & 0 \\ 0 & \frac{\vec{p}\cdot\vec{\sigma}}{|\vec{p}|}\frac{\eta}{2} \end{pmatrix} u_s(\bar{p}) = \begin{pmatrix} \sqrt{\frac{p\cdot\sigma}{m}} & 0 \\ 0 & \sqrt{\frac{p\cdot\bar{\sigma}}{m}} \end{pmatrix} u_s(\overline{p}), \tag{4.52}$$

and

$$v_s(p) = D(\Lambda(p))^{-1}v_s(\bar{p}) = \exp\begin{pmatrix} -\frac{\vec{p}\cdot\vec{\sigma}}{|\vec{p}|}\frac{\eta}{2} & 0 \\ 0 & \frac{\vec{p}\cdot\vec{\sigma}}{|\vec{p}|}\frac{\eta}{2} \end{pmatrix} v_s(\bar{p}) = \begin{pmatrix} \sqrt{\frac{p\cdot\sigma}{m}} & 0 \\ 0 & \sqrt{\frac{p\cdot\bar{\sigma}}{m}} \end{pmatrix} v_s(\overline{p}). \tag{4.53}$$

It is easy to show that

$$\overline{u}_s(p)u_s(p) = u_s(p)^\dagger\gamma^0 u_s(p) = \overline{u}_s(\bar{p})u_s(\bar{p}), \tag{4.54}$$

$$\overline{v}_s(p)v_s(p) = v_s(p)^\dagger\gamma^0 v_s(p) = \overline{v}_s(\bar{p})v_s(\bar{p}). \tag{4.55}$$

The standard normalization of these vectors is chosen as:

$$\overline{u}_s(p)u_s(p) = -\overline{v}_s(p)v_s(p) = 2m, \tag{4.56}$$

which means the vectors in eqs. (4.47) are multiplied by $\sqrt{m}$. It can be checked that

$$\overline{u}_s(p)v_{s'}(p) = \overline{u}_s(\bar{p})v_{s'}(\bar{p}) = 0. \tag{4.57}$$

Some useful relations are the spin sums:

$$\begin{aligned} \sum_{s=1}^{2} u_s(p)\overline{u}_s(p) &= \begin{pmatrix} \sqrt{p\cdot\sigma} \\ \sqrt{p\cdot\overline{\sigma}} \end{pmatrix} (\sqrt{p\cdot\overline{\sigma}}, \sqrt{p\cdot\sigma}) \\ &= \begin{pmatrix} m & p\cdot\sigma \\ p\cdot\bar{\sigma} & m \end{pmatrix} = \not{p} + m, \end{aligned} \tag{4.58}$$

and similarly

$$\sum_{s=1}^{2} v_s(p)\overline{v}_s(p) = \not{p} - m. \tag{4.59}$$

4.4 Helicity

The solutions we have found for the Dirac equation are four-spinors, but we have seen that there are smaller representations of the Lorentz group which are spinors,

ψ_L and ψ_R. We will see now that these are solutions of the massless Dirac equation. Writing the Dirac fermion in terms of $\psi_{L/R}$ using eq. (4.34), the massless Dirac equation reduces to

$$i\partial_0\psi_L - i\sigma_i\partial_i\psi_L = 0, \quad i\partial_0\psi_R + i\sigma_i\partial_i\psi_R = 0, \tag{4.60}$$

Considering plane wave solutions, $\psi_{L/R} = e^{-ipx}u_{L/R}$, with $p^2 = 0$:

$$(p^0 + \sigma\mathbf{p})u_L = 0, \quad (p^0 - \sigma\mathbf{p})u_R = 0, \tag{4.61}$$

or

$$\frac{\vec{\sigma}\mathbf{p}}{|\mathbf{p}|}u_L = -u_L, \quad \frac{\vec{\sigma}\mathbf{p}}{|\mathbf{p}|}u_R = u_R. \tag{4.62}$$

The operator $H \equiv \frac{\vec{\sigma}\mathbf{p}}{|\mathbf{p}|}$ is the helicity, i.e. the spin in the direction of momentum. The L fields are therefore left-handed (spin pointing in the oposite direction to momentum) and the R field are right-handed (spin in the direction of momentum).

From a Dirac spinor we can separate its L/R components using the matrix γ^5:

$$\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}, \tag{4.63}$$

with the following properties:

$$\{\gamma^5, \gamma^\mu\} = 0, \quad (\gamma^5)^2 = I. \tag{4.64}$$

By defining the chiral projectors

$$P_\pm \equiv \frac{1 \pm \gamma^5}{2}, \quad (P_\pm)^2 = P_\pm, \quad P_+P_- = 0, \tag{4.65}$$

we can extract the L/R components of the Dirac spinor:

$$\psi_L = P_-\psi, \quad \psi_R = P_+\psi. \tag{4.66}$$

We define then:

- Chiral states: ψ_L, ψ_R as the eigenstates of the γ_5 with eigenvalue ∓ 1 respectively:

  $$\gamma^5\psi_L = -\psi_L, \quad \gamma^5\psi_R = \psi_R. \tag{4.67}$$

  These states do not mix under Lorentz transformations (they are irreducible representations), they both represent states of spin 1/2 and in the case of massless fermions they are also eigenstates of helicity.

- Helicity states: eigenstates of the helicity operator H. They mix under Lorentz transformations unless they are massless. In other words, a massive spin 1/2 particle might be in two states of helicity.

4.5 Majorana Representation

Can we have a two-dimensional spinor representation, i.e. transforming as $(0, 1/2)$ or $(1/2, 0)$, of a massive spin 1/2 particle? This is the question that Ettore Majorana asked himself and found a clever and surprising solution.

Let us consider the $(1/2, 0)$ representation, ψ_L. There is indeed another bilinear invariant we have not writen down above:

$$\psi_L^T \sigma^2 \psi_L, \tag{4.68}$$

where ^{T} refers to the transpose. The first point to notice is that this contraction would vanish if ψ_L is not an anti-commuting variable.

We can check that this combination is Lorentz invariant, using eq. (4.25) and the property:

$$\sigma^{iT}\sigma^2 + \sigma^2\sigma^i = 0. \tag{4.69}$$

We can therefore include it in a Lorentz invariant Lagrangian:

$$\mathcal{L} = i\psi_L^\dagger \bar{\sigma}^\mu \partial_\mu \psi_L + i\alpha\psi_L^T\sigma^2\psi_L + h.c. \tag{4.70}$$

This is the Lagrangian of a Majorana fermion. It is easy to show that it represents a free massive fermion by solving the equations of motion for ψ_L:

$$\partial_\mu \psi_L^\dagger \bar{\sigma}^\mu \;=\; 2\alpha\psi_L^T\sigma^2 \to \bar{\sigma}^\mu\partial_\mu\psi_L = 2\alpha\sigma^2\psi_L^*. \tag{4.71}$$

Applying $\sigma^\nu\partial_\nu$:

$$\sigma^\nu\partial_\nu\bar{\sigma}^\mu\partial_\mu\psi_L = \partial^\mu\partial_\mu\psi_L = 2\alpha\sigma^2\bar{\sigma}^{\nu *}\partial_\nu\psi_L^*, \tag{4.72}$$

where we have used

$$\sigma^\nu\sigma^2 = \sigma^2\bar{\sigma}^{\nu *}. \tag{4.73}$$

Since

$$\sigma^{\nu *}\partial_\nu\psi_L^* = -2\alpha^*\sigma^2\psi_L, \tag{4.74}$$

we get

$$\partial^\mu\partial_\mu\psi_L = -4|\alpha|^2\psi_L. \tag{4.75}$$

So the field ψ_L satisfies the Klein-Gordon equation identifying

$$\alpha = \frac{m}{2}. \tag{4.76}$$

The quantization of the Majorana field therefore corresponds to a free massive particle.

We often write the Majorana Lagrangian in Dirac notation by defining

$$\psi = \begin{pmatrix} \psi_L \\ i\sigma^2\psi_L^* \end{pmatrix}. \tag{4.77}$$

It is easy to check that the Majorana Lagrangian of eq. (4.70) can be written as

$$\mathcal{L} = \frac{1}{2}\left(i\bar{\psi}\partial\!\!\!/\psi - m\bar{\psi}\psi\right), \tag{4.78}$$

where it must also hold that the anti-commuting variables satisfy:

$$(\psi\psi')^* = \psi'^*\psi^* = -\psi^*\psi'^*. \tag{4.79}$$

Exercise: Show that eq. (4.78) reduces to eq. (4.70) with the definition of eq. (4.77).

A different representation of the gamma matrices for Majorana fields is usually employed. In the Majorana representation

$$\gamma_M^0 = \begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix}, \ \gamma_M^1 = \begin{pmatrix} i\sigma^3 & 0 \\ 0 & i\sigma^3 \end{pmatrix}, \tag{4.80}$$

$$\gamma_M^2 = \begin{pmatrix} 0 & -\sigma^2 \\ \sigma^2 & 0 \end{pmatrix}, \ \gamma_M^3 = -\begin{pmatrix} i\sigma^1 & 0 \\ 0 & i\sigma^1 \end{pmatrix}, \tag{4.81}$$

a Majorana spinor can be shown to be real

$$\psi_M = \psi_M^*. \tag{4.82}$$

Exercise: Find the unitarity transformation, U, that satisfies:

$$U\gamma^\mu U^\dagger = \gamma_M^\mu, \tag{4.83}$$

and show that $U\psi = \psi_M$ is real.

4.6 Parity

Among Lorentz transformations there are two important discrete ones:

$$\begin{aligned} \text{Parity}: \ P &\equiv \text{Diag}(1,-1,-1,-1), \\ \text{Time reversal}: \ T &\equiv \text{Diag}(-1,1,1,1). \end{aligned}$$

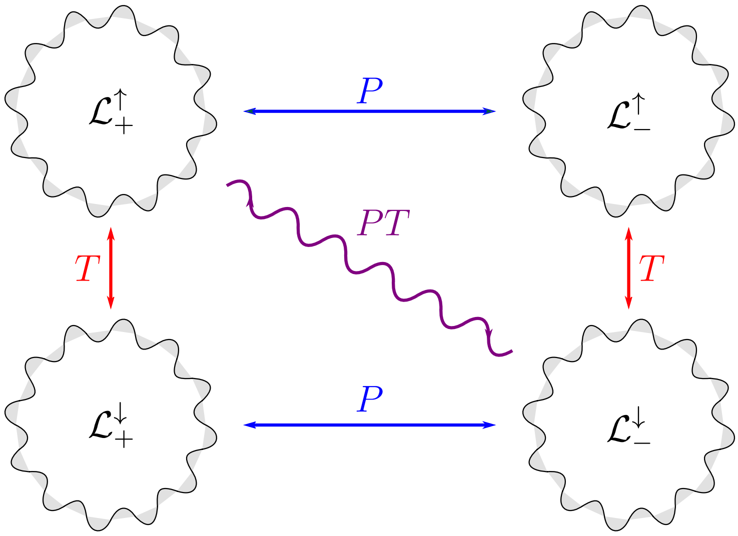


Figure 13: The Lorentz subgroups are connected via parity and time reversal transformations.

All Lorentz transfomations satisfy

$$\det\Lambda = \pm 1, \quad (\Lambda^0{}_0)^2 \geq 1, \tag{4.84}$$

and the sign of $\det\Lambda$ and $\Lambda^0{}_0$ split the Lorentz group into disconnected subgroups:

$$\begin{array}{lcl} \mathcal{L}_+^\uparrow & \text{proper orthochronous} & \det\Lambda = 1,\ \Lambda^0{}_0 \geq 1, \\ \mathcal{L}_-^\uparrow & \text{improper orthochronous} & \det\Lambda = -1,\ \Lambda^0{}_0 \geq 1, \\ \mathcal{L}_-^\downarrow & \text{improper non-orthochronous} & \det\Lambda = -1,\ \Lambda^0{}_0 \leq -1, \\ \mathcal{L}_+^\downarrow & \text{proper non-orthochronous} & \det\Lambda = 1,\ \Lambda^0{}_0 \leq -1, \end{array}$$

All subgroups can be obtained from $\mathcal{L}_+^\uparrow$ by P and/or T (see Fig. 13):

$$\mathcal{L}_+^\uparrow \xrightarrow{P} \mathcal{L}_-^\uparrow, \quad \mathcal{L}_+^\uparrow \xrightarrow{T} \mathcal{L}_+^\downarrow, \quad \mathcal{L}_+^\uparrow \xrightarrow{PT} \mathcal{L}_-^\downarrow. \tag{4.85}$$

Under parity angular momentum does not change, but boosts change sign – therefore under parity the generators J^i remain invariant, while $K^i \to -K^i$ and as a result:

$$\psi_R(x) \xrightarrow{P} \psi_L(x^P), \quad \psi_L(x) \xrightarrow{P} \psi_R(x^P), \tag{4.86}$$

which can be implemented on Dirac spinors with γ^0:

$$\psi(x) \xrightarrow{P} \gamma^0\psi(x^P), \tag{4.87}$$

The four-dimensional Dirac representation is the smallest representation of the Lorentz group and parity. It is easy to check that the Dirac equation is invariant under parity. However we will see that the weak interactions break parity.

It is useful to classify how fermion bilinears transform under Lorentz transformations and parity:

$$\begin{aligned}
\text{scalar}: \quad \bar{\psi}\psi &\xrightarrow{\Lambda} \quad \bar{\psi}\psi, &&\xrightarrow{P} \bar{\psi}\psi \\
\text{pseudo}-\text{scalar}: \quad \bar{\psi}\gamma^5\psi &\xrightarrow{\Lambda} \quad \bar{\psi}\gamma^5\psi, &&\xrightarrow{P} -\bar{\psi}\gamma^5\psi \\
\text{vector}: \quad \bar{\psi}\gamma^\mu\psi &\xrightarrow{\Lambda} \quad \Lambda^\mu{}_\nu\bar{\psi}\gamma^\nu\psi, &&\xrightarrow{P} \eta_\mu\bar{\psi}\gamma^\mu\psi \\
\text{pseudo}-\text{vector}: \quad \bar{\psi}\gamma^\mu\gamma^5\psi &\xrightarrow{\Lambda} \quad \Lambda^\mu{}_\nu\bar{\psi}\gamma^\nu\gamma^5\psi, &&\xrightarrow{P} -\eta_\mu\bar{\psi}\gamma^\mu\gamma^5\psi \\
\text{tensor}: \bar{\psi}[\gamma^\mu,\gamma^\nu]\psi &\xrightarrow{\Lambda} \quad \Lambda^\mu{}_\alpha\Lambda^\nu{}_\beta\bar{\psi}[\gamma^\alpha,\gamma^\beta]\psi &&\xrightarrow{P} -\eta_\mu\eta_\nu\bar{\psi}[\gamma^\mu,\gamma^\nu]\psi
\end{aligned} \tag{4.88}$$

where $\eta_\mu = (1,-1,-1,-1)$.

4.7 Global conserved charge

As well as the invariance under Lorentz transformations and parity, the Dirac Lagrangian is invariant under phase rotations:

$$\psi(x) \to e^{-i\alpha}\psi, \quad \bar{\psi}(x) \to e^{i\alpha}\bar{\psi}. \tag{4.89}$$

For α infinitesimal, we have therefore

$$\psi(x) \to \psi(x) + \alpha\Delta\psi(x) \to \Delta\psi(x) = -i\psi(x), \quad \Delta\bar{\psi}(x) = i\bar{\psi}(x). \tag{4.90}$$

There is a conserved Noether current

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\Delta\psi(x) + \Delta\bar{\psi}\frac{\partial\mathcal{L}}{\partial(\partial_\mu\bar{\psi})} = \bar{\psi}\gamma^\mu\psi. \tag{4.91}$$

This is the conserved vector current. The associated conserved charge is therefore

$$Q = \int d^3x\ \bar{\psi}\gamma^0\psi. \tag{4.92}$$

When we quantize the field we will see that this charge corresponds to fermion number (the number of fermions minus antifermions is conserved), and when we couple a Dirac fermion to the electromagnetic field, this will also be the electromagnetic charge.

In the case when the mass vanishes, there is an additional global symmetry under which

$$\psi \to e^{-i\alpha\gamma^5}\psi, \quad \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \to \begin{pmatrix} e^{i\alpha}\psi_L \\ e^{-i\alpha}\psi_R \end{pmatrix}, \tag{4.93}$$

that is the chiral components are multiplied by opposite phases. It is easy to check that in the limit $m \to 0$, the Dirac Lagrangian is invariant and therefore there is a conserved current that is found to be

$$j_A^\mu = \bar{\psi}\gamma^\mu\gamma^5\psi. \tag{4.94}$$

This is the axial current. For massive Dirac fermions, the current is not conserved and satisfies:

$$i\partial_\mu j_A^\mu = 2m\bar{\psi}\gamma^5\psi. \tag{4.95}$$

We say that the current is partially conserved.

We note that a Majorana fermion carries no conserved charge, neither the vector nor the axial, since the mass term is not invariant.

4.8 Quantization of the Dirac field

We proceed canonically as we did with the scalar fields:

1) Define canonical momenta of ψ:

$$\pi = \frac{\delta\mathcal{L}}{\delta\partial_0\psi} = i\bar{\psi}\gamma^0 = i\psi^\dagger. \tag{4.96}$$

2) Let us assume canonical equal-time commutation relations

$$[\psi_\alpha(\mathbf{x}), \psi_\beta^\dagger(\mathbf{y})] = \delta_{\alpha\beta}\delta(\mathbf{x}-\mathbf{y}), \tag{4.97}$$

We expand in the solutions of the free Dirac equation, that constitute a complete basis:

$$\begin{aligned} \psi(\mathbf{x}) &= \sum_s \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\left[a_p^s u_s(p)e^{i\mathbf{px}} + b_p^{s\dagger}v_s(p)e^{-i\mathbf{px}}\right], \\ \psi(\mathbf{x})^\dagger &= \sum_s \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\left[a_p^{s\dagger} u_s^\dagger(p)e^{-i\mathbf{px}} + b_p^{s}v_s^\dagger(p)e^{i\mathbf{px}}\right]. \end{aligned} \tag{4.98}$$

The commutation relations in eq. (4.97) imply:

$$[a_p^r, a_q^{s\dagger}] = (2\pi)^3\delta(\mathbf{p}-\mathbf{q})\delta^{rs}, \tag{4.99}$$

$$[b_p^r, b_q^{s\dagger}] \quad = \quad -(2\pi)^3\delta(\mathbf{p}-\mathbf{q})\delta^{rs}, \tag{4.100}$$

and all the other combinations vanish. Note the unusual sign in the b commutators.

3) Hamiltonian in terms of the a_p^s and b_p^s:

$$H(t) = \int d^3x \mathcal{H}(\mathbf{x}), \tag{4.101}$$

with the Hamiltonian density:

$$\begin{aligned} \mathcal{H} &= \pi\partial_0\psi - \mathcal{L} = i\psi^+\partial_0\psi - i\bar{\psi}\partial\!\!\!/\psi + m\bar{\psi}\psi, \\ &= -i\bar{\psi}\gamma^i\partial_i\psi + m\bar{\psi}\psi, \end{aligned} \tag{4.102}$$

$$-i\gamma^i\partial_i\psi + m\psi = \int_{\mathbf{p}} \left[(p^i\gamma^i + m)u_s(p)a_p^s e^{i\mathbf{px}} + (-p^i\gamma^i + m)v_s(p)b_p^{s\dagger}e^{-i\mathbf{px}}\right]. \tag{4.103}$$

Using

$$\begin{aligned} (p\!\!\!/ - m)u_s(p) = 0 \quad &\to \quad (p^i\gamma^i + m)u_s(p) = p^0\gamma^0 u_s(p), \\ (p\!\!\!/ + m)v_s(p) = 0 \quad &\to \quad (-p^i\gamma^i + m)v_s(p) = -p^0\gamma^0 v_s(p), \end{aligned} \tag{4.104}$$

where we defined $p\!\!\!/ \equiv p_i\gamma^i$, we obtain

$$-i\gamma^i\partial_i\psi + m\psi \quad = \quad \int_{\mathbf{p}} \omega_p \left[\gamma^0 u_s(p)a_p^s e^{i\mathbf{px}} - \gamma^0 v_s(p)b_p^{s\dagger}e^{-i\mathbf{px}}\right]. \tag{4.105}$$

Multiplying by $\bar{\psi} \equiv \psi^\dagger\gamma^0$ from the left and integrating over $\mathbf{x}$:

$$H = \int \frac{d^3p}{(2\pi)^3}\omega_p \left(a_q^{r\dagger}a_q^r - b_q^r b_q^{r\dagger}\right), \tag{4.106}$$

where we have used

$$u_r(p)^\dagger u_s(p) = v_r(p)^\dagger v_s(p) = 2p^0\delta^{rs}, \quad u_r(p'))^\dagger v_s(p) = v_r(p')^\dagger u_s(p) = 0, \tag{4.107}$$

where $p' = (p_0, -\mathbf{p})$.

Here we encounter a major problem ! The Hamiltonian is not bounded from below. What went wrong? The problem is that spin 1/2 particles are fermions (obeying Fermi-Dirac statistics) and commutators in eqs. (4.97) should be anti-commutators. In this case,

$$\{\psi_\alpha(\mathbf{x}), \psi_\beta^\dagger(\mathbf{y})\} = \delta_{\alpha\beta}\delta(\mathbf{x}-\mathbf{y}), \tag{4.108}$$

that implies

$$\{a_p^r, a_p^{s\dagger}\} = (2\pi)^3\delta(\mathbf{p}-\mathbf{q})\delta^{rs},\ \{b_p^r, b_p^{s\dagger}\} = (2\pi)^3\delta(\mathbf{p}-\mathbf{q})\delta^{rs}. \tag{4.109}$$

Exercise: show that the commutation relations eq. (4.109) imply eq. (4.108).

The Hamiltonian obtained in this case has the standard form

$$:H: = \int \frac{d^3p}{(2\pi)^3}\omega_p \sum_s \left[a_p^{s\dagger}a_p^s + b_p^{s\dagger}b_p^s\right], \tag{4.110}$$

with the commutation relations

$$\begin{aligned}[H, a_p^s] &= -\omega_p a_p^s, \qquad [H, b_p^s] = -\omega_p b_p^s,\\ [H, a_p^{s\dagger}] &= \omega_p a_p^{s\dagger}, \qquad [H, b_p^{s\dagger}] = \omega_p b_p^{s\dagger}.\end{aligned} \tag{4.111}$$

The particle interpretation of this quantum theory should be clear by now: the operators $a_p^{s\dagger}$, $b_p^{s\dagger}$ create particles with fermionic statistics and energy ω_p. The Fock vacuum state satisfies

$$a_p^s|0\rangle = 0, \quad b_p^s|0\rangle = 0, \tag{4.112}$$

and Fock states with N particles of a or b-types can be obtained by operating with N $a_p^{s\dagger}$ or $b_p^{s\dagger}$ operators.

Exercise: complete the proof of eq. (4.110).

The spatial momentum operator is derived from the $0i$ component of the energy-momentum tensor:

$$P^i = \int d^3x\ T^{0i} = -i\int d^3x\bar{\psi}\gamma^0\partial_i\psi, \tag{4.113}$$

and using eq. (4.98) we obtain

$$P^i = \int \frac{d^3p}{(2\pi)^3}p^i \sum_r (a_p^{r\dagger}a_p^r + b_p^{r\dagger}b_p^r), \tag{4.114}$$

with

$$\begin{aligned}[P^i, a_p^s] &= -p^i a_p^s, \qquad [P^i, b_p^s] = -p^i b_p^s,\\ [P^i, a_p^{s\dagger}] &= p^i a_p^{s\dagger}, \qquad [P^i, b_p^{s\dagger}] = p^i b_p^{s\dagger}.\end{aligned} \tag{4.115}$$

The spatial momentum of the one-particle states of a or b types is therefore $\mathbf{p}$.

We can also check that the created particles have spin 1/2. Angular momentum is the conserved charge associated to rotations. Under an infinitesimal rotation the fermion field transforms as

$$\psi(x) \to D[\Lambda]\psi(\Lambda^{-1}x) = \psi(x) - i\omega_{ij}S^{ij}\psi(x) - i\omega_{ij}\left(J^{ij})^{\mu}_{\nu}x^{\nu}\right)\partial_{\mu}\psi(x), \quad (4.116)$$

where S^{ij} is defined in eq. (4.41) and J^{ij} in eqs. (4.16) and (4.17). The conserved current corresponding to a rotation of parameter ω_{ij} is therefore

$$j^{\mu}_{ij} \equiv \frac{\partial\mathcal{L}}{\partial\partial_{\mu}\psi}\Delta_{ij}\psi = \bar{\psi}\gamma^{\mu}\left[S^{ij} + (J^{ij})^{\rho}_{\sigma}x^{\sigma}\partial_{\rho}\right]\psi. \quad (4.117)$$

The first term is the spin contribution while the second is the orbital contribution to the angular momentum.

Let us consider particles created at rest, that is from $a^{s\dagger}_{\mathbf{0}}$ or $b^{s\dagger}_{\mathbf{0}}$. In this case, there is no orbital angular momentum and the spin contribution to the conserved charge is just

$$Q_{ij} \equiv \int d^3x\ \bar{\psi}\gamma^0 S_{ij}\psi. \quad (4.118)$$

If we consider $ij = 12$, corresponding to a rotation around the z axis, the corresponding charge must be the spin in the z direction. It is easy to show that

$$[Q_{12}, a^{1\dagger}_{\mathbf{0}}] = \frac{1}{2}a^{1\dagger}_{\mathbf{0}}, \quad [Q_{12}, a^{2\dagger}_{\mathbf{0}}] = -\frac{1}{2}a^{2\dagger}_{\mathbf{0}}, \quad (4.119)$$

therefore $a^{1\dagger}_{\mathbf{0}}$ creates a particle with spin 1/2, while $a^{2\dagger}_{\mathbf{0}}$ creates a particle with spin -1/2.

Finally, the Fermi-Dirac statistics of the Fock space follows from the anti-commutation relations, eq. (4.109).

As in the case of the complex scalar field, the particles of a and b types can be distinguished by the conserved charge associated with the symmetry of rephasing invariance, i.e. eq. (4.92):

$$Q = \int d^3x j^0 = \int \frac{d^3p}{(2\pi)^3}\left(a^{r\dagger}_p a^r_p - b^{r\dagger}_p b^r_p\right), \quad (4.120)$$

leading to the following commutation relations:

$$[Q, a^{r\dagger}_p] = a^{r\dagger}_p, \quad [Q, b^{r\dagger}_p] = -b^{r\dagger}_p, \quad (4.121)$$

which imply that the one-particle states $a^{r\dagger}_p|0\rangle$ and $b^{r\dagger}_p|0\rangle$ have charges +1 and -1 respectively. As in the complex KG field, the a and b fields create particles and antiparticles.

The time dependence of the quantum Dirac field is easily solved from the Heisenberg equation for the operators a_p^s and b_p^s, and in complete analogy to what we did for the KG field, we obtain:

$$\psi(x) \quad = \quad \sum_s \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} \left[a_p^s u_s(p) e^{-ipx} + b_p^{s\dagger} v_s(p) e^{ipx} \right]. \tag{4.122}$$

4.9 Scattering matrix in the Dirac theory

The scattering matrix of eq. (3.18) and the Dyson series eq. (3.27) are valid for a theory with fermions. The computation of transition amplitudes from some initial state $|i\rangle$ to some final state $|f\rangle$, requires the evaluation of matrix elements of the form

$$\langle f | T(\mathcal{H}_I(x_1)....\mathcal{H}_I(x_n)) | i \rangle, \tag{4.123}$$

where $\mathcal{H}_I$ is the interaction-picture interaction Hamiltonian density. Any initial or final state corresponds to a bunch of free one-particle states. If they are fermions they will be characterized by a momentum, spin and charge (particle + versus antiparticle −):

$$|i\rangle = |p_1, s_1, \pm; ...; p_m, s_m, \pm\rangle, \quad |f\rangle = |q_1, s_1, \pm; ...; q_l, s_l, \pm\rangle, \tag{4.124}$$

with

$$|p, s, +\rangle \equiv \sqrt{2\omega_p} a_p^{s\dagger} |0\rangle, \quad |p, s, -\rangle \equiv \sqrt{2\omega_p} b_p^{s\dagger} |0\rangle. \tag{4.125}$$

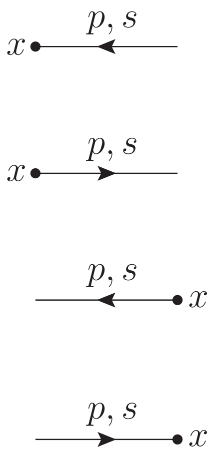
The evaluation of these matrix elements is simplified by Wick's theorem that states that the time-ordered product of a product of fermion fields can be writen as:

$$\begin{aligned} T(\psi_{\alpha_1}(x_1)..\psi_{\alpha_n}(x_n)\bar{\psi}_{\alpha_1}(y_1)\bar{\psi}_{\alpha_n}(y_n)) \quad &= \quad :\psi_{\alpha_1}(x_1)..\psi_{\alpha_n}(x_n)\bar{\psi}_{\alpha_1}(y_1)\bar{\psi}_{\alpha_n}(y_n): \\ &+ \quad \text{all possible terms with pairs of fields}(\psi, \bar{\psi}) \\ &\quad \text{contracted and normal ordered otherwise,} \end{aligned} \tag{4.126}$$

When evaluating the vacuum to vacuum transition, only products of propagators remain. That is, what remains from normal-ordering a time-ordered product of a ψ and a $\bar{\psi}$ field is S_F:

$$\begin{aligned} T(\psi(x)\bar{\psi}(y)) &= :\psi(x)\bar{\psi}(y): + S_F(x-y), \\ S_F(x-y) &\equiv \overline{\psi(x)\bar{\psi}}(y) \equiv \langle 0 | T(\psi(x)\bar{\psi}(y)) | 0 \rangle. \end{aligned} \tag{4.127}$$

The propagator of two ψ, or two $\bar{\psi}$ fields vanishes.



$$\langle p, \overline{s, -|\psi}(x) = v^s(p)e^{ipx},$$

$$\langle p, \overline{s, +|\bar{\psi}}(x) = \bar{u}^s(p)e^{ipx},$$

$$\overline{\bar{\psi}(x)|p, s}, -\rangle = \bar{v}^s(p)e^{-ipx},$$

$$\overline{\psi(x)|p, s}, +\rangle = u^s(p)e^{-ipx}.$$

Figure 14: Wick contractions with external lines.

When initial or final states contain particles:

$$\begin{aligned} &\langle q_1, s_1, \pm; ...q_l, s_l, \pm|T(\psi_{\alpha_1}(x_1)..\psi_{\alpha_n}(x_n)\bar{\psi}_{\alpha_1}(y_1)\bar{\psi}_{\alpha_n}(y_n))|p_1, s_1, \pm; ...; p_m, s_m, \pm\rangle = \\ :(...): \quad + \quad &\text{all possible terms with pairs of fields}(\psi, \bar{\psi}), (\psi/\bar{\psi}, \mathrm{a}^{(\dagger)}/\mathrm{b}^{(\dagger)}) \\ &\text{contracted and normaled ordered otherwise} \end{aligned} \tag{4.128}$$

Care should be taken in ordering the fermions fields, since they anti-commute and it is necessary to keep track of the minus signs.

Contractions of a fermion field with an external creation or annihilation operator is what remains from normal ordering the combination. Only the following four possibilities do not vanish.

4.10 Fermionic Feynman propagator

The Feynman propagator for the Dirac field is given by

$$S_F(x-y) = \begin{cases} \langle 0|\psi_\alpha(x)\bar{\psi}_\beta(y)|0\rangle, & x^0 > y^0, \\ -\langle 0|\bar{\psi}_\beta(y)\psi_\alpha(x)|0\rangle, & y^0 > x^0. \end{cases} \tag{4.129}$$

and is depicted as follows:

$$S_F(x-y)$$

Let us assume $x^0 > y^0$:

$$\begin{aligned} S_F(x-y) &= \sum_{r,s} \int \frac{d^3p}{(2\pi)^3} \int \frac{d^3q}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}\sqrt{2\omega_q}} \\ &\langle 0| \left(a_q^r u^r(q) e^{-iqx} + b_q^{r\dagger} v^r(q) e^{iqx} \right) \left(a_p^{s\dagger} \bar{u}^s(p) e^{ipy} + b_p^r \bar{v}^s(p) e^{-ipy} \right) |0\rangle. \end{aligned} \tag{4.130}$$

Using the anti-commutation relations of the a and b operators and performing the integral over d^3q, only one term survives:

$$\begin{aligned} S_F(x-y) &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_p} \sum_s u^s(p)\bar{u}^s(p) e^{-ip(x-y)} = \int \frac{d^3p}{(2\pi)^3} \frac{\not{p}+m}{2\omega_p} e^{-ip(x-y)} \\ &= (i\not{\partial}_x + m) \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_p} e^{-ip(x-y)}. \end{aligned} \tag{4.131}$$

For the case $y_0 > x_0$, similarly, we obtain:

$$\begin{aligned} S_F(x-y) &= -\int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_p} \sum_s v^s(p)\bar{v}^s(p) e^{ip(x-y)} = -\int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_p} \frac{\not{p}-m}{2\omega_p} e^{ip(x-y)} \\ &= (i\not{\partial}_x + m) \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_p} e^{ip(x-y)}. \end{aligned} \tag{4.132}$$

Combining both cases:

$$\begin{aligned} S_F(x-y) &= (i\not{\partial}_x + m) \left(\theta(x^0-y^0) \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_p} e^{-ip(x-y)} + \theta(y^0-x^0) \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_p} e^{ip(x-y)} \right) \\ &= (i\not{\partial}_x + m)\Delta_F(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i(\not{p}+m)}{p^2-m^2+i\epsilon} e^{-ip(x-y)} \\ &= \int \frac{d^4p}{(2\pi)^4} \frac{i}{\not{p}-m+i\epsilon} e^{-ip(x-y)}, \end{aligned} \tag{4.133}$$

where we have used the scalar Feynman propagator from eq. (3.38) and eq. (3.37).

It is possible to check that microcausality also holds in the case of the fermion QFT:

$$\left\{\psi_\alpha(x), \bar{\psi}_\beta(y)\right\} = (i\not{\partial}_x + m)[\phi(x), \phi(y)] = 0, \qquad (x-y)^2 < 0. \tag{4.134}$$

4.11 Fermion-scalar interactions

The Yukawa Lagrangian describes a theory with a real scalar with mass m and a Dirac fermion with mass M:

$$\mathcal{L}_{\text{Yukawa}} = \bar{\psi}(i\not{\partial} - M)\psi - \lambda i\bar{\psi}\phi\gamma^5\psi + \frac{1}{2}(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2). \tag{4.135}$$

The interaction Hamiltonian density is

$$\mathcal{H}_I = \lambda i \bar{\psi} \phi \gamma^5 \psi. \tag{4.136}$$

This theory was first proposed by Yukawa to describe the interaction of nucleons (protons and neutrons) with pions. The interaction is renormalizable since

$$[\psi] = \frac{3}{2}, \quad [\phi] = 1 \to [\lambda] = 0. \tag{4.137}$$

Another example of fermion-scalar interaction is the interaction of fermions, ψ, with the Higgs field, H, in the Standard Model:

$$\mathcal{L}_{\text{Higgs}} = \bar{\psi}(i\partial\!\!\!/ - M)\psi - \lambda \bar{\psi} H \psi + \frac{1}{2}(\partial_\mu H \partial^\mu H - m^2 H^2). \tag{4.138}$$

4.11.1 Feynman rules

We should follow the same procedure to draw the Feynman diagrams as explained in sec. 3.4. The external lines must have an arrow to indicate if they are particles or antiparticles, as explained above.

The interaction vertex involves two fermion fields and a scalar. To derive the Feynman rule corresponding to the vertex we can consider the simplest amplitude and eliminate the factors corresponding to the external particles, e.g. an incoming particle and anti-particle with momenta, p, p' and a final scalar with momentum q:

$$\begin{aligned} \langle q|(-i)\mathcal{H}_I(x)|p,s,+;p',r,-\rangle &= \sqrt{2\omega_q}\sqrt{2\omega_p}\sqrt{2\omega_{p'}}\langle 0|\, a_q \lambda \phi(x)\, \bar{\psi}(x)\gamma^5 \psi\, (x) a_p^{s\dagger}\, b_{p'}^{r\dagger}\, |0\rangle \\ &= \lambda e^{i(q-p-p')x} \bar{v}^r(p')\gamma_5 u^s(p). \end{aligned} \tag{4.139}$$

If we factorize the factors associated to the external lines, the bare vertex is $\lambda\gamma^5$ (see fig. 15 left).

The amplitude for each fully connected diagram is obtained therefore from the following elements:

- A factor $\lambda\gamma^5 \int d^4 x_i$ for each vertex.
- A Feynman propagator $S_F(x_i - x_j)$ for each line from vertex x_i to x_j.
- A factor $u^s(p)e^{-ipx}$ for each incoming particle with momenta p linked to x
- A factor $\bar{v}^r(p)e^{ipx}$ for each incoming antiparticle with momenta p linked to x
- A factor $\bar{u}^s(p)e^{ipx}$ for each outgoing particle with momenta p linked to x
- A factor $v^s(p)e^{ipx}$ for each outgoing particle with momenta p linked to x

In the case of the Higgs-fermion interaction of eq. (4.138) the only difference is the vertex, which becomes $-i\lambda$, without the γ^5 (see Fig. 15 right).

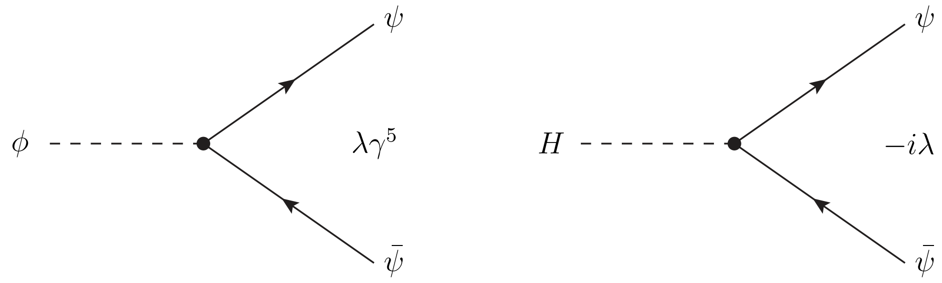


Figure 15: Interaction vertex of the Yukawa theory (left) and the Higgs theory (right).

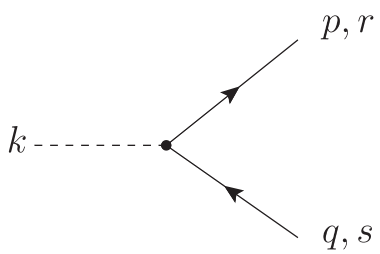


Figure 16: Feynman diagram of the process $H \to f\bar{f}$

4.11.2 Scalar decay into a fermion pair

We are interested in computing the decay width of the Higgs into a pair of fermions to the lowest order in perturbation theory. The only Feynman diagram contributing to this process is shown in Fig. 16.

The total width in the rest frame is given by

$$\Gamma(H \to f\bar{f}) = \int \frac{d^3p}{(2\pi)^3 2\omega_p} \int \frac{d^3q}{(2\pi)^3 2\omega_q} (2\pi)^4 \delta(p+q-k) |\mathcal{M}(H \to f\bar{f})|^2. \quad (4.140)$$

where $k = (m_H, 0, 0, 0)$ is the momentum of the Higgs, and according to the Feynman rules above the amplitude is

$$A = -i \int d^4x \; \bar{u}^r(p) v^s(q) e^{i(p+q-k)x} = -i(2\pi)^4 \delta(p+q-k) \bar{u}^r(p) v^s(q), \quad (4.141)$$

and therefore

$$\mathcal{M} = -i\lambda \bar{u}^r(p) v^s(q), \quad |\mathcal{M}|^2 = \lambda^2 \bar{u}^r(p) v^s(q) \bar{v}^s(q) u^r(p). \quad (4.142)$$

We are interested in the total decay width and therefore we should add the two spin polarizations: $s, r = 1, 2$. Using the spin sums of eqs. (4.58) and (4.59), we obtain:

$$\begin{aligned}\sum_{r,s} |\mathcal{M}|^2 &= \lambda^2 \sum_{r,s} \bar{u}^r_\alpha(p) v^s_\alpha(q) \bar{v}^s_\beta(q) u^r_\beta(p) = \lambda^2 \sum_s v^s_\alpha(q)\bar{v}^s_\beta(q) \sum_r u^r_\beta(p)\bar{u}^r_\alpha(p) \\ &= \lambda^2 (\not{q} - M)_{\beta\alpha} (\not{p} + M)_{\alpha\beta} = \lambda^2 \mathrm{Tr}[(\not{q} - M)(\not{p} + M)]. \end{aligned} \tag{4.143}$$

Using the properties of the γ matrices,

$$\mathrm{Tr}[\gamma^\mu] = 0, \quad \mathrm{Tr}[\not{p}\not{q}] = 4pq, \tag{4.144}$$

the trace can be easily simplified:

$$\sum_{r,s} |\mathcal{M}|^2 = 4\lambda^2 \left(pq - M^2\right). \tag{4.145}$$

Finally, integrating over phase space,

$$\begin{aligned}\Gamma(H \to f\bar{f}) &= \frac{1}{2m} \int \frac{d^3p}{(2\pi)^3 2\omega_p} \int \frac{d^3q}{(2\pi)^3 2\omega_q} (2\pi)^4 \delta(p + q - k) |\mathcal{M}(H \to f\bar{f})|^2 \\ &= \frac{\lambda^2}{2m} \int \frac{d^3p}{(2\pi)^2 \omega_p^2} \delta(m - 2\omega_p) 2\mathbf{p}^2 = \frac{\lambda^2}{m\pi} \int \frac{|\mathbf{p}|^3}{\omega_p} d\omega_p \delta(m - 2\omega_p) \\ &= \frac{\lambda^2 m}{8\pi} \left(\sqrt{1 - 4\frac{M^2}{m^2}}\right)^3. \end{aligned} \tag{4.146}$$

V. Spin 1 Quantum Fields: the photon and Proca field

Classical electrodynamics can be formulated in a Lorentz covariant way in terms of the gauge potential A_μ and the antisymmetric electromagnetic tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\nu. \tag{5.2}$$

The electric and magnetic fields are related to the electromagnetic tensor as

$$F_{0i} = E^i, \quad F_{ij} = -\epsilon_{ijk} B^k, \tag{5.3}$$

with all other components vanishing.

Two of the Maxwell equations in the absence of external charges or currents can be derived as the equations of motion from the Lagrangian density

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \tag{5.4}$$

and are given by

$$\partial_\mu F^{\mu\nu} = 0, \tag{5.5}$$

or

$$\nabla \cdot \mathbf{E} = 0, \qquad \nabla \times \mathbf{B} = \frac{\partial \mathbf{E}}{\partial t}. \tag{5.6}$$

The other two Maxwell equations derive from the Bianchi identities

$$\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0, \tag{5.7}$$

and read

$$\nabla \cdot \mathbf{B} = 0, \qquad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \tag{5.8}$$

The gauge potential A_μ has four components and transforms as a 4-vector under Lorentz transformations. This is the irreducible representation $(1/2, 1/2)$ and has spin one. However, not all of the four components are independent, because any *gauge* transformation of the form

$$A_\mu \to A_\mu + \partial_\mu \alpha(x), \tag{5.9}$$

where $\alpha(x)$ an arbitrary function leaves the electromagnetic tensor invariant, and therefore the Lagrangian. This symmetry is called *gauge invariance*.

We can choose α to simplify our calculations. This is called *fixing the gauge*. An appropriate choice in the absence of sources is the Lorentz gauge that is defined by the condition

$$\partial_\mu A^\mu = 0. \tag{5.10}$$

It is easy to check that in this gauge the equation of motion, eq. (5.5), reduces to

$$\partial_\mu \partial^\mu A_\nu = 0, \tag{5.11}$$

which is the massless Klein-Gordon equation for each component, see eq. (2.8).

Canonical quantization however is problematic due to the redundancy of the degrees of freedom.

For example, the time derivative of A_0 does not appear in the Lagrangian, and therefore the canonical momentum associated with A_0 vanishes:

$$\pi_0 = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_0)} = 0. \tag{5.12}$$

We can bypass this problem by working instead in Coulomb gauge:

$$A_0 = 0, \quad \boldsymbol{\nabla} \cdot \mathbf{A} = 0. \tag{5.13}$$

This gauge is not Lorentz invariant and therefore not very practical, but it has no redundancy and canonical quantization can be carried out without encountering problems. We will first perform the quantization in Coulomb gauge and then consider a covariant gauge.

5.1 Quantization of the electromagnetic field in Coulomb gauge

In Coulomb gauge, the field variables are A_i for $i = 1, 2, 3$, since $A_0 = 0$. However not all are independent since they must satisfy the constraint

$$\nabla \cdot \mathbf{A} = 0. \tag{5.14}$$

The canonical momenta are obtained in the usual way:

$$\pi^i = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_i)} = -F^{0i} = E^i. \tag{5.15}$$

The classical Hamiltonian is therefore

$$H = \int d^3x \; \pi^i \partial_0 A_i - \mathcal{L} = \int d^3x \; \mathbf{E}^2 - \mathcal{L}, \tag{5.16}$$

where we have used $\partial_0 A_i = F_{0i}$ in Coulomb gauge, and $\mathcal{L}$ is given in eq. (5.4). Using eq. (5.3) the Lagrangian density can also be writen as

$$\mathcal{L} = \frac{1}{2}(\mathbf{E}^2 - \mathbf{B}^2), \tag{5.17}$$

with $\mathbf{B} = \nabla \times \mathbf{A}$. Therefore the classical Hamiltonian is

$$H = \frac{1}{2} \int d^3x \; (\mathbf{E}^2 + \mathbf{B}^2) \tag{5.18}$$

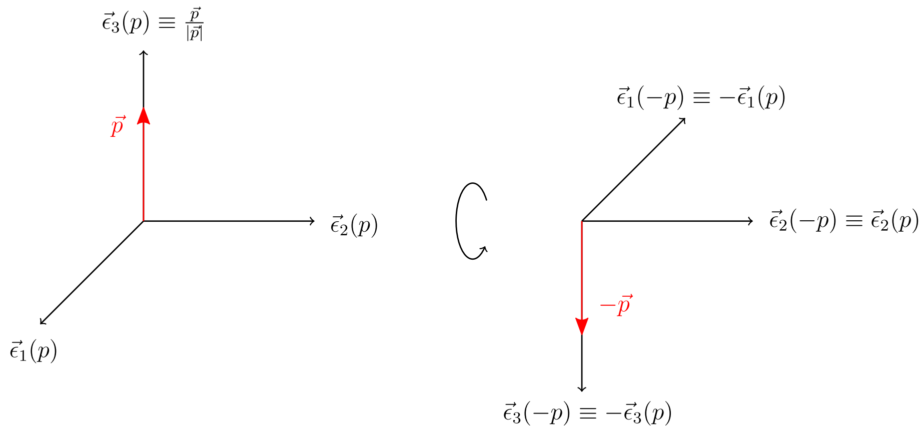


Figure 17: The Lorentz subgroups are connected via parity and time reversal transformations.

Before imposing the canonical commutation relations for A_i and π_i, we need to solve the constraint of eq. (5.14). For that we first go to Fourier space

$$\mathbf{A}(t,\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\tilde{\mathbf{A}}_p(t)e^{i\mathbf{p}\mathbf{x}}, \qquad \boldsymbol{\pi}(t,\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\tilde{\boldsymbol{\pi}}_p(t)e^{i\mathbf{p}\mathbf{x}}. \tag{5.19}$$

The vectors $\tilde{\mathbf{A}}_p, \tilde{\boldsymbol{\pi}}_p$ can always be writen in an orthonormal basis, $\mathbf{e}_r(p)$, as:

$$\tilde{\mathbf{A}}_p = \sum_{r=1-3} A_p^r \mathbf{e}_r(p), \quad \tilde{\boldsymbol{\pi}}_p = \sum_{r=1-3} \pi_p^r \mathbf{e}_r(p) \tag{5.20}$$

where A_p^r, π_p^r, are complex numbers. Plugging this into eq. (5.19), the constraint can be satisfied if

$$\mathbf{e}_r(p)\cdot\mathbf{p} = 0. \tag{5.21}$$

Therefore for each Fourier component $\mathbf{p}$ we can choose a basis of two unitary vectors $r = 1, 2$ orthogonal to $\mathbf{p}$, such that

$$\mathbf{e}_r(p)\cdot\mathbf{p} = 0, \qquad \mathbf{e}_r(p)\cdot\mathbf{e}_s(p) = \delta_{rs}, \qquad r = 1, 2. \tag{5.22}$$

It is easy to see that the Hamiltonian reduces to

$$H = \frac{1}{2}\sum_{r=1,2}\int \frac{d^3p}{(2\pi)^3}\left[\pi_p^r\pi_p^{r*} + |\mathbf{p}|^2 A_p^r A_p^{r*}\right], \tag{5.23}$$

which for each r has the form of the real scalar field in sec. 2.3.

To conserve $\epsilon^1 \times \epsilon^2 = \epsilon^3$ under $\vec{p} \to -\vec{p}$ we can choose these vectors to satisfy

$$\mathbf{e}^1(-p) = -\mathbf{e}^1(p), \qquad \mathbf{e}^2(-p) = \mathbf{e}^2(p), \tag{5.24}$$

which can be seen as a rotation of π around the $\mathbf{e}^2(p)$-axis. Therefore, the reality of the gauge field implies

$$A^1_p = -A^{1*}_{-p}, \quad A^2_p = A^{2*}_{-p}. \tag{5.25}$$

Introducing the raising and lowering operators analogously to eq. (2.40),

$$a^r_p = \sqrt{\frac{\omega_p}{2}} \left(A^r_p + i \frac{\pi^r_p}{\omega_p} \right), \tag{5.26}$$

we can write the field and the momenta as

$$\begin{aligned}
\mathbf{A}(t, \mathbf{x}) &= \sum_{r=1,2} \int \frac{d^3p}{(2\pi)^3 \sqrt{2|\mathbf{p}|}} \left[a^r_p(t) \mathbf{e}_r(p) e^{i\mathbf{px}} + a^{r\dagger}_p(t) \mathbf{e}_r(p) e^{-i\mathbf{px}} \right], \\
\boldsymbol{\pi}(t, \mathbf{x}) &= \sum_{r=1,2} \int \frac{d^3p}{(2\pi)^3} (-i) \sqrt{\frac{|\mathbf{p}|}{2}} \left[a^r_p(t) \mathbf{e}_r(p) e^{i\mathbf{px}} - a^{r\dagger}_p(t) \mathbf{e}_r(p) e^{-i\mathbf{px}} \right].
\end{aligned} \tag{5.27}$$

We can now proceed to quantization by postulating the equal-time commutation relations for the $a^r_p(t)$ operators:

$$[a^r_p, a^{s\dagger}_q] = (2\pi)^3 \delta(\mathbf{p} - \mathbf{q}) \delta^{rs}. \tag{5.28}$$

The quantum electromagnetic field and momentum are therefore given by eq. (5.27) promoting a^r_p and $a^{r\dagger}_p$ to operators.

The quantum Hamiltonian is as usual the classical one in terms of the quantum operators:

$$\begin{aligned}
\hat{H} &= \frac{1}{2} \int d^3x \left[\hat{\boldsymbol{\pi}}^2 + (\boldsymbol{\nabla} \times \hat{\boldsymbol{A}})^2 \right], \\
:\hat{H}: &= \sum_{r=1,2} \int \frac{d^3p}{(2\pi)^3} |\mathbf{p}| \;\; a^{r\dagger}_p a^r_p.
\end{aligned} \tag{5.29}$$

The time evolution of the field operators is easy to derive from the commutation relations

$$[H, a^r_p] = -|\mathbf{p}| a^r_p, \;\; [H, a^{r\dagger}_p] = |\mathbf{p}| a^{r\dagger}_p. \tag{5.30}$$

The Heisenberg equation for the a^r_p operator is:

$$\frac{da^r_p}{dt} = i[H, a^r_p] = -i|\mathbf{p}| a^r_p, \tag{5.31}$$

which can be easily solved,

$$a^r_p(t) = e^{-i|\mathbf{p}|t} a^r_p(0). \tag{5.32}$$

The quantum fields can be therefore be writen in terms of the initial time operators:

$$\mathbf{A}(t,\mathbf{x}) \quad = \quad \sum_{r=1,2} \int \frac{d^3p}{(2\pi)^3\sqrt{2|\mathbf{p}|}} \left[a_p^r \mathbf{e}_r(p) e^{-ipx} + a_p^{r\dagger} \mathbf{e}_r(p) e^{ipx} \right], \qquad (5.33)$$

with $p = (|\mathbf{p}|, \mathbf{p})$.

The particle interpretation should be clear by now. The quantum field creates and annihilates particles in well defined position states. The particles in this Fock space are Einstein's photons. A one photon state is therefore

$$a_p^{r\dagger}|0\rangle, \qquad (5.34)$$

with energy $|\mathbf{p}|$. The momentum is $\mathbf{p}$ as can be checked by applying the momentum operator (i.e. the conserved current associated with space translations) to the state:

$$:\hat{P}^i: = :\int d^3x \frac{\partial \mathcal{L}}{\partial(\partial_0 A_j)} \partial^i A_j: = \sum_r \int \frac{d^3p}{(2\pi)^3} p^i a_p^{r\dagger} a_p^r. \qquad (5.35)$$

It is therefore a massless particle.

The spin operator is obtained from the conserved current associated with space rotations. The field transforms under Lorentz transformations the gauge potential transforms as

$$A^\mu(x) \to \Lambda^\mu_\nu A^\nu(\Lambda^{-1}x). \qquad (5.36)$$

As shown in eq. (4.16), for an infinitesimal transformation

$$\Lambda^\mu_\nu = \delta^\mu_\nu + i\omega_{\alpha\beta}(J^{\alpha\beta})^\mu_\nu \qquad (5.37)$$

the change in the field,

$$\Lambda^\mu_\nu A^\nu(\Lambda^{-1}x) - A^\mu \equiv \omega_{\alpha\beta}\Delta_{\alpha\beta}A^\mu, \qquad (5.38)$$

has therefore two contributions, the part coming from Λ and from the coordinate change $\Lambda^{-1}x$. If Λ corresponds to a rotation, the former is the spin, while the second is the orbital angular momentum.

Let us consider the spin part:

$$\Delta_{\alpha\beta}A^\mu = i(J^{\alpha\beta})^\mu_\nu A^\nu = g^{\alpha\mu}A^\beta - g^{\beta\mu}A^\alpha, \qquad (5.39)$$

where we have used eq. (4.17). The associated conserved current is therefore

$$j^\nu_{\alpha\beta} = \frac{\partial \mathcal{L}}{\partial(\partial_\nu A^\mu)}\Delta_{\alpha\beta}A^\mu = -F^{\nu\alpha}A^\beta + F^{\nu\beta}A^\alpha. \qquad (5.40)$$

The conserved charge associated with a rotation around ij is therefore

$$S_{ij} = \int d^3x \;\; j^0_{ij} = \int d^3x \left(-F^{0i}A^j + F^{0j}A^i\right) = \int d^3x \left(\pi^i A^j - \pi^j A^i\right). \qquad (5.41)$$

After some algebra we obtain

$$:S^{ij}: = i\sum_{rs}\int \frac{d^3p}{(2\pi)^3} a^{s\dagger}_p a^r_p \left[\epsilon^i_s(p)\epsilon^j_r(p) - (i \leftrightarrow j)\right]. \qquad (5.42)$$

Then

$$:S^{ij}:a^{u\dagger}_q|0\rangle = i\sum_s \left[\epsilon^i_u(q)\epsilon^j_s(q) - (i \leftrightarrow j)\right] a^{s\dagger}_q|0\rangle. \qquad (5.43)$$

Let us assume $\mathbf{q} = (0,0,q)$, then $\epsilon_1(q) = (1,0,0)$ and $\epsilon_2(q) = (0,1,0)$, that is $\epsilon^i_r(q) = \delta^i_r$. Therefore for a rotation around the z axis:

$$:S^{12}:a^{1\dagger}_q|0\rangle = ia^{2\dagger}_q|0\rangle, \;\; :S^{12}:a^{2\dagger}_q|0\rangle = -ia^{1\dagger}_q|0\rangle, \qquad (5.44)$$

and

$$:S^{12}:(a^{1\dagger}_q \pm ia^{2\dagger}_q)|0\rangle = \pm(a^{1\dagger}_q \pm ia^{2\dagger}_q)|0\rangle. \qquad (5.45)$$

Therefore the states $(a^{1\dagger}_q \pm ia^{2\dagger}_q)|0\rangle$ are eigenstates of helicity (spin in the direction of the momentum) with eigenvalues ± 1. These are the two photon polarizations.

5.2 Covariant quantization: Gupta-Bleuer

The quantization we carried out has an important drawback: it is not covariant. For example the gauge propagator is easily obtained (in complete analogy with the derivation of the scalar propagator):

$$\langle 0|T(A^i(x)A^j(y))|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{e^{-ip(x-y)}}{2|\mathbf{p}|}\sum_r \epsilon^i_r(p)\epsilon^j_r(p), \qquad (5.46)$$

and using the completeness of the ϵ_r basis with $r = 1,2,3$,

$$\sum_{r=1,2,3} \epsilon^i_r(p)\epsilon^j_r(p) = \delta^{ij}, \qquad \epsilon^i_3(p) = \frac{p^i}{|\mathbf{p}|}, \qquad (5.47)$$

we have

$$\sum_{r=1,2} \epsilon^i_r(p)\epsilon^j_r(p) = \delta^{ij} - \frac{p^ip^j}{|\mathbf{p}|^2}, \qquad (5.48)$$

and the propagator in the Coulomb gauge is therefore

$$\langle 0|T(A^i(x)A^j(y))|0\rangle = \int \frac{d^4p}{(2\pi)^4}\frac{\delta^{ij} - p^ip^j/|\mathbf{p}|^2}{p^2+i\epsilon}e^{-ip(x-y)}. \tag{5.49}$$

This non-covariant propagator makes computations in perturbation theory much more cumbersome and it is difficult to prove properties such as renormalizability.

To overcome this difficulty, we often work in a covariant gauge which involves unphysical degrees of freedom that can be shown not to contribute to physical processes.

Let us consider Lorentz gauge defined by the condition eq. (1.71). In this gauge we can include an extra term in the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2\alpha}\left(\partial_\mu A^\mu\right)^2, \tag{5.50}$$

where α is arbitrary. Since we can always go to the Lorentz gauge, this extra term should have no physical effect.

By an abuse of language, the choice $\alpha = 1$ is referred to as Feynman gauge, while $\alpha = 0$ is termed Landau gauge. For $\alpha = 1$, it is easy to check that the equations of motion is the massless Klein-Gordon for all the four components

$$\partial_\mu\partial^\mu A^\nu = 0. \tag{5.51}$$

Let us consider Feynman gauge. We can formally perform a canonical quantization, and the canonical momentum of A^μ does not vanish. In this gauge, the Lagrangian of eq. (5.50) simplifies to

$$\mathcal{L} = -\frac{1}{2}\partial_\mu A_\nu\partial^\mu A^\nu + \text{total derivatives}, \tag{5.52}$$

and the canonical momenta are simply

$$\pi^\mu = -\partial^0 A^\mu. \tag{5.53}$$

The Hamiltonian density takes the form

$$\mathcal{H} = \pi^0\partial_0 A_0 + \pi^i\partial_0 A_i - \mathcal{L} = -\frac{(\pi^0)^2}{2} + \frac{(\pi^i)^2}{2} - \frac{(\partial_i A_0)^2}{2} + \frac{(\partial_i A_j)^2}{2}. \tag{5.54}$$

Going to Fourier space,

$$\begin{aligned} A_\mu(t,\mathbf{x}) &= \int \frac{d^2p}{(2\pi)^3}\tilde{A}_\mu(t,p)e^{i\mathbf{px}}, \\ \pi^\mu(t,\mathbf{x}) &= \int \frac{d^2p}{(2\pi)^3}\tilde{\pi}^\mu(t,p)e^{i\mathbf{px}}, \end{aligned} \tag{5.55}$$

the Hamiltonian is that of harmonic oscillators for each p and μ. However the μ component has the wrong sign. We will come back to this later. For the moment we continue as if this was not a problem.

As usual, to quantize we impose the canonical equal-time commutation relations

$$[A_\mu(\mathbf{x}), \pi^\nu(\mathbf{y})] = i\delta_\mu^\nu \delta(\mathbf{x}-\mathbf{y}). \tag{5.56}$$

Using eq. (5.53), this is equivalent to

$$[A_\mu(\mathbf{x}), \dot{A}^\nu(\mathbf{y})] = -i\delta_\mu^\nu \delta(\mathbf{x}-\mathbf{y}). \tag{5.57}$$

In all generalities we can write the four-vector $\tilde{A}_\mu$ on a basis of four unitary four vectors, $\epsilon_\mu(p,\lambda)$, for $\lambda = 0-3$:

$$\tilde{A}_\mu(t,p) = \sum_{\lambda=0,1,2,3} \epsilon_\mu(p,\lambda) A_p^\lambda. \tag{5.58}$$

An useful basis is

$$\epsilon^\mu(p,\lambda) = \begin{cases} (1,0,0,0) & \lambda = 0 \\ (0, \boldsymbol{e}_\lambda(p)) & \lambda = 1-3 \end{cases} \tag{5.59}$$

where $\mathbf{e}_{1,2}$ have been defined in the previous section, while $\mathbf{e}_3 = \mathbf{p}/|\mathbf{p}|$.

The following properties hold:

$$\epsilon_\mu(p,\lambda)\epsilon^\mu(p,\lambda') = -\delta_{\lambda\lambda'}\eta_\lambda = g^{\lambda\lambda'}, \tag{5.60}$$

with $\eta_0 = -1, \eta_{1-3} = 1$. Similarly we have

$$\sum_{\lambda=0-3} \epsilon^\mu(p,\lambda)\epsilon^\nu(p,\lambda)\eta_\lambda = -g^{\mu\nu}. \tag{5.61}$$

In complete analogy with previous section, we can introduce

$$a_p^\lambda = \sqrt{\frac{|\mathbf{p}|}{2}} \left(A_p^\lambda + i\frac{\pi_p^\lambda}{|\mathbf{p}|} \right). \tag{5.62}$$

The commutation relations of eq. (5.57) imply

$$[A_p^\lambda, \pi_q^{\lambda'\dagger}] = i(2\pi)^3 \eta^\lambda \delta^{\lambda\lambda'} \delta(\mathbf{p}-\mathbf{q}), \tag{5.63}$$

and

$$[a_p^\lambda, a_q^{\lambda'\dagger}] = -g^{\lambda\lambda'}(2\pi)^3 \delta(\mathbf{p}-\mathbf{q}). \tag{5.64}$$

We arrive to the quantum field operator which is given by

$$A_\mu(t, \mathbf{x}) = \sum_\lambda \int \frac{d^3p}{(2\pi)^3\sqrt{2|\mathbf{p}|}} \left(a_p^\lambda(t)\epsilon_\mu(p,\lambda)e^{i\mathbf{px}} + h.c \right). \tag{5.65}$$

The quantum Hamiltonian is

$$:\hat{H}: = \sum_\lambda \int \frac{d^3p}{(2\pi)^3}|\mathbf{p}| \ \ \eta_\lambda \ a_p^{\lambda\dagger} a_p^\lambda, \tag{5.66}$$

and the temporal dependence of $a_p^\lambda(t)$ can be obtained as usual from the Heisenberg equation.[10] The quantum operator is:

$$A_\mu(x) = \sum_\lambda \int \frac{d^3p}{(2\pi)^3\sqrt{2|\mathbf{p}|}} \left(a_p^\lambda \epsilon_\mu(p,\lambda)e^{-ipx} + h.c \right). \tag{5.67}$$

Exercise: Starting from eq. (5.67) and (5.54), show eq. (5.66)

The wrong sign of the temporal contribution to the Lagrangian in the classical theory, has resulted in a contribution to the quantum Hamiltonian that is not positive definite. For example, imagine the state

$$\hat{a}_p^{0,\dagger}|0\rangle = |p, 0\rangle \tag{5.68}$$

and assume we have a wave packet $\int f(p)\hat{a}_p^{0\dagger}|0\rangle$. Its normalization would be

$$\int_q\int_p f^*(p)f(q)\langle 0|[\hat{a}_p^0, \hat{a}_q^{0\dagger}]|0\rangle = -\int_p |f(p)|^2 = -1, \tag{5.69}$$

which does not make sense because a negative norm would mean negative probabilities. This is however unsurprising since in the covariant method, we know there are unphysical degrees of freedom.

The Gupta-Bleuer solution to this problem relies on a restriction of physical states to those that satisfy the Lorentz condition. More precisely, the expectation value of the operator $\partial_\mu A^\mu$ in any physical state $|\Psi\rangle$ vanishes,

$$\langle\Psi|\partial_\mu A^\mu|\Psi\rangle = 0, \tag{5.70}$$

which is equivalent to

$$(\partial_\mu A^\mu)^+|\Psi\rangle = 0, \tag{5.71}$$

[10] The opposite sign of the $\lambda = 0$ term in the Hamiltonian is compensated by the one in the commutator of eq. (5.64).

where $(\partial_\mu A^\mu)^+$ is the annihilation operator part. This ensures the Lorentz condition and Maxwell's equations on physical states.

Substituting eq. (5.67) in eq. (5.70), implies

$$(\partial_\mu A^\mu)^+|\Psi\rangle = -i\sum_\lambda \int \frac{d^3p}{(2\pi)^3\sqrt{2|\mathbf{p}|}} p_\mu \epsilon^\mu(p,\lambda) \;\; a_p^r e^{-ipx}|\Psi\rangle \tag{5.72}$$

Since

$$p_\mu \epsilon^\mu(p,\lambda) = \begin{cases} |\mathbf{p}| & \lambda = 0 \\ 0 & \lambda = 1,2 \\ -|\mathbf{p}| & \lambda = 3 \end{cases} , \tag{5.73}$$

eq. (5.72) implies

$$(a_p^0 - a_p^3)|\Psi\rangle = 0, \tag{5.74}$$

for all p and Ψ. This also implies

$$\langle\Psi|a_p^{0\dagger}(a_p^0 - a_p^3)|\Psi\rangle = \langle\Psi|(a_p^{0\dagger}a_p^0 - a_p^{3\dagger}a_p^3)|\Psi\rangle = 0. \tag{5.75}$$

The last equality implies that there is no contribution to the energy of any physical state from the $\lambda = 0$ and $\lambda = 3$ polarizations.

In summary, the price of maintaining Lorentz covariance is to include the unphysical polarizations $\lambda = 0, 3$, but they will have no contribution to the energy or any other physical observable as long as the Gupta-Bleuer condition, eq. (5.74) is satisfied.

5.3 Wick contractions for gauge fields

The Feynman propagator in this gauge is given by

$$\langle 0|T(A_\mu(x)A_\nu(y))|0\rangle = \lim_{\epsilon\to 0^+} \int \frac{d^4p}{(2\pi)^4} \frac{-ig_{\mu\nu}}{p^2 + i\epsilon} e^{-ip(x-y)}. \tag{5.76}$$

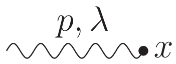
Up to the metric tensor it is the propagator of a massless scalar field. It is usually denoted by a wiggly line, see Fig. 18.

Exercise: Demonstrate eq. (5.76).

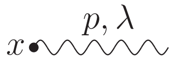
Had we used a different choice of α, or a different gauge, in eq. (5.50), the quantization would be more complicated, but the physics would be the same. Working in a different gauge ammounts to a change of propagator:

$$\langle 0|T(A_\mu(x)A_\nu(y))|0\rangle = -i \lim_{\epsilon\to 0^+} \int \frac{d^4p}{(2\pi)^4} \frac{g_{\mu\nu} + (\alpha - 1)p_\mu p_\nu/p^2}{p^2 + i\epsilon} e^{-ip(x-y)}. \tag{5.77}$$

$$\langle 0|T(A_\mu(x)A_\nu(y))|0\rangle$$



$$A_\mu(x)|p,\lambda\rangle = \eta_\lambda \epsilon_\mu(p,\lambda)e^{-ipx}$$



$$\langle p,\lambda|A_\mu(x) = \eta_\lambda \epsilon^*_\mu(p,\lambda)e^{ipx},$$

Figure 18: Gauge propagator and Wick contractions with external states.

The theory we have quantized is non-interacting. The one-particle states are the photons, characterized by a momentum and a physical polarization, $\lambda = 1, 2$. The properly normalized one-particle state is:

$$|p,\lambda\rangle = \sqrt{2|\mathbf{p}|}a_p^{\lambda\dagger}|0\rangle, \quad \lambda = 1, 2. \tag{5.78}$$

When we include interactions, the computation of the Dyson series for S-matrix elements requires the application of Wick's theorem.

Besides propagators, contractions of fields with external states will enter. The Feynman rules are shown in Fig. 18

As in the case of fermions, there are no renormalizable interactions of the electromagnetic field with itself, but there are with scalars or fermions. The theory that includes the electromagnetic field and a complex scalar is called scalar quantum electrodynamics (SQED) while the theory with a Dirac fermion is the famous quantum electrodynamics (QED), the theory that accurately describes the electromagnetic interactions of electrons. We will consider both theories in the next chapter.

5.4 Massive photons (Proca field)

The classical Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{m^2}{2}A_\mu A^\mu. \tag{5.79}$$

The equations of motion read

$$\partial_\rho F^{\rho\mu} + m^2 A^\mu = 0. \tag{5.80}$$

By applying ∂_μ to this equation we obtain:

$$\partial_\mu \partial_\rho F^{\rho\mu} + m^2 \partial_\mu A^\mu = 0. \tag{5.81}$$

The first term is identically zero and therefore the Lorentz condition follows from the equation of motion:

$$\partial_\mu A^\mu = 0. \tag{5.82}$$

This is not surprising since gauge invariance of eq. (5.9) is broken by the mass term.

By considering the Lorentz condition into account the equation of motion simplifies to

$$\partial_\rho \partial^\rho A^\mu + m^2 A^\mu = 0, \tag{5.83}$$

meaning that each component satisfies the massive Klein-Gordon equation.

Canonical quantization can proceed as above, starting with the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{m^2}{2} A_\mu A^\mu - \frac{1}{2} (\partial_\mu A^\mu)^2, \tag{5.84}$$

we obtain

$$A_\mu(x) = \sum_\lambda \int \frac{d^3p}{(2\pi)^3 \sqrt{2\omega_p}} \left(a_p^\lambda \epsilon_\mu(p,\lambda) e^{-ipx} + h.c \right), \tag{5.85}$$

with

$$p_0 = \omega_p = \sqrt{|\mathbf{p}|^2 + m^2}, \tag{5.86}$$

and the canonical commutation relations:

$$[a_p^\lambda, a_{p'}^{\lambda'}] = \delta_{\lambda\lambda'} (2\pi)^3 \delta(\mathbf{p} - \mathbf{p}'). \tag{5.87}$$

There is only one unphysical polarization in this case, which can be decoupled completely by choosing a basis of polarization vectors that satisfy:

$$p \cdot \epsilon(p,\lambda) = 0. \tag{5.88}$$

The basis can be chosen for momentum in the z direction as

$$\epsilon^\mu(p,1) = (0,1,0,0), \quad \epsilon^\mu(p,2) = (0,0,1,0), \quad \epsilon^\mu(p,3) = (|\mathbf{p}|,0,0,E). \tag{5.89}$$

For any other direction, the appropriate rotation should be performed. It is not hard to show that

$$\sum_\lambda \epsilon^\mu(p,\lambda) \epsilon_\nu(p,\lambda) = -(g_{\mu\nu} - p_\mu p_\nu / m^2), \qquad \epsilon_\mu(p,\lambda) \epsilon^\mu(p,\lambda') = \delta^{\lambda\lambda'} \eta_\lambda. \tag{5.90}$$

The Proca field propagator is found to be

$$\langle 0 | T(A_\mu(x) A_\nu(y)) | 0 \rangle = -i \lim_{\epsilon \to 0+} \int \frac{d^4p}{(2\pi)^4} \frac{g_{\mu\nu} - p_\mu p_\nu / m^2}{p^2 - m^2 + i\epsilon} \, e^{-ip(x-y)}. \tag{5.91}$$

Exercise: Prove eq. (5.91).

VI. Quantum Electrodynamics

Gauge fields can interact with scalar and fermion fields. These interactions are essentially fixed by the gauge symmetry

$$A_\mu(x) \quad \to \quad A_\mu(x) + \partial_\mu \alpha(x), \tag{6.2}$$

combined with a local rotation of the scalar or fermion fields of the form

$$\phi \to e^{ie\alpha(x)}\phi. \tag{6.3}$$

We consider both theories in turn.

6.1 Scalar QED

The theory necessarily contains a complex scalar so that its phase can be rotated locally. The free Lagrangian is the combination of eq. (5.50) and eq. (2.71):

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \partial_\mu\phi^\dagger\partial^\mu\phi - m^2\phi^\dagger\phi + \mathcal{L}_{\text{int}} \tag{6.4}$$

while the interaction Lagrangian $\mathcal{L}_{\text{int}}$ is fixed by imposing the gauge symmetry of eqs. (6.2) and (6.3). The scalar kinetic term is not invariant, but the following change

$$\begin{aligned} \partial_\mu\phi &\quad \to \quad D_\mu\phi \equiv (\partial_\mu - ieA_\mu)\phi, \\ \partial_\mu\phi^\dagger &\quad \to \quad D_\mu\phi^\dagger \equiv (\partial_\mu + ieA_\mu)\phi^\dagger, \end{aligned} \tag{6.5}$$

makes the kinetic term invariant. It is easy to see that under a gauge transformation

$$D_\mu\phi \to e^{ie\alpha(x)}D_\mu\phi, \quad D_\mu\phi^\dagger \to e^{-ie\alpha(x)}D_\mu\phi^\dagger, \tag{6.6}$$

and therefore the gauge invariant Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + D_\mu\phi^\dagger D^\mu\phi - m^2\phi^\dagger\phi. \tag{6.7}$$

and

$$\mathcal{L}_{\text{int}} = -ieA^\mu\left(\partial_\mu\phi^\dagger\phi - \phi^\dagger\partial_\mu\phi\right) + e^2A_\mu A^\mu\phi^\dagger\phi = -\mathcal{H}_{\text{int}}. \tag{6.8}$$

There are two type of interaction vertices: one involving two scalars and one gauge field, and the other involving two gauge fields and two scalar fields.

The Feynman rules for the Wick contractions are summarized in Fig. 19 for the scalar field and in Fig. 18 for the gauge fields.

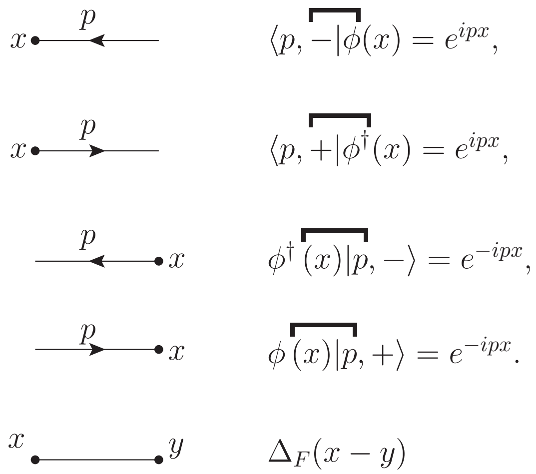


Figure 19: Wick contractions for the complex scalar.

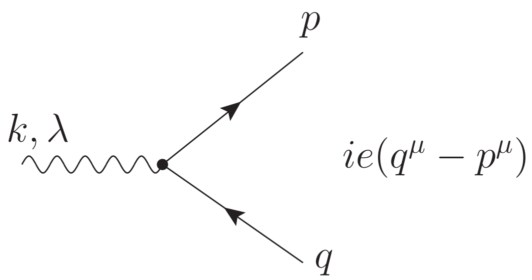


Figure 20: Feynman rule for vertex 1

To obtain the Feynman rules for these vertices we evaluate the amputated amplitude for a photon going into a particle and antiparticle scalar pair. Using the results in Fig. (18) and eqs. (3.48)

$$(-i)\langle p,+;q,-|\mathcal{H}_{\text{int}}|k,\lambda\rangle = ie(p^\mu - q^\mu)[\epsilon^{\lambda *}_{\mu}(k)e^{i(p+q-k)x}]. \tag{6.9}$$

The term inside the brackets is the contribution from the external contractions, therefore the vertex is the rest, as shown in Fig. 20.

The second vertex can be obtained by considering the amplitude of two photons going into a pair:

$$(-i)\langle p,+;q,-|\mathcal{H}_{\text{int}}|k,\lambda;k,\lambda'\rangle = ie^2[\epsilon^{\mu\lambda}(k)g_{\mu\nu}\epsilon^{\nu\lambda}(k')e^{i(p+q-k-k')x}]. \tag{6.10}$$

After factorizing the external factors the vertex is shown in Fig. 21.

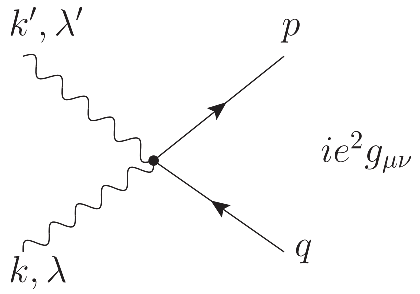


Figure 21: Feynman rule for vertex 2

6.2 QED

The theory contains a Dirac fermion representing the electron, e^-, and positron, e^+, with mass m_e, interacting with the electromagnetic field. As in the previous case, the interaction is fixed by the requirement of gauge invariance. The Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}(i\,\partial\!\!\!/ - m_e)\psi + \mathcal{L}^{\text{int}}. \tag{6.11}$$

The fermion kinetic term does not satisfy gauge invariance, but it does with the change $\partial_\mu \to D_\mu$ since

$$D_\mu\psi \to e^{ie\alpha(x)}D_\mu\psi. \tag{6.12}$$

The most general gauge invariant and renormalizable Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}(i\,D\!\!\!\!/ - m_e)\psi, \tag{6.13}$$

with

$$\mathcal{L}^{\text{int}} = e\bar{\psi}\gamma^\mu\psi A_\mu = -\mathcal{H}_{\text{int}}. \tag{6.14}$$

The Wick contractions are those in Fig. 14 and Fig. 18. There is only one interaction vertex including the fields $A_\mu, \psi, \bar{\psi}$. Considering the amplitude of a photon decaying into an electron and a positron

$$(-i)\langle p, s, +; q, r, -|\mathcal{H}_{\text{int}}|k, \lambda\rangle = ie\bar{u}^s(p)\gamma^\mu v^r(q)\epsilon^\lambda_\mu(k)e^{i(p+q-k)x}. \tag{6.15}$$

Factorizing the external particle factors, the vertex is shown in Fig. 22. Note that we refer to + as the electron which in our convention has charge $-e$.

Obviously in a theory with more than one type of charged particle, for example if we want to describe electrons and protons, we just need an additional fermionic contribution of the same type with the appropriate mass and charge.

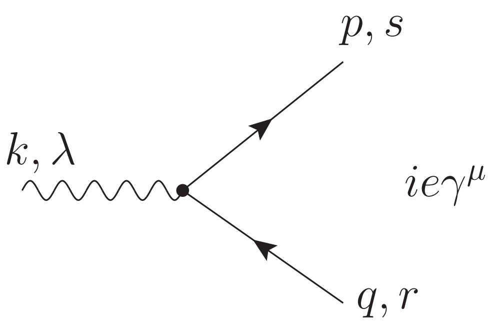


Figure 22: Feynman rule for QED vertex.

6.3 Non-relativistic limit of the Dirac equation

The equation of motion from $\mathcal{L}$ is the Dirac equation

$$(i \not\partial + e \not A - m_e)\psi = 0. \tag{6.16}$$

We want to analyze the non-relativistic limit of this equation, which should lead us to the Schrodinger equation. Multiplying by γ^0 from the left we obtain

$$(i\partial_0 + i\gamma^0\gamma^i\partial_i + eA_\mu\gamma^0\gamma^\mu - m_e\gamma^0)\psi = 0. \tag{6.17}$$

We define

$$\psi'(x) = e^{imt}\psi(x), \tag{6.18}$$

which is equivalent to subtracting m from the energy. The equation for ψ' is:

$$(i\partial_0 + i\gamma^0\gamma^i\partial_i + eA_\mu\gamma^0\gamma^\mu + m(1-\gamma^0))\psi' = 0. \tag{6.19}$$

Defining

$$\psi' = \begin{pmatrix} \chi \\ \phi \end{pmatrix}, \quad \eta \equiv \chi + \phi, \quad \rho \equiv \chi - \phi, \tag{6.20}$$

the equation can be written as

$$\begin{aligned} i\partial_0\eta &= i\sigma^i\partial_i\rho - eA_0\eta + eA_i\sigma^i\rho \\ i\partial_0\rho &= -2m\rho + i\sigma^i\partial_i\eta - eA_0\rho + eA_i\sigma^i\eta. \end{aligned} \tag{6.21}$$

The non-relativistic limit implies $i\partial_0 \ll m$ and $eA_0 \ll m$. In the second equation we can then approximate

$$\rho = \frac{i\sigma^i\partial_i\eta + eA_i\sigma^i\eta}{2m}. \tag{6.22}$$

Substituting in the first equation

$$(i\partial_0 + eA_0)\eta = \frac{[\sigma^i(i\partial_i - eA^i)]^2}{2m}\eta. \tag{6.23}$$

Simplifying

$$\begin{aligned} \sigma^i\sigma^j(i\partial_i - eA^i)(i\partial_j - eA^j) &= \frac{1}{2}\left([\sigma^i,\sigma^j] + \{\sigma^i,\sigma^j\}\right)(i\partial_i - eA^i)(i\partial_j - eA^j) \\ &= (i\partial_i - eA^i)^2 + eB^k\sigma^k. \end{aligned} \tag{6.24}$$

Therefore we obtain a Schödinger equation for η:

$$i\partial_0\eta = H\eta, \tag{6.25}$$

with the Hamiltonian

$$H = \frac{(i\boldsymbol{\nabla} - e\mathbf{A})^2}{2m} + \frac{e}{2m}\boldsymbol{\sigma}\mathbf{B} - eA_0 \tag{6.26}$$

a few observations are in order.

- η is a two-component spinor. In the non-relativistic limit, particles and antiparticles decouple, so half of the degrees of freedom are unnecessary.
- The magnetic term is the Pauli term[11]; it represents the interaction of the magnetic moment of the electron with the magnetic field contributing to the energy,

  $$-\boldsymbol{\mu}\cdot\mathbf{B}. \tag{6.27}$$

  Classically, a charged particle with charge q, mass m and angular momentum $\mathbf{L}$ has a magnetic moment

  $$\boldsymbol{\mu} = \frac{q}{2m}\mathbf{L}. \tag{6.28}$$

  The angular momentum of the electron, with charge $-e$, is the spin and therefore

  $$\boldsymbol{\mu} = -\frac{e}{2m}\mathbf{S} = -\frac{e}{2m}\frac{\boldsymbol{\sigma}}{2}. \tag{6.29}$$

  We see that this classical analysis is wrong by a factor of 2. According to the eq. (6.26), the electron magnetic moment is

  $$\boldsymbol{\mu} = -\frac{e}{2m}\sigma = -\frac{g\,e}{2m}\mathbf{S}, \tag{6.30}$$

  where we have introduced the so-called gyromagnetic ratio, which is $g = 2$.

  The unexpected prediction of $g = 2$ was the major triumph of the Dirac equation. The modification of this prediction by higher orders in perturbation theory, as first computed by J. Schwinger, was a major success of QED, as we will see in the section 6.6

[11] Pauli derived this term phenomenologically to explain experimental data, but it was not justified in non-relativistic quantum mechanics.

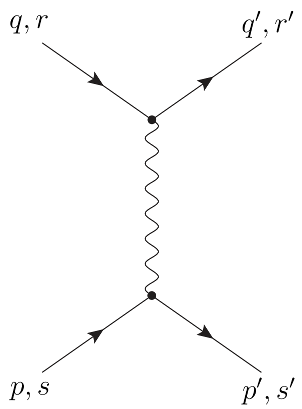


Figure 23: Rutherford scattering.

6.4 Rutherford Scattering

The first process we can consider is Rutherford scattering, that is the elastic scattering of electrons on protons, see Fig. 1:

$$e^- p^+ \to e^- p^+. \tag{6.31}$$

The Rutherford cross-section can be computed classically and in non-relativistic quantum mechanics, the result is the same:

$$\frac{d\sigma}{d\Omega} = \frac{m_e^2 e^4}{64\pi^2 p^4 \sin^4\frac{\theta}{2}}. \tag{6.32}$$

We want to compute this cross section in QED to the leading order in perturbation theory.

The theory contains two types of Dirac fermions: electrons and protons (with different masses and opposite charges). The interaction Hamiltonian is therefore

$$\mathcal{H}^{\text{int}} = -e\bar{\psi}_e \not{A}\psi_e + e\bar{\psi}_p \not{A}\psi_p. \tag{6.33}$$

Let us call the proton momenta and polarization, p, s in the initial state and p', s' in the final state. Similarly the initial and final momenta and polarization of the electron is q, r and q', r' respectively. We want to compute the leading order S-matrix element

$$\mathcal{A} = \langle p', s'; q', r'|T\exp\left(-i\int d^4x\mathcal{H}_{\text{int}}(x)\right)|p, s; q, r\rangle. \tag{6.34}$$

The leading contribution comes necessarily from the second order

$$\mathcal{A} = e^2\int d^4x\int d^4y\ \langle p', s'; q', r'|T(\bar{\psi}_e(x)\not{A}(x)\psi_e(x)\bar{\psi}_p(y)\not{A}(y)\gamma^\mu\psi_p(y))|p, s; q, r\rangle. \tag{6.35}$$

There is only one Wick contraction that gives

$$\mathcal{A} = e^2 \int d^4x \int d^4y \, \langle p', s'; q', r'|T(\bar{\psi}_e(x) \, \not{A}(x)\psi_e(x)\bar{\psi}_p(y) \, \not{A}(y)\psi_p(y))|p, s; q, r\rangle \tag{6.36}$$

There is only one Feynman diagram as shown in Fig. 23.

From the Feynman rules in Fig. 23 and 18 the amplitude is

$$\mathcal{A} = \int d^4x \int d^4y e^{ip'x} e^{iq'y} e^{-ipx} e^{-iqy} \; \bar{u}_{s'}(p')(-ie)\gamma^\mu u_s(p)\bar{u}_{r'}(q')(ie)\gamma^\nu u_r(q) \times \int \frac{d^4k}{(2\pi)^4} \frac{-ig_{\mu\nu}}{k^2} e^{ik(x-y)}, \tag{6.37}$$

and

$$\begin{aligned} \mathcal{A} &= -ie^2(2\pi)^4 \; \bar{u}_{s'}(p')\gamma^\mu u_s(p)\bar{u}_{r'}(q')\gamma_\mu u_r(q) \int d^4k \frac{1}{k^2}\delta(p'-p+q)\delta(q'-q-k) \\ &= (2\pi)^4\delta(p'+q'-p-q)\left[-ie^2\frac{\bar{u}_{s'}(p')\gamma^\mu u_s(p)\bar{u}_{r'}(q')\gamma_\mu u_r(q)}{(p'-p)^2}\right] \\ &= (2\pi)^4\delta(p'+q'-p-q)\mathcal{M}. \end{aligned} \tag{6.38}$$

To compute the cross-section, eq. (3.82), we need $|\mathcal{M}|^2$, but we are interested in the unpolarized cross-section, which implies summing over the final spin polarizations, s' and r' and averaging over the initial ones, s and r:

$$\langle|\mathcal{M}|^2\rangle_{\text{unpol}} = \frac{1}{4}\sum_{s,s',r,r'} |\mathcal{M}|^2, \tag{6.39}$$

with the factor $\frac{1}{4}$ coming from the 4 possibilities of the initial spin polarizations (2 for the electron and for the proton). The result is

$$\begin{aligned} \langle|\mathcal{M}|^2\rangle &= \frac{1}{4}\sum_{s,s'} \bar{u}_{s'}(p')\gamma^\mu u_s(p)\bar{u}_s(p)\gamma^\nu u_{s'}(p') \sum_{r,r'} \bar{u}_{r'}(q')\gamma_\mu u_r(q)\bar{u}_r(q)\gamma_\nu u_{r'}(q') \\ &= \frac{1}{4}\text{Tr}[(\not{p}+m_p)\gamma^\nu(\not{p}'+m_p)\gamma^\mu]\text{Tr}[(\not{q}+m_e)\gamma_\nu(\not{q}'+m_e)\gamma_\mu]. \end{aligned} \tag{6.40}$$

Using the relations

$$\text{Tr}[\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma] = 4(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma} + g^{\mu\sigma}g^{\nu\rho}), \quad \text{Tr}[\gamma^\mu\gamma^\nu] = 4g^{\mu\nu}, \tag{6.41}$$

and the fact that the trace of an odd number of gamma matrices vanishes, we can easily compute the traces:

$$\begin{aligned} \text{Tr}[(\not{p}+m_p)\gamma^\nu(\not{p}'+m_p)\gamma^\mu] &= 4\left[(m_p^2 - pp')g^{\mu\nu} + p^\nu p'^\mu + p^\mu p'^\nu\right], \\ \text{Tr}[(\not{q}+m_e)\gamma^\nu(\not{q}'+m_e)\gamma^\mu] &= 4\left[(m_e^2 - qq')g_{\mu\nu} + q_\nu q'_\mu + q_\mu q'_\nu\right]. \end{aligned} \tag{6.42}$$

Introducing the Madelstam variables,

$$\begin{aligned} s &\equiv (p+q)^2 = (p'+q')^2, && (6.43)\\ t &\equiv (p-p')^2 = (q-q')^2, && (6.44)\\ u &\equiv (p-q')^2 = (q-p')^2, && (6.45)\end{aligned}$$

we can simplify the amplitude to

$$\begin{aligned} \langle|\mathcal{M}|^2\rangle &= \frac{4e^2}{t^2}\left[\frac{t}{2}g^{\mu\nu} + p^\nu p'^\mu + p^\mu p'^\nu\right]\left[\frac{t}{2}g_{\mu\nu} + q^\nu q'^\mu + q^\mu q'^\nu\right] \\ &= \frac{4e^2}{t^2}\left[(m_e^2+m_p^2)t + \frac{(s-m_e^2-m_p^2)^2}{2} + \frac{(u-m_e^2-m_p^2)^2}{2}\right]. \quad (6.46)\end{aligned}$$

Using the relation

$$s+t+u = 2m_e^2 + 2m_p^2, \quad (6.47)$$

we arrive to

$$\langle|\mathcal{M}|^2\rangle = \frac{4e^4}{t^2}\left[\frac{s^2+u^2}{2} + 2(m_e^2+m_p^2)t - (m_e^2+m_p^2)^2\right]. \quad (6.48)$$

Now that we have the amplitude in terms of Lorentz invariant quantities, we can study it in any frame. Particularly, in the proton's rest frame we have

$$p = (m_p, 0, 0, 0), \quad q = (E, \mathbf{q}), \quad q' = (E', \mathbf{q}'). \quad (6.49)$$

In the limit in which $m_p \gg E$ we can approximate

$$E' \simeq E - \frac{|\mathbf{q}|^2}{m_p}(1-\cos\theta), \quad |\mathbf{q}'| \simeq |\mathbf{q}|. \quad (6.50)$$

Then

$$\begin{aligned} s &= m_p^2 + m_e^2 + 2m_pE, \\ u &= m_p^2 + m_e^2 - 2m_pE', \\ t &= 2m_e^2 - 2qq' = 2m_e^2 - 2E^2\left(1 - \frac{|\mathbf{q}|^2}{E^2}\cos\theta\right) = 2m_e^2 - 2E^2 + 2v^2E^2\cos\theta \\ &= -2v^2E^2(1-\cos\theta) = -4v^2E^2\sin^2\frac{\theta}{2}. \quad (6.51)\end{aligned}$$

and

$$\langle|\mathcal{M}|^2\rangle \simeq \frac{16m_p^2E^2e^4}{t^2}\left(1 - v^2\sin^2\frac{\theta}{2}\right). \quad (6.52)$$

Finally the cross-section is given by

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{unpol}} = \frac{1}{64\pi^2 s}\langle|\mathcal{M}|^2\rangle. \tag{6.53}$$

In the non-relativistic limit of the electron, $|\mathbf{q}| \ll m_e$ ($v \to 0$), we recover the Rutherford cross section.

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{unpol}} = \frac{e^4 m_e^2}{64\pi^2|\mathbf{q}|^4 \sin^4\frac{\theta}{2}}. \tag{6.54}$$

6.5 Compton scattering

We next consider Compton scattering, that is the scattering of photons on electrons,

$$e^-(p,s)\gamma(k,\lambda) \to e^-(p',s)\gamma(k',\lambda'), \tag{6.55}$$

which, as we discussed in the introduction, has no classical nor non-relativistic quantum physics analog, since the concept of photon as a particle requires the quantization of the electromagnetic field. QED gives an accurate prediction of this cross-section.

The S-matrix element describing the process is

$$\mathcal{A} = \langle p', s'; k', \lambda'|T\exp\left(-i\int d^4x\ \mathcal{H}_{\text{int}}(x)\right)|p, s; k, \lambda\rangle, \tag{6.56}$$

and the first non-trivial contribution is obtained at the second order,

$$\mathcal{A} = -\frac{e^2}{2}\int d^4x \int d^4y\ \langle p', s'; k', \lambda'|T(\bar{\psi}_e(x)\ \not{A}(x)\psi_e(x)\bar{\psi}_e(y)\ \not{A}(y)\psi_e(y))|p, s; k, \lambda\rangle. \tag{6.57}$$

There are four possible Wick contractions, half of them correspond to the exchange of the vertices and simply cancel the factor of 2. Once the vertices are fixed, there are still two contractions:

$$\begin{aligned}\mathcal{A}^{(1)} &= -e^2\int d^4x \int d^4y\ \langle p', s'; k', \lambda'|T(\bar{\psi}_e(x)\ \not{A}(x)\psi_e(x)\bar{\psi}_e(y)\ \not{A}(y)\psi_e(y))|p, s; k, \lambda\rangle,\\ \mathcal{A}^{(2)} &= -e^2\int d^4x \int d^4y\ \langle p', s'; k', \lambda'|T(\bar{\psi}_e(x)\ \not{A}(x)\psi_e(x)\bar{\psi}_e(y)\ \not{A}(y)\psi_e(y))|p, s; k, \lambda\rangle.\end{aligned} \tag{6.58}$$

The corresponding Feynman diagrams are shown in Fig. 24.

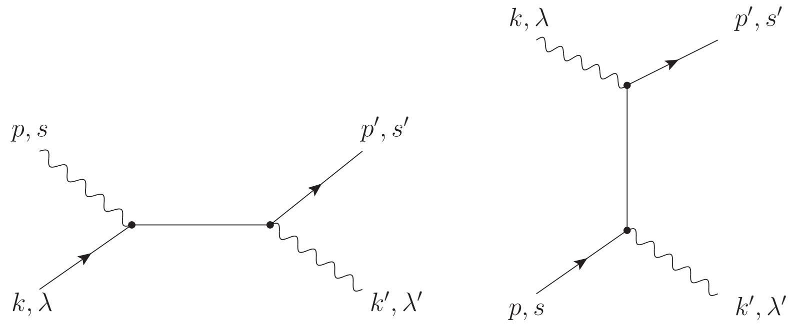


Figure 24: Feynman diagrams for Compton scattering. Left: s-channel, Right: t-channel.

The Feynman rules imply:

$$\begin{aligned}\mathcal{M}^{(1)} &= -ie^2\frac{\bar{u}_{s'}(p')\gamma_\mu(\not{p}+\not{k}+m_e)\gamma_\nu u_s(p)}{(p+k)^2-m_e^2}\epsilon^{\mu*}_{\lambda'}(k')\epsilon^{\nu}_{\lambda}(k) \equiv M^{(1)}_{\mu\nu}\epsilon^{\mu*}_{\lambda'}(k')\epsilon^{\nu}_{\lambda}(k),\\ \mathcal{M}^{(2)} &= -ie^2\frac{\bar{u}_{s'}(p')\gamma_\nu(\not{p}-\not{k}'+m_e)\gamma_\mu u_s(p)}{(p-k')^2-m_e^2}\epsilon^{\mu*}_{\lambda'}(k')\epsilon^{\nu}_{\lambda}(k) \equiv M^{(2)}_{\mu\nu}\epsilon^{\mu*}_{\lambda'}(k')\epsilon^{\nu}_{\lambda}(k).\end{aligned} \tag{6.59}$$

We are interested in the unpolarized cross-section which requires averaging over initial spin polarizations and summing over the final ones. There are four possibilities in the initial polarizations, two for the photon and two for the electron, therefore

$$\langle|\mathcal{M}|^2\rangle = \frac{1}{4}\sum_{\lambda,\lambda',s,s'}|\mathcal{M}^{(1)}+\mathcal{M}^{(2)}|^2 = \frac{1}{4}\sum_{\lambda,\lambda',s,s'}|M_{\mu\nu}\epsilon^{\mu*}_{\lambda'}(k')\epsilon^{\nu}_{\lambda}(k)|^2, \tag{6.60}$$

where $M_{\mu\nu} \equiv M^{(1)}_{\mu\nu} + M^{(2)}_{\mu\nu}$.

The photon polarization sums can be done immediately,

$$\langle|\mathcal{M}|^2\rangle = \frac{1}{4}\sum_{s,s'}M_{\mu\nu}M^*_{\alpha\beta}\sum_{\lambda'}\epsilon^{\mu*}_{\lambda'}(k')\epsilon^{\alpha}_{\lambda'}(k')\sum_{\lambda}\epsilon^{\nu}_{\lambda}(k)\epsilon^{\beta*}_{\lambda}(k). \tag{6.61}$$

From the spin sums for $\lambda = 0-3$ in eq. (5.60):

$$-g^{\mu\nu} = -\epsilon^{\mu*}_0(k)\epsilon^{\nu}_0(k) + \epsilon^{\mu*}_3(k)\epsilon^{\nu}_3(k) + \sum_{\lambda=1,2}\epsilon^{\mu*}_{\lambda}(k)\epsilon^{\nu}_{\lambda}(k), \tag{6.62}$$

Using

$$\epsilon^{\mu}_0(k) = n^\mu = (1,0,0,0), \quad \epsilon^{\mu}_3(k) = \frac{k^\mu-(kn)n^\mu}{kn}, \tag{6.63}$$

$$-\epsilon_0^{\mu *}(k)\epsilon_0^{\nu}(k) + \epsilon_3^{\mu *}(k)\epsilon_3^{\nu}(k) = \frac{k^\mu k^\nu - (kn)k^\mu n^\nu - (kn)n^\mu k^\nu}{(kn)^2}. \tag{6.64}$$

$$\sum_{\lambda=1,2} \epsilon_\lambda^{\mu *}(k)\epsilon_\lambda^{\nu}(k) = -g^{\mu\nu} + \mathcal{O}(k^\mu, k^\nu), \tag{6.65}$$

where $\mathcal{O}(k^\mu, k^\nu)$ refers to terms linear in k^μ or k^ν. But these terms vanish, since gauge invariance implies that

$$k^\mu M_{\mu\nu} = k^\nu M_{\mu\nu} = k'^\mu M_{\mu\nu} = k'^\nu M_{\mu\nu} = 0. \tag{6.66}$$

We can see this at the level of eq. (6.58),

$$\int d^3x \bar{\psi}_e(x)\gamma^\mu A_\mu(x)\psi_e(x)|k,\lambda\rangle = \int d^3x \bar{\psi}_e(x)\gamma^\mu \psi_e(x)\epsilon_\mu(k,\lambda)e^{-ikx}, \tag{6.67}$$

cancelling ϵ^μ we obtain a part of the contribution to $M^{\mu\nu}$. If we contract with k_μ, we obtain

$$\begin{aligned} k_\mu \int d^3x \bar{\psi}_e(x)\gamma^\mu \psi_e(x) e^{-ikx} &= i\int d^3x \bar{\psi}_e(x)\gamma^\mu \psi_e(x)\partial_\mu(e^{-ikx}) \\ &= -i\int d^3x \partial_\mu(\bar{\psi}_e(x)\gamma^\mu\psi_e(x))e^{-ikx} = 0, \end{aligned} \tag{6.68}$$

because $i\bar{\psi}_e\gamma^\mu\psi_e$ is a conserved current, see eq. (4.91).

In practice photon spin sums within physical amplitudes reduce to the metric tensor. Therefore

$$\langle|\mathcal{M}|^2\rangle = \frac{1}{4}\sum_{s,s'} M_{\mu\nu}M^*_{\mu\nu} \equiv \sum_{i,j=1,2}(ij), \tag{6.69}$$

where

$$(ij) = \frac{1}{4}\sum_{s,s'} M^{(i)}_{\mu\nu}M^{(j)*}_{\mu\nu}. \tag{6.70}$$

By performing the fermionic spin sums we obtain

$$\begin{aligned}
(11) &= \frac{e^4}{(s-m_e^2)^2}\mathrm{Tr}[\gamma^\mu(\not{p}+\not{k}+m_e)\gamma^\nu(\not{p}+m_e)\gamma_\nu(\not{p}+\not{k}+m_e)\gamma_\mu(\not{p}'+m_e)], \\
(22) &= \frac{e^4}{(t-m_e^2)^2}\mathrm{Tr}[\gamma^\mu(\not{p}-\not{k}'+m_e)\gamma^\nu(\not{p}+m_e)\gamma_\nu(\not{p}-\not{k}'+m_e)\gamma_\mu(\not{p}'+m_e)], \\
(12) &= \frac{e^4}{(s-m_e^2)(t-m_e^2)}\mathrm{Tr}[\gamma^\mu(\not{p}+\not{k}+m_e)\gamma^\nu(\not{p}+m_e)\gamma_\mu(\not{p}-\not{k}'+m_e)\gamma_\mu(\not{p}'+m_e)], \\
(21) &= \frac{e^4}{(s-m_e^2)(t-m_e^2)}\mathrm{Tr}[\gamma^\nu(\not{p}-\not{k}'+m_e)\gamma^\mu(\not{p}+m_e)\gamma_\nu(\not{p}+\not{k}+m_e)\gamma_\mu(\not{p}'+m_e)].
\end{aligned} \tag{6.71}$$

Using the following relations:

$$\gamma^\nu(\not{p}+m)\gamma_\nu = -2\not{p}+4m_e, \qquad \gamma^\nu\not{p}\not{q}\not{k}\gamma_\nu = -2\not{k}\not{q}\not{p}, \tag{6.72}$$

and using $k^2 = k'^2 = 0$, $p^2 = p'^2 = m_e^2$ and momentum conservation, $p+k = p'+k'$, all terms can be reduced to masses and the scalar products pk and pk'. For example:

$$pp' = p(p+k-k') = m^2 + pk - pk'. \tag{6.73}$$

After some painful algebra we obtain

$$\begin{aligned}
(11) &= \frac{8e^4}{(s-m_e^2)^2}\left[m_e^4 + m_e^2(pk) + (pk)(pk')\right], \\
(22) &= \frac{8e^4}{(t-m_e^2)^2}\left[m_e^4 - m_e^2(pk) + (pk)(pk')\right], \\
(12) &= (21) = \frac{4e^4 m_e^2}{(s-m_e^2)(t-m_e^2)}\left[2m_e^2 + (pk) - (pk')\right].
\end{aligned} \tag{6.74}$$

Using $s - m_e^2 = 2pk$, $t - m_e^2 = -2pk'$ we finally arrive to

$$\langle|\mathcal{M}|^2\rangle = 8e^4\left[m_e^4\left(\frac{1}{pk}-\frac{1}{pk'}\right)^2 + 2m_e^2\left(\frac{1}{pk}-\frac{1}{pk'}\right) + \frac{pk'}{pk} + \frac{pk}{pk'}\right]. \tag{6.75}$$

In the laboratory frame

$$\begin{aligned}
k &= (\omega, 0, 0, \omega), \\
k' &= (\omega', 0, \omega'\sin\theta, \omega'\cos\theta), \\
pk &= m_e\omega, \\
pk' &= m_e\omega'.
\end{aligned} \tag{6.76}$$

And using momentum conservation,

$$\frac{1}{\omega'} - \frac{1}{\omega} = \frac{1}{m_e}(1-\cos\theta), \tag{6.77}$$

which is the same equation as the Compton relation, eq. (1.4), since $\omega = \lambda^{-1}$, we finally arrive at

$$\left(\frac{d\sigma}{d\Omega}\right)_{unpol} = \frac{\alpha^2}{2m_e^2}\frac{\omega'^2}{\omega^2}\left(\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \sin^2\theta\right). \tag{6.78}$$

This is the so-called Klein-Nishima formula.

In the limit $\omega \ll m$, $\omega' \simeq \omega$, the cross-section reduces to the Thomson scattering cross section,

$$\left(\frac{d\sigma}{d\Omega}\right)_{unpol} = \frac{\alpha^2}{2m_e^2}\left(1+\cos^2\theta\right). \tag{6.79}$$

6.6 *Anomalous magnetic moment*

Soon after Dirac's success in predicting the gyromagnetic ratio of the electron, $g = 2$, various experiments in the 1940s showed that the experimental value differed at the per mil level. This effect was called *anomalous magnetic moment*, that is

$$\Delta g = g - 2. \tag{6.80}$$

Such a contribution could be explained if the Dirac Lagrangian would included a term like

$$\Delta\mathcal{L} = ae\bar{\psi}\sigma^{\mu\nu}F_{\mu\nu}\psi. \tag{6.81}$$

The non-relativistic limit of the corresponding equations of motion, as done in sec. 6.3, leads to a modification of eq. (6.26) by the term

$$\Delta H = 2ea\boldsymbol{\sigma}\mathbf{B}, \tag{6.82}$$

that can be absorbed in

$$\Delta g = 8ma. \tag{6.83}$$

This term induces a modification of the QED Hamiltonian:

$$-i\Delta\mathcal{H}_{\text{int}} = iae\;\bar{\psi}\sigma^{\mu\nu}F_{\mu\nu}\psi. \tag{6.84}$$

The photon-electron-positron interaction is therefore modified by

$$\langle q, r; k, \lambda| - i\Delta\mathcal{H}_{\text{int}}(x)|p, s\rangle = 2ea\bar{u}_r(q)k_\nu\sigma^{\mu\nu}u_s(p)\epsilon_\mu(k,\lambda)e^{i(q+k-p)x}. \tag{6.85}$$

Using momentum conservation at the vertex, $p = k + q$ and the Gordon identity

$$\bar{u}(q)\gamma^\mu u(p) = \bar{u}(q)\left[\frac{p^\mu + q^\mu}{2m} + i\sigma^{\mu\nu}\frac{q_\nu - p_\nu}{2m}\right]u(p), \tag{6.86}$$

we obtain

$$\langle q, r; k, \lambda| - i\Delta\mathcal{H}_{\text{int}}(x)|p, s\rangle = i\bar{u}_r(q)\left[\Delta e\gamma^\mu - 2ea(p^\mu + q^\mu)\right]u_s(p)\epsilon_\mu(k,\lambda)e^{i(q+k-p)x}, \tag{6.87}$$

where we have defined $\Delta e \equiv 4mae$. The Feynman rule for the anomalous vertex is therefore

$$\delta\Gamma^\mu = i\Delta e\gamma^\mu - i\frac{e\Delta g}{4m}(p^\mu + q^\mu), \tag{6.88}$$

where we have used eq. (6.83). The first term has the same form as the standard QED vertex, so it can be seen as a modification of the coupling Δe, while the second term has a completely different structure.

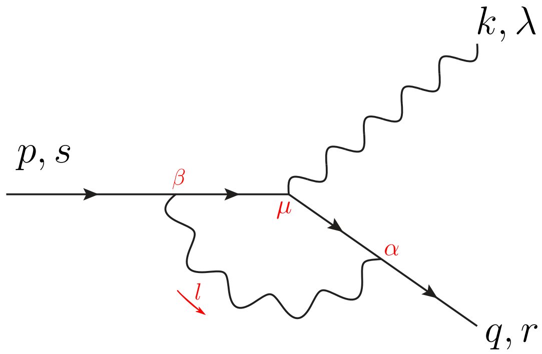


Figure 25: Anomalous vertex from eq. (6.81).

The next step is to see that radiative corrections to the QED vertex, such as those induced by the diagram in Fig. 25 generate a correction of the form in eq. (6.88) and therefore a correction to g.

The amplitude for this process is

$$\begin{aligned}
\mathcal{M} &= \int \frac{d^4 l}{(2\pi)^4} \bar{u}_r(q)(ie\gamma^\alpha) \frac{i}{\not{q} - \not{l} - m} (ie\gamma^\mu) \frac{i}{\not{p} - \not{l} - m} (ie\gamma^\beta) \frac{-i g_{\alpha\beta}}{l^2} u_s(p) \epsilon_{\lambda\mu}(k) \\
&= e^3 \int \frac{d^4 l}{(2\pi)^4} \frac{\bar{u}_r(q)\gamma^\alpha \left(\not{q} - \not{l} + m\right) \gamma^\mu \left(\not{p} - \not{l} + m\right) \gamma_\alpha u_s(p)}{[(q-l)^2 - m^2][(p-l)^2 - m^2] k^2} \epsilon_{\lambda\mu}(k) \\
&\equiv \int \frac{d^4 l}{(2\pi)^4} \frac{N^\mu(p,q,l)\epsilon_{\lambda\mu}(k)}{D(q-l,m)D(p-l,m)D(l,0)},
\end{aligned} \tag{6.89}$$

with

$$D(q,m) \equiv q^2 - m^2, \tag{6.90}$$

and

$$N^\mu(p,q,l) \equiv e^3 \bar{u}_r(q)\gamma^\alpha \left(\not{q} - \not{l} + m\right) \gamma^\mu \left(\not{p} - \not{l} + m\right) \gamma_\alpha u_s(p) \tag{6.91}$$

Feynman's trick enables writing the denominator as a power of a single propagator:

$$\frac{1}{D_1 ... D_n} = (n-1)! \int_0^1 d\alpha_1 \int_0^1 d\alpha_2 .... \int_0^1 d\alpha_n \frac{\delta(1 - \sum_i \alpha_i)}{(D_1\alpha_1 + ... + D_n\alpha_n)^n} \tag{6.92}$$

which applied to the case of three factors leads to

$$\frac{1}{D_1 D_2 D_3} = 2 \int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2 \; \frac{1}{[D_1\alpha_1 + D_2\alpha_2 + D_3(1 - \alpha_1 - \alpha_2)]^3} \tag{6.93}$$

Applying this to eq. (6.89) we find

$$D(q-l,m)\alpha_1 + D(p-l,m)\alpha_2 + D(l,0)(1-\alpha_1-\alpha_2) = D(\bar{l},\Delta), \tag{6.94}$$

with

$$\bar{l} \equiv l - \alpha_1 q - \alpha_2 p, \quad \Delta \equiv m(\alpha_1+\alpha_2), \tag{6.95}$$

where we have used the kinematic relations $k^2 = 0, p^2 = q^2 = m^2, 2pq = 2m^2$. The amplitude can then be simplified to

$$\mathcal{M} = \int \frac{d^4l}{(2\pi)^4} \frac{N^\mu(p,q,l)\epsilon_{\lambda\mu}(k)}{[D(\bar{l},\Delta)]^3}. \tag{6.96}$$

We can now change the integration variable $l \to \bar{l}$,

$$\mathcal{M} = 2\int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2 \int \frac{d^4\bar{l}}{(2\pi)^4} \frac{N^\mu(p,q,\bar{l}+\alpha_1 q+\alpha_2 p)\epsilon_{\lambda\mu}(k)}{D(\bar{l},\Delta)^3}. \tag{6.97}$$

Using the following properties

$$\begin{aligned} \not{p} u_s(p) &= m u_s(p),\ \bar{u}_r(q)\not{q} = m\bar{u}_r(q),\ \{\gamma^\mu, \not{p}\} = 2p^\mu, \\ \gamma^\alpha\gamma^\mu\gamma_\alpha &= -2\gamma^\mu,\ \ \gamma^\alpha\gamma^\mu\gamma^\nu\gamma_\alpha = 4g^{\mu\nu},\ \ \gamma^\alpha\gamma^\mu\gamma^\nu\gamma^\rho\gamma_\alpha = -2\gamma^\rho\gamma^\nu\gamma^\mu \end{aligned} \tag{6.98}$$

it is possible to simplify

$$\begin{aligned} N^\mu(p,q,\bar{l}+\alpha_1 q+\alpha_2 p) = -4e^3 m\bar{u}_r(q) \quad & \big[q^\mu\left((1-\alpha_2)(1-\alpha_1) - (1-\alpha_1-\alpha_1^2)\right) \\ & +p^\mu\left((1-\alpha_2)(1-\alpha_1) - (1-\alpha_2-\alpha_2^2)\right) \\ & +(...)\gamma^\mu + \mathcal{O}(\bar{l}^\mu)\big]\, u_s(p), \end{aligned} \tag{6.99}$$

where we have isolated the terms that are not of the standard form, γ^μ, nor linear in $\bar{l}^\mu$ since the latter will vanish upon integration in $d^4\bar{l}$.

Defining

$$I_n(\Delta) \equiv \int \frac{d^4l}{(2\pi)^4} \frac{1}{D(l,\Delta)^n}. \tag{6.100}$$

We will compute these integrals in the next chapter. Here we simply give the result that we need,

$$I_3(\Delta) = -\frac{i}{32\pi^2\Delta^2}. \tag{6.101}$$

We finally obtain

$$\mathcal{M} = \frac{ie^3}{4\pi^2 m}\bar{u}_r(q)(q^\mu C_1 + p^\mu C_2)u_s(p)\epsilon_\mu(k,\lambda), \tag{6.102}$$

where we have defined

$$C_1 = \int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2 \frac{(1-\alpha_1)(1-\alpha_2)-(1-\alpha_1-\alpha_1^2)}{(\alpha_1+\alpha_2)^2} = -\frac{1}{4}, \quad (6.103)$$

$$C_2 = \int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2 \frac{(1-\alpha_1)(1-\alpha_2)-(1-\alpha_2-\alpha_2^2)}{(\alpha_1+\alpha_2)^2} = -\frac{1}{4}. \quad (6.104)$$

Therefore

$$\mathcal{M} = -ie\frac{\alpha}{4\pi m}(q^\mu + p^\mu). \quad (6.105)$$

By comparing with eq. (6.88) we find the induced anomalous magnetic moment to be

$$\Delta g = \frac{\alpha}{\pi}. \quad (6.106)$$

This simple result was first obtained by J. Schwinger and is also the epitaph printed on his gravestone. It nicely matched the experimental result at the time:

$$\left.\frac{\Delta g}{2}\right|_{\text{Schwinger}} = 0.00116, \quad (6.107)$$

This quantity has since been determined experimentally with impressive accuracy[12]

$$\left.\frac{\Delta g}{2}\right|_{\text{exp}} = 0.00115965218073(28), \quad (6.108)$$

and this precision has been matched by the theoretical prediction in QED –up to $\mathcal{O}(\alpha^5)$:

$$\left.\frac{\Delta g}{2}\right|_{\text{QED}} = 0.001159652181643(764). \quad (6.109)$$

It is clear that QED describes nature!

[12] Hanneke, D. Fogwell Hoogerheide, S. and Gabrielse, G., Physical Review A. 83 (5).

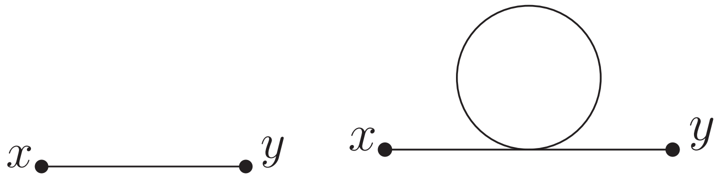


Figure 26: Leading order contribution to the propagator (left) and next-to-leading order (right).

VII. Brief Introduction to Renormalization

In the previous determination of the anomalous magnetic moment we obtained a finite result; however, in the process we ignored corrections that modify the standard coupling, Δe. Had we computed those terms, we would have found unbounded integrals. This is a generic problem in QFT, namely when higher order corrections in perturbation theory are considered, loops appear and the momentum integrals of the form of eq. (6.100) are often divergent. We will see that such ill-defined quantities can be absorbed order by order in a redefinition of the bare parameters in the Lagrangian, such as the fields, the masses and the constant couplings. This is the procedure termed *renormalization.*

To deal with these effects we need to first regularize the perturbative expressions, that is, find a regulator or cutoff for the momentum integrals that make them finite. We can then take the regulator away, after performing the necessary parameter redefinitions.

7.1 One loop corrections in the scalar theory

We will consider for simplicity the real scalar $\lambda\phi^4$ theory,

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m^2}{2}\phi^2 - \frac{\lambda}{4!}\phi^4. \tag{7.2}$$

Let us consider two physical amplitudes at next-to-leading (NLO) order in the Dyson series: the scalar propagator and the two particle scattering amplitude.

The 0-th and 1-st order in the Dyson series for the propagator are depicted in Fig. 26. By applying the Feynman rules it is easy to obtain the result:

$$\begin{aligned}&\langle 0|T(\phi(x)\mathrm{Texp}(-i\int d^4z\mathcal{H}_{\mathrm{int}}(z))\phi(y))|0\rangle_{\mathrm{NLO}}\\ &= \int \frac{d^4q}{(2\pi)^4}\frac{ie^{-iq(x-y)}}{q^2-m^2+i\epsilon}\left(1+\frac{i\Sigma}{q^2-m^2+i\epsilon}\right),\end{aligned} \tag{7.3}$$

with

$$\Sigma \equiv \frac{\lambda}{2} \int \frac{d^4 l}{(2\pi)^4} \frac{1}{l^2 - m^2 + i\epsilon} = \frac{\lambda}{2} I_1(m). \tag{7.4}$$

The integral $I_1(m)$, defined in eq. (6.100), is quadratically divergent, meaning that if the momentum integral is cutoff at a scale Λ, the integral scales as Λ^2

The NLO contributions to the scattering amplitude are given by the sum of the Feynman diagrams in Fig. 27

$$\begin{aligned} &\langle p'; k'| T \exp\left(-i \int d^4 z \mathcal{H}_{\rm int}(z)\right) |p; k\rangle_{\rm NLO} \\ &= (2\pi)^4 \delta(p' + k' - p - k) \left[-i\lambda + \frac{\lambda^2}{2} \left(J^{(4)}(p+k) + J^{(4)}(p-p') + J^{(4)}(p-k')\right)\right], \end{aligned} \tag{7.5}$$

with

$$J^{(4)}(q) \equiv \int \frac{d^4 l}{(2\pi)^4} \frac{1}{(l^2 - m^2)((l-q)^2 - m^2)}. \tag{7.6}$$

Using the Feynman trick

$$\frac{1}{D_1 D_2} = \int_0^1 d\alpha \frac{1}{(\alpha D_1 + (1-\alpha) D_2)^2}. \tag{7.7}$$

The integrals can be reduced to the basic integrals, eq. (6.100)

$$J^{(4)}(q) = \int_0^1 d\alpha \; I_2(\Delta_q), \qquad \Delta_q \equiv m^2 - q^2 \alpha (1-\alpha). \tag{7.8}$$

The integral $I_2(q)$ is logarithmically divergent. Before we try to give a meaning to these divergences, we need a method to regularize the integrals or make them finite. The most popular method changes the space-time dimension to $d = 4 - \epsilon$.

7.2 Dimensional regularization

We want to compute the integrals defined in eq. (6.100). For $n \leq 2$ they are divergent. Firstly we do a Wick rotation by changing the integration contour in the temporal momentum direction to the complex axis, as depicted in Fig. 28, and defining $l^0 = ik^0$. In terms of the Euclidean momentum $k \equiv (k^0, l^i)$ the integral becomes

$$I_n(\Delta) = i(-1)^n \int \frac{d^4 k}{(2\pi)^4} \frac{1}{(k^2 + \Delta^2)^n}, \tag{7.9}$$

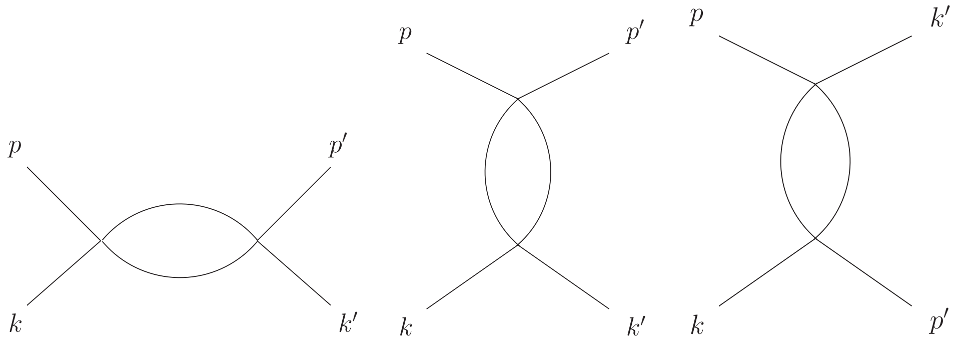


Figure 27: Next-to-leading order contributions to the scattering amplitude.

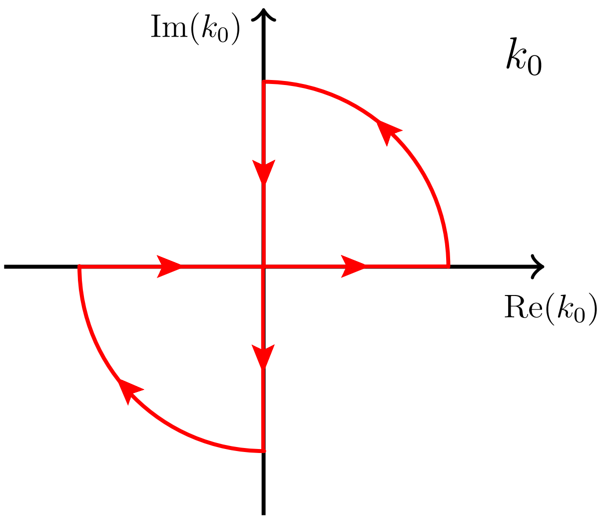


Figure 28: Wick rotation.

where $k^2 = k_0^2 + k_i^2$. To solve this integral we can go to spherical coordinates, $|k| \equiv \sqrt{k_0^2 + k_1^2 + k_2^3 + k_3^2}$:

$$\begin{aligned}\int d^4k \frac{1}{(k^2+m^2)^n} &= \int d\Omega_3 \int_0^\infty d|k| \frac{|k|^3}{(|k|^2+m^2)^n} = C_3 \int_0^\infty d|k| \frac{|k|^3}{(|k|^2+m^2)^n} \\ &= \frac{C_3}{2(n-1)(n-2)(m^2)^{n-2}}, \end{aligned} \tag{7.10}$$

where we have assumed that $n > 2$, since otherwise it diverges.

We can compute C_3 by using for example a simpler function to integrate

$$\int d^4k e^{-k^2/2} = \prod_{i=0-3} \int_{-\infty}^{\infty} dk_i e^{-k_i^2/2} = (\sqrt{2\pi})^4, \tag{7.11}$$

but also

$$\int d^4k e^{-k^2/2} = C_3 \int_0^\infty d|k||k|^3 e^{-|k|^2/2} = 2C_3, \tag{7.12}$$

and therefore

$$C_3 = 2\pi^2. \tag{7.13}$$

The integral for $n = 3$ is then

$$I_3(\Delta) = -\frac{i}{32\pi^2\Delta^2}. \tag{7.14}$$

What happens if $n \leq 2$? The integrals diverge but can be formally evaluated if we take the dimension of the Euclidean space-time to be $d \neq 4$:

$$\int \frac{d^dk}{(2\pi)^d} \frac{1}{(k^2+m^2)^n} = \frac{C_{d-1}}{(2\pi)^d} \int d|k| \frac{|k|^{d-1}}{(|k|^2+m^2)^n} = \frac{C_{d-1}}{(2\pi)^d} \frac{\Gamma\left(\frac{d}{2}\right)\Gamma\left(n-\frac{d}{2}\right)}{2\Gamma(n)(m^2)^{n-d/2}}. \tag{7.15}$$

and where by the same argument as above

$$C_{d-1} = \frac{(2\pi)^{d/2}}{2^{d/2-1}\Gamma\left(\frac{d}{2}\right)}, \tag{7.16}$$

and

$$I_n(\Delta) = \frac{i(-1)^n \Gamma\left(n-\frac{d}{2}\right)}{(4\pi)^{d/2}\Gamma(n)(\Delta^2)^{n-d/2}}. \tag{7.17}$$

Obviously as $d \to 4$, the function should diverge for $n \leq 2$.

For $n = 1$, and $d = 4 - \epsilon$, we obtain

$$I_1(\Delta) = -i\frac{\Gamma(-1+\epsilon/2)(\Delta^2)^{1-\epsilon/2}}{(4\pi)^{2-\epsilon/2}} + \mathcal{O}(\epsilon), \tag{7.18}$$

and expanding for small ϵ,

$$\begin{aligned}\Gamma(-1+\epsilon/2) &= \frac{-2}{\epsilon} - \gamma + 1 + \mathcal{O}(\epsilon), \\ \left(\frac{\Delta^2}{4\pi}\right)^{-\epsilon/2} &= 1 - \frac{\epsilon}{2}\log\left(\frac{\Delta^2}{4\pi}\right) + \mathcal{O}(\epsilon),\end{aligned} \tag{7.19}$$

so finally

$$I_1(\Delta) = -i\frac{\Delta^2}{(4\pi)^2}\left(-\frac{2}{\epsilon} - \gamma + 1 + 2\log\Delta - \log 4\pi\right). \tag{7.20}$$

For $n = 2$ instead

$$I_2(\Delta) = i\frac{\Gamma\left(\frac{\epsilon}{2}\right)}{(4\pi)^{2-\epsilon/2}(\Delta^2)^{\epsilon/2}}, \tag{7.21}$$

and using

$$\Gamma\left(\frac{\epsilon}{2}\right) = \frac{2}{\epsilon} - \gamma + \mathcal{O}(\epsilon), \tag{7.22}$$

we find

$$I_2(\Delta) = \frac{i}{(4\pi)^2}\left(\frac{2}{\epsilon} - \gamma - 2\log\Delta + \log 4\pi\right). \tag{7.23}$$

7.3 Renormalization

We can now compute the correction of the propagator in eq. (7.3),

$$i\Sigma = \frac{\lambda m^2}{32\pi^2}\left(-\frac{2}{\epsilon} - \gamma + 1 + 2\log m - \log 4\pi\right), \tag{7.24}$$

and at the level we are working, we can write it as:

$$\langle 0|T(\phi(x)\phi(y))|0\rangle_{\rm NLO} \simeq \int \frac{d^4q}{(2\pi)^4} e^{-iq(x-y)} \frac{i}{q^2 - m^2 - i\Sigma + i\epsilon} + \mathcal{O}(\lambda^2). \tag{7.25}$$

If we add a term in the Lagrangian, eq. (7.2), of the form

$$\Delta\mathcal{L}_{c.t} \supset -\frac{1}{2}\delta m^2\phi^2, \tag{7.26}$$

we can absorb the divergence in δm^2, by requiring

$$\delta m^2 = -i\Sigma. \tag{7.27}$$

The bare mass parameter is, after this redefinition, $m^2 + \delta m^2$, which is divergent, but the physical mass is m (the pole of the propagator at NLO).

Regarding the scattering amplitude in eq. (7.5)

$$J^{(4)}(q) = \frac{i}{(4\pi)^2}\left(\frac{2}{\epsilon} - \gamma + \log 4\pi - \int_0^1 d\alpha \log[m^2 - q^2\alpha(1-\alpha)]\right). \tag{7.28}$$

we can expand for $q^2 \ll m^2$ to get

$$\begin{aligned} J^{(4)}(q) &= \frac{i}{(4\pi)^2}\left(\frac{2}{\epsilon} - \gamma + \log 4\pi - \log m^2 + \frac{q^2}{m^2}\int_0^1 d\alpha\ \alpha(1-\alpha) + ...\right) \\ &= \frac{i}{(4\pi)^2}\left(\frac{2}{\epsilon} - \gamma + \log 4\pi - \log m^2 + \frac{q^2}{6m^2} + ...\right). \end{aligned} \tag{7.29}$$

Therefore

$$\mathcal{M} = -i\lambda + i\frac{3\lambda^2}{32\pi^2}\left(\frac{2}{\epsilon} - \gamma + \log 4\pi\right) - i\frac{3\lambda^2}{32\pi^2}\log m^2 + i\frac{\lambda^2}{196\pi^2}\frac{s+t+u}{m^2} + ... \tag{7.30}$$

We can absorb the divergence by another counterterm of the form:

$$\Delta\mathcal{L}_{c.t} \supset -\frac{\delta\lambda}{4!}\phi^4. \tag{7.31}$$

with the minimal subtraction scheme choice:

$$\delta\lambda = \frac{3\lambda^2}{32\pi^2}\left(\frac{2}{\epsilon} - \gamma + \log 4\pi\right). \tag{7.32}$$

The bare coupling in the Lagrangian is now $\lambda + \delta\lambda$, and therefore divergent, but the scattering amplitude is now well defined because the modification introduced by the counterterm adds a $-i\delta\lambda$ to the amplitude that precisely cancels the second term in eq. (7.30).

There is arbitrariness in how we choose the counterterms δm^2 and $\delta\lambda$. Each choice is called a scheme and physics should not depend on which scheme we choose.

For example we could also define the physical coupling directly from $\mathcal{M}(s,t,u)$ at some kinematical point, for example:

$$-i\lambda_R^{\text{NLO}} \equiv \mathcal{M}(s,t,u)|_{s=4m^2,t=0,u=0}. \tag{7.33}$$

This would force λ to absorb the divergence term.

Independently of the choice of scheme, it is possible to show that any other physical observable computed at the same order in the perturbative expansion will be finite when expressed in terms of the renormalized couplings.

In our minimal subtraction scheme, the counterterms, eq. (7.26) and eq. (7.31), with the choice of eq. (7.27) and eq. (7.32), will suffice to make any other quantity finite when computed at the same order in perturbation theory, and expressed in terms of λ, m (i.e. the renormalized parameters in our scheme).

What we have found in this simple model and at NLO can be generalized to all orders in the perturbative expansion, and to more complex models. This is the property of quantum field theories that we refer to as renormalizability.

In this theory, this means that at any order in the perturbative expansion, the divergences disappear if counterterms of the most general form of the original Lagrangian are appropriately tuned,

$$\Delta\mathcal{L}_{c.t} = \delta Z \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}\delta m^2\phi^2 - \frac{\delta\lambda}{4!}\phi^2, \tag{7.34}$$

that ammounts to a rescaling of the field, and a mass and coupling renormalization.

The proof to all orders is complicated but we can give some heuristic arguments based on power counting.

Consider a diagram with E external legs, L loops, V vertices and I propagators. Each vertex has four legs and since all legs are contracted,

$$4V = E + 2I. \tag{7.35}$$

The superficial degree of divergence of a diagram comes from the loop integrations, so when all loop momenta are large we expect the following behavior

$$\left(\int_{-\Lambda}^{\Lambda} d^4p\right)^L \left(\frac{1}{p^2}\right)^I \sim \Lambda^{4L-2I} \equiv \Lambda^\omega, \tag{7.36}$$

where Λ is the cutoff of the momentum integrals and $\omega \equiv 4L - 2I$ is the superficial degree of divergence. The finiteness of the diagram requires that $\omega < 0$, although this is a necessary but not sufficient condition.

However, the number of loops is related to the number of momentum integrals that cannot be performed by using momentum conservation at the vertices. For each vertex there is a delta function, while for each internal line there is four-momentum integration. One of the deltas in the diagram is always the delta of external momentum conservation, and therefore there are effectively only $V - 1$ deltas, that can be used to reduce the propagator momentum integrals. Therefore we have

$$L = I - (V - 1) = I - V + 1. \tag{7.37}$$

Using eqs. (7.35) and (7.37), we get

$$\omega = 4L - 2I = 2I - 4V + 4 = 2I - (E + 2I) + 4 = 4 - E. \tag{7.38}$$

The superficially divergent diagrams are those with $E \leq 4$. The divergent part of a diagram with E external legs can be represented by local operators with E external fields, and the only terms with $E \leq 4$ and Lorentz invariant are precisely the counterterms in eq. (7.34).

A similar analysis in the case of a theory with non-renormalizable terms such as ϕ^6 will imply that ω grows with the number of loops and that diagrams with $E > 4$ can also be superficially divergent.

The formal proof of renormalizability is much more complicated that this naive analysis because diagrams can have sub-divergences, and a recursive procedure is needed to ensure the finiteness to all orders.

Renormalizability has been proven for all theories of interest in physics: $\lambda\phi^4$ and its complex analogue, QED, as well as Yang-Mills theories such as those that describe the weak and strong interactions.

7.4 Wilsonian renormalization or emerging renormalizability

Renormalizable theories contain a very small subset of all possible interactions compatible with Lorentz and gauge symmetries. Based on these symmetries we could have added many more interactions:

$$\mathcal{L} = \mathcal{L}_{\text{renorm.}} + \sum_{d>4} \frac{1}{\Lambda^{d-4}} O^d. \tag{7.39}$$

The renormalizable theory in this larger context is seen as an effective theory valid for $E \ll \Lambda$, since by dimensional analysis the higher dimensional operators induce effects of

$$\mathcal{O}\left(\frac{E}{\Lambda}\right)^{d-4}. \tag{7.40}$$

In fact, the theory defined with a cutoff is generically of the form eq. (7.39). For example we can think of defining the QFT on a discretized space-time, that is a lattice of spacing a, which implies a cutoff of the form a^{-1}. Quantum fields can only take values on the sites of this lattice. The QFT is then reduced to a spin system as in a condense matter model. Generically any discretized version of the continuum QFT Lagrangian will contain higher dimensional operators suppressed with a, and for example, if we define the derivative as discrete difference:

$$\phi(x + a\hat{\mu}) - \phi(x) = a\partial_\mu\phi(x) + a^2\partial_\mu^2\phi + ... \tag{7.41}$$

and will therefore be of the form eq. (7.39).

What is the meaning of renormalizability in this regularized definition of the theory?

Renormalizability in this context is related to the property of criticality and universality of the statistical system. A critical point corresponds to a value of the couplings, where long-range correlations develop. When this happens, the physics at long distances does not depend on the details of the interactions we choose to include in the Lagrangian but fall in large universality classes that are essentially fixed by symmetries. This means that we can modify the Lagrangian by changing the couplings of the different terms, but providing we do not change the symmetries, the long-range physics at the critical point will be the same. This universal limit is what a renormalizable QFT is expected to describe.

$$\text{Universality classes} \leftrightarrow \text{Renormalizable QFTs}$$

From this new viewpoint, there is nothing special about a renormalizable Lagrangian, as long as the higher dimensional operators are suppressed by the cutoff – and therefore irrelevant at the critical point. We could repeat the power-counting argument we did above, but for the theory in eq. (7.39) and, providing we keep track of the powers of Λ in the couplings, as well as those coming from momentum integrations, the result will still be that of eq. (7.38).