Quantum Field Theory II

François Gelis

Abstract

The first part of this course will be devoted to the path integral approach in quantum field theory. Starting with elementary examples in quantum mechanics, we will firstly revisit the theory of an interacting scalar field with functional methods. Then, we will discuss the case of fermionic fields and introduce Grassman variables. Focusing on quantum electrodynamics, we will use these functional techniques in order to derive the Schwinger-Dyson equations.

The part on non-abelian gauge fields will start with a discussion of non-abelian Lie groups and algebras, their representations and of the field structures that admit such a local symmetry. Then, we will proceed with the path integral quantization of a vector field with an SU(N) local gauge symmetry, following the Fadeev-Popov method. After having derived the Feynman rules, we will discuss the unitarity and renormalizability of such a field theory, focusing on the aspects that make them different from simpler theories such as QED. An important step for this will be the BRST symmetry. The course will finish with an introduction to lattice field theory.

Contents

Chapter 1

Reminders of QFT I

1.1 Special relativity

1.1.1 Lorentz transformations

Special relativity plays a crucial role in quantum field theories[1]. Various observers in frames that are moving at a constant speed relative to each other should be able to describe physical phenomena using the same laws of Physics. This does not imply that the equations governing these phenomena are independent of the observer's frame, but that these equations transform in a constrained fashion –depending on the nature of the objects they contain– under a change of reference frame.

Let us consider two frames $\mathcal{F}$ and $\mathcal{F}'$, in which the coordinates of a given event are respectively x^μ and x'^μ. A *Lorentz transformation* is a linear transformation such that the interval $ds^2 \equiv dt^2 - d\mathbf{x}^2$ is the same in the two frames[2]. If we denote the coordinate transformation by

$$x'^\mu = \Lambda^\mu{}_\nu \, x^\nu \,, \tag{1.1}$$

the matrix Λ of the transformation must obey

$$g^{\mu\nu} = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \, g^{\rho\sigma} \tag{1.2}$$

where $g_{\mu\nu}$ is the Minkowski metric tensor

$$g_{\mu\nu} \equiv \begin{pmatrix} +1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \,. \tag{1.3}$$

---

[1] An exception to this assertion is for quantum field models applied to condensed matter physics, where the basic degrees of freedom are to a very good level of approximation described by Galilean kinematics.

[2] The physical premises of special relativity require that the speed of light be the same in all inertial frames, which implies solely that $ds^2 = 0$ be preserved in all inertial frames. The group of transformations that achieves this is called the *conformal group*. In four space-time dimensions, the conformal group is 15 dimensional, and in addition to the 6 orthochronous Lorentz transformations it contains dilatations as well as non-linear transformations called *special conformal transformations*.

If we consider an infinitesimal Lorentz transformation,

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu \tag{1.4}$$

(with all components of ω much smaller than unity), this implies that

$$\omega_{\mu\nu} = -\omega_{\nu\mu} \tag{1.5}$$

(with all indices down). Consequently, there are 6 independent Lorentz transformations, three of which are ordinary rotations and three are boosts. Note that the infinitesimal transformations (1.4) have a determinant[3] equal to $+1$ (they are called *proper* transformations), and do not change the direction of the time axis since $\Lambda^0{}_0 = 1 \geq 0$ (they are called *orthochronous*). Any combination of such infinitesimal transformations shares the same properties, and their set forms a subgroup of the full group of transformations that preserve the Minkowski metric.

1.1.2 Representations of the Lorentz group

More generally, a Lorentz transformation acts on a quantum system via a transformation $U(\Lambda)$, that forms a representation of the Lorentz group, i.e.

$$U(\Lambda\Lambda') = U(\Lambda)U(\Lambda') . \tag{1.6}$$

For an infinitesimal Lorentz transformation, we can write

$$U(1 + \omega) = I + \frac{i}{2}\omega_{\mu\nu}M^{\mu\nu} . \tag{1.7}$$

(The prefactor $i/2$ in the second term of the right hand side is conventional.) Since the $\omega_{\mu\nu}$ are antisymmetric, the generators $M^{\mu\nu}$ can also be chosen antisymmetric. By using eq. (1.6) for the Lorentz transformation $\Lambda^{-1}\Lambda'\Lambda$, we arrive at

$$U^{-1}(\Lambda)M^{\mu\nu}U(\Lambda) = \Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma M^{\rho\sigma} , \tag{1.8}$$

indicating that $M^{\mu\nu}$ transforms as a rank-2 tensor. When used with an infinitesimal transformation $\Lambda = 1 + \omega$, this identity leads to the commutation relation that defines the Lie algebra of the Lorentz group

$$\left[M^{\mu\nu}, M^{\rho\sigma}\right] = i(g^{\mu\rho}M^{\nu\sigma} - g^{\nu\rho}M^{\mu\sigma}) - i(g^{\mu\sigma}M^{\nu\rho} - g^{\nu\sigma}M^{\mu\rho}) . \tag{1.9}$$

In a fashion similar to eq. (1.8), we can obtain the transformation of the 4-impulsion P^μ,

$$U^{-1}(\Lambda)P^\mu U(\Lambda) = \Lambda^\mu{}_\rho P^\rho , \tag{1.10}$$

which leads to the following commutation relation between P^μ and $M^{\mu\nu}$,

$$\left[P^\mu, M^{\rho\sigma}\right] = i(g^{\mu\sigma}P^\rho - g^{\mu\rho}P^\sigma) . \tag{1.11}$$

A *scalar field* $\phi(x)$ is a (number or operator valued) object that depends on a spacetime coordinate x and is invariant under a Lorentz transformation, except for the change of coordinate induced by the transformation:

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

---

[3]From eq. (1.2), the determinant may be equal to ± 1.

This formula just reflects the fact that the point x where the transformed field is evaluated was located at the point $\Lambda^{-1}x$ before the transformation. The first derivative $\partial^\mu\phi$ of the field transforms as a 4-vector,

$$U^{-1}(\Lambda)\partial^\mu\phi(x)U(\Lambda) = \Lambda^\mu_{\ \nu}\partial^{\overline{\nu}}\phi(\Lambda^{-1}x) \ , \tag{1.13}$$

where the bar in $\partial^{\overline{\nu}}$ indicates that we are differentiating with respect to the whole argument of ϕ, i.e. $\Lambda^{-1}x$. Likewise, the second derivative $\partial^\mu\partial^\nu\phi$ transforms like a rank-2 tensor, but the D'Alembertian $\Box\phi$ transforms as a scalar.

1.2 Free scalar fields, Mode decomposition

1.2.1 Quantum harmonic oscillators

Let us consider a continuous collection of quantum harmonic oscillators, each of them corresponding to particles with a given momentum $\mathbf{p}$. These harmonic oscillators can be defined by a pair of creation and annihilation operators $a_\mathbf{p}^\dagger, a_\mathbf{p}$, where $\mathbf{p}$ is a 3-momentum that labels the corresponding mode. Note that the energy of the particles is fixed from their 3-momentum by the relativistic dispersion relation,

$$p^0 = E_\mathbf{p} \equiv \sqrt{\mathbf{p}^2 + m^2} \ . \tag{1.14}$$

The operators creating or destroying particles with a given momentum $\mathbf{p}$ obey usual commutation relations,

$$\left[a_\mathbf{p}, a_\mathbf{p}\right] = \left[a_\mathbf{p}^\dagger, a_\mathbf{p}^\dagger\right] = 0 \ , \quad \left[a_\mathbf{p}, a_\mathbf{p}^\dagger\right] \sim 1 \ . \tag{1.15}$$

(in the last commutator, the precise normalization will be defined later.) In contrast, operators acting on different momenta always commute:

$$\left[a_\mathbf{p}, a_\mathbf{q}\right] = \left[a_\mathbf{p}^\dagger, a_\mathbf{q}^\dagger\right] = \left[a_\mathbf{p}, a_\mathbf{q}^\dagger\right] = 0 \ . \tag{1.16}$$

If we denote by $\mathcal{H}$ the Hamiltonian operator of such a system, the property that $a_\mathbf{p}^\dagger$ creates a particle of momentum $\mathbf{p}$ (and therefore of energy $E_\mathbf{p}$) implies that

$$\left[\mathcal{H}, a_\mathbf{p}^\dagger\right] = +E_\mathbf{p}a_\mathbf{p}^\dagger \ . \tag{1.17}$$

Likewise, since $a_\mathbf{p}$ destroys a particle with the same energy, we have

$$\left[\mathcal{H}, a_\mathbf{p}\right] = -E_\mathbf{p}a_\mathbf{p} \ . \tag{1.18}$$

(Implicitly in these equations is the fact that particles are non-interacting, so that adding or removing a particle of momentum $\mathbf{p}$ does not affect the rest of the system.) In these lectures, we will adopt the following normalization for the free Hamiltonian[4],

$$\mathcal{H} = \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_\mathbf{p}} \ E_\mathbf{p} \ \left(a_\mathbf{p}^\dagger a_\mathbf{p} + \mathcal{V}\,E_\mathbf{p}\right) \ , \tag{1.19}$$

---

[4]In a relativistic setting, the measure $d^3\mathbf{p}/(2\pi)^3 2E_\mathbf{p}$ has the important benefit of being Lorentz invariant. Moreover, it results naturally from the 4-dimensional momentum integration $d^4p/(2\pi)^4$ constrained by the positive energy mass-shell condition $2\pi\,\theta(p^0)\,\delta(p^2 - m^2)$.

where $\mathcal{V}$ is the volume of the system. To make contact with the usual treatment[5] of a harmonic oscillator in quantum mechanics, it is useful to introduce the *occupation number* $f_{\mathbf{p}}$ defined by,

$$2E_{\mathbf{p}}\,\mathcal{V}\,f_{\mathbf{p}} \equiv a_{\mathbf{p}}^{\dagger}a_{\mathbf{p}}\;. \tag{1.20}$$

In terms of $f_{\mathbf{p}}$, the above Hamiltonian reads

$$\mathcal{H} = \mathcal{V}\int \frac{d^{3}\mathbf{p}}{(2\pi)^{3}}\; E_{\mathbf{p}}\;\left(f_{\mathbf{p}} + \tfrac{1}{2}\right)\;. \tag{1.21}$$

The expectation value of $f_{\mathbf{p}}$ has the interpretation of the number of particles par unit of phase-space (i.e. per unit of volume in coordinate space and per unit of volume in momentum space), and the $1/2$ in $f_{\mathbf{p}} + \tfrac{1}{2}$ is the ground state occupation of each oscillator[6]. Of course, this additive constant is to a large extent irrelevant since only energy differences have a physical meaning. Given eq. (1.19), the commutation relations (1.17) and (1.18) are fulfilled provided that

$$\left[a_{\mathbf{p}}, a_{\mathbf{q}}^{\dagger}\right] = (2\pi)^{3}\,2E_{\mathbf{p}}\,\delta(\mathbf{p} - \mathbf{q})\;. \tag{1.22}$$

1.2.2 Scalar field operator, Canonical commutation relations

Note that in quantum mechanics, a particle with a well defined momentum $\mathbf{p}$ is not localized at a specific point in space, due to the uncertainty principle. Thus, when we say that $a_{\mathbf{p}}^{\dagger}$ creates a particle of momentum $\mathbf{p}$, this production process may happen anywhere in space and at any time since the energy is also well defined. Instead of using the momentum basis, one may introduce an operator that depends on space-time in order to give preeminence to the time and position at which a particle is created or destroyed. It is possible to encapsulate all the $a_{\mathbf{p}}, a_{\mathbf{p}}^{\dagger}$ into the following Hermitean operator[7]

$$\phi(x) \equiv \int \frac{d^{3}\mathbf{p}}{(2\pi)^{3}2E_{\mathbf{p}}}\;\left[a_{\mathbf{p}}^{\dagger}\,e^{+ip\cdot x} + a_{\mathbf{p}}\,e^{-ip\cdot x}\right]\;, \tag{1.23}$$

where $p\cdot x \equiv p_{\mu}x^{\mu}$ with $p^{0} = +E_{\mathbf{p}}$. In the following, we will also need the time derivative of this operator, denoted $\Pi(x)$,

$$\Pi(x) \equiv \partial_{0}\phi(x) = i\int \frac{d^{3}\mathbf{p}}{(2\pi)^{3}2E_{\mathbf{p}}}\;E_{\mathbf{p}}\;\left[a_{\mathbf{p}}^{\dagger}\,e^{+ip\cdot x} - a_{\mathbf{p}}\,e^{-ip\cdot x}\right]\;. \tag{1.24}$$

---

[5]In relativistic quantum field theory, it is customary to use a system of units in which $\hbar = 1$, $c = 1$ (and also $k_{\mathrm{B}} = 1$ when the Boltzmann constant is needed to relate energies and temperature). In this system of units, the action $\mathcal{S}$ is dimensionless. Mass, energy, momentum and temperature have the same dimension, which is the inverse of the dimension of length and duration:

$$[\text{mass}] = [\text{energy}] = [\text{momentum}] = [\text{temperature}] = [\text{length}^{-1}] = [\text{duration}^{-1}]\;.$$

Moreover, in four dimensions, the creation and annihilation operators introduced in eq. (1.19) have the dimension of an inverse energy:

$$[a_{\mathbf{p}}] = [a_{\mathbf{p}}^{\dagger}] = [\text{energy}^{-1}]$$

(the occupation number $f_{\mathbf{p}}$ is dimensionless.)

[6]This is reminiscent of the fact that the energy of the level n in a quantized harmonic oscillator of base energy ω is $E_{n} = (n + \tfrac{1}{2})\omega$.

[7]In four space-time dimensions, this field has the same dimension as energy:

$$[\phi(x)] = [\text{energy}]\;.$$

Given the commutation relation (1.22), we obtain the following *equal-time* commutation relations for ϕ and Π,

$$\left[\phi(x), \phi(y)\right]_{x^0=y^0} = \left[\Pi(x), \Pi(y)\right]_{x^0=y^0} = 0 \ , \ \left[\phi(x), \Pi(y)\right]_{x^0=y^0} = i\delta(\mathbf{x}-\mathbf{y}) \ . \ (1.25)$$

These are called the *canonical field commutation relations*. In this approach (known as *canonical quantization*), the quantization of a field theory corresponds to promoting the classical Poisson bracket between a dynamical variable and its conjugate momentum to a commutator:

$$\left\{P_i, Q_j\right\} = \delta_{ij} \quad \rightarrow \quad \left[\hat{P}_i, \hat{Q}_j\right] = i\hbar\,\delta_{ij} \ . \tag{1.26}$$

In addition to these relations that hold for equal times, one may prove that $\phi(x)$ and $\Pi(y)$ commute for space-like intervals $(x-y)^2 < 0$. Physically, this is related to the absence of causal relation between two measurements performed at space-time points with a space-like separation.

It is possible to invert eqs. (1.23) and (1.24) in order to obtain the creation and annihilation operators given the operators ϕ and Π. These inversion formulas read

$$a_{\mathbf{p}}^{\dagger} = -i\int d^3\mathbf{x}\, e^{-ip\cdot x}\left[\Pi(x) + iE_{\mathbf{p}}\phi(x)\right] = -i\int d^3\mathbf{x}\, e^{-ip\cdot x}\ \overset{\leftrightarrow}{\partial_0}\ \phi(x) \ ,$$

$$a_{\mathbf{p}} = +i\int d^3\mathbf{x}\, e^{+ip\cdot x}\left[\Pi(x) - iE_{\mathbf{p}}\phi(x)\right] = +i\int d^3\mathbf{x}\, e^{+ip\cdot x}\ \overset{\leftrightarrow}{\partial_0}\ \phi(x) \ , \ (1.27)$$

where the operator $\overset{\leftrightarrow}{\partial_0}$ is defined as

$$A\ \overset{\leftrightarrow}{\partial_0}\ B \equiv A\left(\partial_0 B\right) - \left(\partial_0 A\right) B \ . \tag{1.28}$$

Note that these expressions, although they appear to contain x^0, do not actually depend on time. Using these formulas, we can rewrite the Hamiltonian in terms of ϕ and Π,

$$\mathcal{H} = \int d^3\mathbf{x}\ \left\{\tfrac{1}{2}\Pi^2(x) + \tfrac{1}{2}(\boldsymbol{\nabla}\phi(x))^2 + \tfrac{1}{2}m^2\phi^2(x)\right\} \ . \tag{1.29}$$

From this Hamiltonian, one may obtain equations of motion in the form of Hamilton-Jacobi equations. Formally, they read

$$\partial_0\phi(x) = \frac{\delta\mathcal{H}}{\delta\Pi(x)} = \Pi(x) \ ,$$

$$\partial_0\Pi(x) = -\frac{\delta\mathcal{H}}{\delta\phi(x)} = \left(\boldsymbol{\nabla}^2 + m^2\right)\phi(x) \ . \tag{1.30}$$

1.2.3 Lagrangian formulation

One may also obtain a Lagrangian $\mathcal{L}(\phi, \partial_0\phi)$ that leads to the Hamiltonian (1.29) by the usual manipulations. Firstly, the momentum canonically conjugated to $\phi(x)$ should be given by as

$$\Pi(x) \equiv \frac{\delta\mathcal{L}}{\delta\partial_0\phi(x)} \ . \tag{1.31}$$

For this to be consistent with the first Hamilton-Jacobi equation, the Lagrangian must contain the following kinetic term

$$\mathcal{L} = \int d^3\mathbf{x} \, \tfrac{1}{2} \left(\partial_0 \phi(x)\right)^2 + \cdots \tag{1.32}$$

The missing potential term of the Lagrangian is obtained by requesting that we have

$$\mathcal{H} = \int d^3\mathbf{x} \, \Pi(x)\partial_0\phi(x) - \mathcal{L} \, . \tag{1.33}$$

This gives the following Lagrangian,

$$\mathcal{L} = \int d^3\mathbf{x} \, \left\{ \tfrac{1}{2}(\partial_\mu\phi(x))(\partial^\mu\phi(x)) - \tfrac{1}{2}m^2\phi^2(x) \right\} \, . \tag{1.34}$$

Note that the action.

$$S = \int dx^0 \, \mathcal{L} \, , \tag{1.35}$$

is a Lorentz scalar (this is not true of the Hamiltonian, which may be considered as the time component of a 4-vector from the point of view of Lorentz transformations). The Lagrangian (1.34) leads to the following Euler-Lagrange equation of motion,

$$\left(\Box_x + m^2\right)\phi(x) = 0 \, , \tag{1.36}$$

which is known as the *Klein-Gordon equation*. This equation is of course equivalent to the pair of Hamilton-Jacobi equations derived earlier.

1.3 Interacting scalar fields, Interaction representation

1.3.1 Interaction term

Until now, we have only considered non-interacting particles, which is of course of very limited use in practice. That the Hamiltonian (1.19) does not contain interactions follows from the fact that the only non-trivial term it contains is of the form $a_\mathbf{p}^\dagger a_\mathbf{p}$, that destroys a particle of momentum $\mathbf{p}$ and then creates a particle of momentum $\mathbf{p}$ (hence nothing changes in the state of the system under consideration). By momentum conservation, this is the only allowed Hermitian operator which is quadratic in the creation and annihilation operators. Therefore, in order to include interactions, we must include in the Hamiltonian terms of higher degree in the creation and annihilation operators. The additional term must be Hermitean, since $\mathcal{H}$ generates the time evolution, which must be unitary.

The simplest Hermitean addition to the Hamiltonian is a term of the form

$$\mathcal{H}_{\scriptscriptstyle I} = \int d^3\mathbf{x} \, \frac{\lambda}{n!} \phi^n(x) \, , \tag{1.37}$$

where n is a power larger than 2. The constant λ is called a *coupling constant* and controls the strength of the interactions, while the denominator $n!$ is a symmetry factor that will prove convenient later on. At this point, it seems that any degree n may provide a reasonable interaction term. However, theories with an odd n have an unstable vacuum, and theories

with $n > 4$ are non-renormalizable in four space-time dimensions, as we shall see later. For these reasons, $n = 4$ is the only case which is widely studied in practice, and we will stick to this value in the rest of this chapter.

With this choice, the Hamiltonian and Lagrangian read

$$\mathcal{H} = \int d^3\mathbf{x} \ \left\{ \tfrac{1}{2}\Pi^2(x) + \tfrac{1}{2}(\boldsymbol{\nabla}\phi(x))^2 + \tfrac{1}{2}m^2\phi^2(x) + \tfrac{\lambda}{4!}\phi^4(x) \right\} \ ,$$

$$\mathcal{L} = \int d^3\mathbf{x} \ \left\{ \tfrac{1}{2}(\partial_\mu\phi(x))(\partial^\mu\phi(x)) - \tfrac{1}{2}m^2\phi^2(x) - \tfrac{\lambda}{4!}\phi^4(x) \right\} \ , \tag{1.38}$$

and the Klein-Gordon equation is modified into

$$\left(\Box_x + m^2\right)\phi(x) + \frac{\lambda}{6}\phi^3(x) = 0 \ . \tag{1.39}$$

1.3.2 Interaction representation

A field operator that obeys this non-linear equation of motion can no longer be represented as a linear superposition of plane waves such as (1.23). Let us assume that the coupling constant is very slowly time-dependent, in such a way that

$$\lim_{x^0 \to \pm\infty} \lambda = 0 \ . \tag{1.40}$$

What we have in mind here is that λ goes to zero *adiabatically* at asymptotic times, i.e. much slower than all the physically relevant timescales of the theory under consideration. Therefore, at $x^0 = \pm\infty$, the theory is a free theory whose spectrum is made of the eigenstates of the free Hamiltonian. Likewise, the field $\phi(x)$ should be in a certain sense "close to a free field" in these limits. In the case of the $x^0 \to -\infty$ limit, let us denote this by[8]

$$\lim_{x^0 \to -\infty} \phi(x) = \phi_{\text{in}}(x) \ , \tag{1.41}$$

where ϕ_{in} is a free field operator that admits a Fourier decomposition similar to eq. (1.23),

$$\phi_{\text{in}}(x) \equiv \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_{\mathbf{p}}} \left[a_{\mathbf{p},\text{in}}^\dagger \, e^{+ip\cdot x} + a_{\mathbf{p},\text{in}} \, e^{-ip\cdot x} \right] \ . \tag{1.42}$$

Eq. (1.41) can be made more explicit by writing

$$\phi(x) = U(-\infty, x^0) \, \phi_{\text{in}}(x) \, U(x^0, -\infty) \ , \tag{1.43}$$

where U is a unitary *time evolution* operator defined as a time ordered exponential of the interaction term in the Lagrangian, evaluated with the ϕ_{in} field:

$$U(t_2, t_1) \equiv T \exp i \int_{t_1}^{t_2} dx^0 d^3\mathbf{x} \, \mathcal{L}_{\text{I}}(\phi_{\text{in}}(x)) \ , \tag{1.44}$$

---

[8]In this equation, we ignore for now the issue of field renormalization, onto which we shall come back later (see the section 1.8).

where

$$\mathcal{L}_{\scriptscriptstyle I}(\varphi(x)) \equiv -\frac{\lambda}{4!}\ \varphi^4(x) \ . \tag{1.45}$$

This time evolution operator satisfies the following properties

$$
\begin{aligned}
U(t,t) &= 1 \\
U(t_3,t_1) &= U(t_3,t_2)\,U(t_2,t_1) && \text{(for all } t_2) \\
U(t_1,t_2) &= U^{-1}(t_2,t_1) = U^\dagger(t_2,t_1) \ .
\end{aligned}
\tag{1.46}
$$

One can then prove that

$$(\Box_x + m^2)\varphi(x) + \frac{\lambda}{6}\varphi^3(x) = U(-\infty,x^0)\left[(\Box_x + m^2)\varphi_{\text{in}}(x)\right]U(x^0,-\infty) \ . \tag{1.47}$$

This equation shows that φ_{in} obeys the free Klein-Gordon equation if φ obeys the non-linear interacting one, and justifies a posteriori our choice of the unitary operator U that connects φ and φ_{in}.

1.3.3 In and Out states

The *in* creation and annihilation operators can be used to define a space of eigenstates of the free Hamiltonian, starting from a ground state (vacuum) denoted $|0_{\text{in}}\rangle$. For instance, one particle states would be defined as

$$|\mathbf{p}_{\text{in}}\rangle = a^\dagger_{\mathbf{p},\text{in}}|0_{\text{in}}\rangle \ . \tag{1.48}$$

The physical interpretation of these states is that they are states with a definite particle content at $x^0 = -\infty$, before the interactions are turned on[9].

In the same way as we have constructed *in* field operators, creation and annihilation operators and states, we may construct *out* ones such that the field $\varphi_{\text{out}}(x)$ is a free field that coincides with the interacting field $\varphi(x)$ in the limit $x^0 \to +\infty$ (with the same caveat about field renormalization). Starting from a vacuum state $|0_{\text{out}}\rangle$, we may also define a full set of states, such as $|\mathbf{p}_{\text{out}}\rangle$, that have a definite particle content at $x^0 = +\infty$. It is crucial to observe that the in and out states are not identical:

$$|0_{\text{out}}\rangle \neq |0_{\text{in}}\rangle \quad \text{(they differ by the phase } \langle 0_{\text{out}}|0_{\text{in}}\rangle) \ , \quad |\mathbf{p}_{\text{out}}\rangle \neq |\mathbf{p}_{\text{in}}\rangle \,, \cdots \tag{1.49}$$

Taking the limit $x^0 \to +\infty$ in eq. (1.43), we first see that[10]

$$a_{\mathbf{p},\text{out}} = U(-\infty,+\infty)\,a_{\mathbf{p},\text{in}}\,U(+\infty,-\infty) \,, \quad a^\dagger_{\mathbf{p},\text{out}} = U(-\infty,+\infty)\,a^\dagger_{\mathbf{p},\text{in}}\,U(+\infty,-\infty) \,, \tag{1.50}$$

from which we deduce that the in and out states must be related by

$$|\alpha_{\text{out}}\rangle = U(-\infty,+\infty)\,|\alpha_{\text{in}}\rangle \ . \tag{1.51}$$

The two sets of states are identical for a free theory, since the evolution operator reduces to the identity in this case.

---

[9]For an interacting system, it is not possible to enumerate the particle content of states, because of quantum fluctuations that may temporarily create additional virtual particles.

[10]The evolution operator from $x^0 = -\infty$ to $x^0 = +\infty$ is sometimes called the S-matrix: $S \equiv U(+\infty,-\infty)$.

1.4 LSZ reduction formulas

Among the most interesting physical quantities are the transition amplitudes

$$\langle q_1 q_2 \cdots {}_{\text{out}} | p_1 p_2 \cdots {}_{\text{in}} \rangle \, , \tag{1.52}$$

whose squared modulus enters in cross-sections that are measurable in scattering experiments. Up to a normalization factor, the square of this amplitude gives the probability that particles with momenta $p_1 p_2 \cdots$ in the initial state evolve into particles with momenta $q_1 q_2 \cdots$ in the final state.

A first step in view of calculating transition amplitudes is to relate them to expectation values involving the field operator $\phi(x)$. In order to illustrate the main steps in deriving such a relationship, let us consider the simple case of the transition amplitude between two 1-particle states,

$$\langle q_{\text{out}} | p_{\text{in}} \rangle \, . \tag{1.53}$$

Firstly, we write the state $|p_{\text{in}}\rangle$ as the action of a creation operator on the corresponding vacuum state, and we replace the creation operation by its expression in terms of ϕ_{in},

$$
\begin{aligned}
\langle q_{\text{out}} | p_{\text{in}} \rangle &= \langle q_{\text{out}} | a_{p,\text{in}}^\dagger | 0_{\text{in}} \rangle \\
&= -i \int d^3 x \; e^{-ip \cdot x} \langle q_{\text{out}} | \Pi_{\text{in}}(x) + iE_p \phi_{\text{in}}(x) | 0_{\text{in}} \rangle \, .
\end{aligned} \tag{1.54}
$$

Next, we use the fact that $\phi_{\text{in}}, \Pi_{\text{in}}$ are the limits when $x^0 \to -\infty$ of the interacting fields ϕ, Π, and we express this limit by means of the following trick:

$$\lim_{x^0 \to -\infty} F(x^0) = \lim_{x^0 \to +\infty} F(x^0) - \int_{-\infty}^{+\infty} dx^0 \; \partial_{x^0} F(x^0) \, . \tag{1.55}$$

The term with the limit $x^0 \to +\infty$ produces a term identical to the r.h.s. of the first line of eq. (1.54), but with an $a_{p,\text{out}}^\dagger$ instead of $a_{p,\text{in}}^\dagger$. At this stage we have

$$
\begin{aligned}
\langle q_{\text{out}} | p_{\text{in}} \rangle &= \langle 0_{\text{out}} | a_{q,\text{out}} a_{p,\text{out}}^\dagger | 0_{\text{in}} \rangle \\
&\quad + i \int d^4 x \; \partial_{x^0} \; e^{-ip \cdot x} \langle q_{\text{out}} | \Pi(x) + iE_p \phi(x) | 0_{\text{in}} \rangle \, .
\end{aligned} \tag{1.56}
$$

In the first line, we use the commutation relation between creation and annihilation operators to obtain

$$\langle 0_{\text{out}} | a_{q,\text{out}} a_{p,\text{out}}^\dagger | 0_{\text{in}} \rangle = (2\pi)^3 2E_p \, \delta(p - q) \, . \tag{1.57}$$

This term does not involve any interaction, since the initial state particle simply goes through to the final state (in other words, this particle just acts as a spectator in the process). Such trivial terms always appear when expressing transition amplitudes in terms of the field operator, and they are usually dropped since they do not carry any interesting physical information. We can then perform explicitly the time derivative in the second line to obtain[11]

$$\langle q_{\text{out}} | p_{\text{in}} \rangle \doteq i \int d^4 x \; e^{-ip \cdot x} \, (\Box_x + m^2) \langle q_{\text{out}} | \phi(x) | 0_{\text{in}} \rangle \, , \tag{1.58}$$

---

[11] We use here the dispersion relation $p_0^2 - p^2 = m^2$ of the incoming particle to arrive at this expression. The mass that should enter in this formula is the physical mass of the particles. This remark will become important when we discuss renormalization.

where we use the symbol $\doteq$ to indicate that the trivial non-interacting terms have been dropped.

Next, we repeat the same procedure for the final state particle: (i) replace the annihilation operator $a_{\mathbf{q},\text{out}}$ by its expression in terms of ϕ_{out}, (ii) write ϕ_{out} as a limit of ϕ when $x^0 \to +\infty$, (iii) write this limit as an integral of a time derivative plus a term at $x^0 \to -\infty$, that we rewrite as the annihilation operator $a_{\mathbf{q},\text{in}}$:

$$
\begin{aligned}
\langle \mathbf{q}_{\text{out}} | \mathbf{p}_{\text{in}} \rangle \;\doteq\; & i \int d^4x \, e^{-ip\cdot x} \, (\Box_x + m^2) \Big\{ \langle 0_{\text{out}} | a_{\mathbf{q},\text{in}} \phi(x) | 0_{\text{in}} \rangle \\
& + i \int d^4y \, \partial_{y^0} \, e^{iq\cdot y} \, \langle 0_{\text{out}} | \left(\Pi(y) - iE_{\mathbf{q}} \phi(y) \right) \phi(x) | 0_{\text{in}} \rangle \Big\}
\end{aligned}
\tag{1.59}
$$

However, at this point we are stuck because we would like to bring the $a_{\mathbf{q},\text{in}}$ to the right where it would annihilate $|0_{\text{in}}\rangle$, but we do not know the commutator between $a_{\mathbf{q},\text{in}}$ and the *interacting* field operator $\phi(x)$. The remedy is to go one step back, and note that we are free to insert a T-product in

$$
\left(\Pi_{\text{out}}(y) - iE_{\mathbf{q}} \phi_{\text{out}}(y) \right) \phi(x) = \lim_{y^0 \to +\infty} T \left(\left(\Pi(y) - iE_{\mathbf{q}} \phi(y) \right) \phi(x) \right)
\tag{1.60}
$$

since the time $y^0 \to +\infty$ is obviously larger than x^0. Then the boundary term at $y^0 \to -\infty$ will automatically lead to the desired ordering $\phi(x) a_{\mathbf{q},\text{in}}$,

$$
\begin{aligned}
\langle \mathbf{q}_{\text{out}} | \mathbf{p}_{\text{in}} \rangle \;\doteq\; & i \int d^4x \, e^{-ip\cdot x} \, (\Box_x + m^2) \Big\{ \langle 0_{\text{out}} | \phi(x) \underbrace{a_{\mathbf{q},\text{in}} | 0_{\text{in}} \rangle}_{0} \\
& + i \int d^4y \, \partial_{y^0} \, e^{iq\cdot y} \, \langle 0_{\text{out}} | T \left(\Pi(y) - iE_{\mathbf{q}} \phi(y) \right) \phi(x) | 0_{\text{in}} \rangle \Big\} \,.
\end{aligned}
\tag{1.61}
$$

Performing the derivative with respect to y^0, we finally arrive at

$$
\langle \mathbf{q}_{\text{out}} | \mathbf{p}_{\text{in}} \rangle \doteq i^2 \int d^4x \, d^4y \, e^{i(q\cdot y - p\cdot x)} \, (\Box_x + m^2)(\Box_y + m^2) \, \langle 0_{\text{out}} | T \, \phi(x)\phi(y) | 0_{\text{in}} \rangle \,.
\tag{1.62}
$$

Such a formula is known as a (Lehmann-Symanzik-Zimmermann) *reduction formula.*

The method that we have exposed above on a simple case can easily be applied to the most general transition amplitude, with the following result for the part of the amplitude that does not involve any spectator particle:

$$
\begin{aligned}
\langle q_1 \cdots q_n {}_{\text{out}} | p_1 \cdots p_m {}_{\text{in}} \rangle \doteq\; & i^{m+n} \int \prod_{i=1}^{m} d^4x_j \, e^{-ip_i\cdot x_i} \, (\Box_{x_i} + m^2) \\
& \times \int \prod_{j=1}^{n} d^4y_j \, e^{iq_j\cdot x_j} \, (\Box_{y_j} + m^2) \, \langle 0_{\text{out}} | T \, \phi(x_1) \cdots \phi(x_m) \phi(y_1) \cdots \phi(y_n) | 0_{\text{in}} \rangle \,.
\end{aligned}
\tag{1.63}
$$

The bottom line is that an amplitude with $m+n$ particles is related to the vacuum expectation value of a time-ordered product of $m + n$ interacting field operators (a slight but important modification to this formula will be introduced in the section 1.8, in order to account for field renormalization). Note that the vacuum states on the left and on the right of the expectation value are respectively the *out* and the *in* vacua.

10

1.5 Generating functional

1.5.1 Definition

To facilitate the bookkeeping, it is useful to introduce a generating functional that encapsulates all the expectation values, by defining

$$
\begin{aligned}
Z[j] &\equiv \sum_{n=0}^{\infty} \frac{1}{n!} \int d^4x_1 \cdots d^4x_n \; ij(x_1) \cdots ij(x_n) \, \langle 0_{out}| T \, \phi(x_1) \cdots \phi(x_n) |0_{in}\rangle \\
&= \langle 0_{out}| T \, \exp i \int d^4x \, j(x)\phi(x) |0_{in}\rangle .
\end{aligned}
\tag{1.64}
$$

Note that

$$
Z[0] = \langle 0_{out}|0_{in}\rangle \neq 1
\tag{1.65}
$$

in an interacting theory (but if the vacuum state is stable, then this vacuum to vacuum transition amplitude must be a pure phase whose squared modulus is one). From this functional, the relevant expectation values are obtained by functional differentiation

$$
\langle 0_{out}| T \, \phi(x_1) \cdots \phi(x_n) |0_{in}\rangle = \left. \frac{\delta^n Z[j]}{i\delta j(x_1) \cdots i\delta j(x_n)} \right|_{j=0} .
\tag{1.66}
$$

The knowledge of $Z[j]$ would therefore give access to all the transition amplitudes. However, it is in general not possible to derive $Z[j]$ in closed form, and we need to resort to *perturbation theory*, in which the answer is obtained as an expansion in powers of the coupling constant.

1.5.2 Relation between the free and interacting generating functionals

The generating functional can be brought to a more useful form by first writing

$$
\phi(x_1) \cdots \phi(x_n) = U(-\infty, x_1^0) \, \phi_{in}(x_1) \, U(x_1^0, x_2^0) \, \phi_{in}(x_2) \cdots \phi_{in}(x_n) \, U(x_n^0, \infty) . \tag{1.67}
$$

For convenience, we split the leftmost evolution operator as

$$
U(-\infty, x_1^0) = U(-\infty, +\infty) \, U(+\infty, x_1^0) .
\tag{1.68}
$$

Noticing that the formula (1.67) is true for any ordering of the times x_i^0 and using the expression of the U's as a time-ordered exponential, we have

$$
T \, \phi(x_1) \cdots \phi(x_n) = U(-\infty, +\infty) \, T \, \phi_{in}(x_1) \cdots \phi_{in}(x_n) \, \exp i \int d^4x \, \mathcal{L}_I(\phi_{in}(x)) , \tag{1.69}
$$

where the time-ordering in the right-hand side applies to all the operators on its right. This leads to the following representation of the generating functional

$$
\begin{aligned}
Z[j] &= \underbrace{\langle 0_{out}|U(-\infty, +\infty)}_{\langle 0_{in}|} \, T \, \exp i \int d^4x \, \Big[j(x)\phi_{in}(x) + \mathcal{L}_I(\phi_{in}(x)) \Big] |0_{in}\rangle \\
&= \exp i \int d^4x \, \mathcal{L}_I\left(\frac{\delta}{i\delta j(x)} \right) \underbrace{\langle 0_{in}| T \, \exp i \int d^4x \, j(x)\phi_{in}(x) |0_{in}\rangle}_{Z_0[j]} .
\end{aligned}
\tag{1.70}
$$

This expression of $Z[j]$ is the most useful, since it factorizes the interactions into a (functional) differential operator acting on $Z_0[j]$, the generating functional for the non-interacting theory.

1.5.3 Free generating functional

It turns out that the latter is calculable analytically. The main difficulty in evaluating $Z_0[j]$ is to deal with the non-commuting objects contained in the exponential. A central mathematical result that we shall need is a particular case of the *Baker-Campbell-Hausdorff formula*,

$$\text{if} \quad [A,[A,B]] = [B,[A,B]] = 0 \,, \qquad e^A \, e^B = e^{A+B} \, e^{\frac{1}{2}[A,B]} \,. \tag{1.71}$$

This formula is applicable to our problem because commutators $[a, a^\dagger]$ are c-numbers that commute with everything else. In order to apply it, let us slice the time axis into an infinite number of small intervals, by writing

$$T \exp \int_{-\infty}^{+\infty} d^4x \, O(x) = \prod_{i=-\infty}^{+\infty} T \exp \int_{x_i^0}^{x_{i+1}^0} d^4x \, O(x) \,, \tag{1.72}$$

where the intermediate times are ordered according to $\cdots x_i^0 < x_{i+1}^0 < \cdots$. The product in the right hand side should be understood with the convention that the factors are ordered from left to right when the index i decreases. When the size $\Delta \equiv x_{i+1}^0 - x_i^0$ of these intervals goes to zero, the time-ordering can be removed in the individual factors[12]:

$$T \exp \int_{-\infty}^{+\infty} d^4x \, O(x) = \lim_{\Delta \to 0^+} \prod_{i=-\infty}^{+\infty} \exp \int_{x_i^0}^{x_{i+1}^0} d^4x \, O(x) \,. \tag{1.73}$$

A first application of the Baker-Campbell-Hausdorff formula leads to

$$T \exp i \int d^4x \, j(x)\phi_{in}(x) = \exp\left\{ i \int d^4x \, j(x)\phi_{in}(x) \right\}$$
$$\times \exp\left\{ -\frac{1}{2} \int d^4x d^4y \, \theta(x^0 - y^0) \, j(x)j(y) \left[\phi_{in}(x), \phi_{in}(y)\right] \right\} . \tag{1.74}$$

Note that the exponential in the second line is a c-number. In the end, we will need to evaluate the expectation value of this operator in the $\left|0_{in}\right\rangle$ vacuum state. Therefore, it is desirable to transform it in such a way that the annihilation operators are on the right and the annihilation operators are on the left. This can be achieved by writing

$$\phi_{in}(x) = \phi_{in}^{(+)}(x) + \phi_{in}^{(-)}(x) \,,$$
$$\phi_{in}^{(+)}(x) \equiv \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_\mathbf{p}} \, a_{\mathbf{p},in}^\dagger \, e^{+ip\cdot x} \,,$$
$$\phi_{in}^{(-)}(x) \equiv \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_\mathbf{p}} \, a_{\mathbf{p},in} \, e^{-ip\cdot x} \,, \tag{1.75}$$

and by using once again the Baker-Campbell-Hausdorff formula. We obtain

$$T \exp i \int d^4x \, j(x)\phi_{in}(x) = \exp\left\{ i \int d^4x \, j(x)\phi_{in}^{(+)}(x) \right\} \exp\left\{ i \int d^4x \, j(x)\phi_{in}^{(-)}(x) \right\}$$
$$\times \exp\left\{ \frac{1}{2} \int d^4x d^4y \, j(x)j(y) \left[\phi_{in}^{(+)}(x), \phi_{in}^{(-)}(y)\right] \right\}$$
$$\times \exp\left\{ -\frac{1}{2} \int d^4x d^4y \, \theta(x^0 - y^0) \, j(x)j(y) \left[\phi_{in}(x), \phi_{in}(y)\right] \right\} . \tag{1.76}$$

---

[12]Field operators commute for space-like intervals,
$$\left[O(x), O(y)\right] = 0 \qquad \text{if } (x-y)^2 < 0 \,.$$
Moreover, when $\Delta \to 0$, the separation between any pair of points x, y with $x_i^0 < x^0, y^0 < x_{i+1}^0$ is always space-like.

The operator that appears in the right hand side of the first line is called a *normal-ordered exponential*, and is denoted by bracketing the exponential with colons (: ⋯ :):

$$: \exp i \int d^4x \, j(x)\phi_{in}(x) : \equiv \exp\left\{ i \int d^4x \, j(x)\phi_{in}^{(+)}(x) \right\} \exp\left\{ i \int d^4x \, j(x)\phi_{in}^{(-)}(x) \right\} .$$

(1.77)

A crucial property of the normal ordered exponential is that its *in*-vacuum expectation value is equal to unity:

$$\langle 0_{in}| : \exp i \int d^4x \, j(x)\phi_{in}(x) : |0_{in}\rangle = 1 .$$

(1.78)

Therefore, we have proven that the generating functional of the free theory is a Gaussian in $j(x)$,

$$Z_0[j] = \exp\left\{ -\frac{1}{2} \int d^4 d^4 y \, j(x)j(y) \, G_F^0(x,y) \right\} ,$$

(1.79)

where $G_F^0(x,y)$ is a 2-point function called the *free Feynman propagator* and defined as

$$G_F^0(x,y) = \theta(x^0 - y^0)\big[\phi_{in}(x), \phi_{in}(y)\big] - \big[\phi_{in}^{(+)}(x), \phi_{in}^{(-)}(y)\big] .$$

(1.80)

1.5.4 Feynman propagator

Since the commutators in the right hand side of eq. (1.80) are c-numbers, we can also write

$$\begin{aligned}
G_F^0(x,y) &= \langle 0_{in}|\theta(x^0 - y^0)\big[\phi_{in}(x), \phi_{in}(y)\big] - \big[\phi_{in}^{(+)}(x), \phi_{in}^{(-)}(y)\big]|0_{in}\rangle \\
&= \langle 0_{in}|T \, \phi_{in}(x)\phi_{in}(y)|0_{in}\rangle .
\end{aligned}$$

(1.81)

In other words, the free Feynman propagator is the *in*-vacuum expectation value of the time-ordered product of two free fields. Using the Fourier mode decomposition of ϕ_{in} and the commutation relation between creation and annihilation operators, the Feynman propagator can be rewritten as follows

$$G_F^0(x,y) = \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_{\mathbf{p}}} \left\{ \theta(x^0 - y^0) \, e^{-ip\cdot(x-y)} + \theta(y^0 - x^0) \, e^{+ip\cdot(x-y)} \right\} . \quad (1.82)$$

In the following, we will also make an extensive use of the Fourier transform of this propagator (with respect to the difference of coordinates $x^\mu - y^\mu$, since it is translation invariant):

$$\begin{aligned}
\widetilde{G}_F^0(k) &\equiv \int d^4(x-y) \, e^{ik\cdot(x-y)} \, G_F^0(x,y) \\
&= \frac{1}{2E_k}\left\{ \int_0^{+\infty} dz^0 \, e^{i(k^0 - E_k)z^0} + \int_{-\infty}^0 dz^0 \, e^{i(k^0 + E_k)z^0} \right\} .
\end{aligned}$$

(1.83)

The remaining Fourier integrals over z^0 are not defined as ordinary functions. Instead, they are distributions, that can also be viewed as the limiting value of a family of ordinary functions. In order to see this, let use write

$$\int_0^{+\infty} dz^0 \, e^{iaz^0} = \lim_{\epsilon \to 0^+} \int_0^{+\infty} dz^0 \, e^{i(a+i\epsilon)z^0} = \frac{i}{a + i0^+} .$$

(1.84)

13

Likewise

$$\int_{-\infty}^{0} dz^0 \ e^{i a z^0} = \lim_{\epsilon \to 0^+} \int_{+\infty}^{0} dz^0 \ e^{i(a - i\epsilon)z^0} = -\frac{i}{a - i0^+} \ . \tag{1.85}$$

Therefore, the Fourier space Feynman propagator reads

$$\widetilde{G}_{\text{F}}^0(k) = \frac{i}{k^2 - m^2 + i0^+} \ . \tag{1.86}$$

Note that $\widetilde{G}_{\text{F}}^0(k)$ is Lorentz invariant. Henceforth, $G_{\text{F}}^0(x, y)$ is also Lorentz invariant[13]. It is sometimes useful to have a representation of eq. (1.86) in terms of distributions. This is provided by the following identity:

$$\frac{i}{z + i0^+} = i \, P\left(\frac{1}{z}\right) + \pi \delta(z) \ , \tag{1.87}$$

where $P(1/z)$ is the *principal value* of $1/z$ (i.e. the distribution obtained by cutting out – symmetrically– an infinitesimal interval around $z = 0$). As far as integration over the variable z is concerned, this prescription amounts to shifting the pole slightly below the real axis, or equivalently to going around the pole at $z = 0$ from above (the term in $\pi \delta(z)$ can be viewed as the result of the integral on the infinitesimally small half-circle around the pole):
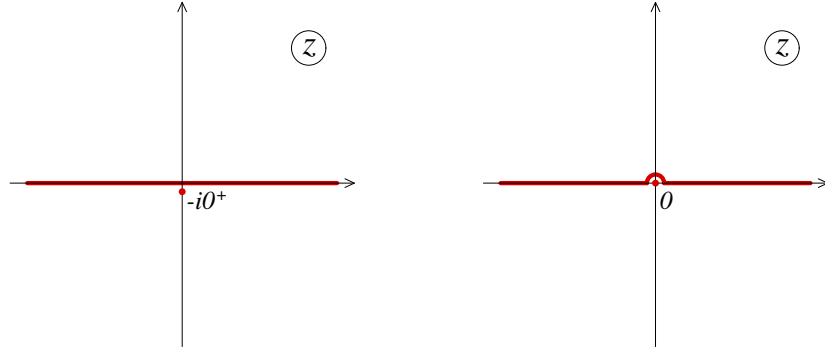
From eq. (1.86), it is trivial to check that $G_{\text{F}}^0(x, y)$ is a Green's function of the operator $\Box_x + m^2$ (up to a normalization factor $-i$):

$$(\Box_x + m^2) \, G_{\text{F}}^0(x, y) = -i\delta(x - y) \ . \tag{1.88}$$

Strictly speaking, the operator $\Box_x + m^2$ is not invertible, since it admits as zero modes all the plane waves $\exp(\pm ik \cdot x)$ with an on-shell momentum $k_0^2 = \mathbf{k}^2 + m^2$. The $i0^+$ prescription in the denominator of eq. (1.86) amounts to shifting infinitesimally the zeroes of $k_0^2 = \mathbf{k}^2 + m^2$ in the complex k_0 plane, in order to have a well defined inverse. The

---

[13]This is somewhat obfuscated by the fact that the step functions $\theta(\pm(x^0 - y^0))$ that enter in the definition of the time-ordered product are not Lorentz invariant. The Lorentz invariance of time-ordered products follows from the following properties:

- if $(x - y)^2 < 0$, then the two fields commute and the time ordering is irrelevant,
- if $(x - y)^2 \geq 0$, then the sign of $x^0 - y^0$ is Lorentz invariant.

regularization of eq. (1.86) is specific to the time-ordered propagator. Other regularizations would provide different propagators; for instance the *free retarded propagator* is given by

$$\widetilde{G}^0_{\scriptscriptstyle R}(k) = \frac{i}{(k_0 + i0^+)^2 - (\mathbf{k}^2 + m^2)} \; .$$ (1.89)

One can easily check that its inverse Fourier transform is a function $G^0_{\scriptscriptstyle R}(x, y)$ that satisfies

$$
\begin{aligned}
(\square_x + m^2)\, G^0_{\scriptscriptstyle R}(x, y) &= -i\delta(x - y) \; , \\
G^0_{\scriptscriptstyle R}(x, y) &= 0 \;\; \text{if } x^0 < y^0 \; .
\end{aligned}
$$ (1.90)

In other words, $G^0_{\scriptscriptstyle R}$ is also a Green's function of the operator $\square_x + m^2$, but with boundary conditions that differ from those of $G^0_{\scriptscriptstyle F}$

1.6 Perturbative expansion and Feynman rules

The generating functional $Z[j]$ is usually not known analytically in closed form, but is given indirectly by eq. (1.70) as the action of a functional differential operator that acts on the generating functional of the free theory. The latter is a Gaussian in j, whose variance is given by the free Feynman propagator $G^0_{\scriptscriptstyle F}$. Although not explicit, this formula provides a straightforward method for obtaining vacuum expectation values of T-products of fields to a given order in the coupling constant λ.

1.6.1 Examples

Let us first illustrate this by computing to order λ^1 the following two functions: $\langle 0_{\text{out}} | 0_{\text{in}} \rangle$ and $\langle 0_{\text{out}} | T\, \phi(x)\phi(y) | 0_{\text{in}} \rangle$. In order to make the notations a bit lighter, we denote $G^0_{xy} \equiv G^0_{\scriptscriptstyle F}(x, y)$. At order one in λ, we have

$$
\begin{aligned}
\langle 0_{\text{out}} | 0_{\text{in}} \rangle &= Z[0] = \left[1 - i\frac{\lambda}{4!} \int d^4z \left(\frac{\delta}{i\delta j(z)} \right)^4 + \mathcal{O}(\lambda^2) \right] Z_0[j]\big|_{j=0} \\
&= 1 - i\frac{\lambda}{8} \int d^4z\, G^{0\,2}_{zz} + \mathcal{O}(\lambda^2) \; ,
\end{aligned}
$$ (1.91)

and

$$
\begin{aligned}
\langle 0_{\text{out}} | T\, \phi(x)\phi(y) | 0_{\text{in}} \rangle &= \left[1 - i\frac{\lambda}{4!} \int d^4z \left(\frac{\delta}{i\delta j(z)} \right)^4 + \mathcal{O}(\lambda^2) \right] \frac{\delta^2 Z_0[j]}{i^2 \delta j(x)\delta j(y)} \bigg|_{j=0} \\
&= G^0_{xy} - i\, G^0_{xy} \frac{\lambda}{8} \int d^4z\, G^{0\,2}_{zz} - i\frac{\lambda}{2} \int d^4z\, G^0_{xz} G^0_{zz} G^0_{zy} + \mathcal{O}(\lambda^2) \\
&= \underbrace{\left[1 - i\frac{\lambda}{8} \int d^4z\, G^{0\,2}_{zz} + \mathcal{O}(\lambda^2) \right]}_{Z[0]} \\
&\quad \times \left[G^0_{xy} - i\frac{\lambda}{2} \int d^4z\, G^0_{xz} G^0_{zz} G^0_{zy} + \mathcal{O}(\lambda^2) \right] \; .
\end{aligned}
$$ (1.92)

Although the final expressions at order one are rather simple, the intermediate steps are quite cumbersome due to the necessity of taking a large number of functional derivatives. Moreover, the expression of the 2-point function $\langle 0_{\text{out}} | T \phi(x)\phi(y) | 0_{\text{in}} \rangle$ becomes simpler after we notice that one can factor out $Z[0]$. This property is in fact completely general; all transition amplitudes contain a factor $Z[0]$. From the remark made after eq. (1.65), this factor is a pure phase and its squared modulus is one and will have no effect in transition probabilities. Therefore, it would be desirable to identify from the start the terms that lead to this prefactor, to avoid unnecessary calculations.

1.6.2 Diagrammatic representation

This simplification follows a quite transparent rule if we represent the above expressions diagrammatically, by introducing the following notation

$$G^0_{xy} \equiv x \underline{\hspace{1cm}} y \ . \tag{1.93}$$

The functions considered above can be represented as follows:

$$Z[0] = 1 + \tfrac{1}{8} \ \underset{z}{\bowtie} \ + \ \mathcal{O}(\lambda^2)$$

$$\langle 0_{\text{out}} | T \phi(x)\phi(y) | 0_{\text{in}} \rangle = x \underline{\hspace{0.5cm}} y + \tfrac{1}{8} x \underline{\hspace{0.5cm}} y \underset{z}{\bowtie} + \tfrac{1}{2} x \underline{\hspace{0.3cm}}_z \underline{\hspace{0.3cm}} y + \mathcal{O}(\lambda^2)$$
$$\tag{1.94}$$

The graphs that appear in the right hand side of these equations are called *Feynman diagrams*. By adding to eq. (1.93) the rule that each vertex should have a factor $-i\lambda$ and an integration over the entire space-time, then these graphs are in one-to-one correspondence with the expressions of eqs. (1.91) and (1.92). For now, we have recalled explicitly the numerical prefactors (1/8, 1/2,...) but they can in fact be recovered simply from the symmetries of the graphs.

In the second of eqs. (1.94), the second term of the right hand side contains a factor which is not *connected* to any of the points x and y. These disconnected graphs are precisely the ones responsible for the factor $Z[0]$ that appears in all transition amplitudes. We can therefore disregard these type of graphs altogether.

1.6.3 Feynman rules

The diagrammatic representation of eqs. (1.94) can in fact be used to completely bypass the explicit calculation of the functional derivatives of $Z_0[j]$. The rules that govern this construction are called *Feynman rules*. The contributions of order λ^p to a n-point function $\langle 0_{\text{out}} | T\phi(x_1) \cdots \phi(x_n) | 0_{\text{in}} \rangle$ can be obtained as follows:

1. Draw all the graphs (with only vertices of valence 4) that connect the n points x_1 to x_n and have exactly p vertices. Graphs that contain a subgraph which is not connected to any of the x_i's should be ignored.

2. Each line of a graph represents a free Feynman propagator G^0_F.

3. Each vertex represents a factor $-i\lambda$ and an integral over the space-time coordinate assigned to this vertex.

4. The numerical prefactor for a given graph is the inverse of the order of its discrete symmetry group. As an illustration, we indicate below the generators of these symmetry groups and their order for the graphs that appear in eqs. (1.94):

$$
\longrightarrow \quad \text{order } 8 \quad \longrightarrow \quad \frac{1}{8} \,,
$$

$$
\longrightarrow \quad \text{order } 2 \quad \longrightarrow \quad \frac{1}{2} \,. \tag{1.95}
$$

Note that this rule for obtaining the *symmetry factor* associated to a given graph is correct only if the corresponding term in the Lagrangian has been properly symmetrized. For instance, the operator ϕ^4 should appear in the Lagrangian with a prefactor $1/4!$.

1.6.4 Connected graphs

At the step **1**, graphs made of several disconnected subgraphs can usually appear in certain functions, provided that each subgraph is connected to at least one of the points x_i. For instance, a 4-point function contains a piece which is simply made of the product of two 2-point functions. In addition, it contains terms that correspond to a genuine 4-point function, not factorizable in a product of 2-point functions. The factorizable pieces are usually less interesting because they can be recovered from already calculated simpler building blocks. For this reason, it is sometimes useful to introduce the *generating function of the connected graphs*, denoted $W[j]$. This functional is very simply related to $Z[j]$ by

$$
W[j] = \log Z[j] \,. \tag{1.96}
$$

To give a glimpse of this identity, let us write

$$
W[j] = \sum_{n=1}^{\infty} \frac{1}{n!} \int d^4x_1 \cdots d^4x_n \, C_n(x_1, \cdots, x_n) \, j(x_1) \cdots j(x_n) \,, \tag{1.97}
$$

where the $C_n(x_1, \cdots, x_n)$ are n-point functions whose diagrammatic representation contain only connected graphs. If we expand $Z[j] = \exp W[j]$, we obtain

$$
\begin{aligned}
Z[j] \;=\; & 1 + \int d^4x \, C_1(x) \, j(x) + \frac{1}{2!} \int d^4x d^4y \, \underbrace{\left[C_2(x,y) + C_1(x)C_1(y) \right]}_{\left\langle 0_{\text{out}} \middle| T \, \phi(x)\phi(y) \middle| 0_{\text{in}} \right\rangle} j(x)j(y) \\
& + \frac{1}{3!} \int d^4x d^4y d^4z \, \Big[C_3(x,y,z) + C_2(x,y)C_1(z) \\
& \qquad\qquad\qquad + C_2(y,z)C_1(x) + C_2(z,x)C_1(y) \\
& \qquad\qquad\qquad\qquad \underbrace{+ C_1(x)C_1(y)C_1(z)}_{\left\langle 0_{\text{out}} \middle| T \, \phi(x)\phi(y)\phi(z) \middle| 0_{\text{in}} \right\rangle} \Big] j(x)j(y)j(z) + \cdots
\end{aligned}
$$

$$
\tag{1.98}
$$

This expansion highlights how the vacuum expectation values of time-ordered products of fields can be factorize into products of connected contributions.

1.6.5 Feynman rules in momentum space

Until now, we have obtained Feynman rules in terms of objects that depend on space-time coordinates, leading to expressions for the perturbative expansion of the vacuum expectation value of time-ordered products of fields. However, in most practical applications, we need subsequently to use the LSZ reduction formula (1.63) to turn these expectation values into transition amplitudes. This involves the application of the operator $i(\Box+m^2)$ to each external point, and a Fourier transform. Firstly, note that thanks to eq. (1.88), the application of $i(\Box + m^2)$ simply removes the external line to which it is applied:

$$(\Box_x + m^2) \left[x \!\!-\!\!\overset{z}{\longrightarrow}\!\! \bigcirc \right] \quad = \quad x\!\!-\!\!\bigcirc \quad . \tag{1.99}$$

Thus, these operators just produce Feynman graphs that are *amputated* of all their external lines. Then, the Fourier transform can be propagated to all the internal lines of the graph, leading to an expression that involves propagators and vertices that depend only on momenta. The Feynman rules for obtaining directly these momentum space expressions are:

1′. The graph topologies that must be considered is of course unchanged. The momenta of the initial state particles are entering into the graph, and the momenta of the final state particles are going out of the graph

2′. Each line of a graph represents a free Feynman propagator in momentum space $\widetilde{G}^0_F(k)$

3′. Each vertex represents a factor $-i\lambda(2\pi)^2\delta(k_1 + \cdots + k_4)$, where the k_i are the four momenta entering into this vertex

3″. All the internal momenta that are not constrained by these delta functions should be integrated over with a measure $d^4k/(2\pi)^4$

4′. Symmetry factors are computed as before.

For instance, these rules lead to:

$$\overset{P}{\underset{\longrightarrow}{\bigcirc}} \quad = \quad -i\frac{\lambda}{2}\int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2-m^2+i0^+}$$

$$\underset{p_1 \quad\quad q_1}{\overset{p_2 \quad k \quad q_2}{\bowtie}} \quad = \quad = \frac{(-i\lambda)^2}{2}\int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2-m^2+i0^+}\frac{i}{(p_1+p_2-k)^2-m^2+i0^+} \quad . \tag{1.100}$$

1.6.6 Counting the powers of λ and $\hbar$

The order in λ of a (connected) graph $\mathcal{G}$ is of course related to the number of vertices n_V in the graph,

$$\mathcal{G} \sim \lambda^{n_V} \quad . \tag{1.101}$$

This can also be related to the number of loops of the graph, which is a better measure of its complexity since it determines how many momentum integrals it contains. Let us denote n_E

the number of external lines, n_I the number of external lines and n_L the number of loops. These parameters are related by the following two identities:

$$\begin{aligned} 4n_V &= 2n_I + n_E \\ n_L &= n_I - n_V + 1 . \end{aligned} \tag{1.102}$$

The first of these equations equates the number of "handles" carried by the vertices, and the number of propagator endpoints that must attached to them. The right hand side of the second equation counts the number of internal momenta that are not constrained by the delta functions of momentum conservation carried the vertices (the $+1$ comes from the fact that not all these delta functions are independent - a linear combination of them must simply tell that the sum of the external momenta must be zero, and therefore does not constrain the internal ones in any way). From these two identities, one obtains

$$n_V = n_L - 1 + \frac{n_E}{2} , \tag{1.103}$$

and the order in λ of the graph is also

$$\mathcal{G} \sim \lambda^{n_L - 1 + n_E/2} . \tag{1.104}$$

According to this formula, the order of a graph depends only on the number of external lines n_E (i.e. on the number of particles involved in the transition amplitude under consideration), and on the number of loops. Thus, the perturbative expansion is also a *loop expansion*, with the leading order being given by *tree diagrams*, the first correction in λ by *one-loop graphs*, etc...

It turns out that the number of loops also counts the order in the Planck constant $\hbar$ of a graph. Although we have been using a system of units in which $\hbar = 1$, it is easy to reinstate $\hbar$ by the substitution

$$S \quad \rightarrow \quad \frac{S}{\hbar} = -\int d^4x \left\{ \frac{1}{2}\phi(x) \frac{\Box_x + m^2}{\hbar} \phi(x) + \frac{\lambda}{4!\hbar} \phi^4(x) \right\} . \tag{1.105}$$

From this, we see that $\hbar$ enters in the Feynman rules as follows

$$\begin{aligned} \text{Propagator :} \quad & \frac{i\hbar}{p^2 - m^2 + i0^+} , \\ \text{Vertex :} \quad & -i\frac{\lambda}{\hbar} , \end{aligned} \tag{1.106}$$

and the order in $\hbar$ of a graph is given by

$$\mathcal{G} \sim \hbar^{n_I - n_V} \sim \hbar^{n_L - 1} . \tag{1.107}$$

Therefore, each additional loop brings a power of $\hbar$, and the loop expansion can also be viewed as an expansion in powers of $\hbar$.

1.7 Calculation of loop integrals

1.7.1 Wick's rotation

Let us consider the first of the examples given in eq. (1.100) and define

$$-i\Sigma(P) \equiv -i\frac{\lambda}{2} \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i0^+} . \tag{1.108}$$

In order to calculate the momentum integral, it is useful to perform a *Wick rotation*, in which we rotate the k_0 integration axis by 90 degrees to bring it along the imaginary axis, as illustrated in the following figure:
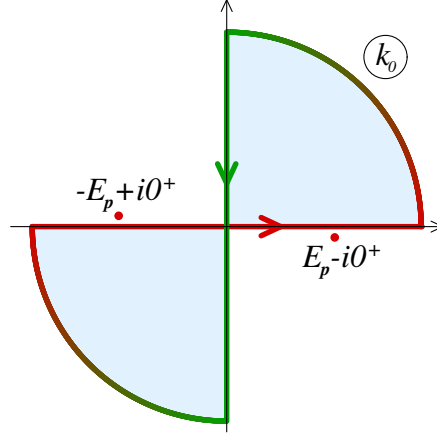
The integrals along the horizontal and vertical axis are opposite because the shaded domain does not contain any of the poles of the Feynman propagator, and because the propagator vanishes as k_0^{-2} when $|k_0| \to \infty$. The integral along the vertical axis amounts to writing $k_0 = -i\kappa$ with κ varying from $-\infty$ to $+\infty$. After this transformation, the integral of eq. (1.108) becomes

$$\Sigma(P) = \frac{\lambda}{2} \int \frac{d^4 k_E}{(2\pi)^4} \frac{1}{k_E^2 + m^2} , \qquad (1.109)$$

where k_E is the *Euclidean* 4-vector defined by $k_E^i = \mathbf{k}$ $(i = 1, 2, 3)$ and $k_E^4 = \kappa$, with squared norm $k_E^2 = \mathbf{k}^2 + \kappa^2$.

1.7.2 Volume element in D dimensions

When the integrand depends only on the norm $|k_E|$, we can separate the radial integration on $|k_E|$ from the angular integration over the orientation of the vector in 4-dimensional Euclidean space. In D dimensions, the volume measure for a rotationally invariant integrand reads

$$d^D k_E = D \, \mathcal{V}_D(1) \, k_E^{D-1} \, dk_E , \qquad (1.110)$$

where $\mathcal{V}_D(k_E)$ is the volume of the D-dimensional ball of radius k_E. These volumes can be determined recursively by

$$\mathcal{V}_1(k_E) = 2k_E , \qquad \mathcal{V}_D(k_E) = k_E \int_0^\pi d\theta \, \sin\theta \, \mathcal{V}_{D-1}(k_E \sin\theta) . \qquad (1.111)$$

Therefore, we have

$$\mathcal{V}_2(k_E) = \pi k_E^2 , \quad \mathcal{V}_3(k_E) = \frac{4\pi}{3} k_E^3 , \quad \mathcal{V}_4(k_E) = \frac{\pi^2}{4} k_E^4 . \qquad (1.112)$$

20

Although knowing $\mathcal{V}_4(k_{\text{E}})$ is sufficient for performing a radial momentum integral in four dimensions, it is interesting to have the formula for an arbitrary dimension, in view of applications to *dimensional regularization*. More generally, we have

$$\mathcal{V}_{\text{D}+1}(1) = \mathcal{V}_{\text{D}}(1)\,\pi^{1/2}\,\frac{\Gamma(\frac{\text{D}}{2}+1)}{\Gamma(\frac{\text{D}}{2}+\frac{3}{2})} \quad \text{and} \quad \mathcal{V}_{\text{D}}(1) = \frac{2\,\pi^{\text{D}/2}}{\text{D}\,\Gamma(\frac{\text{D}}{2})}\,. \tag{1.113}$$

1.7.3 Feynman parameterization of denominators

Let us now consider the second diagram of eq. (1.100) (with the notation $\text{P} \equiv p_1 + p_2$),

$$-i\Gamma_4(\text{P}) \equiv \frac{(-i\lambda)^2}{2}\int\frac{\mathrm{d}^4k}{(2\pi)^4}\,\frac{i}{k^2-m^2+i0^+}\,\frac{i}{(\text{P}-k)^2-m^2+i0^+}\,. \tag{1.114}$$

In this more complicated example, an extra difficulty is that the integrand is not rotationally invariant. The following trick, known as *Feynman parameterization* can be used to rearrange the denominators[14]:

$$\frac{1}{AB} = \int_0^1\frac{\mathrm{d}x}{[xA+(1-x)B]^2}\,. \tag{1.115}$$

The denominator resulting from this transformation is

$$x(k^2-m^2+i0^+)+(1-x)((\text{P}-k)^2-m^2+i0^+) = l^2-m^2-\Delta(x,\text{P})+i0^+\,, \tag{1.116}$$

where we denote $l \equiv k-(1-x)\text{P}$ and $\Delta(x,\text{P}) \equiv -x(1-x)\text{P}^2$. At this point, we can apply a Wick rotation[15] to the shifted integration variable l, in order to obtain

$$\Gamma_4(\text{P}) = -\frac{\lambda^2}{2}\int_0^1\mathrm{d}x\int\frac{\mathrm{d}^4l_{\text{E}}}{(2\pi)^4}\,\frac{1}{[l_{\text{E}}^2+m^2+\Delta(x,\text{P})]^2}\,, \tag{1.117}$$

where the integrand is again invariant by rotation in 4-dimensional Euclidean space.

1.8 Källen-Lehmann spectral representation

As we shall see now, the limit in eq. (1.41) that relates the interacting field ϕ and the free field of the interaction picture ϕ_{in} is too naive. One of the consequences is that we will have to make a slight modification to the reduction formula (1.63).

Consider the time-ordered 2-point function,

$$\langle 0_{\text{out}}|T\,\phi(x)\phi(y)|0_{\text{in}}\rangle = \theta(x^0-y^0)\,\langle 0_{\text{out}}|\phi(x)\phi(y)|0_{\text{in}}\rangle+\theta(y^0-x^0)\,\langle 0_{\text{out}}|\phi(y)\phi(x)|0_{\text{in}}\rangle\,. \tag{1.118}$$

---

[14]For n denominators, this formula can be generalized into

$$\frac{1}{A_1A_2\cdots A_n} = \Gamma(n)\int_0^1\mathrm{d}x_1\cdots\mathrm{d}x_n\,\delta(1-\sum_i x_i)\,\frac{1}{[x_1A_1+\cdots+x_nA_n]^n}\,.$$

[15]It is allowed because the integration axis can be rotated counterclockwise without passing through the poles in the variable l_0.

For each of the expectation values in the right hand side, let us insert an identity operator between the two field operators, written in the form of a sum over all the possible physical states,

$$1 = \sum_{\text{states } \lambda} |\lambda\rangle\langle\lambda| .$$
(1.119)

The states λ can be arranged into classes inside which the states differ only by a boost. A class of states, that we will denote α, is characterized by its particle content and by the relative momenta of these particles. Within a class, the total momentum of the state can be varied by applying a Lorentz boost. For a class α, we will denote $|\alpha_{\mathbf{p}}\rangle$ the state of total momentum $\mathbf{p}$. Each class of states has an *invariant mass* m_α, such that the total energy p^0 and total momentum $\mathbf{p}$ of the states in this class obey $p_0^2 - \mathbf{p}^2 = m_\alpha^2$. In addition, it is useful to isolate the vacuum in the sum over the states. Therefore, the identity operator can be rewritten as

$$1 = |0\rangle\langle0| + \sum_{\text{classes } \alpha} \int \frac{d^3\mathbf{p}}{(2\pi)^3 2\sqrt{\mathbf{p}^2 + m_\alpha^2}} |\alpha_{\mathbf{p}}\rangle\langle\alpha_{\mathbf{p}}| ,$$
(1.120)

where we have written the integral over the total momentum of the states in a Lorentz invariant fashion. (We need not specify if we are using *in* or *out* states here.)

When we insert this identity operator between the two field operators, the vacuum does not contribute. For instance

$$\langle 0_{\text{out}}|\phi(x)|0\rangle = 0 .$$
(1.121)

(ϕ creates or destroys a particle, and therefore has a vanishing matrix element between vacuum states.) Using the momentum operator $\hat{\mathbf{P}}$, we can write

$$
\begin{aligned}
\langle 0_{\text{out}}|\phi(x)|\alpha_{\mathbf{p}}\rangle &= \langle 0_{\text{out}}|e^{i\hat{\mathbf{P}}\cdot x}\phi(0)e^{-i\hat{\mathbf{P}}\cdot x}|\alpha_{\mathbf{p}}\rangle \\
&= \langle 0_{\text{out}}|\phi(0)|\alpha_{\mathbf{p}}\rangle e^{-ip\cdot x} \\
&= \langle 0_{\text{out}}|\phi(0)|\alpha_0\rangle e^{-ip\cdot x} .
\end{aligned}
$$
(1.122)

The second line uses the fact that the total momentum in the vacuum state is zero, and is $\mathbf{p}$ for the state $\alpha_{\mathbf{p}}$. In the last equality, we have applied a boost that cancels the total momentum $\mathbf{p}$, and used the fact that the vacuum is invariant, as well as the scalar field $\phi(0)$. Therefore, we obtain the following representation for the time-ordered 2-point function

$$\langle 0_{\text{out}}|T\,\phi(x)\phi(y)|0_{\text{in}}\rangle = \sum_{\text{classes } \alpha} \langle 0_{\text{out}}|\phi(0)|\alpha_0\rangle\langle\alpha_0|\phi(0)|0_{\text{in}}\rangle$$

$$\times \underbrace{\int \frac{d^3\mathbf{p}}{(2\pi)^3 2\sqrt{\mathbf{p}^2 + m_\alpha^2}}\left\{\theta(x^0 - y^0)e^{-ip\cdot(x-y)} + \theta(y^0 - x^0)e^{ip\cdot(x-y)}\right\}}_{G_F^0(x,y;m_\alpha^2)} ,$$
(1.123)

where the underlined integral, $G_F^0(x, y; m_\alpha^2)$, is the Feynman propagator for a hypothetical scalar field of mass m_α (compare this integral with eq. (1.82)). It is customary to rewrite the above representation as

$$\langle 0_{\text{out}}|T\,\phi(x)\phi(y)|0_{\text{in}}\rangle = \int_0^\infty \frac{dM^2}{2\pi}\,\rho(M^2)\,G_F^0(x,y;M^2) ,$$
(1.124)

where $\rho(m^2)$ is the *spectral function* defined as

$$\rho(M^2) \equiv 2\pi \sum_{\text{classes } \alpha} \delta(M^2 - m_\alpha^2) \langle 0_{\text{out}} | \phi(0) | \alpha_0 \rangle \langle \alpha_0 | \phi(0) | 0_{\text{in}} \rangle . \tag{1.125}$$

This function describes the invariant mass distribution of the non-empty states of the theory under consideration, and the exact Feynman propagator is a sum of free Feynman propagators with varying masses, weighted by this mass distribution.

In a theory of massive particles, the spectral function has a delta function corresponding to states containing a single particle of mass m, and a continuum distribution[16] that starts at the minimal invariant mass $(2m)$ of a 2-particle state:

$$\rho(M^2) = 2\pi Z \, \delta(M^2 - m^2) + \text{ continuum for } M^2 \geq 4m^2 , \tag{1.126}$$

where Z is the product of matrix elements that appear in eq. (1.125), in the case of 1-particle states. In a theory with interactions, Z in general differs from unity (in fact, it may be infinite). Note that in this equation, m must be the *physical* mass of the particles, as it would be inferred from the simultaneous measurement of their energy and momentum. As we shall see shortly, this is not the same as the parameter we denoted m in the Lagrangian.

Taking the Fourier transform of eq. (1.124) and using eq. (1.126) for the spectral function, we obtain the following pole structure for the exact Feynman propagator:

$$\widetilde{G}_{\text{\tiny F}}(p) = \frac{i Z}{p^2 - m^2 + i0^+} + \text{ terms without poles} . \tag{1.127}$$

Therefore, the parameter Z that appears in the spectral function has also the interpretation of the residue of the single particle pole in the exact Feynman propagator.

The fact that $Z \neq 1$ calls for a slight modification of the LSZ reduction formulas. Eq. (1.126) implies that a factor $\sqrt{Z}$ appears in the overlap between the state $\phi(x) | 0_{\text{in}} \rangle$ and the 1-particle state $| p_{\text{in}} \rangle$. In other words, $\phi(x)$ creates a particle with probability Z rather than 1. Therefore, there should be a factor $Z^{-1/2}$ for each incoming and outgoing particle in the LSZ reduction formulas that relate transition amplitudes to products of fields ϕ:

$$\langle q_1 \cdots q_{n \text{ out}} | p_1 \cdots p_{m \text{ in}} \rangle \doteq \left(\frac{i}{Z^{1/2}} \right)^{m+n} \int \prod_{i=1}^{m} d^4 x_j \, e^{-i p_i \cdot x_i} (\Box_{x_i} + m^2)$$

$$\times \int \prod_{j=1}^{n} d^4 y_j \, e^{i q_j \cdot x_j} (\Box_{y_j} + m^2) \langle 0_{\text{out}} | T \, \phi(x_1) \cdots \phi(x_m) \phi(y_1) \cdots \phi(y_n) | 0_{\text{in}} \rangle . \tag{1.128}$$

In practical calculations, the factor Z at a given order of perturbation theory is obtained by studying the 1-particle pole of the dressed propagator, as the residue of this pole. It is common to introduce a *renormalized field* ϕ_r defined as a rescaling of ϕ,

$$\phi \equiv \sqrt{Z} \, \phi_r . \tag{1.129}$$

By construction, the Feynman propagator defined from the 2-point time-ordered product of ϕ_r has a single-particle pole of residue 1. In other words, we may replace in the right hand side of the LSZ reduction formula (1.128) all the fields by renormalized fields, and at the same time remove all the factors $Z^{-1/2}$.

---

[16]Between the 1-particle delta function and the 2-particle continuum, there may be additional delta functions corresponding to multi-particle bound states (to have a stable bound state, the binding energy should decrease the mass of the state compared to the mass $2m$ of two free particles at rest).

1.9 Ultraviolet divergences and renormalization

Until now, we have not attempted to calculate explicitly the integrals over the Euclidean momentum k_E in eqs. (1.109) and (1.117). In fact, these integrals do not converge when $|k_E| \to \infty$, and as such they are therefore infinite. These infinities are called *ultraviolet divergences*.

1.9.1 Regularization of divergent integrals

As we shall see shortly, this has very deep implications on how we should interpret the theory. However, before we can discuss this, it is crucial to make the integrals temporarily finite in order to secure the subsequent manipulations. This procedure, called *regularization*, amounts to altering the theory to make all the integrals finite. There is no unique method for achieving this, and the most common ones are the following:

- *Pauli-Villars method* : modify the Feynman propagator according to

$$\frac{i}{k^2 - m^2 + i0^+} \quad \to \quad \frac{i}{k^2 - m^2 + i0^+} - \frac{i}{k^2 - M^2 + i0^+} \,. \tag{1.130}$$

  When $|k_E| \gg M$, this modified propagator decreases as $|k_E|^{-4}$ instead of $|k_E|^{-2}$ for the unmodified propagator, which is usually sufficient to render the integrals convergent. The original theory (and its ultraviolet divergences) are recovered in the limit $M \to \infty$.

- *Lattice regularization* : replace continuous space-time by a regular lattice of points, for instance a cubic lattice with a spacing a between the nearest neighbor sites. On such a lattice, the momenta are themselves discrete, with a maximal momentum of order a^{-1}. Therefore, the momentum integrals are replaced by discrete sums that are all finite. The original theory is recovered in the limit $a \to 0$. A shortcoming of lattice regularization is that the discrete momentum sums are usually much more difficult to evaluate than continuum integrals, and that it breaks the usual space-time symmetries such as translation and rotation invariance. This is nevertheless the basis of numerical Monte-Carlo methods (*lattice field theory*).

- *Cutoff regularization* : cut the integration over the norm of the Euclidean momentum by $|k_E| \geq \Lambda$. The underlying theory is recovered in the limit $\Lambda \to \infty$. This is a commonly used regularization in scalar theories, due to its simplicity and because it preserves all the symmetries of the theory.

- *Dimensional regularization* : this method is based on the observation that the integral

$$\int_0^\infty dk_E \; \frac{k_E^{D-1}}{[k_E^2 + \Delta]^n} \;\; = \;\; \frac{1}{2} \int_0^\infty du \; \frac{u^{\frac{D}{2}-1}}{[u+\Delta]^n}$$

$$= \;\; \frac{1}{2} \Delta^{\frac{D}{2}-n} \underbrace{\int_0^1 dx \, x^{n-\frac{D}{2}-1} (1-x)^{\frac{D}{2}-1}}_{\frac{\Gamma\left(n-\frac{D}{2}\right)\Gamma\left(\frac{D}{2}\right)}{\Gamma(n)}} \,. \tag{1.131}$$

is well defined for almost any D except for $D = 2n, 2n + 2, 2n + 4, \cdots$ and $D = 0, -2, -4, \cdots$ thanks to the analytical properties of the Gamma function[17]. Dimensional regularization keeps the number of space-time dimensions D arbitrary in all the intermediate calculations, and at the end one usually writes $D = 4 - 2\epsilon$ with $\epsilon \ll 1$. This regularization does not break any of the symmetries of the theory, including gauge invariance (which is not the case of cutoff regularization). There is an extra complication: the coupling constant λ is a priori dimensionless only when $D = 4$. In order to keep the dimension of λ unchanged, we must introduce a parameter μ that has the dimension of a mass, and replace λ by $\lambda\mu^{4-D}$. Note that the field $\phi(x)$ has the dimension of a mass to the power $(D - 2)/2$. Setting $D = 4 - 2\epsilon$, the singular part of the integrals $\Sigma(P)$ and $\Gamma_4(P)$ introduced above as examples is

$$\Sigma(P) = -\frac{\lambda}{2}\frac{m^2}{(4\pi)^2}\frac{1}{\epsilon} + \mathcal{O}(1) \quad , \qquad \Gamma_4(P) = -\frac{\lambda^2}{2}\frac{1}{(4\pi)^2}\frac{1}{\epsilon} + \mathcal{O}(1) \,. \quad (1.132)$$

1.9.2 Mass renormalization

Let us now make a few observations:

- The above divergent terms are momentum independent[18],

- They appear in 2-point and 4-point functions only.

Moreover, it is important to realize that the parameters (m^2 and λ) in the Lagrangian are not directly observable quantities by themselves[19]. For instance, the mass of a particle is a measurable property of the particle (e.g. by measuring both its energy and its momentum, via $p_0^2 - \mathbf{p}^2$). In quantum field theory, this definition of the mass corresponds to the location of the poles of the propagator in the complex p_0 plane. However, as we shall see, loop corrections modify substantially the propagator, and it turns out that the parameter m in the free propagator has in fact little to do with this physical mass. If we *dress* the propagator by summing the multiple insertions of the 1-loop correction $-i\Sigma$,

$$\widetilde{G}_F(P) \equiv \quad \xrightarrow{P} \quad + \quad \xrightarrow{P} \quad + \quad \xrightarrow{P} \quad + \quad \xrightarrow{P} \quad + \cdots , \tag{1.133}$$

we obtain

$$\widetilde{G}_F(P) = \frac{i}{p_0^2 - \mathbf{p}^2 - m^2 - \Sigma + i0^+} \,, \tag{1.134}$$

from which it is immediate to see that this loop correction alters the location of the pole, now given by

$$p_0^2 - \mathbf{p}^2 = \underbrace{m^2 + \Sigma}_{\text{new squared mass}} \,. \tag{1.135}$$

---

[17] $\Gamma(z)$ is analytic in the complex plane, at the exception of a discrete series of simple poles, located at $z_n = -n$ for $n \in \mathbb{N}$, with residues $(-1)^n/n!$.

[18] These examples are not completely general. As we shall see later, divergent terms proportional to P^2 may also appear in the 2-point function.

[19] In this regard, it is important to realize that the renormalization of the parameters of the Lagrangian would be necessary even in a theory that has no divergent loop integrals.

Since the propagator given in eq. (1.134) includes loop corrections, its poles ought to give a value of the mass closer to the physical one. Therefore, it is tempting to write:

$$m_{phys}^2 = m^2 + \Sigma + \mathcal{O}(\lambda^2) \,. \tag{1.136}$$

Of course, since Σ is infinite, the only way this can be satisfied is that the parameter m^2 that appears in the Lagrangian be itself infinite, with an opposite sign in order to cancel the infinity from Σ. To further distinguish it from the physical mass, the parameter m in the Lagrangian is usually called the *bare mass*, while m_{phys} is the physical –or *renormalized*– mass.

1.9.3   Field renormalization

Note that the 1-loop function Σ in a theory with a ϕ^4 interaction is somewhat special, because at this order it is independent of the momentum P. Being a constant, the infinity it contains can be absorbed entirely into a redefinition of the bare mass, but the residue of the pole remains equal to 1. However, starting at two loops, the 2-point functions that correct the propagator are usually momentum dependent, as is the case for instance with this graph:

It is convenient to expand $\Sigma(P^2)$ around the physical mass:

$$\Sigma(P^2) = \Sigma(m_{phys}^2) + (P^2 - m_{phys}^2)\,\Sigma'(m_{phys}^2) + \tfrac{1}{2}(P^2 - m_{phys}^2)\,\Sigma''(m_{phys}^2) + \cdots \tag{1.137}$$

For the resummed propagator $\widetilde{G}_F$ to have a pole at $P^2 = m_{phys}^2$, we need to impose

$$m_{phys}^2 = m^2 + \Sigma(m_{phys}^2) \,, \tag{1.138}$$

that generalizes eq. (1.136) to a momentum-dependent Σ. Then, in the vicinity of the pole, the dressed propagator behaves as

$$\widetilde{G}_F(P) \underset{P^2 \to m_{phys}^2}{\approx} \frac{i}{(1 - \Sigma'(m_{phys}^2))\,(P^2 - m_{phys}^2) + i0^+} \,. \tag{1.139}$$

This indicates that the field renormalization factor Z cannot be equal to 1 when the propagator is corrected by a momentum-dependent loop. Instead, we have

$$Z = \frac{1}{1 - \Sigma'(m_{phys}^2)} \,. \tag{1.140}$$

Moreover, *Weinbergs's theorem* states that the ultraviolet divergences of the 2-point function $\Sigma(P^2)$ arise only in $\Sigma(m_{phys}^2)$ and in the first derivative $\Sigma'(m_{phys}^2)$, while higher derivatives are all finite. Eqs. (1.138) and (1.140) therefore indicate that these infinities can be "hidden" in the bare mass m^2 and in the field renormalization factor Z.

1.9.4   Ultraviolet power counting

From the above considerations, it appears crucial that Σ has divergences only in its 0th and 1st order Taylor coefficients and Γ_4 only in the 0th order, in order to be able to absorb the

divergences by a proper definition of m^2, Z and λ. A simple dimensional argument gives plausibility to this assertion (of which *Weinberg's theorem* provides a more rigorous justification). Let us assume that we scale up all the internal momenta of a graph by some factor ξ. In doing this, a graph $\mathcal{G}$ with n_V vertices and n_I internal lines will scale as

$$\mathcal{G} \sim \xi^{D\,n_L - 2n_I} \, , \tag{1.141}$$

assuming D space-time dimensions for more generality. The exponent $\omega(\mathcal{G}) \equiv D\,n_L - 2n_I$ is called the *superficial degree of divergence* of the graph. This exponent characterizes how the graph diverges when all its internal momenta are rescaled uniformly:

- $\omega \geq 0$: The graph has an intrinsic divergence.

- $\omega < 0$: The graph may be finite, or may contain a divergent subgraph. However, in the renormalization process, subgraphs will have been dealt with earlier since they occur at a lower order of the perturbative expansion.

The superficial degree of divergence signals all the n-point functions that may have ultraviolet divergences of their own. Using eqs. (1.102), $\omega(\mathcal{G})$ can be rewritten in the following way

$$\omega(\mathcal{G}) = 4 - n_E + (D-4)\,n_L \, . \tag{1.142}$$

An important consequence of this formula is that in 4 dimensions the superficial degree of divergence of a graph does not depend on the number of loops, but only on the number of external lines. When $D = 4$, the only functions that have a non-negative ω are the 2-point function and the 4-point function[20]. It is important to realize that this does not mean that a 6-point cannot be divergent. However, it can diverge only if it contains a divergent 2-point or 4-point subgraph. Moreover, the value of the superficial degree of divergence indicates the maximal power of the ultraviolet cutoff that may appear in these functions:

- 2-point: up to Λ^2

- 4-point: up to $\log(\Lambda)$

Note also that if we differentiate a graph with respect to the invariant norm P^2 of one of its external momenta, we get

$$\omega\left(\frac{\partial \mathcal{G}}{\partial P^2}\right) = 2 - n_E + (D-4)\,n_L \, . \tag{1.143}$$

(ω further decreases by two units with each additional derivative with respect to P^2.) Therefore, the momentum derivative $\Sigma'(P^2)$ of the 2-point function has $\omega = 0$ in $D = 4$, and its higher derivatives all have $\omega < 0$. The fact that only $\Gamma_4(m_{\text{phys}}^2)$, $\Sigma(m_{\text{phys}}^2)$ and $\Sigma'(m_{\text{phys}}^2)$ have $\omega \geq 0$ is the very reason why it is possible to get rid of all the divergences of this theory (in 4 dimensions) by a redefinition of the parameters of the Lagrangian. This theory is said to be *renormalizable*.

---

[20]Functions with an odd number of external lines vanish in the theory under consideration. Note also that 0-point functions (vacuum graphs) have a superficial degree of divergence equal to 4, indicating that they may contain up to quartic divergences $\sim \Lambda^4$.

1.9.5 Ultraviolet classification of quantum field theories

In dimensions lower than 4, $\omega(\mathcal{G})$ is a strictly decreasing function of the number of loops, which indicates that graphs with a given n_E do not develop new divergences beyond a certain loop order. Such theories are said *super renormalizable* because they only have a finite number of divergent graphs. Conversely, in dimensions higher than 4, $\omega(\mathcal{G})$ increases with the number of loops, and any function will eventually become divergent at some loop order. These theories are usually[21] *non renormalizable*. One may think of introducing, as they become necessary, additional operators in the Lagrangian with a coupling constant adjusted to cancel the new divergences that arise at a given loop order. However, an infinite number of such parameters would need to be introduced, thereby reducing to nil the predictive power of this type of theory[22].

As we have seen, the renormalizability of a field theory depends both on the interaction terms it contains, and on the dimensionality of space-time. In fact, a simpler equivalent criterion is the mass dimension of the coupling constant in front of the interaction term:

- dim > 0 : super-renormalizable,

- dim $= 0$: renormalizable,

- dim < 0 : non-renormalizable.

For instance, the "coupling constant" m^2 in front of the mass term has always a mass dimension equal to two, and this term is therefore super-renormalizable. In contrast, the coupling constant λ in front of a ϕ^4 interaction has a mass dimension $4 - D$, and is (super)renormalizable in dimensions less than or equal to four.

1.9.6 Renormalization in perturbation theory, Counterterms

A convenient setup for casting the renormalization procedure within perturbation theory is to write the bare Lagrangian,

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu \phi_b\right)\left(\partial^\mu \phi_b\right) - \frac{1}{2}m_b^2 \phi_b^2 - \frac{\lambda_b}{4!}\phi_b^4 \,, \tag{1.144}$$

(here we denote ϕ_b, m_b and λ_b the bare field, mass and coupling, to stress that they are not the physical ones) as the sum of a renormalized Lagrangian and a correction:

$$\begin{aligned}
\mathcal{L} &= \mathcal{L}_r + \Delta\mathcal{L} \\
\mathcal{L}_r &\equiv \frac{1}{2}\left(\partial_\mu \phi_r\right)\left(\partial^\mu \phi_r\right) - \frac{1}{2}m_r^2 \phi_r^2 - \frac{\lambda_r}{4!}\phi_r^4 \\
\Delta\mathcal{L} &\equiv \frac{1}{2}\Delta_z \left(\partial_\mu \phi_r\right)\left(\partial^\mu \phi_r\right) - \frac{1}{2}\Delta_m \phi_r^2 - \frac{1}{4!}\Delta_\lambda \phi_r^4 \,.
\end{aligned} \tag{1.145}$$

$\mathcal{L}_r$ contains the renormalized (i.e. physical) mass m_r and coupling constant λ_r (the latter may be defined from the measurement of some cross-section chosen as reference). In $\Delta\mathcal{L}$, the

---

[21]It may happen that an internal symmetry, such as a gauge symmetry, renders a function finite while its superficial degree of divergence is non negative.

[22]Non-renormalizable field theories may nevertheless be used as low energy effective field theories, where they approximate below a certain cutoff a more fundamental –possibly unknown– theory supposedly valid above the cutoff.

coefficients $\Delta_z, \Delta_m, \Delta_\lambda$ are called *counterterms*. Recalling that $\phi_b = \sqrt{Z}\,\phi_r$, the bare and physical parameters and the counterterms must be related by

$$
\begin{aligned}
\Delta_z &= Z - 1 \\
\Delta_m &= Z m_b^2 - m_r^2 \\
\Delta_\lambda &= Z^2 \lambda_b - \lambda_r \, .
\end{aligned}
\tag{1.146}
$$

The terms in $\Delta\mathcal{L}$ are treated as a perturbation to $\mathcal{L}_r$, and one may introduce extra Feynman rules for the various terms it contains:

$$
\frac{1}{2}\Delta_z \left(\partial_\mu \phi_r\right)\left(\partial^\mu \phi_r\right) - \frac{1}{2}\Delta_m \phi_r^2 \;\; \rightarrow \;\; \xrightarrow{\;P\;}\!\!\!\!\bullet\!\!-\!\!- \;\; = \;\; -i\left(\Delta_z P^2 + \Delta_m\right)
$$

$$
-\frac{1}{4!}\Delta_\lambda \phi_r^4 \;\; \rightarrow \;\; \times \;\; = \;\; -i\Delta_\lambda
\tag{1.147}
$$

At tree level, only the term $\mathcal{L}_r$ is used, and by construction the physical quantities computed at this order will depend only on physical parameters. Higher orders involve divergent loop corrections. The counterterms $\Delta_z, \Delta_m, \Delta_\lambda$ should be adjusted at every order to cancel the new divergences that arise at this order. In particular, after having included the contribution of the counterterms, the self-energy $\Sigma(P^2)$ are usually required to satisfy the following conditions[23]:

$$
\Sigma(m_r^2) = 0 \, , \quad \Sigma'(m_r^2) = 0 \, .
\tag{1.148}
$$

With this choice, it is not necessary to dress the external lines in the LSZ reduction formulas for transition amplitudes. Indeed, the renormalization conditions (1.148) imply that

$$
i(\Box + m_r^2)\, G_F = 1 \, , \quad \lim_{p^2 \to m_r^2} (-i\Sigma) G_F = 0 \, .
\tag{1.149}
$$

For each external line, the reduction formula contains an operator $i(\Box_x + m_r^2)$ acting on the dressed propagator of this external line:

$$
i(\Box + m_r^2)\left\{ \underbrace{G_F + G_F(-i\Sigma)G_F + G_F(-i\Sigma)G_F(-i\Sigma)G_F + \cdots}_{\text{dressed propagator}} \right\} = 1 \, .
\tag{1.150}
$$

Therefore, all the terms except the first one cancel, and we can ignore self-energy corrections on the external lines.

1.10 Spin 1/2 fields

1.10.1 Dimension-2 representation of the rotation group

In ordinary quantum mechanics, the spin s is related to the dimension n of representations of the rotation group by

$$
n = 2s + 1 \, .
\tag{1.151}
$$

---

[23]Strictly speaking, the only requirement is that the counterterms cancel the infinities, which does not fix uniquely their finite part. Various *renormalization schemes* are possible, that differ in how these finite parts are chosen.

Thus, spin 1/2 corresponds to representations of dimension 2. Such a representation is based on the (Hermitean) Pauli matrices:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{1.152}$$

from which we can construct the following unitary 2×2 matrices

$$U \equiv \exp\left(-\tfrac{i}{2}\theta^i\sigma^i\right). \tag{1.153}$$

That the Pauli matrices (up to a factor 2) are generators of the Lie algebra of rotations can be seen from

$$\left[J^i, J^j\right] = i\,\epsilon^{ijk}\,J^k \quad \text{with } J^i \equiv \frac{\sigma^i}{2}. \tag{1.154}$$

1.10.2   Spinor representation of the Lorentz group

This idea can be extended to quantum field theory in order to encompass all the Lorentz transformations rather than just the spatial rotations. We are therefore seeking a dimension 2 representation of the commutation relations (1.9). Firstly, let us assume that we know a set of four $n \times n$ matrices γ^μ that satisfy the following anti-commutation relation:

$$\{\gamma^\mu, \gamma^\nu\} = 2\,g^{\mu\nu}\,\mathbb{1}_{n \times n}. \tag{1.155}$$

Such matrices are called *Dirac matrices*. From these matrices, it is easy to check that the matrices

$$M^{\mu\nu} \equiv \frac{i}{4}\left[\gamma^\mu, \gamma^\nu\right] \tag{1.156}$$

form an n-dimensional representation of the Lorentz algebra. However, an exhaustive search indicates that the smallest matrices that fulfill eqs. (1.155) (in four space-time dimensions, i.e. for $\mu, \nu = 0, \cdots, 3$) are 4×4. Several unitarily equivalent choices are possible for these matrices. A possible representation (known as the Weyl or chiral representation) is the following[24]

$$\gamma^0 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma^i \equiv \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}. \tag{1.157}$$

In this representation, the generators for the boosts and for the rotations are

$$M^{0i} = -\frac{i}{2}\begin{pmatrix} \sigma^i & 0 \\ 0 & -\sigma^i \end{pmatrix}, \quad M^{ij} = \frac{1}{2}\,\epsilon^{ijk}\begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}. \tag{1.158}$$

Given a Lorentz transformation Λ defined by the parameters $\omega_{\mu\nu}$, let us define

$$U_{1/2}(\Lambda) \equiv \exp\left(-\frac{i}{2}\omega_{\mu\nu}M^{\mu\nu}\right). \tag{1.159}$$

---

[24]Although it is sometimes convenient to have an explicit representation of the Dirac matrices, most manipulations only rely on the fact that the obey the anti-commutation relations (1.155).

A *Dirac spinor* is a 4-component field $\psi(x)$ that transforms as follows:

$$\psi(x) \quad \rightarrow \quad U_{1/2}(\Lambda)\,\psi(\Lambda^{-1}x)\,. \tag{1.160}$$

In other words, the matrix $U_{1/2}$ defines how the four components of this field transform under a Lorentz transformation (since these four components mix, $\psi(x)$ is not the juxtaposition of four scalar fields). The fact that the lowest dimension for the Dirac matrices is 4 indicates that the spinor $\psi(x)$ describes two spin-$1/2$ particles: a particle and its *antiparticle*, that are distinct from each other.

1.10.3 Dirac equation and Lagrangian

Let us now determine an equation of motion obeyed by this field, such that it is invariant under Lorentz transformations. Since the $M^{\mu\nu}$'s act only on the Dirac indices, a trivial answer could be the Klein-Gordon equation,

$$\left(\Box_x + m^2\right)\psi(x) = 0\,. \tag{1.161}$$

But there is in fact a stronger equation that remains invariant when ψ is transformed according to eq. (1.160). Notice first that

$$U_{1/2}^{-1}(\Lambda)\gamma^{\mu}U_{1/2}(\Lambda) = \Lambda^{\mu}{}_{\nu}\gamma^{\nu}\,. \tag{1.162}$$

This equation indicates that rotating the Dirac indices of γ^{μ} with $U_{1/2}$ is equivalent to transforming the μ index as one would do for a normal 4-vector. Using this identity, we can check that under the same Lorentz transformation we have

$$\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi(x) \quad \rightarrow \quad U_{1/2}(\Lambda)\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi(\Lambda^{-1}x)\,. \tag{1.163}$$

Therefore, the *Dirac equation*,

$$\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi(x) = 0\,, \tag{1.164}$$

is Lorentz invariant. This equation implies the Klein-Gordon equation (to see it, apply the operator $i\gamma^{\mu}\partial_{\mu} + m$ on the left), and is therefore stronger.

The Dirac matrices are not Hermitean. Instead, they satisfy

$$\left(\gamma^{\mu}\right)^{\dagger} = \gamma^0\gamma^{\mu}\gamma^0\,. \tag{1.165}$$

Therefore, the Hermitic conjugate of $U_{1/2}(\Lambda)$ is

$$U_{1/2}^{\dagger}(\Lambda) = \exp\left(\frac{i}{2}\omega_{\mu\nu}(M^{\mu\nu})^{\dagger}\right) = \gamma^0\exp\left(\frac{i}{2}\omega_{\mu\nu}M^{\mu\nu}\right)\gamma^0 = \gamma^0\,U_{1/2}^{-1}(\Lambda)\,\gamma^0\,. \tag{1.166}$$

Because of this, the simplest Lorentz scalar bilinear combination of ψ's is $\psi^{\dagger}\gamma^0\psi$ (instead of the naive $\psi^{\dagger}\psi$). It is common to denote $\overline{\psi} \equiv \psi^{\dagger}\gamma^0$. From this, we conclude that the Lorentz scalar Lagrangian density that leads to the Dirac equation reads

$$\mathcal{L} = \overline{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi(x)\,. \tag{1.167}$$

31

1.10.4 Basis of free spinors

Before quantizing the spinor field in a similar fashion as the scalar field, we need to find plane wave solutions of the Dirac equation. There are two types of solutions:

$$
\begin{aligned}
\psi(x) &= u(\mathbf{p})\, e^{-ip\cdot x} \quad \text{with } (p_\mu\gamma^\mu - m)\, u(\mathbf{p}) = 0 \,, \\
\psi(x) &= v(\mathbf{p})\, e^{+ip\cdot x} \quad \text{with } (p_\mu\gamma^\mu + m)\, v(\mathbf{p}) = 0 \,.
\end{aligned}
\tag{1.168}
$$

The solutions $u(\mathbf{p})$ and $v(\mathbf{p})$ each form a 2-dimensional linear space, and it is customary to denote a basis by $u_s(\mathbf{p})$ and $v_s(\mathbf{p})$ (the index s, that takes two values $s = \pm$, is interpreted as the two spin states for a spin $1/2$ particle). A convenient normalization of the base vectors is

$$
\begin{aligned}
\bar{u}_r(\mathbf{p})u_s(\mathbf{p}) &= 2m\delta_{rs} \,, \quad u_r^\dagger(\mathbf{p})u_s(\mathbf{p}) = 2E_\mathbf{p}\delta_{rs} \,, \\
\bar{v}_r(\mathbf{p})v_s(\mathbf{p}) &= -2m\delta_{rs} \,, \quad v_r^\dagger(\mathbf{p})v_s(\mathbf{p}) = 2E_\mathbf{p}\delta_{rs} \,, \\
\bar{u}_r(\mathbf{p})v_s(\mathbf{p}) &= \bar{v}_r(\mathbf{p})u_s(\mathbf{p}) = 0 \,.
\end{aligned}
\tag{1.169}
$$

When summing over the spin states, we have:

$$
\sum_{s=\pm} u_s(\mathbf{p})\bar{u}_s(\mathbf{p}) = \not{p} + m \,, \quad \sum_{s=\pm} v_s(\mathbf{p})\bar{v}_s(\mathbf{p}) = \not{p} - m \,,
\tag{1.170}
$$

where we have introduced the notation $\not{p} \equiv p_\mu\gamma^\mu$.

1.10.5 Canonical quantization

From the Lagrangian (1.167), the momentum canonically conjugated to $\psi(x)$ is

$$
\Pi(x) = i\psi^\dagger(x) \,.
\tag{1.171}
$$

Trying to generalize the canonical commutation relation of scalar field operators (1.25) would lead to

$$
\left[\psi_a(x), \psi_b^\dagger(y)\right]_{x^0=y^0} = \delta(\mathbf{x} - \mathbf{y})\delta_{ab} \,,
\tag{1.172}
$$

where we have written explicitly the Dirac indices a, b. However, by decomposing $\psi(x)$ on a basis of plane waves by introducing creation and annihilation operators,

$$
\psi(x) \equiv \sum_{s=\pm} \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_\mathbf{p}} \left\{ a_{s\mathbf{p}}^\dagger\, v_s(\mathbf{p})e^{+ip\cdot x} + b_{s\mathbf{p}}\, u_s(\mathbf{p})e^{+ip\cdot x} \right\} \,,
\tag{1.173}
$$

one would find a Hamiltonian which is not bounded from below. The resolution of this paradox is that the commutation relation (1.172) is incorrect, and should be replaced by an *anti-commutation* relation,

$$
\left\{\psi_a(x), \psi_b^\dagger(y)\right\}_{x^0=y^0} = \delta(\mathbf{x} - \mathbf{y})\delta_{ab} \,,
\tag{1.174}
$$

which leads to anti-commutation relations for the creation and annihilation operators

$$
\left\{a_{r\mathbf{p}}, a_{s\mathbf{q}}^\dagger\right\} = \left\{b_{r\mathbf{p}}, b_{s\mathbf{q}}^\dagger\right\} = (2\pi)^3 2E_\mathbf{p}\delta(\mathbf{p} - \mathbf{q})\delta_{rs} \,.
\tag{1.175}
$$

(All other combinations are zero.) These anti-commutation relations imply that the square of creation operators is zero, which means that it is not possible to have two particles with the same momentum and spin in a quantum state. This is nothing but the *Pauli exclusion principle*. This is the simplest example of the *spin-statistics theorem*, which states that half-integer spin particles must obey Fermi statistics.

1.10.6 Free spin-1/2 propagator

From eq. (1.173), we obtain the following expression for the free Feynman propagator of the Dirac field[25]

$$
\begin{aligned}
S_F^0(x,y) &\equiv \langle 0 | \underbrace{\theta(x^0 - y^0)\psi_a(x)\overline{\psi}_b(y) - \theta(y^0 - x^0)\overline{\psi}_b(y)\psi_a(x)}_{T(\psi_a(x)\overline{\psi}_b(y))} | 0 \rangle \\
&= \int \frac{d^4 p}{(2\pi)^4}\, e^{-ip\cdot(x-y)}\, \underbrace{\frac{i(\not{p} + m)}{p^2 - m^2 + i0^+}}_{S_F^0(p)} \; .
\end{aligned}
\tag{1.176}
$$

The diagrammatic representation of this propagator is a line with an arrow:

$$
S_F^0(p) \quad = \quad \xrightarrow{\;\;p\;\;} \; .
\tag{1.177}
$$
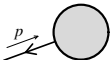
1.10.7 LSZ reduction formula for spin-1/2

The LSZ reduction formula for transition amplitudes with fermions and/or anti-fermions in the initial and final states reads:

$$
\begin{aligned}
\langle \underbrace{q_\sigma \overline{q}_{\overline{\sigma}} \cdots}_{n\text{ particles}} \,_{\text{out}} | \underbrace{p_s \overline{p}_{\overline{s}} \cdots}_{m\text{ particles}} \,_{\text{in}} \rangle &\doteq \left(\frac{i}{Z^{1/2}}\right)^{m+n} \int d^4 x\, e^{-ip\cdot x} \int d^4 \overline{x}\, e^{-i\overline{p}\cdot\overline{x}} \cdots \\
&\times \int d^4 y\, e^{+iq\cdot y} \int d^4 \overline{y}\, e^{+i\overline{q}\cdot\overline{y}} \cdots \overline{v}_{\overline{s}}(\overline{p})(i\overset{\rightarrow}{\not\partial}_{\overline{x}} - m)\,\overline{u}_\sigma(q)(-i\overset{\rightarrow}{\not\partial}_y + m) \\
&\times \langle 0_{\text{out}} | T\,\overline{\psi}(x)\overline{\psi}(\overline{y})\psi(\overline{x})\psi(y)\cdots | 0_{\text{in}}\rangle (i\overset{\leftarrow}{\not\partial}_x + m)u_s(p)\,(-i\overset{\leftarrow}{\not\partial}_{\overline{y}} - m)v_{\overline{\sigma}}(\overline{q}) \; ,
\end{aligned}
\tag{1.178}
$$

where we give examples for fermions and anti-fermions (indicated by a bar over the momentum and spin), both for the initial and final states. Besides the requirement that the external lines of the Feynman graphs should be amputated, this formula leads to the following prescriptions for the open ends of fermionic lines:

Incoming fermion :
 $\quad = \quad u(\mathbf{p})$

Incoming anti-fermion :
 $\quad = \quad \overline{v}(\mathbf{p})$

Outgoing fermion :
 $\quad = \quad \overline{u}(\mathbf{p})$

Outgoing anti-fermion :
 $\quad = \quad v(\mathbf{p}) \; .$

---

[25]We have introduced a minus sign in the definition of the time-ordered product of Dirac fields. One would have to mimic the derivation of the section 1.5 in order to see that this is the propagator that naturally appears in the generating functional for the amplitudes with fermions.

Note that when writing the expression corresponding to a given Feynman graph, the fermion lines it contains must be read in the direction opposite to the arrow carried by the lines.

1.11 Spin 1 fields

1.11.1 Classical electrodynamics

The best known spin-1 particle is the photon. In *classical electrodynamics*, the electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$ obey *Maxwell's equations*,

$$
\begin{aligned}
\boldsymbol{\nabla} \cdot \mathbf{E} &= \rho \\
\boldsymbol{\nabla} \times \mathbf{B} - \partial_t \mathbf{E} &= \mathbf{J} \\
\boldsymbol{\nabla} \times \mathbf{E} + \partial_t \mathbf{B} &= 0 \\
\boldsymbol{\nabla} \cdot \mathbf{B} &= 0 \,,
\end{aligned}
\tag{1.179}
$$

written here in terms of charge density ρ and current $\mathbf{J}$. The local conservation of electrical charge implies the following continuity equation

$$
\partial_t \rho + \boldsymbol{\nabla} \cdot \mathbf{J} = 0 \,.
\tag{1.180}
$$

The last two Maxwell's equations are automatically satisfied if we write the $\mathbf{E}, \mathbf{B}$ fields in terms of potentials V and $\mathbf{A}$,

$$
\mathbf{E} \equiv \partial_t \mathbf{A} + \boldsymbol{\nabla} V \quad , \qquad \mathbf{B} \equiv -\boldsymbol{\nabla} \times \mathbf{A} \,.
\tag{1.181}
$$

This representation is not unique, since $\mathbf{E}$ and $\mathbf{B}$ are unchanged if we transform the potentials as follows:

$$
V \to V + \partial_t \chi \quad , \qquad \mathbf{A} \to \mathbf{A} - \boldsymbol{\nabla} \chi \,,
\tag{1.182}
$$

where χ is an arbitrary function of space and time. Eq. (1.182) is called a (Abelian) *gauge transformation*. Quantities that do not change under (1.182) are said to be *gauge invariant*. For instance, the electrical and magnetic fields are invariant.

1.11.2 Classical electrodynamics in Lorentz covariant form

In order to make manifest the properties of Maxwell's equations under Lorentz transformations, let us firstly rewrite them in covariant form. Introduce a 4-vector A^μ and a rank-2 tensor $F^{\mu\nu}$,

$$
A^\mu \equiv (V, \mathbf{A}) \quad , \qquad F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu \,.
\tag{1.183}
$$

($F^{\mu\nu}$ is called the *field strength*.) Recalling that $\partial^\mu = (\partial_t, -\boldsymbol{\nabla})$, gauge transformations take the following form

$$
A^\mu \to A^\mu + \partial^\mu \chi \,,
\tag{1.184}
$$

and $F^{\mu\nu}$ is gauge invariant. Moreover, we see that

$$
E^i = F^{0i} \quad , \qquad B^i = \tfrac{1}{2} \epsilon^{ijk} F^{jk} \,.
\tag{1.185}
$$

If we also encapsulate ρ and $\mathbf{J}$ in a 4-vector,

$$J^\mu \equiv (\rho, \mathbf{J}) \,, \tag{1.186}$$

the first two Maxwell's equations and the continuity equation read

$$\partial_\mu F^{\mu\nu} = -J^\nu \quad , \qquad \partial_\mu J^\mu = 0 \,. \tag{1.187}$$

The last two Maxwell's equations become

$$\epsilon_{\mu\nu\rho\sigma} \partial^\nu F^{\rho\sigma} = 0 \,. \tag{1.188}$$

(It is automatically satisfied thanks to the antisymmetric structure of $F^{\mu\nu}$.)

A Lorentz scalar Lagrangian density whose Euler-Lagrange equations of motion are the Maxwell's equations is

$$\mathcal{L} \equiv -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + J^\mu A_\mu \,. \tag{1.189}$$

Because of the term $J^\mu A_\mu$ that couples the potential to the sources, this Lagrangian density is not gauge invariant, but the action (integral of $\mathcal{L}$ over all space-time) is, provided that the current is conserved (i.e. satisfies the continuity equation). Indeed, we have

$$\int d^4x\, J^\mu A_\mu \quad \rightarrow \quad \int d^4x\, J^\mu (A_\mu + \partial_\mu \chi) = \int d^4x\, J^\mu A_\mu - \int d^4x\, \chi \underbrace{\partial_\mu J^\mu}_{0} + \begin{matrix} \text{boundary} \\ \text{term} \end{matrix} \,. \tag{1.190}$$

(The boundary term is zero if we assume that there are no sources at infinity.)

1.11.3 Canonical quantization in Coulomb gauge

Although it leads to Maxwell's equations, the above Lagrangian has an unusual property, related to gauge invariance: the conjugate momentum of the potential A^0 is identically zero,

$$\Pi^0(x) \equiv \frac{\delta \mathcal{L}}{\delta \partial_0 A^0(x)} = 0 \,. \tag{1.191}$$

Therefore, we cannot quantize electrodynamics simply by promoting the Poisson bracket between A^0 and its conjugate momentum to a commutator. However, this problem is not intrinsic to quantum mechanics: the very same issue arises when trying to formulate classical electrodynamics in Hamilton form. The resolution of this problem is to fix the gauge, i.e. to impose an extra condition on the potential A^μ such that a unique A^μ corresponds to given $\mathbf{E}$ and $\mathbf{B}$ fields. Possible *gauge conditions* are:

Axial gauge :	$n^\mu A_\mu = 0$	(n^μ is a fixed 4-vector) ,
Lorenz gauge :	$\partial^\mu A_\mu = 0$,	
Coulomb gauge :	$\boldsymbol{\nabla} \cdot \mathbf{A} = 0$.	(1.192)

Let us illustrate this procedure in Coulomb gauge[26]. Firstly, let us decompose the vector potential $\mathbf{A}^i$ into longitudinal and transverse components:

$$\mathbf{A}^i = \mathbf{A}^i_\parallel + \mathbf{A}^i_\perp \,, \tag{1.193}$$

with

$$\mathbf{A}^i_\parallel \equiv \frac{\partial^i \partial^j}{\partial^2} \mathbf{A}^j \quad , \qquad \mathbf{A}^i_\perp \equiv \left(\delta^{ij} - \frac{\partial^i \partial^j}{\partial^2} \right) \mathbf{A}^j \,. \tag{1.194}$$

The Coulomb gauge condition is equivalent to $\mathbf{A}^i_\parallel = 0$. The remaining components of A^μ are therefore A^0 and the two components of $\mathbf{A}^i_\perp$, in terms of which the Lagrangian reads:

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}(\partial_t \mathbf{A}^i_\perp)(\partial_t \mathbf{A}^i_\perp) - \frac{1}{2}(\partial_j \mathbf{A}^i_\perp)(\partial_j \mathbf{A}^i_\perp) + \frac{1}{2}(\partial_i A^0)(\partial_i A^0) \\ &\quad + \underline{(\partial_t \mathbf{A}^i_\perp)(\partial_i A^0) + \frac{1}{2}(\partial_i \mathbf{A}^j_\perp)(\partial_j \mathbf{A}^i_\perp)} + J^0 A^0 - J^i \mathbf{A}^i_\perp \,. \end{aligned} \tag{1.195}$$

Note that the two underlined terms will vanish in the action, after an integration by parts (thanks to the transversality of $\mathbf{A}^i_\perp$). The Euler-Lagrange equation for the field A^0 is

$$\partial^2 A^0 = J^0 \,, \tag{1.196}$$

i.e. the Poisson equation with source term J^0. Note that this equation has no time derivative. Therefore, A^0 reflects instantaneously the changes of the charge density J^0 (this does not contradict special relativity, since A^0 is not an observable – only $\mathbf{E}$ and $\mathbf{B}$ are). Ignoring all the terms that would vanish in the action upon integration by parts, we may thus rewrite the Lagrangian as

$$\mathcal{L} = \frac{1}{2}(\partial_t \mathbf{A}^i_\perp)(\partial_t \mathbf{A}^i_\perp) - \frac{1}{2}(\partial_j \mathbf{A}^i_\perp)(\partial_j \mathbf{A}^i_\perp) - J^i \mathbf{A}^i_\perp + \frac{1}{2} J^0 \frac{1}{\partial^2} J^0 \,, \tag{1.197}$$

and obtain the following Euler-Lagrange equation of motion for the field $\mathbf{A}^i_\perp$:

$$\square \mathbf{A}^i_\perp = -\left(\delta^{ij} - \frac{\partial^i \partial^j}{\partial^2} \right) J^j \,, \tag{1.198}$$

i.e. a massless Klein-Gordon equation with the transverse projection of the charge current as source term.

In this form, electrodynamics has no redundant degrees of freedom, and can now be quantized in the vacuum ($J^0 = J^i = 0$) in the canonical way. Firstly, we define the momentum conjugated to $\mathbf{A}^i_\perp$,

$$\Pi^i_\perp(x) \equiv \frac{\delta \mathcal{L}}{\delta \partial_t \mathbf{A}^i_\perp(x)} = \partial_t \mathbf{A}^i_\perp(x) \,. \tag{1.199}$$

The, we promote $\mathbf{A}^i_\perp$ and $\Pi^i_\perp$ to quantum operators, and we impose on them the following canonical equal-time commutation relations,

$$\begin{aligned} \left[\mathbf{A}^i_\perp(x), \Pi^j_\perp(y) \right]_{x^0=y^0} &= i \left(\delta^{ij} - \frac{\partial^i \partial^j}{\partial^2} \right) \delta(\mathbf{x} - \mathbf{y}) \,, \\ \left[\mathbf{A}^i_\perp(x), \mathbf{A}^j_\perp(y) \right]_{x^0=y^0} &= \left[\Pi^i_\perp(x), \Pi^j_\perp(y) \right]_{x^0=y^0} = 0 \,. \end{aligned} \tag{1.200}$$

---

[26]One may start from another gauge condition, and follow a similar line of reasoning in order to derive a quantized theory of the photon field in another gauge. However, as we shall see later, we can make the gauge fixing much more transparent by using functional quantization.

(In the first of these relations, the transverse projector in the right hand side follows from the fact that $\mathbf{A}_\perp^i$ and $\mathbf{\Pi}_\perp^j$ are both transverse.) These commutation relations can be realized by decomposing $\mathbf{A}_\perp^i$ on a basis of solutions of the Klein-Gordon equation, i.e. plane waves:

$$\mathbf{A}_\perp^i(x) \equiv \sum_{\lambda=1,2} \int \frac{d^3\mathbf{p}}{(2\pi)^3 2|\mathbf{p}|} \left[\boldsymbol{\epsilon}_\lambda^i(\mathbf{p})\, a_{\lambda\mathbf{p}}^\dagger\, e^{+ip\cdot x} + \boldsymbol{\epsilon}_\lambda^{i*}(\mathbf{p})\, a_{\lambda\mathbf{p}}\, e^{-ip\cdot x} \right] , \qquad (1.201)$$

where the two vectors $\boldsymbol{\epsilon}_{1,2}^i(\mathbf{p})$ are *polarization vectors* orthogonal to $\mathbf{p}$,

$$\mathbf{p} \cdot \boldsymbol{\epsilon}_\lambda(\mathbf{p}) = 0 . \qquad (1.202)$$

In 3 spatial dimensions, a basis of such vectors has two elements, that we have labeled with $\lambda = 1, 2$. In addition, it is convenient to normalize the polarization vectors as follows

$$\boldsymbol{\epsilon}_\lambda(\mathbf{p}) \cdot \boldsymbol{\epsilon}_{\lambda'}^*(\mathbf{p}) = \delta_{\lambda\lambda'} \quad , \qquad \sum_{\lambda=1,2} \boldsymbol{\epsilon}_\lambda^{i*}(\mathbf{p}) \boldsymbol{\epsilon}_\lambda^j(\mathbf{p}) = \delta^{ij} - \frac{p^i p^j}{\mathbf{p}^2} . \qquad (1.203)$$

With this choice, the commutation relations of eqs. (1.200) are equivalent to the following commutation relations between creation and annihilation operators:

$$\begin{aligned}
\left[a_{\lambda\mathbf{p}}, a_{\lambda'\mathbf{q}} \right] &= \left[a_{\lambda\mathbf{p}}^\dagger, a_{\lambda'\mathbf{q}}^\dagger \right] = 0 , \\
\left[a_{\lambda\mathbf{p}}, a_{\lambda'\mathbf{q}}^\dagger \right] &= (2\pi)^3\, 2|\mathbf{p}|\, \delta_{\lambda\lambda'}\, \delta(\mathbf{p} - \mathbf{q}) .
\end{aligned} \qquad (1.204)$$

1.11.4 Feynman rules for photons

Eq. (1.201) can be inverted to obtain the creation and annihilation operators as

$$\begin{aligned}
a_{\lambda\mathbf{p}}^\dagger &= -i\, \boldsymbol{\epsilon}_\lambda^{i*}(\mathbf{p}) \int d^3\mathbf{x}\, e^{-ip\cdot x}\, \overset{\leftrightarrow}{\partial_0}\, \mathbf{A}_\perp^i(x) , \\
a_{\lambda\mathbf{p}} &= +i\, \boldsymbol{\epsilon}_\lambda^i(\mathbf{p}) \int d^3\mathbf{x}\, e^{+ip\cdot x}\, \overset{\leftrightarrow}{\partial_0}\, \mathbf{A}_\perp^i(x) ,
\end{aligned} \qquad (1.205)$$

With these formulas, it is easy to derive the LSZ reduction formulas for photons in the initial and final states,

$$\begin{aligned}
\langle \underbrace{\mathbf{q}_{\lambda'}\cdots}_{n\text{ photons}}{}_{\text{out}} | \underbrace{\mathbf{p}_\lambda\cdots}_{m\text{ photons}}{}_{\text{in}} \rangle &\doteq \left(\frac{i}{Z^{1/2}} \right)^{m+n} \int d^4x\, e^{-ip\cdot x}\, \boldsymbol{\epsilon}_\lambda^{i*}(\mathbf{p})\, \Box_x \cdots \\
&\times \int d^4y\, e^{+iq\cdot y}\, \boldsymbol{\epsilon}_{\lambda'}^j(\mathbf{q})\, \Box_y \cdots \langle 0_{\text{out}} | T\left(\mathbf{A}_\perp^i(x) \mathbf{A}_\perp^j(y) \cdots \right) | 0_{\text{in}} \rangle .
\end{aligned} \qquad (1.206)$$

The free Feynman propagator of the photon (in Coulomb gauge) can be read off the quadratic part of the Lagrangian (1.197). In momentum space, it reads

$$G_F^{0\ ij}(\mathbf{p}) = i\ \overset{p}{\underset{i \sim\sim\sim\sim\sim j}{\longrightarrow}} = \frac{i\left(\delta^{ij} - \frac{p^i p^j}{\mathbf{p}^2} \right)}{p^2 + i0^+} . \qquad (1.207)$$

The operator $\epsilon_\lambda^i(\mathbf{p}) \square_x$ in the reduction formula simply amputates the external photon line to which it is applied[27]. Transition amplitudes with incoming and outgoing photons are therefore given by amputated graphs, with a polarization vector contracted to the Lorentz index of each external photon.

1.12 Abelian gauge invariance, Quantum Electrodynamics

So far, we have derived a quantized field theory for spin $1/2$ fermions and a quantized field theory of photons (in the absence of charged sources), but they appear as unrelated constructions. The next step is to combine the two into a quantum theory of charged fermions that interact electromagnetically via photon exchanges.

1.12.1 Global $U(1)$ symmetry of the Dirac Lagrangian

Firstly, note that the fermion Lagrangian is invariant under the following transformation of the fermion field

$$\psi \quad \rightarrow \quad \Omega^\dagger \psi \, , \tag{1.208}$$

where Ω is a phase (i.e. an element of the group $U(1)$), provided that we consider only rigid transformations (i.e. independent of the space-time point x). By *Noether's theorem*, this continuous symmetry corresponds to the existence of a conserved current,

$$J^\mu = \overline{\psi} \gamma^\mu \psi \, . \tag{1.209}$$

It is indeed straightforward to check from Dirac's equation that

$$\partial_\mu J^\mu = 0 \, . \tag{1.210}$$

A more physical interpretation of this current emerges from the spatial integral of the component J^0,

$$Q \equiv \int d^3\mathbf{x} \, J^0(x) \, . \tag{1.211}$$

Using the Fourier mode decomposition (1.173) of the spinor $\psi(x)$, we obtain the following expression:

$$\begin{aligned} Q &= \sum_{s=\pm} \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_\mathbf{p}} \left\{ a_{s\mathbf{p}} a_{s\mathbf{p}}^\dagger + b_{s\mathbf{p}}^\dagger b_{s\mathbf{p}} \right\} \\ &= \sum_{s=\pm} \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_\mathbf{p}} \left\{ b_{s\mathbf{p}}^\dagger b_{s\mathbf{p}} - a_{s\mathbf{p}}^\dagger a_{s\mathbf{p}} \right\} + \text{(infinite) constant} \, . \end{aligned} \tag{1.212}$$

Thus, the operator Q counts the number of particles created by $b^\dagger$ minus the number of particles created by $a^\dagger$. If we assign a *charge* $+1$ to the former and -1 to the latter, wee can interpret Q as the operator that measures the total charge in the system.

---

[27]Note that

$$\left(\delta^{ij} - \frac{p^i p^j}{\mathbf{p}^2} \right) \epsilon_\lambda^j(\mathbf{p}) = \epsilon_\lambda^i(\mathbf{p}) \, .$$

Therefore, the transverse projectors attached to the external photon lines can be dropped.

1.12.2 Minimal coupling to a spin-1 field

Secondly, note that the gauge transformation of the potential A^μ, given in eq. (1.184), can also be written in the following form[28]:

$$A^\mu \quad \to \quad \Omega^\dagger A^\mu \Omega + i\,\Omega^\dagger \partial^\mu \Omega \,, \tag{1.213}$$

with

$$\Omega(x) \equiv e^{-i\,\chi(x)} \,. \tag{1.214}$$

When written in this form, the gauge transformation of the photon field appears to be also generated by the group $U(1)$. Unlike the quantum field theory for fermions, the photon Lagrangian is invariant under *local gauge transformations*, i.e. where Ω depends on x in an arbitrary fashion. Therefore, at this point we have two disjoint quantum field theories: a theory of non-interacting charged fermions that has a global $U(1)$ invariance, and a theory of non-interacting photons that has a local $U(1)$ invariance, but no coupling between the two.

Let us see what minimal modification would be necessary in order to promote the $U(1)$ symmetry of the fermion sector into a local symmetry. An immediate obstacle is that

$$\Omega(x)\,\partial^\mu\,\Omega^\dagger(x) \neq \partial^\mu \,. \tag{1.215}$$

Equivalently, the problem comes from the fact that the derivative $\partial_\mu \psi$ does not transform in the same way as ψ itself when Ω depends on x. Instead, we have

$$\partial_\mu \psi \quad \to \quad \partial_\mu \Omega^\dagger\,\psi = \Omega^\dagger\,\partial_\mu \psi + (\partial_\mu \Omega^\dagger)\,\psi \,. \tag{1.216}$$

But we see that the second term can be connected to the variation of a photon field under the same transformation. This suggests that the combination $(\partial_\mu - iA_\mu)\psi$ has a simpler transformation law:

$$\begin{aligned}
\left(\partial_\mu - iA_\mu\right)\psi \quad \to \quad & \left(\partial_\mu - i\left(\Omega^\dagger A_\mu \Omega + i\Omega^\dagger \partial_\mu \Omega\right)\right)\Omega^\dagger\psi \\
= & \ \Omega^\dagger\left(\partial_\mu - iA_\mu\right)\psi + \Omega^\dagger\Big(\underbrace{\Omega(\partial_\mu \Omega^\dagger) + (\partial_\mu \Omega)\Omega^\dagger}_{\partial_\mu(\Omega\Omega^\dagger)=0}\Big)\,\psi \,.
\end{aligned} \tag{1.217}$$

The operator $D_\mu \equiv \partial_\mu - iA_\mu$ is called a *covariant derivative*. The above calculation shows that $\overline{\psi}D_\mu\psi$ is invariant under local gauge transformations.

1.12.3 Abelian gauge theories

This observation is the basis of (Abelian) gauge theories: the minimal change to the Dirac Lagrangian that makes it locally gauge invariant introduces a coupling $\overline{\psi}A_\mu\psi$ between two fermion fields and a spin-1 field such as the photon. The complete Lagrangian of this theory therefore reads:

$$\mathcal{L} = -\frac{1}{4}\,F_{\mu\nu}F^{\mu\nu} + \overline{\psi}\big(i\slashed{D} - m\big)\,\psi \,. \tag{1.218}$$

---

[28]Naturally, $\Omega^\dagger A^\mu \Omega = A^\mu$. We have used this somewhat more complicated form to highlight the analogy with the non-Abelian gauge theories that we will study later.

We already know the Feynman rules for the photon and fermion propagators, and the pre-scription for external photon and fermion lines. The only additional Feynman rules is the following interaction vertex,

$$= -i\gamma^\mu \, , \qquad (1.219)$$

that can be read off directly from the Lagrangian.

Quantum Electrodynamics (QED) is a quantum field theory that describes the interactions between electromagnetic radiation (photons) and charged particles (electrons and positrons for instance), whose Lagrangian is of the form (1.218). The only necessary generalization compared to the previous discussion is to introduce a parameter e that represents the (bare) electrical charge of the electron, which leads to the following changes:

$$\begin{aligned}
\text{Covariant derivative :} \quad & D_\mu \equiv \partial_\mu - i\,e\,A_\mu \\
\text{Gauge transformation of the photon :} \quad & A_\mu \;\to\; \Omega^\dagger A^\mu \Omega + \frac{i}{e}\,\Omega^\dagger \partial^\mu \Omega \\
\text{Electrical current :} \quad & e\,\overline{\psi}\gamma^\mu\psi \\
\text{Photon-electron vertex :} \quad & -i\,e\,\gamma^\mu
\end{aligned} \qquad . \quad (1.220)$$

1.13 Charge conservation and Ward-Takahashi identities

1.13.1 Charge of 1-particle states

The charge operator Q defined in eq. (1.211) is invariant by translation in time (because J^μ is a conserved current) and in space (because it is integrated over all space). Since the current J^μ is a 4-vector, Q is also invariant under Lorentz transformations. Therefore Q conserves the energy and momentum of the states on which it acts. When acting on the vacuum state, one has

$$Q\,|0\rangle = 0 \, . \qquad (1.221)$$

When acting on a 1-particle state $|\alpha_{\mathbf{p}}\rangle$, Q gives another state with the same 4-momentum, and therefore the same invariant mass. But since single particle states are separated from states with a higher occupancy in the spectral function of the theory, $Q\,|\alpha_{\mathbf{p}}\rangle$ must in fact be proportional to $|\alpha_{\mathbf{p}}\rangle$ itself,

$$Q\,|\alpha_{\mathbf{p}}\rangle = q_{\alpha,\mathbf{p}}\,|\alpha_{\mathbf{p}}\rangle \, . \qquad (1.222)$$

In other words, 1-particle states are eigenvectors of the charge operator. Since Q is Lorentz invariant, the eigenvalue $q_{\alpha,\mathbf{p}}$ cannot depend on the momentum $\mathbf{p}$ (nor on the spin state of the particle), and it can only depend on the species of particle α. We will thus denote it q_α, and call it the electrical charge of the particle of type α.

In theories with 1-particle states that do not correspond to the fundamental fields of the Lagrangian (e.g. composite bound states made of several elementary particles), one may go a bit further. The canonical anti-commutation relations imply

$$\left[J^0(x),\psi(y)\right]_{x^0=y^0} = -e\,\psi(x)\,\delta(\mathbf{x}-\mathbf{y}) \, , \quad \left[Q,\psi(y)\right] = -e\,\psi(y) \, . \qquad (1.223)$$

More generally, for any local function $F(\psi(x), \psi^\dagger(x))$, we have

$$\left[Q, F(\psi(y), \psi^\dagger(y))\right]_{x^0 = y^0} = -e\,(n_+ - n_-)\,F(\psi(y), \psi^\dagger(y)) \qquad (1.224)$$

where n_+ is the number of ψ's in F and n_- the number of $\psi^\dagger$'s. If we evaluate this identity between the vacuum and a 1-particle state $|\alpha_{\mathbf{p}}\rangle$, we obtain

$$\langle 0|F(\psi(y), \psi^\dagger(y))|\alpha_{\mathbf{p}}\rangle\,(q_\alpha - (n_+ - n_-)e) = 0\,. \qquad (1.225)$$

Therefore, if the operator $F(\psi, \psi^\dagger)$ can create the particle α from the vacuum (i.e. the matrix element in the left hand side is non-zero), then we must have

$$q_\alpha = (n_+ - n_-)\,e\,. \qquad (1.226)$$

In other words, the charge of the particle α is the number of ψ's it contains, minus the number of $\psi^\dagger$'s, times the electrical charge e of the field ψ (as it appears in the Lagrangian). The non-trivial aspect of this assertion comes from the fact that it does not depend on the (usually complicated and non-perturbative) interactions that produce the binding.

So far, we have not discussed the renormalization of the parameter e. Its renormalized value e_r should be such that the covariant derivative retains its form[29] when expressed in terms of the renormalized photon field A_r^μ, i.e.

$$\partial^\mu - i\,e_r A_r^\mu\,. \qquad (1.227)$$

Since the field A_r^μ is related to the bare photon field A_b^μ by

$$A_b^\mu = Z_3^{1/2}\,A_r^\mu\,, \qquad (1.228)$$

the bare and renormalized charges must be related by

$$e_b = Z_3^{-1/2}\,e_r\,. \qquad (1.229)$$

In combination with eq. (1.226), this means that the charges of all 1-particle states are renormalized by the same factor $Z_3^{-1/2}$, regardless of the species of particle contained in the state. For this to work, cancellations between various Feynman graphs are necessary. These cancellations are a consequence of the local gauge invariance of the theory, and in their simplest form they can be encapsulated in the *Ward-Takahashi identities*, that we shall derive now.

1.13.2 Ward-Takahashi identities

Amplitudes with amputated external photon lines can be obtained as follows:

$$\begin{aligned} M^{\mu_1 \mu_2 \cdots}(q_1, q_2, \cdots) &= \int d^4 x_1\, d^4 x_2 \cdots e^{-iq_1 \cdot x_1}\,e^{-iq_2 \cdot x_2} \cdots \\ &\quad \times \langle \beta_{\text{out}}|T\left\{J^{\mu_1}(x_1) J^{\mu_2}(x_2) \cdots\right\}|\alpha_{\text{in}}\rangle\,, \end{aligned} \qquad (1.230)$$

where only electromagnetic currents appear inside the T-product, and all the external charged particles are kept in the initial and final states α and β (and are therefore on-shell).

---

[29]The implicit assumption of this sentence is that the renormalization of QED preserves its local gauge invariance.

Let us contract the Lorentz index μ_1 with the momentum $q_1^{\mu_1}$ of the first photon. After an integration by parts, this reads

$$
\begin{aligned}
q_{1,\mu_1}\, M^{\mu_1\mu_2\cdots}(q_1, q_2, \cdots) \;=\; & -i \int d^4x_1\, d^4x_2 \cdots e^{-iq_1\cdot x_1}\, e^{-iq_2\cdot x_2}\cdots \\
& \times \big\langle 0_{\text{out}}\big|\partial_{\mu_1}\, T\big\{J^{\mu_1}(x_1)J^{\mu_2}(x_2)\cdots\big\}\big|0_{\text{in}}\big\rangle \,.
\end{aligned}
\tag{1.231}
$$

The derivative of the T-product involves two types of terms: (i) terms where the derivative acts directly on the current $J^{\mu_1}(x_1)$, that are zero thanks to current conservation, and (ii) terms where it acts on the theta functions that order the times inside the T-product. With two currents, the latter term reads[30]

$$
\frac{\partial}{\partial x^\mu}\, T\big\{J^\mu(x)J^\nu(y)\big\} = \delta(x^0 - y^0)\,\big[J^0(x), J^\nu(y)\big] = 0 \,.
\tag{1.232}
$$

This generalizes to more than two currents, and we therefore have quite generally

$$
q_{1,\mu_1}\, M^{\mu_1\mu_2\cdots}(q_1, q_2, \cdots) = 0 \,.
\tag{1.233}
$$

The same property would hold for all the external photon lines of the amplitude. This equation is known as the *Ward-Takahashi identity*.

A consequence of eq. (1.233) is that QED transition amplitudes are unchanged if the photon propagators or polarization vectors are modified by terms proportional to the momentum p^μ,

$$
\begin{aligned}
G^{0\ \mu\nu}_{\text{F}}(p) \quad &\to \quad G^{0\ \mu\nu}_{\text{F}}(p) + a^\mu\, p^\nu + b^\nu\, p^\mu \\
\epsilon^\mu_\lambda(\mathbf{p}) \quad &\to \quad \epsilon^\mu_\lambda(\mathbf{p}) + c\, p^\mu \,.
\end{aligned}
\tag{1.234}
$$

This is precisely the modification of the Feynman rules one would encounter by using a different gauge fixing in the quantization of the theory. Thus, the Ward-Takahashi identities imply the gauge invariance of the transitions amplitudes in QED.

1.14 Perturbative Unitarity

Unitarity is one of the pillars of quantum mechanics, since it is tightly related to the conservation of probability. A completely general consequence of unitarity is the *optical theorem*, whose perturbative translation becomes manifest in the so-called *Cutkosky's cutting rules*.

1.14.1 Optical theorem

The "S-matrix" is the name given to the evolution operator that relates the *in* and *out* states:

$$
\big\langle \alpha_{\text{out}}\big| \equiv \big\langle \alpha_{\text{in}}\big|\, S \,.
\tag{1.235}
$$

---

[30]This step of the argument would fail if we had kept charged field operators inside the T-product, because their equal-time commutator with J^0 is non-zero. Therefore, the Ward-Takahashi identities are valid provided all the external charged particles are on-shell, but there is no such requirement for the neutral external particles (e.g. the photons).

In a unitary field theory, the S matrix is a unitary operator *on the space of physical states*:

$$SS^\dagger = S^\dagger S = 1 \,. \tag{1.236}$$

This property means that for a properly normalized initial physical state $\left|\alpha_{\text{in}}\right\rangle$, we have

$$\sum_{\text{states } \beta} |\langle\beta_{\text{out}}|\alpha_{\text{in}}\rangle|^2 = 1 \,, \tag{1.237}$$

where the sum includes only physical states. In other words, in any interaction process, the state α must evolve with probability one into other physical states. In general, one subtracts from the S-matrix the identity operator, that corresponds to the absence of interactions, and one writes:

$$S \equiv 1 + iT \,. \tag{1.238}$$

Therefore, one has

$$1 = (1 + iT)(1 - iT^\dagger) = 1 + iT - iT^\dagger + TT^\dagger \,, \tag{1.239}$$

or equivalently

$$-i(T - T^\dagger) = TT^\dagger \,. \tag{1.240}$$

Let us now take the expectation value of this identity in the state $\left|\alpha_{\text{in}}\right\rangle$, and insert the identity operator written as a complete sum over physical states between T and $T^\dagger$ in the right hand side. This leads to:

$$-i\left\langle\alpha_{\text{in}}|T - T^\dagger|\alpha_{\text{in}}\right\rangle = \sum_{\text{states } \beta} \left|\left\langle\alpha_{\text{in}}|T|\beta_{\text{in}}\right\rangle\right|^2 \,. \tag{1.241}$$

Equivalently, this identity reads

$$\text{Im}\left\langle\alpha_{\text{in}}|T|\alpha_{\text{in}}\right\rangle = \frac{1}{2}\sum_{\text{states } \beta} \left|\left\langle\alpha_{\text{in}}|T|\beta_{\text{in}}\right\rangle\right|^2 \,. \tag{1.242}$$

This identity is known as the *optical theorem*. It implies that the total probability to evolve from the state α to any state β equals twice the imaginary part of the forward transition amplitude $\alpha \to \alpha$.

1.14.2 Cutkosky's cutting rules

Eq. (1.242) is valid to all orders in the interactions. But as we shall see it also manifests itself in some properties of the perturbative expansion. Let us first consider as an example a scalar field theory, with a cubic interaction in $-\frac{i}{3!}\lambda\phi(x)^3$.

Firstly, decompose the free Feynman propagator in two terms, depending on the ordering between the times at the two endpoints:

$$G_F^0(x,y) \equiv \theta(x^0 - y^0)G_{-+}^0(x,y) + \theta(y^0 - x^0)G_{+-}^0(x,y) \,. \tag{1.243}$$

The 2-point functions G_{-+}^0 and G_{+-}^0 are therefore defined as

$$G_{-+}^0(x,y) \equiv \left\langle 0_{\text{in}}\left|\phi_{\text{in}}(x)\phi_{\text{in}}(y)\right|0_{\text{in}}\right\rangle \,, \quad G_{+-}^0(x,y) \equiv \left\langle 0_{\text{in}}\left|\phi_{\text{in}}(y)\phi_{\text{in}}(x)\right|0_{\text{in}}\right\rangle \,. \tag{1.244}$$

In order to streamline the notations, it is convenient to rename G_F^0 by G_{++}^0, and to introduce another propagator with a reversed time ordering:

$$G_{--}^0(x,y) \equiv \theta(x^0 - y^0)G_{+-}^0(x,y) + \theta(y^0 - x^0)G_{-+}^0(x,y) \ . \tag{1.245}$$

The usual Feynman rules in coordinate space amount to connect a vertex at x and a vertex at y by the propagator $G_{++}^0(x,y)$. The coordinate x of each vertex is integrated out over all space-time, and a factor $-i\lambda$ is attached to each vertex. We will call $+$ this type of vertex. Thus, the Feynman rules for calculating transition amplitudes involve only the $+$ vertex and the G_{++}^0 propagator.

Let us then introduce a vertex of type $-$, to which a factor $+i\lambda$ is assigned (instead of $-i\lambda$ for the vertex of type $+$). The integrand of a Feynman graph $\mathcal{G}$ is a function $\mathcal{G}(x_1, x_2, \cdots)$ of the coordinates x_i of its vertices. We will generalize this function by assigning $+$ or $-$ indices to all the vertices,

$$\mathcal{G}(x_1, x_2, \cdots) \quad \rightarrow \quad \mathcal{G}_{\epsilon_1 \epsilon_2 \ldots}(x_1, x_2, \cdots) \ , \tag{1.246}$$

where the indices $\epsilon_i = \pm$ indicate which is the type of the i-th vertex. The usual Feynman rules thus correspond to the function $\mathcal{G}_{++\ldots}$. These generalized integrands are constructed according to the following rules:

$$
\begin{array}{rcl}
+ \text{ vertex :} & -i\lambda & , \\
- \text{ vertex :} & +i\lambda & , \\
\text{Propagator from } \epsilon \text{ to } \epsilon' : & G_{\epsilon \epsilon'}^0(x,y) & .
\end{array}
\tag{1.247}
$$

Let us assume that the i-th vertex carries the largest time among all the vertices of the graph. Since x_i^0 is largest than all the other times, then the propagator that connects this vertex to an adjacent vertex of type ϵ at the position x is given by

$$G_{\pm \epsilon}^0(x_i, x) = G_{-\epsilon \epsilon}^0(x_i, x) \ . \tag{1.248}$$

In other words, this propagator depends only on the type ϵ of the neighboring vertex, but not on the type of the i-th vertex. Therefore, we have

$$\mathcal{G}_{\ldots[+_i]\ldots}(x_1, x_2, \cdots) + \mathcal{G}_{\ldots[-_i]\ldots}(x_1, x_2, \cdots) = 0 \ , \tag{1.249}$$

where the notation $[\pm_i]$ indicates that the i-th vertex has type $+$ or $-$ (the types of the vertices not written explicitly are the same in the two terms, but otherwise arbitrary). This identity, known as the *largest time equation*, follows from eq. (1.248) and from the sign change when a vertex changes from $+$ to $-$.

A similar identity also applies to the sum extended to all the possible assignments of the $+$ and $-$ indices:

$$\sum_{\{\epsilon_i = \pm\}} \mathcal{G}_{\epsilon_1 \epsilon_2 \ldots}(x_1, x_2, \cdots) = 0 \ . \tag{1.250}$$

This is obtained by pairing the terms and using eq. (1.249). It is crucial to observe that this identity is now valid for any ordering of the times at the vertices of the graph. Therefore, it is also valid in momentum space after a Fourier transform. If we isolate the two terms where all the vertices are of type $+$ or all of type $-$, this also reads

$$\mathcal{G}_{++\ldots} + \mathcal{G}_{--\ldots} = - \sum_{\{\epsilon_i = \pm\}'} \mathcal{G}_{\epsilon_1 \epsilon_2 \ldots} \ , \tag{1.251}$$

where the symbol $\{\epsilon_i = \pm\}'$ indicates the set of all the vertex assignments, except $++\cdots$ and $--\cdots$.

Using eq. (1.82),

$$G^0_{++}(x,y) = \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_\mathbf{p}} \left\{ \theta(x^0-y^0)\, e^{-ip\cdot(x-y)} + \theta(y^0-x^0)\, e^{+ip\cdot(x-y)} \right\}, \quad (1.252)$$

and comparing with eq. (1.245), we can read off the following representations for G^0_{-+} and G^0_{+-},

$$\begin{aligned}
G^0_{-+}(x,y) &= \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_\mathbf{p}}\, e^{-ip\cdot(x-y)} \\
G^0_{+-}(x,y) &= \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_\mathbf{p}}\, e^{+ip\cdot(x-y)}.
\end{aligned} \quad (1.253)$$

Likewise, we obtain

$$G^0_{--}(x,y) = \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_\mathbf{p}} \left\{ \theta(x^0-y^0)\, e^{+ip\cdot(x-y)} + \theta(y^0-x^0)\, e^{-ip\cdot(x-y)} \right\}, \quad (1.254)$$

Note that $G^0_{++} + G^0_{--} = G^0_{-+} + G^0_{+-}$.

Using the following representation for step functions:

$$\begin{aligned}
\theta(x^0-y^0) &= \int \frac{dp_0}{2\pi} \frac{-i}{p_0 - i0^+} e^{ip_0(x^0-y^0)}, \\
\theta(y^0-x^0) &= \int \frac{dp_0}{2\pi} \frac{i}{p_0 + i0^+} e^{ip_0(x^0-y^0)},
\end{aligned} \quad (1.255)$$

we can derive the momentum space expressions of these propagators:

$$\begin{aligned}
G^0_{++}(p) &= \frac{i}{p^2 - m^2 + i0^+}, \\
G^0_{--}(p) &= \frac{-i}{p^2 - m^2 - i0^+} = \left[G^0_{++}(p) \right]^*, \\
G^0_{-+}(p) &= 2\pi\, \theta(+p_0)\delta(p^2 - m^2), \\
G_{+-}(p) &= 2\pi\, \theta(-p_0)\delta(p^2 - m^2).
\end{aligned} \quad (1.256)$$

Therefore, the momentum space Feynman rules for the $-$ sector are the complex conjugate of those for the $+$ sector, since we have also $+i\lambda = (-i\lambda)^*$. Note that for this assertion to be true, it is crucial that the coupling constant λ be real, which is a condition for unitarity.

The Fourier transform of an amputated Feynman graph $\mathcal{G}$ gives a contribution to a transition amplitude (recall the LSZ reduction formula), i.e. a matrix element of the S operator. Therefore, $\Gamma \equiv i\mathcal{G}$ gives a matrix element of the T operator. Therefore, after Fourier transform, eq. (1.251) becomes

$$\mathrm{Im}\, \Gamma_{++\cdots} = \frac{1}{2} \sum_{\{\epsilon_i = \pm\}'} \left[i\Gamma \right]_{\epsilon_1 \epsilon_2 \cdots}. \quad (1.257)$$

45

If the graph contains N vertices, there are a priori $2^N - 2$ terms in the right hand side of this equation. However, this number is considerably reduced if we notice that the $+-$ and $-+$ propagators can carry energy only in one direction (from the $-$ vertex to the $+$ vertex), because of the factors $\theta(\pm p^0)$. This constraint on energy flow forbids "islands" of vertices of type $+$ surrounded by only type $-$ vertices, or the reverse. From the LSZ reduction formula (1.62) and the definition (1.83) of the Fourier transformed propagators, we see that the notation $G_{-+}(p)$ implies a momentum p defined as flowing from the $+$ endpoint to the $-$ endpoint:

$$G_{-+}(p) = \underset{+ \qquad -}{\xrightarrow{\hspace{1cm} p \hspace{1cm}}} \; . \tag{1.258}$$

Thus, the proportionality $G_{-+}(p) \propto \theta(p^0)$ indicates that the energy flows from the $+$ endpoint to the $-$ endpoint.

Let us consider the example of a very simple 1-loop two-point function[31] $\Gamma(p)$,

$$-i\Gamma(p) = \quad \xrightarrow{\; p \;} \bigcirc \!\!\!\!\text{---} \quad . \tag{1.259}$$

Because of the constrained energy flow direction in the propagators G_{-+}, G_{+-}, if the momentum p is entering into the graph from the left with $p^0 > 0$, the only assignments that mix $+$ and $-$ vertices must divide the graph into two connected subgraphs: a connected part made only of $+$ vertices that comprises the vertex where $p^0 > 0$ enters in the graph, and a connected part containing only $-$ vertices comprising the vertex where the energy leaves the graph. For the topology shown in eq. (1.259), there is only one possibility,

$$-i\Gamma_{+-}(p) = \quad \xrightarrow{\; p \;} \bigcirc\!\!\!|\!\!\!\oplus\!\!\!\text{---} \quad , \tag{1.260}$$

where the vertex of type $-$ is circled in the diagrammatic representation. The division of the graph into these two subgraphs may be materialized by drawing a line (shown in gray above) through the graph. This line is called a *cut*, and the rules for calculating the value of a graph with a given assignment of $+$ and $-$ vertices are called *Cutkosky's cutting rules*. For instance, in the case of the above example, they lead immediately to the following expression[32] for the imaginary part of Γ_{++},

$$\text{Im}\,\Gamma_{++}(p) = \frac{\lambda^2}{2}\frac{1}{2}\int \frac{d^4k}{(2\pi)^4}\, G_{-+}(k)G_{-+}(p-k) \,, \tag{1.261}$$

that can be rewritten as

$$\begin{aligned} \text{Im}\,\Gamma_{++}(p) \;=\; & \frac{\lambda^2}{4}\int \frac{d^4k_1}{(2\pi)^4} 2\pi\theta(k_1^0)\delta(k_1^2 - m^2) \\ & \times \int \frac{d^4k_2}{(2\pi)^4} 2\pi\theta(k_2^0)\delta(k_2^2 - m^2)(2\pi)^4\delta(p - k_1 - k_2) \,. \end{aligned} \tag{1.262}$$

---

[31]Momentum conservation implies that it depends on a single momentum p.

[32]The first factor $1/2$ comes from eq. (1.257), and the second $1/2$ is the symmetry factor of the graph for a scalar loop. In the formula for $\text{Im}\,\Gamma_{++}$, it has the interpretation of the factor that symmetrizes a 2-particle final state.

In the right hand side of this equation, we recognize the square of the transition amplitude $\langle k_1 k_{2\,\text{out}} | p_{\text{in}} \rangle$ (whose value at tree level is simply λ), integrated over the (symmetrized) accessible phase-space for a 2-particle final state. We can therefore view this equation as a perturbative realization of the optical theorem at order λ^2. Indeed, at this order, the only states β that may be included in the sum over final states are 2-particle states.

The considerations developed on this example can be generalized to the 2-point function at any loop order. We can write

$$\text{Im}\,\Gamma_{++}(p) = \frac{1}{2} \sum_{\text{cuts}\,\gamma} (i\Gamma_\gamma(p)) , \tag{1.263}$$

where the sum is now limited reduced to a sum over all the possible cuts (with the $+$ vertices left of the cut and the $-$ vertices right of the cut). As an illustration, let us consider the following 2-loop example, for which three cuts are possible:

$$\text{Im}\,\Gamma_{++}(p) = \quad \text{—} \!\!\!\bigcirc\!\!\!\!\text{—} \quad . \tag{1.264}$$

At this order start to appear various contributions to the right hand side of eq. (1.242): the central cut corresponds to a 3-body final state, while the other two cuts correspond to an interference between the tree level and the 1-loop correction to a 2-body decay.

1.14.3 Fermions

In the case of spin $1/2$ fermions, the propagators connecting the various types of vertices are given by

$$S^0_{++}(p) = \frac{i(\not{p} + m)}{p^2 - m^2 + i0^+} ,$$
$$S^0_{--}(p) = \frac{-i(\not{p} + m)}{p^2 - m^2 - i0^+} ,$$
$$S^0_{-+}(p) = 2\pi(\not{p} + m)\theta(-p_0)\delta(p^2 - m^2) ,$$
$$S^0_{+-}(p) = 2\pi(\not{p} + m)\theta(+p_0)\delta(p^2 - m^2) . \tag{1.265}$$

The cutting rules for fermions are therefore similar to those for scalar particles. The possibility to interpret the cut fermion propagators in terms of on-shell final state fermions is a consequence of the following identities:

$$\not{p} + m = \sum_{\text{spin}\,s} u_s(\mathbf{p})\overline{u}_s(\mathbf{p}) ,$$
$$\not{p} - m = \sum_{\text{spin}\,s} v_s(\mathbf{p})\overline{v}_s(\mathbf{p}) , \tag{1.266}$$

that are valid when $p_0 = \sqrt{\mathbf{p}^2 + m^2} > 0$. In the case of the propagator $S^0_{-+}(p)$, we may attach the spinor $u_s(\mathbf{p})$ to the amplitude on the right of the cut, and the spinor $\overline{u}_s(\mathbf{p})$ to

the amplitude on the left, which are precisely the spinors required by the LSZ formula for a fermion of momentum $\mathbf{p}$ in the final state. In the case of $S^0_{+-}(p)$, for which $p^0 < 0$, we should first write

$$
\begin{aligned}
S^0_{+-}(p) &= -2\pi(-\not{p} - m)\theta(-p_0)\delta(p^2 - m^2) \\
&= -2\pi \sum_{\text{spin } s} v_s(-\mathbf{p})\bar{v}_s(-\mathbf{p})\theta(-p_0)\delta(p^2 - m^2) ,
\end{aligned}
\tag{1.267}
$$

in order to see that it corresponds to an anti-fermion in the final state.

1.14.4 Photons

Coulomb gauge : For photons in Coulomb gauge, the reasoning is very similar to the case of fermions. Firstly, the four different types of propagators read

$$
\begin{aligned}
G^{0\,ij}_{++}(p) &= \frac{i\left(\delta^{ij} - \frac{p^i p^j}{\mathbf{p}^2}\right)}{p^2 + i0^+} , \\
G^{0\,ij}_{--}(p) &= \frac{-i\left(\delta^{ij} - \frac{p^i p^j}{\mathbf{p}^2}\right)}{p^2 - i0^+} , \\
G^{0\,ij}_{-+}(p) &= 2\pi\,\theta(+p^0)\left(\delta^{ij} - \frac{p^i p^j}{\mathbf{p}^2}\right)\delta(p^2) , \\
G^{0\,ij}_{+-}(p) &= 2\pi\,\theta(-p^0)\left(\delta^{ij} - \frac{p^i p^j}{\mathbf{p}^2}\right)\delta(p^2) .
\end{aligned}
\tag{1.268}
$$

Recalling also that

$$
\sum_{\lambda=\pm} \epsilon^{i*}_\lambda(\mathbf{p})\epsilon^j_\lambda(\mathbf{p}) = \delta^{ij} - \frac{p^i p^j}{\mathbf{p}^2} ,
\tag{1.269}
$$

we see that the projector that appears in the cut propagators can be interpreted as the polarization vectors that should attached to amplitudes for each final state photon. Therefore, the cutting rules in Coulomb gauge have a direct interpretation in terms of the optical theorem. This simplicity follows from the fact that the only propagating modes are physical modes in Coulomb gauge.

Feynman gauge : This interpretation is not so direct in covariant gauges, such as the Feynman gauge. In this gauge, the free photon propagator is given by:

$$
G^{0\,\mu\nu}_{++}(p) = -g^{\mu\nu}\frac{i}{p^2 + i0^+} .
\tag{1.270}
$$

The factor $-g^{\mu\nu}$ does not change anything to the cutting rules, and simply appears as a prefactor in all propagators:

$$
\begin{aligned}
G^{0\,\mu\nu}_{--}(p) &= -g^{\mu\nu}\frac{-i}{p^2 - i0^+} , \\
G^{0\,\mu\nu}_{-+}(p) &= -2\pi\,g^{\mu\nu}\theta(+p_0)\delta(p^2) , \\
G^{0\,\mu\nu}_{+-}(p) &= -2\pi\,g^{\mu\nu}\theta(-p_0)\delta(p^2) .
\end{aligned}
\tag{1.271}
$$

Let us assume for definiteness that the photon momentum $\mathbf{p}$ is in the $\hat{z}$ direction, i.e. $\mathbf{p} = (0, 0, p)$. Therefore, the two physical polarizations vectors, orthogonal to $\mathbf{p}$, can be chosen as follows

$$
\begin{aligned}
\epsilon_1^\mu(\mathbf{p}) &\equiv (0, 1, 0, 0) \,, \\
\epsilon_2^\mu(\mathbf{p}) &\equiv (0, 0, 1, 0) \,.
\end{aligned}
\tag{1.272}
$$

They are orthonormal

$$
\epsilon_\lambda(\mathbf{p}) \cdot \epsilon_{\lambda'}(\mathbf{p}) = -\delta_{\lambda\lambda'} \,,
\tag{1.273}
$$

and transverse: $p_\mu \epsilon_{1,2}^\mu(\mathbf{p}) = 0$. However, the tensor $-g^{\mu\nu}$ that appears in the cut photon propagators cannot be written as a sum over physical polarizations:

$$
-g^{\mu\nu} \neq \sum_{\lambda=1,2} \epsilon_\lambda^\mu(\mathbf{p}) \epsilon_\lambda^\nu(\mathbf{p})^* \,.
\tag{1.274}
$$

Only the $\mu = 1, 2$ components of these tensors are equal. As a consequence, it seems that Cutkosky's cutting rules may not lead to terms that we cannot interpret as physical final photon states, which would violate the optical theorem. If this was the case, then perturbation theory would not be consistent with unitarity. To see how this paradox is resolved, let us introduce two more (unphysical) polarization vectors [33]:

$$
\begin{aligned}
\epsilon_+^\mu(\mathbf{p}) &\equiv \frac{1}{\sqrt{2}}(1, 0, 0, 1) \,, \\
\epsilon_-^\mu(\mathbf{p}) &\equiv \frac{1}{\sqrt{2}}(1, 0, 0, -1) \,,
\end{aligned}
\tag{1.276}
$$

thanks to which we may now write

$$
g^{\mu\nu} = \epsilon_+^\mu(\mathbf{p}) \epsilon_-^\nu(\mathbf{p})^* + \epsilon_-^\mu(\mathbf{p}) \epsilon_+^\nu(\mathbf{p})^* - \sum_{\lambda=1,2} \epsilon_\lambda^\mu(\mathbf{p}) \epsilon_\lambda^\nu(\mathbf{p})^* \,.
\tag{1.277}
$$

In other words, the physical polarization sum in the right hand side of eq. (1.274) is equal to $-g^{\mu\nu}$, plus some extra terms that are proportional to p^μ of p^ν.

When we use Cutkosky's cutting rules in order to calculate the imaginary part of graph, a cut photon line carrying the momentum p^μ leads to an expression that has the following structure:

$$
iM_1^\mu(p) \left[-g^{\mu\nu}\right] (iM_2^\nu(p))^* \,,
\tag{1.278}
$$

where iM_1^μ and iM_2^ν are the amplitudes on the left and on the right of the cut, respectively. Here, we have highlighted only one of the cut photons, and the other cut lines have not been written explicitly since they do not play any role in the argument. Moreover, only the tensor

---

[33]For an arbitrary momentum $\mathbf{p}$, these polarization vectors read:

$$
\begin{aligned}
\epsilon_+^\mu(\mathbf{p}) &\equiv \frac{1}{\sqrt{2}|\mathbf{p}|}(p_0, \mathbf{p}) \,, \\
\epsilon_-^\mu(\mathbf{p}) &\equiv \frac{1}{\sqrt{2}|\mathbf{p}|}(p_0, -\mathbf{p}) \,.
\end{aligned}
\tag{1.275}
$$

structure of the cut propagator matters, and we have therefore only written the factor $-g^{\mu\nu}$. The above quantity can be rewritten as

$$
i M_1^\mu(p) \left[\sum_{\lambda=1,2} \epsilon_\lambda^\mu(\mathbf{p}) \epsilon_\lambda^\nu(\mathbf{p})^* - \epsilon_+^\mu(\mathbf{p}) \epsilon_-^\nu(\mathbf{p})^* - \epsilon_-^\mu(\mathbf{p}) \epsilon_+^\nu(\mathbf{p})^* \right] (i M_2^\nu(p))^*
$$

$$
= i M_1^\mu(p) \left[\sum_{\lambda=1,2} \epsilon_\lambda^\mu(\mathbf{p}) \epsilon_\lambda^\nu(\mathbf{p})^* \right] (i M_2^\nu(p))^* . \tag{1.279}
$$

Indeed, the last two terms are zero thanks to the Ward identity satisfied[34] by the amplitudes $i M_1^\mu$ and $i M_2^\nu$:

$$
p_\mu M_1^\mu(p) = p_\nu M_2^\nu(p) = 0 , \tag{1.280}
$$

and because $\epsilon_+^\mu(\mathbf{p})$ is proportional to p^μ. Therefore, the non-physical photon degrees of freedom, that may appear in the cutting rules in covariant gauges, are in fact canceled by the Ward identities satisfied by QED amplitudes.

1.14.5 Schwinger-Keldysh formalism

Perturbation theory provides a way of computing transition amplitudes like $\langle \mathbf{p}' \mathbf{q}'_{\text{out}} | \mathbf{p} \mathbf{q}_{\text{in}} \rangle$. The calculation of these matrix elements is amenable via the LSZ reduction formulas to the expectation value of time-ordered products of field operators, between the in- and out-vacuum states, for instance $\langle 0_{\text{out}} | T \phi(x_1) \phi(x_2) \phi(x_3) \phi(x_4) | 0_{\text{in}} \rangle$, the calculation of which can be performed with the usual Feynman rules.

However, there is a class of more general problems that cannot be addressed by this standard perturbation theory. One of the simplest problems of that kind is the evaluation of the expectation value of the number operator $\langle \alpha_{\text{in}} | a_{\text{out}}^\dagger(\mathbf{p}) a_{\text{out}}(\mathbf{p}) | \alpha_{\text{in}} \rangle$, that counts the particles of momentum $\mathbf{p}$ in the final state, given that the initial state was the state α. To evaluate this matrix element, one needs to calculate the amplitude $\langle \alpha_{\text{in}} | \phi(x) \phi(y) | \alpha_{\text{in}} \rangle$, that has no time ordering, and where one has *in* states on both sides. More generally, one sometimes needs the amplitudes $\langle 0_{\text{in}} | \overline{T} (\phi(x_1) \cdots \phi(x_n)) \, T (\phi(y_1) \cdots \phi(y_p)) | 0_{\text{in}} \rangle$, where $\overline{T}$ denotes the anti-time ordering. The Schwinger-Keldysh formalism is tailored for addressing these more general questions. Moreover, as we shall see, it is formally identical to Cutkosky's cutting rules.

Schwinger-Keldysh perturbation theory :   Consider the expectation value

$$
\langle 0_{\text{in}} | \overline{T} (\phi(x_1) \cdots \phi(x_n)) \, T (\phi(y_1) \cdots \phi(y_p)) | 0_{\text{in}} \rangle . \tag{1.281}
$$

As we did in the derivation of ordinary perturbation theory, let us first replace each Heisenberg field operator by its counterpart in the interaction representation, using eq. (1.43). After some

---

[34]When an amplitude has external charged particles, the Ward identity is satisfied only if these particles are on-shell. This is indeed the case here, because all the cut lines are on-shell, as well as all the incoming particles.

rearrangement of the evolution operators, we get :

$$\langle 0_{in}|\overline{T}\,\phi(x_1)\cdots\phi(x_n)\,T\phi(y_1)\cdots\phi(y_p)|0_{in}\rangle =$$

$$= \langle 0_{in}|\overline{T}\left[\phi_{in}(x_1)\cdots\phi_{in}(x_n)\exp i\int_{-\infty}^{+\infty}d^4x\,\mathcal{L}_I(\phi_{in}(x))\right]$$

$$\times T\left[\phi_{in}(y_1)\cdots\phi_{in}(y_p)\exp i\int_{-\infty}^{+\infty}d^4x\,\mathcal{L}_I(\phi_{in}(x))\right]|0_{in}\rangle .$$

$$(1.282)$$

Here, we have exploited the fact that the factor $U(-\infty, +\infty)$ that appears in these manipulations is the anti-time ordered exponential of the interaction term, in order to write this formula in a more symmetric way. To go further, it is useful to imagine that the time axis is in fact a contour $\mathcal{C}$ made of two branches labeled $+$ and $-$ running parallel to the real axis, as illustrated in figure 1.1. This contour is oriented, with the $+$ branch running in the direction

Figure 1.1: *Time contour in the Schwinger-Keldysh formalism.*
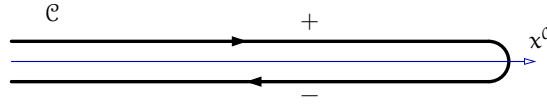
of increasing time, followed by the $-$ branch running in the direction of decreasing time. Then, it is convenient to introduce a *path ordering*, denoted by P and defined as a standard ordering along the contour $\mathcal{C}$. In more detail, one has

$$P\,A(x)B(y) = \begin{cases} T\,A(x)B(y) & \text{if} & x^0, y^0 \in \mathcal{C}_+ , \\ \overline{T}\,A(x)B(y) & \text{if} & x^0, y^0 \in \mathcal{C}_- , \\ A(x)B(y) & \text{if} & x^0 \in \mathcal{C}_- , \; y^0 \in \mathcal{C}_+ , \\ B(y)A(x) & \text{if} & x^0 \in \mathcal{C}_+ , \; y^0 \in \mathcal{C}_- . \end{cases}$$

$$(1.283)$$

One can use this contour ordering to write the previous equations in a much more compact way. In particular, eq. (1.282) can be generalized into :

$$\langle 0_{in}|P\,\phi^-(x_1)\cdots\phi^-(x_n)\phi^+(y_1)\cdots\phi^+(y_p)|0_{in}\rangle =$$

$$= \langle 0_{in}|P\,\phi_{in}^-(x_1)\cdots\phi_{in}^-(x_n)\phi_{in}^+(y_1)\cdots\phi_{in}^+(y_p)\exp i\int_{\mathcal{C}}d^4x\,\mathcal{L}_I(\phi_{in}(x))|0_{in}\rangle .$$

$$(1.284)$$

The differences compared to eq. (1.282) are threefold :

i. A single overall path ordering takes care automatically of both the time ordering and the anti-time ordering contained in the original formula,

ii. For this trick to work, one must (temporarily) assume that the fields on the upper and lower branch of the contour $\mathcal{C}$ are distinct: ϕ^+ and ϕ^- respectively,

iii. The time integration in the exponential is now running over both branches of the contour $\mathcal{C}$.

The advantage of having introduced this more complicated time contour is that it leads to a expressions that are formally identical to those of ordinary perturbation theory, provided one replaces the time ordering by the path ordering and provided one extends the time integration from $\mathbb{R}$ to $\mathcal{C}$. In particular, one can first define a generating functional,

$$Z^{\text{SK}}[j] \equiv \langle 0_{\text{in}} | \text{T} \exp i \int_{\mathcal{C}} d^4x \, j(x)\phi(x) \, | 0_{\text{in}} \rangle , \tag{1.285}$$

that encodes all the correlators considered in this section, provided the external source j has distinct values j_+ and j_- on the two branches of the contour (the superscript SK is used to distinguish this generating functional from the standard one). As in the case of Feynman perturbation theory, one can write this generating functional as:

$$Z^{\text{SK}}[j] = \exp i \int_{\mathcal{C}} d^4x \, \mathcal{L}_{\text{I}} \left(\frac{\delta}{i\delta j(x)} \right) \underbrace{\langle 0_{\text{in}} | \text{T} \exp i \int_{\mathcal{C}} d^4x \, j(x)\phi_{\text{in}}(x) | 0_{\text{in}} \rangle}_{Z_0^{\text{SK}}[j]} , \tag{1.286}$$

with

$$
\begin{aligned}
Z_0^{\text{SK}}[j] &= \exp \left\{ -\frac{1}{2} \int_{\mathcal{C}} d^4x d^4y \, j(x)j(y) \, G_{\mathcal{C}}^0(x,y) \right\} \\
G_{\mathcal{C}}^0(x,y) &\equiv \langle 0_{\text{in}} | \text{P} \, \phi_{\text{in}}(x)\phi_{\text{in}}(y) | 0_{\text{in}} \rangle .
\end{aligned}
\tag{1.287}
$$

The free propagator $G_{\mathcal{C}}^0$, defined on the contour $\mathcal{C}$, is a natural extension of the Feynman propagator (in particular, it coincides with the Feynman propagator if the two time arguments are on the $+$ branch of the contour). Besides the propagator, the other change to the perturbative expansion in the Schwinger-Keldysh formalism is that the time integration at the vertices of a diagram must run over the contour $\mathcal{C}$ instead of the real axis.

The connection with Cutkosky's cutting rules appears when we break down the propagator into 4 components $G_{\pm\pm}^0(x,y)$, depending on whether the times x^0, y^0 are on the upper or lower branch of the contour. An explicit calculation of these free propagators leads to

$$
\begin{aligned}
G_{++}^0(x,y) &= i \int \frac{d^4p}{(2\pi)^4} \frac{e^{-ip\cdot(x-y)}}{p^2 - m^2 + i\epsilon} , \\
G_{--}^0(x,y) &= -i \int \frac{d^4p}{(2\pi)^4} \frac{e^{-ip\cdot(x-y)}}{p^2 - m^2 - i\epsilon} , \\
G_{+-}^0(x,y) &= \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} 2\pi\theta(-p^0)\delta(p^2 - m^2) , \\
G_{-+}^0(x,y) &= \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} 2\pi\theta(+p^0)\delta(p^2 - m^2) .
\end{aligned}
\tag{1.288}
$$

The time integration on the contour $\mathcal{C}$ is also split into two terms, the upper branch corresponding to a vertex $+ (-i\lambda)$ and the lower branch to a vertex $- (+i\lambda$, because of the minus sign due to integrating from $+\infty$ to $-\infty$).

52

In the Schwinger-Keldysh formalism, the vacuum-vacuum diagrams are simpler than in conventional perturbation theory. Here, one has

$$Z^{\text{SK}}[0] = \langle 0_{\text{in}} | 0_{\text{in}} \rangle = 1 \ , \tag{1.289}$$

which means that all the connected vacuum-vacuum diagrams are zero. This is due to the fact that in this formalism one is calculating correlators that have the in- vacuum on both sides. This cancellation works individually for each diagram topology, and results from a cancellation between the various ways of assigning the $+$ and $-$ indices to the vertices of a diagram (a vacuum-vacuum diagram with a fixed assignment of $+$ and $-$ vertices is not zero in general). This cancellation can be viewed as a consequence of eq. (1.250).

Relation between the functionals $Z[j]$ **and** $Z^{\text{SK}}[j]$ **:**   There is a useful functional relation between the generating functional of conventional perturbation theory $Z[j]$, and that of the Schwinger-Keldysh formalism :

$$Z^{\text{SK}}[j_+, j_-] = \exp\left[\iint d^4x d^4y \ G^0_{+-}(x,y) \, \square_x \square_y \, \frac{\delta^2}{\delta j_+(x) \delta j_-(y)} \right] Z[j_+] \, Z^*[j_-] \ . \tag{1.290}$$

(Here, in order to avoid any confusion, we write explicitly the two components $+$ and $-$ of the source j in the Schwinger-Keldysh generating functional.) Thanks to this formula, one can construct diagrams in the Schwinger-Keldysh formalism by *stitching* an ordinary Feynman diagram and the complex conjugate of another Feynman diagram. In order to prove this relation, it is sufficient to establish it for the free theory, since the interactions are always trivially factorizable (see eqs. (1.70) and (1.286)).

Chapter 2

Functional quantization

2.1 Path integral in quantum mechanics

Let us consider a quantum mechanical system with a single degree of freedom, whose Hamiltonian is

$$\mathcal{H} \equiv \frac{P^2}{2m} + V(Q) \,. \tag{2.1}$$

The position and momentum operators Q and P obey the following commutation relation $\big[Q, P\big] = i$. We would like to calculate the probability for the system to start at the position q_i at a time t_i and end at the position q_f at the time t_f. The answer may be obtained as $\big|\psi(q_f, t_f)\big|^2$ by solving Schrödinger's equation with an initial wavefunction localized at q_i,

$$i\partial_t \psi(q, t) = \mathcal{H}\,\psi(q, t) \quad, \qquad \psi(q, t_i) \equiv \delta(q - q_i) \,. \tag{2.2}$$

More formally, in the Schrödinger picture, it is given by the squared modulus of the following transition amplitude

$$\big\langle q_f \big| e^{-i\mathcal{H}(t_f - t_i)} \big| q_i \big\rangle \,, \tag{2.3}$$

where $\big| q \big\rangle$ denote the eigenstate of the position operator with eigenvalue q. Let us subdivide the time interval $[t_i, t_f]$ into N equal sub-intervals, by introducing:

$$\Delta \equiv \frac{t_f - t_i}{N} \quad, \qquad t_n \equiv t_i + n\,\Delta \,. \tag{2.4}$$

(Therefore, we have $t_0 = t_i$ and $t_{_N} = t_f$.) The time evolution operator can be factorized as

$$e^{-i\mathcal{H}(t_f - t_i)} = e^{-i\mathcal{H}(t_{_N} - t_{_{N-1}})} \times e^{-i\mathcal{H}(t_{_{N-1}} - t_{_{N-2}})} \times \cdots \times e^{-i\mathcal{H}(t_1 - t_0)} \,. \tag{2.5}$$

Between the successive factors in the right hand side, we can insert the identity operator written as a complete sum over the position eigenstates:

$$1 = \int_{-\infty}^{+\infty} dq \, \big| q \big\rangle\big\langle q \big| \,, \tag{2.6}$$

and the transition amplitude (2.3) becomes

$$\langle q_f | e^{-i\mathcal{H}(t_f - t_i)} | q_i \rangle \;=\; \int \prod_{j=1}^{N-1} dq_j \, \langle q_f | e^{-i\Delta\mathcal{H}} | q_{N-1} \rangle \langle q_{N-1} | e^{-i\Delta\mathcal{H}} | q_{N-2} \rangle \cdots$$

$$\cdots \langle q_1 | e^{-i\Delta\mathcal{H}} | q_i \rangle \,. \tag{2.7}$$

Note that this formula, illustrated in the figure 2.1, is exact for any value of N. In the Hamilto-

Figure 2.1: *Illustration of eq. (2.7) with 10 and 200 intermediate points. The endpoints (green) are fixed, while the intermediate points (red) are integrated over. The line segments connecting the points are just a help to guide the eye, but there is no "path" at this stage.*
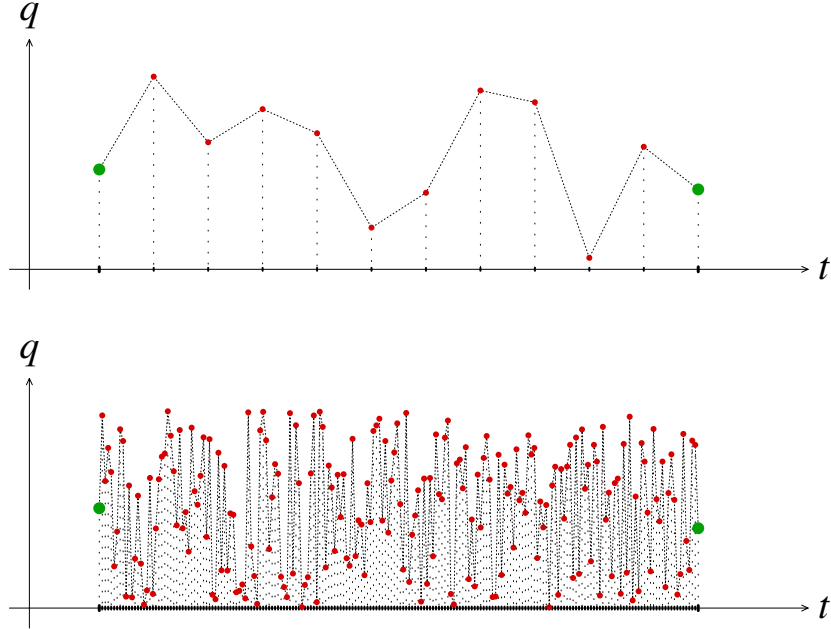
nian (2.1), the kinetic energy and potential energy terms do not commute, which complicates the evaluation of its exponential. We can remedy this situation by using the *Baker-Campbell-Hausdorff formula*, that we shall write here as follows

$$e^{\Delta(A+B)} = e^{\Delta A} e^{\Delta B} \, e^{-\frac{\Delta^2}{2}[A,B] + \mathcal{O}(\Delta^3)} \,. \tag{2.8}$$

In the limit $\Delta \to 0$ (i.e. $N \to \infty$), we may neglect the last factor since the product of all such factors goes to unity[1] when $N \to \infty$. Therefore, each elementary factor of eq. (2.7) is

---

[1] We use

$$\lim_{N\to\infty} e^{\alpha_1/N^2} e^{\alpha_2/N^2} \cdots e^{\alpha_N/N^2} = 1 \,,$$

provided that the sum $\sum_i \alpha_i$'s does not diverge too quickly.

rewritten as

$$\langle q_{i+1}|e^{-i\Delta\mathcal{H}}|q_i\rangle \approx \langle q_{i+1}|e^{-i\Delta\frac{P^2}{2m}}\,e^{-i\Delta V(Q)}|q_i\rangle$$

$$= \int \frac{dp_i}{2\pi}\,\langle q_{i+1}|e^{-i\Delta\frac{P^2}{2m}}|p_i\rangle\,\langle p_i|e^{-i\Delta V(Q)}|q_i\rangle\,, \qquad (2.9)$$

where we have introduced the identity operator, written this time as a complete sum over momentum eigenstates:

$$\mathbb{1} \equiv \int \frac{dp}{2\pi}\,|p\rangle\langle p|\,. \qquad (2.10)$$

In the two factors, the exponential operator depends only on P or Q, and the matrix elements are trivial to evaluate by using the fact that the operators are enclosed between momentum and position eigenstates:

$$\langle q_{i+1}|e^{-i\Delta\frac{P^2}{2m}}|p_i\rangle = e^{-i\Delta\frac{p_i^2}{2m}}\,\langle q_{i+1}|p_i\rangle\,,$$
$$\langle p_i|e^{-i\Delta V(Q)}|q_i\rangle = e^{-i\Delta V(q_i)}\,\langle p_i|q_i\rangle\,. \qquad (2.11)$$

Using now

$$\langle q|p\rangle = e^{ipq}\,, \qquad (2.12)$$

we arrive at the formula[2]

$$\langle q_{i+1}|e^{-i\Delta\mathcal{H}}|q_i\rangle = e^{-i\Delta\mathcal{H}(p_i,q_i)}\,e^{i\,p_i(q_{i+1}-q_i)}\left(1+\mathcal{O}(\Delta^2)\right)\,. \qquad (2.13)$$

If we define $\dot{q}_i \equiv (q_{i+1}-q_i)/\Delta$ the slope of the line segments in the figure 2.1, and we take the limit $N\to\infty$, we may write the transition amplitude as a *path integral*:

$$\langle q_f|e^{-i\mathcal{H}(t_f-t_i)}|q_i\rangle = \int\limits_{\substack{q(t_i)=q_i\\q(t_f)=q_f}} \left[Dp(t)Dq(t)\right]\,\exp\left\{i\int_{t_i}^{t_f}dt\,\left(p(t)\dot{q}(t)-\mathcal{H}(p(t),q(t))\right)\right\}\,.$$

$$(2.14)$$

One should be aware of the fact that the functional measure $\left[Dq(t)Dp(t)\right]$ in general lacks solid mathematical foundations, although it allows for some powerful manipulations that would be extremely cumbersome to perform at the level of quantum operators. Note that at the boundaries $t_{i,f}$ the position is well defined, and therefore the momentum is not constrained (by the uncertainty principle). A crucial aspect of eq. (2.14) is that all the objects that appear in the right hand side are ordinary c-numbers that commute, while the left hand side is made of quantum operators and states. In this section, we have started from the conventional formulation of transition amplitudes in quantum mechanics, in order to arrive at the formula (2.14). However, one may now "forget" the canonical formalism and view the path integral expression of transition amplitudes as another way of going from a classical Hamiltonian $\mathcal{H}$ to a quantized theory.

---

[2]A bit more care is necessary for Hamiltonians that are not separable into a sum of a P-dependent term and a Q-dependent term. A proper treatment should use Weyl's prescription for defining the quantum Hamiltonian operator from the classical Hamiltonian. In eq. (2.13), one would obtain $\mathcal{H}(p_i,\frac{1}{2}(q_i+q_{i+1}))$ instead of $\mathcal{H}(p_i,q_i)$.

For a Hamiltonian where the P dependence has no powers higher than quadratic, as in the example of eq. (2.1), it is possible to perform exactly the integral over $p(t)$. This type of integral is called a *Gaussian path integral*. Gaussian path integrals can be evaluated in the same way as their ordinary counterparts, using the following formulas,

$$\int_{-\infty}^{+\infty} dx\ e^{-x^2/(2\sigma)} = \sqrt{2\pi\sigma}\ , \quad \int_{-\infty}^{+\infty} dx\ e^{\pm i x^2/(2\sigma)} = e^{\pm i\frac{\pi}{4}} \sqrt{2\pi\sigma}\ , \tag{2.15}$$

and treating each $p(t)$ as an independent variable. In the present case, we need the integral

$$\int dp\ e^{i\Delta(p\dot{q} - \frac{p^2}{2m})} = \underbrace{e^{-i\frac{\pi}{4}} \sqrt{\frac{2\pi m}{\Delta}}}_{\substack{\text{prefactor} \\ \text{independent of } q,\dot{q}}}\ e^{i\Delta \frac{m\dot{q}^2}{2}}\ . \tag{2.16}$$

The (infinite in the limit $\Delta \to 0$) prefactors can be hidden in the measure $\left[Dq(t)\right]$ since they do not depend on the path, and we are therefore led to the following formula:

$$\left\langle q_f \middle| e^{-i\mathcal{H}(t_f - t_i)} \middle| q_i \right\rangle = \int\limits_{\substack{q(t_i)=q_i \\ q(t_f)=q_f}} \left[Dq(t)\right] \exp\left\{i \int_{t_i}^{t_f} dt\ \mathcal{L}(q(t))\right\} = \int\limits_{\substack{q(t_i)=q_i \\ q(t_f)=q_f}} \left[Dq(t)\right]\ e^{i\mathcal{S}[q(t)]}\ , \tag{2.17}$$

where $\mathcal{L}(q)$ is the classical Lagrangian:

$$\mathcal{L}(q) \equiv \frac{m\dot{q}^2}{2} - V(q) \tag{2.18}$$

and $\mathcal{S}[q]$ the corresponding classical action.

2.2 Classical limit, Least action principle

In the previous section, we have written all the formulas with $\hbar = 1$. Had we kept the Planck constant, the final formula would have been:

$$\left\langle q_f \middle| e^{-i\mathcal{H}(t_f - t_i)} \middle| q_i \right\rangle = \int\limits_{\substack{q(t_i)=q_i \\ q(t_f)=q_f}} \left[Dq(t)\right]\ e^{\frac{i}{\hbar}\mathcal{S}[q(t)]}\ . \tag{2.19}$$

(This can be guessed a posteriori based on the fact that $\hbar$ has the dimension of an action.) Because of the factor i inside the exponential, this integral is wildly oscillating, except in the immediate vicinity of the function $q_c(t)$ that realizes the extremum of the action. Note that this function is precisely the solution of the classical Euler-Lagrange equations of motion. Roughly speaking, the phase oscillations become significant when

$$\left| \mathcal{S}[q(t)] - \mathcal{S}[q_c(t)] \right| \geq 2\pi\hbar\ , \tag{2.20}$$

and paths that fulfill this inequality do not contribute to the path integral. Therefore, in the limit $\hbar \to 0$, the path integral is dominated by the unique path $q_c(t)$, i.e. by the classical trajectory of the system. The path integral formalism thus provides a very intuitive way of connecting smoothly quantum and classical mechanics.

Figure 2.2: *Illustration of eq. (2.19). The paths whose action is far apart from the classical extremum are plotted in fainter colors. The solid black line is the classical trajectory.*
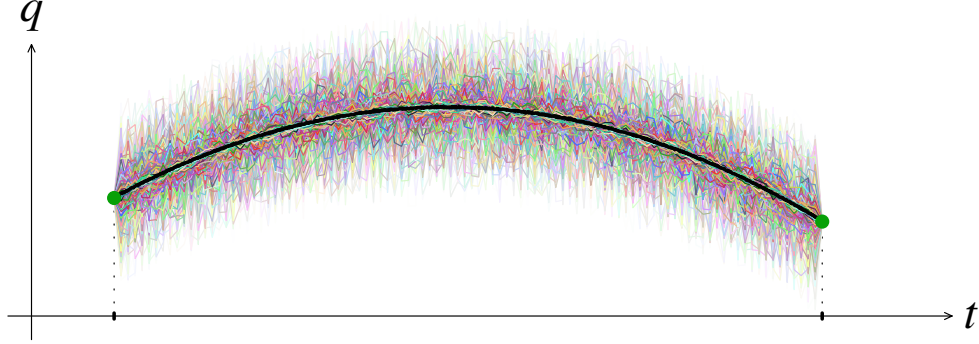
2.3 More functional machinery

2.3.1 Time-ordered products

Consider the matrix element

$$\left\langle q_f \middle| e^{-i\mathcal{H}(t_f-t_1)} \, Q \, e^{-i\mathcal{H}(t_1-t_i)} \middle| q_i \right\rangle , \qquad (2.21)$$

that measures the expectation value of the position at the time t_1. In order to evaluate this object, we need to insert on either side of the position operator Q an identity operator written as a complete sum over position eigenstates, i.e.

$$Q \quad \rightarrow \quad \int dq \, dq' \, |q\rangle \underbrace{\langle q|Q|q'\rangle}_{q \, \delta(q-q')} \langle q'| = \int dq \, q \, |q\rangle\langle q| . \qquad (2.22)$$

This leads immediately to the following path integral representation:

$$\left\langle q_f \middle| e^{-i\mathcal{H}(t_f-t_1)} \, Q \, e^{-i\mathcal{H}(t_1-t_i)} \middle| q_i \right\rangle = \int\limits_{\substack{q(t_i)=q_i \\ q(t_f)=q_f}} \left[Dq(t)\right] q(t_1) \, e^{iS[q(t)]} . \qquad (2.23)$$

Likewise, if $t_2 > t_1$, we have:

$$\left\langle q_f \middle| e^{-i\mathcal{H}(t_f-t_2)} \, Q \, e^{-i\mathcal{H}(t_2-t_1)} \, Q \, e^{-i\mathcal{H}(t_1-t_i)} \middle| q_i \right\rangle =$$

$$= \int\limits_{\substack{q(t_i)=q_i \\ q(t_f)=q_f}} \left[Dq(t)\right] q(t_1) \, q(t_2) \, e^{iS[q(t)]} . \qquad (2.24)$$

If we introduce a time-dependent position operator

$$Q(t) \equiv e^{i\mathcal{H}t} \, Q \, e^{-i\mathcal{H}t} , \qquad (2.25)$$

and its eigenstates

$$\big|q, t\big\rangle \equiv e^{i\mathcal{H}t}\big|q\big\rangle, \tag{2.26}$$

the previous equation takes a much more compact form

$$\big\langle q_f, t_f\big|Q(t_2)Q(t_1)\big|q_i, t_i\big\rangle \underset{t_2 > t_1}{=} \int\limits_{\substack{q(t_i)=q_i \\ q(t_f)=q_f}} \big[Dq(t)\big]\, q(t_1)\, q(t_2)\, e^{iS[q(t)]}. \tag{2.27}$$

The condition $t_2 > t_1$ is crucial here, because the left hand side would is quite different if the times are ordered differently. In contrast, the objects $q(t_1)$ and $q(t_2)$ in the right hand side are ordinary numbers that commute. One may render this formula true for any ordering between t_1 and t_2 by introducing a T-product, that ensures that the operator with the largest time is always on the left:

$$\big\langle q_f, t_f\big|T\big(Q(t_1)Q(t_2)\big)\big|q_i, t_i\big\rangle = \int\limits_{\substack{q(t_i)=q_i \\ q(t_f)=q_f}} \big[Dq(t)\big]\, q(t_1)\, q(t_2)\, e^{iS[q(t)]}. \tag{2.28}$$

This formula generalizes to n factors:

$$\big\langle q_f, t_f\big|T\big(Q(t_1)\cdots Q(t_n)\big)\big|q_i, t_i\big\rangle = \int\limits_{\substack{q(t_i)=q_i \\ q(t_f)=q_f}} \big[Dq(t)\big]\, q(t_1)\cdots q(t_n)\, e^{iS[q(t)]}. \tag{2.29}$$

This result is extremely important in applications to quantum field theory, since time-ordered products of field operators are the central objects that appear in the LSZ reduction formulas. One may also apply differential operators containing time derivatives on this equation, for instance:

$$\frac{\partial}{\partial t_1}\big\langle q_f, t_f\big|T\big(Q(t_1)\cdots Q(t_n)\big)\big|q_i, t_i\big\rangle = \int\limits_{\substack{q(t_i)=q_i \\ q(t_f)=q_f}} \big[Dq(t)\big]\, \dot{q}(t_1)\cdots q(t_n)\, e^{iS[q(t)]} \tag{2.30}$$

In other words, a time derivative in the integrand of the path integral also applies to the step functions that ensure the time ordering in the left hand side.

2.3.2 Functional sources and derivatives

The amplitudes of the form (2.29) can all be encapsulated into the following *generating functional*:

$$Z_{fi}[j(t)] \equiv \big\langle q_f, t_f\big|T\exp i\int_{t_i}^{t_f} dt\, j(t)\, Q(t)\big|q_i, t_i\big\rangle, \tag{2.31}$$

where $j(t)$ is some arbitrary function of time. From $Z_{fi}[j]$, the amplitudes can be recovered by *functional differentiation*:

$$\big\langle q_f, t_f\big|T\big(Q(t_1)\cdots Q(t_n)\big)\big|q_i, t_i\big\rangle = \frac{\delta^n\, Z_{fi}[j]}{i^n \delta j(t_1)\cdots \delta j(t_n)}\bigg|_{j\equiv 0}. \tag{2.32}$$

Functional derivatives obey the usual rules of differentiation, with the additional property that the values of the function $j(t)$ at different times should be viewed as independent variables, i.e.

$$\frac{\delta j(t)}{\delta j(t')} = \delta(t - t') \,.$$
(2.33)

From this formula, one may also read the dimension of a functional derivative:

$$\dim\left[\frac{\delta}{\delta j(t)}\right] = -\dim\left[j(t)\right] - \dim\left[t\right] \,.$$
(2.34)

From eq. (2.29), we can derive an expression of the generating functional Z_{fi} as a path integral,

$$Z_{fi}[j(t)] = \int\limits_{\substack{q(t_i)=q_i \\ q(t_f)=q_f}} \left[Dq(t)\right] e^{iS[q(t)]+i\int_{t_i}^{t_f} dt\, j(t)q(t)} \,,$$
(2.35)

that involves only the commuting c-number $q(t)$ and no time-ordering. Note also that there ia an Hamiltonian version of this path integral:

$$Z_{fi}[j(t)] = \int\limits_{\substack{q(t_i)=q_i \\ q(t_f)=q_f}} \left[Dp(t)Dq(t)\right] \exp\left\{i\int_{t_i}^{t_f} dt\, \left(p(t)\dot{q}(t) - \mathcal{H}(p(t), q(t)) + j(t)q(t)\right)\right\} \,.$$
(2.36)

2.3.3 Projection on the ground state at asymptotic times

So far in this section, we have considered amplitudes where the initial and final states are position eigenstates. However, the path integral formalism is not limited to this situation. Let us assume for instance that the system is in a state $\left|\psi_i\right\rangle$ at the time t_i and in the state $\left|\psi_f\right\rangle$ at the time t_f. For any operator $\mathcal{O}$, the expectation value between these two states can be related to transitions between position eigenstates by writing

$$\left\langle\psi_f, t_f\left|\mathcal{O}\right|\psi_i, t_i\right\rangle = \int dq_i dq_f\, \psi_f^*(q_f)\,\psi_i(q_i)\,\left\langle q_f, t_f\left|\mathcal{O}\right|q_i, t_i\right\rangle \,,$$
(2.37)

where

$$\psi(q) \equiv \left\langle q\middle|\psi\right\rangle$$
(2.38)

is the position representation of the wavefunction of the state $\left|\psi\right\rangle$. However, the use of this formula is cumbersome in practice, because of the integrations over $q_{i,f}$.

In the special case where the initial and final states are the ground state of the Hamiltonian, $\left|0\right\rangle$, and the initial and final times are $-\infty$ and $+\infty$, there is trick to circumvent this difficulty. Let us introduce the eigenstates $\left|n\right\rangle$ of the Hamiltonian, with eigenvalue E_n and eigenfunction $\psi_n(q) \equiv \left\langle q\middle|n\right\rangle$, and write

$$\begin{aligned}
\left|q_i, t_i\right\rangle &= e^{i\mathcal{H}t_i}\left|q_i\right\rangle \\
&= \sum_{n=0}^{\infty} e^{i\mathcal{H}t_i}\left|n\right\rangle\left\langle n\middle|q_i\right\rangle \\
&= \sum_{n=0}^{\infty} \psi_n^*(q_i)\, e^{iE_n t_i}\left|n\right\rangle \,.
\end{aligned}$$
(2.39)

We will assume that the Hamiltonian is shifted by a constant so that the energy of the ground state $|0\rangle$ is $E_0 = 0$. Now, we multiply the Hamiltonian by $1 - i0^+$, where 0^+ denotes some positive infinitesimal number. All the factors $\exp(i(1-i0^+)E_n t_i)$ go to zero when $t_i \to -\infty$, except for $n = 0$. Therefore, after this alteration of the Hamiltonian, we have:

$$\lim_{t_i \to -\infty} |q_i, t_i\rangle = \psi_0^*(q_i) |0\rangle . \tag{2.40}$$

We can then weight this equation by a function $\varphi(q_i)$,

$$\lim_{t_i \to -\infty} \varphi(q_i)|q_i, t_i\rangle = \underbrace{\int dq_i \varphi(q_i)\psi_0^*(q_i)}_{\langle 0|\varphi\rangle} |0\rangle . \tag{2.41}$$

For instance, the constant function $\varphi(q) = 1$, corresponding to the momentum eigenstate $p = 0$, would be appropriate in this role. But this would also be the case of any function $\varphi(q)$ such that the state $|\varphi\rangle$ has a non-zero overlap with the ground state $\langle 0|$. Likewise, changing $\mathcal{H} \to (1-i0^+)\mathcal{H}$ has a similar effect on the final state in the limit $t_f \to +\infty$,

$$\lim_{t_f \to +\infty} \langle q_f, t_f| = \psi_0(q_f) \langle 0| . \tag{2.42}$$

From these considerations, when the initial and final states at $\pm\infty$ are the ground state, we can write the generating functional in the following simple path integral form:

$$Z[j(t)] = \int \left[Dp(t)Dq(t)\right] \exp\left\{i\int dt \left(p(t)\dot{q}(t)-(1-i0^+)\mathcal{H}(p(t),q(t))+j(t)q(t)\right)\right\} . \tag{2.43}$$

From the discussion after eq. (2.41), we see that the boundary conditions on the paths are not important. They only affect an overall prefactor, that can be adjusted in such a way that $Z[0] = 1$. After performing the Gaussian functional integral over $p(t)$, we can rewrite this expression in Lagrangian form:

$$Z[j(t)] = \int \left[Dq(t)\right] \exp\left\{i\int dt \left((1+i0^+)\frac{m\dot{q}^2(t)}{2} - (1-i0^+)V(q(t))+j(t)q(t)\right)\right\} . \tag{2.44}$$

The term in $(i0^+)\dot{q}^2$ may be viewed as contributing to the convergence of the integral at large velocities. Likewise, for a confining potential such that $V(q) \to +\infty$ when $|q| \to \infty$, the term in $(i0^+)V(q)$ contributes to the convergence at large coordinates.

2.3.4 Functional Fourier transform

Given a functional $F[q(t)]$, its *functional Fourier transform* is defined by

$$\widetilde{F}[p(t)] \equiv \int \left[Dq(t)\right] F[q(t)] \exp\left\{i\int dt\, p(t)q(t)\right\} . \tag{2.45}$$

In other words, the Fourier conjugate of the "variable" $q(t)$ is another function of time, $p(t)$. Eq. (2.45) may be inverted by

$$F[q(t)] \equiv \int \left[Dp(t)\right] \widetilde{F}[p(t)] \exp\left\{-i\int dt\, p(t)q(t)\right\} . \tag{2.46}$$

The usual properties of ordinary Fourier transforms extend to the functional case, e.g.:

- The Fourier transform of a constant is a delta function,

- The Fourier transform of a Gaussian is another Gaussian,

- The Fourier transform of a product is the convolution product of the Fourier transforms.

2.3.5 Functional translation operator

The functional derivative $\delta/\delta j(t)$ may be viewed as the generator of translations in the space of the functions $j(t)$. Its exponential provides a translation operator:

$$\exp\left\{\int dt\, a(t)\,\frac{\delta}{\delta j(t)}\right\} F[j(t)] = F[j(t) + a(t)]\,, \tag{2.47}$$

for any functional $F[j(t)]$.

Another extremely important formula is

$$\underbrace{\exp\left\{\lambda\int dt\,\left(\frac{\delta}{\delta j(t)}\right)^n\right\}\,\exp\left\{\int dt\, j(t)q(t)\right\}}_{A[j,q;\lambda]} = \underbrace{\exp\left\{\int dt\,\left(j(t)q(t) + \lambda\, q^n(t)\right)\right\}}_{B[j,q;\lambda]}\,.$$

$$\tag{2.48}$$

The proof of this formula consists in noticing that $A[j, q; \lambda = 0] = B[j, q; \lambda = 0]$, and in comparing their (ordinary) derivatives with respect to λ:

$$\partial_\lambda A[j, q; \lambda] = \lambda\int dt\,\left(\frac{\delta}{\delta j(t)}\right)^n A[j, q; \lambda] = \lambda\int dt\, q^n(t)\, A[j, q; \lambda]\,,$$

$$\partial_\lambda B[j, q; \lambda] = \lambda\int dt\, q^n(t)\, B[j, q; \lambda]\,. \tag{2.49}$$

Therefore $A[j, q; \lambda]$ and $B[j, q; \lambda]$ are equal at $\lambda = 0$ and obey the same differential equation.

2.3.6 Functional diffusion operator

It is sometimes useful to evaluate the action of an operator which is quadratic in functional derivatives. The result is given by

$$\exp\left\{\int dt\,\frac{\sigma(t)}{2}\left(\frac{\delta}{\delta j(t)}\right)^2\right\} F[j] = \int [Da(t)]\,\exp\left\{-\int dt\,\frac{a^2(t)}{2\sigma(t)}\right\} F[j + a]\,. \tag{2.50}$$

In order to establish this formula, consider the following differential equation,

$$\partial_z F[j(t); z] = \left\{\int dt\,\frac{\sigma(t)}{2}\left(\frac{\delta}{\delta j(t)}\right)^2\right\} F[j; z]\,, \tag{2.51}$$

where z is an ordinary real-valued variable. One may view this equation as a diffusion equation in the space of the functions $j(t)$, and $F[j; z]$ as a density functional on this space. The left hand side of eq. (2.50) is the formal expression of the solution of this equation at $z = 1$, if we interpret $F[j]$ as its initial condition at $z = 0$. In order to show that it is equal to the right

hand side, one should first transform the diffusion equation (2.51) by performing a functional Fourier transform,

$$\widetilde{F}[k(t); z] \equiv \int [Dj(t)] \exp\left\{i\int dt\, j(t)k(t)\right\} F[j(t)]$$

$$\partial_z \widetilde{F}[k(t); z] = -\left\{\int dt\, \frac{\sigma(t)}{2}k^2(t)\right\} \widetilde{F}[k(t); z] . \tag{2.52}$$

The solution of the latter equation is simply

$$\widetilde{F}[k(t); z = 1] = \exp\left\{-\int dt\, \frac{\sigma(t)}{2}k^2(t)\right\} \widetilde{F}[k(t); z = 0] , \tag{2.53}$$

and the inverse Fourier transform of this solution leads to the right hand side of eq. (2.50).

2.4 Path integral in scalar field theory

The functional formalism that we have exposed in the context of quantum mechanics can now be extended to quantum field theory. The main change is that the functions over which one integrates are functions of time and space (as opposed to functions of time only in quantum mechanics). All the result of the previous section can be translated into analogous formulas in quantum field theory, thanks to the following correspondence:

$$\begin{aligned} q(t) &\longleftrightarrow \phi(x) \\ p(t) &\longleftrightarrow \Pi(x) \\ j(t) &\longleftrightarrow j(x) \end{aligned}$$

$$\tag{2.54}$$

The main results of the previous section, namely that time-ordered products of operators in the canonical formalism become simple products of ordinary functions in the path integral representation, and that the ground state at $\pm\infty$ can be obtained by relaxing the boundary conditions and multiplying the Hamiltonian by $1 - i0^+$, remain true in this new context. Thus, the analogue of eq. (2.43) in a real scalar field theory is:

$$Z[j] = \int [D\Pi(x)D\phi(x)] \exp\left\{i\int d^4x\, \left(\Pi(x)\dot{\phi}(x) - (1 - i0^+)\mathcal{H}(\Pi, \phi) + j(x)\phi(x)\right)\right\}. \tag{2.55}$$

Since the Hamiltonian is quadratic in Π,

$$\mathcal{H} = \frac{1}{2}\Pi^2 + \frac{1}{2}(\boldsymbol{\nabla}\phi)\cdot(\boldsymbol{\nabla}\phi) + \frac{1}{2}m^2\phi^2 + V(\phi) , \tag{2.56}$$

it is easy to perform the (Gaussian) functional integration on Π, to obtain:

$$Z[j] = \int [D\phi(x)] \exp\left\{i\int d^4x\, \left(\mathcal{L}(\phi) + j(x)\phi(x)\right)\right\}, \tag{2.57}$$

where

$$\mathcal{L}(\phi) \equiv \frac{1}{2}(1 + i0^+)\dot{\phi}^2 - \frac{1}{2}(1 - i0^+)\left((\boldsymbol{\nabla}\phi)\cdot(\boldsymbol{\nabla}\phi) + m^2\phi^2\right) - (1 - i0^+)V(\phi) . \tag{2.58}$$

Note that the $1 - i0^+$ in front of the interaction potential plays no role if we turn off adiabatically the coupling constant when $|x^0| \to \infty$. Using the analogue of eq. (2.48), we can separate the interactions as follows

$$Z[j] = \exp\left\{ -i \int d^4x \, V\left(\frac{\delta}{i\delta j(x)}\right) \right\} Z_0[j] \,, \tag{2.59}$$

with

$$\begin{aligned} Z_0[j] &\equiv \int [D\phi(x)] \, \exp\left\{ i \int d^4x \, \big(\mathcal{L}_0(\phi) + j(x)\phi(x)\big) \right\}, \\ \mathcal{L}_0(\phi) &= \frac{1}{2}(1 + i0^+)\dot\phi^2 - \frac{1}{2}(1 - i0^+)\big((\boldsymbol{\nabla}\phi)\cdot(\boldsymbol{\nabla}\phi) + m^2\phi^2\big)\,. \end{aligned} \tag{2.60}$$

The functional integral that gives $Z_0[j]$ in eq. (2.60) is Gaussian in ϕ and can be performed in a straightforward manner, giving

$$Z_0[j] = \exp\left\{ -\frac{1}{2} \int d^4x d^4y \, j(x)j(y) \, G_F^0(x,y) \right\} \,, \tag{2.61}$$

where $G_F^0(x,y)$ is the inverse of the operator

$$i\left[(1 + i0^+)\partial_0^2 - (1 - i0^+)(\boldsymbol{\nabla}^2 + m^2)\right]\,. \tag{2.62}$$

Note that the terms in $i0^+$ ensure the existence of this inverse. Going to momentum space, we see that the Fourier transform of this inverse is

$$\frac{i}{(1 + i0^+)k_0^2 - (1 - i0^+)(\mathbf{k}^2 + m^2)}\,, \tag{2.63}$$

which after some rearrangement of the $i0^+$'s appears to be nothing but eq. (1.86). Although the canonical quantization of a scalar field theory was tractable, we see on this example that the path integral approach provides a much quicker way of obtaining the expression of the free generating functional, with the correct pole prescription for the free Feynman propagator.

2.5 Functional determinants

In the earlier sections of this chapter, we have been a bit cavalier with Gaussian integrations, since we have disregarded the constant prefactors they produce. This was legitimate in the problems we were considering, since the normalization of the generating functional can be fixed by hand. However, in certain situations, these prefactors depend crucially on quantities that have a physical significance, e.g. on a background field.

In order to compute this prefactor, let us start from a simple 1-dimensional Gaussian integral,

$$\int_{-\infty}^{+\infty} dx \, e^{-\frac{1}{2}ax^2} = \sqrt{\frac{2\pi}{a}}\,. \tag{2.64}$$

The first stage of generalization is to replace x by an n-component vector $\mathbf{x} \equiv (x_1, \cdots, x_n)$, and the positive number a by a positive definite symmetric matrix $\mathbf{A}$, and to consider the integral

$$I(\mathbf{A}) \equiv \int \prod_{i=1}^{n} dx_i \, e^{-\frac{1}{2}\mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x}}\,. \tag{2.65}$$

65

This integral can be calculated by representing the vector $\mathbf{x}$ in the orthonormal basis made of the eigenvectors of $\mathbf{A}$ (such a basis exists, since $\mathbf{A}$ is symmetric). The measure $\prod_i dx_i$ is unchanged, because the diagonalization of the matrix can be done by an orthogonal transformation. Therefore, the above integral also reads

$$I(\mathbf{A}) = \int \prod_{i=1}^{n} dy_i \; e^{-\frac{1}{2}\sum_i a_i y_i^2} = \prod_{i=1}^{n} \sqrt{\frac{2\pi}{a_i}} \; , \tag{2.66}$$

where the numbers a_i are the eigenvalues of $\mathbf{A}$. This result can be written in a much more compact form:

$$I(\mathbf{A}) = \frac{(2\pi)^{n/2}}{\sqrt{\det \mathbf{A}}} \; . \tag{2.67}$$

This reasoning can be generalized to the functional case by writing:

$$\int \left[D\phi(x) \right] \; \exp \left\{ -\frac{1}{2} \int d^4x d^4y \; \phi(x) A(x,y) \phi(y) \right\} = \left[\det(A) \right]^{-1/2} \; , \tag{2.68}$$

where $A(x,y)$ is a symmetric operator. In this formula, we have still disregarded some truly constant (and infinite) prefactors, made of powers of 2π. One can also generalize this Gaussian integral to the case where the vector $\mathbf{x}$ is complex,

$$J(\mathbf{A}) \equiv \int \prod_{i=1}^{n} dx_i dx_i^* \; e^{-x^\dagger \mathbf{A} x} = \frac{(2\pi)^n}{\det \mathbf{A}} \; , \tag{2.69}$$

where $\mathbf{A}$ is a Hermitean matrix. The functional analogue of this integral is

$$\int \left[D\phi(x) D\phi^*(x) \right] \; \exp \left\{ -\int d^4x d^4y \; \phi^*(x) A(x,y) \phi(y) \right\} = \left[\det(A) \right]^{-1} \; , \tag{2.70}$$

Zeta function regularization : Despite the elegance of this formula, one should keep in mind that the functional determinant $\det A$ is most often infinite, because the spectrum of the operator extends to infinity. A common regularization technique for functional determinants is based on a generalization of Riemann's ζ function. Let the λ_n be the eigenvalues of A, and define:

$$\zeta_A(s) \equiv \mathrm{tr}\left(A^{-s} \right) = \sum_n \frac{1}{\lambda_n^s} \; . \tag{2.71}$$

(The function ζ_A is called the zeta function of the operator A.) The determinant of A is related to this function by

$$\det(A) = \exp\left(-\zeta_A'(0) \right) \; . \tag{2.72}$$

The sum over n in the definition of ζ_A usually converges only if $\mathrm{Re}\,(s)$ is large enough (how large depends on the distribution of eigenvalues at large n), but not for $s = 0$. However, like in the case of Riemann's zeta function, $\zeta_A(s)$ can be analytically continued to most of the complex s-plane, which provides a regularized definition of the determinant.

Diagrammatic interpretation :   Let us consider as an example the operator $A_\varphi \equiv \square + \frac{\lambda}{2}\varphi^2$, where $\varphi(x)$ is a background field. The inverse of this operator is the propagator of a scalar particle (with a ϕ^4 interaction) over the background field φ. We can skip the regularization step if we make a ratio with the determinant of the similar operator with no background field:

$$R \equiv \frac{\det\left(\square\right)}{\det\left(\square + \frac{\lambda}{2}\varphi^2\right)} \ . \tag{2.73}$$

A very useful formula relates the determinant of an operator to the trace of its logarithm,

$$\det\left(A\right) = \exp\left(\mathrm{Tr}\,\log\left(A\right)\right)\ . \tag{2.74}$$

This formula can be proven (heuristically, since the objects we are manipulating may not be finite) by writing both sides of the equation in terms of the eigenvalues of A:

$$\det\left(A\right) = \prod_n \lambda_n = \exp\sum_n \log\lambda_n = \exp\left(\mathrm{Tr}\,\log\left(A\right)\right)\ . \tag{2.75}$$
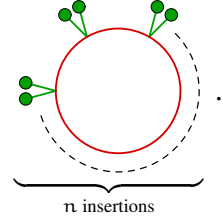
Therefore, the ratio defined in eq. (2.73) can be rewritten as

$$R = \exp\left(-\mathrm{Tr}\,\log\left(1 + \tfrac{\lambda\varphi^2}{2}\square^{-1}\right)\right)\ . \tag{2.76}$$

Writing $\square^{-1} = iG_{\scriptscriptstyle F}^0$, and expanding the logarithm gives

$$R = \exp\left\{\sum_{n=1}^\infty \frac{1}{n}\mathrm{Tr}\left(\left[-i\tfrac{\lambda\varphi^2}{2}G_{\scriptscriptstyle F}^0\right]^n\right)\right\}\ . \tag{2.77}$$

The argument of the exponential has a simple interpretation as a 1-loop diagram made of a line dressed with insertions of the background field, the index n being the number of such insertions:

$$\frac{1}{n}\mathrm{Tr}\left(\left[-i\tfrac{\lambda\varphi^2}{2}G_{\scriptscriptstyle F}^0\right]^n\right) = \qquad\qquad . \tag{2.78}$$

n insertions

Each of the insertions of the background field (shown in green in the above diagram) corresponds to a factor $-i\frac{\lambda}{2}\varphi^2$. The prefactor $1/n$ is the symmetry factor for the cyclic permutations of the n insertions. The argument of the exponential is a sum of connected 1-loop diagrams. Taking the exponential to obtain the ratio R simply produces all the multiply connected graphs made of products of such 1-loop diagrams.

2.6   Quantum effective action

2.6.1   Definition

The action $S[\phi]$ that enters in the path integral representation of the generating functional $Z[j]$ is the *classical action*. Its parameters reflect the interactions among the constituents of

the system at tree level, but in order to express higher order corrections loop corrections are necessary. The *quantum effective action*, denoted $\Gamma[\phi]$, is defined as the functional that would produce the all-orders value of the interactions solely from tree-level contributions. $\Gamma[\phi]$ should coincide with the classical action at lowest order of perturbation theory, but also encapsulates all the higher order corrections. One may write $\Gamma[\phi]$ formally as

$$\Gamma[\phi] \equiv \sum_{n=2}^{\infty} \frac{1}{n!} \int d^4x_1 \cdots d^4x_n \; \phi(x_1) \cdots \phi(x_n) \; \Gamma_n(x_1, \cdots, x_n) \, . \tag{2.79}$$

$\Gamma_2(x_1, x_2)$ is therefore the inverse of the exact propagator, $\Gamma_4(x_1, \cdots, x_4)$ is the exact 4-point function (in coordinate space), etc...

An important class of diagrams when discussing the quantum effective action are the *one-particle irreducible (1PI)* diagrams, that are defined as diagrams that remain connected if one cuts any one of their internal propagators. For instance, the first of these diagrams is 1PI while the second one is not:
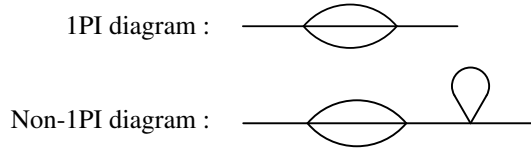
The concept of 1PI diagrams is most useful when discussing two-point functions, since it allows to isolate the elementary building blocks that allow to construct the full propagator by the resummation of a geometrical series. Thus, the Γ_2 that enters in the expansion of the quantum effective action is made of 1PI graphs.

2.6.2   Relation between $\Gamma[\phi]$ and $W[j]$

Until now, we have introduced the generating functional of the vacuum expectation value of time-ordered products of fields, $Z[j]$, as well as the functional $W[j] \equiv \log Z[j]$ that generates the subset made of connected Feynman graphs. Recall that in term of path integrals,

$$Z[j] = e^{W[j]} = \int \left[D\phi(x) \right] \; \exp\left[iS[\phi(x)] + i \int d^4x \; j(x)\phi(x) \right] \, . \tag{2.80}$$

Let us replace the classical action $S[phi]$ by the quantum effective action $\Gamma[\phi]$ in the previous formula, to define

$$Z_\Gamma[j] = e^{W_\Gamma[j]} = \int \left[D\phi(x) \right] \; \exp\left[i\Gamma[\phi(x)] + i \int d^4x \; j(x)\phi(x) \right] \, . \tag{2.81}$$

This functional generates graphs whose building blocks are the exact propagator (Γ_2^{-1}), and the exact vertices ($\Gamma_3, \Gamma_4. \cdots$). From the definition of $\Gamma[\phi]$ as the "action" that would generate the exact theory at tree level, we conclude that

$$W_\Gamma[j]\big|_{\text{tree}} = W[j] \, . \tag{2.82}$$

In other words, the tree diagrams of $W_\Gamma[j]$ should be equal to the all-orders $W[j]$. The tree diagrams may be isolated by reintroducing Planck's constant in the definition of $Z_\Gamma[j]$ as follows

$$Z_\Gamma[j;\hbar] = e^{W_\Gamma[j;\hbar]} = \int \left[D\phi(x) \right] \; \exp\left[\frac{i}{\hbar} \left(\Gamma[\phi(x)] + \int d^4x \; j(x)\phi(x) \right) \right] \, . \tag{2.83}$$

As we have discussed in the section 1.6.6, the order in $\hbar$ of a connected graph is

$$\hbar^{n_L - 1} \, , \tag{2.84}$$

where n_L is the number of loops of the graph. Therefore, the functional $W_\Gamma[j; \hbar]$ has the following loop expansion:

$$W_\Gamma[j; \hbar] = \sum_{n_L = 0}^{\infty} \hbar^{n_L - 1} \underbrace{W_{\Gamma, n_L}[j]}_{n_L \text{ loops}} \, , \tag{2.85}$$

and the tree level contributions in $W_\Gamma[j]$ are the terms that dominate in the formal limit $\hbar \to 0$:

$$W_\Gamma[j]\big|_{\text{tree}} = \lim_{\hbar \to 0} \hbar\, W_\Gamma[j; \hbar] \, . \tag{2.86}$$

But from our discussion of the classical limit of path integrals in section 2.2, we know that the limit $\hbar \to 0$ corresponds to the extremum of the argument of the exponential, i.e.

$$\frac{\delta \Gamma[\phi]}{\delta \phi(x)} + j(x) = 0 \, . \tag{2.87}$$

Note that this equation is the analogue of the usual Euler-Lagrange equation of motion, with the quantum effective action in place of the classical action. This equation implicitly defines ϕ as a function of j, that we will denote ϕ_j, in terms of which we can write

$$e^{W_\Gamma[j\hbar]} \underset{\hbar \to 0}{\approx} \exp\left[\frac{i}{\hbar} \left(\Gamma[\phi_j(x)] + \int d^4x\, j(x) \phi_j(x) \right) \right] \, , \tag{2.88}$$

which leads to the following relationship between the quantum effective action and the generating functional of connected graphs:

$$\Gamma[\phi_j] = -i\, W[j] - \int d^4x\, j(x) \phi_j(x) \, . \tag{2.89}$$

Note that the "quantum equation of motion" (2.87) may also be viewed as defining j in terms of ϕ, that we shall denote j_ϕ. Eq. (2.89) may therefore also be written as

$$\Gamma[\phi] = -i\, W[j_\phi] - \int d^4x\, j_\phi(x) \phi(x) \, . \tag{2.90}$$

Taking a functional derivative of this equation with respect to $\phi(y)$ and using the chain rule, we obtain

$$\underbrace{\frac{\delta \Gamma[\phi]}{\delta \phi(y)}}_{-j_\phi(y)} = -i \int d^4x\, \frac{\delta W[j]}{\delta j(x)} \bigg|_{j = j_\phi} \frac{\delta j_\phi(x)}{\delta \phi(y)} - j_\phi(y) - \int d^4x\, \frac{\delta j_\phi(x)}{\delta \phi(y)} \phi(x) \, . \tag{2.91}$$

This leads to

$$\phi(x) = -i\, \frac{\delta W[j]}{\delta j(x)} \bigg|_{j = j_\phi} \, , \quad \text{or equivalently} \quad \phi_j(x) = -i\, \frac{\delta W[j]}{\delta j(x)} = \langle \phi(x) \rangle_j \, . \tag{2.92}$$

In other words, ϕ_j is the connected 1-point function (i.e. the vacuum expectation value of the field) in the presence of the source j.

Differentiating eq. (2.87) with respect to $j(y)$ gives:

$$
\begin{aligned}
\delta(x-y) &= -\frac{\delta}{\delta j(y)} \frac{\delta \Gamma[\phi_j]}{\delta \phi_j(x)} \\
&= -\int d^4 z \, \frac{\delta \phi_j(z)}{\delta j(y)} \frac{\delta^2 \Gamma[\phi_j]}{\delta \phi_j(x) \delta \phi_j(z)} \\
&= i \int d^4 z \, \underbrace{\frac{\delta^2 W[j]}{\delta j(y) \delta j(z)}}_{G(y,z)_{\text{connected}}} \underbrace{\frac{\delta^2 \Gamma[\phi_j]}{\delta \phi_j(z) \delta \phi_j(x)}}_{\Gamma_2(z,x)} .
\end{aligned}
\tag{2.93}
$$

This formula shows a posteriori that (up to a factor i), the coefficient Γ_2 in the expansion (2.79) is indeed the inverse of the exact connected 2-point function, as was expected from our request that the effective action $\Gamma[\phi]$ reproduces the full content of the theory. By similar (but much more cumbersome) manipulations, one could check that the higher functional derivatives of $\Gamma[\phi]$ are the connected 1PI amplitudes.

2.6.3 One-loop effective action

At one loop, one may obtain a closed expression for the quantum effective action. For this, write the Lagrangian as a renormalized Lagrangian plus counter-terms:

$$
\mathcal{L} \equiv \mathcal{L}_r(\phi_r) + \Delta \mathcal{L}(\phi_r) ,
\tag{2.94}
$$

both depending on the renormalized field ϕ_r. We will denote $\mathcal{S}_r$ and $\Delta \mathcal{S}$ the corresponding actions. Likewise, we write the external source $j = j_r + \delta j$, where j_r is the current that solves the following equation:

$$
\left. \frac{\delta \mathcal{S}_r[\phi_r]}{\delta \phi_r(x)} \right|_\varphi + j_r(x) = 0 ,
\tag{2.95}
$$

i.e. the current that solves at lowest order the defining equation of the effective action. The correction Δj is then adjusted order by order so that the expectation value of the field remains equal to φ at all orders,

$$
\varphi(x) = \left\langle \phi_r(x) \right\rangle_{j_r + \Delta j} .
\tag{2.96}
$$

In the path integral representation of the generating functional $Z[j]$, we write the field as $\phi_r = \varphi + \eta$:

$$
Z[j] = \int \left[D\eta(x) \right] e^{i \left\{ \mathcal{S}_r[\varphi+\eta] + \Delta \mathcal{S}[\varphi+\eta] + \int d^4x \, (j_r + \Delta j)(\varphi+\eta) \right\}} ,
\tag{2.97}
$$

and we expand the argument of the exponential in powers of η up to quadratic order:

$$
\begin{aligned}
\mathcal{S}_r[\varphi+\eta] + \int d^4x \, j_r(\varphi+\eta) &= \mathcal{S}_r[\varphi] + \int d^4x \, j_r(x) \varphi(x) \\
&\quad + \int d^4x \left(\left. \frac{\delta \mathcal{S}_r[\phi_r]}{\delta \phi_r(x)} \right|_\varphi + j_r \right) \eta(x) \\
&\quad + \frac{1}{2} \int d^4x d^4y \, \eta(x) \left(\left. \frac{\delta^2 \mathcal{S}_r[\phi_r]}{\delta \phi_r(x) \delta \phi_r(y)} \right|_\varphi \right) \eta(y) \\
&\quad + \cdots
\end{aligned}
\tag{2.98}
$$

Note that the term linear in η is zero by virtue of eq. (2.95). Therefore, we may rewrite $Z[j]$ as follows

$$Z[j] \quad = \quad e^{i\left\{S_r[\varphi]+\Delta S[\varphi]+\int d^4x\, j\varphi\right\}} \int \left[D\eta(x)\right]\, e^{i\left\{S_\varphi[\eta]+\Delta S_\varphi[\eta]\right\}}\,, \tag{2.99}$$

where we denote

$$S_\varphi[\eta] \equiv \frac{1}{2}\int d^4x\, d^4y\,\eta(x)\left(\left.\frac{\delta^2 S_r[\phi_r]}{\delta\phi_r(x)\delta\phi_r(y)}\right|_\varphi\right)\eta(y) + \cdots \tag{2.100}$$

(Likewise, $\Delta S_\varphi[\eta]$ results from the expansion in powers of η of the counter-terms.) At one loop, it is sufficient to keep only the quadratic terms in η, and the path integral gives a determinant:

$$\left[\det\left(-\frac{i}{2}\left.\frac{\delta^2 S_r[\phi_r]}{\delta\phi_r(x)\delta\phi_r(y)}\right|_\varphi\right)\right]^{-1/2} = \exp\left[-\frac{1}{2}\mathrm{tr}\,\ln\left(-\frac{i}{2}\left.\frac{\delta^2 S_r[\phi_r]}{\delta\phi_r(x)\delta\phi_r(y)}\right|_\varphi\right)\right]\,. \tag{2.101}$$

At this order, the generating functional of connected graphs reads

$$W[j] = i\left\{S_r[\varphi]+\Delta S[\varphi]+\int d^4x\, j\varphi\right\} - \frac{1}{2}\,\mathrm{tr}\,\ln\left(-\frac{i}{2}\left.\frac{\delta^2 S_r[\phi_r]}{\delta\phi_r(x)\delta\phi_r(y)}\right|_\varphi\right) + \cdots \tag{2.102}$$

from which we obtain the following quantum effective action

$$\Gamma[\varphi] = S_r[\varphi] + \Delta S[\varphi] + \frac{i}{2}\,\mathrm{tr}\,\ln\left(-\frac{i}{2}\left.\frac{\delta^2 S_r[\phi_r]}{\delta\phi_r(x)\delta\phi_r(y)}\right|_\varphi\right) + \cdots \tag{2.103}$$

Note that the object inside the logarithm is the inverse of the propagator dressed by the background field φ.

2.7 Euclidean path integral and Statistical mechanics

2.7.1 Statistical mechanics in path integral form

A path integral formalism also exists for statistical mechanics. In order to illustrate this, let us consider again the quantum mechanical system described by the Hamiltonian of eq. (2.1). Our goal is to calculate the *partition function* in the canonical ensemble[3],

$$\mathcal{Z}_\beta \equiv \mathrm{Tr}\left(e^{-\beta\mathcal{H}}\right)\,, \tag{2.104}$$

where β is the inverse temperature (it is customary to use a system of units in which Boltzmann's constant k_{B} is equal to unity – therefore temperature has the same dimension as energy.) More generally, one may want to calculate the following canonical ensemble expectation values,

$$\left\langle\mathcal{O}\right\rangle_\beta \equiv \mathcal{Z}_\beta^{-1}\,\mathrm{Tr}\left(e^{-\beta\mathcal{H}}\,\mathcal{O}\right)\,. \tag{2.105}$$

---

[3]In theories with a conserved quantity, it is also possible to study the grand canonical ensemble. One needs to substitute $\mathcal{H}\to\mathcal{H}-\mu Q$ in the definition of the partition function, where Q is the operator of the conserved charge and μ the associated chemical potential.

The cyclicity of the trace leads to an important identity for expectation values of products of operators:

$$
\begin{aligned}
\langle \mathcal{O}_1(t)\mathcal{O}_2(t')\rangle_\beta &\equiv \mathcal{Z}_\beta^{-1}\, \mathrm{Tr}\left(e^{-\beta\mathcal{H}}\,\mathcal{O}_1(t)\mathcal{O}_2(t')\right) \\
&= \mathcal{Z}_\beta^{-1}\, \mathrm{Tr}\left(e^{-\beta\mathcal{H}}\,\mathcal{O}_1(t)\underbrace{e^{+\beta\mathcal{H}}e^{-\beta\mathcal{H}}}_{1}\,\mathcal{O}_2(t')\right) \\
&= \mathcal{Z}_\beta^{-1}\, \mathrm{Tr}\left(e^{-\beta\mathcal{H}}\mathcal{O}_2(t')\,\underbrace{e^{-\beta\mathcal{H}}\,\mathcal{O}_1(t)e^{+\beta\mathcal{H}}}_{\mathcal{O}_1(t+i\beta)}\right) \\
&= \langle \mathcal{O}_2(t')\mathcal{O}_1(t+i\beta)\rangle_\beta\,,
\end{aligned}
\tag{2.106}
$$

where we have formally identified the density operator $\exp(-\beta\mathcal{H})$ with a time evolution operator for an imaginary time $i\beta$. This relationship is called the *Kubo-Martin-Schwinger (KMS)* identity. Although we have established it for an expectation value of a product of two operators, it is completely general.

The identification of the density operator with an imaginary time evolution operator is at the heart of the formalism to evaluate canonical ensemble expectation values. If we represent the trace that appears in the partition function in the coordinate basis,

$$
\mathcal{Z}_\beta = \int dq\,\langle q|e^{-\beta\mathcal{H}}|q\rangle\,,
\tag{2.107}
$$

the integrand in the right hand side is a transition amplitude similar to eq. (2.3), except that initial and final coordinates are identical, and the time interval is imaginary. We can nevertheless formally reproduce all the manipulations of the section 2.1, with an initial time $t_i \equiv 0$ and a final time $t_f \equiv -i\beta$. It is common to introduce the *Euclidean time* $\tau \equiv it$, with τ varying from 0 to β. The only changes to our original derivation of the path integral is that the path $q(t)$ must be replaced by a path $q(\tau)$ whose time derivative is the *Euclidean velocity* $\dot{q}_E$, related to the usual velocity $\dot{q}$ by

$$
\dot{q} \equiv \frac{dq}{dt} = i\,\underbrace{\frac{dq}{d\tau}}_{\dot{q}_E}\,.
\tag{2.108}
$$

We obtain the following path integral representation of the partition function:

$$
\begin{aligned}
\mathcal{Z}_\beta &= \int dq \int_{\substack{q(0)=q \\ q(\beta)=q}} \left[Dp(\tau)Dq(\tau)\right]\,\exp\left\{\int_0^\beta d\tau\left(i\,p(\tau)\dot{q}_E(\tau)-\mathcal{H}(p(\tau),q(\tau))\right)\right\} \\
&= \int_{q(0)=q(\beta)} \left[Dp(\tau)Dq(\tau)\right]\,\exp\left\{\int_0^\beta d\tau\left(i\,p(\tau)\dot{q}_E(\tau)-\mathcal{H}(p(\tau),q(\tau))\right)\right\}.
\end{aligned}
\tag{2.109}
$$

In the second line, we have simplified the boundary conditions of the path $q(\tau)$, since the only constraint it must obey is to be β-periodic in imaginary time. The integration over the momentum $p(\tau)$ is again Gaussian, and after performing it we obtain the following expression

$$
\mathcal{Z}_\beta = \int_{q(0)=q(\beta)} \left[Dq(\tau)\right]\,\exp\left\{-\underbrace{\int_0^\beta d\tau\left(\frac{m}{2}\dot{q}_E^2(\tau)+V(q(\tau))\right)}_{S_E[q(\tau)]}\right\}.
\tag{2.110}
$$

The quantity $\mathcal{S}_E[q]$ is called the *Euclidean action*.

Then, we can generalize this formalism to calculate ensemble averages of time-ordered (in imaginary time) products of position operators. For instance, the analogue of eq. (2.28) is

$$\mathrm{Tr}\left(e^{-\beta\mathcal{H}}\,T_\tau\left(Q(\tau_1)Q(\tau_2)\right)\right) = \int_{q(0)=q(\beta)}\left[Dq(\tau)\right]\,e^{-\mathcal{S}_E[q(\tau)]}\,q(\tau_1)q(\tau_2)\,, \quad (2.111)$$

where the symbol T_τ denotes the time-ordering in the imaginary time τ. Likewise, we may define a generating functional for these expectation values

$$\mathrm{Tr}\left(e^{-\beta\mathcal{H}}\,T_\tau\,\exp\int_0^\beta d\tau\,j(\tau)Q(\tau)\right) = \int_{q(0)=q(\beta)}\left[Dq(\tau)\right]\,e^{-\mathcal{S}_E[q(\tau)]+\int_0^\beta d\tau\,j(\tau)q(\tau)}\,.$$

$$(2.112)$$

2.7.2   Statistical field theory

This formalism can be extended readily to a quantum field theory. In this context, it can be used to calculate canonical ensemble expectation values of operators for a system of relativistic particles. One can write directly the following generalization of eq. (2.112),

$$\underbrace{\mathrm{Tr}\left(e^{-\beta\mathcal{H}}\,T_\tau\,\exp\int_0^\beta d^4x_E\,j(x)\phi(x)\right)}_{Z[j;\beta]} = \int_{\phi(0,\boldsymbol{x})=\phi(\beta,\boldsymbol{x})}\left[D\phi(x)\right]\,e^{-\mathcal{S}_E[\phi(x)]+\int_0^\beta d^4x_E\,j(x)\phi(x)}\,,$$

$$(2.113)$$

where the measure d^4x_E stands for $d\tau\,d^3\boldsymbol{x}$. Like in the case of ordinary QFT in Minkowski space-time, we can isolate the interactions by writing:

$$Z[j;\beta] = \exp\left\{-\int d^4x_E\,\mathcal{L}_{E,I}\left(\frac{\delta}{\delta j(x)}\right)\right\}\,Z_0[j;\beta]\,, \quad (2.114)$$

where $\mathcal{L}_{E,I}$ is the interaction term in the Euclidean Lagrangian density, and $Z_0[j;\beta]$ is the generating functional of the non-interacting theory:

$$Z_0[j;\beta] = \int_{\phi(0,\boldsymbol{x})=\phi(\beta,\boldsymbol{x})}\left[D\phi(x)\right]\,\exp\left[-\int_0^\beta d^4x_E\,\left(\frac{1}{2}\left((\partial_\tau\phi)^2+(\boldsymbol{\nabla}\phi)^2+m^2\phi^2\right)-j\phi\right)\right]\,.$$

$$(2.115)$$

The Gaussian path integral in this expression leads to:

$$Z_0[j;\beta] = \exp\left\{\frac{1}{2}\int_0^\beta d^4x_E\,d^4y_E\,j(x)\,G_E^0(x,y)\,j(y)\right\}\,, \quad (2.116)$$

where the free Euclidean propagator $G_E^0(x,y)$ is the inverse of the operator $m^2-\partial_\tau^2-\boldsymbol{\nabla}^2$ over the space of functions that are β-periodic in the imaginary time variable. Because of this periodicity, the "energy" variable, conjugate to the Euclidean time, is discrete:

$$\omega_n \equiv \frac{2\pi n}{\beta} \qquad (n\in\mathbb{Z})\,. \quad (2.117)$$

In terms of these energies, called *Matsubara frequencies*, the free Euclidean propagator in momentum space reads

$$\widetilde{G}_E^0(\omega_n, \mathbf{p}) = \frac{1}{\omega_n^2 + \mathbf{p}^2 + m^2} \; .$$ (2.118)

Note that the denominator cannot vanish, and therefore this propagator does not need an $i0^+$ prescription for being fully defined. Eqs. (2.114) and (2.116) lead to a perturbative expansion that can be cast into an expansion in terms of Feynman diagrams. The Feynman rules associated to these graphs are very similar to those already encountered when calculating scattering amplitudes, with only a few modifications:

$$\text{Propagators :} \qquad \frac{1}{\omega_n^2 + \mathbf{p}^2 + m^2} \; , \qquad (2.119)$$

$$\text{Vertices :} \qquad -\lambda \, 2\pi \, \delta\Big(\sum_i \omega_{n_i}\Big) \, (2\pi)^3 \delta\Big(\sum_i \mathbf{p}_i\Big) \, , \qquad (2.120)$$

$$\text{Loops :} \qquad \frac{1}{\beta} \sum_{n \in \mathbb{Z}} \int \frac{d^3\mathbf{p}}{(2\pi)^3} \; . \qquad (2.121)$$

In other words, the main difference with the usual perturbative expansion is that the energies are replaced by the discrete Matsubara frequencies, and that the loop integration on p^0 is replaced by a discrete sum over these frequencies.

Chapter 3

Fermion path integral

In the previous chapter, we have learned that the quantization of a scalar field may be performed by means of the path integral representation. This leads to a much more concise derivation of the generating functional, and of the free propagator, compared to the canonical approach. In this chapter, we will therefore seek a similar path integral formalism for other types of fields, in view of the functional quantization of a gauge theory such as QED (and later, of non-Abelian gauge theories, for which a canonical approach would be extremely difficult to implement).

3.1 Grassmann variables

3.1.1 Definition

In the functional formulation of a scalar field theory, we saw that time-ordered products of field operators correspond to the ordinary product of the integration variable in the integrand of the path integral (see the eq. (2.29)). Ultimately, a path integral representation of the time-ordered product of fermion field operators should allow the same, but with a catch: the T-product for fermions involves a minus sign when two operators are exchanged (see 1.176), that we need to be able to generate in the integrand of a would-be fermionic path integral. This can be achieved with *Grassmann numbers*[1], that are anti-commuting variables. In a sense, Grassmann numbers are the classical analogue of anti-commuting quantum operators. For a set of Grassmann variables ψ_i ($i = 1 \cdots N$), we have

$$\{\psi_i, \psi_j\} = 0 . \tag{3.1}$$

The linear space spanned by the ψ_i's is called a *Grassmann algebra*.

---

[1]Although we call them "numbers", they are not representable as scalar (e.g. real or complex) variables. A Grassmann number may be represented by a nilpotent 2×2 matrix, and the Grassmann algebra with N generators admits a representation in terms of $2^N \times 2^N$ matrices, that may be viewed as operators acting on the Hilbert space of N identical fermions of spin $1/2$ (of dimension 2^N since each spin has two states). For instance, when $N = 2$, one may represent the Grassmann numbers $\psi_{1,2}$ as

$$\psi_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} , \qquad \psi_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} .$$

3.1.2 Functions of a single Grassmann variable

Consider first the case $N = 1$. The square of a Grassmann number ψ is therefore zero, $\psi^2 = 0$, and by induction all higher powers of ψ are also zero. The Taylor expansion of a function of ψ is therefore limited to the first two terms,

$$f(\psi) = a + \psi b \,. \tag{3.2}$$

In general, we need to deal with functions $f(\psi)$ that are themselves commuting objects. Therefore, the coefficient a is an ordinary number, while b is another Grassmann number, $\{b, b\} = \{b, \psi\} = 0$. This implies that

$$f(\psi) = a + \psi b = a - b\psi \,. \tag{3.3}$$

Because of the non-commuting nature of b and ψ, we may define *left* and *right* derivatives, denoted by:

$$\overrightarrow{\partial}_\psi \, f(\psi) = b \quad , \qquad f(\psi) \, \overleftarrow{\partial}_\psi = -b \,. \tag{3.4}$$

One may define a linear mapping on functions of a Grassmann variable, that behaves for most purposes as an integration (although it is not an integral in the Lebesgue sense), called the *Berezin integral*. We require two basic axioms:

- Linearity :

$$\int d\psi \, \alpha \, f(\psi) = \alpha \int d\psi \, f(\psi) \,, \tag{3.5}$$

- The integral of a total derivative is zero:

$$\int d\psi \, \partial_\psi f(\psi) = 0 \,. \tag{3.6}$$

The only definition consistent with these requirements is

$$\int d\psi \, f(\psi) = b \,, \tag{3.7}$$

up to an overall constant that should be the same for all functions. Thus, integration and differentiation of functions of a Grassmann variable are essentially the same thing. In particular, the Berezin integral satisfies:

$$\int d\psi \, 1 = 0 \quad , \qquad \int d\psi \, \psi = 1 \,. \tag{3.8}$$

3.1.3 Functions of N Grassmann variables

Taylor expansion : We will denote collectively $\psi \equiv (\psi_1, \cdots, \psi_N)$. The most general function of N Grassmann variables can be written as

$$f(\psi) = \sum_{p=0}^{N} \frac{1}{p!} \psi_{i_1} \psi_{i_2} \cdots \psi_{i_p} \, C_{i_1 i_2 \cdots i_p} \,, \tag{3.9}$$

with implicit summations on the indices i_n. Terms of degree higher than N cannot exist because they would contain the square of at least one of the ψ_i's, and therefore be zero. We have chosen to write the Grassmann variables on the left of the coefficients in order to simplify the calculation of the left derivatives $\overrightarrow{\partial}_\psi$. Note that the last coefficient $C_{i_1 \cdots i_N}$ must be proportional to the Levi-Civita tensor:

$$C_{i_1 \cdots i_N} \equiv \gamma\, \epsilon_{i_1 \cdots i_N}\ . \tag{3.10}$$

Note that this last term can also be written as:

$$\tfrac{1}{N!}\, \psi_{i_1} \cdots \psi_{i_N}\, \gamma\, \epsilon_{i_1 \cdots i_N} = \psi_1 \cdots \psi_N\, \gamma\ . \tag{3.11}$$

Integration : In order to be consistent with eqs. (3.8), the integral of $f(\psi)$ over the N Grassmann variables $\psi_1, \cdots, \psi_N$, must be given by

$$\int d^N\psi\, f(\psi) = \gamma\ . \tag{3.12}$$

The terms of degree 0 through $N-1$ in the "Taylor expansion" of $f(\psi)$ cannot contribute to the integral, since at least one of the ψ_i is absent in these terms, and the integral over this ψ_i will therefore give zero. A somewhat more explicit formulation of an integral over N Grassmann variables is to write the measure as $d^N\psi \equiv d\psi_N\, d\psi_{N-1} \cdots d\psi_1$ (in this order), and to perform the N integrals successively, starting with the innermost one (i.e. $d\psi_1$). Therefore

$$\int d^N\psi\, \psi_1 \cdots \psi_N = \int d\psi_N \cdots \left(\int d\psi_2 \left(\underbrace{\int d\psi_1\, \psi_1}_{1} \right) \psi_2 \right) \cdots \psi_N = 1\ . \tag{3.13}$$

Change of variables : Let us now consider a linear change of variables:

$$\psi_i \equiv J_{ij}\, \theta_j\ , \tag{3.14}$$

where $\theta_1 \cdots \theta_N$ are N Grassmann variables. The last term of the expansion of $f(\psi)$, the only one relevant for integration, can be rewritten as

$$
\begin{aligned}
\psi_{i_1} \cdots \psi_{i_N}\, \epsilon_{i_1 \cdots i_N}\, \gamma &= \left(J_{i_1 j_1} \theta_{j_1} \right) \cdots \left(J_{i_N j_N} \theta_{j_N} \right) \epsilon_{i_1 \cdots i_N}\, \gamma \\
&= \det(J)\, \theta_{j_1} \cdots \theta_{j_N}\, \epsilon_{j_1 \cdots j_N}\, \gamma\ .
\end{aligned}
\tag{3.15}
$$

From this relationship, we conclude that

$$\underbrace{\int d^N\psi\, f(\psi)}_{\gamma} = \left[\det(J) \right]^{-1} \underbrace{\int d^N\theta\, f(\psi(\theta))}_{\det(J)\,\gamma}\ . \tag{3.16}$$

Thus, a change of variables in a Grassmann integral involves the inverse of the Jacobian that would normally appear in the same change of variables for a scalar integral.

Gaussian integrals : Let $\psi_1, \cdots, \psi_N$ be N Grassmann variables, and consider the following integral

$$I(\mathbf{M}) \equiv \int d^N \psi \, \exp\left(\tfrac{1}{2} \psi_i M_{ij} \psi_j\right), \tag{3.17}$$

where $\mathbf{M}$ is an antisymmetric $N \times N$ matrix made of commuting numbers (real or complex). Firstly, note that such an integral is non-zero only if N is even. For $N = 2$, this matrix is of the form

$$\mathbf{M} = \begin{pmatrix} 0 & \mu \\ -\mu & 0 \end{pmatrix}, \tag{3.18}$$

and the exponential in the integral reads

$$\exp\left(\tfrac{1}{2} \psi_i M_{ij} \psi_j\right) = 1 + \mu \, \psi_1 \psi_2 . \tag{3.19}$$

(Recall that functions of two Grassmann variables are in fact polynomials of degree two.) Therefore, in the case $N = 2$, the Gaussian integral (3.17) reads[2]

$$I(\mathbf{M}) = \mu = \left[\det(\mathbf{M})\right]^{1/2} . \tag{3.20}$$

In the case of a general even N, the matrix $\mathbf{M}$ may be written in the following block diagonal form,

$$\mathbf{M} = \mathbf{Q} \underbrace{\begin{pmatrix} 0 & \mu_1 & & & \\ -\mu_1 & 0 & & & \\ & & 0 & \mu_2 & \\ & & -\mu_2 & 0 & \\ & & & & \ddots \end{pmatrix}}_{\mathbf{D}} \mathbf{Q}^\top , \tag{3.21}$$

where $\mathbf{Q}$ is a special[3] orthogonal matrix. Defining $\mathbf{Q}^\top \psi \equiv \theta$, we have

$$I(\mathbf{M}) = \left[\det(\mathbf{Q})\right]^{-1} \underbrace{\int d^N \theta \, \exp\left(\tfrac{1}{2} \theta^\top \mathbf{D} \theta\right)}_{\mu_1 \mu_2 \cdots = [\det(\mathbf{D})]^{1/2}} . \tag{3.22}$$

But since $\det(\mathbf{Q}) = +1$, this becomes

$$I(\mathbf{M}) = \left[\det(\mathbf{D})\right]^{1/2} = \left[\det(\mathbf{M})\right]^{1/2} . \tag{3.23}$$

Contrast this with the result of a Gaussian integral in the case of ordinary real variables, eq. (2.67), where the square root of the determinant appeared in the denominator.

It is often necessary to perform a Gaussian integral in the presence of a source that shifts the minimum of the quadratic form in the exponential,

$$I(\mathbf{M}, \eta) \equiv \int d^N \psi \, \exp\left(\tfrac{1}{2} \psi_i M_{ij} \psi_j + \eta_i \psi_i\right), \tag{3.24}$$

where η is a set of N Grassmann sources. By introducing the new Grassmann variable $\psi_i' \equiv \psi_i - M_{ij}^{-1} \eta_j$, this integral falls back to the previous type, and we obtain:

$$I(\mathbf{M}, \eta) = \left[\det(\mathbf{M})\right]^{1/2} \exp\left(-\tfrac{1}{2} \eta^\top \mathbf{M}^{-1} \eta\right) . \tag{3.25}$$

---

[2]The determinant of a real antisymmetric matrix of even size is the square of its *Pfaffian* and is therefore positive.
[3]Orthogonal matrices have determinant $+1$ or -1. The special orthogonal matrices are the subgroup of those that have determinant $+1$.

Gaussian integral with $2N$ **variables :**   Another useful type of Gaussian integral is

$$J(\mathbf{M}) \equiv \int d^N\xi d^N\psi \, \exp\left(\psi_i M_{ij}\xi_j\right) , \tag{3.26}$$

where $\mathbf{M}$ is an $N \times N$ matrix of commuting numbers, and ψ and ξ are independent Grassmann variables. The only non-zero contribution to this integral comes from the term of order N in the Taylor expansion of the exponential,

$$
\begin{aligned}
J(\mathbf{M}) &= \frac{1}{N!} \int d^N\xi d^N\psi \, \left(\psi_{i_1} M_{i_1 j_1} \xi_{j_1}\right) \cdots \left(\psi_{i_N} M_{i_N j_N} \xi_{j_N}\right) \\
&= \frac{(-1)^{\frac{N(N-1)}{2}}}{N!} \int d^N\xi d^N\psi \, \left(\psi_{i_1} \cdots \psi_{i_N}\right)\left(\xi_{j_1} \cdots \xi_{j_N}\right) M_{i_1 j_1} \cdots M_{i_N j_N} \\
&= \frac{(-1)^{\frac{N(N-1)}{2}}}{N!} \, \epsilon_{i_1 \cdots i_N} \, \epsilon_{j_1 \cdots j_N} \, M_{i_1 j_1} \cdots M_{i_N j_N} .
\end{aligned}
\tag{3.27}
$$

In the second line, we have reordered the Grassmann variables in order to bring all the ψ_i's on the left, and the sign in the prefactor keeps track of the number of permutations that are necessary to achieve this. To give a non-zero result, the indices $\{i_n\}$ and $\{j_n\}$ must be permutations of $[1 \cdots N]$:

$$
\begin{aligned}
J(\mathbf{M}) &= \frac{(-1)^{\frac{N(N-1)}{2}}}{N!} \sum_{\sigma,\rho\in\mathfrak{S}_n} \epsilon(\sigma)\epsilon(\rho) \, M_{\sigma(1)\rho(1)} \cdots M_{\sigma(N)\rho(N)} \\
&= \frac{(-1)^{\frac{N(N-1)}{2}}}{N!} \sum_{\sigma,\tau\in\mathfrak{S}_n} \epsilon(\sigma)\epsilon(\tau\sigma) \, M_{1\tau(1)} \cdots M_{N\tau(N)}
\end{aligned}
\tag{3.28}
$$

where $\epsilon(\sigma)$ is the signature of the permutation σ, and with $\tau \equiv \rho\sigma^{-1}$ in the second line. Using $\epsilon(\sigma)\epsilon(\tau\sigma) = \epsilon(\tau)$, this becomes:

$$J(\mathbf{M}) = (-1)^{\frac{N(N-1)}{2}} \underbrace{\left(\frac{1}{N!}\sum_{\sigma\in\mathfrak{S}_n} 1\right)}_{1} \underbrace{\sum_{\tau\in\mathfrak{S}_n} \epsilon(\tau) \, M_{1\tau(1)} \cdots M_{N\tau(N)}}_{\det(\mathbf{M})} . \tag{3.29}$$

Note that this overall sign may be absorbed into a reordering of the measure, since:

$$d^N\xi d^N\psi = (-1)^{\frac{N(N-1)}{2}} \, d\xi_N d\psi_N \cdots d\xi_1 d\psi_1 . \tag{3.30}$$

Therefore, we have

$$\int d\xi_N d\psi_N \cdots d\xi_1 d\psi_1 \, \exp\left(\psi_i M_{ij}\xi_j\right) = \det\left(\mathbf{M}\right) . \tag{3.31}$$

3.1.4   Complex Grassmann variables

Now, let us define *complex Grassmann variables*, from two of the previously defined Grassmann variables ψ and ξ:

$$\chi \equiv \frac{\psi + i\xi}{\sqrt{2}} \quad , \qquad \overline{\chi} \equiv \frac{\psi - i\xi}{\sqrt{2}} . \tag{3.32}$$

Conversely, we have

$$\psi = \frac{\overline{\chi} + \chi}{\sqrt{2}} \quad , \qquad \xi = \frac{i\,(\overline{\chi} - \chi)}{\sqrt{2}} \, , \tag{3.33}$$

and the integrations over these variables are related by

$$d\psi_2 d\psi_1 = i\,d\chi d\overline{\chi} \, ,$$
$$\psi_1 \psi_2 = -i\,\overline{\chi}\chi \, ,$$
$$\int d\chi d\overline{\chi}\ \overline{\chi}\chi = \int d\psi_2 d\psi_1\ \psi_1 \psi_2 = 1 \, . \tag{3.34}$$

From this, we obtain

$$\int d\chi d\overline{\chi}\ \exp\left(\mu\,\overline{\chi}\chi\right) = \mu \, , \tag{3.35}$$

that can be generalized into

$$\int d\chi_N d\overline{\chi}_N \cdots d\chi_1 d\overline{\chi}_1\ \exp(\overline{\chi}^{\mathsf{T}} M\chi) = \det\left(M\right) \, . \tag{3.36}$$

In the presence of sources η and $\overline{\eta}$, we obtain the following Gaussian integral:

$$\int d\chi_N d\overline{\chi}_N \cdots d\chi_1 d\overline{\chi}_1\ \exp\left(\overline{\chi}^{\mathsf{T}} M\chi + \overline{\eta}^{\mathsf{T}}\chi + \overline{\chi}^{\mathsf{T}}\eta\right) = \det\left(M\right)\ \exp\left(-\overline{\eta}^{\mathsf{T}} M^{-1}\eta\right) \, . \tag{3.37}$$

3.2 Path integral for fermions

We now have all the ingredients for constructing a path integral for spin $1/2$ fermions. Let us work our way backwards, starting from a generating functional that generates the free time-ordered products of spinors,

$$Z_0[\overline{\eta}, \eta] \equiv \exp\left\{ -\int d^4x d^4y\ \overline{\eta}(x) S_F^0(x, y)\eta(y)\right\} \, , \tag{3.38}$$

where $S_F^0(x, y)$ is the free Dirac time-ordered propagator and $\overline{\eta}$ and η are a pair of complex Grassmann-valued sources. Indeed, we have

$$\frac{\overrightarrow{\delta}}{i\delta\overline{\eta}(x)}\ Z_0[\overline{\eta}, \eta]\ \frac{\overleftarrow{\delta}}{i\delta\eta(y)}\bigg|_{\overline{\eta}=\eta=0} = S_F^0(x, y) \, . \tag{3.39}$$

Taking more than two derivatives (but with an equal number of derivatives with respect to $\overline{\eta}$ and with respect to η) will lead to all the contributions in the free time-ordered product of spinors, with the correct signs to account for their anti-commuting nature. Note that using Grassmann-valued sources was necessary in order to get these signs.

Then, by comparing eqs. (3.37) and (3.38), we can represent this free generating function as a path integral over Grassmann variables:

$$\begin{aligned} Z_0[\overline{\eta}, \eta] &= \int \left[D\psi(x)D\overline{\psi}(x)\right]\ \exp\left\{i\int d^4x\left(\overline{\psi}(x)(i\slashed{\partial} - m)\psi(x)\right.\right. \\ &\qquad\qquad\qquad\qquad\qquad \left.\left. +\overline{\eta}(x)\psi(x) + \overline{\psi}(x)\eta(x)\right)\right\} \\ &= \int \left[D\psi(x)D\overline{\psi}(x)\right]\ e^{iS[\overline{\psi}, \psi]}\ e^{i\int d^4x\left(\overline{\eta}(x)\psi(x) + \overline{\psi}(x)\eta(x)\right)} \, . \end{aligned} \tag{3.40}$$

We have ignored the determinant, since it is independent of the sources. Instead, one simply adjusts the normalization of the generating functional so that $Z_0[0,0] = 1$. The second line shows that the path integral formulation of a field theory of spin $1/2$ fermions takes the same form as that of scalar fields, provided we use Grassmann variables instead of commuting c-numbers.

In quantum electrodynamics, fermions interact only by their minimal coupling to the photon fields,

$$\mathcal{L}_I = -i\,e\,\overline{\psi}\gamma^\mu\,A_\mu\,\psi \;. \tag{3.41}$$

As in the scalar case, this interaction can be factored out of the generating functional, by writing:

$$Z[\overline{\eta},\eta] = \exp\left\{ -ie\int d^4x\, A_\mu(x)\frac{\overrightarrow{\delta}}{i\delta\overline{\eta}(x)}\gamma^\mu\frac{\overleftarrow{\delta}}{i\delta\eta(x)}\right\} Z_0[\overline{\eta},\eta]\;. \tag{3.42}$$

Here, we are treating the photon field as a fixed background. When we consider the path integral representation of dynamical photons in the next section, the $A_\mu(x)$ inside the exponential will also be replaced by a functional derivative.

3.3 Path integral for photons

3.3.1 Problems with the naive path integral

In the case of photons, the difficulties encountered in the path integral formulation are of a different nature. Since photons are bosons, we expect that they can be represented by a functional integration over commuting functions $A_\mu(x)$. But the gauge invariance of the theory implies that there is an unavoidable redundancy in this representation: the naive path integral over $[DA_\mu(x)]$ would integrate over infinitely many copies of the same physical configurations. Therefore, we need a way to cut through this redundancy, which is achieved by *gauge fixing*.

In order to better see the nature of this difficulty, let us assume that we can treat $A_\mu(x)$ as four scalar fields, and write the following path integral,

$$Z_0[j^\mu] \equiv \int [DA_\mu(x)]\; \exp\left\{i\int d^4x\,\left(-\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + j^\mu A_\mu\right)\right\}\;. \tag{3.43}$$

This is a Gaussian integral, since $F^{\mu\nu}F_{\mu\nu}$ is quadratic in the field A_μ,

$$\begin{aligned}
-\frac{1}{4}\int d^4x\, F^{\mu\nu}F_{\mu\nu} &= -\frac{1}{4}\int d^4x\,\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right)\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right)\\
&= +\frac{1}{2}\int d^4x\, A^\mu\left(g_{\mu\nu}\Box - \partial_\mu\partial_\nu\right)A^\nu\\
&= -\frac{1}{2}\int \frac{d^4k}{(2\pi)^4}\,\widetilde{A}^\mu(k)\left(g_{\mu\nu}k^2 - k_\mu k_\nu\right)\widetilde{A}^\nu(-k)\;. \tag{3.44}
\end{aligned}$$

Performing this Gaussian integral requires the inverse of the object $g_{\mu\nu}k^2 - k^\mu k^\nu$, that one may seek as a linear combination of the metric tensor $g^{\mu\nu}$ and $k^\mu k^\nu/k^2$, i.e. we are looking for coefficients α and β such that:

$$\underbrace{\left(g_{\mu\nu}k^2 - k_\mu k_\nu\right)\left(\alpha\,g^{\nu\rho} + \beta\,\tfrac{k^\nu k^\rho}{k^2}\right)}_{\alpha\,k^2\delta^\rho_\mu - \alpha\,k_\mu k^\rho} = \delta^\rho_\mu\;. \tag{3.45}$$

This equation has clearly no solution, and therefore it is impossible to invert $g_{\mu\nu}k^2 - k^\mu k^\nu$. This means that some eigenvalues of this operator are zero, and that the quadratic form $\widetilde{A}^\mu(k)\big(g_{\mu\nu}k^2 - k_\mu k_\nu\big)\widetilde{A}^\nu(-k)$ has flat directions. Along these flat directions, the exponential in the path integral (3.43) does not decrease, which spoils its convergence. These flat directions correspond to the projection of $\widetilde{A}^\mu(k)$ along k^μ. Note that they also do not contribute to the linear term $j^\mu A_\mu$, for a conserved current that satisfies $\partial_\mu j^\mu = 0$. Therefore, one should not integrate over these components of A^μ in eq. (3.43).

3.3.2 Path integral in Landau gauge

A simple way out it to decompose A^μ as follows:

$$
\begin{aligned}
A^\mu &= A^\mu_\perp + A^\mu_\parallel\,,\\
\widetilde{A}^\mu_\perp(k) &\equiv \left(g^{\mu\nu} - \frac{k^\mu k^\nu}{k^2}\right)\widetilde{A}_\nu(k)\,,\\
\widetilde{A}^\mu_\parallel(k) &\equiv \left(\frac{k^\mu k^\nu}{k^2}\right)\widetilde{A}_\nu(k)\,.
\end{aligned}
\tag{3.46}
$$

The functional measure can be factorized as follows

$$
\big[DA^\mu\big] = \big[DA^\mu_\perp\big]\big[DA^\mu_\parallel\big]\,,
\tag{3.47}
$$

and since nothing depends on $A^\mu_\parallel$ the photon kinetic term, we can write

$$
\begin{aligned}
Z_0[j^\mu] &\equiv \int \big[DA^\mu_\parallel(x)\big]\,\exp\left\{i\int d^4x\, j_\mu A^\mu_\parallel\right\}\\
&\quad\times \int \big[DA^\mu_\perp(x)\big]\,\exp\left\{i\int d^4x\,\big(-\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + j_\mu A^\mu_\perp\big)\right\}\,.
\end{aligned}
\tag{3.48}
$$

By integrating by parts the argument of the exponential in the integral on $A^\mu_\parallel$, we obtain a delta function of $\partial_\mu j^\mu$. Thus, for external currents that obey the continuity equation $\partial_\mu j^\mu = 0$, this prefactor is an infinite constant that can be ignored. When restricted to the subspace of $A^\mu_\perp$, the operator $g_{\mu\nu}k^2 - k^\mu k^\nu$ is invertible, and we can now perform the Gaussian integral, to obtain:

$$
Z_0[j^\mu] = \exp\left\{-\frac{1}{2}\int d^4x\,d^4y\, j_\mu(x)\, G^{0\ \mu\nu}_{\scriptscriptstyle F}(x,y)\, j_\nu(y)\right\}\,,
\tag{3.49}
$$

with the free photon propagator in momentum space given by

$$
G^{0\ \mu\nu}_{\scriptscriptstyle F}(p) \equiv \frac{-i}{p^2 + i0^+}\left(g^{\mu\nu} - \frac{p^\mu p^\nu}{p^2}\right)\,.
\tag{3.50}
$$

(We have introduced the $i0^+$ prescription that selects the ground state at $x^0 \to \pm\infty$, using the same argument as in the section 2.3.3.) The procedure used here is equivalent to imposing the gauge fixing condition $\partial_\mu A^\mu = 0$, called *Lorenz gauge* or *Landau gauge*. As one can see, the resulting propagator (3.50) differs from the Coulomb gauge propagator given in eq. (1.207).

3.3.3 General covariant gauges

All gauge fixings amount to constrain in some way the quantity $\partial_\mu A^\mu$, since it does not appear in the integrand of the photon path integral. Instead of imposing $\partial_\mu A^\mu = 0$, one may instead impose the more general condition

$$\partial_\mu A^\mu(x) = \omega(x) \,, \tag{3.51}$$

where $\omega(x)$ is some arbitrary function of space-time. This can be done by introducing a functional delta function, $\delta[\partial_\mu A^\mu - \omega]$, inside the path integral. However, the introduction of the function $\omega(x)$ breaks Lorentz invariance. To mitigate this problem, one integrates over all the functions $\omega(x)$, with a Gaussian weight. This amounts to defining the generating functional as follows[4],

$$
\begin{aligned}
Z_0[j^\mu] \ \equiv \ & \int \left[D\omega(x)\right] \, \exp\left\{-i\frac{\xi}{2}\int d^4x \, \omega^2(x)\right\} \\
& \times \int \left[DA_\mu(x)\right] \, \delta\left[\partial_\mu A^\mu - \omega\right] \, \exp\left\{i\int d^4x \left(-\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + j^\mu A_\mu\right)\right\} \,,
\end{aligned}
\tag{3.52}
$$

where ξ is an arbitrary constant. Performing the integration on $\omega(x)$ thanks to the delta functional, and integrating by parts, this becomes

$$Z_0[j^\mu] = \int \left[DA_\mu(x)\right] \, \exp\left\{i\int d^4x \left(\tfrac{1}{2}A^\mu(g_{\mu\nu}\Box - (1-\xi)\partial_\mu\partial_\nu)A^\nu + j^\mu A_\mu\right)\right\} . \tag{3.53}$$

From this formula, a standard Gaussian integration tells us that the corresponding photon propagator in momentum space should be the inverse of

$$i\left(g_{\mu\nu}p^2 - (1-\xi)p_\mu p_\nu\right) . \tag{3.54}$$

Looking for an inverse of the form $\alpha \, g^{\nu\rho} + \beta \, \frac{p^\nu p^\rho}{p^2}$, we find

$$G_F^{0\,\mu\nu}(p) = \frac{-i\,g^{\mu\nu}}{p^2 + i0^+} + \frac{i}{p^2 + i0^+}\left(1 - \frac{1}{\xi}\right)\frac{p^\mu p^\nu}{p^2} \,. \tag{3.55}$$

The gauge fixing parameter ξ appears in the propagator, but only in the term proportional to $p^\mu p^\nu$. Thanks to the Ward-Takahashi identities, it does not have any incidence on physical results, provided that all the external charged particles are on mass-shell. The Landau gauge of the previous subsection corresponds to $\xi \to \infty$. Another popular choice is the *Feynman gauge*, obtained for $\xi = 1$,

$$G_F^{0\,\mu\nu}(p) \underset{\xi=1}{=} \frac{-i\,g^{\mu\nu}}{p^2 + i0^+} \,. \tag{3.56}$$

Note that one could also introduce a non Lorentz covariant condition inside the delta function, such as $\delta[\partial_i A^i - \omega]$, in order to derive the photon propagator in Coulomb gauge via the path integral.

---

[4]Since the argument of the delta function is linear in the variable $A_\parallel^\mu$ that does not appear in the integrand, we do not need a Jacobian. It is possible to impose non-linear gauge conditions of the form $\delta[F(\partial_\mu A^\mu) - \omega]$, but this should be done by writing the path integral as follows

$$\int \left[D\omega(x)\right] \, \exp\left\{-i\frac{\xi}{2}\int d^4x \, \omega^2(x)\right\} \int \left[DA_\mu(x)\right] \underbrace{F'(\partial_\mu A^\mu)}_{\text{Jacobian}} \, \delta\left[F(\partial_\mu A^\mu) - \omega\right] \cdots$$

In general, the Jacobian cannot be ignored since it depends on the gauge field, but it can be expressed in terms of *ghost fields*. Doing this would be an useless complication in QED, but is an essential step in the quantization of non-Abelian gauge theories.

3.4 Schwinger-Dyson equations

3.4.1 Functional derivation

Consider a Lagrangian density $\mathcal{L}(\phi, \partial_\mu \phi)$ (ϕ may be a collection of fields, but we do not write any index on it to keep the notation light), and $\mathcal{S} \equiv \int_x \mathcal{L}$ the corresponding action. The generating functional of time-ordered products of fields has the following path integral representation:

$$Z[j] = \int \big[D\phi(x) \big] \; e^{i\mathcal{S}[\phi] + i\int j\phi} \; . \tag{3.57}$$

In the right hand side, $\phi(x)$ should be viewed as a dummy integration variable, and the result of the integral should be unmodified if we change $\phi(x) \to \phi(x) + \delta\phi(x)$. This translates into

$$0 = \delta Z[j] = i \int \big[D\phi(x) \big] \; e^{i\mathcal{S}[\phi] + i\int j\phi} \left\{ \int d^4 x \, \delta\phi(x) \Big(j(x) + \frac{\delta\mathcal{S}}{\delta\phi(x)} \Big) \right\} . \tag{3.58}$$

Taking n functional derivatives of this identity with respect to $ij(x_1),\ldots,ij(x_n)$ and setting then j to zero gives:

$$\begin{aligned}
0 &= \int \big[D\phi(x) \big] \; e^{i\mathcal{S}[\phi]} \int d^4 x \, \delta\phi(x) \Big\{ i \, \phi(x_1)\cdots\phi(x_n) \frac{\delta\mathcal{S}}{\delta\phi(x)} \\
&\qquad\qquad + \sum_{i=1}^{n} \delta(x - x_i) \prod_{j \neq i} \phi(x_j) \Big\} .
\end{aligned} \tag{3.59}$$

Since in this discussion the variation $\delta\phi(x)$ is arbitrary, this implies the following identities

$$0 = \int \big[D\phi(x) \big] \; e^{i\mathcal{S}[\phi]} \left\{ i \, \phi(x_1)\cdots\phi(x_n) \frac{\delta\mathcal{S}}{\delta\phi(x)} + \sum_{i=1}^{n} \delta(x - x_i) \prod_{j \neq i} \phi(x_j) \right\} , \tag{3.60}$$

known as the *Schwinger-Dyson equations* (here written in functional form). For instance, in the case of a scalar field theory with a ϕ^4 interaction term, this leads to

$$\begin{aligned}
&i\big(\Box_x + m^2\big)\big\langle 0_{\text{out}}\big| T \, \phi(x_1)\cdots\phi(x_n)\phi(x)\big|0_{\text{in}}\big\rangle \\
&+i\tfrac{\lambda}{3!}\big\langle 0_{\text{out}}\big| T \, \phi(x_1)\cdots\phi(x_n)\phi^3(x)\big|0_{\text{in}}\big\rangle = \sum_{i=1}^{n} \delta(x - x_i) \big\langle 0_{\text{out}}\big| T \prod_{j \neq i} \phi(x_j)\big|0_{\text{in}}\big\rangle .
\end{aligned} \tag{3.61}$$

(We have used the remark following eq. (2.30) in order to let the operator $\Box + m^2$ act also on the step functions that order the operators in the time-ordered product.) If we convolute this equation with the free Feynman propagator (i.e. the inverse of the operator $\Box_x + m^2$), the above Schwinger-Dyson equation can be represented diagrammatically as follows:

$$\tag{3.62}$$

84

The Schwinger-Dyson equations have several simple consequences. When applied to a free theory ($\lambda = 0$) in the case $n = 1$, we get

$$\left(\Box_x + m^2\right)\langle 0_{\text{out}}|T\,\phi(x_1)\phi(x)|0_{\text{in}}\rangle = -i\delta(x - x_1)\,, \tag{3.63}$$

which is nothing but the equation of motion satisfied by the Feynman propagator. In the general case, if x differs from all the x_i's, we obtain

$$\left(\Box_x + m^2\right)\langle 0_{\text{out}}|T\,\phi(x_1)\cdots\phi(x_n)\phi(x)|0_{\text{in}}\rangle$$
$$+\tfrac{\lambda}{3!}\langle 0_{\text{out}}|T\,\phi(x_1)\cdots\phi(x_n)\phi^3(x)|0_{\text{in}}\rangle = 0\,. \tag{3.64}$$

Thus, in a certain sense[5], we can say that time-ordered products of fields satisfy the Euler-Lagrange equation of motion.

3.4.2 Schwinger-Dyson equations and conserved currents

The functional derivative of the action S with respect to $\phi(x)$ is given by

$$\frac{\delta S}{\delta\phi(x)} = \frac{\partial\mathcal{L}}{\partial\phi(x)} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi(x))}\,. \tag{3.65}$$

When we equate this to zero, we recover the Euler-Lagrange equation of motion. Under an infinitesimal variation $\delta\phi(x)$ of the field, the Lagrangian density varies by

$$
\begin{aligned}
\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi(x)}\,\delta\phi(x) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi(x))}\,\partial_\mu(\delta\phi(x)) \\
&= \partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi(x))}\,\delta\phi(x)\Big) + \frac{\delta S}{\delta\phi(x)}\,\delta\phi(x)\,.
\end{aligned} \tag{3.66}
$$

When the variation $\delta\phi(x)$ corresponds to a symmetry of the Lagrangian, we have $\delta\mathcal{L} = 0$, and therefore

$$\frac{\delta S}{\delta\phi(x)}\,\delta\phi(x) = -\partial_\mu\Big(\underbrace{\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi(x))}\,\delta\phi(x)}_{J^\mu(x)}\Big)\,, \tag{3.67}$$

where J^μ is the *Noether current* associated to this continuous symmetry. In the classical theory, this current is conserved, i.e. $\partial_\mu J^\mu = 0$, if the fields obey the Euler-Lagrange equation of motion. The Schwinger-Dyson equations provide a quantum analogue of this conservation law, at the level of the expectation values of time-ordered products of fields. In eq. (3.59), we can replace $\delta\phi(\delta S/\delta\phi)$ by $-\partial_\mu J^\mu$. When the resulting identity is rewritten in terms of operators, the derivative ∂_μ should go outside the time-ordering, and we obtain

$$\partial_\mu\langle 0_{\text{out}}|T\,J^\mu(x)\phi(x_1)\cdots\phi(x_n)|0_{\text{in}}\rangle$$
$$+i\sum_{i=1}^n \delta(x - x_i)\,\langle 0_{\text{out}}|T\,\delta\phi(x)\prod_{j\neq i}\phi(x_j)|0_{\text{in}}\rangle = 0\,. \tag{3.68}$$

---

[5]I.e., up to the terms in $\delta(x - x_i)$ that may appear in the right hand side, called *contact terms*. These contact terms in fact take care of the action of the time derivative on the theta functions of the time ordering operator T.

Therefore, when a Noether current operator is inserted inside a time-ordered product, it satisfies the continuity equation up to contact terms (coming from the action of ∂_0 on the theta functions of the T product). Eq. (3.68) is a generalization of the Ward-Takahashi identities, already discussed in the context of electric charge conservation.

Note that in some cases, a continuous symmetry does not leave the Lagrangian density invariant, but modifies it by a total derivative,

$$\delta\mathcal{L} = \partial_\mu K^\mu \,, \tag{3.69}$$

so that only the action is invariant. There is still a conserved current, given by

$$J^\mu(x) \equiv \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi(x))}\,\delta\phi(x) - K^\mu(x) \,. \tag{3.70}$$

This however does not modify eqs. (3.68).

3.5 Quantum anomalies

3.5.1 General considerations

It may happen that some symmetries of the Lagrangian (i.e. symmetries of the classical theory) are broken by quantum corrections. This phenomenon is called a *quantum anomaly*. One way this may appear is via the introduction of a regularization (e.g. a cutoff), whose effect leaves an imprint on physical results even after the cutoff has been taken to infinity. Here we will adopt a functional point of view on this issue. In the previous section, a crucial point in the derivation of the Schwinger-Dyson equations is that the functional measure must be invariant under the symmetry under consideration. Quantum anomalies may be viewed as an obstruction in defining a functional measure which is invariant under certain symmetries, e.g. axial symmetry.

Let us consider a set of fermion fields $\psi_n(x)$, that we encapsulate into a multiplet denoted $\psi(x)$, and assume that they interact with a gauge potential $A_\mu^a(x)$ in a non-chiral way (this is the case of electromagnetic interactions and of strong interactions). Consider now the following transformation of the fermion fields:

$$\psi(x) \to U(x)\psi(x) \,. \tag{3.71}$$

The Hermitic conjugate of ψ transforms as:

$$\psi^\dagger(x) \to \psi^\dagger(x)U^\dagger(x) \,, \tag{3.72}$$

so that we have

$$\overline{\psi}(x) \equiv \psi^\dagger(x)\gamma^0 \to \psi^\dagger(x)U^\dagger(x)\gamma^0 = \overline{\psi}(x)\gamma^0 U^\dagger(x)\gamma^0 \,. \tag{3.73}$$

Since they are Grassmann variables, the measure should be transformed with the inverse of the determinant of the transformation. Since the transformation under consideration is local in x, it reads

$$\left[D\psi D\overline{\psi}\right] \to \frac{1}{\det\left(\mathcal{U}\right)\det\left(\overline{\overline{\mathcal{U}}}\right)}\left[D\psi D\overline{\psi}\right] \,, \tag{3.74}$$

where the matrices $\mathcal{U}$ and $\overline{\mathcal{U}}$ carry both indices for the fermion species and space-time indices:

$$\begin{aligned}\mathcal{U}_{xm,yn} &\equiv U_{mn}(x)\,\delta(x-y) \,, \\ \overline{\mathcal{U}}_{xm,yn} &\equiv (\gamma^0 U^\dagger(x)\gamma^0)_{mn}\,\delta(x-y) \,.\end{aligned} \tag{3.75}$$

86

3.5.2 Non-chiral transformations

Let us consider the following transformation:

$$U(x) = e^{i\alpha(x)t} \, , \tag{3.76}$$

where $\alpha(x) \in \mathbb{R}$ and where t is a Hermitean matrix that does not contain $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$. Therefore:

$$U^\dagger(x) = e^{-i\alpha(x)t} \, , \tag{3.77}$$

and

$$
\begin{aligned}
(\overline{\mathcal{U}}\mathcal{U})_{xm,yn} &= \int d^4z \sum_p \overline{\mathcal{U}}_{xm,zp} \, \mathcal{U}_{zp,yn} \\
&= \int d^4z \, \delta(x-z)\delta(z-y) \sum_p \left(e^{-i\alpha(z)t}\right)_{mp} \left(e^{i\alpha(z)t}\right)_{pn} \\
&= \delta_{mn}\delta(x-y) \, .
\end{aligned} \tag{3.78}
$$

Thus $\overline{\mathcal{U}}\mathcal{U} = 1$, which implies $\det\mathcal{U} \det\overline{\mathcal{U}} = 1$, and the fermion measure is invariant under this kind of transformations. This means that this symmetry does not exhibit quantum anomalies.

3.5.3 Chiral transformations

Let us now define the right-handed and left-handed projections of a spinor,

$$\psi_{\text{R}} \equiv \left(\frac{1+\gamma^5}{2}\right)\psi \quad , \qquad \psi_{\text{L}} \equiv \left(\frac{1-\gamma^5}{2}\right)\psi \, , \tag{3.79}$$

and consider a transformation that acts differently on these two components:

$$U(x) = e^{i\alpha(x)\gamma^5 t} \, , \tag{3.80}$$

where t is again a Hermitean matrix. Such transformations are called *chiral transformations*. The matrix $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$ satisfies

$$
\begin{aligned}
\left(\gamma^5\right)^2 &= 1 \, , \\
\gamma^{5\,\dagger} &= \gamma^5 \, , \\
\{\gamma^5, \gamma^0\} &= 0 \, ,
\end{aligned} \tag{3.81}
$$

which implies:

$$\gamma^0 U^\dagger(x)\gamma^0 = \gamma^0 e^{-i\alpha(x)\gamma^5 t}\gamma^0 = e^{i\alpha(x)\gamma^5 t} = U(x) \, . \tag{3.82}$$

Thus $\overline{\mathcal{U}} = \mathcal{U}$, and $\det\mathcal{U} = \det\overline{\mathcal{U}}$. Unless this determinant is equal to one, the measure is not invariant and transforms according to:

$$\left[D\psi D\overline{\psi}\right] \rightarrow \frac{1}{(\det\mathcal{U})^2} \left[D\psi D\overline{\psi}\right] \, . \tag{3.83}$$

Consider an infinitesimal transformation of the form given in eq. (3.80). We can write:

$$(\mathcal{U} - 1)_{xm,yn} = i\,\alpha(x)(\gamma^5 t)_{mn}\,\delta(x - y)\,.\tag{3.84}$$

In order to calculate $\left(\det\mathcal{U}\right)^{-2}$, we use the formula[6]:

$$(\det\mathcal{U})^{-2} = e^{-2\,\mathrm{tr}\,\ln\mathcal{U}}\,.\tag{3.85}$$

In the present case, we have:

$$
\begin{aligned}
(\det\mathcal{U})^{-2} &= \exp\left[-2\,\mathrm{tr}\,\ln\left(1 + i\alpha(x)\,\gamma^5\,t\,\delta(x-y)\right)\right]\\
&\underset{\alpha\ll 1}{\approx} \exp\left[-2\,i\,\mathrm{tr}\left(\alpha(x)\,\gamma^5\,t\,\delta(x-y)\right)\right]\\
&= \exp\left[i\int d^4x\,\alpha(x)\mathcal{A}(x)\right]\,,
\end{aligned}\tag{3.86}
$$

with a function $\mathcal{A}(x)$ whose formal expression is

$$\mathcal{A}(x) \equiv -2\,\mathrm{tr}\left(\gamma^5 t\right)\delta(x - x)\,.\tag{3.87}$$

In this equation, the trace symbol tr denotes both a trace on the indices carried by the Dirac matrices and a trace on the fermion species. In terms of this function, the measure transforms as

$$\left[D\psi D\overline{\psi}\right] \to e^{i\int d^4x\,\alpha(x)\mathcal{A}(x)}\left[D\psi D\overline{\psi}\right]\,.\tag{3.88}$$

The fact that this measure is not invariant under the transformation (3.80) implies that there exists fermion loop corrections that break the invariance under chiral transformations, even if the Dirac Lagrangian itself is invariant (this is the case when one considers a global transformation, i.e. a constant $\alpha(x)$, and the fermions are massless). The prefactor that alters the measure can be absorbed into a redefinition of the Lagrangian,

$$\mathcal{L}(x) \to \mathcal{L}(x) + \alpha(x)\mathcal{A}(x)\,.\tag{3.89}$$

All happens as if the Lagrangian itself was not invariant under this transformation. If one integrates out the fermion fields in order to obtain an effective theory for the other fields, the term in $\alpha(x)\mathcal{A}(x)$ must be included in the Lagrangian of this effective theory in order to correctly account for the quantum anomalies.

3.5.4   Calculation of $\mathcal{A}(x)$

At first sight, the expression (3.87) of the anomaly function $\mathcal{A}(x)$ is very poorly defined: the trace is zero, but it is multiplied by an infinite $\delta(0)$. In order to manipulate finite expressions, we must first regularize the delta function. This can be done by writing:

$$\mathcal{A}(x) = -2\lim_{y\to x, M\to+\infty}\mathrm{tr}\left\{\gamma^5\,t\,\mathcal{F}\left(-\frac{\slashed{D}_x^2}{M^2}\right)\right\}\delta(x - y)\,,\tag{3.90}$$

---

[6]If the λ_i are the eigenvalues of $\mathcal{U}$, we have:

$$\det\mathcal{U} = \prod_i\lambda_i = \exp\left(\sum_i\ln\lambda_i\right) = e^{\mathrm{tr}\,\ln\mathcal{U}}\,.$$

where $\not{D}_x$ is the *Dirac operator*[7]

$$\not{D}_x \equiv \gamma^\mu \left(\partial_\mu - i\, g\, t^a A_\mu^a(x) \right) , \tag{3.91}$$

and where $\mathcal{F}(s)$ is a function such that

$$\begin{aligned}
\mathcal{F}(0) &= 1 , \\
\mathcal{F}(+\infty) &= 0 , \\
s\, \mathcal{F}'(s) &= 0 \ \text{ at } s = 0 \ \text{ and at } s = +\infty .
\end{aligned} \tag{3.92}$$

A covariant derivative is mandatory in eq. (3.90), since an ordinary derivative would break gauge invariance. Then, we replace the delta function by its Fourier representation:

$$\delta(x - y) = \int \frac{d^4k}{(2\pi)^4}\, e^{ik(x-y)} , \tag{3.93}$$

which leads to

$$\begin{aligned}
\mathcal{A}(x) &= -2 \int \frac{d^4k}{(2\pi)^4} \lim_{y \to x,\, M \to +\infty} \mathrm{tr} \left\{ \gamma^5\, t\, \mathcal{F}\left(-\frac{\not{D}_x^2}{M^2} \right) \right\} e^{ik(x-y)} \\
&= -2 \int \frac{d^4k}{(2\pi)^4} \lim_{M \to +\infty} \mathrm{tr} \left\{ \gamma^5\, t\, \mathcal{F}\left(-\frac{(i\not{k} + \not{D}_x)^2}{M^2} \right) \right\} .
\end{aligned} \tag{3.94}$$

The second equality follows from

$$\lim_{y \to x} \mathcal{F}(\partial_x)\, e^{ik\cdot(x-y)} = \mathcal{F}(ik + \partial_x) . \tag{3.95}$$

The function $\mathcal{A}(x)$ can then be rewritten as follows:

$$\mathcal{A}(x) = -2 \lim_{M \to +\infty} M^4 \int \frac{d^4k}{(2\pi)^4} \mathrm{tr} \left\{ \gamma^5\, t\, \mathcal{F}\left(-\left[i\not{k} + \frac{\not{D}_x}{M} \right]^2 \right) \right\} , \tag{3.96}$$

by redefining the integration variable, $k \to Mk$. Then, we can write:

$$-\left[i\not{k} + \frac{\not{D}_x}{M} \right]^2 = k^2 - 2i\frac{k \cdot D_x}{M} - \left(\frac{\not{D}_x}{M} \right)^2 , \tag{3.97}$$

and expand the function $\mathcal{F}(\cdot)$ in powers of $1/M$. The only terms that give a non-zero contribution to $\mathcal{A}(x)$ should not go to zero too quickly when $M \to +\infty$: only the terms decreasing at most as $1/M^4$ should be kept. Moreover, the Dirac trace should be non-zero, which implies that the matrix γ^5 must be accompanied by at least four ordinary γ^μ matrices. The matrices γ^μ come from the term $\not{D}_x^2$ in eq. (3.97), that brings two of them[8], and we therefore need to go to the second order in the Taylor expansion of the function $\mathcal{F}(\cdot)$. In fact, a single term fulfills all these constraints:

$$\mathcal{A}(x) = -\int \frac{d^4k}{(2\pi)^4}\, \mathcal{F}''(k^2)\, \mathrm{tr} \left(\gamma^5\, t\not{D}_x^4 \right) . \tag{3.98}$$

---

[7]We are considering here the case where the fermions are coupled to a non-Abelian gauge field. The index a carried by A_μ^a is a "color" index, and the t^a's are the generators of the Lie algebra representation where the fermions live. g is the coupling of the fermions to the gauge fields. See the next chapter for more details.

[8]In this counting, we assume that the matrix t does not contain Dirac matrices.

By a Wick's rotation ($k \rightarrow i\kappa$, $k^2 \rightarrow \kappa^2$), we obtain[9]:

$$\int d^4k \, \mathcal{F}''(k^2) = 2i\pi^2 \int\limits_0^{+\infty} d\kappa \, \kappa^3 \, \mathcal{F}''(\kappa^2) = i\pi^2 \, . \tag{3.99}$$

The last equality is obtained by two successive integrations by parts. We also have:

$$
\begin{aligned}
\slashed{D}_x^2 &= D_x^\mu D_x^\nu \, \gamma_\mu \gamma_\nu \\
&= \frac{1}{2} D_x^\mu D_x^\nu \left(\{\gamma_\mu, \gamma_\nu\} + [\gamma_\mu, \gamma_\nu] \right) \\
&= D_x^2 + \frac{1}{4} [D_x^\mu, D_x^\nu] \, [\gamma_\mu, \gamma_\nu] \\
&= D_x^2 - \frac{ig}{4} t^a F_a^{\mu\nu} \, [\gamma_\mu, \gamma_\nu] \, .
\end{aligned}
\tag{3.100}
$$

Using

$$\mathrm{tr} \left(\gamma^5 \gamma_\mu \gamma_\nu \gamma_\rho \gamma_\sigma \right) = -4 \, i \, \epsilon_{\mu\nu\rho\sigma} \, , \tag{3.101}$$

we obtain

$$\mathcal{A}(x) = -\frac{g^2}{16\pi^2} \epsilon_{\mu\nu\rho\sigma} \, F_a^{\mu\nu}(x) \, F_b^{\rho\sigma}(x) \, \mathrm{tr} \left(t^a t^b t \right) \, , \tag{3.102}$$

where the trace is now only on the fermion species. When t is the identity matrix, the integral of $\mathcal{A}(x)$ depends only on topological properties of the gauge field configuration and takes discrete values. In the context of anomalies, it is called the *Chern-Pontryagin index*. Moreover, the *Atiyah-Singer theorem* relates this invariant to the zero modes of the Euclidean Dirac operator in this gauge field (see the section 3.5.7).

3.5.5 Anomaly of the axial current

When the action is invariant under global chiral transformations, its variation under local chiral transformation may be written as

$$\delta S = \int d^4x \, J_5^\mu(x) \, \partial_\mu \alpha(x) \, , \tag{3.103}$$

where $J_5^\mu(x)$ is the *axial current*. Integrating by parts, and identifying this variation with the term obtained in the previous section, we should have

$$\langle \partial_\mu J_5^\mu(x) \rangle_A = -\frac{g^2}{16\pi^2} \epsilon_{\mu\nu\rho\sigma} \, F_a^{\mu\nu}(x) \, F_b^{\rho\sigma}(x) \, \mathrm{tr} \left(t^a t^b t \right) \, , \tag{3.104}$$

where $\langle \cdot \rangle_A$ is an average over the fermion fields, in a fixed gauge field configuration.

---

[9]Recall that the rotationally invariant measure in 4-dimensional Euclidean space is $2\pi^2 \kappa^3 \, d\kappa$.

3.5.6 Anomaly in the u and d quarks sector

Strong interaction :   Consider the sector of the two lightest quarks flavors, u et d. If one neglects their mass, the corresponding action is invariant under the following chiral transformations:

$$\delta u = i\,\alpha\gamma^5\,u\,,\quad \delta d = -i\,\alpha\gamma^5\,d\,. \tag{3.105}$$

The matrix t in quark flavor space that corresponds to this transformation is

$$t = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,. \tag{3.106}$$

Through the strong interactions, all quark flavors couple identically with the gluons (i.e. all quarks belong to the same representation of the $SU(3)$ algebra). In other words, the matrices t^a that describe this coupling do not depend on the quark flavor, and the trace that appears in the anomaly function is

$$\text{tr}\,(t_a t_b t) = t_{uu}\text{tr}\,(t_a t_b) + t_{dd}\text{tr}\,(t_a t_b) = 0\,. \tag{3.107}$$

This means that the anomalies that may occur in the gluon-gluon term cancel between the u and d flavors of quarks.

Electromagnetic interaction :   The situation is different with electromagnetic interactions, because the u and d quarks have different electrical charges. The analogue of the matrices t^a is the charge matrix of the $((u,d)$ doublet:

$$Q \equiv \begin{pmatrix} \frac{2}{3} & 0 \\ 0 & -\frac{1}{3} \end{pmatrix} \otimes \mathbf{1}_{\text{color}}\,. \tag{3.108}$$

(The factor $\mathbf{1}_{\text{color}}$ comes from the fact that all quark colors couple to photons in the same way.) Therefore we have

$$\text{tr}\,(Q^2 t) = \frac{N_c}{3}\,, \tag{3.109}$$

where $N_c = 3$ is the number of colors. This leads to

$$\mathcal{A}(x) = -\frac{e^2 N_c}{48\pi^2}\,\epsilon_{\mu\nu\rho\sigma}\,F^{\mu\nu}(x)\,F^{\rho\sigma}(x)\,, \tag{3.110}$$

where $F^{\mu\nu}$ is the electromagnetic field strength.

Decay of the neutral pion in two photons :   At low energy, the strong interactions may be described by an effective theory that couples a doublet of fermions ψ (the u and d quarks), the three pions $\boldsymbol{\pi}$ and a field σ. The interaction term in this model is

$$\mathcal{L}_{\text{I}} \equiv \lambda\,\overline{\psi}(\sigma + i\boldsymbol{\pi}\cdot\boldsymbol{\tau}\gamma^5)\psi\,, \tag{3.111}$$

where τ^i ($i = 1, 2, 3$) are the Pauli matrices. Note that π^3 must be the neutral pion, since it couples diagonally to the two components of the doublet (τ^3 is a diagonal matrix). This

interaction term is invariant under the transformation (3.105) provided that the fields σ and $\boldsymbol{\pi}$ transform as

$$\sigma \to \sigma - \alpha\,\boldsymbol{\pi}^3 \quad , \quad \boldsymbol{\pi}^{1,2} \to \boldsymbol{\pi}^{1,2} \quad , \quad \boldsymbol{\pi}^3 \to \boldsymbol{\pi}^3 + \alpha\sigma \; . \tag{3.112}$$

Moreover, the masses of nucleons are due to a spontaneous breaking of this symmetry, in which the σ field has a non-zero expectation value in the ground state: $\langle \sigma \rangle = f_\pi$. Thus the variation of the field $\boldsymbol{\pi}^3$ is $\delta\boldsymbol{\pi}^3 = f_\pi\,\alpha$.

When photons are added to this model, there is no direct coupling between the neutral pion and the photon. Let us now consider the theory that would result from integrating out the quark fields. The anomaly (3.110) would produce a term

$$\mathcal{L}_{\mathrm{anom}}(x) = -\frac{e^2\,N_c}{48\pi^2}\,\epsilon_{\mu\nu\rho\sigma}\,F^{\mu\nu}(x)\,F^{\rho\sigma}(x)\,\alpha(x) \; . \tag{3.113}$$

in the Lagrangian. This term should be canceled somehow, because we are now talking about an effective theory of pions and photons, that should be chiral invariant. The resolution of this issue is that this effective theory contains a coupling between the neutral pion and two photons, of the form:

$$\mathcal{L}_{\pi^0\gamma\gamma} = -\frac{e^2\,N_c}{48\pi^2 f_\pi}\,\epsilon_{\mu\nu\rho\sigma}\,F^{\mu\nu}(x)\,F^{\rho\sigma}(x)\,\boldsymbol{\pi}^3(x) \; . \tag{3.114}$$

The decay rate of a neutral pion into two photons can be easily determined from the effective coupling (3.114):

$$\Gamma(\pi^0 \to 2\gamma) = \frac{N_c^2\alpha_{\mathrm{em}}^2 m_\pi^3}{144\pi^3 f_\pi^2} \; . \tag{3.115}$$

This result could also be obtained by computing the transition amplitude at one loop from a neutral pion to two photons in the effective model we started from. The present considerations show that this decay is in fact controlled to a large extent by a quantum anomaly.

3.5.7 Atiyah-Singer index theorem

Covariant derivatives $D_\mu = \partial_\mu - i\,g\,A_\mu^a t^a$ are anti-Hermitean, because the gauge potential A_μ^a is real and the color matrices t^a are Hermitean (recall that an ordinary derivative is anti-Hermitean). However, γ^0 is Hermitean, while $\gamma^{1,2,3}$ are anti-Hermitean. Therefore, the Dirac operator $D_\mu\gamma^\mu$ in Minkowski space is neither Hermitean not anti-Hermitean.

Let us introduce an Euclidean time via $x_4 \equiv i x^0$. Likewise, we also have:

$$\partial_4 = i\partial_0 \quad , \quad A_4 = iA_0 \quad , \quad \gamma_4 = i\gamma_0 \; , \tag{3.116}$$

and the measure over space-time becomes $d^4x = i\,d^4x_{\mathrm{E}}$ where d^4x_{E} is the measure over 4-dimensional Euclidean space ($d^4x_{\mathrm{E}} = dx_1\,dx_2\,dx_3\,dx_4$). The Dirac operator becomes:

$$\slashed{D} = \sum_{i=1}^{4}(\partial_i - i\,g\,A_i^a t^a)\gamma^i \; , \tag{3.117}$$

where the index i runs from 1 to 4. Now, the Dirac matrices γ^i are all anti-Hermitean, which implies that the Euclidean Dirac operator is Hermitean. It can therefore be diagonalized in an orthonormal basis of eigenfunctions ϕ_k:

$$\slashed{D}_x \phi_k(x) = \lambda_k \phi_k(x) \,,$$
$$\int d^4 x_E \ \phi_k^\dagger(x)\phi_{k'}(x) = \delta_{kk'} \,, \tag{3.118}$$

with real eigenvalues λ_k.

Moreover, let us consider transformations that act in flavor space, but do not contain Dirac matrices. Therefore, the matrix t commutes with the Dirac operator, and we can choose the eigenfunctions ϕ_k so that they are also eigenfunctions of t:

$$t \, \phi_k(x) = t_k \phi_k(x) \,. \tag{3.119}$$

Note also that these eigenfunctions must obey the following completeness relation:

$$\sum_k \phi_k(x)\phi_k^\dagger(y) = \delta(x - y) \,. \tag{3.120}$$

We can use this completeness identity in order to express the delta function in the anomaly function $\mathcal{A}(x)$ in eq. (3.90):

$$
\begin{aligned}
\mathcal{A}(x) &= -2 \lim_{y \to x, M \to +\infty} \mathrm{tr} \left\{ \gamma^5 \, t \, \mathcal{F}\left(-\frac{\slashed{D}_x^2}{M^2} \right) \sum_k \phi_k(x)\phi_k^\dagger(y) \right\} \\
&= -2 \lim_{y \to x, M \to +\infty} \sum_k \mathrm{tr} \left\{ \phi_k^\dagger(y) \gamma^5 \, t \, \mathcal{F}\left(-\frac{\slashed{D}_x^2}{M^2} \right) \phi_k(x) \right\} \\
&= -2 \lim_{M \to +\infty} \sum_k t_k \mathcal{F}\left(-\frac{\lambda_k^2}{M^2} \right) \phi_k^\dagger(x)\gamma^5 \phi_k(x) \,. \tag{3.121}
\end{aligned}
$$

Specializing to the case where $t = 1$ (i.e. all the eigenvalues t_k are equal to 1), we obtain the following relationship,

$$
\frac{g^2}{32\pi^2} \int d^4 x_E \ \epsilon_{ijkl} \, F_{ij}^a(x) \, F_{kl}^b(x) \, \mathrm{tr}(t^a t^b)
$$
$$
= -\frac{1}{2} \int d^4 x_E \ \mathcal{A}(x) = \lim_{M \to +\infty} \sum_k \mathcal{F}\left(-\frac{\lambda_k^2}{M^2} \right) \int d^4 x_E \ \phi_k^\dagger(x)\gamma^5 \phi_k(x) \,, \tag{3.122}
$$

between an integral that involves the field strength of a gauge field configuration and a sum over the spectrum of the Euclidean Dirac operator (in the same gauge field). Since γ^5 anti-commutes with the Dirac operator,

$$\{\gamma^5, \slashed{D}\} = 0 \,, \tag{3.123}$$

the state $\phi_{k'} \equiv \gamma^5 \phi_k(x)$ is also an eigenfunction of $\slashed{D}_x$ with the eigenvalue $-\lambda_k$:

$$\slashed{D}_x(\gamma^5 \phi_k(x)) = -\lambda_k(\gamma^5 \phi_k(x)) \,. \tag{3.124}$$

When $\lambda_k \neq 0$, the state $\phi_{k'}$ is distinct from the state $\phi_k(x)$. Since $\not{D}_x$ is Hermitean, they are in fact orthogonal:

$$\int d^4 x_E \; \phi_k^\dagger(x) \gamma^5 \phi_k(x) = \int d^4 x_E \, \phi_k^\dagger(x) \phi_{k'}(x) = 0 \; . \tag{3.125}$$

This implies that none of the eigenfunctions ϕ_k with a non-zero eigenvalue can contribute to the right hand side of eq. (3.122). The only contributions to eq. (3.122) come from the eigenfunctions for which $\lambda_k = 0$, i.e. the *zero modes* of the Euclidean Dirac operator. Since we have assumed that $f(0) = 1$, we have:

$$\frac{g^2}{32\pi^2} \int d^4 x_E \; \epsilon_{ijkl} F_{ij}^a(x) F_{kl}^b(x) \operatorname{tr}(t^a t^b) = \sum_{k|\lambda_k = 0} \int d^4 x_E \; \phi_k^\dagger(x) \gamma^5 \phi_k(x) \; . \tag{3.126}$$

Since $\{\gamma^5, \not{D}_x\} = 0$, we can choose these zero modes in such a way that they are also eigenmodes of γ^5, with eigenvalues $+1$ or -1. We can thus divide the zero modes in two families, the right-handed and the left-handed zero modes:

$$\begin{aligned}
\not{D}_x \phi_R(x) &= 0 \; , & \gamma^5 \phi_R(x) &= +\phi_R(x) \; , \\
\not{D}_x \phi_L(x) &= 0 \; , & \gamma^5 \phi_L(x) &= -\phi_L(x) \; .
\end{aligned} \tag{3.127}$$

Using also the fact that the eigenfunctions are normalized as follows,

$$\begin{aligned}
\int d^4 x_E \; \phi_R^\dagger(x) \phi_R(x) &= 1 \; , \\
\int d^4 x_E \; \phi_L^\dagger(x) \phi_L(x) &= 1 \; ,
\end{aligned} \tag{3.128}$$

we obtain the following identity

$$\frac{g^2}{32\pi^2} \int d^4 x_E \; \epsilon_{ijkl} F_{ij}^a(x) F_{kl}^b(x) \operatorname{tr}(t^a t^b) = n_R - n_L \; , \tag{3.129}$$

where n_R and n_L are the numbers of right-handed and left-handed zero modes, respectively. This formula is the Atiyah-Singer theorem. It tells us that the integral in the left hand side is an integer, despite being the integral of a quantity that changes continuously when one deforms the gauge field. Different considerations, from the study of Euclidean gauge field configurations known as *instantons*, provide another insight on this integral by relating it to the third homotopy group of the gauge group, $\pi_3(SU(N_c)) = \mathbb{Z}$.

Chapter 4

Non-Abelian gauge symmetry

4.1 Non-abelian Lie groups and algebras

4.1.1 Reminder: Abelian gauge transformations

Gauge theories are quantum field theories with matter fields (usually spin $1/2$ fermions, but also possibly scalars) and gauges potentials in such a way that the Lagrangian is invariant under the action of a local continuous transformation. Quantum Electrodynamics is the simplest such theory, with a local $U(1)$ invariance. Given $\Omega(x) \in U(1)$, the various objects that enter in the theory transform as follows:

$$
\begin{aligned}
\psi &\rightarrow \Omega^\dagger \psi \,, \\
A^\mu &\rightarrow A^\mu + \frac{i}{e}\Omega^\dagger \partial^\mu \Omega \,, \\
F^{\mu\nu} &\rightarrow F^{\mu\nu} \,, \\
D^\mu &\rightarrow \Omega^\dagger D^\mu \Omega = \partial^\mu - ie\left(A^\mu + \frac{i}{e}\Omega^\dagger \partial^\mu \Omega\right) \,.
\end{aligned}
\tag{4.1}
$$

In this construction, the gauge transformation of A^μ could have been found by requesting that $D^\mu \psi$ transforms as ψ itself,

$$
D^\mu \psi \quad \rightarrow \quad \Omega^\dagger(x) \, D^\mu \psi \,,
\tag{4.2}
$$

with $D^\mu \equiv \partial^\mu - ieA^\mu$ the covariant derivative. The field strength $F^{\mu\nu}$ would then be defined as $\partial^\mu A^\nu - \partial^\nu A^\mu$.

Our goal is now to generalize the concept of gauge theory to more general transformations, in view of applications to the electroweak and to the strong interactions. In these two cases, the internal group of transformations is $SU(2)$ and $SU(3)$, respectively, but we will consider in most of this chapter a general *Lie group* $\mathcal{G}$. Starting from the first of eqs. (4.1), with $\Omega(x) \in \mathcal{G}$, we want to determine the other transformation laws, and construct an invariant action.

4.1.2 Lie groups

Let us start by recalling that a Lie group is a group which is also a smooth manifold. The group operation will be denoted multiplicatively, as in $\Omega_2\Omega_1$, and we will denote the identical element by $\mathbf{1}$ and the inverse of a group element Ω by Ω^{-1}. The fact that a Lie group $\mathcal{G}$ is also a manifold allows to use concepts of differential geometry in their study.

Matrix Lie groups, that will be our main concern in view of applications to quantum field theory, are closed subsets of $GL(n, \mathbb{C})$, the general linear group of $n \times n$ matrices on the field of complex numbers. Here is a list of some classical examples of matrix Lie groups, along with their definition:

Special linear groups :	$SL(n, \mathbb{C})$, $SL(n, \mathbb{R})$	$\det(\Omega) = 1$
Special orthogonal group :	$SO(n)$	$\Omega^{\mathsf{T}}\Omega = \mathbf{1}$, $\det \Omega = 1$
Unitary group :	$U(n)$	$\Omega^{\dagger}\Omega = \mathbf{1}$
Special unitary group :	$SU(n)$	$\Omega^{\dagger}\Omega = \mathbf{1}$, $\det \Omega = 1$

$$(4.3)$$

4.1.3 Lie algebras

Geometrically, the *Lie algebra* $\mathfrak{g}$ is a vector space that may be viewed as tangent to the group at the identity $\mathbf{1}$. Therefore, its dimension is the same as that of the group manifold. The group multiplication induces on the tangent space a non-associative multiplication, the *Lie bracket*, thereby turning it into an algebra. The Lie algebra completely encapsulates the local properties of the underlying Lie group, and if the group is simply connected its Lie algebra defines it globally. Because they are linear spaces, Lie algebras are usually easier to study than their group counterpart, although they provide most of the information.

In the specific case of matrix Lie groups, the corresponding Lie algebra can be defined as the following set of matrices[1]

$$\mathfrak{g} \equiv \left\{ X \,\middle|\, e^{it\,X} \in \mathcal{G}, \forall t \in \mathbb{R} \right\} . \tag{4.4}$$

A crucial property of the matrix exponential is that

$$e^{X+Y} \neq e^X \, e^Y \qquad \text{if } [X, Y] \neq 0 . \tag{4.5}$$

Instead, one may use *Trotter's formula*[2]:

$$e^{X+Y} = \lim_{n \to \infty} \left(e^{X/n} \, e^{Y/n} \right)^n . \tag{4.6}$$

(See the figure 4.2 for a geometrical illustration of this formula.) From the definition (4.4) of the Lie algebra, and using Trotter's formula, one can check that any real linear combination

---

[1]The prefactor i inside the exponential is common in the quantum physics literature, but seldom used in mathematics. Its main benefit is to make X a Hermitean matrix when the group elements are unitary.

[2]A sketch of the proof is the following:

$$e^{X/n} \, e^{Y/n} = 1 + \tfrac{X}{n} + \tfrac{Y}{n} + \mathcal{O}(n^{-2}) = \exp\left(\tfrac{X+Y}{n} + \mathcal{O}(n^{-2}) \right) ,$$

$$\left(e^{X/n} \, e^{Y/n} \right)^n = \exp\left(X + Y + \mathcal{O}(n^{-1}) \right) .$$

of elements of $\mathfrak{g}$ is in $\mathfrak{g}$, i.e. that $\mathfrak{g}$ is a vector space. Therefore, every element of $\mathfrak{g}$ can be written as a linear combination of some basis elements t^a,

$$X = x_a t^a \,, \tag{4.7}$$

with an implicit sum on the index a. The t^a's are also called the *generators* of the algebra.

Thanks to the exponential mapping (4.4), the properties of the Lie groups listed in eqs. (4.3) translate into specific properties of the matrices X in the corresponding algebras:

Special linear groups :	$\mathfrak{sl}(n, \mathbb{C})$, $\mathfrak{sl}(n, \mathbb{R})$	$\operatorname{tr}(X) = 0$	
Special orthogonal group :	$\mathfrak{so}(n)$	$X^{\mathsf{T}} = -X$	
Unitary group :	$\mathfrak{u}(n)$	$X^{\dagger} = X$	
Special unitary group :	$\mathfrak{su}(n)$	$X^{\dagger} = X$, $\operatorname{tr}(X) = 0$	

$$\tag{4.8}$$

Note that the conditions imposed on Ω in eqs. (4.3) are non-linear, in contrast with the linear conditions obeyed by the matrices X in eqs. (4.8). This is why a Lie group is a curved manifold, while a Lie algebra is a linear space.

4.1.4 Geometrical interpretation

First note that we have

$$iX = \left.\frac{\mathrm{d}}{\mathrm{d}t} e^{itX}\right|_{t=0} \,. \tag{4.9}$$

The group elements $\exp(itX)$ form a smooth curve on the group manifold ($t = 0$ corresponds to the identity), and iX may be viewed as the vector tangent to this curve at the identity, as illustrated in the figure 4.1. The non-commutativity of the group is related to the curvature of the corresponding manifold[3]. Because of this curvature, a displacement e^{iX} followed by a displacement e^{iY} does not lead to the same point as the two displacements performed in reverse order. This geometrical representation also provides an interpretation of Trotter's formula for the exponential of a sum, as shown in the figure 4.2. From the figure 4.1, we see that the dimension of the Lie algebra is the number of independent directions on the group manifold. From the conditions listed in (4.8) on the matrices $X \in \mathfrak{g}$, it is easy to determine the dimension of these algebras (viewed as algebras over the field $\mathbb{R}$):

Algebra	Dimension
$\mathfrak{sl}(n, \mathbb{R})$	$n^2 - 1$
$\mathfrak{so}(n)$	$n(n-1)/2$
$\mathfrak{u}(n)$	n^2
$\mathfrak{su}(n)$	$n^2 - 1$

Despite these correspondences, the Lie algebra may not reflect the global properties of the group (e.g. whether it is connected), and distinct Lie groups may have isomorphic Lie algebras. This is for instance the case of $U(1)$ and $SO(2)$, $SO(3)$ and $SU(2)$, or $SU(2) \times SU(2)$ and $SO(4)$.

---

[3]This statement can be made more precise: it is possible to define a metric tensor for the group manifold, and to express the associated curvature tensor in terms of the structure constants.

Figure 4.1: *Lie group and Lie algebra.*

Lie Group manifold

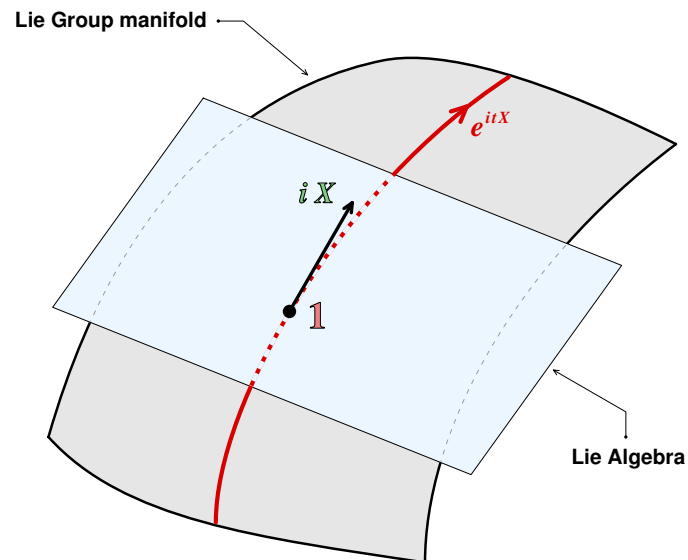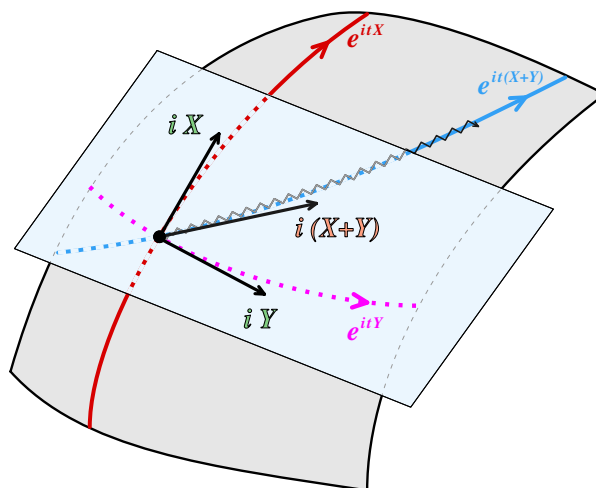e^{itX}

$i\,X$

$\mathbf{1}$

Lie Algebra

Figure 4.2: *Geometrical interpretation of Trotter's formula: the broken path, made of a succession of elementary steps* $e^{itX/n}$ *and* $e^{itY/n}$, *approximates better and better the curve* $e^{it(X+Y)}$ *on the group manifold as* $n \to \infty$.

e^{itX}

$e^{it(X+Y)}$

$i\,X$

$i\,(X+Y)$

$i\,Y$

e^{itY}

4.1.5 Lie bracket and structure constants

Consider an element Ω of the Lie group and an element X of the Lie algebra. For any real number t, we have

$$\exp\left(i\,t\,\Omega^{-1}X\Omega\right) = \Omega^{-1}\underbrace{\underbrace{e^{itX}}_{\in\mathcal{G}}\Omega}_{\in\mathcal{G}}\,, \tag{4.10}$$

where the equality follows from the Taylor series of the exponential. From the definition of the Lie algebra, this implies that $\Omega^{-1}\,X\,\Omega \in \mathfrak{g}$. Therefore, if $X, Y \in \mathfrak{g}$ we also have

$$e^{-itX}\,Y\,e^{itX} \in \mathfrak{g}\,, \tag{4.11}$$

and the derivative with respect to t at $t = 0$ is also an element of the algebra,

$$-i\left[X, Y\right] \in \mathfrak{g}\,. \tag{4.12}$$

In other words, $-i$ times the *commutator* of two elements of a Lie algebra is another element of the algebra. Thus $-i[\cdot, \cdot]$ is the multiplication law[4] in $\mathfrak{g}$ (it is also called the *Lie bracket*). Therefore, the commutators between its generators can be written as

$$\left[t^a, t^b\right] = i\,f^{abc}\,t^c\,, \tag{4.13}$$

where the f^{abc} are real numbers called the *structure constants*. The antisymmetry of the commutator implies that $f^{abc} = -f^{bac}$. Given three elements $X, Y, Z \in \mathfrak{g}$ of the algebra, their commutator satisfies the *Jacobi identity*

$$\left[X, \left[Y, Z\right]\right] + \left[Y, \left[Z, X\right]\right] + \left[Z, \left[X, Y\right]\right] = 0\,, \tag{4.14}$$

which implies the following relationship among the structure constants:

$$f^{ade}f^{bcd} + f^{bde}f^{cad} + f^{cde}f^{abd} = 0\,. \tag{4.15}$$

4.1.6 Baker-Campbell-Hausdorff formula

Given an element $X \in \mathfrak{g}$, we may define a function from $\mathfrak{g}$ to $\mathfrak{g}$ as follows:

$$\mathrm{ad}_X\left(Y\right) \equiv -i\left[X, Y\right]\,. \tag{4.16}$$

The function ad_X is called the *adjoint mapping* at the point X. The exponential of the adjoint mapping plays an important role, thanks to the following formula

$$e^{\mathrm{ad}_X}\,Y = e^{-iX}\,Y\,e^{iX}\,. \tag{4.17}$$

This allows to write the derivative of the exponential of a (matrix-valued) function as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}e^{iX(t)} = i\,e^{iX(t)}\,\frac{e^{\mathrm{ad}_{X(t)}} - 1}{\mathrm{ad}_{X(t)}}\,\frac{\mathrm{d}X(t)}{\mathrm{d}t}\,. \tag{4.18}$$

---

[4]In contrast, the ordinary product of two elements of the algebra is in general not in the algebra.

(This is known as Duhamel's formula[5].) The non-trivial aspect of this formula is that it is true even when $X(t)$ does not commute with its derivative. Then, given $X, Y \in \mathfrak{g}$, let us define a matrix $Z(t)$ by

$$e^{iZ(t)} \equiv e^{iX} e^{itY} . \tag{4.19}$$

Differentiating both sides with respect to t (using eq. (4.18) for the left hand side), we obtain

$$\frac{dZ(t)}{dt} = \left[\frac{e^{\mathrm{ad}_{Z(t)}} - 1}{\mathrm{ad}_{Z(t)}} \right]^{-1} Y . \tag{4.20}$$

From eq. (4.17), we can also see that

$$e^{\mathrm{ad}_{Z(t)}} = e^{t\,\mathrm{ad}_Y} e^{\mathrm{ad}_X} . \tag{4.21}$$

Integrating eq. (4.20) from $t = 0$ to $t = 1$, we obtain the following identity:

$$\ln \left(e^{iX} e^{iY} \right) = iX + i \int_0^1 dt\, F \left(e^{t\,\mathrm{ad}_Y} e^{\mathrm{ad}_X} \right) Y , \tag{4.22}$$

where the function $F(\cdot)$ is defined by

$$F(z) \equiv \frac{\ln(z)}{z-1} . \tag{4.23}$$

Eq. (4.22) is the integral form of the Baker-Campbell-Hausdorff formula. In order to recover the more familiar expansion in nested commutators, note that

$$e^{t\,\mathrm{ad}_Y} e^{\mathrm{ad}_X} = 1 + t\,\mathrm{ad}_Y + \mathrm{ad}_X + \tfrac{1}{2}(t^2 \mathrm{ad}_Y^2 + \mathrm{ad}_X^2) + t\,\mathrm{ad}_Y\,\mathrm{ad}_X + \cdots$$
$$F(z) = 1 - \frac{1}{2}(z-1) + \tfrac{1}{3}(z-1)^2 \cdots . \tag{4.24}$$

This leads to

$$\ln \left(e^{iX} e^{iY} \right) = i(X+Y) - \frac{1}{2}[X, Y] - \frac{i}{12} \Big([X, [X, Y]] - [Y, [X, Y]] \Big) + \cdots \tag{4.25}$$

(Explicit expressions for all the coefficients of this series are given by *Dynkin's formula*.) In applications to quantum field theory, we usually need only the first two terms of this expansion because the commutators we encounter are c-numbers and all the subsequent terms are zero. Besides being an intermediate step in the derivation of eq. (4.25), the integral form (4.22) shows that the group product can be reconstructed from Lie algebra manipulations (since the right hand side of this equation contains only objects that belong to the algebra).

---

[5]Note that this formula is equivalent to:

$$\frac{d}{dt} e^{iX(t)} = i \int_0^1 ds\, e^{isX(t)} \frac{dX(t)}{dt} e^{i(1-s)X(t)} .$$

This latter form can be proven by writing

$$e^{iX(t')} = e^{i\left(X(t)+(t'-t)X'(t)+\cdots\right)} = \lim_{n\to\infty} \left(e^{i\frac{X(t)}{n}} e^{i(t'-t)\frac{X'(t)}{n}+\cdots} \right)^n$$

and by expanding the right hand side to first order in $t' - t$.

4.2 Representations

A real *representation of a Lie group* $\mathcal{G}$ is a group homomorphism from elements of $\mathcal{G}$ to elements of $GL(n, \mathbb{R})$, i.e. a mapping π from $\mathcal{G}$ to $GL(n, \mathbb{R})$ that preserves the group structure:

$$\pi(\mathbb{1}) = \mathbb{1} \quad , \qquad \pi(\Omega_2 \Omega_1) = \pi(\Omega_2)\pi(\Omega_1) \; . \tag{4.26}$$

Likewise, one may define representations of a Lie algebra, as homomorphisms from $\mathfrak{g}$ to $\mathfrak{gl}(n, \mathbb{R})$. A representation is said to be faithful if π is a one-to-one mapping.

Since we are focusing on matrix Lie groups, their elements are already matrices, and one may wonder what representations are good for. In fact, it is often important to know how a given group (e.g. the rotation group $SO(3)$) acts on a more general linear space. In the example of $SO(3)$, even though the "defining" action is on $\mathbb{R}^3$ in terms of 3×3 matrices, the group has many other matrix representations made of objects that act on spaces other than $\mathbb{R}^3$.

Singlet representation : The *singlet representation*, or *trivial representation*, is the representation for which the mapping is $\pi(\Omega) = \mathbb{1}$ for all Ω's. The objects that belong to this representation space are invariant under the transformations of the group $\mathcal{G}$. In quantum field theory, one says that these objects are "neutral" (under the group $\mathcal{G}$).

Fundamental representation : The *fundamental representation*, or *standard representation*, is the smallest faithful representation. It is also the representation obtained when π is the identical map. In other words, in the fundamental representation, the elements of a matrix Lie group are simply represented by themselves. It is customary to normalize the generators of the fundamental representation of the Lie algebra, denoted t_f^a, as follows:

$$\mathrm{tr}\left(t_f^a t_f^b\right) = \frac{\delta^{ab}}{2} \; . \tag{4.27}$$

This choice sets the normalization of the structure constants, through eq. (4.13). Then, one usually normalizes the generators of other representations in such a way that they fulfill the commutation relation (4.13) with the same structure constants (but the trace formula (4.27) will in general be satisfied only in the fundamental representation).

Fierz identity for $\mathfrak{su}(n)$: In the case of $\mathfrak{su}(n)$, there are $n^2 - 1$ generators t_f^a, while the linear space of all $n \times n$ Hermitian matrices has a dimension n^2. A basis of the latter can be obtained by adding the identity matrix to the t_f^a's. Thus, any $n \times n$ Hermitean matrix M can be written as

$$M = m_0 \, \mathbb{1} + m_a \, t_f^a \; . \tag{4.28}$$

Since the t_f^a's are traceless, we have

$$m_0 = \frac{1}{n} \, \mathrm{tr}\,(M) \quad , \qquad m_a = 2 \, \mathrm{tr}\,(M \, t^a) \; . \tag{4.29}$$

Considering the entry ij of the matrix M, we can write

$$
\begin{aligned}
M_{ij} &= \frac{1}{n} M_{kk}\,\delta_{ij} + 2\,M_{lk}\left(t_f^a\right)_{kl}\left(t_f^a\right)_{ij} \\
&= M_{lk}\left[\frac{1}{n}\delta_{kl}\delta_{ij} + 2\left(t_f^a\right)_{kl}\left(t_f^a\right)_{ij}\right].
\end{aligned}
\tag{4.30}
$$

Since this is true for any Hermitean matrix M, we must have

$$
\frac{1}{n}\delta_{kl}\delta_{ij} + 2\left(t_f^a\right)_{kl}\left(t_f^a\right)_{ij} = \delta_{il}\delta_{jk},
\tag{4.31}
$$

which is usually written as follows

$$
\left(t_f^a\right)_{ij}\left(t_f^a\right)_{kl} = \frac{1}{2}\left[\delta_{il}\delta_{jk} - \frac{1}{n}\delta_{ij}\delta_{kl}\right].
\tag{4.32}
$$

This formula is called a *Fierz identity*. It has a convenient diagrammatic representation,

$$
\left(t_f^a\right)_{ij}\left(t_f^a\right)_{kl} = \;\text{(diagram)}\; = \frac{1}{2}\;\text{(diagram)}\; - \frac{1}{2n}\;\text{(diagram)}\;,
\tag{4.33}
$$

in which the solid blobs represent the t_f^a matrices, and the wavy line indicates that the indices a are contracted. In the right hand side, the solid lines indicate how the indices $ijkl$ are connected by the delta symbols. By contracting the indices jk in the Fierz identity (4.32), we obtain:

$$
\left(t_f^a t_f^a\right)_{il} = \frac{n^2 - 1}{2n}\,\delta_{il}.
\tag{4.34}
$$

The quadratic combination $t_f^a t_f^a$, called the *fundamental Casimir operator*, is proportional to the identity (and therefore commutes with everything). The prefactor is sometimes denoted $C_f \equiv (n^2 - 1)/2n$.

The diagrammatic representation (4.33) provides a very convenient way of obtaining certain identities involving the generators of the fundamental representation. As an illustration, let us consider the following example:

$$
\begin{aligned}
t_f^a t_f^b t_f^a = \;\text{(diagram)}\; &= \frac{1}{2}\;\text{(diagram)}\; - \frac{1}{2n}\;\text{(diagram)}\; \\
&= \frac{1}{2}\,\mathrm{tr}\left(t_f^b\right)\mathbf{1} - \frac{1}{2n}\,t^b = -\frac{1}{2n}\,t^b.
\end{aligned}
\tag{4.35}
$$

For the first term, we have used the fact that a closed loop in this diagrammatic representation corresponds to a trace over the color indices, and the tracelessness the generators. Likewise, one would obtain

$$
t_f^a t_f^b t_f^c t_f^a t_f^b = \;\text{(diagram)}\; = \frac{1}{4}\left(1 + \frac{1}{n^2}\right)t_f^c.
\tag{4.36}
$$

Adjoint representation : The *adjoint representation* of a Lie group $\mathcal{G}$ is a representation as linear operators that act on the Lie algebra $\mathfrak{g}$, defined by the following mapping:

$$\Omega \in \mathcal{G} \;\to\; \mathrm{Ad}_\Omega \in \mathrm{GL}(\mathfrak{g}) \qquad \text{such that} \quad \mathrm{Ad}_\Omega(X) = \Omega^{-1} X \Omega . \tag{4.37}$$

If the dimension of the Lie algebra is d, then Ad_Ω may be viewed as an $d \times d$ matrix. We may also define the adjoint representation of the algebra $\mathfrak{g}$, as follows:

$$X \in \mathfrak{g} \;\to\; \mathrm{ad}_X \in \mathrm{GL}(\mathfrak{g}) \qquad \text{such that} \quad \mathrm{ad}_X(Y) = -i[X, Y] . \tag{4.38}$$

It is sufficient to know the adjoint representation of the generators t^a, for which one often uses the following notation

$$-i\, \mathrm{ad}_{t^a} \equiv T^a_{\mathrm{adj}} . \tag{4.39}$$

Note that T^a_{adj} can be represented by a $d \times d$ matrix. Using Jacobi's identity, one may check that $\left[\mathrm{ad}_{t^a}, \mathrm{ad}_{t^b}\right] = -\mathrm{ad}_{i[t^a, t^b]} = f^{abc}\, \mathrm{ad}_{t^c}$. Therefore, the T^a_{adj}'s fulfill the same commutation relations as the t^a's themselves:

$$\left[T^a_{\mathrm{adj}}, T^b_{\mathrm{adj}}\right] = i\, f^{abc}\, T^c_{\mathrm{adj}} . \tag{4.40}$$

Using eq. (4.13), we find that the components of these matrices are given by

$$\left(T^a_{\mathrm{adj}}\right)_{bc} = -i\, f^{abc} . \tag{4.41}$$

In other words, the adjoint representation is a representation by matrices whose size is the dimension of the algebra, and in which the components of the generators are the structure constants. That eqs. (4.40) and (4.41) are consistent is a consequence of the Jacobi identity (4.15) satisfied by the structure constants.

A common use of the adjoint representation is to rearrange expressions such as

$$e^{-iX}\, Y\, e^{iX} = e^{\mathrm{ad}_X}\, Y , \tag{4.42}$$

where X and Y are in some representation r of the Lie algebra. Using $X = X_a t^a_r$ and $Y = Y_a t^a_r$, we can rewrite this as follows

$$e^{\mathrm{ad}_X}\, Y = e^{X_a\, \mathrm{ad}_{t^a}}\, Y_b t^b_r = t^c_r \left[e^{iX_a T^a_{\mathrm{adj}}}\right]_{cb} Y_b . \tag{4.43}$$

Thus, we have

$$\left[e^{-iX}\, Y\, e^{iX}\right]_c = \left[e^{iX_{\mathrm{adj}}}\right]_{cb} Y_b , \tag{4.44}$$

where the left hand side may be in any representation r. In other words, the right and left multiplication by a group element and its inverse can be rewritten as a left multiplication by the adjoint of this group element.

4.2.1 More identities for $\mathfrak{su}(n)$

We list here a few useful identities, that are specific to the case of $\mathfrak{su}(n)$. Contrary to the commutator, the anti-commutator of two matrices of the algebra does not belong to the algebra. However, it can be decomposed as a linear combination of the identity and the generators of the algebra:

$$\{t^a_f, t^b_f\} = \frac{\delta^{ab}}{n}\, \mathbf{1} + d^{abc}\, t^c_f . \tag{4.45}$$

The first term is obtained by taking the trace of the equation, using eq. (4.27) and the fact that the generators are traceless. The constants d^{abc} are sometimes called the *symmetric structure constants*. Therefore, the product of two generators of the fundamental representation can be written as

$$t_f^a t_f^b = \frac{1}{2} \left(\frac{\delta^{ab}}{n} \mathbf{1} + (d^{abc} + i f^{abc}) t_f^c \right) . \tag{4.46}$$

From this, we deduce the following identities

$$
\begin{aligned}
\operatorname{tr} \left(t_f^a t_f^b t_f^c \right) &= \frac{1}{4} \left(d^{abc} + i f^{abc} \right) , \\
\operatorname{tr} \left(t_f^a t_f^b t_f^a t_f^c \right) &= -\frac{1}{4n} \delta_{bc} , \\
f^{acd} f^{bcd} &= n \, \delta_{ab} , \\
d^{acd} d^{bcd} &= \left(\frac{4}{n} - n \right) \delta_{ab} , \\
f^{acd} d^{bcd} &= 0 , \\
f^{ade} f^{bef} f^{cfd} &= \frac{n}{2} \, f^{abc} .
\end{aligned}
\tag{4.47}
$$

Note that the third of these equations provides the trace of the product of two generators in the adjoint representation:

$$\operatorname{tr} \left(T_{\text{adj}}^a T_{\text{adj}}^b \right) = n \, \delta_{ab} . \tag{4.48}$$

4.3 Covariant derivative

Spinors : Consider a spinor ψ (this may be a multiplet made of several elementary spinors). Saying that ψ lives in a representation r of the Lie group $\mathcal{G}$ means that the group $\mathcal{G}$ acts on ψ as follows:

$$\psi \quad \to \quad \Omega^{-1} \psi , \tag{4.49}$$

where Ω is an element of the representation r of the Lie group.

In the Dirac Lagrangian density, the kinetic term $\overline{\psi} \gamma^\mu \partial_\mu \psi$ and the mass term $m \overline{\psi} \psi$ are invariant under such transformations when the Lie group $\mathcal{G}$ is $SU(N)$. The mass term is in fact invariant under *local transformations*, i.e. when Ω depends on space-time:

$$\psi(x) \quad \to \quad \Omega^{-1}(x) \, \psi(x) , \tag{4.50}$$

but this is not true of the kinetic term. Loosely speaking the obstruction comes from the fact that we cannot carry an x-dependent $\Omega(x)$ through the derivative ∂_μ:

$$\Omega(x) \, \partial_\mu \cdots = \partial_\mu \, \Omega(x) \cdots - \underbrace{\left(\partial_\mu \Omega(x) \right) \cdots}_{\text{extra term}} \tag{4.51}$$

104

Gauge fields : A *covariant derivative*, denoted D_μ, is a deformation of the ordinary derivative such that $\overline{\psi} D_\mu \psi$ is invariant under local $SU(N)$ transformations. Given the transformation law of ψ, the covariant derivative must therefore transform as follows:

$$D_\mu \quad \to \quad \Omega^\dagger(x)\, D_\mu\, \Omega(x)\,. \tag{4.52}$$

Let us look for a covariant derivative of the form

$$D_\mu \equiv \partial_\mu - ig\, A_\mu(x)\,, \tag{4.53}$$

where g is a coupling constant similar to the constant e in QED and $A_\mu(x)$ a 4-vector (in quantum field theory, this field is called a gauge field). The transformation law (4.52) is fulfilled provided that $A_\mu(x)$ transforms in a very specific way. Note first that the ordinary derivative ∂_μ is invariant (i.e. it belongs to the singlet representation of $SU(N)$). If we denote $A_\mu^\Omega(x)$ the transformed $A_\mu(x)$, then we must have:

$$
\begin{aligned}
\partial_\mu - ig\, A_\mu^\Omega(x) &= \Omega^\dagger(x)\left[\partial_\mu - ig\, A_\mu(x)\right]\Omega(x) \\
&= \partial_\mu + \Omega^\dagger(x)\big(\partial_\mu \Omega(x)\big) - ig\, \Omega^\dagger(x)\, A_\mu(x)\, \Omega(x)\,,
\end{aligned}
\tag{4.54}
$$

from which we obtain the transformation law[6] of $A_\mu(x)$:

$$A_\mu(x) \quad \to \quad A_\mu^\Omega(x) \equiv \Omega^\dagger(x)\, A_\mu(x)\, \Omega(x) + \frac{i}{g}\, \Omega^\dagger(x)\big(\partial_\mu \Omega(x)\big)\,. \tag{4.55}$$

From eqs. (4.16), (4.17) and (4.18), we see that if Ω is an element of a Lie group $\mathcal{G}$, then $\Omega^\dagger \partial_\mu \Omega$ belongs to the Lie algebra $\mathfrak{g}$. Thus the second term in the right hand side of eq. (4.55) is an element of the representation r of $\mathfrak{su}(N)$. For consistency, the first term should also be in this algebra, which implies that A_μ is also an element of the representation r of the Lie algebra, that we can decompose as follows:

$$A_\mu(x) \equiv A_\mu^a(x)\, t_r^a\,, \tag{4.56}$$

where the t_r^a are the generators of the algebra in this representation.

Like in quantum electrodynamics, we see that imposing the invariance of the Dirac Lagrangian under local transformations (4.52) leads to the necessity of introducing a vector field $A_\mu(x)$ whose transformation law is given in eq. (4.55). This requirement also completely specifies the form of the coupling between the fields ψ and A_μ:

$$\mathcal{L}_I = -ig\, \overline{\psi}_i \gamma^\mu\, A_\mu^a \left(t_r^a\right)_{ij} \psi_j\,, \tag{4.57}$$

where we have written explicitly the $\mathfrak{su}(N)$ indices i, j of all the objects. These indices, that run from 1 to the size of the representation r, label the "charge" (under the group $SU(N)$) carried by the fermions, while the index a may be viewed as the charge carried by the spin-1 particle associated to the vector field A_μ^a (this index runs from 1 to the dimension of the group, i.e. $N^2 - 1$ for $SU(N)$).

---

[6]From this transformation law, we see that field configurations of the form $ig^{-1}\Omega^\dagger \partial_\mu \Omega$ may be transformed into the null field $A_\mu \equiv 0$. Such configurations are called *pure gauge fields*.

Infinitesimal transformations :   Eqs. (4.50) and (4.55) specify how the fields ψ and A_μ change under any transformation of $\mathcal{G}$. However, it is sometimes useful to consider infinitesimal transformations, i.e. Ω close to 1. This is done by writing $\Omega = \exp(i g\, \theta_a t_r^a)$, with $|\theta_a| \ll 1$, and by expanding eqs. (4.50) and (4.55) to order one in θ_a. The variations of the fields ψ and A_μ are given by:

$$\delta\psi(x) = -i g\, \theta_r(x)\, \psi(x)$$
$$\delta A_\mu = -\partial_\mu \theta_r(x) + i\, g\, [A_\mu(x), \theta_r(x)] = -[D_\mu, \theta_r(x)] \,, \tag{4.58}$$

where we have defined $\theta_r \equiv \theta_a t_r^a$. The second one can be written more explicitly as

$$\delta A_\mu^a = -\partial_\mu \theta_a(x) + g\, f^{abc}\, \theta_b(x)\, A_\mu^c(x) = -\left(D_\mu^{\text{adj}}\right)_{ab} \theta_b(x) \,, \tag{4.59}$$

where D_μ^{adj} is the covariant derivative in the adjoint representation.

4.4   Field strength

In the previous section, we have constructed a Dirac Lagrangian which is invariant under local transformations of $\mathcal{G}$, by introducing a vector field A_μ that we shall interpret as a spin-1 particle that mediates the gauge interaction. But so far, this field enters only in the interaction term of eq. (4.57). We therefore need to construct a kinetic term for A_μ, with the constraint that it is invariant under the transformations (4.55). In the case of quantum electrodynamics, this Lagrangian was $-\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$, where the field strength was defined as $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$. However, a direct verification indicates that this expression of the field strength cannot lead to an invariant Lagrangian.

In order to mimic QED, we aim at constructing a Lagrangian with second order derivatives. Indeed, since the field $A_\mu(x)$ has the dimension of a mass, two derivatives and two powers of the field would provide the required dimension 4 for a Lagrangian in four space-time dimensions. A useful intermediate step is the construction of a field that depends only on $A_\mu(x)$ and has a simple transformation law. From the transformation law of the covariant derivative, we find that the commutator $[D_\mu, D_\nu]$ transforms as

$$[D_\mu, D_\nu] \quad \to \quad \Omega^\dagger(x)\, [D_\mu, D_\nu]\, \Omega(x) \,. \tag{4.60}$$

More explicitly, this commutator reads

$$[D_\mu, D_\nu] = -i g\, \Big(\underbrace{\partial_\mu A_\nu - \partial_\nu A_\mu - i g\, [A_\mu, A_\nu]}_{F_{\mu\nu}} \Big) \,. \tag{4.61}$$

This generalizes the *field strength* $F_{\mu\nu}$ to an arbitrary gauge group $\mathcal{G}$. Note the extra term, made of a commutator between gauge fields, that did not exist in QED. By construction, the field strength is an element of algebra, in the same representation as A_μ,

$$F_{\mu\nu}(x) \equiv F_{\mu\nu}^a(x)\, t_r^a \,, \tag{4.62}$$

and its transformation law[7] is

$$F_{\mu\nu}(x) \quad \to \quad \Omega^\dagger(x)\, F_{\mu\nu}(x)\, \Omega(x) \,. \tag{4.63}$$

---

[7]The field strength associated to a pure gauge field is zero, since there exists a transformation Ω for which A_μ becomes the null field.

As in QED, one may define electrical and magnetic fields[8] by

$$\mathbf{E}_a^i = F_a^{0i} \quad , \qquad \mathbf{B}_a^i = \tfrac{1}{2}\,\epsilon^{ijk}\,F_a^{jk} \,, \tag{4.64}$$

but with a few important differences: these $\mathbf{E}$ and $\mathbf{B}$ fields carry a group index, and they are not gauge invariant. Instead, they transform covariantly under a gauge transformation:

$$\mathbf{E}^i(x) \quad \rightarrow \quad \Omega^\dagger(x)\,\mathbf{E}^i(x)\,\Omega(x)\,, \quad \mathbf{B}^i(x) \quad \rightarrow \quad \Omega^\dagger(x)\,\mathbf{B}^i(x)\,\Omega(x)\,. \tag{4.65}$$

In order to build a kinetic term for A_μ from $F_{\mu\nu}$, we must contract all the Lorentz indices to have a Lorentz invariant Lagrangian. This forces us to have at least two F's, since $g^{\mu\nu}\,F_{\mu\nu} = 0$. Therefore, if we restrict to objects of mass dimension 4, this kinetic term should be quadratic in F. Moreover, a trace is necessary in order to obtain a gauge invariant quantity. Therefore, we arrive at the following Lagrangian[9] for the field A_μ,

$$\mathcal{L}_A \equiv -\frac{1}{2}\,\mathrm{tr}\left(F_{\mu\nu}(x)F^{\mu\nu}(x)\right) = -\frac{1}{4}\,F_{\mu\nu}^a(x)F_a^{\mu\nu}(x)\,. \tag{4.66}$$

Despite its resemblance with the photon kinetic term in QED, this Lagrangian has a quite remarkable feature: due to the commutator term in $F_{\mu\nu}$, $\mathcal{L}_A$ contains terms that are cubic in A_μ and terms which are quartic in A_μ. These terms are interactions between three and four of the spin-1 particles described by A_μ, respectively. Thus, unlike in QED, the Lagrangian (4.66) has a very rich structure, and defines in itself a very interesting quantum field theory, called *Yang-Mills theory*.

4.5 Non-Abelian gauge theories

A *non-Abelian gauge theory* is a quantum field theory that has at least a gauge field whose symmetry group is a non-Abelian group $\mathcal{G}$. Thus, the Lagrangian of all non-Abelian gauge theories contains a Yang-Mills term:

$$\mathcal{L}_A \equiv -\frac{1}{2}\,\mathrm{tr}\left(F_{\mu\nu}(x)F^{\mu\nu}(x)\right)\,. \tag{4.67}$$

If the gauge potential A_μ is the only field of the theory, then it is called a *Yang-Mills theory*. However, all useful gauge theories in particle physics also have matter fields. For each spin 1/2 matter field, the Dirac Lagrangian

$$\mathcal{L}_D = \overline{\psi}(x)\left(i\slashed{D}_x - m\right)\psi(x) \tag{4.68}$$

must be added to the Lagrangian. Two important interactions in Nature are described by non-Abelian gauge theories of this type:

- *Quantum chromodynamics*, the quantum field theory of strong interactions, is of this type: the gauge fields are the gluons, and the matter fields are the quarks, of which exist 6 families, or flavors (up, down, strange, charmed, bottom, top). The charge associated to this gauge interaction is called *color*. The gauge group of QCD is $SU(3)$, and the quarks live in the fundamental representation (therefore, they can have three different colors). In QCD, the gluons interact equally with the right-handed and left-handed projections of the spinors: it is said to be *non-chiral*.

---

[8] In Quantum Chromodynamics, these fields are called the *chromo-electric* and *chromo-magnetic* fields.
[9] We ignore for now the operator $\epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\left(F^{\mu\nu}F^{\rho\sigma}\right)$. This term will be discussed later in the section 4.7.

- Likewise, the *Electroweak theory* is a non-Abelian gauge theory with the gauge group $U(1) \times SU(2)$, but with the peculiarity that the $SU(2)$ acts only on the left-handed projection of the fermions. In other words, the right-handed fermions belong to the singlet representation of $SU(2)$ (while the left-handed fermions are arranged in doublets, in representations of dimension 2).

It is also possible to couple a (charged) scalar field $\phi(x)$ to a gauge potential A_μ. Under a local gauge transformation, ϕ transforms as follows

$$\phi(x) \quad \rightarrow \quad \Omega^\dagger(x)\,\phi(x)\,, \tag{4.69}$$

and therefore the following Lagrangian density is invariant under local gauge transformations:

$$\mathcal{L}_{\text{scalar}} = \big(D_\mu \phi(x)\big)^\dagger \big(D^\mu \phi(x)\big) - m^2\,\phi^\dagger(x)\phi(x) - V\big(\phi^\dagger(x)\phi(x)\big)\,. \tag{4.70}$$

(The potential should depend on the scalar field via the combination $\phi^\dagger \phi$ in order to be gauge invariant). The most important example of such scalar in particle physics is the *Higgs boson*.

4.6 Classical equations of motion

From the Lagrangians (4.67), (4.68) and (4.70), it is straightforward to obtain the classical Euler-Lagrange equations of motion. For the fermions, we simply obtain the Dirac equation

$$\big(i\slashed{D} - m\big)\,\psi(x) = 0\,. \tag{4.71}$$

For scalar fields, the classical equation of motion is a deformation of the Klein-Gordon equation, in which the ordinary derivatives are replaced by covariant derivatives:

$$\Big[D_\mu D^\mu + m^2 + V'\big(\phi^\dagger(x)\phi(x)\big)\Big]\,\phi(x) = 0\,. \tag{4.72}$$

For the gauge field A_μ, the derivatives of the various pieces of the Lagrangian read:

$$\partial_\mu \frac{\partial \mathcal{L}_A}{\partial(\partial_\mu A^a_\nu)} = -F^{a\,\mu\nu}\,,$$

$$\frac{\partial \mathcal{L}_A}{\partial A^a_\nu} = g\,f^{abc}\,A^b_\mu\,F^{c\,\mu\nu}\,,$$

$$\frac{\partial \mathcal{L}_D}{\partial A^a_\nu} = g\,\overline{\psi}\,\gamma^\nu\,t^a\,\psi\,,$$

$$\frac{\partial \mathcal{L}_{\text{scalar}}}{\partial A^a_\nu} = ig\Big(\phi^\dagger\,t^a\big(D_\nu\phi\big) - \big(D_\nu\phi\big)^\dagger\,t^a\,\phi\Big)\,. \tag{4.73}$$

This leads to the following equation of motion

$$\big[D_\mu, F^{\mu\nu}\big]_a = -J^\nu_a\,,$$

$$J^\nu_a = g\,\overline{\psi}\,\gamma^\nu\,t^a\,\psi + ig\Big(\phi^\dagger\,t^a\big(D_\nu\phi\big) - \big(D_\nu\phi\big)^\dagger\,t^a\,\phi\Big)\,, \tag{4.74}$$

known as *Yang-Mills equation*. From the Dirac and Klein-Gordon equations, one may check that the color current J^ν_a is *covariantly conserved*:

$$\big[D_\nu, J^\nu\big] = 0\,. \tag{4.75}$$

The field strength also obeys another, homogeneous, equation,

$$\big[D_\mu, F^{\nu\rho}\big] + \big[D_\nu, F^{\rho\mu}\big] + \big[D_\rho, F^{\mu\nu}\big] = 0\,, \tag{4.76}$$

that follows from the Jacobi identity between covariant derivatives.

4.7 θ-term and strong-CP problem

4.7.1 CP-odd gauge invariant operator

In the construction of the Lagrangian of Yang-Mills theory, we have argued that the only dimension four gauge invariant local operator is an operator quadratic in the field strength $F_a^{\mu\nu}$. All the Lorentz indices should be contracted in order to obtain a Lorentz invariant Lagrangian density. An obvious possibility is $F_a^{\mu\nu} F_{\mu\nu}^a$, which is the combination that appears in the Yang-Mills action. However, there exist another Lorentz invariant contraction, obtained by introducing the Levi-Civita tensor,

$$\mathcal{L}_\theta \equiv \frac{g^2\theta}{32\pi^2}\,\epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\,(F^{\mu\nu}\,F^{\rho\sigma})\;. \tag{4.77}$$

The prefactor $1/32\pi^2$ will appear convenient later, and the coupling constant in front of this term is usually denoted θ. Consequently, this term is referred to as the θ-*term*.

4.7.2 Expression as a total derivative

Firstly, we should clarify why we have not considered this term right away when we listed the possible gauge invariant operators that may enter in a non-Abelian gauge theory. As we shall prove now, the θ-term is a total derivative. Therefore, it does not enter in the field equations of motion, and has also no influence on perturbation theory. Since our discussion has been so far centered on the perturbative expansion, this term was irrelevant. However, the θ-term –that we cannot exclude on the grounds of symmetries– may lead to non-perturbative effects that we shall discuss in this section.

Let us consider the following vector[10]:

$$K^\mu \equiv \epsilon^{\mu\nu\rho\sigma}\left[A_\nu^a F_{\rho\sigma}^a - \frac{g}{3}\,f^{abc}\,A_\nu^a A_\rho^b A_\sigma^c\right]\;. \tag{4.78}$$

The divergence of this vector is given by

$$
\begin{aligned}
\partial_\mu K^\mu &= \epsilon^{\mu\nu\rho\sigma}\Big[(\partial_\mu A_\nu^a)(\partial_\rho A_\sigma^a - \partial_\sigma A_\rho^a + g f^{abc} A_\rho^b A_\sigma^c)\\
&\quad + A_\nu^a(\partial_\mu\partial_\rho A_\sigma^a - \partial_\mu\partial_\sigma A_\rho^a + g f^{abc}(\partial_\mu A_\rho^b)A_\sigma^c + g f^{abc}A_\rho^b(\partial_\mu A_\sigma^c))\\
&\quad - \frac{g}{3}f^{abc}(\partial_\mu A_\nu^a)A_\rho^b A_\sigma^c - \frac{g}{3}f^{abc}A_\nu^a(\partial_\mu A_\rho^b)A_\sigma^c - \frac{g}{3}f^{abc}A_\nu^a A_\rho^b(\partial_\mu A_\sigma^c)\Big]\\
&= \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}\Big[F_{\mu\nu}^a F_{\rho\sigma}^a\\
&\quad - g^2 f^{abc}f^{ade}A_\mu^b A_\nu^c A_\rho^d A_\sigma^e\\
&\quad + \frac{g}{3}f^{abc}\Big(A_\mu^b A_\nu^c(\partial_\rho A_\sigma^a - \partial_\sigma A_\rho^a) - A_\rho^b A_\sigma^c(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a)\Big)\Big]\;. \tag{4.79}
\end{aligned}
$$

---

[10]Note that this vector can also be expressed as a trace:

$$K^\mu \equiv 2\epsilon^{\mu\nu\rho\sigma}\,\mathrm{tr}\left[A_\nu F_{\rho\sigma} + \frac{2ig}{3}A_\nu A_\rho A_\sigma\right]\;.$$

The two terms of the third line are antisymmetric under the exchange $(\mu\nu) \leftrightarrow (\rho\sigma)$, while the prefactor $\epsilon^{\mu\nu\rho\sigma}$ is symmetric under this exchange. These terms are therefore zero after summing over the indices $\nu\rho\sigma\mu$. Then, the term on the second line can be written as follows:

$$g^2 \epsilon^{\mu\nu\rho\sigma} \, \text{tr} \, ([A_\mu, A_\nu][A_\rho, A_\sigma])$$
$$= g^2 \epsilon^{\mu\nu\rho\sigma} \, \text{tr} \, (A_\mu A_\nu A_\rho A_\sigma + A_\nu A_\mu A_\sigma A_\rho - A_\nu A_\mu A_\rho A_\sigma - A_\mu A_\nu A_\sigma A_\rho) \,. \tag{4.80}$$

Each term is a trace of four factors, and is invariant under cyclic permutations of the indices. Since cyclic permutations are odd in four dimensions, the $\epsilon_{\mu\nu\rho\sigma}$ tensor changes sign under such a permutation, and the contraction with the trace is zero. Therefore, we obtain:

$$\partial_\mu K^\mu = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} \, F^a_{\mu\nu} \, F^a_{\rho\sigma} \,, \tag{4.81}$$

which is proportional to the θ-term. More precisely, we have

$$\mathcal{L}_\theta = \frac{g^2 \theta}{32\pi^2} \, \partial_\mu K^\mu \,. \tag{4.82}$$

4.7.3   Effect of the θ-term on the Euclidean path integral

We have already encountered the integral of the θ-term over Euclidean spacetime in the context of anomalies and the Atiyah-Singer theorem (see eq. (3.129)):

$$\int d^4 x_{\scriptscriptstyle E} \, \mathcal{L}_\theta = n\theta \quad, \qquad n \in \mathbb{Z} \,, \tag{4.83}$$

where the integer n is related to the zero modes of the Dirac operator in the gauge field configuration. When added to the Yang-Mills action, the integral of the θ-term modifies the Euclidean path integral as follows

$$\int \left[DA_\mu \cdots \right] e^{-\mathcal{S}[A, \cdots]} \quad \rightarrow \quad \int \left[DA_\mu \cdots \right] e^{-\mathcal{S}[A, \cdots] - \int d^4 x_{\scriptscriptstyle E} \, \mathcal{L}_\theta}$$
$$= \sum_{n \in \mathbb{Z}} e^{-n\theta} \int \left[DA_\mu \cdots \right]_n e^{-\mathcal{S}[A, \cdots]} \,, \tag{4.84}$$

where the measure $\left[DA_\mu \right]_n$ is restricted to the gauge fields of index n. Thus, the effect of the θ-term is to reweight the gauge field configurations by a factor $(e^{-\theta})^n$ that depends only on θ and on the index n. Note that since n is an integer, the path integral is periodic in θ, with a period $2i\pi$.

4.7.4   Strong CP-problem

As we have seen in the section 3.5.6, an effective description of the interactions of nucleons with pions is provided by the linear σ model, whose interaction term is

$$\mathcal{L}_{\scriptscriptstyle I} \equiv \lambda \, \overline{\psi}(\sigma + i\boldsymbol{\pi} \cdot \boldsymbol{\tau} \gamma^5)\psi \,. \tag{4.85}$$

110

However, this does not include any CP-violating interactions, such as those that may result by the θ-term. Its effects may be included in the effective theory by generalizing the interaction term into

$$\mathcal{L}_{\text{I}} \equiv \overline{\psi}(\sigma + \boldsymbol{\pi} \cdot \boldsymbol{\tau}(i\lambda\gamma^5 + \overline{\lambda}))\psi \ . \tag{4.86}$$

By a matching with the underlying theory, the new coupling $\overline{\lambda}$ can be related to the parameter θ by the following estimate

$$|\overline{\lambda}| \approx 0.038 \ |\theta| \ . \tag{4.87}$$

Then, the effective theory (4.86) can be used to estimate the neutron electric dipole moment D_{N} (in the chiral limit where the pion mass m_π is much smaller than the nucleon mass m_{N}. This leads to

$$D_{\text{N}} \approx \lambda\overline{\lambda}\,e\,\frac{\ln\left(\frac{m_{\text{N}}}{m_\pi}\right)}{4\pi^2 m_{\text{N}}} \approx 5 \times 10^{-16}\,\theta e \text{ cm} \ . \tag{4.88}$$

Current experimental limits on the neutron electric dipole moment indicate that

$$\left|D_{\text{N}}\right| \leq 3 \times 10^{-26} \text{ e cm} \ , \tag{4.89}$$

implying that

$$|\theta| \lesssim 10^{-10} \ . \tag{4.90}$$

We thus face a paradoxical situation. The gauge symmetry of quantum chromodynamics allows the addition of the θ-term to the Yang-Mills action, and without any prior knowledge of the coupling θ, one may expect that natural values are of order unity. This constitutes the *strong-CP problem*: lacking a symmetry principle that would force θ to be zero, why is it nevertheless extremely small?

Note that there is an interesting interplay between the θ-term and chiral transformations of quark fields:

$$\psi_f \longrightarrow e^{i\gamma_5 \alpha_f}\psi_f \ , \tag{4.91}$$

where f is an index labeling the quark flavors and the α_f are real phases. Under this transformation, the functional measure for the quarks is not invariant, but transforms as follows

$$[D\psi D\overline{\psi}] \longrightarrow \exp\left(-\frac{i}{32\pi^2}\int d^4x \ \epsilon^{\mu\nu\rho\sigma}F^a_{\mu\nu}F^a_{\rho\sigma}\sum_f \alpha_f\right)[D\psi D\overline{\psi}] \ . \tag{4.92}$$

The same effect would have been obtained by a change of the angle θ:

$$\theta \to \theta - 2\sum_f \alpha_f \ . \tag{4.93}$$

For the quarks, we can write generically the following mass term[11]

$$\sum_f M_f \ \overline{\psi}_f\frac{1+\gamma_5}{2}\psi_f + \sum_f M_f^* \ \overline{\psi}_f\frac{1-\gamma_5}{2}\psi_f \ , \tag{4.94}$$

---

[11] If the masses are complex, then the symmetries P and CP are explicitly broken.

that transforms into the following under the above chiral transformation

$$\sum_f e^{2i\alpha_f} M_f \,\overline{\psi}_f \frac{1+\gamma_5}{2} \psi_f + \sum_f e^{-2i\alpha_f} M_f^* \,\overline{\psi}_f \frac{1-\gamma_5}{2} \psi_f \;.\tag{4.95}$$

This is equivalent to transforming the quark masses as follows:

$$M_f \to e^{2i\alpha_f} M_f \;.\tag{4.96}$$

Since any change of θ can be absorbed by a chiral transformation of the quarks, whose effect is to multiply the quark masses by phases, physical quantities cannot depend separately on θ and on the quark masses. Instead, they can depend only on the following combination

$$e^{i\theta} \prod_f M_f \;,\tag{4.97}$$

which is invariant. This discussion indicates that the θ-term has no effect if at least one of the quarks is massless. Unfortunately, a massless up quark (the lightest quark) does not seem consistent with existing experimental and lattice evidence...

4.7.5   Link with the topology of gauge fields

Using Stokes' theorem, the integral of the θ-term over Euclidean spacetime may be rewritten as an integral over a surface localized at infinity:

$$\int d^4 x_{\scriptscriptstyle E} \,\mathcal{L}_\theta = \frac{g^2\theta}{32\pi^2} \int d^4 x_{\scriptscriptstyle E} \,\partial_\mu K^\mu = \frac{g^2\theta}{32\pi^2} \lim_{R\to\infty} \int_{\mathcal{S}_{3,R}} dS_\mu \,K^\mu \;,\tag{4.98}$$

where $\mathcal{S}_{3,R}$ is a 3-dimensional sphere of radius R and dS_μ the measure on this surface.

Let us now assume that the colored objects of the problem are comprised in a finite region of space-times, so that the gauge field configuration goes to a pure gauge at infinity. Such a field can be written as

$$A_\mu(x) = a_\mu(x) + \frac{i}{g}\,\Omega^\dagger(\widehat{x})\,\partial_\mu\,\Omega(\widehat{x}) \;,\tag{4.99}$$

where $\Omega(\widehat{x})$ is an element of the gauge group that depends only on the direction of the vector x^μ, and $a_\mu(x)$ the deviation from the pure gauge form. For the total field to be a pure gauge at infinity, this deviation must decrease faster than $|x|^{-1}$. When $|x| \to +\infty$, $A_\nu(x)$ goes to 0 as $|x|^{-1}$, while $F_{\rho\sigma}(x)$ goes to 0 *faster* than $|x|^{-2}$ (since $A_\nu(x)$ goes to a pure gauge), and we have:

$$K_\mu \xrightarrow[|x|\to+\infty]{} \frac{4ig}{3}\epsilon^{\mu\nu\rho\sigma}\,\mathrm{tr}\left(A_\nu A_\rho A_\sigma\right) \;,\tag{4.100}$$

and

$$\int d^4 x_{\scriptscriptstyle E} \,\mathcal{L}_\theta = \frac{\theta}{24\pi^2} \lim_{R\to\infty} \int_{\mathcal{S}_{3,R}} dS\,\widehat{x}_\mu\,\epsilon^{\mu\nu\rho\sigma}\,\mathrm{tr}\left(\Omega^\dagger(\partial_\nu\Omega)\Omega^\dagger(\partial_\rho\Omega)\Omega^\dagger(\partial_\sigma\Omega)\right) \;,\tag{4.101}$$

where we have used $dS_\mu = \widehat{x}_\mu \, dS$, with dS the element of area on the 3-sphere. Note that the integrand decreases as R^{-3} because of the three derivatives, while $dS \sim R^3$. Therefore, the integral is in fact independent of the radius R and we can drop the limit:

$$\int d^4 x_{\scriptscriptstyle E} \, \mathcal{L}_\theta = \frac{\theta}{24\pi^2} \int_{\mathcal{S}_3} dS \, \widehat{x}_\mu \, \epsilon^{\mu\nu\rho\sigma} \, \mathrm{tr} \left(\Omega^\dagger(\partial_\nu \Omega) \Omega^\dagger(\partial_\rho \Omega) \Omega^\dagger(\partial_\sigma \Omega) \right) . \qquad (4.102)$$

Thus, the integral of the θ-term depends only on the function $\Omega(\widehat{x})$, that maps the 3-dimensional sphere $\mathcal{S}_3$ onto the gauge group:

$$\Omega : \quad \mathcal{S}_3 \quad \longmapsto \quad \mathcal{G} \, . \qquad (4.103)$$

It turns out that these mappings can be grouped in classes of Ω's that can be deformed continuously into one another. On the contrary, Ω's that belong to distinct classes cannot be related by a continuous deformation. The set of these classes possesses a group structure, and is called the *third homotopy group* of $\mathcal{G}$, denoted $\pi_3(\mathcal{G})$. For all $SU(N)$ groups with $N \geq 2$, the third homotopy group is isomorphic to $(\mathbb{Z}, +)$. The interpretation of eq. (4.102) is that the integral of the right hand side depends only on the class to which Ω belongs, and is therefore a topological quantity.

4.8 Non-local gauge invariant operators

4.8.1 Two-fermion non-local operator

The discussion in the previous sections exhausts the *local* gauge invariant objects of dimension less than or equal to 4. However, it is sometimes useful to construct gauge invariant non-local operators, for instance in the definition of *parton distributions*. The simplest operator of this type is an operator with two spinor fields at different space-time positions, $\overline{\psi}(y) \, W(y, x) \, \psi(x)$. Since the transformation laws of the two spinors involve different Ω's, such an operator is gauge invariant only if the non-written object between the spinors transforms as follows:

$$W(y, x) \quad \to \quad \Omega^\dagger(y) \, W(y, x) \, \Omega(x) \, . \qquad (4.104)$$

4.8.2 Wilson lines

In order to construct such an object, let us define a path $\gamma^\mu(s)$ that goes from x to y,

$$\gamma^\mu(0) = x^\mu \, , \quad \gamma^\mu(1) = y^\mu \, , \qquad (4.105)$$

and consider the following differential equation

$$\frac{DW}{ds} \equiv \frac{d\gamma^\mu}{ds} D_\mu(\gamma(s)) \, W = 0 \, , \quad \text{with initial condition } W(0) = 1 \, , \qquad (4.106)$$

where the notation $D_\mu(\gamma(s))$ indicates that the gauge field in the covariant derivative must be evaluated at the point $\gamma^\mu(s)$. In other words, the covariant derivative of W, projected along the tangent vector to the path $\gamma^\mu(s)$, is zero. From this definition, it follows that $W(s)$ is

an element of the representation r of the gauge group if A_μ is in the representation r of the algebra.

Note that when the gauge field A_μ is zero everywhere, then the solution is trivially $W(s) = 1$. The value of the solution[12] at $s = 1$ is a property of the path γ^μ and of the gauge potential A^μ. This object, that we will denote as

$$\mathcal{W}_{yx}[A; \gamma] \equiv W(1) , \tag{4.107}$$

is called a *Wilson line*. Let us now study how it changes under a gauge transformation Ω. From the transformation law of the covariant derivative, the differential equation that defines the transformed $W_\Omega(s)$ is

$$\frac{d\gamma^\mu}{ds} \Omega^\dagger(\gamma(s)) D_\mu(\gamma(s)) \Omega(\gamma(s)) W_\Omega(s) = 0 , \quad \text{with initial condition } W_\Omega(0) = 1 . \tag{4.108}$$

If we define $Z(s) \equiv \Omega(\gamma(s)) W_\Omega(s)$, this equation is equivalent to

$$\frac{d\gamma^\mu}{ds} D_\mu(\gamma(s)) Z(s) = 0 , \quad \text{with initial condition } Z(0) = \Omega(x) . \tag{4.109}$$

Comparing this equation with the original equation (4.106), we obtain

$$Z(s) = W(s) \Omega(x) , \quad \text{i.e. } W_\Omega(s) = \Omega^\dagger(\gamma(s)) W(s) \Omega(x) . \tag{4.110}$$

Looking now at the point $s = 1$, we see that the Wilson line transforms as

$$\mathcal{W}_{yx}[A; \gamma] \quad \rightarrow \quad \Omega^\dagger(y) \mathcal{W}_{yx}[A; \gamma] \Omega(x) . \tag{4.111}$$

Thus, the Wilson line transforms precisely as we wanted in eq. (4.104), and we conclude that the operator $\overline{\psi}(y) \mathcal{W}_{yx}[A; \gamma] \psi(x)$ is gauge invariant. Note that the Wilson line $\mathcal{W}_{yx}[A; \gamma]$, solution of eq. (4.106) at $s = 1$, can also be written as a *path-ordered exponential*,

$$\mathcal{W}_{yx}[A; \gamma] = P \exp \left(ig \int_\gamma dx^\mu A_\mu(x) \right) . \tag{4.112}$$

Although this compact notation is suggestive, it is often useful to revert to the defining differential equation (4.106).

4.8.3 Path dependence

By inserting a Wilson line between the points x and y, we can construct a gauge invariant non-local operator $\overline{\psi}(y) \cdots \psi(x)$. However, in doing so, we have introduced a path γ, for which there are infinitely many possible choices since only its endpoints are fixed. It turns out that in general, the Wilson line depends on the path γ, i.e.

$$\mathcal{W}_{yx}[A; \gamma] \neq \mathcal{W}_{yx}[A; \gamma'] . \tag{4.113}$$

This implies that, although we may define gauge invariant non-local bilinear operators, their definition is not unique and each choice of the path connecting the two points leads to a different operator.

---

[12]Note that if the initial condition is $W(0) = \Omega_0$ instead of 1, then the solution would be changed as follows $W(s) \rightarrow W(s)\Omega_0$.

4.8.4 Case of pure gauge fields

When the gauge potential is a pure gauge field, there exists a function $\Omega(x)$ such that

$$A_\mu^\Omega(x) = \frac{i}{g}\,\Omega^\dagger(x)\,\partial_\mu\Omega(x)\ . \tag{4.114}$$

Since this field is a gauge transformation of the null field $A_\mu \equiv 0$, Wilson lines in this pure gauge field are given by

$$\mathcal{W}_{yx}[A^\Omega;\gamma] = \Omega^\dagger(y)\,\Omega(x)\ . \tag{4.115}$$

In other words, in a pure gauge field, the Wilson lines depend only on their endpoints, but not on the path chosen to connect them. This is the only exception to the remark of the previous paragraph.

Conversely, a gauge potential $A^\mu(x)$ in which the Wilson lines depend only on the endpoints is a pure gauge. A function $\Omega(x)$ that gives this gauge potential through eq. (4.114) can be constructed as a Wilson line from x to some arbitrary base point x_0:

$$\Omega(x) = \mathcal{W}_{x_0 x}[A;\gamma]\ . \tag{4.116}$$

(The path γ can be chosen arbitrarily.)

4.8.5 Wilson loops

A *Wilson loop* is a special kind of Wilson line, where the initial point and endpoint are identical, $x = y$, and therefore the path γ is a closed loop:

$$\mathcal{W}[A;\gamma] = P \exp\left(ig \oint_\gamma dx^\mu\, A_\mu(x)\right)\ . \tag{4.117}$$

Note that they are a property of the closed loop γ, and do not depend on the choice of the starting point x. Because they have identical endpoints, the trace of a Wilson loop is gauge invariant. From the result of the previous paragraph, they are equal to the identity in a pure gauge field, but they depend non-trivially on the path in a generic gauge field[13].

In Abelian gauge theories, the Wilson loop can be rewritten in terms of the integral of the field strength $F_{\mu\nu}$ over a surface Σ of boundary γ, by using Stokes theorem:

$$\exp\left(ig\oint_\gamma dx^\mu A_\mu(x)\right) \underset{\text{Abelian}}{=} \exp\left(i\frac{g}{2}\int_\Sigma dx^\mu \wedge dx^\nu\, F_{\mu\nu}(x)\right)\ . \tag{4.118}$$

Generalizations of this formula to the non-Abelian case exist, that involve a path-ordering in the left hand side (thus giving a Wilson loop) and a *surface-ordering* in the right hand side. For infinitesimally small closed loops, a more direct connection to the field strength may be established. Consider for instance a small square closed path in the (12) plane,

$$\gamma =$$

---
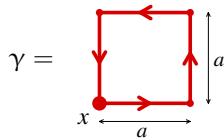
[13]Wilson loops are extensively used in lattice gauge theories. Moreover, Giles' theorem states that all the gauge invariant information contained in a gauge potential A_μ can be reconstructed from the trace of Wilson loops (assuming we know Wilson loops for arbitrary loops).

The Wilson loop along this path may be approximated by

$$
\begin{aligned}
\mathcal{W}[A;\gamma] \ \approx \ & \exp\left(-iga\,A^2(x+\tfrac{a}{2}\hat{\jmath})\right)\exp\left(-iga\,A^1(x+\tfrac{a}{2}\hat{\imath}+a\hat{\jmath})\right) \\
& \times \exp\left(iga\,A^2(x+a\hat{\imath}+\tfrac{a}{2}\hat{\jmath})\right)\exp\left(iga\,A^1(x+\tfrac{a}{2}\hat{\imath})\right), \quad (4.119)
\end{aligned}
$$

where we make an error of order a^3 on each of the Wilson lines at the edges of the square. By expanding the exponentials, we obtain

$$
\begin{aligned}
\mathcal{W}[A;\gamma] \ = \ & 1 + iga^2\left(\partial_1 A^2(x)-\partial_2 A^1(x)\right) \\
& \quad -g^2 a^2\left(A^2(x)A^1(x)-A^1(x)A^2(x)\right)+\mathcal{O}(a^3) \\
= \ & 1 + ig\,a^2\,F^{12}(x)+\mathcal{O}(a^3)\,. \quad (4.120)
\end{aligned}
$$

Thus, the first non-trivial correction to a small Wilson loop is the area of the loop times the field strength projected on the plane of the loop. Since $\mathcal{W}[A;\gamma]$ is an element of the representation r of the group, in the vicinity of the identity, it may be represented as

$$
\begin{aligned}
\mathcal{W}[A;\gamma] \ = \ & \exp\left(i\left(\epsilon\,\alpha^a t_r^a + \epsilon^2\,\beta^a t_r^a + \mathcal{O}(\epsilon^3)\right)\right) \\
= \ & \mathbf{1}_r + i\left(\epsilon\,\alpha^a t_r^a + \epsilon^2\,\beta^a t_r^a\right) - \frac{\epsilon^2}{2}\,\alpha^a\alpha^b\,t_r^a t_r^b + \mathcal{O}(\epsilon^3)\,, \quad (4.121)
\end{aligned}
$$

where ϵ is an infinitesimal parameter quantifying how close $\mathcal{W}[A;\gamma]$ is from the identity, and $\mathbf{1}_r$ is the identity matrix in the representation r. Comparing eqs. (4.120) and (4.121), we see that we must identity $\epsilon \equiv ga^2$ and $\alpha^a \equiv F_a^{12}(x)$. The formula (4.120) is insufficient in order to determine β^a, since this term gives a contribution of order a^4 in the Wilson loop. But we can nevertheless use eq. (4.121) to determine the lowest order correction to the trace of the Wilson loop,

$$
\mathrm{tr}\left(\mathcal{W}[A;\gamma]\right) = \mathrm{tr}\left(\mathbf{1}_r\right) - \frac{g^2\,a^4}{2}\,F_a^{12}(x)F_b^{12}(x)\,\mathrm{tr}\left(t_r^a t_r^b\right)+\mathcal{O}(a^6)\,, \quad (4.122)
$$

where we have used the fact that the generators t_r^a are traceless for the $\mathfrak{su}(N)$ algebra. Eq. (4.122) is the basis of the discretization of the Yang-Mills action, the first step in the formulation of *lattice gauge theories*.

Chapter 5

Quantization of Yang-Mills theory

5.1 Introduction

Generically, the Lagrangian density of non-Abelian gauge theory reads:

$$
\begin{aligned}
\mathcal{L} \;\equiv\; & \left(D_\mu\phi(x)\right)^\dagger\left(D^\mu\phi(x)\right) - m^2\,\phi^\dagger(x)\phi(x) - V\!\left(\phi^\dagger(x)\phi(x)\right) \\
& +\overline{\psi}(x)\left(i\slashed{D}_x - m\right)\psi(x) \\
& -\tfrac{1}{4}\,F^a_{\mu\nu}(x)F^{a\,\mu\nu}(x)\,.
\end{aligned}
\tag{5.1}
$$

The local non-Abelian gauge invariance of this Lagrangian does not change anything to the quantization of the scalar field ϕ and of the spinor ψ, for which we may use the standard canonical or path integral approaches, with the result that the usual Feynman rules still apply. The main complication resides in the pure Yang-Mills part (third term) of this Lagrangian, i.e. with the quantization of the gauge potential A^μ. The identification of the degrees of freedom that are made redundant by the gauge symmetry is much more complicated than in QED, and a lot more care is necessary in order to isolate the genuine dynamical variables of the theory.

In order to get a sense of the difficulty, let us try to mimic the QED case in order to guess the Feynman rules for non-Abelian gauge fields. Using the explicit form of the field strength,

$$
F^a_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + g\,f^{abc}\,A^b_\mu A^c_\nu\,,
\tag{5.2}
$$

we can rewrite the Yang-Mills Lagrangian as follows

$$
\begin{aligned}
\mathcal{L}_A \;=\; & \tfrac{1}{2}\,A^a_\mu\left(g^{\mu\nu}\Box - \partial^\mu\partial^\nu\right)A^a_\nu \\
& -g\,f^{abc}\left(\partial_\mu A^a_\nu\right)A^{b\,\mu}A^{c\,\nu} \\
& -\tfrac{1}{4}\,g^2\,f^{abc}f^{ade}\,A^b_\mu A^c_\nu A^{d\,\mu}A^{e\,\nu}\,,
\end{aligned}
\tag{5.3}
$$

where we have anticipated an integration by parts in the first (kinetic) term. Note that the kinetic term is formally identical to the kinetic term of a photons, except for the color index A carried by the gauge potential. Therefore, one may be tempted to generalize the QED Feynman rules to a non-Abelian gauge boson. As in the QED case, the quadratic part of

the Lagrangian (5.3) poses a difficulty when trying to determine the free propagator, because the operator between $A_\mu \cdots A_\nu$ is not invertible. If we take for granted that a similar gauge fixing procedure (more on this later, as this is in fact the heart of the problem) can be applied here, we may assume that the free gauge boson propagator[1] in Feynman gauge is

$$G^{0\,\mu\nu}_{F\,ab}(p) = \quad \overrightarrow{p} \quad \text{(gluon line)} \quad = \frac{-i\,g^{\mu\nu}\delta_{ab}}{p^2 + i0^+}\,, \tag{5.4}$$

and one may read off directly from the Lagrangian (5.3) the following 3-gluon and 4-gluon vertices:

$$= \quad \begin{aligned} g\,f^{abc}\,\{ & g^{\mu\nu}\,(k-p)^\rho \\ & + g^{\nu\rho}\,(p-q)^\mu + g^{\rho\mu}\,(q-k)^\nu \} \end{aligned} \tag{5.5}$$

$$= \quad \begin{aligned} -i\,g^2\,\{ & f^{abe}f^{cde}\,(g^{\mu\rho}g^{\nu\sigma} - g^{\mu\sigma}g^{\nu\rho}) \\ & + f^{ace}f^{bde}\,(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\sigma}g^{\nu\rho}) \\ & + f^{ade}f^{bce}\,(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma}) \} \end{aligned} \tag{5.6}$$

All this seems fine, except for a rather subtle problem that would appear when using these perturbation theory: these Feynman rules lead to amplitudes that do not fulfill Ward identities, even when all the external colored particles are on their mass-shell. From the discussion of perturbative unitarity for amplitudes with external gauge bosons in 1.14.4, the lack of Ward identities seems to imply a violation of unitarity in perturbation theory. Since unitarity is one of cornerstones of any quantum theory, this is not a conclusion we are ready to accept, and we must conclude that something is missing in the above Feynman rules.

5.2 Gauge fixing

In our naive attempt to guess the Feynman rules appropriate for non-Abelian gauge bosons, we have implicitly assumed that the gauge fixing works in the same way as in QED, namely that the gauge fixing trivially leads to the factorization of an infinite factor in the path integral, with no other change to the degrees of freedom that are not constrained by the gauge condition. It turns out that this assumption is incorrect. Let us start from the path integral representation of the expectation value of some gauge invariant operator $\mathcal{O}(A_\mu)$:

$$\langle \mathcal{O} \rangle \equiv \int \left[DA^a_\mu(x)\right]\,\mathcal{O}(A_\mu)\,\exp\left\{i\underbrace{\int d^4x\,\left(-\frac{1}{4}F^a_{\mu\nu}F^{a\,\mu\nu}\right)}_{S_{YM}[A_\mu]}\right\}\,. \tag{5.7}$$

---

[1] In this chapter, we use the diagrammatic convention of QCD, where the gauge bosons (gluons) are represented as springs in Feynman diagrams. In the electroweak theory, it is more common to represent them as wavy lines, like the photon in QED.

Local gauge transformations of the field A_μ,

$$A_\mu(x) \quad \rightarrow \quad A_\mu^\Omega(x) \equiv \Omega^\dagger(x)\, A_\mu(x)\, \Omega(x) + \frac{i}{g}\, \Omega^\dagger(x)\, \partial_\mu \Omega(x)\,, \tag{5.8}$$

leave the action and the observable unchanged. Moreover, the functional measure is also invariant, since

$$\left[DA_{a\,\mu}^\Omega(x) \right] = \left[DA_{a\,\mu}(x) \right]\, \det\left[\left(\frac{\delta A_{a\,\mu}^\Omega(x)}{\delta A_{b\,\nu}(y)} \right) \right]\,, \tag{5.9}$$

where the determinant is the Jacobian of the "change of coordinates". Using eq. (4.44), this determinant can be rewritten as follows

$$\det\left[\left(\frac{\delta A_{a\,\mu}^\Omega(x)}{\delta A_{b\,\nu}(y)} \right) \right] \quad = \quad \det\left[g_\mu{}^\nu\, \delta(x-y)\, \left[\Omega_{\text{adj}}(x) \right]_{ab} \right] = 1\,, \tag{5.10}$$

since the group element Ω_{adj} is a unitary matrix. Therefore, there is a large amount of redundancy in the above path integral. By applying a gauge transformation, each configuration A_μ develops into a *gauge orbit* (see the figure 5.1), along which the physics is invariant. In order to eliminate this redundancy, we would like to impose a condition at every space-time point x on the gauge fields,

$$G^a\big(A_{a\,\mu}(x) \big) = 0\,, \tag{5.11}$$

in order to select a unique[2] field configuration along each orbit. Geometrically, the *gauge condition* (5.11) defines a manifold that intersects each orbit, as shown in the figure 5.1, and we choose this intersection as the representative of this field configuration.

5.3 Fadeev-Popov quantization and Ghost fields

Thus, we would like to split the integration measure in eq. (5.7) into a physical component in the manifold $G(A) = 0$, and a component along the gauge orbits that we should factor out. Unfortunately, achieving this in a non-Abelian gauge theory is far more complicated than in QED, because the modification of the gauge potential under a gauge transformation is non-linear. In order to see the difficulty, let us define

$$\Delta^{-1}[A_\mu] \equiv \int \left[D\Omega(x) \right]\, \delta[G^a(A_\mu^\Omega)]\,. \tag{5.12}$$

$\Delta[A_\mu]$ is the determinant of the derivative of the constraint $G(A_\mu)$ with respect to the gauge transformation Ω, at the point where $G(A_\mu) = 0$,

$$\Delta(A_\mu) = \det\left(\frac{\delta G^a}{\delta \Omega} \right)_{G^a(A_\mu^\Omega)=0}\,. \tag{5.13}$$

In QED, for linear gauge fixing conditions, this derivative (and therefore the determinant) is independent of the gauge field, and can be trivially factored out of the path integral. This is

---

[2]It turns out that this is not possible, due to the *Gribov ambiguity*: all gauge conditions of the form (5.11) have several solutions, called *Gribov copies*. However, only one of these solutions is a "small field", while the others are proportional to the inverse coupling g^{-1}. Since perturbation theory is an expansion around the vacuum (i.e. in the small field regime), these non-perturbatively large copies do not play any role in perturbation theory.

Figure 5.1: *Illustration of the gauge fixing procedure. The red lines represent the gauge field configurations spanned when varying Ω. The shaded surface is the manifold where the gauge condition is satisfied, and the black dots are the gauge-fixed field configurations.*

A^μ

$G(A^\mu) = 0$

gauge
fixed A^μ

gauge orbit
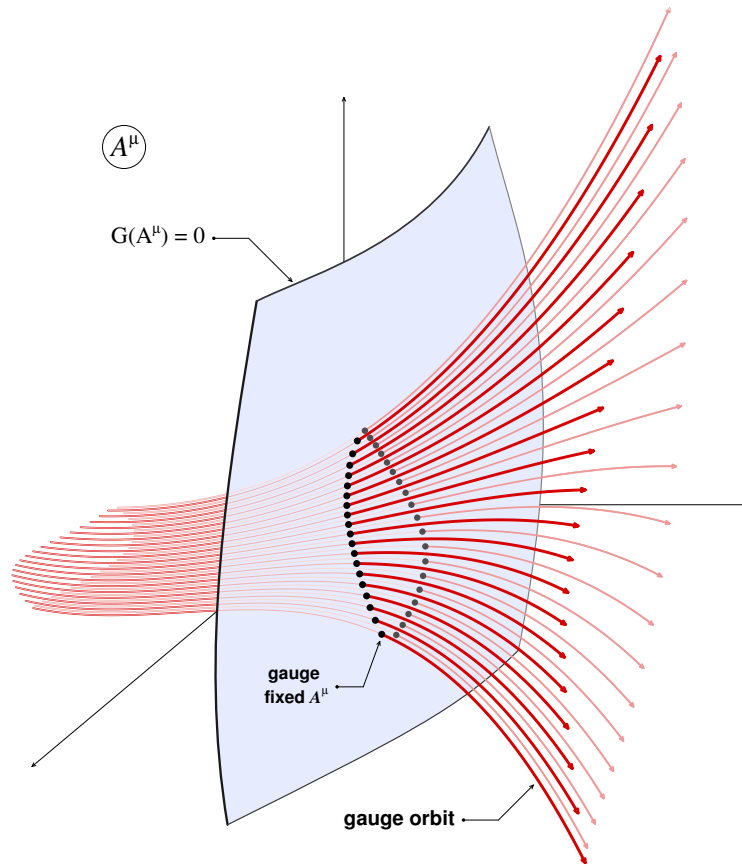
not the case in non-Abelian gauge theories, and this determinant is the source of significant complications. One can first prove that the determinant $\Delta[A_\mu]$ is gauge invariant. Indeed, changing $A_\mu \to A_\mu^\Theta$, we have:

$$
\begin{aligned}
\Delta^{-1}[A_\mu^\Theta] &= \int \left[D\Omega(x)\right]\, \delta[G^a(A_\mu^{\overbrace{\Theta\Omega}^{\Omega'}})] \\
&= \int \left[D(\Theta^\dagger(x)\Omega'(x))\right]\, \delta[G^a(A_\mu^{\Omega'})] \\
&= \int \left[D\Omega'(x)\right]\, \delta[G^a(A_\mu^{\Omega'})] = \Delta^{-1}[A_\mu]\,.
\end{aligned}
\tag{5.14}
$$

By inserting

$$
1 = \Delta[A_\mu] \int \left[D\Omega(x)\right]\, \delta[G^a(A_\mu^\Omega)]
\tag{5.15}
$$

inside the path integral (5.7), we obtain

$$
\langle \mathcal{O} \rangle = \int \left[D\Omega(x)\right] \int \left[DA_\mu^a(x)\right]\, \Delta[A_\mu]\, \delta[G^a(A_\mu^\Omega)]\, \mathcal{O}(A_\mu)\, e^{i\mathcal{S}_{\text{YM}}[A_\mu]}\,.
\tag{5.16}
$$

Now, we change the integration variable of the second integral according to $A_\mu \to A_\mu^{\Omega^\dagger}$. In this transformation, the measure $[DA_\mu]$, the Yang-Mills action $\mathcal{S}_{\text{YM}}[A_\mu]$, the observable $\mathcal{O}(A_\mu)$ and the determinant $\Delta[A_\mu]$ are all unchanged (because they are gauge invariant):

$$
\begin{aligned}
\left[DA_\mu^{\Omega^\dagger}\right] &= \left[DA_\mu\right], \\
\mathcal{S}_{\text{YM}}[A_\mu^{\Omega^\dagger}] &= \mathcal{S}_{\text{YM}}[A_\mu], \\
\mathcal{O}[A_\mu^{\Omega^\dagger}] &= \mathcal{O}[A_\mu], \\
\Delta[A_\mu^{\Omega^\dagger}] &= \Delta[A_\mu],
\end{aligned}
\tag{5.17}
$$

while the field A_μ^Ω becomes A_μ. Therefore, we have

$$
\langle \mathcal{O} \rangle = \int \left[D\Omega(x)\right] \int \left[DA_\mu^a(x)\right]\, \Delta[A_\mu]\, \delta[G^a(A_\mu)]\, \mathcal{O}(A_\mu)\, e^{i\mathcal{S}_{\text{YM}}[A_\mu]}\,.
\tag{5.18}
$$

At this point, the second integral does not contain the gauge transformation Ω anymore, and therefore we have managed to factorize the "integral along the orbits" in the form of the first integral over $\left[D\Omega\right]$. Dropping this constant factor, we can therefore write an integral free of any redundancy:

$$
\langle \mathcal{O} \rangle = \int \left[DA_\mu^a(x)\right]\, \Delta[A_\mu]\, \delta[G^a(A_\mu)]\, \mathcal{O}(A_\mu)\, e^{i\mathcal{S}_{\text{YM}}[A_\mu]}\,.
\tag{5.19}
$$

In the above formula, the determinant $\Delta[A_\mu]$ depends on the gauge field and must therefore have an effect on the Feynman rules. The *Fadeev-Popov method* consists in rewriting this determinant as a path integral. Note that since $\Delta[A_\mu]$ appears in the numerator, we need Grassmann variables in order to represent it as a path integral[3], according to eq. (3.36):

$$
\det\left(i\mathcal{M}\right) = \int \left[D\chi_a(x)D\overline{\chi}_a(x)\right]\, \exp\left\{ i \int d^4x\, d^4y\, \overline{\chi}_a(x)\, \mathcal{M}_{ab}(x,y)\, \chi_b(y) \right\}\,.
\tag{5.20}
$$

---

[3]The factor i in $\det\left(i\mathcal{M}\right)$ has been included for aesthetic reasons, but does not change anything. In fact any rescaling $\mathcal{M} \to \alpha\mathcal{M}$ would leave the results unchanged. Indeed, such a change would alter the ghost propagator according to $S \to \alpha^{-1}S$, and the ghost-gauge boson vertex by $V \to \alpha V$. Since the ghosts appear only in closed loops, that contain an equal number of propagators and vertices, these factors α would cancel out.

An extra generalization, that we have already used in the path integral quantization of the photon (see eq. (3.52)), is to shift the gauge condition from $G^a(A) = 0$ to $G^a(A) = \omega^a$ and to perform a Gaussian integration over ω^a. The final result takes the following form:

$$
\begin{aligned}
\langle \mathcal{O} \rangle \; = \; & \int \left[DA^a_\mu(x) \right] \left[D\chi_a(x) D\overline{\chi}_a(x) \right] \; \mathcal{O}(A_\mu) \\
& \times \exp i \int d^4x \left(\underbrace{-\frac{1}{4} F^a_{\mu\nu} F^{a\,\mu\nu}}_{\mathcal{L}_{YM}} \underbrace{-\frac{\xi}{2} (G^a(A_\mu))^2}_{\mathcal{L}_{GF}} + \underbrace{\overline{\chi}_a \, \mathcal{M}_{ab} \, \chi_b}_{\mathcal{L}_{FPG}} \right) ,
\end{aligned}
$$

(5.21)

where $\mathcal{M}_{ab}$ is the derivative of $G^a(A^\Omega)$ with respect to the gauge transformation Ω, at the point $\Omega = 1$ (here, we use the fact that the determinant is gauge invariant to choose freely the Ω at which we compute the derivative). The unphysical Grassmann fields χ and $\overline{\chi}$ introduced as a trick to express the determinant are called *Fadeev-Popov ghosts*, or simply ghosts. Although physical observables do not depend on these fictitious fields, there is in general a coupling between the ghosts and the gauge fields, because the matrix $\mathcal{M}_{ab}$ may contain the gauge field. This implies that the ghosts may appear in the form of loop corrections in the perturbative expansion. As we shall see shortly, they are in fact crucial for the consistency of perturbation theory in non-Abelian gauge theories. In particular, the ghosts ensure that the theory is unitary.

5.4 Feynman rules for non-abelian gauge theories

Eq. (5.21) contains all the necessary ingredients to complete the Feynman rules that we have started to derive heuristically at the beginning of this chapter. To turn this formula into explicit Feynman rules, we should first choose the gauge fixing function $G^a(A)$, since it enters directly in the term in $\frac{\xi}{2}(G^a(A))^2$, and implicitly in the matrix $\mathcal{M}_{ab}$ that defines the ghost term. In the common situation where this gauge fixing function is linear in A_μ (all our examples will be of this type), then the terms that are quadratic in the gauge field are the same as in QED, and therefore the gauge boson propagator is also the same (except for an extra factor δ_{ab} that expresses the fact that the free propagation of a gluon does not change its color). Thus our guess (5.4) for the Feynman gauge propagator was in fact correct. In addition, the gauge fixing term and the ghost term cannot contain terms of degree 3 or 4 in the gauge field, which implies that the vertices given in eqs. (5.5) and (5.6) are also correct.

5.4.1 Covariant gauge

Let us now consider the general covariant gauge, all known as the R_ξ-gauge, already introduced in eq. (3.51) for QED. This amounts to choosing the gauge fixing function as

$$
G^a(A) \equiv \partial^\mu A^a_\mu - \omega^a(x) .
$$

(5.22)

With this gauge fixing, the free gauge boson propagator is

$$
G^{0\,\mu\nu}_{F\,ab}(p) = \overrightarrow{}^{\;p} \;\;\; = \frac{-i\, g^{\mu\nu}\, \delta_{ab}}{p^2 + i0^+} + \frac{i\, \delta_{ab}}{p^2 + i0^+} \left(1 - \frac{1}{\xi} \right) \frac{p^\mu p^\nu}{p^2} .
$$

(5.23)

(The simplest form is obtained in the limit $\xi \to 1$, giving the Feynman gauge[4].) The matrix $\mathcal{M}_{ab}$ can be calculated by applying an infinitesimal gauge transformation $\Omega = \exp(i\theta_a t^a)$ to A_μ. The variation of the gauge field is

$$\delta A_{a\,\mu}(x) = g\,f^{abc}\,\theta_b(x)\,A_{c\,\mu}(x) - \partial_\mu \theta_a(x)\,, \tag{5.24}$$

and the variation of $G^a(A)$ at the point x is

$$\delta G^a = g\,f^{abc}\left(\partial^\mu \theta_b(x)\right) A_{c\,\mu}(x) + g\,f^{abc}\,\theta_b(x)\left(\partial^\mu A_{c\,\mu}(x)\right) - \Box\,\theta_a(x)\,. \tag{5.25}$$

Therefore, we have

$$\mathcal{M}_{ab} = \frac{\delta G^a(A)}{\delta \theta^b} = g\,f^{abc}\left(\partial^\mu A_{c\,\mu}(x)\right) + g\,f^{abc}\,A_{c\,\mu}(x)\,\partial^\mu - \delta_{ab}\,\Box\,, \tag{5.26}$$
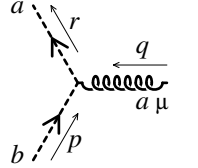
and the terms that depend on the Fadeev-Popov ghosts can be encapsulated in the following effective Lagrangian:

$$\mathcal{L}_{\text{FPG}} = \overline{\chi}_a \left(-\delta_{ab}\,\Box + g\,f^{abc}\left(\partial^\mu A_{c\,\mu}(x)\right) + g\,f^{abc}\,A_{c\,\mu}(x)\,\partial^\mu \right) \chi_b \tag{5.27}$$

The first term leads to the following propagator for the ghosts:

$$G^0_{\text{F}}(p) = \quad \overset{p}{\dashrightarrow}\!\!\!--- \quad = \frac{i\,\delta_{ab}}{p^2 + i0^+}\,. \tag{5.28}$$

Note that it has the form of a scalar propagator, although the ghosts are anti-commuting Grassmann variables. The vertex between ghosts and gauge bosons reads

$$= g\,f^{abc}\,(p_\mu + q_\mu) = g\,f^{abc}\,r_\mu\,. \tag{5.29}$$

The Feynman rules for non-Abelian gauge theories in covariant gauge are summarized in the figure 5.2, where we have added for completeness the rules relative to fermions.

5.4.2 Axial gauge

The *axial gauge* fixing consists in constraining the value of $n^\mu A^a_\mu$, where n^μ is a fixed 4-vector (when this vector is time-like, this gauge is called the *temporal gauge*, and when it is light-like, it is called the *light-cone gauge*). Therefore, the gauge fixing function is

$$G^a(A) \equiv n^\mu A^a_\mu - \omega^a(x)\,. \tag{5.30}$$

After gauge fixing, the quadratic part of the effective Lagrangian reads

$$\frac{1}{2}\,A^a_\mu \left(g^{\mu\nu}\Box - \partial^\mu \partial^\nu - \xi\,n^\mu n^\nu\right) A^a_\nu\,, \tag{5.31}$$

---

[4]Another popular choice is the *Landau gauge*, obtained in the limit $\xi \to +\infty$, that corresponds to a strict enforcement of the condition $\partial_\mu A^\mu = 0$. Indeed, in this limit the exponential of $i\frac{\xi}{2}(\partial_\mu A^\mu)^2$ in the gauge fixed Lagrangian oscillates wildly –and produces cancellations– unless $\partial_\mu A^\mu = 0$. Equivalently, the Gaussian distribution for the function $\omega^a(x)$ has a vanishing width in this limit, which forces the strict equality $\partial^\mu A^a_\mu = 0$.

Figure 5.2: *Feynman rules of non-Abelian gauge theories in covariant gauge. We also list the rules involving fermions for completeness. Latin characters* a, b, c *refer to the adjoint representation, while the letters* i, j *refer to the representation* r *in which the fermions live.*

$$\frac{p}{\text{ceccec}} \;=\; \frac{-i\,g^{\mu\nu}\,\delta_{ab}}{p^2 + i0^+} + \frac{i\,\delta_{ab}}{p^2 + i0^+}\left(1 - \frac{1}{\xi}\right)\frac{p^\mu p^\nu}{p^2}$$

$$\xrightarrow{\;p\;} \;=\; \frac{i\,\delta_{ij}}{\not{p} - m + i0^+}$$

$$\xrightarrow{\;p\;}\text{----} \;=\; \frac{i\,\delta_{ab}}{p^2 + i0^+}$$

$$=\; g\,f^{abc}\left\{g^{\mu\nu}\,(k-p)^\rho \right.$$
$$\left. + g^{\nu\rho}\,(p-q)^\mu + g^{\rho\mu}\,(q-k)^\nu\right\}$$

$$=\; -i\,g^2\left\{f^{abe}f^{cde}\,(g^{\mu\rho}g^{\nu\sigma} - g^{\mu\sigma}g^{\nu\rho})\right.$$
$$+ f^{ace}f^{bde}\,(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\sigma}g^{\nu\rho})$$
$$\left. + f^{ade}f^{bce}\,(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma})\right\}$$

$$=\; -i\,g\,\gamma^\mu\left(t_r^a\right)_{ij}$$

$$=\; g\,f^{abc}\,(p_\mu + q_\mu) = g\,f^{abc}\,r_\mu$$

and the free gauge boson propagator is obtained in momentum space by inverting

$$g^{\mu\nu}p^2 - p^\mu p^\nu + \xi\, n^\mu n^\nu \ . \tag{5.32}$$

The inverse of this matrix must be of the form

$$A\, g^{\mu\nu} + B\, p^\mu p^\nu + C\, n^\mu n^\nu + D\, (n^\mu p^\nu + n^\nu p^\mu) \ . \tag{5.33}$$

(This is the most general symmetric tensor that one may construct with $g^{\mu\nu}$, p^μ and n^μ.) This leads to the following propagator

$$G^{0\ \mu\nu}_{F\ ab}(p) = \frac{-i\,\delta_{ab}}{p^2 + i0^+}\left[g^{\mu\nu} - \frac{p^\mu n^\nu + p^\nu n^\mu}{p\cdot n} + \frac{p^\mu p^\nu}{(p\cdot n)^2}\left(n^2 + \xi^{-1}p^2\right)\right] \ . \tag{5.34}$$

Note that this propagator does not vanish as p^{-2} at large momentum, because of the term proportional to ξ^{-1}, With this gauge fixing, the variation of the gauge fixing function under an infinitesimal gauge transformation is given by

$$\delta G^a = g\, f^{abc}\, \theta_b(x)\, n^\mu A_{c\,\mu}(x) - n^\mu \partial_\mu\, \theta_a(x) \ . \tag{5.35}$$

and the matrix $\mathcal{M}$ reads

$$\mathcal{M}_{ab} = g\, f^{abc}\, n^\mu A_{c\,\mu}(x) - \delta_{ab}\, n^\mu \partial_\mu \ , \tag{5.36}$$

Therefore, the Fadeev-Popov term in the effective Lagrangian is

$$\mathcal{L}_{\mathrm{FPG}} = \overline{\chi}_a\left(-\delta_{ab}\, n^\mu \partial_\mu + g\, f^{abc}\, n^\mu A_{c\,\mu}(x)\right)\chi_b \ , \tag{5.37}$$

which leads to the following expressions for the ghost propagator and its coupling to the gauge boson:

$$G^0_F(p) = \ \cdots\!\!\!\longrightarrow\!\!\!\cdots \ = -\frac{\delta_{ab}}{p\cdot n + i0^+}$$

$$= i\, g\, f^{abc}\, n_\mu \ . \tag{5.38}$$

A significant simplification of these Feynman rules occurs in the limit $\xi \to \infty$ (that one may call the *strict axial gauge*, since the condition $n^\mu A^a_\mu = 0$ holds exactly in this limit). In this limit, the gauge boson propagator becomes

$$G^{0\ \mu\nu}_{F\ ab}(p) = \frac{-i\,\delta_{ab}}{p^2 + i0^+}\left[g^{\mu\nu} - \frac{p^\mu n^\nu + p^\nu n^\mu}{p\cdot n} + \frac{p^\mu p^\nu\, n^2}{(p\cdot n)^2}\right] \ , \tag{5.39}$$

and satisfies

$$n_\mu\, G^{0\ \mu\nu}_{F\ ab}(p) = n_\nu\, G^{0\ \mu\nu}_{F\ ab}(p) = 0 \ . \tag{5.40}$$

Therefore, the gauge boson propagator gives zero when contracted into the ghost-gauge boson vertex, which effectively decouples the ghosts from the gauge bosons. Thus, the limit $\xi \to \infty$ of the axial gauge is ghost-free (but its propagator is arguably much more complicated than the Feynman gauge propagator).

5.5 Ghosts and unitarity

5.5.1 Example

In Abelian gauge theories, we were able to show that cutting rules provide a perturbative realization of the optical theorem, by using the Ward identities obeyed by amplitudes when all the external charged particles are on-shell. These identities were sufficient to conclude that the unphysical polarizations carried by the internal photon lines of a graph cancel when these lines are cut. But in non-Abelian gauge theories, this reasoning faces two difficulties:

i. There are no Ward identities similar to those of QED, that could be used to prove unitarity.

ii. Higher order graphs in general have ghost loops, whose interpretation is at the moment unclear when such loops are cut.

As we shall see, these two issues are in fact related: the cut ghost lines precisely cancel the unphysical polarizations of the cut gluons. Let us first work out an explicit example that illustrates this assertion: the tree level annihilation of a quark and an antiquark into two gluons in QCD. The corresponding diagrams are the following:
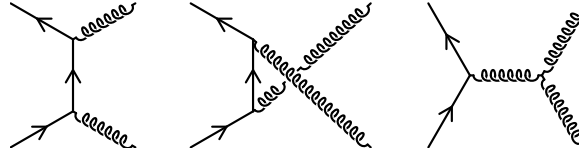
We denote p et q the momenta of the incoming quark and antiquark, respectively, and $k_{1,2}$ the momenta of the outgoing gluons (with Lorentz indices μ, ν and colors a, b, respectively).

The contribution of the first two graphs is very similar to that of the analogous graphs in QED for the emission of two photons, except for the extra color matrices at the quark-gluon vertices:

$$i\,M^{\mu\nu}_{ab}\big|_{1+2}(\mathbf{p},\mathbf{q}|\mathbf{k}_1,\mathbf{k}_2) \;=\; (i\,g)^2\,\bar{v}(\mathbf{q})\left\{\gamma^\mu t^a \frac{i}{\slashed{k}_1-\slashed{q}-m}\gamma^\nu t^b\right.$$
$$\left.+\gamma^\nu t^b \frac{i}{\slashed{p}-\slashed{k}_1-m}\gamma^\mu t^a\right\}u(\mathbf{p})\,. \quad (5.41)$$

By contracting this amplitude with the photon momentum $k_{1\mu}$, we get:

$$k_{1\mu}\;i\,M^{\mu\nu}_{ab}\big|_{1+2}(\mathbf{p},\mathbf{q}|\mathbf{k}_1,\mathbf{k}_2) \;=\; (i\,g)^2\,\bar{v}(\mathbf{q})\Big\{\slashed{k}_1 t^a \frac{i}{\slashed{k}_1-\slashed{q}-m}\gamma^\nu t^b$$
$$+\gamma^\nu t^b \frac{i}{\slashed{p}-\slashed{k}_1-m}\slashed{k}_1 t^a\Big\}u(\mathbf{p})\,. \quad (5.42)$$

In the numerator of the first term, we may write

$$\slashed{k}_1 = (\slashed{k}_1-\slashed{q}-m)+(\slashed{q}+m)\,, \qquad\qquad (5.43)$$

126

and use the Dirac equation $\bar{v}(\mathbf{q})(\not{q} + m) = 0$. Likewise, we may simplify the second term by using

$$
\begin{aligned}
\not{k}_1 &= (\not{p} - m) - (\not{p} - \not{k}_1 - m) , \\
(\not{p} - m)u(\mathbf{p}) &= 0 ,
\end{aligned} \tag{5.44}
$$

which leads to

$$
k_{1\mu} \ i M^{\mu\nu}_{ab}|_{1+2} (\mathbf{p}, \mathbf{q}|k_1, k_2) = i (i g)^2 \bar{v}(\mathbf{q}) \gamma^\nu [t^a, t^b] u(\mathbf{p}) . \tag{5.45}
$$

This is non-zero, because of the non-commutativity of the Lie generators in a non-Abelian gauge theory. However, by using $[t^a, t^b] = i f^{abc} t^c$, this result may be related to the third graph, that contains a 3-gluon vertex. If we use the Feynman gauge for the internal gluon propagator, its contribution can be written as

$$
\begin{aligned}
i M^{\mu\nu}_{ab}|_3 (\mathbf{p}, \mathbf{q}|k_1, k_2) = {}& i g \bar{v}(\mathbf{q}) \gamma_\rho t^c u(\mathbf{p}) \frac{-i}{k_3^2} \\
& \times g f^{abc} [g^{\mu\nu}(k_2 - k_1)^\rho + g^{\nu\rho}(k_3 - k_2)^\mu + g^{\rho\mu}(k_1 - k_3)^\nu] ,
\end{aligned} \tag{5.46}
$$

where we denote $k_3 \equiv -k_1 - k_2$. Contracting this amplitude with $k_{1\mu}$ gives

$$
\begin{aligned}
k_{1\mu} \ i M^{\mu\nu}_{ab}|_3 (\mathbf{p}, \mathbf{q}|k_1, k_2) = {}& i g \bar{v}(\mathbf{q}) \gamma_\rho t^c u(\mathbf{p}) \frac{-i}{k_3^2} \\
& \times g f^{abc} [g^{\nu\rho} k_2^2 - k_2^\nu k_2^\rho - g^{\nu\rho} k_3^2 + k_3^\nu k_3^\rho] .
\end{aligned} \tag{5.47}
$$

In this equation, the term in $k_3^\nu k_3^\rho$ vanishes once contracted with γ_ρ, since we can write

$$
\bar{v}(\mathbf{q}) \gamma_\rho t^c u(\mathbf{p}) k_3^\rho = -\bar{v}(\mathbf{q})[(\not{p} - m) + (\not{q} + m)] t^c u(\mathbf{p}) = 0 . \tag{5.48}
$$

However, this is not sufficient for (5.47) to fully cancel (5.45).

Since gluons are charged, we put them on-shell in trying to check Ward identities. Setting $k_2^2 = 0$ indeed kills another term in eq. (5.47). The term in $k_2^\nu k_2^\rho$ would be canceled if in addition we contract the amplitudes with a *transverse* polarization vector $\epsilon_{1,2\nu}(\mathbf{k}_2)$, since $k_2^\nu \epsilon_{1,2\nu}(\mathbf{k}_2) = 0$. We indeed have:

$$
k_{1\mu} \epsilon_{1,2\nu}(\mathbf{k}_2) \left[i M^{\mu\nu}_{ab}|_{1+2} (\mathbf{p}, \mathbf{q}|k_1, k_2) + i M^{\mu\nu}_{ab}|_3 (\mathbf{p}, \mathbf{q}|k_1, k_2) \right]_{k_2^2 = 0} = 0 . \tag{5.49}
$$

The same cancellation happens if we contract the amplitudes simultaneously with $k_{1\mu}$ et $k_{2\nu}$:

$$
k_{1\mu} k_{2\mu} \left[i M^{\mu\nu}_{ab}|_{1+2} (\mathbf{p}, \mathbf{q}|k_1, k_2) + i M^{\mu\nu}_{ab}|_3 (\mathbf{p}, \mathbf{q}|k_1, k_2) \right] = 0 , \tag{5.50}
$$

even if the momentum k_2 is not on-shell. Thus, we may have for this process a Ward identity similar to the QED one, but only if certain extra conditions are satisfied by the second gluon. These restrictions weaken the resulting identity, and it is not sufficient to eliminate the longitudinal gluon polarizations when we try to recover the amplitude from the imaginary part of the $q\bar{q} \to q\bar{q}$ forward amplitude at one loop. In particular, some unphysical polarizations will not cancel in the following cut:
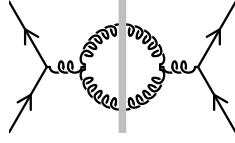
Except for a graph with a quark loop that does not play any role in the present discussion (since it does not give any 2-gluon final state when cut), the complete list of graphs contributing to the $q\bar{q} \to q\bar{q}$ forward amplitude at one loop is shown in the 5.3. The contribution of

Figure 5.3: *One-loop diagrams contributing to $q\bar{q} \to q\bar{q}$.*
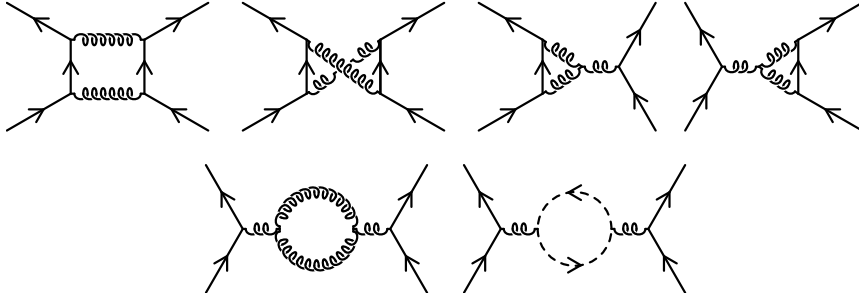
the first 5 graphs (i.e. those with gluon internal lines) to the optical theorem can be calculated easily by noting that it can be expressed in terms of the amplitude we have just calculated:

$$i\, M^{\mu\nu}_{ab}(\mathbf{p}, \mathbf{q}|\mathbf{k}_1, \mathbf{k}_2) \equiv i\, M^{\mu\nu}_{ab}\big|_{1+2}(\mathbf{p}, \mathbf{q}|\mathbf{k}_1, \mathbf{k}_2) + i\, M^{\mu\nu}_{ab}\big|_3(\mathbf{p}, \mathbf{q}|\mathbf{k}_1, \mathbf{k}_2)\,, \qquad (5.51)$$

as follows[5]

$$\frac{1}{2} \int \frac{d^4 k_1}{(2\pi)^4} \int \frac{d^4 k_2}{(2\pi)^4} (2\pi)^4 \delta^{(4)}(p + q - k_1 - k_2)$$
$$\times 2\pi(-g_{\mu\rho})\theta(k_1^0)\delta(k_1^2 - m^2)\, 2\pi(-g_{\nu\sigma})\theta(k_2^0)\delta(k_2^2 - m^2)$$
$$\times i\, M^{\mu\nu}_{ab}(\mathbf{p}, \mathbf{q}|\mathbf{k}_1, \mathbf{k}_2)\,(i\, M^{\rho\sigma}_{ab}(\mathbf{p}, \mathbf{q}|\mathbf{k}_1, \mathbf{k}_2))^*\,. \qquad (5.52)$$

For a successful interpretation of this formula as a physical contribution in the optical theorem, only physical polarizations should survive after we have replaced the tensors $-g_{\mu\rho}$ and $-g_{\nu\sigma}$ by using (see eq. (1.277))

$$g^{\mu\nu} = \epsilon^{\mu}_+(\mathbf{k})\epsilon^{\nu}_-(\mathbf{k})^* + \epsilon^{\mu}_-(\mathbf{k})\epsilon^{\nu}_+(\mathbf{k})^* - \sum_{\lambda=1,2} \epsilon^{\mu}_\lambda(\mathbf{k})\epsilon^{\nu}_\lambda(\mathbf{k})^*\,, \qquad (5.53)$$

where $\epsilon^{\mu}_\pm(\mathbf{k})$ are unphysical polarizations (with $\epsilon^{\mu}_+(\mathbf{k})$ proportional to k^{μ}). After this substitution, several terms are not problematic:

- The terms that contain only the polarizations $\epsilon^{\mu}_{1,2}$ since they are fully physical.

- The terms containing $\epsilon^{\mu}_{1,2}\epsilon^{\nu}_+$ or $\epsilon^{\mu}_+\epsilon^{\nu}_+$ vanish by virtue of eqs. (5.49) and (5.50).

---

[5]The factor $1/2$ is a symmetry factor due to the presence of two identical gluons in the final state.

Thus, we need only study the following term

$$\frac{1}{2}\left[\left(i\,M_{ab}^{\mu\nu}\epsilon_{-\mu}\epsilon_{+\nu}\right)\left(i\,M_{ab}^{\rho\sigma}\epsilon_{+\rho}\epsilon_{-\sigma}\right)^*\right.$$
$$\left.+\left(i\,M_{ab}^{\mu\nu}\epsilon_{+\mu}\epsilon_{-\nu}\right)\left(i\,M_{ab}^{\rho\sigma}\epsilon_{-\rho}\epsilon_{+\sigma}\right)^*\right]\,, \tag{5.54}$$

integrated over the on-shell momenta k_1 and k_2. Using $\epsilon_+^\mu(\mathbf{k})=k^\mu/\sqrt{2}|\mathbf{k}|$ and eqs. (5.45) and (5.47), we obtain

$$\epsilon_{+\mu}(\mathbf{k}_1)\,i\,M_{ab}^{\mu\nu}=-\frac{g^2}{\sqrt{2}|\mathbf{k}_1|}\,\frac{1}{k_3^2}\,\bar{v}(\mathbf{q})\,\not{k}_2\,k_2^\gamma\,f^{abc}\,t^c\,u(\mathbf{p})\,. \tag{5.55}$$

Likewise with the other gluon, we have

$$\epsilon_{+\nu}(\mathbf{k}_2)\,i\,M_{ab}^{\mu\nu}=\frac{g^2}{\sqrt{2}|\mathbf{k}_2|}\,\frac{1}{k_3^2}\,\bar{v}(\mathbf{q})\,\not{k}_1\,k_1^\gamma\,f^{abc}\,t^c\,u(\mathbf{p})\,. \tag{5.56}$$

Using then $\epsilon_-^\mu(\mathbf{k})=(k_0,-\mathbf{k})/\sqrt{2}|\mathbf{k}|$, we get

$$\epsilon_{-\nu}(\mathbf{k}_2)\,\epsilon_{+\mu}(\mathbf{k}_1)\,i\,M_{ab}^{\mu\nu}=-g^2\,\frac{|\mathbf{k}_2|}{|\mathbf{k}_1|}\,\frac{1}{k_3^2}\,\bar{v}(\mathbf{q})\,\not{k}_2\,f^{abc}\,t^c\,u(\mathbf{p})\,,$$
$$\epsilon_{+\nu}(\mathbf{k}_2)\,\epsilon_{-\mu}(\mathbf{k}_1)\,i\,M_{ab}^{\mu\nu}=+g^2\,\frac{|\mathbf{k}_1|}{|\mathbf{k}_2|}\,\frac{1}{k_3^2}\,\bar{v}(\mathbf{q})\,\not{k}_1\,f^{abc}\,t^c\,u(\mathbf{p})\,. \tag{5.57}$$

Furthermore, notice that

$$\bar{v}(\mathbf{q})(\not{k}_1+\not{k}_2)u(\mathbf{p})=\bar{v}(\mathbf{q})(\not{q}+m+\not{p}-m)u(\mathbf{p})=0\,. \tag{5.58}$$

Combining these equations, the non-physical contribution to the optical theorem of the diagrams with a gluon loop, (5.54), can be written as follows:

$$g^4\,\frac{1}{(k_3^2)^2}\,\left[\bar{v}(\mathbf{q})\,\not{k}_1\,f^{abc}\,t^c\,u(\mathbf{p})\right]\,\left[\bar{v}(\mathbf{q})\,\not{k}_1\,f^{abd}\,t^d\,u(\mathbf{p})\right]\,. \tag{5.59}$$

If this was all there is, as the naive Feynman rules we tried to guess at the beginning of this chapter would suggest, then we would have to conclude that Yang-Mills theories are inconsistent because they violate unitarity. Fortunately, there is one more graph in figure 5.3, with a ghost loop. Let us first evaluate the annihilation amplitude of the quark-antiquark pair into a ghost-antighost pair:

$$i\,M_{q\bar{q}\to X\bar{X}}=i\,g\,\bar{v}(\mathbf{q})\,\gamma_\rho\,t^c\,u(\mathbf{p})\,\frac{i}{k_3^2}\,(g\,f^{abc}\,k_1^\rho)\,. \tag{5.60}$$

Squaring this amplitude, and including the $-$ sign[6] associated to a ghost loop[7], the contribution of the last graph of fig. 5.3 to the optical theorem becomes

$$-g^4\,\frac{1}{(k_3^2)^2}\,\left[\bar{v}(\mathbf{q})\,\not{k}_1\,f^{abc}\,t^c\,u(\mathbf{p})\right]\,\left[\bar{v}(\mathbf{q})\,\not{k}_1\,f^{abd}\,t^d\,u(\mathbf{p})\right]\,, \tag{5.61}$$

that exactly cancels the unphysical gluon contribution of eq. (5.59). In other words, the optical theorem is satisfied with only physical modes in the final state sum, thanks to a crucial cancellation that involves ghosts.

---

[6]There is no $1/2$ symmetry factors for a ghost-antighost final state, because they are not identical.

[7]We see here how essential it is that ghosts are anti-commuting fields – otherwise, their contribution would not have the proper sign to cancel the unphysical gluon polarizations in the optical theorem.

5.5.2 Becchi-Rouet-Stora-Tyutin symmetry

The cancellation that occurred in the previous example is in fact general: for every gluon loop, there is a graph of identical topology where this loop is replaced by a ghost loop, that cancels the contribution from the unphysical gluon polarizations in the optical theorem. However, it is difficult to turn the calculation of the previous subsection into a general proof. It turns out that this cancellation originates from a residual symmetry of the gauge fixed Lagrangian: although the gauge fixing term explicitly breaks the gauge symmetry, the effective Lagrangian that appears in eq. (5.21) has a remnant of the original gauge symmetry, known as the *Becchi-Rouet-Stora-Tyutin symmetry* (BRST).

Under an infinitesimal gauge transformation parameterized by $\theta_a(x)$, the gauge field and fermion field vary by

$$
\begin{aligned}
\delta A_\mu^a(x) &= -\left(D_\mu^{\text{adj}}\right)_{ab}\theta_b(x) \\
\delta\psi(x) &= -i\,g\theta_a(x)\,t_r^a\,\psi(x)\,,
\end{aligned}
\tag{5.62}
$$

where r is the representation in which the fermions live. A BRST transformation is similar to the above transformation, but with the substitution $\theta_a(x) \to -\vartheta\,\chi_a(x)$, where ϑ is a Grassmann constant[8],

$$
\begin{aligned}
\delta_{\text{BRST}}A_\mu^a(x) &= \left(D_\mu^{\text{adj}}\right)_{ab}\left[\vartheta\,\chi_b(x)\right] \\
\delta_{\text{BRST}}\psi(x) &= i\,g\left[\vartheta\,\chi_a(x)\right]t_r^a\,\psi(x)\,.
\end{aligned}
\tag{5.63}
$$

Since the BRST transformation is structurally identical to a local gauge transformation, any gauge invariant combination of gauge fields and fermions is also BRST-invariant. This is therefore the case of the Yang-Mills Lagrangian and the Dirac Lagrangian with a minimal coupling of the fermions to the gauge fields. It is customary to introduce a generator $\mathbf{Q}_{\text{BRST}}$ for this transformation, by denoting $\delta_{\text{BRST}} = \vartheta\,\mathbf{Q}_{\text{BRST}}$. Thus

$$
\mathbf{Q}_{\text{BRST}}A_\mu^a(x) = \left(D_\mu^{\text{adj}}\right)_{ab}\chi_b(x)\quad,\qquad \mathbf{Q}_{\text{BRST}}\psi(x) = i\,g\,\chi_a(x)\,t_r^a\,\psi(x)\,.
\tag{5.64}
$$

Eqs. (5.63) do not tell how ghost and antighost fields transform under BRST. For reasons that will become clear later, we shall impose that the BRST transformation is nilpotent, i.e. that $\mathbf{Q}_{\text{BRST}}^2 = 0$ when applied to any of the fields of the theory. This requirement constrains the BRST transformation of the ghosts. Indeed, a double BRST transformation applied to fermions reads

$$
\begin{aligned}
\mathbf{Q}_{\text{BRST}}^2\psi(x) &= i\,g\left\{\left(\mathbf{Q}_{\text{BRST}}\chi_a(x)\right)t_r^a\,\psi(x) - \chi_a(x)\,t_r^a\,\mathbf{Q}_{\text{BRST}}\psi(x)\right\} \\
&= i\,g\left(\mathbf{Q}_{\text{BRST}}\chi_a(x)\right)t_r^a\,\psi(x) + g^2\,\chi_a(x)\chi_b(x)\,t_r^a t_r^b\,\psi(x)\,.
\end{aligned}
\tag{5.65}
$$

(The BRST generator is an anti-commuting object, which leads to a minus sign in the second term of the first line when we push it through the Grassmann field χ_a.) Since χ_a and χ_b anti-commute, we can replace $t_r^a t_r^b$ by $\frac{1}{2}[t_r^a, t_r^b] = \frac{i}{2}f^{abc}t_r^c$. We see that eq. (5.65) will identically vanish provided that

$$
\mathbf{Q}_{\text{BRST}}\chi_a(x) = -\frac{1}{2}\,g\,f^{abc}\,\chi_b(x)\,\chi_c(x)\,.
\tag{5.66}
$$

---

[8]This Grassmann constant makes $\vartheta\,\chi_a(x)$ a commuting object like θ_a.

Then, we can calculate the action of a double BRST transformation on the gauge field,

$$
\begin{aligned}
\mathbf{Q}^2_{\text{BRST}} A^a_\mu &= \left(D^{\text{adj}}_\mu\right)_{ab} \left(\mathbf{Q}_{\text{BRST}} \chi_b\right) - g\, f^{abc} \left(\mathbf{Q}_{\text{BRST}} A^c_\mu\right) \chi_b \\
&= \left(D^{\text{adj}}_\mu\right)_{ab} \left[-\tfrac{g}{2} f^{bcd} \chi_c(x)\chi_d(x)\right] - g\, f^{abc} \left[\partial_\mu \chi_c - g f^{cde} A^e_\mu \chi_d\right] \chi_b
\end{aligned}
$$
$$(5.67)$$

The terms linear in the gauge field cancel by using the anti-commuting nature of the χ's and the Jacobi identity satisfied by the structure constants:

$$
\begin{aligned}
&-\tfrac{1}{2} g^2 f^{abe} f^{bcd} A^e_\mu \chi_c(x)\chi_d(x) + g^2 f^{abc} f^{cde} A^e_\mu \chi_d \chi_b \\
&= \tfrac{1}{2} g^2 \big[\underbrace{-f^{ace} f^{cbd} + f^{abc} f^{cde} - f^{adc} f^{cbe}}_{0}\big] A^e_\mu \chi_b \chi_d \;.
\end{aligned}
$$
$$(5.68)$$

The terms with the derivative ∂_μ read

$$
\begin{aligned}
&-\tfrac{1}{2} g f^{acd} \partial_\mu\left(\chi_c \chi_d\right) - g f^{abc} \left(\partial_\mu \chi_c\right)\chi_b \\
&= \tfrac{1}{2}\, g\, f^{abc}\, \big[\partial_\mu(\chi_c\chi_b) \underbrace{-(\partial_\mu\chi_c)\chi_b + (\partial_\mu\chi_b)\chi_c}_{\substack{-(\partial_\mu\chi_c)\chi_b + \chi_c(\partial_\mu\chi_b) \\ = -\partial_\mu(\chi_c\chi_b)}}\big] = 0 \;.
\end{aligned}
$$
$$(5.69)$$

The double transformation of the ghost field also vanishes

$$
\mathbf{Q}^2_{\text{BRST}} \chi_a = \tfrac{g^2}{4} \big(\underbrace{f^{abc} f^{bde} + f^{acb} f^{bde}}_{0}\big) \chi_c \chi_d \chi_e
$$
$$(5.70)$$

Therefore, the prescription (5.66) for the BRST transformation of a ghost field leads to

$$
\mathbf{Q}^2_{\text{BRST}} \psi = 0 \quad , \qquad \mathbf{Q}^2_{\text{BRST}} A^a_\mu = 0 \quad , \qquad \mathbf{Q}^2_{\text{BRST}} \chi_a = 0 \;.
$$
$$(5.71)$$

We need now to specify the BRST transformation of the antighost field. Note that in the path integral that gives the Fadeev-Popov determinant, the ghost and antighost fields are treated as independent; therefore the BRST transformation of the antighost does not have to be related to that of the ghost. Let us denote:

$$
\mathbf{Q}_{\text{BRST}} \overline{\chi}_a(x) \equiv B_a(x) \;,
$$
$$(5.72)$$

where $B_a(x)$ is a commuting field. For $\mathbf{Q}_{\text{BRST}}$ to be nilpotent, we must have in addition:

$$
\mathbf{Q}_{\text{BRST}} B_a(x) = 0 \;.
$$
$$(5.73)$$

(And of course $\mathbf{Q}^2_{\text{BRST}} B_a(x) = 0$.)

Consider now a local function Ξ of all the fields (including B_a), and add its BRST variation to the Yang-Mills and Dirac Lagrangians:

$$
\mathcal{L} \equiv \underbrace{\mathcal{L}_{\text{YM}} + \mathcal{L}_{\text{D}}}_{\text{BRST-invariant}} + \mathbf{Q}_{\text{BRST}} \Xi \;.
$$
$$(5.74)$$

Since $\mathbf{Q}_{\text{BRST}}$ is nilpotent, this Lagrangian is BRST-invariant. Let us choose

$$
\Xi \equiv \overline{\chi}_a(x) \left[\frac{1}{2\xi} B^a(x) + G^a(A(x))\right] \;,
$$
$$(5.75)$$

where ξ is a parameter and $G^a(A)$ is the gauge fixing function. We can write[9]

$$
\begin{aligned}
\mathbf{Q}_{\text{BRST}}\Xi &= (\mathbf{Q}_{\text{BRST}}\overline{\chi}_a)\left[\frac{1}{2\xi}B^a + G^a\right] - \overline{\chi}_a\left[\frac{1}{2\xi}(\mathbf{Q}_{\text{BRST}}B^a) + \frac{\partial G^a}{\partial A^b_\mu}(\mathbf{Q}_{\text{BRST}}A^b_\mu)\right] \\
&= \frac{1}{2\xi}B^a B^a + B^a G^a + \underbrace{\overline{\chi}_a\frac{\partial G^a}{\partial A^b_\mu}(-D^{\text{adj}}_\mu)_{bc}\chi_c}_{\mathcal{L}_{\text{FPG}}} .
\end{aligned} \tag{5.76}
$$

Note that the last term is nothing but the Fadeev-Popov part of the Lagrangian we have derived earlier in this chapter. Moreover, the field B^a enters only quadratically in this Lagrangian. Therefore, the path integral on B^a can be performed trivially[10],

$$
\int \left[DB^a(x)\right] e^{i\int d^4x \left(\frac{1}{2\xi}B^a B^a + B^a G^a\right)} = e^{-i\frac{\xi}{2}\int d^4x\, G^a G^a} . \tag{5.77}
$$

Therefore, after integrating out the auxiliary field B^a, the resulting theory has exactly the same effective Lagrangian as the one resulting from the Fadeev-Popov procedure:

$$
\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{YM}} + \mathcal{L}_{\text{D}} - \frac{\xi}{2}G^a G^a + \overline{\chi}_a\frac{\partial G^a}{\partial A^b_\mu}(-D^{\text{adj}}_\mu)_{bc}\chi_c . \tag{5.78}
$$

The formal construction we have followed in this section proves that $\mathcal{L}_{\text{eff}}$ is BRST invariant, but in a somewhat obfuscated manner after the auxiliary field B^a has been integrated out. The BRST invariance of eq. (5.78) is realized if we define the BRST variation of the antighost field as follows

$$
\mathbf{Q}_{\text{BRST}}\overline{\chi}_a = -\xi\, G^a , \tag{5.79}
$$

which is reminiscent of the relationship between B^a and G^a when we do the Gaussian integration on B^a.

5.5.3 BRST current and charge

The Lagrangian (5.74), with the choice (5.75) for the function Ξ, possesses the following symmetries:

- Global gauge invariance (because all the color indices are contracted).

- BRST invariance.

- Ghost number conservation, if we assign a ghost number $+1$ to χ's and -1 to $\overline{\chi}$'s.

The BRST invariance implies the existence of a conserved current:

$$
J^\mu_{\text{BRST}} \equiv \sum_{\Phi\in\{A_\mu,\psi,\chi,\overline{\chi},B\}} \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)}\left(\mathbf{Q}_{\text{BRST}}\Phi\right) . \tag{5.80}
$$

---

[9]Note that a minus sign arises when moving $\mathbf{Q}_{\text{BRST}}$ through the anti-commuting field $\overline{\chi}_b$.

[10]Note that this is equivalent to evaluating the argument of the exponential at the stationary point $B^a = -\xi\, G^a$, since the stationary phase approximation is exact for Gaussian integrals.

From the 0-th component of this current, we may obtain the BRST charge

$$Q_{\text{BRST}} \equiv \int d^3\boldsymbol{x} \, J^0_{\text{BRST}}(x^0, \boldsymbol{x}) . \tag{5.81}$$

In fact, this charge generates the BRST transformation in the following sense:

$$i \left[Q_{\text{BRST}}, \Phi \right]_{\pm} = \boldsymbol{Q}_{\text{BRST}} \, \Phi \qquad (\Phi \in \{A_\mu, \psi, \chi, \overline{\chi}, B\}) , \tag{5.82}$$

where $[\cdot, \cdot]_{\pm}$ is a commutator if Φ is a commuting field and an anti-commutator if Φ is anti-commuting. If we consider free fields (i.e. we set $g = 0$), and we Fourier decompose all the fields that appear in the (anti)-commutation relations (5.82),

$$
\begin{aligned}
A^\mu_a(x) &= \sum_{\lambda=1,2,+,-} \int \frac{d^3\boldsymbol{p}}{(2\pi)^3 2|\boldsymbol{p}|} \left\{ \epsilon^\mu_\lambda(\boldsymbol{p}) \, a^\dagger_{a\lambda\boldsymbol{p}} \, e^{+ip\cdot x} + \epsilon^{\mu*}_\lambda(\boldsymbol{p}) \, a_{a\lambda\boldsymbol{p}} \, e^{-ip\cdot x} \right\} \\
\psi(x) &\equiv \sum_{s=\pm} \int \frac{d^3\boldsymbol{p}}{(2\pi)^3 2E_{\boldsymbol{p}}} \left\{ d^\dagger_{s\boldsymbol{p}} \, v_s(\boldsymbol{p}) e^{+ip\cdot x} + b_{s\boldsymbol{p}} \, u_s(\boldsymbol{p}) e^{+ip\cdot x} \right\} \\
\chi_a(x) &\equiv \int \frac{d^3\boldsymbol{p}}{(2\pi)^3 2|\boldsymbol{p}|} \left\{ \alpha^\dagger_{a\boldsymbol{p}} \, e^{+ip\cdot x} + \alpha_{s\boldsymbol{p}} \, e^{+ip\cdot x} \right\} \\
\overline{\chi}_a(x) &\equiv \int \frac{d^3\boldsymbol{p}}{(2\pi)^3 2|\boldsymbol{p}|} \left\{ \beta^\dagger_{a\boldsymbol{p}} \, e^{+ip\cdot x} + \beta_{s\boldsymbol{p}} \, e^{+ip\cdot x} \right\} ,
\end{aligned}
\tag{5.83}
$$

we obtain

$$
\begin{aligned}
\left[Q_{\text{BRST}}, a^\dagger_{a\lambda\boldsymbol{p}} \right] &\propto \delta_{\lambda+} \, \alpha^\dagger_{a\boldsymbol{p}} , \\
\left\{ Q_{\text{BRST}}, \alpha_{a\boldsymbol{p}} \right\} &= 0 , \\
\left\{ Q_{\text{BRST}}, \beta_{a\boldsymbol{p}} \right\} &\propto a^\dagger_{a-\boldsymbol{p}} , \\
\left[Q_{\text{BRST}}, b^\dagger_{s\boldsymbol{p}} \right] &= \left[Q_{\text{BRST}}, d^\dagger_{s\boldsymbol{p}} \right] = 0 .
\end{aligned}
\tag{5.84}
$$

5.5.4   BRST cohomology, Physical states and Unitarity

The fact that the BRST charge is nilpotent, $Q^2_{\text{BRST}} = 0$, has profound implications on the states of the system. The *kernel* of Q_{BRST} is the set of states annihilated by Q_{BRST},

$$\text{Ker} \left(Q_{\text{BRST}} \right) \equiv \left\{ \psi \; \middle| \; Q_{\text{BRST}} |\psi\rangle = 0 \right\} . \tag{5.85}$$

The set of states that can be obtained by the action of Q_{BRST} on another state is called the *image* of Q_{BRST},

$$\text{Im} \left(Q_{\text{BRST}} \right) \equiv \left\{ Q_{\text{BRST}} |\psi\rangle \right\} . \tag{5.86}$$

Because Q_{BRST} is nilpotent, the image is a subset of the kernel,

$$\text{Im} \left(Q_{\text{BRST}} \right) \subset \text{Ker} \left(Q_{\text{BRST}} \right) . \tag{5.87}$$

Note that states in the image cannot be physical states, because they have a null norm:

$$\langle \psi | \psi \rangle = \langle \phi | \underbrace{Q_{\text{BRST}} Q_{\text{BRST}}}_{0} | \phi \rangle = 0 . \tag{5.88}$$

133

Consider now the following equivalence relationship between states in the kernel: two states are considered equivalent if their difference is in the image,

$$|\psi\rangle \sim |\psi'\rangle \qquad \text{if} \qquad |\psi\rangle - |\psi'\rangle \in \text{Im}\left(\mathcal{Q}_{\text{BRST}}\right) . \tag{5.89}$$

The *cohomology* of $\mathcal{Q}_{\text{BRST}}$ is the set of classes of equivalent states,

$$\mathsf{H}\left(\mathcal{Q}_{\text{BRST}}\right) \equiv \text{Ker}\left(\mathcal{Q}_{\text{BRST}}\right) / \text{Im}\left(\mathcal{Q}_{\text{BRST}}\right) . \tag{5.90}$$

It turns out that the physical states are the non-zero norm[11] elements of the cohomology. Indeed, using eqs. (5.84), it is easy to prove that if $|\psi\rangle$ is a state in the cohomology, then

$$
\begin{aligned}
a^{\dagger}_{a\{1,2\}\mathbf{p}}|\psi\rangle &\in& \mathsf{H}\left(\mathcal{Q}_{\text{BRST}}\right) \\
b^{\dagger}_{s\mathbf{p}}|\psi\rangle &\in& \mathsf{H}\left(\mathcal{Q}_{\text{BRST}}\right) \\
d^{\dagger}_{s\mathbf{p}}|\psi\rangle &\in& \mathsf{H}\left(\mathcal{Q}_{\text{BRST}}\right) ,
\end{aligned}
\tag{5.91}
$$

while

$$
\begin{aligned}
a^{\dagger}_{a\pm\mathbf{p}}|\psi\rangle &\notin& \mathsf{H}\left(\mathcal{Q}_{\text{BRST}}\right) \\
\alpha^{\dagger}_{\mathbf{p}}|\psi\rangle &\notin& \mathsf{H}\left(\mathcal{Q}_{\text{BRST}}\right) \\
\beta^{\dagger}_{\mathbf{p}}|\psi\rangle &\notin& \mathsf{H}\left(\mathcal{Q}_{\text{BRST}}\right) .
\end{aligned}
\tag{5.92}
$$

In other words, adding to the state a physical particle (gluon with a physical polarization, or quark or antiquark) gives another state in the cohomology, while adding to the state a nonphysical quantum (gluon with a non-physical polarization, ghost or antighost) takes the state out of the cohomology.

Furthermore, since the effective Lagrangian is BRST invariant, it corresponds to a Hamiltonian $\mathcal{H}$ that commutes with $\mathcal{Q}_{\text{BRST}}$. Therefore, a state in the kernel (i.e. for $\mathcal{Q}_{\text{BRST}}|\psi\rangle = 0$) stays in the kernel under the time evolution generated by this Hamiltonian. Furthermore, the time evolution preserves the norm, and therefore states in the cohomology stay in the cohomology at all times. Therefore, starting from a physical states, the time evolution cannot produce unphysical objects in the final state. This explains why unphysical modes cannot appear in the final states sum in the optical theorem, despite the fact that the internal lines of Feynman graphs may propagate all sorts of unphysical objects.

---

[11]This restriction is necessary, because one of the equivalency classes in $\mathsf{H}\left(\mathcal{Q}_{\text{BRST}}\right)$ is $\text{Im}\left(\mathcal{Q}_{\text{BRST}}\right)$ itself, that we know has only zero-norm states.

Chapter 6

Renormalization of gauge theories

6.1 Ultraviolet power counting

Before studying in more detail the renormalizability of gauge theories, one may assess the plausibility of this renormalizability by calculating the superficial degree of ultraviolet divergence of graphs in such a theory. Furthermore, this will guide us regarding which classes of graphs may contain divergences. For simplicity, we will consider here a pure Yang-Mills theory, without matter fields (keeping fermions would force us to distinguish the fermion propagators from the gluon and ghost propagators in the counting, because they have different behaviors at large momentum). Note that the gluon propagator decreases as $(\text{momentum})^{-2}$ in the ultraviolet, both in covariant and axial gauge. This is also the behavior of the ghost propagator[1]. Moreover, the 3-gluon vertex and the gluon-ghost-antighost vertex have the same scaling with momentum. Therefore, we need not distinguish in the ultraviolet power counting the ghosts and the gluons. Thus, let us consider a generic connected graph $\mathcal{G}$ with the following list of propagators and vertices:

- $n_{\scriptscriptstyle E}$ external lines (gluons or ghosts),

- $n_{\scriptscriptstyle I}$ internal lines (gluons or ghosts),

- n_3 trivalent vertices (3-gluon or gluon-ghost-antighost),

- n_4 four-gluon vertices,

- $n_{\scriptscriptstyle L}$ loops.

These quantities are related by the following identities:

$$n_{\scriptscriptstyle E} + 2n_{\scriptscriptstyle I} = 3n_3 + 4n_4 \tag{6.1}$$

$$n_{\scriptscriptstyle L} = n_{\scriptscriptstyle I} - (n_3 + n_4) + 1 \; . \tag{6.2}$$

The first equation states that each vertex must have all its "handles" attached to the endpoint of a propagator, and the second equation counts the number of internal momenta that are not

---

[1] In axial gauge, the ghost propagator behaves differently, but the ghosts decouple from the gluons.

fixed by energy momentum conservation. In terms of these parameters, the ultraviolet degree of divergence of this graph (in four space-time dimensions) is

$$\omega(\mathcal{G}) = 4n_{\mathrm{L}} - 2n_{\mathrm{I}} + n_3 . \tag{6.3}$$

Note that each trivalent vertex contains one power of momentum and therefore contribute $+1$ to this counting. Adding eq. (6.1) and four times eq. (6.2), we obtain

$$\omega(\mathcal{G}) = 4 - n_{\mathrm{E}} , \tag{6.4}$$

that does not depend on any of the internal details of the graph. Moreover, the only functions that have intrinsic ultraviolet divergences are the 2-point, 3-point and 4-point functions, which suggests that Yang-Mills theories may indeed be renormalizable. However, a Yang-Mills theory is not simply the addition of gluon and ghost kinetic terms, 3- and 4-gluon vertices, and a ghost-antighost-gluon vertex: all these terms of the Lagrangian are tightly constrained by gauge symmetry. For instance (but this is not the only constraint), all the vertices depend on a unique coupling constant g. Therefore, in order to establish the renormalizability of Yang-Mills theories, one needs to prove that the structure of the divergences in the above listed functions is such that they can be absorbed into a redefinition of the classical Lagrangian that does not upset these tight constraints (up to a renormalization of the fields).

6.2 Symmetries of the quantum effective action

6.2.1 Linearly realized symmetries

After fixing the gauge with the Fadeev-Popov procedure, we have obtained the following effective Lagrangian:

$$\mathcal{L}_{\mathrm{eff}} = \mathcal{L}_{\mathrm{YM}} + \mathcal{L}_{\mathrm{D}} - \frac{\xi}{2} G^a G^a + \overline{\chi}_a \frac{\partial G^a}{\partial A^b_\mu} \left(- D^{\mathrm{adj}}_\mu \right)_{bc} \chi_c . \tag{6.5}$$

Although the local gauge invariance of the Yang-Mills Lagrangian is now broken (this was precisely the goal of the gauge fixing procedure), this effective Lagrangian has a number of symmetries. One of them is the BRST symmetry, that we have exhibited in the previous chapter. In addition, $\mathcal{L}_{\mathrm{eff}}$ has the following symmetries:

- *Ghost number conservation* : the effective Lagrangian is invariant under global phase transformations of the ghost and antighost,

$$\chi \to e^{i\alpha}\chi \quad , \qquad \overline{\chi} \to e^{-i\alpha}\overline{\chi} . \tag{6.6}$$

  Therefore, if we assign a *ghost number* $+1$ to the field χ and -1 to the field $\overline{\chi}$, this quantity is conserved by the Feynman rules of the gauge fixed theory.

- *Global gauge invariance* : since all color indices are contracted in the effective Lagrangian, it is invariant under gauge transformations that do not depend on spacetime.

- *Lorentz invariance* is of course also present in the effective Lagrangian.

For these three symmetries, the infinitesimal variation of the fields is linear in the fields (which is not the case of the BRST symmetry). These linearly realized symmetries of the classical action are inherited directly by the quantum effective action.

In order to prove this assertion, let us consider a generic infinitesimal linear transformation of the fields

$$\phi_n(x) \quad \rightarrow \quad \phi_n(x) + \varepsilon \, F_n[x; \phi] \,, \tag{6.7}$$

where $\phi_1, \phi_2, \cdots$ denote the various fields of the theory (gauge fields, ghosts, ...) and $F_n[x; \phi]$ is a local function of the fields (for now, we do not assume that it is linear in the fields). We assume that both the classical action and the functional measure are invariant under this symmetry. Consider now the generating functional $Z[j]$,

$$Z[j] \equiv \int \left[D\phi_n(x)\right] \, e^{i\left(S[\phi_n] + \int d^4x \, j_n(x)\phi_n(x)\right)} \,, \tag{6.8}$$

where there is one external source j_n for each field ϕ_n. Since $\phi_n(x)$ is a dummy integration variable in this path integral, we should obtain the same result after performing the change of variable (6.7). Using the fact that this transformation preserves the measure and the classical action, this implies that

$$\begin{aligned}
Z[j] &= \int \left[D\phi_n(x)\right] \, e^{i\left(S[\phi_n] + \int d^4x \, j_n(x)\phi_n(x) + \varepsilon \int d^4x \, j_n(x)F_n[x;\phi]\right)} \\
&\approx Z[j] + i\varepsilon \int \left[D\phi_n(x)\right] \, e^{i\left(S[\phi_n] + \int d^4x \, j_n(x)\phi_n(x)\right)} \int d^4x \, j_n(x) F_n[x; \phi] \,.
\end{aligned} \tag{6.9}$$

Therefore, for any sources j_n, we must have

$$\int d^4x \, j_n(x) \left\langle F_n[x; \phi(x)] \right\rangle_j = 0 \,, \tag{6.10}$$

where $\left\langle \cdots \right\rangle_j$ denotes the quantum average in the presence of an external source j,

$$\left\langle \mathcal{O}[\phi] \right\rangle_j \equiv \frac{1}{Z[j]} \int \left[D\phi_n(x)\right] \, e^{i\left(S[\phi_n] + \int d^4x \, j_n(x)\phi_n(x)\right)} \, \mathcal{O}[\phi] \,. \tag{6.11}$$

(We have normalized it so that $\left\langle 1 \right\rangle_j = 1$.) Recall now that the sources and field can be related implicitly by using the quantum effective action:

$$j_{n;\phi}(x) = -\frac{\delta \Gamma[\phi]}{\delta \phi_n(x)} \,. \tag{6.12}$$

Therefore, the condition (6.10) is equivalent to

$$\int d^4x \, \left\langle F_n[x; \phi(x)] \right\rangle_{j_\phi} \frac{\delta \Gamma[\phi]}{\delta \phi_n(x)} = 0 \,, \tag{6.13}$$

now satisfied for any fields ϕ_n. In other words, the functional $\Gamma[\phi]$ is invariant under the transformation

$$\phi_n(x) \quad \rightarrow \quad \phi_n(x) + \varepsilon \, \left\langle F_n[x; \phi] \right\rangle_{j_\phi} \,. \tag{6.14}$$

It is crucial to note that, because the quantum average in the right hand side is performed with the external field $j_{n;\phi}$ that depends implicitly on the fields ϕ_n, this is a priori not the same transformation as in eq. (6.7).

Let us now consider the special case of a transformation of type (6.7) which is linear in the fields. In this case, we may write

$$F_n[x;\phi] = \int d^4y \, f_{nm}(x,y) \, \phi_m(y) \,. \tag{6.15}$$

(In most practical cases, the transformation will be local and the coefficients proportional to $\delta(x-y)$, but this restriction is not necessary for the following argument.) For such a linear transformation, we have

$$\left\langle F_n[x;\phi]\right\rangle_{j_\phi} = \int d^4y \, f_{nm}(x,y) \left\langle \phi_m(y)\right\rangle_{j_\phi} \,. \tag{6.16}$$

Recalling that j_ϕ is the configuration of the source j such that the quantum average $\left\langle \phi(x)\right\rangle_j$ precisely equals $\phi(x)$, this in fact reads

$$\left\langle F_n[x;\phi]\right\rangle_{j_\phi} = F_n[x;\phi] \,. \tag{6.17}$$

It is this last step that fails when F_n is nonlinear in the fields. From eq. (6.17), we see that the transformations (6.14) and (6.7) are identical. We have thus proven that all linearly realized symmetries of the classical action are also symmetries of the quantum effective action.

6.2.2 BRST symmetry and Slavnov-Taylor identities

Since an infinitesimal BRST variation is not linear in the fields, the BRST symmetry of the classical action is not inherited simply by the quantum effective action. Instead, it leads to a set of identities known as the *Slavnov-Taylor identities*, that may be viewed as the analogue of Ward identities for the BRST invariance. Their derivation follows the method of the section 3.4.2. Since we need to apply a BRST transformation to the Yang-Mills path integral, we should first study how this transformation affects the measure $\left[DA_\mu D\chi D\overline{\chi}\right]$. Under such a transformation, the fields transform into

$$
\begin{aligned}
A_\mu^a &\rightarrow A_\mu^{a\prime} \equiv A_\mu^a + \vartheta \left(D_\mu^{\mathrm{adj}}\right)_{ab}\chi_b = A_\mu^a + \vartheta \left(\partial_\mu\delta_{ab} + gf^{abc}A_\mu^c\right)\chi_b \\
\chi_a &\rightarrow \chi_a' \equiv \chi_a - g\frac{\vartheta}{2} f^{abc}\chi_b\chi_c \\
\overline{\chi}_a &\rightarrow \overline{\chi}_a' \equiv \overline{\chi}_a + \vartheta B^a = \overline{\chi}_a - \xi\vartheta G^a \,,
\end{aligned}
\tag{6.18}
$$

where ϑ is a Grassmann constant. The Jacobian matrix has the following block structure:

$$
\frac{\partial\left(A_\mu^{a\prime}, \chi_a', \overline{\chi}_a'\right)}{\partial\left(A_\nu^b, \chi_b, \overline{\chi}_b\right)} = \delta(x-y)
\begin{pmatrix}
\delta_\mu^\nu(\delta_{ab} - g\vartheta f^{abc}\chi_c) & * & 0 \\
0 & \delta_{ab} + g\vartheta f^{abc}\chi_c & 0 \\
-\xi\vartheta\frac{\partial G^a}{\partial A_\nu^b} & 0 & \delta_{ab}
\end{pmatrix}
\,, \tag{6.19}
$$

where the $*$ denotes a non-zero element that we do not need to calculate because it does not contribute to the determinant. From this structure, we see that the determinant is given by the product of the diagonal elements, and is therefore equal to 1 (recall that $\vartheta^2 = 0$).

In the derivation of the Slavnov-Taylor identities, it is convenient to introduce sources $j_\mu^a, \overline{\eta}_a, \eta_a$ that couple respectively to $A_\mu^a, \chi_a, \overline{\chi}_a$, but also two extra sources that couple directly to $\mathbf{Q}_{\text{BRST}} A_\mu^a$ and $\mathbf{Q}_{\text{BRST}} \chi_a$:

$$
\begin{aligned}
Z[j, \eta, \overline{\eta}; \zeta, \kappa] &\equiv \int [DA_\mu D\chi D\overline{\chi}] \, \exp\left\{ i \int d^4x \left(\mathcal{L}_{\text{eff}} + j_\mu^a A_a^\mu + \overline{\eta}_a \chi_a + \overline{\chi}_a \eta_a \right.\right. \\
&\qquad\qquad\qquad\qquad \left.\left. + \zeta_a^\mu (\mathbf{Q}_{\text{BRST}} A_\mu^a) - \kappa_a (\mathbf{Q}_{\text{BRST}} \chi_a) \right) \right\} \\
&= \int [DA_\mu D\chi D\overline{\chi}] \, \exp\left\{ i \int d^4x \, \mathcal{L}_{\text{tot}} \right\}, \qquad (6.20)
\end{aligned}
$$

where we use the shorthand $\mathcal{L}_{\text{tot}}$ for the sum of terms inside the exponential. Note that the coefficients of the new sources ζ_a^μ and κ_a are BRST invariant since the BRST transformation is nilpotent. Let us now perform a BRST transformation of the integration variables inside the path integral. This is just a change of variables, that does not change the value of the path integral. Using the fact that measure and $\mathcal{L}_{\text{eff}}$ are BRST invariant, we obtain

$$
\begin{aligned}
Z[j, \eta, \overline{\eta}; \zeta, \kappa] &= \int [DA_\mu D\chi D\overline{\chi}] \, \exp\left\{ i \int d^4x \, \mathcal{L}_{\text{tot}} \right\} \\
&\quad \times \left[1 + i \int d^4x \left(j_\mu^a \vartheta (\mathbf{Q}_{\text{BRST}} A_a^\mu) + \overline{\eta}_a \vartheta (\mathbf{Q}_{\text{BRST}} \chi_a) + \vartheta (\mathbf{Q}_{\text{BRST}} \overline{\chi}_a) \eta_a \right) \right] \\
&= Z[j, \eta, \overline{\eta}; \zeta, \kappa] + i\vartheta \int d^4x \left(j_\mu^a(x) \frac{\delta Z}{i\delta \zeta_\mu^a(x)} + \overline{\eta}_a(x) \frac{\delta Z}{i\delta \kappa_a(x)} \right. \\
&\qquad\qquad\qquad\qquad\qquad\qquad \left. - \xi \, G^a \left(\frac{\delta Z}{i\delta j(x)} \right) \eta_a(x) \right). \qquad (6.21)
\end{aligned}
$$

(Note that ϑ anticommutes with $\overline{\eta}_a$.) Therefore, we conclude that

$$
\int d^4x \left(j_\mu^a(x) \frac{\delta Z}{i\delta \zeta_\mu^a(x)} + \overline{\eta}_a(x) \frac{\delta Z}{i\delta \kappa_a(x)} - \xi \, G^a \left(\frac{\delta Z}{i\delta j(x)} \right) \eta_a(x) \right) = 0. \qquad (6.22)
$$

This is one of the forms of the Slavnov-Taylor identities. In this derivation, we see that having introduced sources specifically coupled to the BRST variation of the gauge field A_μ^a and of the ghost χ_a avoided the need for terms with higher order derivatives (indeed, these variations are non-linear in the fields, and would have required more derivatives to be expressed as functional derivatives with respect to sources coupled to elementary fields). By writing $Z = \exp(W)$, we see that the same identity applies to W,

$$
\int d^4x \left(j_\mu^a(x) \frac{\delta W}{i\delta \zeta_\mu^a(x)} + \overline{\eta}_a(x) \frac{\delta W}{i\delta \kappa_a(x)} - \xi \, G^a \left(\frac{\delta W}{i\delta j(x)} \right) \eta_a(x) \right) = 0. \qquad (6.23)
$$

(Here, we have assumed that the gauge fixing function is linear in the gauge field.)

The next step is to convert this into an identity for the quantum effective action Γ that generates the 1PI graphs. In this transformation, we will keep the auxiliary sources ζ_μ^a and κ_a unmodified, as parameters. Thus, Γ and W are related by

$$
-i W[j, \eta, \overline{\eta}; \zeta, \kappa] = \Gamma[A, \chi, \overline{\chi}; \zeta, \kappa] + \int d^4x \left(j_a^\mu(x) A_\mu^a(x) + \overline{\chi}_a(x) \eta^a(x) + \overline{\eta}^a(x) \chi_a(x) \right).
$$

$$(6.24)$$

Fields and sources are related by the following quantum equations of motion:

$$\frac{\delta\Gamma}{\delta A_\mu^a(x)} + j_a^\mu(x) = 0 \;,$$

$$\frac{\delta\Gamma}{\delta\overline{\chi}_a(x)} + \eta^a(x) = 0 \;,$$

$$\frac{\delta\Gamma}{\delta\chi_a(x)} + \overline{\eta}^a(x) = 0 \;, \tag{6.25}$$

and we also have

$$\frac{\delta W}{\delta j_a^\mu(x)} = i\,A_\mu^a(x) \;,$$

$$\frac{\delta W}{\delta\zeta_\mu^a(x)} = i\,\frac{\delta\Gamma}{\delta\zeta_\mu^a(x)} \;,$$

$$\frac{\delta W}{\delta\kappa_a(x)} = i\,\frac{\delta\Gamma}{\delta\kappa^a(x)} \;. \tag{6.26}$$

Therefore, the Slavnov-Taylor identity expressed in terms of the functional Γ reads

$$\int d^4x \left(\frac{\delta\Gamma}{\delta A_\mu^a(x)}\frac{\delta\Gamma}{\delta\zeta_\mu^a(x)} + \frac{\delta\Gamma}{\delta\chi_a(x)}\frac{\delta\Gamma}{\delta\kappa^a(x)} - \xi\,G^a(A)\,\frac{\delta\Gamma}{\delta\overline{\chi}_a(x)} \right) = 0 \;. \tag{6.27}$$

This equation can be simplified a bit as follows. By inserting a derivative $\delta/\delta\overline{\chi}_a(x)$ under the integral in the definition (6.21) of Z, we obtain zero since we now have the integral of a total derivative. Recalling that the Fadeev-Popov term in the effective Lagrangian is

$$\mathcal{L}_{\text{FPG}} = \overline{\chi}_a\,\frac{\partial G^a}{\partial A_\mu^b}\left(-D_\mu^{\text{adj}} \right)_{bc}\chi_c \;, \tag{6.28}$$

we can perform explicitly this derivative to obtain

$$0 = \int \left[DA_\mu D\chi D\overline{\chi} \right] \left[\frac{\partial G^a}{\partial A_\mu^b}\underbrace{\left(-D_\mu^{\text{adj}} \right)_{bc}\chi_c(x)}_{\substack{-Q_{\text{BRST}}A_\mu^b(x)\\ = i\frac{\delta}{\delta\zeta_\mu^b(x)}}} + \eta_a(x) \right] e^{i\int d^4x\,\mathcal{L}_{\text{tot}}} \;. \tag{6.29}$$

This implies the following functional identity

$$\left[\eta_a(x) + i\,\frac{\partial G^a}{\partial A_\mu^b}\,\frac{\delta}{\delta\zeta_\mu^b(x)} \right] Z = 0 \;, \tag{6.30}$$

or equivalent identities for W or Γ:

$$\eta_a(x) + i\,\frac{\partial G^a}{\partial A_\mu^b}\,\frac{\delta W}{\delta\zeta_\mu^b(x)} = 0 \quad , \qquad \frac{\delta\Gamma}{\delta\overline{\chi}_a(x)} + \frac{\partial G^a}{\partial A_\mu^b}\,\frac{\delta\Gamma}{\delta\zeta_\mu^b(x)} = 0 \;. \tag{6.31}$$

Furthermore, define a slightly modified effective action:

$$\overline{\Gamma} \equiv \Gamma + \frac{\xi}{2}\int d^4x\,G^a(A)G^a(A) \;. \tag{6.32}$$

140

Now the Slavnov-Taylor identity takes the following more compact form:

$$\int d^4x \left(\frac{\delta\overline{\Gamma}}{\delta A_\mu^a(x)} \frac{\delta\overline{\Gamma}}{\delta\zeta_\mu^a(x)} + \frac{\delta\overline{\Gamma}}{\delta\chi_a(x)} \frac{\delta\overline{\Gamma}}{\delta\kappa^a(x)} \right) = 0 \,, \tag{6.33}$$

from which any explicit reference to the gauge fixing function $G^a(A)$ has disappeared, as well as the coupling constant g.

Eq. (6.33) applies to the full quantum effective action, that encapsulates the results from all-order perturbation theory. In the next section, we will show that this identity (combined with the other symmetries of the effective action) completely constrains the structure of its local terms of dimension less than or equal to four, forcing them to be identical to those in the classical action (up to a rescaling of the fields and of the coupling constant).

6.3 Renormalizability

6.3.1 Constraints on the counterterms

By taking the $\hbar \to 0$ limit in eq. (6.33), one immediately concludes that it is also satisfied by the classical action, $\overline{S}$, supplemented with ghosts as well as the sources $\zeta_{\mu a}$ and κ_a:

$$\overline{S}[A, \chi, \overline{\chi}; \zeta, \kappa] = \int d^4x \left[-\frac{1}{4} \overline{F}_a^{\mu\nu} \overline{F}_{\mu\nu}^a + \left(\zeta_a^\mu + \partial^\mu \overline{\chi}_a\right) \left(\overline{D}_\mu^{adj}\right)_{ab} \chi_b \right.$$
$$\left. + \frac{g}{2} f^{abc} \kappa_a \chi_b \chi_c \right] . \tag{6.34}$$

By introducing the following compact notation,

$$(\mathcal{A}, \mathcal{B}) \equiv \int d^4x \left(\frac{\delta\mathcal{A}}{\delta A_\mu^a(x)} \frac{\delta\mathcal{B}}{\delta\zeta_\mu^a(x)} + \frac{\delta\mathcal{A}}{\delta\chi_a(x)} \frac{\delta\mathcal{B}}{\delta\kappa^a(x)} \right) , \tag{6.35}$$

we therefore have

$$\left(\overline{S}, \overline{S}\right) = 0 \,,$$
$$\left(\overline{\Gamma}, \overline{\Gamma}\right) = 0 \,. \tag{6.36}$$

The first equation may be viewed as a constraint on the terms that can appear in the classical action, while the second equation constrains which divergences may appear in higher orders.

Let us now write the effective action as a loop expansion,

$$\overline{\Gamma} \equiv \overline{S} + \sum_{l=1}^\infty \overline{\Gamma}_l \,, \tag{6.37}$$

where $\overline{S}$ is given in eq. (6.34), and the subsequent terms $\overline{\Gamma}_l$ are of order l in $\hbar$. The Slavnov-Taylor identity at order L thus reads

$$\sum_{p+q=L} \left(\overline{\Gamma}_p, \overline{\Gamma}_q\right) = 0 \,. \tag{6.38}$$

The renormalization procedure amounts to correcting order by order with counterterms the classical action $\overline{\mathcal{S}}$,

$$\overline{\mathcal{S}} \quad \to \quad \overline{\mathcal{S}}_{(L)} \, , \tag{6.39}$$

such that $\overline{\mathcal{S}}_{(L)}$ contains counterterms up to order L, and gives finite $\overline{\Gamma}_l$'s for $l \leq L$ (but in general not beyond the order L).

The first step is to prove that it is possible to find counterterms such that the equation $(\overline{\mathcal{S}}, \overline{\mathcal{S}}) = 0$ is preserved at every order. Let us assume that we have achieved this up to the order $L - 1$. All $\overline{\Gamma}_l$ for $l \leq L - 1$ are now finite, while $\overline{\Gamma}_L$ still contains a divergent part, that we denote $\overline{\Gamma}_{L,\text{div}}$. We can rewrite the Slavnor-Taylor identity at order L as follows,

$$\left(\overline{\mathcal{S}}, \overline{\Gamma}_L\right) + \left(\overline{\Gamma}_L, \overline{\mathcal{S}}\right) = -\sum_{l=1}^{L-1} \left(\overline{\Gamma}_l, \overline{\Gamma}_{L-l}\right) . \tag{6.40}$$

Only the left hand side is divergent, and we therefore have

$$\left(\overline{\mathcal{S}}, \overline{\Gamma}_{L,\text{div}}\right) + \left(\overline{\Gamma}_{L,\text{div}}, \overline{\mathcal{S}}\right) = 0 \, , \tag{6.41}$$

which constrains the structure of the divergences at order L. A natural candidate for the counterterm at order L is to simply add $-\overline{\Gamma}_{L,\text{div}}$ to the classical action,

$$\overline{\mathcal{S}} \quad \to \quad \overline{\mathcal{S}} - \overline{\Gamma}_{L,\text{div}} \, , \tag{6.42}$$

since this automatically cancels the superficial divergence of $\overline{\Gamma}_L$ without affecting anything in the lower orders. However, this modified classical action does not obey exactly the Slavnov-Taylor equation, since

$$\left(\overline{\mathcal{S}} - \overline{\Gamma}_{L,\text{div}}, \overline{\mathcal{S}} - \overline{\Gamma}_{L,\text{div}}\right) = \underbrace{\left(\overline{\mathcal{S}}, \overline{\mathcal{S}}\right)}_{\text{0 from lower orders}} - \underbrace{\left[\left(\overline{\mathcal{S}}, \overline{\Gamma}_{L,\text{div}}\right) + \left(\overline{\Gamma}_{L,\text{div}}, \overline{\mathcal{S}}\right)\right]}_{\text{0 from eq. (6.41)}} + \underbrace{\left(\overline{\Gamma}_{L,\text{div}}, \overline{\Gamma}_{L,\text{div}}\right)}_{\neq 0} . \tag{6.43}$$

Note that the non-zero term in the right hand side is of order strictly greater than L. It is possible to make it vanish by adding to the shift of eq. (6.42) some terms of higher order than L, that do not change anything for any order $\leq$ L. The conclusion of this inductive argument is that one can shift the classical action at each order in such a way that the divergences in $\overline{\Gamma}$ are canceled, while always preserving $(\overline{\mathcal{S}}, \overline{\mathcal{S}}) = 0$.

6.3.2 Allowed terms in the classical action

The second step in the discussion of the renormalization of Yang-Mills theory is to determine the terms that are allowed in the classical action. This action must satisfy the constraint $(\overline{\mathcal{S}}, \overline{\mathcal{S}}) = 0$, as well as Lorentz invariance, global gauge symmetry and ghost number conservation. In addition, from the power counting of the section 6.1 and Weinberg's theorem, we know that all the ultraviolet divergences in Yang-Mills theory will occur in *local* operators of dimension 4 at most.

In order to discuss the form of the allowed terms, let us first list the mass dimension and ghost number of the various fields that enter in $\overline{\mathcal{S}}$:

field	A_a^μ	χ_a	$\overline{\chi}_a$	ζ_a^μ	κ_a
mass dimension	1	1	1	2	2
ghost number	0	+1	-1	-1	-2

All the allowed terms in $\overline{S}$ must obey the following conditions:

- mass dimension 4 or less,

- ghost number 0,

- Lorentz invariance,

- global gauge invariance.

In addition, eq. (6.31) implies that the $\overline{\chi}$ and ζ dependences come in the form of a dependence on the combination

$$\zeta_\mu - \overline{\chi} \frac{\partial G}{\partial A_\mu} = \zeta_\mu + \partial_\mu \overline{\chi} \,, \tag{6.44}$$

where in the right hand side we have assumed the gauge condition $G(A) = \partial_\mu A^\mu$ and anticipated an integration by parts. Finally, the Slavnov-Taylor identity $(\overline{S}, \overline{S}) = 0$ must be satisfied.

Since the sources ζ_a^μ and κ_a have mass dimension 2, at most two of them may appear. However, terms with two such sources cannot contain any other field since the mass dimension 4 is already reached, and they cannot have ghost number zero. Therefore, $\overline{S}$ can only contain terms that have degree 0 or 1 in ζ_a^μ and κ_a.

The source ζ_a^μ must be combined with another combination of fields that have one Lorentz index, one color index, mass dimension at most 2, and ghost number +1. The only operators that fulfill these conditions are

$$f^{abc} \zeta_a^\mu A_\mu^b \chi_c \qquad \text{and} \qquad \zeta_a^\mu \partial_\mu \chi_a \,. \tag{6.45}$$

Once the dependence on ζ_a^μ is fixed, the dependence on the antighosts will be completely known from eq. (6.44). Likewise, κ_a must be combined with an object that has one color index, mass dimension at most 2 and ghost number +2. The only possibility is

$$f^{abc} \kappa_a \chi_b \chi_c \,. \tag{6.46}$$

.

From the information gathered so far, the classical action must have the following general form:

$$\overline{S}[A, \chi, \overline{\chi}; \zeta, \kappa] = \overline{\Sigma}[A] + \int d^4x \left[g\alpha \, f^{abc} \left(\zeta_a^\mu + \partial^\mu \overline{\chi}_a \right) A_\mu^b \chi_c \right.$$
$$\left. + \beta \left(\zeta_a^\mu + \partial^\mu \overline{\chi}_a \right) \partial_\mu \chi_a + \tfrac{\gamma}{2} \, f^{abc} \kappa_a \chi_b \chi_c \right], \tag{6.47}$$

where α, β, γ are three arbitrary constants. The term $\overline{\Sigma}$ cannot depend on the sources ζ_a^μ and κ_a because we have already constructed explicitly all the allowed terms that contain these sources, and cannot depend on $\overline{\chi}$ because the antighost dependence is already encapsulated in the combination $\zeta_a^\mu + \partial^\mu \overline{\chi}_a$. A dependence on χ in $\overline{\Sigma}$ is also forbidden because χ would be the only field in $\overline{\Sigma}$ with a non-zero ghost number. Our next step is to constrain the coefficients

$g\alpha, \beta, \gamma$ and the functional $\overline{\Sigma}[A]$ in order to satisfy the Slavnov-Taylor identity (6.33). The functional derivatives that enter in (6.33) are given by:

$$
\begin{aligned}
\frac{\delta \overline{S}}{\delta A_a^\mu} &= \frac{\delta \overline{\Sigma}}{\delta A_a^\mu} - g\alpha\, f^{abc}\, (\zeta_{\mu b} + \partial_\mu \overline{\chi}_b)\, \chi_c \ , \\
\frac{\delta \overline{S}}{\delta \zeta_{\mu a}} &= g\alpha\, f^{ade}\, A_d^\mu \chi_e + \beta\, \partial^\mu \chi_a \ , \\
\frac{\delta \overline{S}}{\delta \chi_a} &= g\alpha\, f^{abc}\, (\zeta_{\mu b} + \partial_\mu \overline{\chi}_b)\, A_c^\mu + \beta\, (\zeta_{\mu a} + \partial_\mu \overline{\chi}_a)\, \partial^\mu + \gamma\, f^{abc}\, \kappa_b\, \chi_c \ , \\
\frac{\delta \overline{S}}{\delta \kappa_a} &= \frac{\gamma}{2}\, f^{ade}\, \chi_d\, \chi_e \ .
\end{aligned}
\tag{6.48}
$$

Thus, the Slavnov-Taylor identity reads

$$
\begin{aligned}
0 = \int d^4 x \Big[& \frac{\delta \overline{\Sigma}}{\delta A_a^\mu} \big[g\alpha\, f^{ade}\, A_d^\mu \chi_e + \beta\, \partial^\mu \chi_a \big] \\
& + (\zeta_{\mu b} + \partial_\mu \overline{\chi}_b) \big[- g\alpha\, f^{abc} \chi_c \big(g\alpha\, f^{ade}\, A_d^\mu \chi_e + \beta\, \partial^\mu \chi_a \big) \\
& \qquad\qquad + \big(g\alpha\, f^{abc}\, A_c^\mu + \beta\, \delta_{ab} \partial^\mu \big) \tfrac{\gamma}{2}\, f^{ade}\, \chi_d\, \chi_e \big] \\
& + \frac{\gamma^2}{2}\, f^{abc} f^{ade}\, \kappa_b\, \chi_c\, \chi_d\, \chi_e \Big] \ .
\end{aligned}
\tag{6.49}
$$

Using the Jacobi identity satisfied by the structure constants, one may first check that the last term, in $\kappa \chi \chi \chi$, is identically zero, and therefore does not provide any constraint. Consider now the terms in $\zeta A \chi \chi$:

$$
\begin{aligned}
g\alpha\, \zeta_{\mu b} \Big(&- g\alpha\, f^{abc} f^{ade} A_d^\mu \chi_c \chi_e + \tfrac{\gamma}{2}\, \underbrace{f^{abc} f^{ade}}_{-f^{abd}f^{aec} - f^{abe}f^{acd}}\, A_c^\mu \chi_d \chi_e \Big) \\
&= g\alpha\, (\gamma - g\alpha)\, f^{abc} f^{ade} \zeta_{\mu b}\, A_d^\mu \chi_c \chi_e \ .
\end{aligned}
\tag{6.50}
$$

Since this is the only term containing this combination of fields, it cannot be canceled by other terms, and therefore we must have

$$
g\alpha = \gamma \ .
\tag{6.51}
$$

Let us now study the terms in $\zeta \chi (\partial \chi)$,

$$
\beta\, \zeta_{\mu b} \Big(- g\alpha\, f^{abc} \chi_c \partial^\mu \chi_a + \tfrac{\gamma}{2}\, f^{bde} \partial^\mu (\chi_d \chi_e) \Big) = \beta\, (\gamma - g\alpha)\, f^{bac} \zeta_{\mu b}\, (\partial^\mu \chi_a)\, \chi_c \ .
\tag{6.52}
$$

Thus, the cancellation of this term does not bring any additional constraint beyond eq. (6.51). At this point, all the terms containing ζ_a^μ have been canceled (and by extension also the terms with $\partial^\mu \overline{\chi}_a$), and the Slavnov-Taylor identity reduces to

$$
0 = \int d^4 x\, \frac{\delta \overline{\Sigma}}{\delta A_a^\mu} \big[g\alpha\, f^{acb}\, A_c^\mu \chi_b + \beta\, \partial^\mu \chi_a \big] \ .
\tag{6.53}
$$

Let us first rewrite the second factor as follows

$$
\beta\, \big(\partial_\mu \delta_{ab} - ig\alpha\beta^{-1}\, \underbrace{(-if^{cab}) A_c^\mu}_{(A_{\mathrm{adj}}^\mu)_{ab}} \big)\, \chi_b \ ,
\tag{6.54}
$$

144

and note that it has the structure of an adjoint covariant derivative acting on χ_b,

$$\left(\overline{D}_{\text{adj}}^{\mu}\right)_{ab} \equiv \partial_{\mu}\delta_{ab} - ig\alpha\beta^{-1}\left(A_{\text{adj}}^{\mu}\right)_{ab} . \tag{6.55}$$

Thus, eq. (6.53) is equivalent to

$$0 = \int d^4x \, \frac{\delta\overline{\Sigma}}{\delta A_a^{\mu}} \left(\overline{D}_{\text{adj}}^{\mu}\right)_{ab} \chi_b . \tag{6.56}$$

The second factor may be viewed as the variation of the gauge field under an infinitesimal gauge transformation,

$$A_a^{\mu} \to A_a^{\mu} + \vartheta \left(\overline{D}_{\text{adj}}^{\mu}\right)_{ab} \chi_b , \tag{6.57}$$

where we have introduced a constant Grassmann variable ϑ to make the second term a commuting object. Therefore, for the integral to be zero for an arbitrary $\chi_b(x)$, the functional $\overline{\Sigma}[A]$ must be invariant under this transformation. Recalling our discussion of the local gauge invariant operators of mass dimension four or less, we conclude that the only possible form for $\overline{\Sigma}$ is

$$\overline{\Sigma}[A] = -\frac{\delta}{4} \int d^4x \, \overline{F}_a^{\mu\nu}\overline{F}_{\mu\nu}^a , \tag{6.58}$$

where $\overline{F}^{\mu\nu}$ is the field strength constructed with the covariant derivative $\overline{D}^{\mu}$ and δ another constant. Given all the above constraints, we must have

$$\overline{S}[A, \chi, \overline{\chi}; \zeta, \kappa] = \int d^4x \left[-\frac{\delta}{4} \, \overline{F}_a^{\mu\nu}\overline{F}_{\mu\nu}^a + \beta \left(\zeta_a^{\mu} + \partial^{\mu}\overline{\chi}_a \right) \left(\overline{D}_{\mu}^{\text{adj}}\right)_{ab}\chi_b \right.$$
$$\left. + \frac{g\alpha}{2} \, f^{abc}\kappa_a\chi_b\chi_c \right] . \tag{6.59}$$

Up to rescalings of the various fields and of the coupling constant g, this is structurally identical to the bare classical action of eq. (6.34). Note that this equation implies that the field renormalization factors for the gauge field A_a^{μ} and for the source κ_a are equal,

$$Z_A = Z_{\kappa} . \tag{6.60}$$

6.4 Background field method

6.4.1 Rescaled fields

In this section, we describe the calculation of the one-loop quantum corrections to the coupling constant by a method based on the quantum effective action combined with the so-called background field method.

The first step of this method is to rescale the gauge field by the inverse of the coupling constant:

$$g\,A^{\mu} \quad \to \quad A^{\mu} . \tag{6.61}$$

By doing this, the various objects that appear in the Yang-Mills action are transformed as follows:

$$
\begin{aligned}
F^{\mu\nu} \quad &\rightarrow \quad \frac{1}{g}\left(\partial^{\mu}A^{\nu} - \partial^{\nu} - i\left[A^{\mu}, A^{\nu}\right]\right) \\
D^{\mu} \quad &\rightarrow \quad \partial^{\mu} - i A^{\mu} .
\end{aligned}
\tag{6.62}
$$

In other words, up to a rescaling in the case of the field strength $F^{\mu\nu}$, these objects are transformed into their counterparts for a coupling equal to unity. In the rest of this section, the notation $A^{\mu}, D^{\mu}, F^{\mu\nu}$ will refer to the rescaled quantities. In terms of the rescaled fields, the Yang-Mills action simply reads

$$
\mathcal{S}_{\text{YM}} = -\frac{1}{4\,g^{2}}\int d^{4}x \underbrace{F_{a}^{\mu\nu}F_{\mu\nu}^{a}}_{\text{no } g} ,
\tag{6.63}
$$

where all the dependence on the coupling constant appears now in the prefactor g^{-2}. This action has a local non-Abelian gauge invariance analogous to the original one, but with $g = 1$:

$$
A_{\mu} \quad \rightarrow \quad A_{\mu}^{\Omega} \equiv \Omega^{\dagger}A_{\mu}\Omega + i\,\Omega^{\dagger}\partial_{\mu}\Omega .
\tag{6.64}
$$

6.4.2 Background field gauge

The *background field method* consists in choosing a background field $\mathcal{A}_{\mu}^{a}(x)$, and in writing the gauge field $A_{\mu}^{a}(x)$ as a deviation around this background

$$
A_{\mu}^{a} \equiv \mathcal{A}_{\mu}^{a} + a_{\mu}^{a} .
\tag{6.65}
$$

In this decomposition, the background field $\mathcal{A}_{\mu}$ is not a dynamical field: it will just act as a parameter that we shall not quantize, and the path integration is thus only on the deviation a_{μ}^{a} (one may thus view this as a shift of the integration variable). In terms of $\mathcal{A}_{\mu}^{a}$ and a_{μ}^{a}, the field strength that enters in the Yang-Mills action can be written as

$$
F^{\mu\nu} = \mathcal{F}^{\mu\nu} + \left(\partial^{\mu}a^{\nu} - ig[\mathcal{A}^{\mu}, a^{\nu}]\right) - \left(\partial^{\nu}a^{\mu} - ig[\mathcal{A}^{\nu}, a^{\mu}]\right) - i\left[a^{\mu}, a^{\nu}\right] ,
\tag{6.66}
$$

where $\mathcal{F}^{\mu\nu}$ is the field strength constructed with the background field. With explicit color indices, this reads

$$
F_{a}^{\mu\nu} = \mathcal{F}_{a}^{\mu\nu} + \left(\mathcal{D}_{\text{adj}}^{\mu}\right)_{ab}a_{b}^{\nu} - \left(\mathcal{D}_{\text{adj}}^{\nu}\right)_{ab}a_{b}^{\mu} + f^{abc}\,a_{b}^{\mu}\,a_{c}^{\nu} ,
\tag{6.67}
$$

where $\mathcal{D}_{\text{adj}}^{\mu} = \partial^{\mu} - i\left[\mathcal{A}^{\mu}, \cdot\right]$ is the adjoint covariant derivative associated to the background field $\mathcal{A}^{\mu}$. If we view the background field as a constant, the original gauge transformation on A^{μ} corresponds to the following transformation on a^{μ},

$$
a^{\mu} \quad \rightarrow \quad \Omega^{\dagger}a^{\mu}\Omega + \Omega^{\dagger}\mathcal{A}^{\mu}\Omega - \mathcal{A}^{\mu} + i\,\Omega^{\dagger}\partial^{\mu}\Omega .
\tag{6.68}
$$

If we parameterize $\Omega = \exp(i\theta_{a}t^{a})$ and expand to first order in θ_{a}, an infinitesimal gauge transformation of a_{a}^{μ} reads

$$
a_{a}^{\mu} \quad \rightarrow \quad a_{a}^{\mu} - \left(\mathcal{D}_{\text{adj}}^{\mu}\right)_{ab}\theta_{b} + f^{abc}\theta_{b}a_{c}^{\mu} .
\tag{6.69}
$$

This invariance leads to the same pathologies as in the original theory, and we must fix the gauge in order to have a well defined path integral. The *background field gauge* corresponds to the following condition on a_μ^a,

$$G^a(A) \equiv \left(\mathcal{D}_{\text{adj}}^\mu\right)_{ab} a_\mu^b \, . \tag{6.70}$$

Let us recall that a gauge fixing function $G^a(A)$ leads to the following terms in the effective Lagrangian:

$$\mathcal{L}_{\text{GF}} = -\frac{\xi}{2\,g^2}\, G^a(A) G^a(A) \qquad \text{(gauge fixing term)}$$

$$\mathcal{L}_{\text{FPG}} = -\overline{\chi}_a\, \frac{\partial G^a}{\partial A_\mu^b} \left(D_\mu^{\text{adj}}\right)_{bc} \chi_c \qquad \text{(Fadeev-Popov ghosts)} \, . \tag{6.71}$$

With the choice of eq. (6.70), the Fadeev-Popov term becomes

$$\mathcal{L}_{\text{FPG}} = -\overline{\chi}_a \left(\mathcal{D}_{\text{adj}}^\mu\right)_{ab} \left(D_\mu^{\text{adj}}\right)_{bc} \chi_c = \left(\mathcal{D}_{\text{adj}}^\mu \overline{\chi}\right)_a \left(D_\mu^{\text{adj}} \chi\right)_a \, , \tag{6.72}$$

where in the second equality we have anticipated an integration by parts and used the notation $(D_\mu^{\text{adj}} \chi)_a \equiv (D_\mu^{\text{adj}})_{ab} \chi_b$ (and a similar notation for $(\mathcal{D}_{\text{adj}}^\mu \overline{\chi})_a$).

6.4.3 Residual symmetry of the gauge fixed Lagrangian

The effective Lagrangian $\mathcal{L}_{\text{YM}} + \mathcal{L}_{\text{GF}} + \mathcal{L}_{\text{FPG}}$ possesses a residual gauge symmetry that corresponds to gauge transforming in the same way the background field $\mathcal{A}^\mu$ and the total field A^μ,

$$\begin{aligned}
A_\mu &\rightarrow \Omega^\dagger A_\mu \Omega + i\,\Omega^\dagger \partial_\mu \Omega \, , \\
\mathcal{A}_\mu &\rightarrow \Omega^\dagger \mathcal{A}_\mu \Omega + i\,\Omega^\dagger \partial_\mu \Omega \, .
\end{aligned} \tag{6.73}$$

Indeed, under this joint transformation we have

$$\begin{aligned}
a_\mu &\rightarrow \Omega^\dagger a_\mu \Omega \, , \\
D_\mu &\rightarrow \Omega^\dagger D_\mu \Omega \, , \\
\mathcal{D}_\mu &\rightarrow \Omega^\dagger \mathcal{D}_\mu \Omega \, , \\
\chi &\rightarrow \Omega^\dagger \chi \, , \\
\overline{\chi} &\rightarrow \overline{\chi} \Omega \, , \\
G(A) &\rightarrow \Omega^\dagger G(A) \Omega \, .
\end{aligned} \tag{6.74}$$

From this, we conclude that the gauge fixing Lagrangian $\mathcal{L}_{\text{GF}}$ and the Fadeev-Popov Lagrangian $\mathcal{L}_{\text{FPG}}$ are both invariant in this transformation, as well as the Yang-Mills Lagrangian. Since the path integration measure over $a_\mu, \chi, \overline{\chi}$ is also invariant under this transformation, the result of the path integral must be invariant under local gauge transformations of the background field $\mathcal{A}^\mu$.

6.4.4 One-loop running coupling

Let us now turn to the calculation of the quantum effective action at one-loop. For this, we use the results of the section 2.6.3, where we have shown that these one-loop corrections are obtained by expanding the classical action to quadratic order in deviations with respect to a background field, and by performing the resulting Gaussian path integration with respect to the deviations (which gives a functional determinant).

The first step is to expand the three terms of the gauge fixed Lagrangian to second order in the deviation a_μ. In this calculation, we choose the gauge fixing parameter $\xi = 1$. The quadratic terms in the combined Yang-Mills and gauge fixing terms read

$$
\begin{aligned}
\mathcal{L}_{\text{YM}} + \mathcal{L}_{\text{GF}} &= -\frac{1}{2g^2}\left\{ \tfrac{1}{2}\left((\mathcal{D}_{\text{adj}}^\mu a^\nu)_a - (\mathcal{D}_{\text{adj}}^\nu a^\mu)_a\right)^2 + f^{abc}\mathcal{F}_a^{\mu\nu}a_\mu^b a_\nu^c + \left((\mathcal{D}_{\text{adj}}^\mu a_\mu)^a\right)^2 \right\} \\
&= -\frac{1}{2g^2}\left\{ a_\mu^a\left[-(\mathcal{D}_{\text{adj}})_{ac}^2 g^{\mu\nu} - 2f^{abc}\mathcal{F}_b^{\mu\nu}\right]a_\nu^c \right\} \\
&= -\frac{1}{2g^2}\left\{ a_\mu^a\left[-(\mathcal{D}_{\text{adj}})_{ac}^2 g^{\mu\nu} + (\mathcal{F}_{\text{adj}}^{\rho\sigma})_{ac}(M_{\rho\sigma}^{(1)})^{\mu\nu}\right]a_\nu^c \right\}, \quad (6.75)
\end{aligned}
$$

where we have introduced $(M_{\rho\sigma}^{(1)})^{\mu\nu} \equiv i(\delta_\rho{}^\mu\delta_\sigma{}^\nu - \delta_\rho{}^\nu\delta_\sigma{}^\mu)$ the generators of the Lorentz transformations for 4-vectors (the Lorentz transformation corresponding to the transformation parameters $\omega^{\rho\sigma}$ reads $\Lambda^{\mu\nu} = \exp(\tfrac{i}{2}\omega^{\rho\sigma}(M_{\rho\sigma}^{(1)})^{\mu\nu})$). For the ghost term, the quadratic part is

$$
\mathcal{L}_{\text{FPG}} = \overline{\chi}_a\left[-(\mathcal{D}_{\text{adj}})_{ab}^2\right]\chi_b. \quad (6.76)
$$

Note that the operator that appears between the two ghost fields is the spin-0 analogue of the one that appears in eq. (6.75), since the generators of Lorentz transformations for spin-0 objects are identically zero ($M_{\rho\sigma}^{(0)} \equiv 0$). Although we have not considered fermions so far in this chapter, the Dirac Lagrangian would give a contribution equal to the determinant of $i\slashed{D}$, or equivalently the square root of the determinant of $(i\slashed{D})^2$. Noting that

$$
\begin{aligned}
(i\slashed{D})^2 &= -\mathcal{D}^2 + i\left(\tfrac{i}{2}[\gamma^\mu,\gamma^\nu]\right)\mathcal{D}_\mu\mathcal{D}_\nu \\
&= -\mathcal{D}^2 + (\mathcal{F}^{\rho\sigma})M_{\rho\sigma}^{(1/2)}, \quad (6.77)
\end{aligned}
$$

where the $M_{\rho\sigma}^{(1/2)} \equiv \tfrac{i}{4}[\gamma_\rho,\gamma_\sigma]$ are the generators of Lorentz transformations for spin-1/2 fields. Note that the covariant derivatives and the field strength are in the fundamental representation (assuming fermions that transform according to the fundamental representation, like quarks). Therefore, for each of the fields that appear in the quantum effective action (gauge fields, ghosts, fermions), we get a determinant $\Delta_{r,s}$ of an operator containing $-\mathcal{D}^2$ (in the representation r corresponding to the field under consideration) plus a "spin connection"[2] made of the contraction of the field strength with the Lorentz generators corresponding to the spin s of the field:

$$
\begin{aligned}
\text{gauge fields}: \quad & \Delta_{\text{adj},s=1} \equiv \det\left(-\mathcal{D}_{\text{adj}}^2 + \mathcal{F}_{\text{adj}}^{\rho\sigma}M_{\rho\sigma}^{(1)}\right) \\
\text{ghosts}: \quad & \Delta_{\text{adj},s=0} \equiv \det\left(-\mathcal{D}_{\text{adj}}^2 + \underbrace{\mathcal{F}_{\text{adj}}^{\rho\sigma}M_{\rho\sigma}^{(0)}}_{=0}\right) \\
\text{fermions}: \quad & \Delta_{\text{f},s=1/2} \equiv \det\left(-\mathcal{D}_{\text{f}}^2 + \mathcal{F}_{\text{f}}^{\rho\sigma}M_{\rho\sigma}^{(1/2)}\right). \quad (6.78)
\end{aligned}
$$

---

[2]This terms describes the coupling between the magnetic moment of the particle and the background field.

In terms of these determinants, the 1-loop quantum effective action is given by

$$\Gamma[A, \chi, \psi] = \mathcal{S}_r + \Delta\mathcal{S} + \frac{i}{2} \ln \Delta_{\text{adj},s=1} - \frac{i\, n_f}{2} \ln \Delta_{f,s=1/2} - i \ln \Delta_{\text{adj},s=0} \ , \qquad (6.79)$$

where $\Delta\mathcal{S}$ denotes the 1-loop counterterms, and n_f is the number of fermion flavors. Using the invariance with respect to local gauge transformations of the background field, we must have

$$\ln \Delta_{r,s} = \frac{i}{4} C_{r,s} \int d^4x \, \mathcal{F}_a^{\mu\nu} \mathcal{F}_{\mu\nu}^a + \cdots \ , \qquad (6.80)$$

where the dots represent higher dimensional gauge invariant operators. Being of dimension higher than four, these operators do not contribute to the renormalization of the coupling. The constant $C_{r,s}$ depends on the group representation r and spin s of the field. These coefficients are ultraviolet divergent,

$$C_{r,s} = c_{r,s} \, \ln \frac{\Lambda^2}{\kappa^2} \ , \qquad (6.81)$$

where Λ is an ultraviolet scale and κ the typical scale of inhomogeneities of the background field. After combining them with the counterterms from $\Delta\mathcal{S}$, the ultraviolet scale is replaced by a renormalization scale μ,

$$C_{r,s} \quad \to \quad C_{r,s} = c_{r,s} \, \ln \frac{\mu^2}{\kappa^2} \ . \qquad (6.82)$$

From eq. (6.79), we see that the 1-loop renormalized coupling at the scale μ and the bare coupling must be related by

$$\begin{aligned}
\frac{1}{g_b^2} &= \frac{1}{g_r^2(\mu)} + \frac{1}{2} C_{\text{adj},1} - \frac{n_f}{2} C_{f,1/2} - C_{\text{adj},0} \\
&= \frac{1}{g_r^2(\mu)} + \left(\frac{1}{2} c_{\text{adj},1} - \frac{n_f}{2} c_{f,1/2} - c_{\text{adj},0}\right) \ln \frac{\mu^2}{\kappa^2} \ .
\end{aligned} \qquad (6.83)$$

The explicit calculation of the constants $c_{r,s}$ requires to expand the logarithm of the functional determinants to second order in the background field strength $\mathcal{F}^{\mu\nu}$. Thanks to the organization of eqs. (6.78), this calculation needs to be performed only once, for generic gauge group and Lorentz representations. This leads to

$$c_{r,s} = \frac{1}{(4\pi)^2} \left[\tfrac{1}{3} d(s) - 4C(s)\right] C(r) \ , \qquad (6.84)$$

where $d(s)$ is the number of spin components (respectively $1, 4, 4$ for scalars, fermions, and vector particles), $C(s)$ is the normalization of the trace of two Lorentz generators[3],

$$\text{tr}\left(M_{\rho\sigma}^{(s)} M_{\alpha\beta}^{(s)}\right) = C(s) \left(g_{\rho\alpha} g_{\sigma\beta} - g_{\rho\beta} g_{\sigma\alpha}\right) \ , \qquad (6.85)$$

and $C(r)$ is the normalization of the trace of two generators of the Lie algebra in representation r,

$$\text{tr}\left(t_r^a t_r^b\right) = C(r) \, \delta^{ab} \ . \qquad (6.86)$$

---

[3]For spin-0, $\frac{1}{2}$ and 1, this constant is respectively 0, 1 and 2.

For the fundamental and adjoint representations of $\mathfrak{su}(N)$, we have $C(f) = \frac{1}{2}$ and $C(\text{adj}) = N$. Therefore, the constants involved in the 1-loop running coupling are

$$c_{\text{adj},0} = \frac{N}{3(4\pi)^2} \, , \quad c_{\text{adj},1} = -\frac{20\,N}{3(4\pi)^2} \quad c_{f,1/2} = -\frac{4}{3(4\pi)^2} \, , \tag{6.87}$$

and the coupling evolves according to

$$\frac{1}{g_r^2(\mu)} = \frac{1}{g_b^2} + \frac{1}{(4\pi)^2} \bigg(\underbrace{\tfrac{11}{3}\,N - \tfrac{2}{3}\,n_f}_{>0 \text{ for } n_f \le \frac{11\,N}{2}} \bigg) \ln \frac{\mu^2}{\kappa^2} \, . \tag{6.88}$$

Given two scales μ and μ_0, the renormalized couplings at these scales are related by

$$\frac{1}{g_r^2(\mu)} - \frac{1}{g_r^2(\mu_0)} = \frac{1}{(4\pi)^2} \left(\tfrac{11}{3}\,N - \tfrac{2}{3}\,n_f \right) \ln \frac{\mu^2}{\mu_0^2} \, , \tag{6.89}$$

which may be rewritten as

$$g^2(\mu) = \frac{g^2(\mu_0)}{1 + \frac{g_r^2(\mu_0)}{(4\pi)^2} \left(\tfrac{11}{3}\,N - \tfrac{2}{3}\,n_f \right) \ln \frac{\mu^2}{\mu_0^2}} \, . \tag{6.90}$$

In quantum chromodynamics, where the gauge group is $SU(3)$ (i.e. $N = 3$) and where there are 6 flavors of quarks in the fundamental representation, the coefficient in front of the logarithm is positive, which indicates that the coupling constant decreases as the scale μ increases. The coupling constant in fact goes to zero when $\mu \to \infty$, a property known as *asymptotic freedom*. Thanks to the formula (6.84), it would have been easy to determine the one-loop running of the coupling in the presence of matter fields in arbitrary representations.

Chapter 7

Lattice field theory

7.1 Discretization of space-time

In the previous chapter, we have seen that the running coupling in non-Abelian gauge theories decreases at large energy (provided the number of quark flavors is less than $11N/2$). The counterpart of asymptotic freedom is that the coupling increases towards lower energies, precluding the use of perturbation theory to study phenomena in this regime. Among such properties is that of color confinement, i.e. the fact that colored states cannot exist as asymptotic states. Instead the quarks and gluons arrange themselves into color neutral bound states, that can be mesons (e.g. pions, kaons) made of a quark and an antiquark or baryons (e.g. protons, neutrons) made of three quarks[1]. A legitimate question would be to determine the mass spectrum of the asymptotic states of QCD from its Lagrangian.

Figure 7.1: *Discretization of Euclidean space-time on a hyper-cubic lattice (here shown in three dimensions).*

---
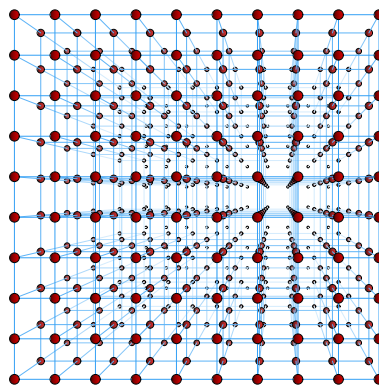
[1]More exotic bound states made of four (tetraquarks) or five (pentaquarks) have also been speculated, but the experimental evidence for these states is so far not fully conclusive. Likewise, there may exist bound states without valence quarks, the *glueballs*.

Since the perturbative expansion is not applicable for this type of problem, one would like to be able to attack it via some non-perturbative approach. By *non-perturbative*, we mean a method by which observables would directly be obtained to all orders in the coupling constant, without any expansion. Such a method, known as *lattice field theory*, consists in discretizing space-time in order to evaluate numerically the path integral. The continuous space-time is replaced by a discrete grid of points, the simplest arrangement being a hyper-cubic lattice such as the one shown in the figure 7.1. The distance between nearest neighbor sites is called the lattice spacing, and usually denoted a. The lattice spacing, being the smallest distance that exists in this setup, therefore provides a natural ultraviolet regularization. Indeed, on a lattice of spacing a, the largest conjugate momentum is of order a^{-a}. Moreover, one usually uses periodic boundary conditions; if the lattice has N spacings in a given direction, then we have $\phi(x + Na) = \phi(x)$ for bosonic fields and $\phi(x + Na) = -\phi(x)$ for fermionic fields.

7.2 Scalar field theory

As an illustration of some of the issues involved in the discretization of a quantum field theory, let us consider a simple scalar field theory with a local interaction in ϕ^4, whose action in continuous space-time is

$$S = \int d^4x \left\{ -\frac{1}{2}\phi(x)(\partial_\mu\partial^\mu + m^2)\phi(x) - \frac{\lambda}{4!}\phi^4(x) \right\} . \tag{7.1}$$

A natural choice is to replace the integral over space-time by a discrete sum over the sites of the lattice, weighted by the volume a^4 of each elementary cell of the lattice,

$$a^4 \sum_{x \in \text{ lattice}} \underset{a \to 0}{\to} \int d^4x . \tag{7.2}$$

Then we replace the continuous function $\phi(x)$ by a discrete set of real numbers that live on the lattice nodes. For simplicity, we keep denoting $\phi(x)$ the value of the field on the lattice site x. The discretization of the mass and interaction terms is trivial, but the discretization of the derivatives that appear in the D'Alembertian operator is not unique. Using only two nearest neighbors, one may define forward or backward finite differences,

$$\nabla_{\text{F}} f(x) \equiv \frac{f(x + a) - f(x)}{a}$$
$$\nabla_{\text{B}} f(x) \equiv \frac{f(x) - f(x - a)}{a} , \tag{7.3}$$

that both go to the continuum derivative in the limit $a \to 0$. However, unlike the continuous derivative, ∇_{F} and ∇_{B} are not anti-adjoint. Instead, assuming periodic boundary conditions, one has

$$\sum_{x \in \text{ lattice}} f(x) \left(\nabla_{\text{F}} g(x) \right) = - \sum_{x \in \text{ lattice}} \left(\nabla_{\text{B}} f(x) \right) g(x) . \tag{7.4}$$

In other words, $\nabla_{\text{F}}^\dagger = -\nabla_{\text{B}}$. From this, we may construct a self-adjoint discrete second derivative as follows:

$$\nabla_{\text{B}} \nabla_{\text{F}} f(x) = \frac{f(x + a) + f(x - a) - 2f(x)}{a^2} \underset{a \to 0}{\to} f''(x) . \tag{7.5}$$

Thus, a self-adjoint discretization of the scalar Lagrangian leads to

$$S_{\text{lattice}} = a^4 \sum_{x \in \text{lattice}} \left\{ -\frac{1}{2}\phi(x)\left(\nabla_{\text{B}\,\mu}\nabla_{\text{F}}^{\mu} + m^2\right)\phi(x) - \frac{\lambda}{4!}\phi^4(x) \right\} . \qquad (7.6)$$

Let us make a few remarks concerning the errors introduced by the discretization. Firstly, the continuous spacetime symmetries (translation and rotation invariance) of the underlying theory are now reduced to the subgroup of the discrete symmetries of a cubic lattice. They are recovered in the $a \to 0$ limit. Another source of discrepancy between the continuum and discrete theories is the dispersion relation that relates the energy and momentum of an on-shell particle. In the continuum theory, this relation is of course

$$E^2 = \mathbf{p}^2 + m^2 , \qquad (7.7)$$

where $-\mathbf{p}^2$ is an eigenvalue of the Laplacian. In order to find its counterpart with the above discretization, we must determine the spectrum of the finite difference operator $\nabla_{\text{B}}\nabla_{\text{F}}$. On a lattice with $\mathcal{N}$ sites and periodic boundary conditions, its eigenfunctions are given by

$$\phi_k(x) \equiv e^{2i\pi\frac{kx}{\mathcal{N}a}} \qquad \text{with } k \in \mathbb{Z} , \quad -\frac{\mathcal{N}}{2} \leq k \leq \frac{\mathcal{N}}{2} . \qquad (7.8)$$

The associated eigenvalue is

$$\lambda_k \equiv \frac{2}{a^2}\left(\cos\frac{2\pi k}{\mathcal{N}} - 1\right) = -\frac{4}{a^2}\sin^2\frac{\pi k}{\mathcal{N}} . \qquad (7.9)$$

Thus, the one dimensional discrete analogue of the continuum $\mathbf{p}^2 + m^2$ is

$$m^2 + \frac{4}{a^2}\sin^2\frac{\pi k}{\mathcal{N}} . \qquad (7.10)$$

As long as $k \ll \mathcal{N}$, this leads us to a linear relationship between the energy and the index k that labels the eigenfunctions, but this correspondence becomes nonlinear for larger values of k. This discrepancy is illustrated in the figure 7.2. This mismatch does not improve by increasing the number of lattice points: only the center of the Brillouin zone has a dispersion relation that agrees with the continuum one. In order to mitigate this problem, one should choose the parameters of the lattice in such a way that the physically relevant scales correspond to values of k for which the distortion of the dispersion curve is not important.

7.3 Gluons and Wilson action

Non-Abelian gauge theories pose an additional difficulty: since the local gauge invariance plays a central role in their properties, any attempt at discretizing gauge fields should preserve this symmetry. It turns out that there exists a discretization of the Yang-Mills action that goes to the continuum action in the limit where $a \to 0$, and has an exact gauge invariance. The main ingredient in this construction is eq. (4.122), that relates the trace of a Wilson loop along a small square,

$$[\square]_{x;\mu\nu} \equiv U_{\nu}^{\dagger}(x)\,U_{\mu}^{\dagger}(x + \hat{\nu})\,U_{\nu}(x + \hat{\mu})\,U_{\mu}(x) , \qquad (7.11)$$

to the squared field strength. These elementary lattice Wilson loops are called *plaquettes*. In the fundamental representation, we have

$$\text{tr}\left([\square]_{x;\mu\nu}\right) = N - \frac{g^2 a^4}{4}\,F_a^{\mu\nu}(x)F_{\mu\nu}^a(x) + \mathcal{O}(a^6) . \qquad (7.12)$$

153

Figure 7.2: *Discrepancy between the continuous (solid curve) and discrete (points) dispersion relations, on a one-dimensional lattice with $\mathcal{N} = 40$.*
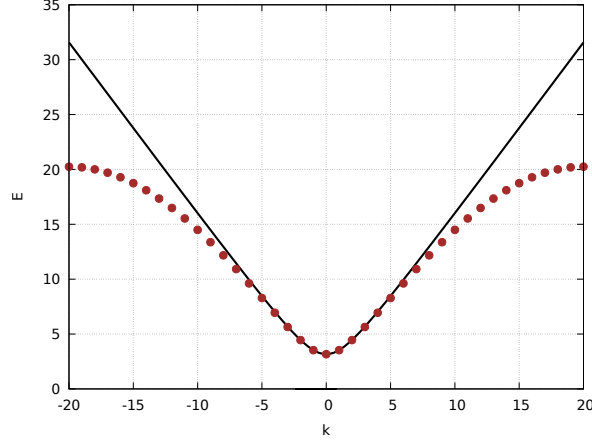
Note that, although the first two terms in the right hand side are real valued, the remainder (terms of order a^6 and beyond) may be complex. Therefore, it is convenient to take the real part of the trace of the Wilson loop in order to construct a real valued discrete action. By summing this equation over all the lattice points x and all the pairs of distinct directions (μ, ν), we obtain

$$a^4 \sum_{x \in \text{lattice}} \left(-\frac{1}{4} F_a^{\mu\nu}(x) F_{\mu\nu}^a(x) \right) = \underbrace{\frac{N}{g^2} \sum_{x \in \text{lattice}} \sum_{(\mu,\nu)} \left(N^{-1} \text{tr} \left(\text{Re} \, [\square]_{x;\mu\nu} \right) - 1 \right)}_{\text{Wilson action, denoted } \frac{1}{g^2} S_W[U]} + \mathcal{O}(a^2) .$$

(7.13)

Note that the error term of order a^6 becomes a term of order a^2 after summation over the lattice sites, since the number of sites grows like a^{-4} if the volume is hold fixed. Thus, the sum of the traces of the Wilson loops over all the elementary plaquettes of the lattice provides a discretization of the Yang-Mills action. In this discrete formulation, the natural variables are not the gauge potentials $A^\mu(x)$ themselves, but the Wilson lines $U_\mu(x)$ that live on the edges of the lattice, called *link variables*. In this notation, x is the starting point and μ the direction of the Wilson line, as illustrated in the left panel of figure 7.3. Note that the Wilson line oriented in the $-\hat{\mu}$ direction, i.e. starting at the point $x + \hat{\mu}$ and ending at the point x, is simply the Hermitic conjugate of $U_\mu(x)$. Under a local gauge transformation, the link variables are changed as follows:

$$U_\mu(x) \quad \rightarrow \quad \Omega^\dagger(x + \hat{\mu}) \, U_\mu(x) \, \Omega(x) . \tag{7.14}$$

The plaquette variable, shown in the right panel of figure 7.3, can then be obtained by multiplying four link variables, as shown by eq. (7.11), and its trace is obviously invariant under the transformation of eq. (7.14).

154

Figure 7.3: *Left: link variable. Right: plaquette on an elementary square of the lattice.*
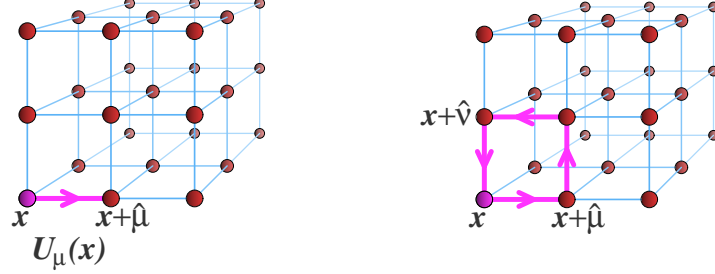
At this stage, the discrete analogue of the path integral that gives the expectation value of a gauge invariant operator reads,

$$\langle \mathcal{O} \rangle = \int \prod_{x,\mu} dU_\mu(x) \; \mathcal{O}[U] \; \exp\left\{ i \frac{N}{g^2} \sum_x \sum_{(\mu,\nu)} \left(N^{-1} \mathrm{tr}\left(\mathrm{Re}\, \square_{x;\mu\nu} \right) - 1 \right) \right\}. \quad (7.15)$$

Since there exists a left- and right-invariant[2] group measure $dU_\mu(x)$, the left hand side of this formula is gauge invariant. Moreover, it goes to the expectation value of the continuum theory in the limit of vanishing lattice spacing.

7.4 Monte-Carlo sampling

Thanks to the discretization, the path integral of the original theory is replaced by an ordinary integral over each of the link variables $U_\mu(x)$, whose number is finite. A non-perturbative answer could be obtained if one were able to evaluate these integrals numerically. However, because of the prefactor i inside the exponential in eq. (7.15), the integrand is a strongly oscillating function, whose numerical evaluation is practically impossible except on lattices with a very small number of sites. In order to be amenable to a numerical calculation, this integral must be transformed into an Euclidean one,

$$\langle \mathcal{O} \rangle_{\text{\tiny E}} = \int \prod_{x,\mu} dU_\mu(x) \; \mathcal{O}[U] \; \exp\left\{ \frac{N}{g^2} \sum_x \sum_{(\mu,\nu)} \left(N^{-1} \mathrm{tr}\left(\mathrm{Re}\, \square_{x;\mu\nu} \right) - 1 \right) \right\}. \quad (7.16)$$

The exponential under the integral is now real-valued, and thus positive definite. Note that numerical quadratures such as Simpson's rule, are not practical for this problem, given the huge number of dimensions of the integral to be evaluated. For instance, for the 8-dimensional Lie group $SU(3)$, in 4 space-time dimensions, on a lattice with N^4 points, this dimension is $8 \times 4 \times N^4$. For $N = 32$, the path integral is thus transformed into a 2^{25}-dimensional ordinary integral. Instead, one views the exponential of the Wilson action as a probability distribution

---

[2]This means that:

$$\int dU \; f[U] = \int dU \; f[\Omega U] = \int dU \; f[U\Omega] \,.$$

Such a measure, known as the Haar measure, exists for compact Lie groups, such as $SU(N)$.

(up to a normalization constant) for the link variables, that may be sampled by a Monte-Carlo algorithm (e.g. the *Metropolis algorithm*) in order to estimate the integral.

In this approach, as long as one is evaluating the expectation value of gauge invariant observables, it is not necessary to fix the gauge in lattice QCD calculations. Gauge fixing is necessary when calculating non-gauge invariant quantities, such as propagators for instance. The Landau gauge is the most commonly used, because the Landau gauge condition is realized at the extrema of a functional of the link variables, However, the comparison between gauge-fixed lattice calculations and analytical calculations is very delicate, because of the existence of Gribov copies (the problem stems from the fact that the two setups may not select the same Gribov copy).

Although considering the Euclidean path-integral instead of the Minkowski one allows for a numerical calculation by Monte-Carlo sampling, this leads to a serious limitation: only quantities that can be expressed as an Euclidean expectation value are directly calculable. Others could in principle be reached by an analytic continuation from imaginary to real time, but this turns out to be practically impossible numerically. For instance, the masses of hadrons are accessible to lattice QCD calculations (see the section 7.6 for an example), while scattering amplitudes cannot be calculated by this method.

7.5 Fermions

7.5.1 Discretization of the Dirac action

Consider now the Dirac action, whose expression in continuum space reads

$$\mathcal{S}_{\mathrm{D}} = \int d^4x \, \overline{\psi}(x) \left(i \gamma^\mu \, D_\mu - m \right) \psi(x) \,. \tag{7.17}$$

In the discretization, we assign one spinor $\psi(x)$ to each site of the lattice. Under a gauge transformation $\Omega(x)$, these spinors transform in the way as in the continuous theory,

$$\psi(x) \; \to \; \Omega^\dagger(x) \, \psi(x) \quad , \qquad \overline{\psi}(x) \; \to \; \Omega(x) \, \overline{\psi}(x) \,. \tag{7.18}$$

The main difficulty in defining a discrete covariant derivative that transforms appropriately under a gauge transformation is that $\psi(x)$ and $\psi(x \pm a)$ transform differently when $\Omega(x)$ depends on space-time. This problem can be remedied by using a link variable between the point x and its neighbors. Like with the ordinary derivatives, one may define forward and backward discrete derivatives,

$$D_{\mathrm{F}}^\mu \psi(x) \equiv \frac{U_\mu^\dagger(x) \psi(x+a) - \psi(x)}{a} \,,$$
$$D_{\mathrm{B}}^\mu \psi(x) \equiv \frac{\psi(x) - U_\mu(x-a) \psi(x-a)}{a} \,. \tag{7.19}$$

However, none of these two operators is anti-adjoint, and therefore they would not give a Hermitean Lagrangian density. This may be achieved by using instead $\frac{1}{2}\left(D_{\mathrm{F}}^\mu + D_{\mathrm{B}}^\mu\right)$, whose action is given by a symmetric forward-backward difference

$$\frac{1}{2}\left(D_{\mathrm{F}}^\mu + D_{\mathrm{B}}^\mu\right) \psi(x) = \frac{U_\mu^\dagger(x) \psi(x+a) - U_\mu(x-a) \psi(x-a)}{2a} \,. \tag{7.20}$$

7.5.2 Fermion doublers

Let us now study how the dispersion relation of fermions is modified by this discretization. This can be done in the vacuum, i.e. by setting all the link variables to the identity. In this case, the eigenfunctions of the operator $\frac{1}{2}\left(D_F^\mu + D_B^\mu\right)$ are

$$\psi_k(x) = e^{2i\pi \frac{(k+1/2)x}{Na}} \qquad \text{with } k \in \mathbb{Z}, \quad -\frac{N}{2} \le k \le \frac{N}{2}, \tag{7.21}$$

with the eigenvalue

$$\lambda_k = \frac{i}{a} \sin \frac{2\pi(k+1/2)}{N}, \tag{7.22}$$

and the corresponding dispersion relation is $E^2 = |\lambda_k|^2 + m^2$. This dispersion relation is shown in the figure 7.4. Like in the bosonic case, the discrete dispersion relation agrees with

Figure 7.4: *Discrepancy between the continuous (solid curve) and discrete (points) dispersion curves in the fermionic case, on a one-dimensional lattice with $N = 40$.*
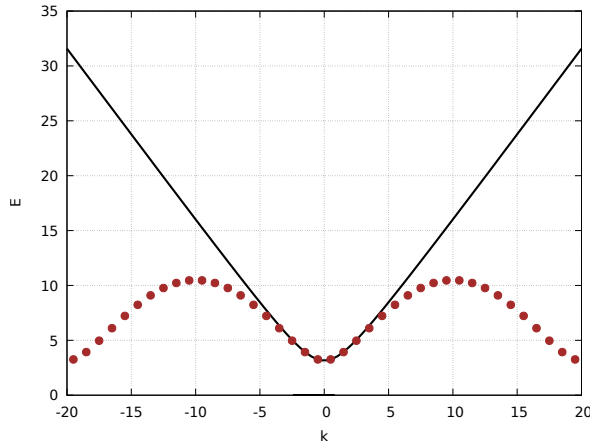
the continuous one only for small enough k. However, the discrepancy at large k is now much more serious, because the discrete dispersion curve has another minimum at the edge of the Brillouin zone. This additional minimum indicates the existence of a second propagating mode of mass m. This spurious mode is called a *fermion doubler*. In d dimensions, the number of these fermionic modes is 2^d, while our goal was to have only one. This problem is quite serious, because it affects all quantities that depend on the number of quark flavors. In particular, this is the case of the running of the coupling constant, whenever quark loops are included.

7.5.3 Wilson term

Modifications of the discretized Dirac action have been proposed to remedy the problem of fermion doublers. One of these modifications, known as the *Wilson term*, consists in adding

to the Lagrangian the following term (written here for the direction μ),

$$-\frac{1}{2a}\, \overline{\psi}(x)\left[U^\dagger_\mu(x)\,\psi(x+a) + U_\mu(x-a)\,\psi(x-a) - 2\psi(x)\right],$$ (7.23)

which is nothing but a D'Alembertian (or a Laplacian in the Euclidean theory) constructed with covariant derivatives

$$\frac{a}{2}\, \overline{\psi}\left(D_\mu D^\mu\right)\psi\,.$$ (7.24)

Note that the denominator in eq. (7.23) has a single power of the lattice spacing a, hence the prefactor a in the previous equation. Therefore, this term goes to zero in the limit $a \to 0$, and it should have no effect in the continuum limit. In the absence of gauge field ($U_\mu(x) \equiv 1$), the functions of eq. (7.21) are still eigenfunctions after adding the Wilson term, but with modified eigenvalues,

$$\lambda_k = \frac{i}{a}\, \sin\frac{2\pi\,(k+1/2)}{N} + \frac{1}{a}\left(1 - \cos\frac{2\pi\,(k+1/2)}{N}\right)\,.$$ (7.25)

Thus, the Wilson term does not modify the spectrum at small k, but lifts the spurious mini-

Figure 7.5: *Discrepancy between the continuous (solid curve) and discrete (points) dispersion curves in the fermionic case, on a one-dimensional lattice with $N = 40$, after inclusion of the Wilson term.*
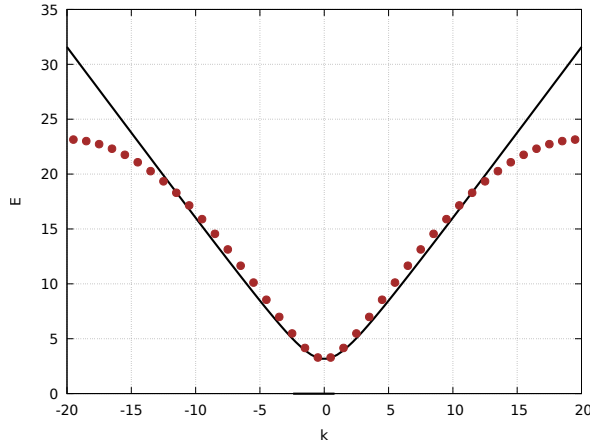
mum that existed at the edge of the Brillouin zone, as shown in the figure 7.5. However, the Wilson term has an important drawback: there is no Dirac matrix γ^μ in eqs. (7.23) and (7.24) since the Lorentz indices are contracted between the two covariant derivatives. Therefore, the Wilson term –like an ordinary mass term– breaks explicitly the chiral symmetry of the Dirac Lagrangian. The fermion doublers are in fact intimately related to chiral symmetry. Without the Wilson term, lattice QCD with massless quarks has an exact chiral symmetry unbroken by the lattice regularization, and therefore there cannot be a chiral anomaly. In fact, this absence of anomaly is precisely due to a cancellation of anomalies among the multiple copies (the doublers) of the fermion modes. This argument is completely general and not

specific to the Wilson term: any mechanism that lifts the degeneracy among the doublers will spoil the anomaly cancellation and thus break chiral symmetry. For this reason, the study of phenomena related to chiral symmetry is always delicate in lattice QCD.

7.5.4 Evaluation of the fermion path integral

The path integral representation for fermions uses anti-commuting Grassmann variables. However, such variables are not representable as ordinary numbers in a numerical implementation. To circumvent this difficulty, one exploits the fact that the Dirac action is quadratic in the fermion fields (this remains true after adding the Wilson term to remove the fermion doublers). Therefore, the path integral over the fermion fields can be done exactly. In addition to the fermion fields contained in the action, there may be ψ's and $\overline{\psi}$'s (in equal numbers) in the operator whose expectation value is being evaluated. The result of such a fermionic path integral is given by

$$\int \left[D\psi D\overline{\psi} \right] \left(\psi(x_1)\overline{\psi}(x_2) \right) e^{iS_D[\psi,\overline{\psi}]} = S(x_1, x_2) \times \det \left(i\gamma^\mu D_\mu - m \right) ; \quad (7.26)$$

where $S(x_1, x_2)$ is the inverse of the Dirac operator $i\slashed{D} - m$ between the points x_1 and x_2. When there is more than one $\psi\overline{\psi}$ pair in the operator, one must sum over all the ways of connecting the ψ's and the $\overline{\psi}$'s by the propagators $S(x, y)$. The same can be done in the lattice formulation. In this case, the Dirac operator is simply a (very large) matrix that depends on the configuration of link variables. Therefore one needs the inverse of this matrix, and its determinant.

Figure 7.6: *Illustration of the two types of quark contributions. In red: quark propagators (i.e. inverse of the Dirac operator) that connect the ψ's and $\overline{\psi}$'s in the operator being evaluated. In purple: quark loops coming from the determinant of the Dirac operator.*
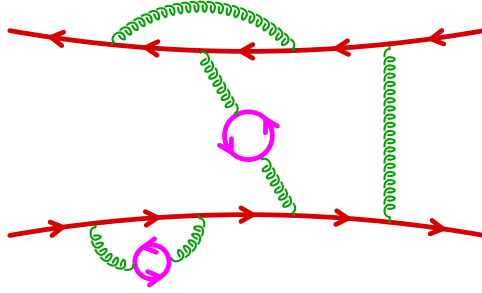
In eq. (7.26), the determinant provides closed quark loops, while the propagator $S(x_1, x_2)$ connects the external points of the operator under consideration. This observation, illustrated in the figure 7.6, clarifies the meaning of the *quenched approximation*, in which the determinant of the Dirac operator is replaced by 1. This approximation, motivated primarily by the computational difficulty of evaluating the Dirac determinant, was widely used in lattice QCD computations until advances in computer hardware made it unnecessary. Note that, although quark loops are not included in the quenched approximation, gluon loops are present to their full extent. In contrast, lattice QCD calculations that include the Dirac determinant, and thus the effect of quark loops, as said to use *dynamical fermions*.

7.6 Hadron mass determination on the lattice

Let us consider a hadronic state $|h\rangle$. Any operator $\mathcal{O}$ that carries the same quantum numbers as this hadron leads to a non-zero matrix element $\langle h|\mathcal{O}|0\rangle$. The vacuum expectation value of the product of two $\mathcal{O}$ at different times 0 and T can be rewritten as follows,

$$
\begin{aligned}
\langle 0|\mathcal{O}^\dagger(0)\mathcal{O}(T)|0\rangle &= \sum_n \langle 0|\mathcal{O}^\dagger(0)|\Psi_n\rangle\langle\Psi_n|\mathcal{O}(T)|0\rangle \\
&= \sum_n \langle 0|\mathcal{O}^\dagger(0)|\Psi_n\rangle\langle\Psi_n|\mathcal{O}(0)|0\rangle \, e^{-M_n T} \\
&= \sum_n \left|\langle\Psi_n|\mathcal{O}(0)|0\rangle\right|^2 e^{-M_n T} .
\end{aligned}
\tag{7.27}
$$

In the first equality, we have inserted a complete basis of eigenstates of the QCD Hamiltonian, and the second equality follows from the fact that $\langle\Psi_n|$ is an eigenstate of rest energy M_n (these is no factor i inside the exponential because of the Euclidean time used in lattice QCD). The sum in the last equality receives non-zero contributions from all the states Ψ_n that possess the quantum numbers carried by the operator $\mathcal{O}$. However, taking the limit $T \to \infty$ selects the one among these eigenstates that has the smallest mass. This observation can be turned into a method to determine hadron masses in lattice QCD:

1. Choose an operator $\mathcal{O}$ that has the quantum numbers of the hadron of interest. The choice of the operator is not crucial, as long as the overlap $\langle h|\mathcal{O}|0\rangle$ is not zero. However, eq. (7.27) suggests that a better, i.e. less noisy with limited statistics, result may be obtained by trying to maximize this overlap.

2. Evaluate the vacuum expectation value of $\mathcal{O}^\dagger(0)\mathcal{O}(T)$ by Monte-Carlo sampling, as a function of T.

3. Fit the large T tail of this expectation value. The slope of the exponential gives the mass of the lightest hadron that carries these quantum numbers.

The discretized QCD Lagrangian contains several dimensionful parameters: the lattice spacing a and the quark masses m_f (one for each quark flavor), whose values need to be fixed before novel predictions can be made. One must choose (at least) an equal number of physical quantities that are known experimentally. Their computed values depend on a, m_f, and one should adjust these parameters so that they match the experimental values. After this has been done, quantities computed in lattice QCD do not contain any free parameter anymore and are thus predictions. The figure 7.7 shows results of the determination of hadron masses using lattice QCD.

7.7 Wilson loops and confinement

7.7.1 Strong coupling expansion

While perturbation theory is an expansion in powers of g^2, it is possible to use the lattice formulation of a Yang-Mills theory in order to perform an expansion in powers of the quantity

Figure 7.7: *Hadron mass determination from lattice QCD. Blue: masses used as input in order to set the lattice parameters. Red: predictions of lattice QCD. Boxes: experimental values.*
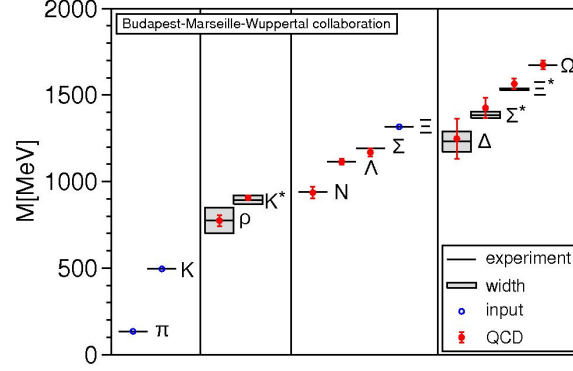
$\beta \equiv g^{-2}$ that appears as a prefactor in the Wilson action. This is called a *strong coupling expansion*, since it becomes exact in the limit of infinite coupling.

This expansion produces integrals over the gauge group such as

$$\int dU\, U_{i_1 j_1} \cdots U_{i_n j_n}\, U^\dagger_{k_1 l_1} \cdots U^\dagger_{k_m l_m} \,. \tag{7.28}$$

The simplest of these integrals,

$$\int dU = 1 \tag{7.29}$$

is simply a choice of normalization of the group measure. From the unitarity of the group elements, one then obtains

$$\int dU\, U_{ij} U^\dagger_{kl} = \frac{1}{N}\, \delta_{jk}\delta_{il} \,. \tag{7.30}$$

In these integrals, the link variables on different edges of the lattice are independent variables, and there is a separate integral for each of them. This is completely general: integrals of the form (7.28) are non-zero only if the integrands contains an equal number of U's and $U^\dagger$'s, i.e. for $n = m$. Therefore, each link variable U that appears in such a group integral must be matched by a corresponding $U^\dagger$. For instance, the group integral of the Wilson loop defined on an elementary plaquette is zero,

$$\int \prod_{x,\mu} dU_\mu(x)\, \mathrm{tr}\left(\underbrace{U^\dagger_\nu(x)\, U^\dagger_\mu(x+\hat\nu)\, U_\nu(x+\hat\mu)\, U_\mu(x)}_{[\Box]_{x;\mu\nu}} \right) = 0 \,, \tag{7.31}$$

because the four link variables live on four distinct edges of the lattice. In contrast, the integral of the trace of a plaquette time the trace of the conjugate plaquette is non-zero:

$$\int \prod_{x,\mu} dU_\mu(x)\, \left(\mathrm{tr}\,[\Box]_{x;\mu\nu}\right)\, \left(\mathrm{tr}\,[\Box]^\dagger_{x;\mu\nu}\right) = 1 \,. \tag{7.32}$$

161

Using these results, we can calculate to order β the expectation value of the trace of a plaquette:

$$\langle \mathrm{tr}\,[\Box]_{x;\mu\nu} \rangle \equiv \frac{\int \prod_{x,\mu} dU_\mu(x) \left(\mathrm{tr}\,[\Box]_{x;\mu\nu} \right) \exp\left\{ \beta N \sum_{y;\rho\sigma} \left(N^{-1} \mathrm{tr}\,\mathrm{Re}\,[\Box]_{y;\rho\sigma} - 1 \right) \right\}}{\int \prod_{x,\mu} dU_\mu(x) \, \exp\left\{ \beta N \sum_{y;\rho\sigma} \left(N^{-1} \mathrm{tr}\,\mathrm{Re}\,[\Box]_{y;\rho\sigma} - 1 \right) \right\}}$$

$$= \frac{\beta}{2} + \mathcal{O}(\beta^2) \, . \tag{7.33}$$

Consider now the trace of a more general Wilson loop along a path γ (planar, to simplify the discussion). Each U and $U^\dagger$ in the Wilson loop must be compensated by a link variable coming from the β expansion of the exponential of the Wilson action. The lowest order term in β corresponds to a minimal tiling of the Wilson loop by elementary plaquettes, as illustrated in the figure 7.8. The corresponding contribution is

Figure 7.8: *Tiling of a closed loop by elementary plaquettes.*
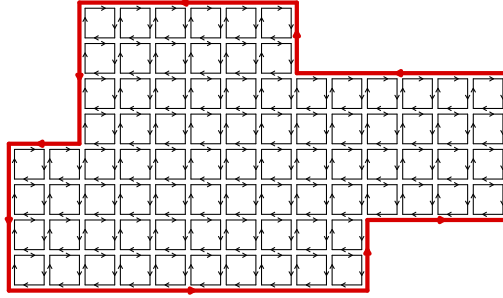
$$\langle \mathrm{tr}\,W_\gamma \rangle = \left(\frac{\beta}{2} \right)^{\mathrm{Area}\,(\gamma)} + \cdots \, , \tag{7.34}$$

where the dots are terms of higher order in β (that can be constructed from non-minimal tilings of the contour γ, such that all the U's and $U^\dagger$'s are still paired appropriately).

7.7.2 Heavy quark potential

Let us consider now a rectangular loop, with an extent R in the spatial direction 1 and an extent T in the Euclidean time direction 4. The previous result indicates that the expectation value of the trace of the corresponding Wilson loop has the following form,

$$\langle \mathrm{tr}\,W_\gamma \rangle \sim e^{-\sigma RT} + \cdots \, , \tag{7.35}$$

where σ is a constant. Although it is gauge invariant, this expectation value is easier to interpret in an axial gauge where $A_4 \equiv 0$. Indeed, in this gauge, the Wilson loop receives only contributions from gauge links along the spatial direction, as shown in the figure 7.9.

162

Figure 7.9: *Rectangular Wilson loop in the $A_4 \equiv 0$ gauge.*
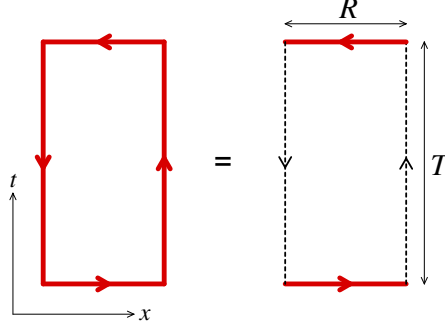
Note that the remaining Wilson lines are precisely those that are needed to make a non-local operator made of a quark at $x = R$ and an antiquark at $x = 0$ gauge invariant,

$$\mathcal{O}_{q\bar{q}}(t) \equiv \overline{\psi}(t, 0) \, W_{[0,R]} \, \psi(t, R) \,, \tag{7.36}$$

where $W_{[0,R]}$ is a (spatial) Wilson line going from (t, R) to $(t, 0)$. Consider now the vacuum expectation value $\langle 0 | \mathcal{O}_{q\bar{q}}^{\dagger}(0) \mathcal{O}_{q\bar{q}}(T) | 0 \rangle$. In this expectation value, the fermionic path integral produces two quark propagators that connect the ψ's to the $\overline{\psi}$'s. However, in the limit of infinite quark mass, the quarks propagate on straight lines at constant velocity and their propagator reduces to a Wilson line along this trajectory. For the propagation between $(0, x)$ and (T, x), this is a temporal Wilson line, that reduces to the identity in the $A_4 = 0$ gauge (represented by the dotted lines in the figure 7.9). Thus, we have

$$\langle \operatorname{tr} W_{\gamma} \rangle \propto \lim_{M \to \infty} \langle 0 | \mathcal{O}_{q\bar{q}}^{\dagger}(0) \mathcal{O}_{q\bar{q}}(T) | 0 \rangle \,. \tag{7.37}$$

Inserting a complete basis of eigenstates of the Hamiltonian in the right hand side of eq. (7.37) and taking the limit $T \to \infty$, we find a result dominated by the quark-antiquark state of lowest energy E_0,

$$\lim_{T \to \infty} \langle 0 | \mathcal{O}_{q\bar{q}}^{\dagger}(0) \mathcal{O}_{q\bar{q}}(T) | 0 \rangle = \left| \langle 0 | \mathcal{O}_{q\bar{q}}^{\dagger}(0) | \Psi_0 \rangle \right|^2 e^{-E_0 T} \,. \tag{7.38}$$

Moreover, in the limit of large mass, the energy E_0 of this state is dominated by the potential energy $V(R)$ between the quark and the antiquark (the quark and antiquark are non-relativistic, and their kinetic energy behaves as $\mathbf{P}^2 / 2M \to 0$),

$$\lim_{M, T \to \infty} \langle 0 | \mathcal{O}_{q\bar{q}}^{\dagger}(0) \mathcal{O}_{q\bar{q}}(T) | 0 \rangle = \left| \langle 0 | \mathcal{O}_{q\bar{q}}^{\dagger}(0) | \Psi_0 \rangle \right|^2 e^{-V(R) T} \tag{7.39}$$

By comparing this result with that of the strong coupling expansion, eq. (7.35), we conclude that

$$V(R) = \sigma R \,. \tag{7.40}$$

This linear potential indicates that the force between the quark and antiquark is constant at large distance, in sharp contrast with a Coulomb potential in electrodynamics. This is a consequence of the *color confinement* property of QCD.